ACCORDING TO HEC SYLLABUS

TOOLS FOR REASONING SKILLS

QUANTITATIVE REASONING II

2024

Muhammad Usman Hamid

Preface

In an era where data is abundant and insights are scarce, the ability to reason quantitatively has become a vital skill for navigating the complexities of modern life. From the scientist evaluating evidence to the business leader making informed decisions, quantitative reasoning is the foundation upon which wise choices are built.

"Tools for Quantitative Reasoning" is designed to equip readers with the conceptual frameworks, practical methods, and critical thinking skills necessary to excel in a data-driven world. This book provides a comprehensive toolkit for working with numbers, understanding statistical analysis, and applying quantitative techniques to real-world problems.

Through its pages, readers will discover how to:

Evaluate evidence and arguments using quantitative criteria

Apply statistical methods to understand complex phenomena

Make informed decisions based on data-driven insights

Communicate quantitative ideas with clarity and precision

This course is based on quantitative reasoning I course. It will enhance the quantitative reasoning skills learned in quantitative reasoning I course. Students will be introduced to more tools necessary for quantitative reasoning skills to live in the fast paced 21st century. Students will be introduced to importance of statistical and mathematical skills in different professional settings, social and natural sciences. These quantitative reasoning skills will help students to better participate in national and international issues like political and health issues. This course will prepare the students to apply quantitative reasoning tools more efficiently in their professional and daily life activities. This course will help them to better understand the information in form of numeric, graphs, tables, and functions.

Whether you are a student, professional, or simply someone seeking to enhance your analytical abilities, "Tools for Quantitative Reasoning" is your guide to unlocking the power of quantitative thinking. Join us on this journey to develop the skills and knowledge necessary to succeed in an increasingly quantitative world.

Muhammad Usman Hamid
Superior Group of Colleges
Sillanwali, Sargodha
0323 – 6032785

CHAPTER

1

INTRODUCTION

TO ENUMERATION AND ITS APPLICATION

Welcome to Quantitative Skills and Reasoning! Just what are quantitative skills and reasoning? The simple answer is working with numbers, making sense of data, and using your brain to figure things out. This will cover fundamental concepts from problem solving, statistics, probability, graphs, logic, sets, measurements, and finance.

In this chapter we will learn about;

- Introduction to quantitative reasoning
- Types of quantitative reasoning
- Types of standard numbers system & basic arithmetic operations
- The Base numbers system
- Overview of contributions of Mathematicians and Statisticians especially Muslim scholars
- Inductive reasoning
- Deductive reasoning
- Abductive Reasoning

What is Reasoning?

Reasoning is the ability to assess things rationally by applying logic based on new or existing information when making a decision or solving a problem. It allows you to weigh the benefits and disadvantages of two or more courses of action before choosing the one with the most benefit or the one that suits your needs.

Types of Reasoning

- **Deductive Reasoning:** Reasoning that uses formal logic and observations to prove a theory or hypothesis. It can be used to apply a general law to a specific case or test an induction. Its results typically have a logical certainty.
- **Inductive Reasoning:** Inductive reasoning is the process of reaching a general conclusion by examining specific examples. It uses theories and assumptions to validate observations. It can be used to apply a specific law to a general. Its results are not always certain because it uses conclusions from observations to make generalizations. It is helpful for extrapolation, prediction and part – to – whole arguments.
- **Analogical Reasoning:** Form of thinking that finds similarities between two or more things and then use those characteristics to find other qualities common to them. It is based on brain tendency. It can help you expand your understanding by looking for similarities between different things.
- **Abductive Reasoning:** Type of reasoning that uses an observation or set of observations to reach a logical conclusion. It is similar to inductive reasoning; however it permits making best guesses to arrive at the simplest conclusions.
- **Cause and Effect Reasoning:** Type of thinking in which you show the linkage between two events. It explains what may happen if an action takes place or why things happen when some conditions are present.
- **Critical Thinking:** It involves extensive rational thought about a specific object in order to come to a definitive conclusion. It is helpful in logic, computing and social sciences.
- **Decompositional Reasoning:** It is the process of breaking things into constituent parts to understand the function of each component and how it contributes to the operation of the item as a whole. It is helpful in logic, computing, game theory, product development, marketing and social sciences.

Quantitative Skills

Any skills that use or manipulate numbers are called quantitative skills. They help to make sense of numerical, categorical or ordinal data and scientific concepts. It is helpful in statistics, economics, algebra, finance, business, logic and social sciences.

Quantitative Reasoning / Quantitative Literacy / Enumeration / Numeracy

Quantitative Reasoning is the ability to assess mathematical ideas or things rationally by applying logic based on new or existing information when making a decision or solving a problem. It is application of mathematical concepts or skills to solve real world problems.

Importance of Quantitative Skills / Numeracy

Numeracy is simply the application of critical thinking skills like analysis and interpretation along with mathematical basics like algebra to quantitative information. It refers to the ability to solve quantitative reasoning problems, or to making judgment derived from quantitative reasoning in a variety of context. It helps to make sense of numerical, categorical or ordinal data and scientific concepts. It is helpful in statistics, economics, algebra, finance, business, logic and social sciences.

Quantitative Reasoning Examples

- **Statistical Analysis:** Analysts apply quantitative reasoning when they assess large dataset to derive meaningful conclusions. They use statistical methods like regression analysis and hypothesis testing to interpret data and distinguish patterns.
- **Financial Planning:** A financial planner utilizes quantitative reasoning for a client's investment strategy. This involves analyzing expected returns, tax implications and risk factors.

What is Mathematics?

The branch of science that deals with the numbers is called Mathematics. The world "Mathematics" is derived from the Greek word "Mathematikos" which means "inclined to learn".

Mathematics is based on deductive reasoning though man's first experience with mathematics was of an inductive nature. This means that the foundation of mathematics is the study of some logical and philosophical notions. We elaborate in simple terms that the deductive system involves four things:

Known Branches of Mathematics

- **Logic:** The Study of Principles of Reasoning.
- **Arithmetic:** Method for operating on numbers.
- **Algebra:** Method for working with unknown quantities.
- **Geometry:** The study of size and shape.
- **Trigonometry:** The study of triangles and their uses.
- **Probability:** The study of chance.
- **Statistics:** Method for analyzing data.
- **Calculus:** The study of quantities that change.

Number System

A number is a mathematical object used to count, measure, & label. It is the mathematical notation for representing numbers of a given set by using digits or other symbols in a consistent manner. It provides a unique representation of every number and represents the arithmetic and algebraic structure of the figures.

Number System: A system of writing to express numbers. It presents a unique representation of numbers.

Types of Standard Numbers

1. Natural Numbers

Common counting numbers. Natural numbers are also called "counting numbers" which contains the set of positive integers that start at 1 and continue infinitely. The set of natural numbers is represented by the letter "N". i.e. N = {1,2,3,4,5….}.

2. Whole Numbers

In math, whole numbers are positive integers, including zero, that do not have any decimal or fractional parts. The symbol for whole numbers is "W". i.e. W = {0, 1, 2, 3, 4, 5...}.

3. Integers

Integers, also known as whole numbers or round numbers, or positive or negative numbers that don't have fractional or decimal parts. The symbol for integers is Z

i.e. Z = {…,−1,−2,0,1,2,...}.

4. Rational Numbers

The set of rational numbers includes all the integers, each of which can be written as a quotient with the integer as the numerator and 1 as the denominator. Rational number, in arithmetic, a number that can be represented as the quotient p/q of two integers such that $q \neq 0$. i.e. $Q = \left\{x : x = \frac{p}{q}, q \neq 0, p, q \in Z\right\}$.

5. Irrational Numbers

An irrational number is a real number that cannot be written as a fraction of two integers, or in the form of p/q, where p and q are integers and $q \neq 0$. e.g. $F = \{\sqrt{2}, \sqrt{3}, \sqrt{5}, \pi\}$.

Decimal Representation of Rational and Irrational Numbers

1) Terminating decimals: A decimal which has only a finite number of digits in its decimal part, is called a terminating decimal. Thus 202.04, 0.0000415, 100000.41237895 are examples of terminating decimals. Since a terminating decimal can be converted into a common fraction, so every terminating decimal represents a rational number.

2) Recurring Decimals: A decimal in which one or more digits repeat indefinitely is called a recurring or periodic decimal. e.g., 1.333…, 21.134134… are both recurring decimals.

3) Non-terminating and Non-recurring decimal: A non-terminating, non-recurring decimal is a decimal which neither terminates nor it is recurring. It is not possible to convert such a decimal into a common fraction. Thus a non-terminating, non-recurring decimal represents an irrational number.

Examples:

i) $.25\left(=\frac{25}{100}\right)$ is a rational number.

ii) $.333\ldots(\frac{1}{3}=)$ is a recurring decimal, it is a rational number.

iii) $2.\overline{3}(=2.333\ldots)$ is a rational number.

iv) $0.142857142857\ldots\left(=\frac{1}{7}\right)$ is a rational number.

v) 0.01001000100001 ... is a non-terminating, non-periodic decimal, so it is an irrational number.

vi) 214.121122111222 1111 2222 ... is also an irrational number.

vii) 1.4142135 ... is an irrational number.

viii) 7.3205080 ... is an irrational number.

ix) 1.709975947 ... is an irrational number.

x) 3.141592654... is an important irrational number called it π(Pi) which denotes the constant ratio of the circumference of any circle to the length of its diameter i.e.,

$$\pi=\frac{circumference\ of\ any\ circle}{length\ of\ its\ diameter}$$

6. Real Numbers

Real numbers can be positive or negative & include fractions, integers& irrational numbers. They can be used in arithmetic operations and represented on a number line. Real numbers includes rationals and irrationals. i.e. $R=Q\cup Q'$.

7. Prime Numbers

A prime number is a whole number that is greater than 1 and can only be divided by itself and 1 without a remainder. For example, 19 is a prime number because it can only be divided by 1 and 19 these are P ={2,3,5,7,11…}.

8. Composite Numbers

A number that is divisible by a number other than 1 and the number itself, is called a composite number. For example, 4 and 6 are composite numbers.

9. Complex Numbers

A complex number is a number that has both real and imaginary parts, and is written in the form $C = \{a + ib: a, b \in R\}$. For example $C \ = \{2 + 0i \ = 2, 1 + 3i\}$.

10. Even & Odd Numbers

Even numbers are numbers that can be divided into two equal parts, while odd numbers are numbers that cannot.

Even numbers: End in 0, 2, 4, 6, or 8.

Odd numbers: End in 1, 3, 5, 7, or 9.

Arithmetic

Arithmetic is a field of mathematic that studies the characteristics of classical operations on numbers, such as addition, subtraction, multiplication, division, exponentiation and root extraction.

Arithmetic Operations

Arithmetic is the fundamental of mathematics that includes the operations of numbers. These operations are addition, subtraction, multiplication and division. Arithmetic is one of the important branches of mathematics that lays the foundation of the subject 'Mathematics', for students.

Addition: Combines objects into a larger collection, or increases a value. It is represented by the plus sign (+) and the answer is called the sum. For example, 4 + 7 = 11.

Subtraction: Finds the difference between numbers or quantities, or decreases a value. It is represented by the minus sign (-) and the answer is called the difference.

For example, 9 - 7 = 2.

Multiplication: Multiplication is represented by the multiplication signs (×) or (*). For example, 8 multiplied by 4 is equal to 32, which can be written as 8 × 4 = 32.

Division: Division is a method of dividing or distributing a number into equal parts. For example, 16 divided by 4 is equal to 4, which can be written as 16 ÷ 4 = 4.

Base Number System

A base number system, also known as a radix or numeral system, is a mathematical notation that represents numbers using a specific set of digits or symbols. Each base number system has its own unique characteristics and applications, and is used in various fields such as mathematics, computer science, and engineering.

Some common base number systems are as follows:

Base 2 - Binary Number System

The binary number system is a base-2 number system that uses only two digits: 0 and 1.

Base 8 - Octal Number System

The octal number system is a base-8 number system that uses eight digits: 0-7.

Base 10 - Decimal Number System

The decimal number system is a base-10 number system that uses ten digits: 0-9.

Hexadecimal Number System (base 16)

The hexadecimal number system is a base-16 number system that uses ten digits: 0-15.

Note that usually, the hexadecimal digits used are 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, and F, where the letters A through F represent the digits corresponding to the numbers 10 through 15 (in decimal notation).

1. **What is the decimal expansion of the integer that has $(1\ 0101\ 1111)_2$ as its binary expansion?**

Solution: We have

$$(1\ 0101\ 1111)_2 = 1 \cdot 2^8 + 0 \cdot 2^7 + 1 \cdot 2^6 + 0 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 351.$$

2. **What is the decimal expansion of the number with octal expansion $(7016)_8$?**

Solution: Using the definition of a base b expansion with b = 8 tells us that

$$(7016)_8 = 7 \cdot 8^3 + 0 \cdot 8^2 + 1 \cdot 8 + 6 = 3598.$$

3. **What is the decimal expansion of the number with hexadecimal expansion $(2AE0B)_{16}$?**

Solution: Using the definition of a base b expansion with b = 16 tells us that

$$(2AE0B)_{16} = 2 \cdot 16^4 + 10 \cdot 16^3 + 14 \cdot 16^2 + 0 \cdot 16 + 11 = 175627.$$

Find the octal expansion of $(12345)_{10}$.

Solution:

8	12345
8	1543 - 1
8	192 – 7
8	24 – 0
	3 – 0

Hence, $(12345)_{10} = (30071)_8$

4. Find the hexadecimal expansion of $(177130)_{10}$.
Solution:

16	173130
16	11070 - 10
16	691 – 14
16	43 – 3
	2 – 11

Hence, $(177130)_{10} = (2B3EA)_{16}$

(Recall that the integers 10, 11, and 14 correspond to the hexadecimal digits A, B, and E, respectively.)

5. Find the binary expansion of $(241)_{10}$.
Solution:

2	241
2	120 - 1
2	60 – 0
2	30 – 0
2	15 – 0
2	7 – 1
2	3 – 1
2	1 – 1

Hence, $(241)_{10} = (1111\ 0001)_2$

Here, a table is given for the integers 0 to 15 with their expansions in Decimal, Binary, Octal and Hexadecimal systems.

TABLE 1 **Hexadecimal, Octal, and Binary Representation of the Integers 0 through 15.**

Decimal	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Hexadecimal	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Octal	0	1	2	3	4	5	6	7	10	11	12	13	14	15	16	17
Binary	0	1	10	11	100	101	110	111	1000	1001	1010	1011	1100	1101	1110	1111

6. Find the octal expansions of $(11\ 1110\ 1011\ 1100)_2$.
Solution:
To convert $(11\ 1110\ 1011\ 1100)_2$ into octal notation we group the binary digits into blocks of three, adding initial zeros at the start of the leftmost block if necessary.These blocks, from left to right, are 011, 111, 010, 111, and 100, corresponding to 3, 7, 2, 7, and 4, respectively.
Consequently, $(11\ 1110\ 1011\ 1100)_2 = (37274)_8$.

7. **Find the hexadecimal expansions of $(11\ 1110\ 1011\ 1100)_2$.**
   Solution:
   To convert $(11\ 1110\ 10111100)_2$ into hexadecimal notation we group the binary digits into blocks of four, adding initial zeros at the start of the leftmost block if necessary. These blocks, from left to right, are 0011,1110, 1011, and 1100, corresponding to the hexadecimal digits 3, E, B, and C, respectively. Consequently, $(11\ 1110\ 1011\ 1100)_2 = (3EBC)_{16}$.
8. **Find the binary expansions of $(765)_8$ and $(A8D)_{16}$.**
   Solution:
   To convert $(765)_8$ into binary notation, we replace each octal digit by a block of three binary digits. These blocks are 111, 110, and 101. Hence, $(765)_8 = (1\ 1111\ 0101)_2$. To convert $(A8D)_{16}$ into binary notation, we replace each hexadecimal digit by a block of four binary digits. These blocks are 1010, 1000, and 1101. Hence, $(A8D)_{16} = (1010\ 1000\ 1101)_2$.

Contributions of Mathematicians and Statisticians
Especially Muslim Scholars

Here is the list with era, date of birth, and date of death:

Mathematicians:

Isaac Newton (1643-1727)

Developed calculus, laws of motion, and universal gravitation.
Published "Philosophiæ Naturalis Principia Mathematica" (1687).
Laid the foundation for classical mechanics and modern physics.

Archimedes (c. 287 BC - c. 212 BC)

Discovered the principle of buoyancy and developed fluid mechanics.
Made significant contributions to geometry and the study of spheres, cylinders, and cones.
Invented various machines, including the Claw of Archimedes and the Archimedes' screw.

Euclid(fl. 300 BC)

Authored the famous book "Elements," systematizing geometry and establishing axioms.
Developed the concept of theorems and proofs.
Introduced the concept of irrational numbers.

Pierre-Simon Laplace (1749-1827)

Developed probability theory and the concept of expected value.
Made significant contributions to celestial mechanics and the study of planetary motion.
Published "A Philosophical Essay on Probabilities" (1812).

Albert Einstein (1879-1955)
Developed the theory of special relativity and the famous equation $E = mc^2$.
Introduced the concept of spacetime and the speed of light as a universal constant.
Made significant contributions to the development of quantum mechanics.

Statisticians:

Ronald Fisher (1890-1962)
Developed modern statistical inference and experimental design.
Introduced the concept of null hypothesis testing.
Made significant contributions to the development of maximum likelihood estimation.

Karl Pearson (1857-1936)
Developed the correlation coefficient and principal component analysis.
Introduced the concept of the p-value.
Published "The Grammar of Science" (1892).

William Gosset (1876-1937)
Developed the t-distribution and statistical hypothesis testing.
Introduced the concept of the T-test.
Made significant contributions to quality control and statistical process control.

John Turkey (1915-2000)
Developed exploratory data analysis and the Fast Fourier Transform.
Introduced the concept of the box plot.
Made significant contributions to statistical graphics and visualization.

Florence Nightingale (1820-1910)
Developed statistical graphics and applied statistics to medicine.
Introduced the concept of the polar area chart.
Made significant contributions to hospital sanitation and public health.

Muslim Scholars

Muhammad ibn Musa al-Khwarizmi

Muhammad ibn Musa al-Khwarizmi (780AD–850AD) was a Persian mathematician, astronomer, astrologer geographer and a scholar in the House of Wisdom in Baghdad. He was born in Persia of that time around 780. Al-Khwarizmi was one of the learned men who worked in the House of Wisdom. The House of Wisdom was a scientific research and teaching center. Al-Khwarizmi developed the concept of the algorithm in mathematics Al-Khwarizmi's algebra is regarded as the foundation and cornerstone of the sciences. He is known as the "father of algebra", a word derived from the title of his book, Kitab al-Jabr. Muhammad ibn Musa al-Khwarizmi died in c. 850 being remembered as one of the most seminal scientific minds of early Islamic culture.

Ibn al-Haytham

Ibn al-Haytham Latinised as Alhazen (965AD–1040AD) was a medieval mathematician, astronomer, and physicist of the Islamic Golden Age from present-day Iraq. Referred to as "the father of modern optics".he made significant contributions to the principles of optics and the use of scientific experiments. His most influential work is titled Kitāb al-Manāẓir "Book of Optics" in Latin Edition. Ibn al-Haytham, who lived a thousand years ago, is finally being recognized as the world's first true scientist.

Omar Khayyam

Omar Khayyam (1048AD–1131AD) was a Persian mathematician, astronomer, and poet. He made great contributions to these areas. He lived during the period of the Seljuk dynasty, around the time of the First Crusade. As a mathematician, he is most notable for his work on the classification and solution of cubic equations. He is best known for his work in geometric algebra, the Jalil calendar, and his poetry collected as, The Rubaiyat.

Ibrahim ibn Sinan

Ibrahim ibn Sinan (908AD-946AD) was born in Baghdad. He was a mathematician and astronomer who belonged to a family of scholars originally from Harran in northern Mesopotamia. He belonged to a religious sect of star worshippers known as the Sabians of Harran. Ibrahim ibn Sinan studied geometry, in particular tangents to circles. He made advances in the quadrature of the parabola and the theory of integration, generalizing the work of Archimedes, which was unavailable at the time. Ibrahim ibn Sinan is often considered to be one of the most important mathematicians of his time.

Sharaf al-Din al-Tusi

Sharaf al-Din al-Tusi (1135AD-1213AD) was an Iranian mathematician and astronomer of the Islamic Golden Age (during the Middle Ages). Sharaf al-Tusi was an Islamic mathematician who wrote a treatise on cubic equations. Al-Tusi is best known for his mathematically impressive study of the conditions under which cubic equations have a positive real root and of numerical methods for finding a solution of such equations. He made significant contributions to development of Algebraic geometry & cubic equation.

Abu Rayhan Muhammad ibn Ahmad al-Biruni

Abu Rayhan Muhammad ibn Ahmad al-Biruni (973AD–1050AD) known as al-Biruni, was a Khwarazmian Iranian scholar and polymath during the Islamic Golden Age. He has been called variously "Father of Comparative Religion", "Father of modern geodesy", Founder of Indology and the first anthropologist. Al-Biruni was well versed in physics, mathematics, astronomy, and natural sciences, and also distinguished himself as a historian, chronologist, and linguist. In 1017, he travelled to the Indian subcontinent and wrote a treatise on Indian culture entitled Tārīkh al-Hind ("The History of India"). Al-Biruni developed many instruments for astronomy and geography measurements. He was also a very good encyclopedia writer. His famous achievements were, studying geography of India, accurately measuring Earth's radius, comparing different calendars & He enabled direction of Qibla.

Inductive and Deductive Reasoning

Inductive Reasoning

Inductive reasoning is the process of reaching a general conclusion by examining specific examples.

Conjecture: The conclusion formed by using inductive reasoning is often called a **conjecture**, since it may or may not be correct.

Deductive Reasoning

Deductive reasoning is the process of reaching a conclusion by applying general assumptions, procedures, or principles.

9. Use Inductive Reasoning to Predict a Number

Use inductive reasoning to predict the most probable next number in each of the following lists.

a. 3, 6, 9, 12, 15, ?

b. 1, 3, 6, 10, 15, ?

Solution.

a. Each successive number is 3 larger than the preceding number. Thus we predict that the most probable next number in the list is 3 larger than 15, which is **18**.

b. The first two numbers differ by 2. The second and the third numbers differ by 3. It appears that the difference between any two numbers is always 1 more than the preceding difference. Since 10 and 15 differ by 5, we predict that the next number in the list will be 6 larger than 15, which is **21**.

10. Use inductive reasoning to predict the most probable next number in each of the following lists.

a. 5, 10, 15, 20, 25, ?

b. 2, 5, 10, 17, 26, ?

Solution.

a. Each successive number is 5 larger than the preceding number. Thus we predict that the next number in the list is 5 larger than 25, which is 30.
b. The first two numbers differ by 3. The second and third numbers differ by 5. It appears that the difference between any two numbers is always 2 more than the preceding difference. Since 17 and 26 differ by 9, we predict that the next number will be 11 more than 26, which is 37.

11. Use Inductive Reasoning to Make a Conjecture

Consider the following procedure: Pick a number. Multiply the number by 8, add 6 to the product, divide the sum by 2, and subtract 3.

Complete the above procedure for several different numbers. Use inductive reasoning to make a conjecture about the relationship between the size of the resulting number and the size of the original number.

Solution.

Suppose we pick 5 as our original number. Then the procedure would produce the following results:

Original number:	5
Multiply by 8:	$8 \times 5 = 40$
Add 6:	$40 + 6 = 46$
Divide by 2:	$46 \div 2 = 23$
Subtract 3:	$23 - 3 = 20$

We started with 5 and followed the procedure to produce 20. Starting with 6 as our original number produces a final result of 24. Starting with 10 produces a final result of 40. Starting with 100 produces a final result of 400. In each of these cases the resulting number is four times the original number. We conjecture that following the given procedure will produce a resulting number that is four times the original number.

12. Consider the following procedure: Pick a number. Multiply the number by 9, add 15 to the product, divide the sum by 3, and subtract 5.

Complete the above procedure for several different numbers. Use inductive reasoning to make a conjecture about the relationship between the size of the resulting number and the size of the original number.

Solution.

If the original number is 2, then $\frac{2\times9+15}{3} - 5 = 6$ which is three times the original number. If the original number is 7, then $\frac{7\times9+15}{3} - 5 = 21$ which is three times the original number.

If the original number is -12 then $\frac{-12\times9+15}{3} - 5 = -36$ which is three times the original number. It appears, by inductive reasoning, that the procedure produces a number that is three times the original number.

13. Examine the table below. Identify a pattern and then use that pattern to find the missing terms in the sequence.

Term	1	2	3	4	5	6	7	8
Value	1	3	9	27	81			

Solution.

With this problem we see that the pattern to get the next number in the sequence is to multiply the previous term in the sequence by 3. So to find the 6th, 7th, and 8th terms in the sequence we will use this pattern. The 5^{th} term is 81. The 6^{th} term is $3 \times 81 = 243$, the 7^{th} term is $3 \times 243 = 729$, and the 8^{th} term is $3 \times 729 = 2187$.

14. Examine the table below. Identify a pattern and then use that pattern to find the missing terms in the sequence.

Term	1	2	3	4	5	6	7	8
Value	58	46	34	22	10			

Solution.

With this problem we see that the pattern to get the next number in the sequence is to subtract 12 from the previous term in the sequence. To find the 6th, 7th, and 8th terms in the sequence we will use this pattern. The 5^{th} term is 10. The 6^{th} term is $10 - 12 = -2$, 7^{th} term is $-2 - 12 = -14$, and 8^{th} term is $-14 - 12 = -26$.

15. Examine the table below. Identify a pattern and then use that pattern to find the missing terms in the sequence.

Term	1	2	3	4	5	6	7	8	9	10
Value	5	10	30	120	240	720	2880			

Solution.

With this sequence we see to go from 5 to 10 we multiply by 2. To go from 10 to 30 we multiply by 3. To go from 30 to 120 we multiply by 4. Then we see that this pattern repeats to get the next three terms in the sequence. $2 \times 120 = 240$, $3 \times 240 = 720$, and $4 \times 720 = 2880$. So we will use this same pattern to get the 8^{th}, 9^{th}, and 10^{th} terms. The 8^{th} term is $2 \times 2880 = 5760$, the 9^{th} term is $3 \times 5760 = 17280$, and the 10th term is $4 \times 17280 = 69120$.

16. Use the data in the table and inductive reasoning to answer each of the following.

Length of Pendulum, in Units	Period of Pendulum, in Heartbeats
1	1
4	2
9	3
16	4

a. If a pendulum has a length of 25 units, what is its period?
b. If the length of a pendulum is quadrupled, what happens to its period?
Solution.
a. In the table, each pendulum has a period that is the square root of its length. Thus we conjecture that a pendulum with a length of 25 units will have a period of 5 heartbeats.
b. In the table, a pendulum with a length of 4 units has a period that is twice that of a pendulum with a length of 1 unit. A pendulum with a length of 16 units has a period that is twice that of a pendulum with a length of 4 units. It appears that quadrupling the length of a pendulum doubles its period.

17. A tsunami is a sea wave produced by an under-water earthquake. The velocity of a tsunami as it approaches land depends on the height of the tsunami. Use the table at the left and inductive reasoning to answer each of the following questions.

Height of Tsunami, in Feet	Velocity of Tsunami, in Feet Per Second
4	6
9	9
16	12
25	15
36	18
49	21
64	24

a. What happens to the height of a tsunami when its velocity is doubled?
b. What should be the height of a tsunami if its velocity is 30 feet per second?
Solution.
a. It appears that when the velocity of a tsunami is doubled, it height is quadrupled.
b. A tsunami with a velocity of 30 feet per second will have a height that is four times that of a tsunami with a speed of 15 feet per second. Thus, we predict a height of $4 \times 25 = 100$ feet for a tsunami with a velocity of 30 feet per second.

18. The last four times I have driven downtown at 6pm there has been traffic. Use inductive reasoning to draw your conclusion.
Solution.
My conclusion is that there is always traffic downtown around 6pm.

19. Consider the statement and determine if it is inductive or deductive:
"Every month has 30 days in it. July is a month. Therefore it has 30 days in it."
Solution.
This statement starts with a generalization and it's then applied to a specific case. This follows the pattern of **deductive reasoning**. The statements are not necessarily true, but if every month has 30 days in it, then it would be true.

20. Use deductive reasoning to show that the following procedure produces a number that is four times the original number.
Procedure: Pick a number. Multiply the number by 8, add 6 to the product, divide the sum by 2, and subtract 3.
Solution.
Let n represent the original number.

Multiply the number by 8: $8n$

Add 6 to the product: $8n + 6$

Divide the sum by 2: $\frac{8n + 6}{2} = 4n + 3$

Subtract 3: $4n + 3 - 3 = 4n$

We started with n and ended with 4n. The procedure given in this example produces a number that is four times the original number.

21. Use deductive reasoning to show that the following procedure produces a number that is three times the original number.

Procedure: Pick a number. Multiply the number by 6, add 10 to the product, divide the sum by 2, and subtract 5. Hint: Let n represent the original number.

Solution.

Let n represent the original number.

$6n$

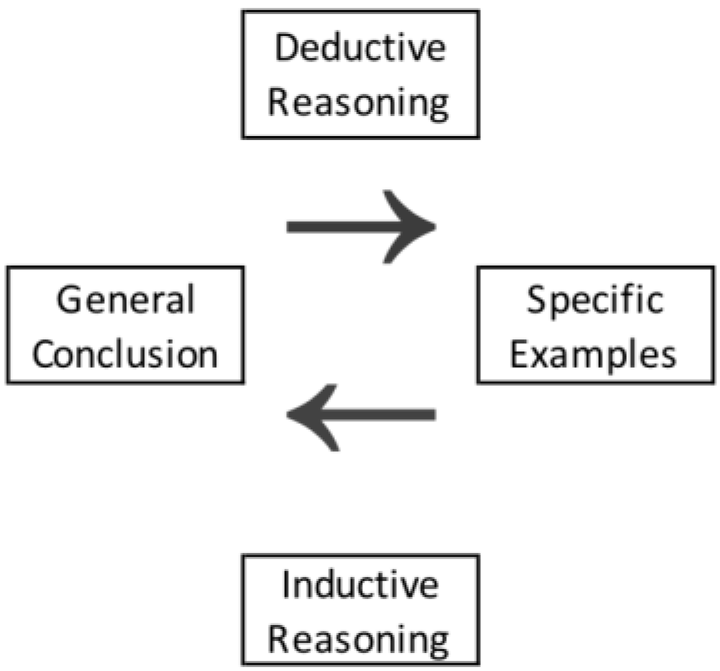
$6n + 10$

$\frac{6n + 10}{2} = 3n + 5$

$3n + 5 - 5 = 3n$

The procedure always produces a number that is three times the original number.

Inductive Reasoning vs. Deductive Reasoning



- For Inductive Reasoning we start with examples or cases, and then draw general conclusions.
- For Deductive Reasoning we start with a general statement and apply it to examples or cases.

22. Determine whether each of the following arguments is an example of inductive reasoning or deductive reasoning.

a. During the past 10 years, a tree has produced plums every other year. Last year the tree did not produce plums, so this year the tree will produce plums.

b. All home improvements cost more than the estimate. The contractor estimated my home improvement will cost $35,000. Thus my home improvement will cost more than $35,000.

Solution.

a. This argument reaches a conclusion based on specific examples, so it is an example of inductive reasoning.

b. Because the conclusion is a specific case of a general assumption, this argument is an example of deductive reasoning.

23. Determine whether each of the following arguments is an example of inductive reasoning or deductive reasoning.

a. All Janet Evanovich novels are worth reading. The novel *To the Nines* is a Janet Evanovich novel. Thus *To the Nines* is worth reading.

b. I know I will win a jackpot on this slot machine in the next 10 tries, because it has not paid out any money during the last 45 tries.

Solution.

a. The conclusion is a specific case of a general assumption, so the argument is an example of deductive reasoning.
b. The argument reaches a conclusion based on specific examples, so the argument is an example of inductive reasoning.

24. Consider the statement and determine if it is inductive or deductive:

" Pizza Hut has a lunch buffet. Stevi B’s has a lunch buffet. Therefore all pizza restaurants have a lunch buffet."

Solution.

This statement starts with two examples about pizza restaurants having lunch buffets. Based on these examples a generalization is made. This follows the pattern of **inductive reasoning**.

25. Consider the statement and determine if it is inductive or deductive:

" All pro wrestlers have a catch phrase. Macho Man Randy Savage was a pro wrestler. Therefore he had a catch phrase. ”

Solution.

This statement starts with a generalization about pro wrestlers having catch phrases. It’s then applied to the specific case of Macho Man Randy Savage. This follows the pattern of **deductive reasoning**.

Abductive Reasoning

Type of reasoning that uses an observation or set of observations to reach a logical conclusion. It is similar to inductive reasoning; however it permits making best guesses to arrive at the simplest conclusions. It is a form of logical reasoning that involves making an educated guess or hypothesis based on incomplete or limited information. It involves:

1. Observing a phenomenon or pattern
2. Identifying possible explanations
3. Selecting the most plausible explanation
4. Testing and refining the hypothesis

Abductive reasoning is essential in mathematics, science, and problem-solving.

26. What is the next number in the sequence: 2, 4, 8, 16, ?

Solution. 32 (recognizing a geometric progression).

27. A bakery sells 250 loaves of bread per day. If each loaf costs $2, how much money does the bakery make daily?

Solution: $500 (assuming each loaf sells at the given price).

28. A car travels 250 miles in 5 hours. What is its average speed?

Solution. 50 mph (using distance = rate × time).

29. What is the sum of the interior angles of a triangle?

Solution. 180° (using geometric properties).

30. A survey shows 3/5 of students prefer pizza. If 100 students participated, how many prefer pizza?

Solution: 60 (applying proportionality).

31. Solve for x: $2x + 5 = 11$

Solution. $x = 3$ (using algebraic manipulation).

32. A rectangle has a perimeter of 24 cm. If its length is 8 cm, what is its width?
Solution: 4 cm (using perimeter = 2(length + width)).

33. What is the probability of rolling a 6 on a fair six-sided die?
Solution. 1/6 (using probability theory).

34. A water tank fills at 0.5 liters/minute. How long to fill a 30-liter tank?
Solution. 60 minutes (using rate × time)

35. Find the missing value: 3, 6, 12, ?, 48
Solution. 24 (recognizing a geometric progression).

36. What is the next number in the sequence: 1, 2, 4, 7, 11, ?
Solution. 16 (recognizing a quadratic progression).

37. A snail moves 3 cm/hour. How far will it move in 5 hours?
Solution. 15 cm (using rate × time).

38. Solve for x: $x^2 + 5x - 6 = 0$
Solution. $x = -6$ or $x = 1$ (using quadratic formula).

39. A circle has a circumference of 20π cm. What is its radius?
Solution. 10 cm (using circumference = $2\pi r$).

40. What is the sum of the exterior angles of a polygon?
Solution. 360° (using geometric properties).

41. A survey shows 2/3 of students prefer math. If 150 students participated, how many prefer math?
Solution. 100 (applying proportionality).

42. Find the missing value: 2, 6, 12, 20, ?
Solution. 30 (recognizing a quadratic progression).

43. A cylinder has a volume of 40π cm^3. If its height is 10 cm, what is its radius?
Solution. 2 cm (using volume = $\pi r^2 h$).

44. What is the probability of drawing an ace from a standard deck of cards?
Solution. 4/52 = 1/13 (using probability theory).

45. Solve for x: $3x - 2 = 14$
Solution. x = 16/3 (using algebraic manipulation).

46. Explain the concept of abductive reasoning and its significance in mathematical problem-solving.

Solution.

Abductive reasoning involves making educated guesses or hypotheses based on incomplete information. In mathematics, it's essential for:

1. Pattern recognition: Identifying relationships between numbers or shapes.
2. Hypothesis formation: Proposing solutions to problems.
3. Logical inference: Drawing conclusions from available data.

Abductive reasoning facilitates mathematical discovery, fosters critical thinking, and enhances problem-solving skills.

47. Discuss how abductive reasoning differs from deductive and inductive reasoning.

Solution.

Abductive reasoning differs from:

1. Deductive reasoning: Abductive reasoning involves uncertainty, whereas deductive reasoning draws definitive conclusions.
2. Inductive reasoning: Abductive reasoning focuses on hypothesis formation, whereas inductive reasoning seeks generalizations.

Abductive reasoning bridges the gap between deductive and inductive reasoning, enabling mathematicians to navigate uncertainty.

48. Describe a real-world scenario where abductive reasoning is essential.

Solution.

Medical diagnosis: Doctors use abductive reasoning to:

1. Identify symptoms
2. Formulate hypotheses
3. Test and refine diagnoses

Abductive reasoning enables doctors to make informed decisions amidst uncertainty, saving lives.

49. Analyze the role of abductive reasoning in resolving mathematical paradoxes.

Solution.

Abductive reasoning helps resolve paradoxes by:

1. Identifying underlying assumptions
2. Formulating alternative hypotheses
3. Evaluating evidence

Examples: Russell's Paradox, the Liar Paradox. Abductive reasoning facilitates deeper understanding and resolution.

Exercise

1. Is the following statement inductive or deductive reasoning? " All Noble prize winners get a monetary award. Jennifer Doudna won a Noble Prize, so she must have received money."
2. Is the following statement inductive or deductive reasoning? " My friend and my brother graduated from Harvard and immediately got great jobs. Therefore, everyone who graduates from Harvard will immediately get a great job. "
3. Find a counter example to disprove the hypothesis: If two even numbers are divided, the quotient is a whole number.
4. Find a counter example to disprove the hypothesis: If a number is added to itself, the sum is greater than the original number.
5. Describe the pattern found in the following sequence of numbers and then find in the next two values: 1, 2, 4, 7, 11, 16.
6. Describe the pattern found in the following sequence of days and then find in the next two values: Monday, Thursday, Sunday, Wednesday, Saturday.
7. In Exercises i–x, use inductive reasoning to predict the most probable next number in each list.
    - i. 4, 8, 12, 16, 20, 24, ?
    - ii. 5, 11, 17, 23, 29, 35, ?
    - iii. 3, 5, 9, 15, 23, 33, ?
    - iv. 1, 8, 27, 64, 125, ?
    - v. 1, 4, 9, 16, 25, 36, 49, ?
    - vi. 80, 70, 61, 53, 46, 40, ?
    - vii. $\frac{3}{5}, \frac{5}{7}, \frac{7}{9}, \frac{9}{11}, \frac{11}{13}, \frac{13}{15}$, ?
    - viii. $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}$, ?
    - ix. $2, 7, -3, 2, -8, -3, -13, -8, -18$, ?
    - x. 1, 5, 12, 22, 35, ?
8. Determine whether the argument is an example of inductive reasoning or deductive reasoning.
    - i. Andrea enjoyed reading the Dark Tower series by Stephen King, so I know she will like his next novel.
    - ii. All pentagons have exactly five sides. Figure A is a pentagon. Therefore, Figure A has exactly five sides.
    - iii. Every English setter likes to hunt. Duke is an English setter, so Duke likes to hunt.
    - iv. Cats don’t eat tomatoes. Scat is a cat. Therefore, Scat does not eat tomatoes.
    - v. A number is a “neat” number if the sum of the cubes of its digits equals the number. Therefore, 153 is a “neat” number.
    - vi. The Atlanta Braves have won five games in a row.

9. Convert the decimal expansion of each of these integers to a binary expansion.

a) 231 b) 4532 c) 97644

10. Convert the binary expansion of each of these integers to a decimal expansion.

a) $(1\ 1111)_2$ b) $(10\ 0000\ 0001)_2$

c) $(1\ 0101\ 0101)_2$ d) $(110\ 1001\ 0001\ 0000)_2$

11. Convert the octal expansion of each of these integers to a binary expansion.

a) $(572)_8$ b) $(1604)_8$

c) $(423)_8$ d) $(2417)_8$

12. Convert the binary expansion of each of these integers to an octal expansion.

a) $(1111\ 0111)_2$

b) $(1010\ 1010\ 1010)_2$

c) $(111\ 0111\ 0111\ 0111)_2$

d) $(101\ 0101\ 0101\ 0101)_2$

13. Convert the hexadecimal expansion of each of these integers to a binary expansion.

a) $(80E)_{16}$ b) $(135AB)_{16}$

c) $(ABBA)_{16}$ d) $(DEFACED)_{16}$

CHAPTER

2

LOGIC

AND ITS APPLICATION

In today's complex world, it is not so easy to summarize the topic of logic. For lawyers and business people, logic is the science of correct reasoning. They often use logic to construct valid arguments, analyze legal contracts, and solve complicated problems. The principles of logic can also be used as a production tool. For example, programmers use logic to design computer software, engineers use logic to design the electronic circuits in computers, and mathematicians use logic to solve problems and construct mathematical proofs.

In this chapter we will learn about;

- Logical reasoning, importance and their application in modern age
- Propositions, prepositional variables
- Induction, deduction, logical connectives
- Compound statements, truth values, truth tables and applications
- Negation, conjunction, disjunction
- Tautologies, contradiction and contingencies, equivalent statements
- The conditional and related statements
- Equivalent statements and logical equivalence
- Truth sets, a link between set theory and logic
- Arguments and their validity
- Categorical prepositions and Venn diagram
- Categorical propositions and Euler diagrams
- Quantifiers and predicates
- Logical fallacies and its types

The rules of logic give precise meaning to mathematical statements. These rules are used to distinguish between valid and invalid mathematical arguments. Because a major goal of this chapter is to teach the reader how to understand and how to construct correct mathematical arguments, we begin our study of discrete mathematics with an introduction to logic. Besides the importance of logic in understanding mathematical reasoning, logic has numerous applications to computer science. These rules are used in the design of computer circuits, the construction of computer programs, the verification of the correctness of programs, and in many other ways. Furthermore, software systems have been developed for constructing some, but not all, types of proofs automatically. We will discuss these applications of logic in this and later chapters.

Logic: Logic is the discipline that deals with the method of reasoning.

In this chapter, we will encounter several facets of logic. Specifically, we will use logic to

- Analyze information and relationships between statements.
- Determine the validity of arguments and theorems.
- Determine valid conclusions based on given assumptions, and
- Analyze electronic circuits.
- In computer science to verify the correctness of programs.
- In physical science to draw conclusion from experiments.

Propositions/Statements

A proposition/statement is a declarative sentence that is either true or false, but not both simultaneously.

It may not be necessary to determine whether a sentence is true or false to determine whether it is a statement. For instance, the following sentence is either true or false:

Every even number greater than 2 can be written as the sum of two prime numbers.

At this time mathematicians have not determined whether the sentence is true or false, but they do know that it is either true or false and that it is not both true and false. Thus the sentence is a statement.

Arguments: Arguments are a set of statements.

Premises: The premises are the statements being offered in support for the conclusion.

Conclusion: The statements being argued for.

Proportional Variables: The letters $(p, q, r, \ldots)$ that can be replaced by statements are called proportional variables. e.g. p: Today is Friday, q: It is raining, r: I am going to a movie.

1. Determine whether each sentence is a statement.
   a. Florida is a state in the United States.
   b. The word dog has four letters.
   c. How are you?
   d. $9^{(9^9)} + 2$ is a prime number.
   e. $x + 1 = 5$

Solution.

a. Florida is one of the 50 states in the United States, so this sentence is true and it is **a statement**.
b. The word dog consists of exactly three letters, so this sentence is false and it is **a statement**.
c. The sentence "How are you?" is a question; it is not a declarative sentence. Thus it is **not a statement**.
d. You may not know whether $9^{(9^9)} + 2$ is a prime number; however, you do know that it is a whole number larger than 1, so it is either a prime number or it is not a prime number. The sentence is either true or false, and it is not both true and false simultaneously, so it is **a statement**.
e. $x + 1 = 5$ is a statement. It is known as **an open statement**. It is true for $x = 4$ and it is false for any other value of x. For any given value of x, it is true or false but not both.

2. Determine whether each sentence is a statement.
   a. Open the door.
   b. 7055 is a large number.
   c. $4 + 5 = 8$
   d. In the year 2019, the president of the United States will be a woman.
   e. $x > 3$

   Solution.

   a. The sentence "Open the door" is a command. It is not a statement.
   b. The word large is not a precise term. It is not possible to determine whether the sentence "7055 is a large number" is true or false and thus the sentence is not a statement.
   c. The sentence $4 + 5 = 8$ is a false statement.
   d. At this time we do not know whether the given sentence is true or false, but we know that the sentence is either true or false and that it is not both true and false. Thus the sentence is a statement.
   e. The sentence $x > 3$ is a statement because for any given value of x, the inequality $x > 3$ is true or false, but not both.

Induction

To draw general conclusion from limited number of observations or experiences is called induction.

e.g. A person gets penicillin injection once or twice and experiences reaction soon afterwards. He generalizes that he is allergic to penicillin.

سر درد کی دوائی ڈاکٹر کے کہنے پر لینا

deduction

سر درد کی دوائی وکیل کے کہنے پر لینا

induction

Deduction

To draw general conclusion from well known facts is called deduction.

e.g. All men are mortal. We are men. Therefore, we are all mortal.

Logical Connectives

Symbols that are used to combine statements or proportional variables.

George Boole used symbols such as p, q, r, and s to represent statements and the symbols $\wedge, \vee, \sim, \rightarrow$ and $\leftrightarrow$ to represent connectives. See Table,

Original Statement	Connective	Statement in Symbolic Form	Type of Compound Statement
not *p*	not	$\sim p$	negation
p and *q*	and	$p \wedge q$	conjunction
p or *q*	or	$p \vee q$	disjunction
If *p*, then *q*	If ... then	$p \rightarrow q$	conditional
p if and only if *q*	if and only if	$p \leftrightarrow q$	biconditional

Compound Statements

Two or more sentences are connected to form a compound statement. Connecting statements with words and phrases such as and, or, not, if ... then, and if and only if creates a **compound statement**. For instance, “I will attend the meeting or I will go to school” is a **compound statement**. It is composed of the two component statements “I will attend the meeting” and “I will go to school.” The word **or** is a **connective** for the two component statements.

Truth Value and Truth Tables

The **truth value** of a statement is true (T) if the statement is true and false (F) if the statement is false. A **truth table** is a table that shows the truth values of a statement for all possible truth values of its components.

Negation

The negation of the statement "Today is Friday" is the statement "Today is not Friday." In symbolic logic, the tilde symbol ~ is used to denote the negation of a statement. If a statement p is true, its negation ~ p is false, and if a statement p is false, its negation ~ p is true. See the table. The negation of the negation of a statement is the original statement. Thus, ~ (~ p) can be replaced by p in any statement.

The Truth Table for ~ p

p	$\sim p$
T	F
F	T

3. Write the negation of each statement.

a. Bill Gates has a yacht.
b. The number 10 is a prime number.
c. The Dolphins lost the game.

Solution.

a. Bill Gates does not have a yacht.
b. The number 10 is not a prime number.
c. The Dolphins did not lose the game.

4. Write the negation of each statement.

a. 1001 is divisible by 7.
b. 5 is an even number.
c. The fire engine is not red.

Solution.

a. 1001 is not divisible by 7.
b. 5 is not an even number.
c. The fire engine is red.

Remark

The notation for the negation operator is not standardized. Although ¬p and $\bar{p}$ are the most common notations used in mathematics to express the negation of p, other notations you might see are ∼p, −p, p′ ,Np, and !p.

5. Find the negation of the proposition
   "Michael's PC runs Linux"
   and express this in simple English.
   Solution.
   The negation is "It is not the case that Michael's PC runs Linux."
   This negation can be more simply expressed as
   "Michael's PC does not run Linux."

6. Find the negation of the proposition
   "Vandana's smartphone has at least 32 GB of memory"
   and express this in simple English.
   Solution.
   The negation is
   "It is not the case that Vandana's smartphone has at least 32 GB of memory."
   This negation can also be expressed as
   "Vandana's smartphone does not have at least 32 GB of memory"
   or even more simply as
   "Vandana's smartphone has less than 32 GB of memory."

7. Consider the following statements.

p: Today is Friday.
q: It is raining.
r: I am going to a movie.
s: I am not going to the basketball game.

Write the following compound statements in symbolic form.

a. Today is Friday and it is raining.
b. It is not raining and I am going to a movie.
c. I am going to the basketball game or I am going to a movie.
d. If it is raining, then I am not going to the basketball game.

Solution.

a. $p \wedge q$ **b.** $\sim q \wedge r$ **c.** $\sim s \vee r$ **d.** $q \rightarrow s$

8. Consider the following statements.

p: Today is Friday.
q: It is raining.
r: I am going to a movie.
s: I am not going to the basketball game.

Write the following compound statements in symbolic form.

a. Today is not Friday and I am going to a movie.
b. I am going to the basketball game and I am not going to a movie.
c. I am going to a movie if and only if it is raining.
d. If today is Friday, then I am not going to a movie

Solution.

a. $\sim p \wedge r$ **b.** $\sim s \wedge \sim r$ **c.** $r \leftrightarrow q$ **d.** $p \rightarrow \sim r$

9. Consider the following statements.

p: The game will be played in Atlanta.
q: The game will be shown on CBS.
r: The game will not be shown on ESPN.
s: The Dodgers are favored to win.

Write each of the following symbolic statements in words.

a. $q \wedge p$ **b.** $\sim r \wedge s$ **c.** $s \leftrightarrow \sim p$

Solution.

a. The game will be shown on CBS and the game will be played in Atlanta.
b. The game will be shown on ESPN and the Dodgers are favored to win.
c. The Dodgers are favored to win if and only if the game will not be played in Atlanta.

Conjunction/ The logic of AND

The conjunction of two statements p and q is true if and only if both p and q are true otherwise false. It is denoted by $p \wedge q$ and read as p and q.

Disjunction/ The logic of OR

The disjunction of two statements p and q is false if and only if both p and q are false otherwise true. It is denoted by $p \vee q$ and read as p or q. The disjunction $p \vee q$ is true if p is true, if q is true, or if both p and q are true.

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

10. Determine whether each statement is true or false.

a. $7 \geq 5$

b. 5 is a whole number and 5 is an even number.

c. 2 is a prime number and 2 is an even number.

Solution.

a. $7 \geq 5$ means $7 > 5$ or $7 = 5$. Because $7 > 5$ is true, the statement $7 \geq 5$ is a true statement.

b. This is a false statement because 5 is not an even number.

c. This is a true statement because each component statement is true.

11. Determine whether each statement is true or false.

a. 21 is a rational number and 21 is a natural number.

b. $4 \leq 9$

c. $-7 \geq -3$

Solution.

a. True. A conjunction is true provided both components are true.

b. True. A disjunction is true provided at least one component is true.

c. False. If both components of a disjunction are false, then the disjunction is false.

12. Find the conjunction of the propositions p and q where p is the proposition "Rebecca's PC has more than 16 GB free hard disk space" and q is the proposition "The processor in Rebecca's PC runs faster than 1 GHz."

Solution.

The conjunction of these propositions, $p \wedge q$, is the proposition "Rebecca's PC has more than 16 GB free hard disk space, and the processor in Rebecca's PC runs faster than 1 GHz." This conjunction can be expressed more simply as "Rebecca's PC has more than 16 GB free hard disk space, and its processor runs faster than 1 GHz." For this conjunction to be true, both conditions given must be true. It is false when one or both of these conditions are false.

13. Translate the statement "Students who have taken calculus or introductory computer science can take this class" in a statement in propositional logic using the propositions p: "A student who has taken calculus can take this class" and q: "A student who has taken introductory computer science can take this class."

Solution.

We assume that this statement means that students who have taken both calculus and introductory computer science can take the class, as well as the students who have taken only one of the two subjects. Hence, this statement can be expressed as $p \vee q$, the inclusive or, or disjunction, of p and q.

14. What is the disjunction of the propositions p and q, where p is the proposition "Rebecca's PC has more than 16 GB free hard disk space" and q is the proposition "The processor in Rebecca's PC runs faster than 1 GHz."

Solution.

The disjunction of p and q, $p \vee q$, is the proposition
"Rebecca's PC has at least 16 GB free hard disk space, or the processor in Rebecca's PC runs faster than 1 GHz."
This proposition is true when Rebecca's PC has at least 16 GB free hard disk space, when the PC's processor runs faster than 1 GHz, and when both conditions are true. It is false when both of these conditions are false, that is, when Rebecca's PC has less than 16 GB free hard disk space and the processor in her PC runs at 1 GHz or slower.

15. Write the negation of each of the following statements.

a. Some baseball players are worth a million dollars.
b. All movies are worth the price of admission.
c. No odd numbers are divisible by 2.

Solution.

a. No baseball player is worth a million dollars.
b. Some movies are not worth the price of admission.
c. Some odd numbers are divisible by 2

16. Write the negations of the following statements.

a. All bears are brown.
b. No math class is fun.
c. Some vegetables are not green.

Solution.

a. Some bears are not brown.
b. Some math classes are fun.
c. All vegetables are green.

Tautologies

A tautology is a statement that is always true. For example $p \to q \leftrightarrow \sim q \to \sim p$

Contradiction / Self-Contradiction / Absurdity

A self-contradiction is a statement that is always false. For example $p \to \sim p$

Contingency

A statement which can be true or false depending upon the truth values of the variables involved in it is called a contingency e.g., $(p \to q) \wedge (p \vee q)$ is a contingency.

17. Is the statement $x + 2 = 5$ a tautology or a self-contradiction?

Answer.

Neither. The statement is not true for all values of x, and it is not false for all values of x.

18. Show that $p \vee (\sim p \vee q)$ is a tautology.

Solution.

The table shows that $p \vee (\sim p \vee q)$ is always true. Thus $p \vee (\sim p \vee q)$ is a tautology.

p	q	p	∨	(~p	∨	q)
T	T	T	T	F	T	T
T	F	T	T	F	F	F
F	T	F	T	T	T	T
F	F	F	T	T	T	F

19. Show that $p \wedge (\sim p \wedge q)$ is a self-contradiction.

Solution.

p	q	p	∧	(~p	∧	q)
T	T	T	F	F	F	T
T	F	T	F	F	F	F
F	T	F	F	T	T	T
F	F	F	F	T	F	F

Table shows that $p \wedge (\sim p \wedge q)$ is always false. Thus $p \wedge (\sim p \wedge q)$ is a self-contradiction.

20. Construct a table for ∼ (∼ p ∨ q) ∨ q.

Use the truth table from part a to determine the truth value of ∼ (∼ p ∨ q) ∨ q, given that p is true and q is false.

Solution.

p	q	$\sim p$	$\sim p \vee q$	$\sim(\sim p \vee q)$	$\sim(\sim p \vee q) \vee q$
T	T	F	T	F	T
T	F	F	F	T	T
F	T	T	T	F	T
F	F	T	T	F	F

In row 2 of the above truth table, we see that when p is true and q is false, the statement ∼ (∼ p ∨ q) ∨ q in the rightmost column is true.

21. Construct a truth table for (p ∧∼ q) ∨ (∼ p ∨ q).

Use the truth table that you constructed in part a to determine the truth value Of (p ∧∼ q) ∨ (∼ p ∨ q) given that p is true and q is false.

Solution.

p	q	$\sim p$	$\sim q$	$p \wedge \sim q$	$\sim p \vee q$	$(p \wedge \sim q) \vee (\sim p \vee q)$
T	T	F	F	F	T	T
T	F	F	T	T	F	T
F	T	T	F	F	T	T
F	F	T	T	F	T	T

p is true and q is false in row 2 of the above truth table. The truth value of
(p ∧∼ q) ∨ (∼ p ∨ q)
in row 2 is T (true).

22. Construct a table for (p ∧ q) ∧ (∼ r ∨ q).

Use the truth table that you constructed in part a to determine the truth value of (p ∧ q) ∧ (∼ r ∨ q), given that p is true, q is true and r is false.

Solution.

p	q	r	p ∧ q	∼r	∼r ∨ q	(p ∧ q) ∧ (∼r ∨ q)
T	T	T	T	F	T	T
T	T	F	T	T	T	T
T	F	T	F	F	F	F
T	F	F	F	T	T	F
F	T	T	F	F	T	F
F	T	F	F	T	T	F
F	F	T	F	F	F	F
F	F	F	F	T	T	F

In row 2 of the above truth table, we see that (p ∧ q) ∧ (∼ r ∨ q) is true when p is true, q is true, and r is false.

23. Construct a table for (∼ p ∧ r) ∨ (q ∧∼ r).

Use the truth table that you constructed in part a to determine the truth value Of (∼ p ∧ r) ∨ (q ∧∼ r) given that p is false, q is true, and r is false.

Solution.

p	q	r	∼p	∼r	∼p ∧ r	q ∧ ∼r	(∼p ∧ r) ∨ (q ∧ ∼r)	
T	T	T	F	F	F	F	F	Row 1
T	T	F	F	T	F	T	T	Row 2
T	F	T	F	F	F	F	F	Row 3
T	F	F	F	T	F	F	F	Row 4
F	T	T	T	F	T	F	T	Row 5
F	T	F	T	T	F	T	T	Row 6
F	F	T	T	F	T	F	T	Row 7
F	F	F	T	T	F	F	F	Row 8

p is false, q is true, and r is false in row 6 of the above truth table. The truth value of (∼ p ∧ r) ∨ (q ∧∼ r) in row 6 is T (true).

24. Construct a truth table for p ∨ [∼ (p ∧∼ q)].

Solution.

The given statement p ∨ [∼ (p ∧∼ q)] has the two simple statements p and q. Thus we start with a standard form that has $2^2 = 4$ rows.

In each column, enter the truth values for the statements p and ∼ q as show in the columns numbered 1, 2, and 3 of the following table. Use the truth values in columns 2 and 3 to determine the truth values to enter under the "and" connective. See the column numbered 4. Now negate the truth values in the column numbered 4 to produce the truth values in the column numbered 5. Use the truth values in the columns numbered 1 and 5 to determine the truth values to enter under the "or" connective. See the column numbered 6, which is the truth table for given.

p	*q*	*p*	∨	[∼	(*p*	∧	∼*q*)]
T	T	T	T	T	T	F	F
T	F	T	T	F	T	T	T
F	T	F	T	T	F	F	F
F	F	F	T	T	F	F	T

25. Construct a truth table for ∼ p ∨ (p ∧ q).

Solution.

The given statement has two simple statements. Thus you should use a standard form that has $2^2 = 4$ rows. **Step 1:** Enter the truth values for each simple statement and their negations. See columns 1, 2, and 3 in the table following step 3. **Step 2:** Use the truth values in columns 2 and 3 to determine the truth values to enter under the "and" connective. See column 4 in the table following step 3. **Step 3:** Use the truth values in columns 1 and 4 to determine the truth values to enter under the "or" connective. See column 5 in the table below.

p	*q*	∼*p*	∨	(*p*	∧	*q*)
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	T	F	F	T
F	F	T	T	F	F	F
		1	5	2	4	3

The truth table for ∼ p ∨ (p ∧ q) is displayed in column 5.

Implication or Conditional Statement

A compound statement of the form if p then q, also written **p implies q**, is called a **conditional** or an **implication**, denoted by **p → q**. p is called the **antecedent** or **hypothesis** and q is called the **consequent** or the **conclusion**.

For instance, in the conditional statement,

If our school was this nice, I would go there more than once a week, the antecedent is "our school was this nice" and the consequent is "I would go there more than once a week."

Conditional statements can be written in if p, then q form or in if p, q form. For instance, all of the following are conditional statements.

- If we order pizza, then we can have it delivered.
- If you go to the movie, you will not be able to meet us for dinner.
- If n is a prime number greater than 2, then n is an odd number.

Truth Value of the Conditional

The conditional is false if p is true and q is false. It is true in all other cases.

A conditional is regarded as false only when the antecedent is true and consequent is false. In all other cases it is considered to be true. Its truth table is, therefore, of the adjoining form;

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

26. Determine the truth value of each of the following.

a. If 2 is an integer, then 2 is a rational number.

b. If 3 is a negative number, then 5 > 7.

c. If 5 > 7 then 2 + 7 = 4.

Solution.

a. Because the consequent is true, this is a true statement.

b. Because the antecedent is false, this is a true statement.

c. Because the antecedent is true and the consequent is false, this is a false statement.

27. Determine the truth value of each of the following.

a. If $4 \geq 3$ then $2 + 5 = 6$

b. If $5 > 9$ then $4 > 9$

c. If Tuesday follows Monday, then April follows March.

Solution.

a. Because the antecedent is true and the consequent is false, the statement is a false statement.

b. Because the antecedent is false, the statement is a true statement.

c. Because the consequent is true, the statement is a true statement.

28. Construct a Truth Table for a Statement Involving a Conditional

$[p \wedge (q \vee \sim p)] \rightarrow \sim p$

Solution.

p	*q*	[*p*	∧	(*q*	∨	~*p*)]	→	~*p*
T	T	T	T	T	T	F	F	F
T	F	T	F	F	F	F	T	F
F	T	F	F	T	T	T	T	T
F	F	F	F	F	T	T	T	T

29. Construct a Truth Table for a Statement Involving a Conditional

$[p \wedge (p \rightarrow q)] \rightarrow q$

Solution.

p	*q*	[*p*	∧	(*p*	→	*q*)]	→	*q*
T	T	T	T	T	T	T	T	T
T	F	T	F	T	F	F	T	F
F	T	F	F	F	T	T	T	T
F	F	F	F	F	T	F	T	F
		1	6	2	5	3	7	4

30. Let p be the statement "Maria learns discrete mathematics" and q the statement "Maria will find a good job." Express the statement $p \rightarrow q$ as a statement in English.

Solution.

From the definition of conditional statements, we see that when p is the statement "Maria learns discrete mathematics" and q is the statement

"Maria will find a good job," $p \rightarrow q$ represents the statement

"If Maria learns discrete mathematics, then she will find a good job."

There are many other ways to express this conditional statement in English.

Among the most natural of these are

"Maria will find a good job when she learns discrete mathematics."

"For Maria to get a good job, it is sufficient for her to learn discrete mathematics."

And

"Maria will find a good job unless she does not learn discrete mathematics."

31. What is the value of the variable x after the statement

if $2 + 2 = 4$ then $x := x + 1$

if $x = 0$ before this statement is encountered?

(The symbol := stands for assignment.

The statement $x := x + 1$ means the assignment of the value of $x + 1$to x.)

Solution.

Because $2 + 2 = 4$ is true, the assignment statement $x := x + 1$ is executed.

Hence, x has the value $0 + 1 = 1$ after this statement is encountered.

Equivalent Forms of the Conditional

Every conditional statement can be stated in many equivalent forms. It is not even necessary to state the antecedent before the consequent. For instance, the conditional "If I live in Boston, then I must live in Massachusetts" can also be stated as

I must live in Massachusetts, if I live in Boston.

Table lists some of the various forms that may be used to write a conditional statement.

Common Forms of p → q

Every conditional statement $p \rightarrow q$ can be written in the following equivalent forms.	
If *p*, then *q*.	Every *p* is a *q*.
If *p*, *q*.	*q*, if *p*.
p only if *q*.	*q* provided *p*.
p implies *q*.	*q* is a necessary condition for *p*.
Not *p* or *q*.	*p* is a sufficient condition for *q*.

32. Write a Statement in an Equivalent Form

Write each of the following in "If p, then q" form.

a. The number is an even number provided it is divisible by 2.

b. Today is Friday, only if yesterday was Thursday.

Solution.

a. The statement "The number is an even number provided it is divisible by 2" is in "q provided p" form. The antecedent is "it is divisible by 2," and the consequent is "the number is an even number." Thus its "If p, then q" form is If it is divisible by 2, then the number is an even number.

b. The statement "Today is Friday, only if yesterday was Thursday" is in "p only if q" form. The antecedent is "today is Friday." The consequent is "yesterday was Thursday." Its "If p, then q" form is

If today is Friday, then yesterday was Thursday

33. Write each of the following in "If p, then q" form.

a. Every square is a rectangle.

b. Being older than 30 is sufficient to show I am at least 21.

Solution.

a. If it is a square, then it is a rectangle.

b. If I am older than 30, then I am at least 21.

The Converse, the Inverse, and the Contrapositive

Every conditional statement has three related statements. They are called the converse, the inverse, and the contrapositive.

The converse of $\mathbf{p \to q}$ is $\mathbf{q \to p}$

The inverse of $\mathbf{p \to q}$ is $\mathbf{\sim p \to \sim q}$

The contrapositive of $\mathbf{p \to q}$ is $\mathbf{\sim q \to \sim p}$

The above definitions show the following:

- The converse of $p \to q$ is formed by interchanging the antecedent p with the consequent q.
- The inverse of $p \to q$ is formed by negating the antecedent p and negating the consequent q.
- The contrapositive of $p \to q$ is formed by negating both the antecedent p and the consequent q and interchanging these negated statements.

p	q	Conditional $p \to q$	Converse $q \to p$	Inverse $\sim p \to \sim q$	Contrapositive $\sim q \to \sim p$
T	T	T	T	T	T
T	F	F	T	T	F
F	T	T	F	F	T
F	F	T	T	T	T

34. Write the converse, inverse, and contrapositive of

If I get the job, then I will rent the apartment.

Solution.

Converse: If I rent the apartment, then I get the job.

Inverse: If I do not get the job, then I will not rent the apartment.

Contrapositive: If I do not rent the apartment, then I did not get the job.

35. Write the converse, inverse, and contrapositive of

If we have a quiz today, then we will not have a quiz tomorrow.

Solution.

Converse: If we are not going to have a quiz tomorrow, then we will have a quiz today.

Inverse: If we don’t have a quiz today, then we will have a quiz tomorrow.

Contrapositive: If we have a quiz tomorrow, then we will not have a quiz today.

36. Determine whether the given statements are equivalent.

a. If a number ends with a 5, then the number is divisible by 5.
If a number is divisible by 5, then the number ends with a 5.

b. If two lines in a plane do not intersect, then the lines are parallel.
If two lines in a plane are not parallel, then the lines intersect

Solution.

a. The second statement is the converse of the first. The statements are not equivalent.

b. The second statement is the contrapositive of the first. The statements are equivalent.

37. Determine whether the given statements are equivalent.

a. If $a = b$ then $a.c = b.c$
If $a \neq b$ then $a.c \neq b.c$

b. If I live in Nashville, then I live in Tennessee.
If I do not live in Tennessee, then I do not live in Nashville.

Solution.

a. The second statement is the inverse of the first statement. Thus the statements are not equivalent. This can also be demonstrated by the fact that the first statement is true for c = 0 and the second statement is false for c = 0.

b. The second statement is the contrapositive of the first statement. Thus the statements are equivalent.

Remark

In mathematics, it is often necessary to prove statements that are in "If p, then q" form. If a proof cannot readily be produced, mathematicians often try to prove the contrapositive, "If $\sim q$ then $\sim p$" Because a conditional and its contrapositive are equivalent statements, a proof of either statement also establishes the proof of the other statement.

38. A mathematician wishes to prove the following statement about the integer x.
If x^2 is an odd integer, then x is an odd integer. (I)
If the mathematician is able to prove the statement, "If x is an even integer, then x^2 is an even integer," does this also prove statement (I)?

Answer.

Yes, because the second statement is the contrapositive of (I).

39. Write the contrapositive of each statement and use the contrapositive to determine whether the original statement is true or false.

a. If a + b is not divisible by 5, then a and b are not both divisible by 5.
b. If x^3 is an odd integer, then x is an odd integer. (Assume x is an integer.)
c. If a geometric figure is not a rectangle, then it is not a square.

Solution.

a. If a and b are both divisible by 5, then a + b is divisible by 5. This is a true statement, so the original statement is also true.
b. If x is an even integer, then x^3 is an even integer. This is a true statement, so the original statement is also true.
c. If a geometric figure is a square, then it is a rectangle. This is a true statement, so the original statement is also true.

40. Write the contrapositive of each statement and use the contrapositive to determine whether the original statement is true or false.

a. If 3 + x is an odd integer, then x is an even integer. (Assume x is an integer.)
b. If two triangles are not similar triangles, then they are not congruent triangles. Note: Similar triangles have the same shape. Congruent triangles have the same size and shape.
c. If today is not Wednesday, then tomorrow is not Thursday.

Solution.

a. Contrapositive: If x is an odd integer, then 3 + x is an even integer. The contrapositive is true and so the original statement is also true.
b. Contrapositive: If two triangles are congruent triangles, then the two triangles are similar triangles. The contrapositive is true and so the original statement is also true.
c. Contrapositive: If tomorrow is Thursday, then today is Wednesday. The contrapositive is true and so the original statement is also true.

41. Find the contrapositive, the converse, and the inverse of the conditional statement "The home team wins whenever it is raining."

Solution. Because "q whenever p" is one of the ways to express the conditional statement p → q, the original statement can be rewritten as "If it is raining, then the home team wins."

Consequently, the contrapositive of this conditional statement is "If the home team does not win, then it is not raining."

The converse is "If the home team wins, then it is raining."

The inverse is "If it is not raining, then the home team does not win."

Only the contrapositive is equivalent to the original statement.

Biconditional

The statement **p ↔ q** is called a biconditional and is read as "p if and only if q." It is equivalent to the following form;

$$p \leftrightarrow q \equiv [(p \rightarrow q) \wedge (q \rightarrow p)]$$

p	q	$p \leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

42. State whether each biconditional is true or false.

a. $x + 4 = 7$ if and only if $x = 3$.
b. $x^2 = 36$ if and only if $x = -6$.

Solution.

a. Both components are true when $x = 3$ and both are false when $x \neq 3$. Both components have the same truth value for any value of x, so this is a true statement.
b. If $x = -6$ the first component is true and the second component is false. Thus, this is a false statement.

43. State whether each biconditional is true or false.

c. $x > 7$ if and only if $x > 6$.
d. $x + 5 > 7$ if and only if $x > 2$.

Solution.

a. Let x = 6.5. Then the first component of the biconditional is false and the second component of the biconditional is true. Thus the given biconditional statement is false.

b. Both components of the biconditional are true for $x > 2$, and both components are false for $x \leq 2$. Because both components have the same truth value for any real number x, the given biconditional is true.

44. Let p be the statement "You can take the flight," and let q be the statement "You buy a ticket." Then find p ↔ q.

Solution.

Then p ↔ q is the statement

"You can take the flight if and only if you buy a ticket."

This statement is true if p and q are either both true or both false, that is, if you buy a ticket and can take the flight or if you do not buy a ticket and you cannot take the flight. It is false when p and q have opposite truth values, that is, when you do not buy a ticket, but you can take the flight (such as when you get a free trip) and when you buy a ticket but you cannot take the flight (such as when the airline bumps you).

Equivalent Statements

Two statements are equivalent if they both have the same truth values for all possible truth values of their component statements. Equivalent statements have identical truth values in the final columns of their truth. The notation p $\equiv$ q is used to indicate that the statements p and q are equivalent.

1. Show that $\sim(p \vee \sim q)$ and $\sim p \wedge q$ are equivalent statements.
   Solution.
   Construct two truth tables and compare the results. The truth tables below show that $\sim(p \vee \sim q)$ and $\sim p \wedge q$ have the same truth values for all possible truth values of their component statements. Thus the statements are equivalent.

p	q	$\sim(p \vee \sim q)$
T	T	F
T	F	F
F	T	T
F	F	F

p	q	$\sim p \wedge q$
T	T	F
T	F	F
F	T	T
F	F	F

Identical truth values

Thus $\sim(p \vee \sim q) \equiv \sim p \wedge q$.

2. Show that p and $(p \wedge \sim q)$ are equivalent statements.
   Solution.

p	q	p	$\vee$	$(p$	$\wedge$	$\sim q)$
T	T	T	T	T	F	F
T	F	T	T	T	T	T
F	T	F	F	F	F	F
F	F	F	F	F	F	T

This shows that $p \equiv p \wedge \sim q$

De Morgan's Laws for Statements

For any statements p and q,

$\sim(p \vee q) \equiv \sim p \wedge \sim q$ & $\sim(p \wedge q) \equiv \sim p \vee \sim q$

p	q	$\sim(p \vee q)$	$\sim p \wedge \sim q$
T	T	F	F
T	F	F	F
F	T	F	F
F	F	T	T

p	q	$\sim(p \wedge q)$	$\sim p \vee \sim q$
T	T	F	F
T	F	T	T
F	T	T	T
F	F	T	T

3. Use one of De Morgan's laws to restate the following sentence in an equivalent form. *It is not the case that I graduated or I got a job*.

Solution.

Let p represent the statement "I graduated." Let q represent the statement "I got a job." In symbolic form, the original sentence is $\sim(p \vee q)$. One of De Morgan's laws states that this is equivalent to $\sim p \wedge \sim q$. Thus a sentence that is equivalent to the original sentence is "I did not graduate and I did not get a job."

4. Use one of De Morgan's laws to restate the following sentence in an equivalent form. *It is not true that I am going to the dance and I am going to the game*.

Solution.

Let d represent "I am going to the dance." Let g represent "I am going to the game." The original sentence in symbolic form is $\sim(d \wedge g)$, Applying one of De Morgan's laws, we find that $\sim(d \wedge g) \equiv \sim d \vee \sim g$. Thus an equivalent form of "It is not true that I am going to the dance and I am going to the game" is "I am not going to the dance or I am not going to the game."

Logically Equivalent

The compound propositions p and q are called logically equivalent if $p \leftrightarrow q$ is a tautology. The notation $p \equiv q$ denotes that p and q are logically equivalent.

5. Show that $p \rightarrow q$ and $\neg p \vee q$ are logically equivalent. (This is known as **the conditional- disjunction equivalence**.)

   Solution.

   We construct the truth table for these compound propositions in Table. Because the truth values of $\neg p \vee q$ and $p \rightarrow q$ agree, they are logically equivalent.

p	q	$\neg p$	$\neg p \vee q$	$p \rightarrow q$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

6. Show that $p \vee (q \wedge r)$ and $(p \vee q) \wedge (p \vee r)$ are logically equivalent. This is the distributive law of disjunction over conjunction.

   Solution.

   We construct the truth table for these compound propositions in Table. Because the truth values of $p \vee (q \wedge r)$ and $(p \vee q) \wedge (p \vee r)$ agree, these compound propositions are logically equivalent.

p	q	r	$q \wedge r$	$p \vee (q \wedge r)$	$p \vee q$	$p \vee r$	$(p \vee q) \wedge (p \vee r)$
T	T	T	T	T	T	T	T
T	T	F	F	T	T	T	T
T	F	T	F	T	T	T	T
T	F	F	F	T	T	T	T
F	T	T	T	T	T	T	T
F	T	F	F	F	T	F	F
F	F	T	F	F	F	T	F
F	F	F	F	F	F	F	F

Logical Equivalences Involving Conditional Statements.

$$p \to q \equiv \neg p \lor q$$
$$p \to q \equiv \neg q \to \neg p$$
$$p \lor q \equiv \neg p \to q$$
$$p \land q \equiv \neg(p \to \neg q)$$
$$\neg(p \to q) \equiv p \land \neg q$$
$$(p \to q) \land (p \to r) \equiv p \to (q \land r)$$
$$(p \to r) \land (q \to r) \equiv (p \lor q) \to r$$
$$(p \to q) \lor (p \to r) \equiv p \to (q \lor r)$$
$$(p \to r) \lor (q \to r) \equiv (p \land q) \to r$$

Logical Equivalences Involving Biconditional Statements

$$p \leftrightarrow q \equiv (p \to q) \land (q \to p)$$
$$p \leftrightarrow q \equiv \neg p \leftrightarrow \neg q$$
$$p \leftrightarrow q \equiv (p \land q) \lor (\neg p \land \neg q)$$
$$\neg(p \leftrightarrow q) \equiv p \leftrightarrow \neg q$$

Truth Sets, A link between Set Theory and Logic.

Logical propositions p, q etc., are formulae expressed in terms of some variables. For the sake of simplicity and convenience we may assume that they are all expressed in terms of a single variable x where x is a real variable. Thus $p = p(x)$ where, $x \in U$. All those values of x which make the formula p(x) true form a set, say P. Then P is the truth set of p. Similarly, **the truth set**, Q, of q may be defined. We can extend this notion and apply it in other cases.

i) **Truth set of ~p:** Truth set of ~p will evidently consist of those values of the variable for which p is false i.e., they will be members of P', the complement of P.

ii) **p ∨ q:** Truth set of $p \lor q = p(x) \lor q(x)$ consists of those values of the variable for which p(x) is true or q(x) is true or both p(x) and q(x) are true.

Therefore, truth set of p v q will be: $p \lor q = P \cup Q = \{x \mid p(x) \text{ is true or } q(x) \text{ is true}\}$

iii) **p∧q:** Truth set of $p(x) \land q(x)$ will consist of those values of the variable for which both p(x) and q(x) are true. Evidently truth set of

$p \land q = P \cap Q = \{x \mid p(x) \text{ is true} \land q(x) \text{ is true}\}$

iv) **p→q:** We know that p→q is equivalent to ~p ∨ q therefore truth set of p→q will be $P' \cup Q$

v) **p↔q:** We know that p ↔ q means that p and q are simultaneously true or false. Therefore, in this case truth sets of p and q will be the same i.e. P = Q

Note

(1) Evidently truth set of a tautology is the relevant universal set and that of an absurdity is the empty set .

(2) With the help of the above results we can express any logical formula in set- theoretic form and vice versa.

7. Give logical proofs of the following theorems:
$(A \cup B)' = A' \cap B'$

Solution.

The corresponding formula of logic is $\sim (p \vee q) = \sim p \wedge \sim q$

We construct truth table of the two sides.

p	p	$\sim p$	$\sim q$	$p \vee q$	$\sim(p \vee q)$	$\sim p \wedge \sim q$
T	T	F	F	T	F	F
T	F	F	T	T	F	F
F	T	T	F	T	F	F
F	F	T	T	F	T	T

The last two columns of the table establish the equality of the two sides of equation.

8. Give logical proofs of the following theorems:
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Solution.

Logical form of the theorem is

$p \wedge (q \vee r) = (p \wedge q) \vee (p \wedge r)$

We construct the table for the two sides of this equation

p	p	r	$q \vee r$	$p \wedge (q \vee r)$	$p \wedge q$	$p \wedge r$	$(p \wedge q) \vee (p \wedge r)$
T	T	T	T	T	T	T	T
T	T	F	T	T	T	F	T
T	F	T	T	T	F	T	T
T	F	F	F	F	F	F	F
F	T	T	T	F	F	F	F
F	T	F	T	F	F	F	F
F	F	T	T	F	F	F	F
F	F	F	F	F	F	F	F

Comparison of the entries of columns 5 and 8 is sufficient to establish the desired result.

Argument and Valid Argument

An **argument** consists of a set of statements called **premises** and another statement called the **conclusion**. An argument is **valid** if the conclusion is true whenever all the premises are assumed to be true. An argument is **invalid** if it is not a valid argument.

Remember

In the argument about Aristotle, the two premises and the conclusion are shown below. It is customary to place a horizontal line between the premises and the conclusion.

First Premise: If Aristotle was human, then Aristotle was mortal.
Second Premise: Aristotle was human.
Conclusion: Therefore, Aristotle was mortal.

Arguments can be written in symbolic form. For instance, if we let h represent the statement "Aristotle was human" and m represent the statement "Aristotle was mortal," then the argument can be expressed as

$$\begin{array}{l} h \rightarrow m \\ \underline{h \qquad} \\ \therefore m \end{array}$$

The three dots ∴ are a symbol for "therefore."

9. Write the following argument in symbolic form.
The fish is fresh or I will not order it. The fish is fresh. Therefore I will order it.
Solution.
Let *f* represent the statement "The fish is fresh." Let o represent the statement "I will order it." The symbolic form of the argument is

$$\begin{array}{l} f \vee \sim o \\ \underline{f \qquad} \\ \therefore o \end{array}$$

10. Write the following argument in symbolic form.
If she doesn't get on the plane, she will regret it. She does not regret it.
Therefore, she got on the plane
Solution.
Let p represent the statement "She got on the plane." Let r represent the statement "She will regret it." Then the symbolic form of the argument is

$$\begin{array}{l} \sim p \rightarrow r \\ \underline{\sim r \qquad} \\ \therefore p \end{array}$$

Arguments and Truth Tables

The following truth table procedure can be used to determine whether an argument is valid or invalid.

Truth Table Procedure to Determine the Validity of an Argument

1. Write the argument in symbolic form.
2. Construct a truth table that shows the truth value of each premise and the truth value of the conclusion for all combinations of truth values of the component statements.
3. If the conclusion is true in every row of the truth table in which all the premises are true, the argument is valid. If the conclusion is false in any row in which all the premises are true, the argument is invalid.

11. If Aristotle was human, then Aristotle was mortal. Aristotle was human. Therefore, Aristotle was mortal. Using procedure show that argument is valid or not?

Solution.

Let h represent the statement "Aristotle was human" and m represent the statement "Aristotle was mortal." In symbolic form, the argument is

$h \rightarrow m$	First premise
h	Second premise
$\therefore m$	Conclusion

Construct a truth table as shown below.

h	m	First Premise $h \rightarrow m$	Second Premise h	Conclusion m	
T	T	T	T	T	Row 1
~~T~~	~~F~~	~~F~~	~~T~~	~~F~~	~~Row 2~~
~~F~~	~~T~~	~~T~~	~~F~~	~~T~~	~~Row 3~~
~~F~~	~~F~~	~~T~~	~~F~~	~~F~~	~~Row 4~~

Row 1 is the only row in which all the premises are true, so it is the only row that we examine. Because the conclusion is true in row 1, the argument is valid.

12. Determine whether the following argument is valid or invalid.

If it rains, then the game will not be played. It is not raining. Therefore, the game will be played.

Solution.

If we let r represent "it rains" and g represent "the game will be played," then the symbolic form is

$$\begin{array}{l} r \rightarrow \sim g \\ \underline{\sim r \quad\quad} \\ \therefore g \end{array}$$

The truth table for this argument is as follows:

r	g	First Premise $r \rightarrow \sim g$	Second Premise $\sim r$	Conclusion g	
~~T~~	~~T~~	~~F~~	~~F~~	~~T~~	~~Row 1~~
~~T~~	~~F~~	~~T~~	~~F~~	~~F~~	~~Row 2~~
F	T	T	T	T	Row 3
F	F	T	T	F	Row 4

Because the conclusion in row 4 is false and the premises are both true, we know the argument is invalid.

13. Why do we need to examine only rows 3 and 4?

Solution.

Rows 3 and 4 are the only rows in which all of the premises are true.

14. Determine whether the following argument is valid or invalid.

If the stock market rises, then the bond market will fall.

<u>The bond market did not fall.</u>

∴ The stock market did not rise.

Solution.

Let r represent the statement "The stock market rises." Let f represent the statement "The bond market will fall." Then the symbolic form of the argument is

$$\begin{array}{l} r \rightarrow f \\ \underline{\sim f \quad} \\ \therefore \sim r \end{array}$$

The truth table for this argument is as follows:

		First Premise	Second Premise	Conclusion	
r	f	$r \to f$	$\sim f$	$\sim r$	
T	T	T	F	F	Row 1
T	F	F	T	F	Row 2
F	T	T	F	T	Row 3
F	F	T	T	T	Row 4

Row 4 is the only row in which all the premises are true, so it is the only row that we examine.
Because the conclusion is true in row 4, the argument is valid.

15. Determine whether the following argument is valid or invalid.
If I am going to run the marathon, then I will buy new shoes.
<u>If I buy new shoes, then I will not buy a television.</u>
∴ If I buy a television, I will not run the marathon.

Solution.
Label the statements
m: I am going to run the marathon.
s: I will buy new shoes.
t: I will buy a television.
The symbolic form of the argument is given as;

$$\begin{array}{l} m \to s \\ \underline{s \to \sim t} \\ \therefore t \to \sim m \end{array}$$

The truth table for this argument is as follows:

m	s	t	First Premise $m \to s$	Second Premise $s \to \sim t$	Conclusion $t \to \sim m$	
~~T~~	~~T~~	~~T~~	~~T~~	~~F~~	~~F~~	~~Row 1~~
T	T	F	T	T	T	Row 2
~~T~~	~~F~~	~~T~~	~~F~~	~~T~~	~~F~~	~~Row 3~~
~~T~~	~~F~~	~~F~~	~~F~~	~~T~~	~~T~~	~~Row 4~~
~~F~~	~~T~~	~~T~~	~~T~~	~~F~~	~~T~~	~~Row 5~~
F	T	F	T	T	T	Row 6
F	F	T	T	T	T	Row 7
F	F	F	T	T	T	Row 8

The only rows in which both premises are true are rows 2, 6, 7, and 8. Because the conclusion is true in each of these rows, the argument is valid.

16. Determine whether the following argument is valid or invalid.

If I arrive before 8 A.M., then I will make the flight.

If I make the flight, then I will give the presentation.

∴ If I arrive before 8 A.M., then I will give the presentation.

Solution.

Let a represent the statement "I arrive before 8 A.M." Let f represent the statement "I will make the flight." Let p represent the statement "I will give the presentation." Then the symbolic form of the argument is

$$\begin{array}{l} a \to f \\ \underline{f \to p} \\ \therefore a \to p \end{array}$$

The truth table for this argument is as follows:

			First Premise	Second Premise	Conclusion	
a	f	p	$a \to f$	$f \to p$	$a \to p$	
T	T	T	T	T	T	Row 1
T	T	F	T	F	F	Row 2
T	F	T	F	T	T	Row 3
T	F	F	F	T	F	Row 4
F	T	T	T	T	T	Row 5
F	T	F	T	F	T	Row 6
F	F	T	T	T	T	Row 7
F	F	F	T	T	T	Row 8

The only rows in which all the premises are true are rows 1, 5, 7, and 8. In each of these rows the conclusion is also true. Thus the argument is a valid argument.

Standard Forms

Some arguments can be shown to be valid if they have the same symbolic form as an argument that is known to be valid. For instance, we have shown that the argument

$$\begin{array}{l} h \rightarrow m \\ \underline{h \qquad} \\ \therefore m \end{array}$$

is valid. This symbolic form is known as **modus ponens** or the **law of detachment**.
All arguments that have this symbolic form are valid. Table shows **four symbolic forms** and the name used to identify each form. Any argument that has a symbolic form identical to one of these symbolic forms is a valid argument.

Modus Ponens	Modus Tollens	Law of Syllogism	Disjunctive Syllogism
$p \rightarrow q$ $\underline{p \qquad}$ $\therefore q$	$p \rightarrow q$ $\underline{\sim q \qquad}$ $\therefore \sim p$	$p \rightarrow q$ $\underline{q \rightarrow r}$ $\therefore p \rightarrow r$	$p \vee q$ $\underline{\sim p \qquad}$ $\therefore q$

The law of syllogism can be extended to include more than two conditional premises. For example, if the premises of an argument are $a \rightarrow b, b \rightarrow c, \ldots, y \rightarrow z$ then a valid conclusion for the argument is $a \rightarrow z$.We will refer to any argument of this form with more than two conditional premises as the **extended law of syllogism**. Table shows two **symbolic forms** associated with **invalid arguments**. Any argument that has one of these symbolic forms is invalid.

Fallacy of the Converse	Fallacy of the Inverse
$p \rightarrow q$ $\underline{q \qquad}$ $\therefore p$	$p \rightarrow q$ $\underline{\sim p \qquad}$ $\therefore \sim q$

17. Use a standard form to determine whether the following argument is valid or invalid.

The program is interesting or I will watch the basketball game.

<u>The program is not interesting.</u>

∴ I will watch the basketball game.

Solution.

Label the statements.

i: The program is interesting.

w: I will watch the basketball game.

In symbolic form the argument is:

$$\begin{array}{l} i \vee w \\ \underline{\sim i \quad\quad} \\ \therefore w \end{array}$$

This symbolic form matches the standard form known as disjunctive syllogism. Thus the argument is valid.

18. Use a standard form to determine whether the following argument is valid or invalid.

If I go to Florida for spring break, then I will not study.

<u>I did not go to Florida for spring break.</u>

∴ I studied.

Solution.

Let f represent "I go to Florida for spring break." Let $\sim s$ represent "I will not study." Then the symbolic form of the argument is

$$\begin{array}{l} f \rightarrow \sim s \\ \underline{\sim f \quad\quad} \\ \therefore s \end{array}$$

This argument has the form of the fallacy of the inverse. Thus the argument is invalid.

19. Determine whether the following argument is valid.

If the movie was directed by George Lucas (l), then I want to see it (w). If I want to see a movie, then the movie's production costs must have exceeded 20 million dollars (c). The movie's production costs were less than 20 million dollars. Therefore, the movie was not directed by George Lucas.

Solution.

In symbolic form the argument is:

$l \to w$	Premise 1
$w \to c$	Premise 2
$\sim c$	Premise 3
$\therefore \sim l$	Conclusion

Applying the law of syllogism to Premises 1 and 2 produces

$l \to w$	Premise 1
$w \to c$	Premise 2
$\therefore l \to c$	Law of syllogism

Combining the above conclusion $l \to$ c with Premise 3 gives us

$l \to c$	Conclusion from above
$\sim c$	Premise 3
$\therefore \sim l$	Modus tollens

This sequence of valid arguments has produced the conclusion given in the original argument. Thus the original argument is a valid argument.

20. Determine whether the following argument is valid.

I start to fall asleep if I read a math book. I drink a soda whenever I start to fall asleep. If I drink a soda, then I must eat a candy bar. Therefore, I eat a candy bar whenever I read a math book.

Solution.

Let r represent "I read a math book." Let *f* represent "I start to fall asleep." Let d represent "I drink a soda." Let e represent "I eat a candy bar." Then the symbolic form of the argument is

$$\begin{array}{l} r \to f \\ f \to d \\ \underline{d \to e} \\ \therefore r \to e \end{array}$$

The argument has the form of the extended law of syllogism. Thus the argument is valid.

Categorical Propositions

In logic, categorical propositions are statements that express a relationship between two categories or classes of things. They are a fundamental concept in Aristotelian logic and are used to make assertions about the properties or characteristics of a particular group or category.

A categorical proposition typically has the following structure:

Subject-Predicate-Quantity-Quality

Subject: The category or class being referred to (e.g., "All humans").

Predicate: The property or characteristic being attributed to the subject (e.g., "are mortal").

Quantity: The extent or scope of the statement (e.g., "all," "some," "no").

Quality: The affirmation or negation of the statement (e.g., "are" or "are not").

Types of Categorical Propositions

There are four types of categorical propositions, based on the combinations of quantity and quality:

Universal Affirmative (A): All S are P. (e.g., "All humans are mortal.")

Universal Negative (E): No S are P. (e.g., "No humans are immortal.")

Particular Affirmative (I): Some S are P. (e.g., "Some humans are wise.")

Particular Negative (O): Some S are not P. (e.g., "Some humans are not wise.")

Here's a breakdown of each type:

A (Universal Affirmative): Asserts that all members of the subject class have the predicate property.

E (Universal Negative): Asserts that no members of the subject class have the predicate property.

I (Particular Affirmative): Asserts that at least one member of the subject class has the predicate property.

O (Particular Negative): Asserts that at least one member of the subject class does not have the predicate property.

Categorical propositions are used to make arguments and inferences in logic, and they play a crucial role in evaluating the validity of arguments.

Following examples illustrate how categorical propositions can be applied to mathematical statements, using quantifiers like "all", "some", "no", and "there exists" to make assertions about sets, functions, and relations.

Universal Affirmative (A)

- All integers are rational numbers.

- For all real numbers x, $x^2 \geq 0$.

- Every even number is divisible by 2.

Universal Negative (E)

-No integer is equal to its own square.

- There is no real number x such that $x^2 + 1 = 0$.

- No odd number is divisible by 2.

Particular Affirmative (I)

- Some integers are prime numbers.

- There exists a real number x such that $x^2 = 2$.

- Some even numbers are perfect squares.

Particular Negative (O)

- Some integers are not rational numbers.

- There exists a real number x such that $x^2 < 0$.

- Some odd numbers are not prime.

Examples with Sets

- Let A = {1, 2, 3, 4, 5} and B = {2, 4, 6, 8, 10}. Then:

- All elements of A are less than 6. (A)

- Some elements of A are not in B. (I)

- No element of A is greater than 5. (E)

- Some elements of B are not in A. (O)

Examples with Functions

- Let $f(x) = x^2$ and $g(x) = 2x$. Then:

- For all real numbers x, $f(x) \geq 0$. (A)

- There exists a real number x such that $f(x) = g(x)$. (I)

- No real number x satisfies $f(x) = -1$. (E)

- Some real numbers x satisfy $f(x) \neq g(x)$. (O)

Following examples illustrate how categorical propositions can be applied to various fields of study, including education, environmental science, history, economics, philosophy, sociology, psychology, and business.

Universal Affirmative (A)

All students who graduate from a university have a degree.

In this example, the subject is "students who graduate from a university", and the predicate is "have a degree". This proposition is asserting that every single student who graduates from a university has a degree.

Universal Negative (E)

No fossil fuel is renewable.

In this example, the subject is "fossil fuel", and the predicate is "renewable". This proposition is asserting that not a single fossil fuel is renewable.

Particular Affirmative (I)

Some ancient civilizations built pyramids.

In this example, the subject is "ancient civilizations", and the predicate is "built pyramids". This proposition is asserting that at least one ancient civilization built pyramids.

Particular Negative (O)

Some countries do not have a minimum wage.

In this example, the subject is "countries", and the predicate is "have a minimum wage". This proposition is asserting that there exists at least one country that does not have a minimum wage.

Universal Affirmative (A)

All moral actions are guided by reason.

In this example, the subject is "moral actions", and the predicate is "guided by reason". This proposition is asserting that every single moral action is guided by reason.

Universal Negative (E)

No social institution is completely egalitarian.

In this example, the subject is "social institution", and the predicate is "completely egalitarian". This proposition is asserting that not a single social institution is completely egalitarian.

Particular Affirmative (I)

Some people are born with a talent for music.

In this example, the subject is "people", and the predicate is "born with a talent for music". This proposition is asserting that at least one person is born with a talent for music.

Particular Negative (O)

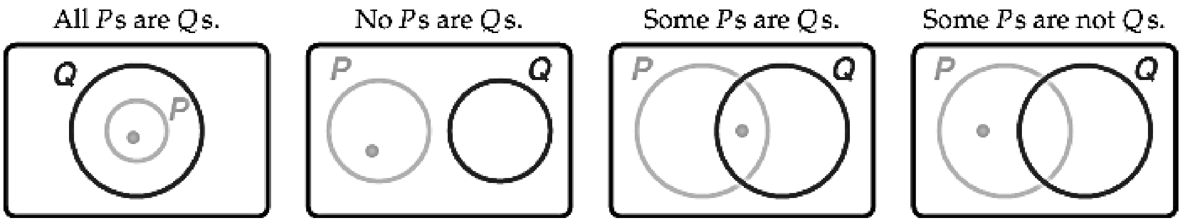
Some companies do not prioritize profit over social responsibility.

In this example, the subject is "companies", and the predicate is "prioritize profit over social responsibility". This proposition is asserting that there exists at least one company that does not prioritize profit over social responsibility.

Euler Diagrams

Many arguments involve sets whose elements are described using the quantifiers all, some, and none. The mathematician Leonhard Euler used diagrams to determine whether arguments that involved quantifiers were valid or invalid.

The following figures show **Euler diagrams** that illustrate the **four possible relationships** that can exist **between two sets**.



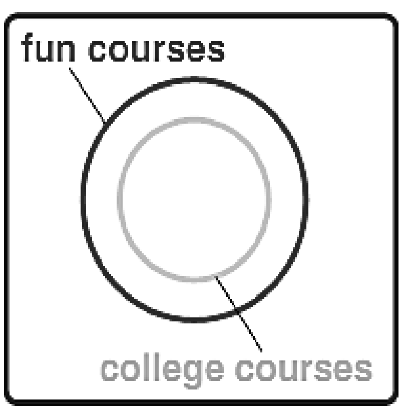
Euler used diagrams to illustrate logic concepts. Some 100 years later, John Venn extended the use of Euler's diagrams to illustrate many types of mathematics. In this section, we will construct diagrams to determine the validity of arguments. We will refer to these diagrams as Euler diagrams.

21. Use an Euler diagram to determine whether the following argument is valid or invalid.

All college courses are fun.
This course is a college course.
∴ This course is fun.

Solution.

The first premise indicates that the set of college courses is a subset of the set of fun courses. We illustrate this subset relationship with an Euler diagram, as shown in Figure 1. The second premise tells us that "this course" is an element of the set of college courses. If we use c to represent "this course," then c must be placed inside the set of college courses, as shown in Figure 2.



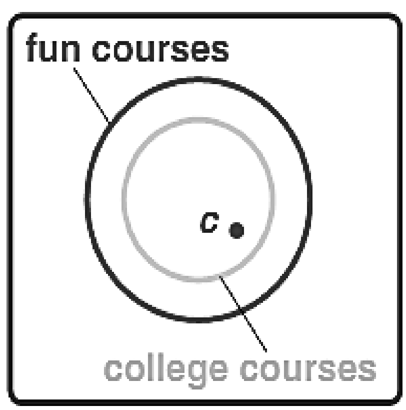


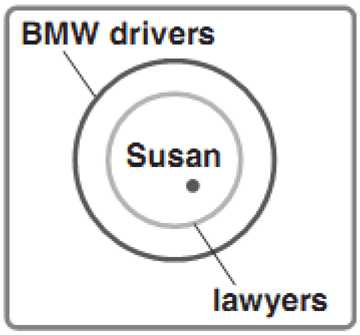
Figure 2 illustrates that c must also be an element of the set of fun courses. Thus the argument is valid.

22. Use an Euler diagram to determine whether the following argument is valid or invalid.

All lawyers drive BMWs.
<u>Susan is a lawyer.</u>
∴ Susan drives a BMW.

Solution.

The following Euler diagram shows that the argument is valid.



Note

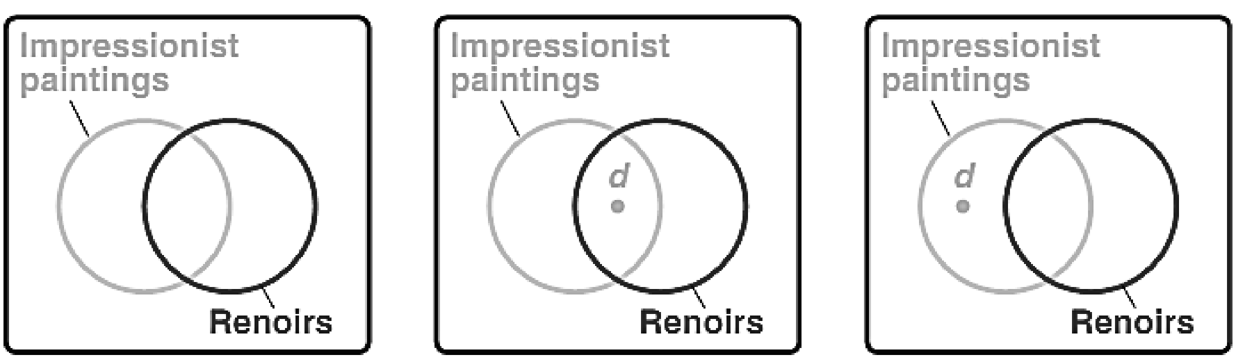
If an Euler diagram can be drawn so that the conclusion does not necessarily follow from the premises, then the argument is invalid.

23. Use an Euler diagram to determine whether the following argument is valid or invalid.

Some impressionist paintings are Renoirs.
<u>Dance at Bougival is an impressionist painting.</u>
∴ Dance at Bougival is a Renoir.

Solution.

The Euler diagram in Figure 1 illustrates the premise that some impressionist paintings are Renoirs. Let d represent the painting Dance at Bougival. Figures 2 and 3 show that d can be placed in one of two region.



Although Figure 2 supports the argument, Figure 3 shows that the conclusion does not necessarily follow from the premises, and thus the argument is invalid.

24. Use an Euler diagram to determine whether the following argument is valid or invalid.

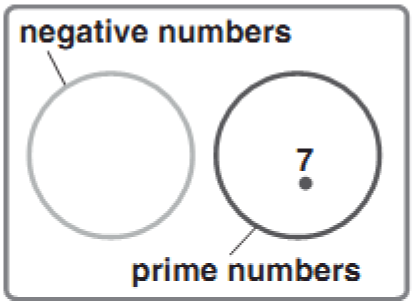
No prime numbers are negative.

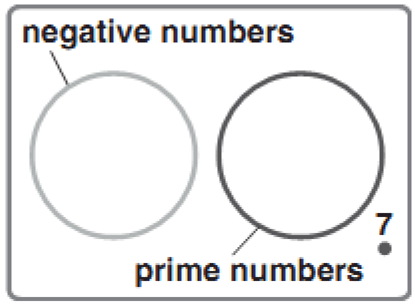
<u>The number 7 is not negative.</u>

∴ The number 7 is a prime number.

Solution.

From the given premises we can conclude that 7 may or may not be a prime number. Thus the argument is invalid.





25. If one particular example can be found for which the conclusion of an argument is true when its premises are true, must the argument be valid?

Answer.

No. To be a valid argument, the conclusion must be true whenever the premises are true. Just because the conclusion is true for one specific example, it does not mean the argument is a valid argument.

26. Use an Euler diagram to determine whether the following argument is valid or invalid.

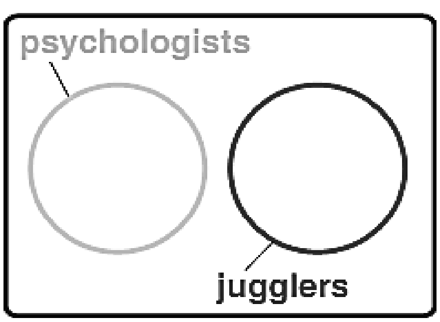
No psychologist can juggle.

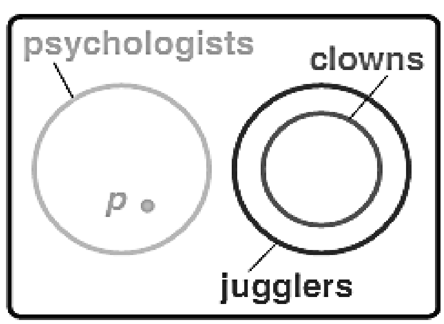
<u>All clowns can juggle.</u>

∴ No psychologist is a clown.

Solution.

The Euler diagram in Figure 1 shows that the set of psychologists and the set of jugglers are disjoint sets. Figure 2shows that because the set of clowns is a subset of the set of jugglers, no psychologists p are elements of the set of clowns. Thus the argument is valid.





27. Use an Euler diagram to determine whether the following argument is valid or invalid.

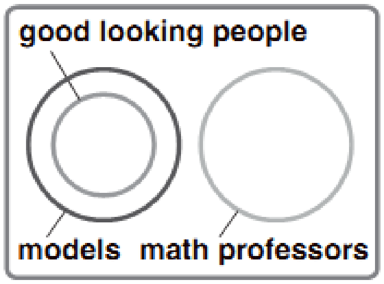
No mathematics professors are good-looking.

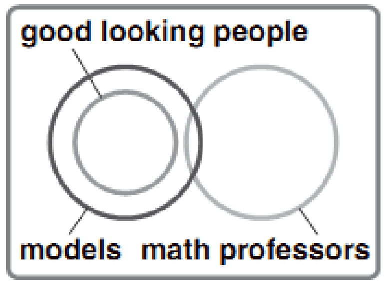
<u>All good-looking people are models.</u>

∴ No mathematics professor is a model.

Solution.

From the given premises we can construct two possible Euler diagrams.





From the rightmost Euler diagram we can determine that the argument is invalid.

Euler Diagrams and the Extended Law of Syllogism

28. Use an Euler diagram to determine whether the following argument is valid or invalid.

All fried foods are greasy.

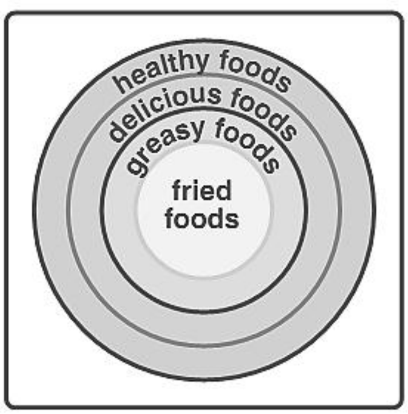
All greasy foods are delicious.

<u>All delicious foods are healthy.</u>

∴ All fried foods are healthy.

Solution.

The figure at the left illustrates that every fried food is an element of the set of healthy foods, so the argument is valid.

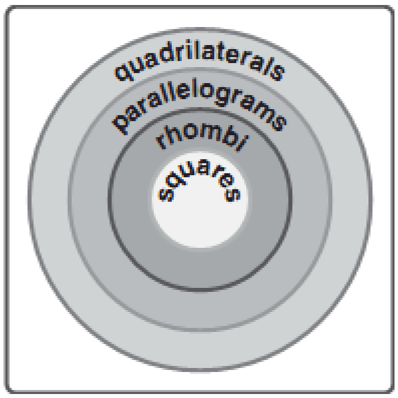


29. Use an Euler diagram to determine whether the following argument is valid or invalid.

All squares are rhombi.
All rhombi are parallelograms.
All parallelograms are quadrilaterals.
∴ All squares are quadrilaterals.

Solution.

The following Euler diagram illustrates that all squares are quadrilaterals, so the argument is a valid argument.



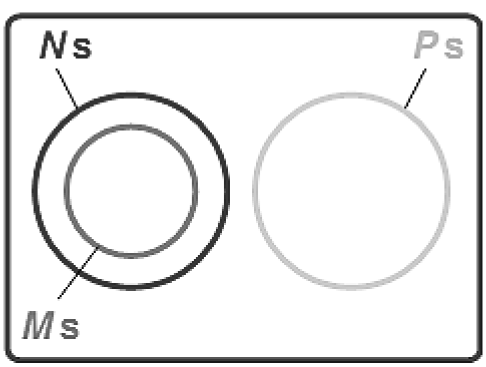
Using Euler Diagrams to Form Conclusions

30. Use an Euler diagram and all of the premises in the following argument to determine a valid conclusion for the argument.

All Ms are Ns.
No Ns are Ps.
∴ ?

Solution

The first premise indicates that the set of Ms is a subset of the set of Ns. The second premise indicates that the set of Ns and the set of Ps are disjoint sets. The following Euler diagram illustrates these set relationships. An examination of the Euler diagram allows us to conclude that no Ms are Ps.



31. Use an Euler diagram and all of the premises in the following argument to determine a valid conclusion for the argument.

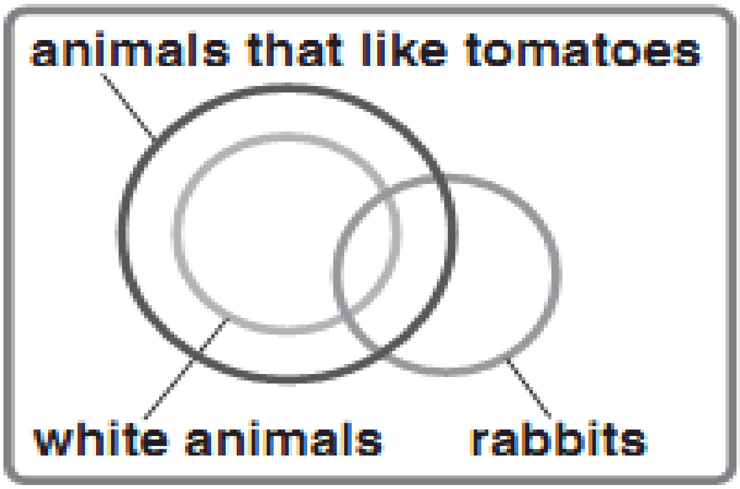
Some rabbits are white.

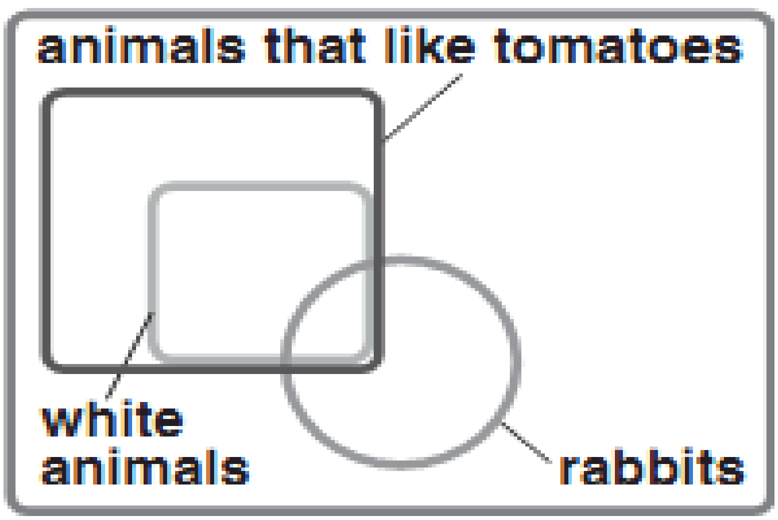
<u>All white animals like tomatoes.</u>

∴ ?

Solution.

The following Euler diagrams illustrate two possible cases. In both cases we see that all white rabbits like tomatoes.





Quantifiers

The word or symbol, which conveys the idea of quantity or numbers, is called quantifier. In mathematics and logic, quantifiers specify the scope of a predicate or statement.

- Universal Quantifier (∀): "For all"
- Existential Quantifier (∃): "There exists"
- In a statement, the word *some* and the phrases *there exists* and *at least one* are called existential quantifiers. Existential quantifiers are used as prefixes to assert the existence of something.
- In a statement, the words *none, no, all*, and *every* are called universal quantifiers.
- The universal quantifiers n*one* and n*o* deny the existence of something, whereas the universal quantifiers *all* and *every* are used to assert that every element of a given set satisfies some condition.

Universal Quantification

The universal quantification of P(x) is the statement

"P(x) for all values of x in the domain."

The notation ∀xP(x) denotes the universal quantification of P(x). Here ∀ is called the universal quantifier. We read ∀xP(x) as "for all xP(x)" or "for every xP(x)." An element for which P(x) is false is called a counterexample to ∀xP(x).

Statement	*When True?*	*When False?*
$\forall x P(x)$	$P(x)$ is true for every x.	There is an x for which $P(x)$ is false.
$\exists x P(x)$	There is an x for which $P(x)$ is true.	$P(x)$ is false for every x.

32. Let P(x) be the statement "x + 1 > x." What is the truth value of the quantification ∀xP(x), where the domain consists of all real numbers?
Solution.
Because P(x) is true for all real numbers x, the quantification ∀xP(x) is true.

Remark

- Generally, an implicit assumption is made that all domains of discourse for quantifiers are nonempty. Note that if the domain is empty, then ∀xP(x) is true for any propositional function P(x) because there are no elements x in the domain for which P(x) is false.
- A statement ∀xP(x) is false, where P(x) is a propositional function, if and only if P(x)isnot always true when x is in the domain. One way to show that P(x) is not always true when x is in the domain is to find a counterexample to the statement ∀xP(x). Note that a single counterexample is all we need to establish that ∀xP(x) is false.

33. Let Q(x) be the statement "x < 2." What is the truth value of the quantification ∀xQ(x), where the domain consists of all real numbers?

Solution.

Q(x) is not true for every real number x, because, for instance, Q(3) is false. That is, x = 3 is a counterexample for the statement ∀xQ(x). Thus, ∀xQ(x) is false.

34. Suppose that P(x)is"$x^2 > 0$." To show that the statement ∀xP(x) is false where the universe of discourse consists of all integers.

Solution.

Suppose that P(x)is"$x^2 > 0$." To show that the statement ∀xP(x) is false where the universe of discourse consists of all integers, we give a counterexample. We see that x = 0 is a counterexample because $x^2 = 0$ when x = 0, so that x^2 is not greater than 0 when x = 0.

35. What does the statement ∀xN(x) mean if N(x) is "Computer x is connected to the network" and the domain consists of all computers on campus?

Solution.

The statement ∀xN(x) means that for every computer x on campus, that computer x is connected to the network. This statement can be expressed in English as "Every computer on campus is connected to the network."

36. What is the truth value of $\forall x(x^2 \geq x)$ if the domain consists of all real numbers? What is the truth value of this statement if the domain consists of all integers?

Solution.

The universal quantification $\forall x(x^2 \geq x)$, where the domain consists of all real numbers, is false. For example, $\left(\frac{1}{2}\right)^2 \ngeq \left(\frac{1}{2}\right)$.

Note that $x^2 \geq x$ if and only if $x^2 - x = x(x - 1) \geq 0$. Consequently, $x^2 \geq x$ if and only if $x \leq 0$ or $x \geq 1$. It follows that $\forall x(x2 \geq x)$ is false if the domain consists of all real numbers (because the inequality is false for all real numbers x with $0 < x < 1$). However, if the domain consists of the integers, $\forall x(x^2 \geq x)$ is true, because there are no integers x with $0 < x < 1$.

Existential Quantification

The existential quantification of P(x) is the proposition

"There exists an element x in the domain such that P(x)."

We use the notation $\exists xP(x)$ for the existential quantification of P(x).

Here $\exists$ is called the existential quantifier.

37. Let P(x) denote the statement "$x > 3$." What is the truth value of the quantification $\exists xP(x)$, where the domain consists of all real numbers?

Solution.

Because "$x > 3$" is sometimes true—for instance, when $x = 4$—the existential quantification of P(x), which is $\exists xP(x)$, is true.

Observe that the statement $\exists xP(x)$ is false if and only if there is no element x in the domain for which P(x) is true. That is, $\exists xP(x)$ is false if and only if P(x) is false for every element of the domain. We illustrate this observation in next Example.

38. Let Q(x) denote the statement "$x = x + 1$."What is the truth value of the quantification $\exists xQ(x)$, where the domain consists of all real numbers?

Solution.

Because Q(x) is false for every real number x, the existential quantification of Q(x), which is $\exists xQ(x)$, is false.

Predicates

In mathematics and logic, a predicate is a statement that contains variables and asserts a property or relationship about those variables. Predicates are used to define sets, functions, and relationships.

Types of Predicates

- Universal predicate (∀): "For all"
- Existential predicate (∃): "There exists"
- Equality predicate (=): "Equal to"

39. Let P(x) be the predicate "x is an even number." What is P(4)?
Solution. True

40. Let Q(x) be the predicate "x > 5." What is Q(3)?
Solution: False

41. $\exists x\ (x^2 = 16)$
Solution. True ($x = \pm 4$)

42. $\forall x\ (x + 0 = x)$
Solution. True (additive identity)

43. Let P(x) denote the statement "x > 3." What are the truth values of P(4) and P(2)?
Solution.
We obtain the statement P(4) by setting x = 4 in the statement "x > 3." Hence, P(4), which is the statement "4 > 3," is true. However, P(2), which is the statement "2 > 3," is false.

44. Let A(x) denote the statement "Computer x is under attack by an intruder." Suppose that of the computers on campus, only CS2 and MATH1 are currently under attack by intruders. What are truth values of A(CS1), A(CS2), and A(MATH1)?
Solution.
We obtain the statement A(CS1) by setting x = CS1 in the statement "Computer x is under attack by an intruder." Because CS1 is not on the list of computers currently under attack, we conclude that A(CS1) is false. Similarly, because CS2 and MATH1 are on the list of computers under attack, we know that A(CS2) and A(MATH1) are true.

We can also have statements that involve more than one variable. For instance, consider the statement “x = y + 3.”We can denote this statement by Q(x, y), where x and y are variables and Q is the predicate. When values are assigned to the variables x and y, the statement Q(x, y) has a truth value.

45. Let Q(x, y) denote the statement “x = y + 3.” What are the truth values of the propositions Q(1, 2) and Q(3, 0)?
Solution.
To obtain Q(1, 2), set x = 1 and y = 2 in the statement Q(x, y). Hence, Q(1, 2) is the statement “1 = 2 + 3,” which is false. The statement Q(3, 0) is the proposition “3 = 0 + 3,” which is true.

46. Let A(c, n) denote the statement “Computer c is connected to network n,” where c is a variable representing a computer and n is a variable representing a network. Suppose that the computer MATH1 is connected to network CAMPUS2, but not to network CAMPUS1. What are the values of A(MATH1, CAMPUS1) and A(MATH1, CAMPUS2)?
Solution.
BecauseMATH1 is not connected to the CAMPUS1 network, we see that A(MATH1, CAMPUS1) is false. However, because MATH1 is connected to the CAMPUS2 network, we see that A(MATH1, CAMPUS2) is true.
Similarly, we can let R(x, y, z) denote the statement “x + y = z.” When values are assigned to the variables x, y, and z, this statement has a truth value.

47. What are the truth values of the propositions R(1, 2, 3) and R(0, 0, 1)?
Solution.
The proposition R(1, 2, 3) is obtained by setting x = 1, y = 2, and z = 3 in the statement R(x, y, z). We see that R(1, 2, 3) is the statement “1 + 2 = 3,” which is true. Also note that R(0, 0, 1), which is the statement “0 + 0 = 1,” is false.

Key Concepts

1. Quantifiers (∀, ∃)
2. Variables (x, y, z)
3. Predicates (P(x), Q(x), R(x, y))
4. Logical operators (∧, ∨, ¬)
5. Set notation ({x | P(x)})

Logical Fallacies

A fallacy is a deceptive argument—an argument in which the conclusion is not well supported by the premises. While Logical fallacies are errors in reasoning that undermine the validity of an argument, leading to misleading or false conclusions. These fallacies occur when arguments are based on flawed assumptions, invalid inferences, or misleading evidence.

In logic, the term **argument** refers to a reasoned or thoughtful process. Specifically, an argument uses a set of facts or assumptions, called **premises**, to support a **conclusion**. A **fallacy** is a deceptive argument, an argument in which the conclusion is not well supported by the premises.

Types of Logical Fallacies

- **Informal Fallacies (error in the content of an argument)**
  - Ad Hominem: Attacking the person, not the argument.
  - Straw Man: Misrepresenting or exaggerating an opposing argument.
  - False Dilemma: Presenting only two options when more exist.
  - Slippery Slope: Assuming a chain of events without evidence.
  - Begging the Question: Assuming the truth of the conclusion.
- **Formal Fallacies (error in the logical structure of an argument)**
  - Appeal to Authority: Using authority to support an unjustified claim.
  - Appeal to Emotion: Using emotions rather than logic.
  - Appeal to Popularity: Arguing that popularity validates an argument.
  - False Cause: Assuming causation without evidence.
  - Slanting: Selectively presenting information.
- **Fallacies of Presumption**
  - Assuming the Conclusion: Using a conclusion as a premise.
  - Shifting the Burden of Proof: Requiring others to disprove an argument.
  - Lack of Evidence: Arguing from ignorance.
- **Fallacies of Ambiguity**
  - Equivocation: Using ambiguous language.
  - Vagueness: Using unclear language.
- **Fallacies of Relevance**
  - Red Herring: Introducing irrelevant information.
  - Non Sequitur: Drawing unrelated conclusions.

Characteristics of Logical Fallacies

- Misleading or false premises
- Flawed assumptions
- Invalid inferences
- Misleading evidence
- Emotional appeals

Consequences of Logical Fallacies

- Misleading conclusions
- Invalid arguments
- Poor decision-making
- Miscommunication
- Erosion of critical thinking

Examples and Countermeasures

- Ad Hominem: "John's argument is wrong because he's biased."
  Countermeasure: Address the argument, not the person.
- Strawman: "You want to raise taxes, so you want socialism."
  Countermeasure: Clarify and address the actual argument.
- False Dilemma: "You're either with us or against us."
  Countermeasure: Identify and propose alternative options.

Recognizing Logical Fallacies

- Identify assumptions.
- Evaluate evidence.
- Check for ambiguity.
- Ensure relevance.
- Consider alternative perspectives.

Avoiding Logical Fallacies

- Use clear language.
- Support claims with evidence.
- Address counterarguments.
- Avoid emotional appeals.
- Encourage critical thinking.

Appeal to Popularity

"Ford makes the best pickup trucks in the world. More people drive Ford pickups than any other light truck."

Analysis The first step in dealing with any argument is recognizing which statements are premises and which are conclusions. This argument tries to make the case that Ford makes the best pickup trucks in the world, so this statement is its conclusion. The only evidence it offers to support this conclusion is the statement more people drive Ford pickups than any other light truck. This is the argument's only premise. Overall, this argument has the form

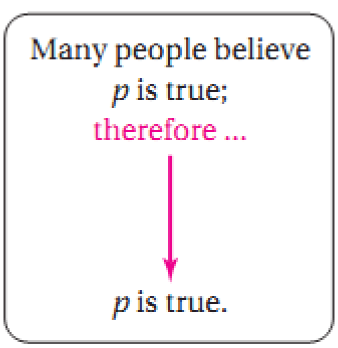
Premise: More people drive Ford pickups than any other light truck.

Conclusion: Ford makes the best pickup trucks in the world.

Note that the original written argument states the conclusion before the premise. Such "backward" structures are common in everyday speech and are perfectly legitimate as long as the argument is well reasoned. In this case, however, the reasoning is faulty.

The fact that more people drive Ford pickups does not necessarily mean that they are the best trucks.

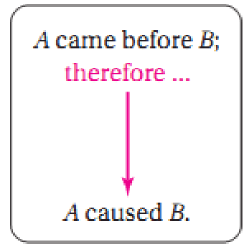
This argument suffers from the fallacy of appeal to popularity (or appeal to majority), in which the fact that large numbers of people believe or act some way is used inappropriately as evidence that the belief or action is correct. We can represent the general form of this fallacy with a diagram in which the letter p stands for a particular statement (Figure). In this case, p stands for the statement Ford makes the best pickup trucks in the world.



False Cause

"I placed the quartz crystal on my forehead, and in five minutes my headache was gone. The crystal made my headache go away."

Analysis We identify the premises and conclusion of this argument as follows:



Premise: I placed the quartz crystal on my forehead.

Premise: Five minutes later my headache was gone.

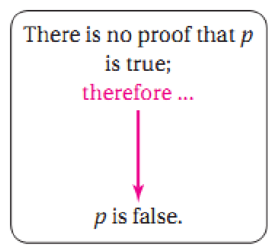
Conclusion: The crystal made my headache go away.

The premises tell us that one thing (crystal on forehead) happened before another (headache went away), but they don't prove any connection between them. That is, we cannot conclude that the crystal caused the headache to go away.

This argument suffers from the fallacy of false cause, in which the fact that one event came before another is incorrectly taken as evidence that the first event caused the second event. We can represent this fallacy with a diagram in which A and B represent two different events (Figure). In this case, A is the event of putting the crystal on the forehead and B is the event of the headache going away.

Appeal to ignorance

"Scientists have not found any concrete evidence of aliens visiting Earth. Therefore, anyone who claims to have seen a UFO must be hallucinating."



Analysis If we strip the argument to its core, it says this:

Premise: There's no proof that aliens have visited Earth.

Conclusion: Aliens have not visited Earth.

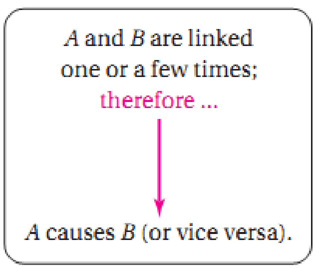
The fallacy should be clear: A lack of proof of alien visits does not mean that visits have not occurred. This fallacy is called appeal to ignorance because it uses ignorance (lack of knowledge) about the truth of a proposition to conclude the opposite (Figure). We sometimes sum up this fallacy with the statement "An absence of evidence is not evidence of absence."

Hasty generalization

"Two cases of childhood leukemia have occurred along the street where the high-voltage power lines run. The power lines must be the cause of these illnesses."

Analysis The premise of this argument cites two cases of leukemia, but two cases are not enough to establish a pattern, let alone to conclude that the power lines caused the illnesses.

The fallacy here is hasty generalization, in which a conclusion is drawn from an inadequate number of cases or cases that have not been sufficiently analyzed. If any connection between power lines and leukemia exists, it would have to be established with far more evidence than is provided in this argument. (In fact, decades of research have found no connection between power lines and illness.) We can represent this fallacy with a diagram in which A and B represent two linked events (Figure).



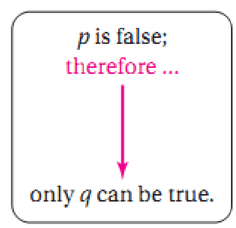
Limited Choice

"You don't support the President, so you are not a patriotic American."

Analysis This argument has the form

Premise: You don't support the President.

Conclusion: You are not a patriotic American.



The argument suggests that there are only two types of Americans: patriotic ones who support the President and unpatriotic ones who don't. But there are many other possibilities, such as being patriotic while disliking a particular President.

This fallacy is called limited choice (or false dilemma) because it artificially precludes choices that ought to be considered. Figure shows one common form of this fallacy.

Limited choice also arises with questions such as "Have you stopped smoking?" Because both yes and no answers imply that you smoked in the past, the question precludes the possibility that you never smoked. (In legal proceedings, questions of this type are disallowed because they attempt to "lead the witness.") Another simple and common form of this fallacy is "You're wrong, so I must be right."

Appeal to Emotion

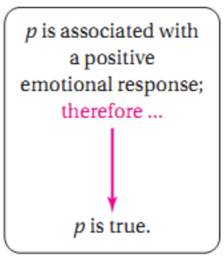
In ads for Michelin tires, a picture of a baby is shown with the words “because so much is riding on your tires.”

Analysis If we can consider this an argument at all, it has the form

Premise: You love your baby.

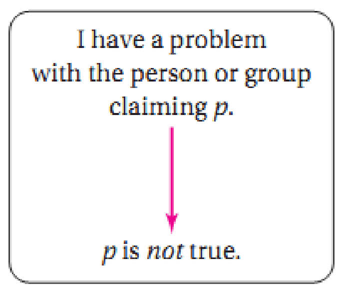
Conclusion: You should buy Michelin tires.

The advertisers hope that the love you feel for a baby will make you want to buy their tires. This attempt to evoke an emotional response as a tool of persuasion represents the fallacy of appeal to emotion. Figure shows its form when the emotional response is positive. Sometimes the appeal is to negative emotions. For example, the statement if my opponent is elected, your tax burden will rise tries to convince you that electing the other candidate will lead to consequences you won’t like. (In this negative form, the fallacy is sometimes called appeal to force.)



Personal attack

Gwen: You should stop drinking because it’s hurting your grades, endangering people when you drink and drive, and destroying your relationship with your family.



Merle: I’ve seen you drink a few too many on occasion yourself!

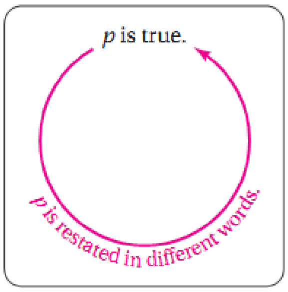
Analysis Gwen’s argument is well reasoned, with premises offering strong support for her conclusion that Merle should stop drinking. Merle rejects this argument by noting that Gwen sometimes drinks too much herself. Even if Merle’s claim is true, it is irrelevant to Gwen’s point. Merle has resorted to attacking Gwen personally rather than arguing logically, so we call this fallacy personal attack (Figure). (It is also called ad hominem, Latin for “to the person.”)

The fallacy of personal attack can also apply to groups. For example, someone might say, “This new bill will be an environmental disaster because its sponsors received large campaign contributions from oil companies.” This argument is fallacious because it doesn’t challenge the provisions of the bill, but only questions the motives of the sponsors.

Circular reasoning

"Society has an obligation to provide health insurance because health care is a right of citizenship."

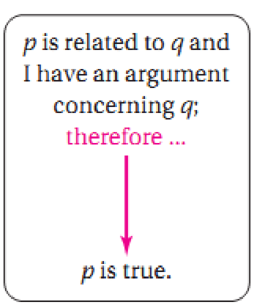
Analysis This argument states the conclusion (society has an obligation to provide health insurance) before the premise (health care is a human right). But if you read carefully, you'll recognize that the premise and the conclusion both say essentially the same thing, as social obligations are generally based on definitions of accepted rights. This argument therefore suffers from circular reasoning (Figure).



Diversion (red herring)

"We should not continue to fund cloning research because there are so many ethical issues involved. Decisions are based on ethics, and we cannot afford to have too many ethical loose ends."

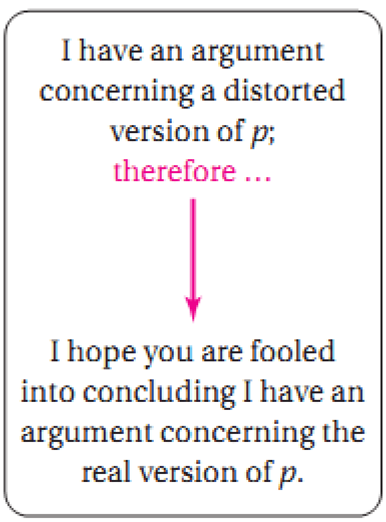
Analysis The argument begins with its conclusion—we should not continue to fund cloning research. However, the discussion is all about ethics. This argument represents the fallacy of diversion (Figure) because it attempts to divert attention from the real issue (funding for cloning research) by focusing on another issue (ethics). The issue to which attention is diverted is sometimes called a red herring. (A herring is a fish that turns red when rotten. Use of the term red herring to mean a diversion can be traced back to the 19th century, when British fugitives discovered that they could divert bloodhounds from their pursuit by rubbing a red herring across their trail.)



Straw Man

Suppose that the mayor of a large city proposes decriminalizing drug possession in order to reduce overcrowding in jails and save money on enforcement. His challenger in the upcoming election says, “The mayor doesn’t think there’s anything wrong with drug use, but I do.”

Analysis The mayor did not say that drug use is acceptable. His proposal for decriminalization is designed to solve another problem—overcrowding of jails—and tells us nothing about his general views on drug use. The speaker has distorted the mayor’s views. Any argument based on a distortion of someone’s words or beliefs is called a straw man (Figure). The term comes from the idea that the speaker has used a poor representation of a person’s beliefs in the same way that a straw man is a poor representation of a real man. A straw man is similar to a diversion. The primary difference is that a diversion argues against an unrelated issue, while the straw man argues against a distorted version of the real issue.



Exercise

1. In Exercises a – j, determine whether each sentence is a statement.
   a) West Virginia is west of the Mississippi River.
   b) 1031 is a prime number.
   c) The area code for Storm Lake, Iowa, is 512.
   d) Some negative numbers are rational numbers.
   e) Have a fun trip.
   f) Do you like to read?
   g) All hexagons have exactly five sides.
   h) If x is a negative number, then x^2 is a positive number.
   i) Mathematics courses are better than history courses.
   j) Every real number is a rational number.
2. In Exercises a – d, write the negation of each statement.
   a) The Giants lost the game.
   b) The lunch was served at noon.
   c) The game did not go into overtime.
   d) The game was not shown on ABC.
3. Which of these sentences are propositions? What are the truth values of those that are propositions?
   a) Boston is the capital of Massachusetts. b) Miami is the capital of Florida.
   c) $2 + 3 = 5$.
   d) $5 + 7 = 10$.
   e) $x + 2 = 11$.
   f) Answer this question.
4. Which of these are propositions? What are the truth values of those that are propositions?
   a) Do not pass go.
   b) What time is it?
   c) There are no black flies in Maine.
   d) $4 + x = 5$.
   e) The moon is made of green cheese.
   f) $2^n \geq 100$.
5. What is the negation of each of these propositions?
   a) Linda is younger than Sanjay.
   b) Mei makes more money than Isabella.
   c) Moshe is taller than Monica.
   d) Abby is richer than Ricardo.
6. What is the negation of each of these propositions?
   a) Janice has more Facebook friends than Juan.
   b) Quincy is smarter than Venkat.
   c) Zelda drives more miles to school than Paola.
   d) Briana sleeps longer than Gloria.

7. What is the negation of each of these propositions?
a) Mei has an MP3 player.
b) There is no pollution in New Jersey. c) 2 + 1 = 3.
d) The summer in Maine is hot and sunny.

8. What is the negation of each of these propositions?
a) Jennifer and Teja are friends.
b) There are 13 items in a baker's dozen.
c) Abby sent more than 100 text messages yesterday.
d) 121 is a perfect square.

9. What is the negation of each of these propositions?
a) Steve has more than 100 GB free disk space on his laptop.
b) Zach blocks e-mails and texts from Jennifer.
c) $7 \cdot 11 \cdot 13 = 999$.
d) Diane rode her bicycle 100 miles on Sunday.

10. Suppose that Smartphone A has 256 MB RAM and 32 GB ROM, and the resolution of its camera is 8 MP; Smartphone B has 288 MB RAM and 64 GB ROM, and the resolution of its camera is 4 MP; and Smartphone C has 128 MB RAM and 32 GB ROM, and the resolution of its camera is 5 MP. Determine the truth value of each of these propositions.
a) Smartphone B has the most RAM of these three smartphones.
b) Smartphone C has more ROM or a higher resolution camera than Smartphone B.
c) Smartphone B has more RAM, more ROM, and a higher resolution camera than Smartphone A.
d) If Smartphone B has more RAM and more ROM than Smartphone C, then it also has a higher resolution camera.
e) Smartphone A has more RAM than Smartphone B if and only if Smartphone B has more RAM than Smart phone A.

11. Suppose that during the most recent fiscal year, the annual revenue of Acme Computer was 138 billion dollars and its net profit was 8 billion dollars, the annual revenue of Nadir Software was 87 billion dollars and its net profit was 5 billion dollars, and the annual revenue of Quixote Media was 111 billion dollars and its net profit was 13 billion dollars. Determine the truth value of each of these propositions for the most recent fiscal year.
a) Quixote Media had the largest annual revenue.
b) Nadir Software had the lowest net profit and Acme Computer had the largest annual revenue.
c) Acme Computer had the largest net profit or Quixote Media had the largest net profit.
d) If Quixote Media had the smallest net profit, then Acme Computer had the largest annual revenue.
e) Nadir Software had the smallest net profit if and only if Acme Computer had the largest annual revenue.

12. Let p and q be the propositions

p: You drive over 65 miles per hour.

q: You get a speeding ticket.

Write these propositions using p and q and logical connectives (including negations).

a) You do not drive over 65 miles per hour.

b) You drive over 65 miles per hour, but you do not get a speeding ticket.

c) You will get a speeding ticket if you drive over 65 miles per hour.

d) If you do not drive over 65 miles per hour, then you will not get a speeding ticket.

e) Driving over 65miles per hour is sufficient for getting a speeding ticket.

f) You get a speeding ticket, but you do not drive over 65 miles per hour.

g) Whenever you get a speeding ticket, you are driving over 65 miles per hour.

13. State the converse, contrapositive, and inverse of each of these conditional statements.

a) If it snows today, I will ski tomorrow.

b) I come to class whenever there is going to be a quiz.

c) A positive integer is a prime only if it has no divisors other than 1 and itself.

14. State the converse, contrapositive, and inverse of each of these conditional statements.

a) If it snows tonight, then I will stay at home.

b) I go to the beach whenever it is a sunny summer day.

c) When I stay up late, it is necessary that I sleep until noon.

15. Construct truth table of converse, inverse and contra positive of given conditional.

Conditional	**Converse**	**Inverse**	**Contra Positive**
$\sim p \rightarrow q$			
$q \rightarrow p$			
$\sim p \rightarrow \sim q$			
$\sim q \rightarrow \sim p$			

16. Construct truth table for the following statements.

a) $(p \rightarrow \sim p) \vee (p \rightarrow q)$

b) $(p \wedge \sim p) \rightarrow q$

c) $\sim(p \rightarrow q) \leftrightarrow (p \wedge \sim q)$

17. Determine whether each of the following is a tautology or not?

a) $(p \wedge q) \rightarrow p$

b) $p \rightarrow (p \vee q)$

c) $\sim(p \rightarrow q) \rightarrow p$

d) $\sim q \wedge (p \rightarrow q) \rightarrow \sim p$

18. Show that each of these conditional statements is a tautology by using truth tables.

a) $(p \land q) \rightarrow p$ b) $p \rightarrow (p \lor q)$
c) $\neg p \rightarrow (p \rightarrow q)$ d) $(p \land q) \rightarrow (p \rightarrow q)$
e) $\neg(p \rightarrow q) \rightarrow p$ f) $\neg(p \rightarrow q) \rightarrow \neg q$

19. Show that each of these conditional statements is a tautology by using truth tables.

a) $[\neg p \land (p \lor q)] \rightarrow q$
b) $[(p \rightarrow q) \land (q \rightarrow r)] \rightarrow (p \rightarrow r)$
c) $[p \land (p \rightarrow q)] \rightarrow q$
d) $[(p \lor q) \land (p \rightarrow r) \land (q \rightarrow r)] \rightarrow r$

20. Determine whether each of the following is a tautology, a contingency or an absurdity?

a) $p \land \sim p$
b) $p \rightarrow (q \rightarrow p)$
c) $q \lor (\sim q \lor p)$

21. Prove that $p \lor (\sim p \land \sim q) \lor (p \land q) = p \lor (\sim p \land \sim q)$
22. Show that $p \leftrightarrow q$ and $(p \land q) \lor (\neg p \land \neg q)$ are logically equivalent.
23. Show that $\neg(p \leftrightarrow q)$ and $p \leftrightarrow \neg q$ are logically equivalent.
24. Show that $p \rightarrow q$ and $\neg q \rightarrow \neg p$ are logically equivalent.
25. Show that $\neg p \leftrightarrow q$ and $p \leftrightarrow \neg q$ are logically equivalent.
26. Show that $\neg(p \oplus q)$ and $p \leftrightarrow q$ are logically equivalent.
27. Show that $\neg(p \leftrightarrow q)$ and $\neg p \leftrightarrow q$ are logically equivalent.
28. Show that $(p \rightarrow q) \land (p \rightarrow r)$ and $p \rightarrow (q \land r)$ are logically equivalent.
29. Show that $(p \rightarrow r) \land (q \rightarrow r)$ and $(p \lor q) \rightarrow r$ are logically equivalent.
30. Show that $(p \rightarrow q) \lor (p \rightarrow r)$ and $p \rightarrow (q \lor r)$ are logically equivalent.
31. Show that $(p \rightarrow r) \lor (q \rightarrow r)$ and $(p \land q) \rightarrow r$ are logically equivalent.
32. Show that $\neg p \rightarrow (q \rightarrow r)$ and $q \rightarrow (p \lor r)$ are logically equivalent.
33. Show that $p \leftrightarrow q$ and $(p \rightarrow q) \land (q \rightarrow p)$ are logically equivalent.
34. Show that $p \leftrightarrow q$ and $\neg p \leftrightarrow \neg q$ are logically equivalent.
35. Show that $(p \rightarrow q) \land (q \rightarrow r) \rightarrow (p \rightarrow r)$ is a tautology.
36. Show that $(p \lor q) \land (\neg p \lor r) \rightarrow (q \lor r)$ is a tautology.
37. Show that $(p \rightarrow q) \rightarrow r$ and $p \rightarrow (q \rightarrow r)$ are not logically equivalent.
38. Show that $(p \land q) \rightarrow r$ and $(p \rightarrow r) \land (q \rightarrow r)$ are not logically equivalent.
39. Show that $(p \rightarrow q) \rightarrow (r \rightarrow s)$ and $(p \rightarrow r) \rightarrow (q \rightarrow s)$ are not logically equivalent.
40. Let $P(x)$ denote the statement "$x \leq 4$." What are these truth values?

a) $P(0)$ b) $P(4)$ c) $P(6)$

41. Let $P(x)$ be the statement "The word x contains the letter a." What are these truth values?

a) P(orange) b) P(lemon)
c) P(true) d) P(false)

42. Let $Q(x, y)$ denote the statement "x is the capital of y." What are these truth values?

a) Q(Denver, Colorado) b) Q(Detroit, Michigan)
c) Q(Massachusetts, Boston) d) Q(New York, New York)

43. State the value of x after the statement if P(x) then x := 1 is executed, where P(x) is the statement “x > 1,” if the value of x when this statement is reached is
a) x = 0. b) x = 1. c) x = 2.

44. Let P(x) be the statement “x spends more than five hours every weekday in class,” where the domain for x consists of all students. Express each of these quantifications in English.
a) $\exists x P(x)$ b) $\forall x P(x)$
c) $\exists x \neg P(x)$ d) $\forall x \neg P(x)$

45. Let N(x) be the statement “x has visited North Dakota,” where the domain consists of the students in your school. Express each of these quantifications in English.
a) $\exists x N(x)$ b) $\forall x N(x)$ c) $\neg \exists x N(x)$
d) $\exists x \neg N(x)$ e) $\neg \forall x N(x)$ f) $\forall x \neg N(x)$

46. Construct the following theorems to logical form and prove them by constructing truth tables.
a) $(A \cap B)' = A' \cup B'$
b) $(A \cup B) \cup C = A \cup (B \cup C)$
c) $(A \cap B) \cap C = A \cap (B \cap C)$
d) $(A \cup B) \cap C = (A \cup B) \cap (A \cup C)$

47. In Exercises i–viii, use a truth table to determine whether the argument is valid or invalid.

i
$$\begin{array}{l} p \lor \sim q \\ \underline{\sim q \qquad} \\ \therefore p \end{array}$$

ii
$$\begin{array}{l} \sim p \land q \\ \underline{\sim p \qquad} \\ \therefore q \end{array}$$

iii
$$\begin{array}{l} p \rightarrow \sim q \\ \underline{\sim q \qquad} \\ \therefore p \end{array}$$

iv
$$\begin{array}{l} p \rightarrow \sim q \\ \underline{p \qquad} \\ \therefore \sim q \end{array}$$

v
$$\begin{array}{l} \sim p \rightarrow \sim q \\ \underline{\sim p \qquad} \\ \therefore \sim q \end{array}$$

vi
$$\begin{array}{l} \sim p \rightarrow q \\ \underline{p \qquad} \\ \therefore \sim q \end{array}$$

vii
$$\begin{array}{l} (p \rightarrow q) \land (\sim p \rightarrow q) \\ \underline{q \qquad\qquad\qquad\qquad} \\ \therefore p \end{array}$$

viii
$$\begin{array}{l} (p \lor q) \land (p \land q) \\ \underline{p \qquad\qquad\qquad\qquad} \\ \therefore q \end{array}$$

48. In Exercises a–f, use the indicated letters to write the argument in symbolic form. Then use a truth table to determine whether the argument is valid or invalid.

a) If you finish your homework (h) you may attend the reception (r). You did not finish your homework. Therefore, you cannot go to the reception.
b) The X Games will be held in Oceanside (o) if and only if the city of Oceanside agrees to pay $100,000 in prize money (a). If San Diego agrees to pay $200,000 in prize money (s), then the city of Oceanside will not agree to pay $100,000 in prize money. Therefore, if the X Games were held in Oceanside, then San Diego did not agree to pay $200,000 in prize money.
c) If I can't buy the house (~b) then at least I can dream about it (d). I can buy the house or at least I can dream about it. Therefore, I can buy the house.
d) If the winds are from the east (e) then we will not have a big surf (~s). We do not have a big surf. Therefore, the winds are from the east.
e) If I master college algebra (c) then I will be prepared for trigonometry I am prepared for trigonometry (t). Therefore, I mastered college algebra.
f) If it is a blot (b) then it is not a clot (~c). If it is a zlot (z) then it is a clot. It is a blot. Therefore, it is not a zlot.

49. In Exercises a–d, draw an Euler diagram that illustrates the relationship between the given sets. Also use a dot to show an element of the first set that satisfies the given relationship.

a) All cats (C) are nimble (N).
b) Some mathematicians (M) are extroverts (E).
c) Some actors (A) are not famous (F).
d) No alligators (A) are trustworthy (T).

50. In Exercises i–xvi, use an Euler diagram to determine whether the argument is valid or invalid.

i. All frogs are poetical.
<u>Kermit is a frog.</u>
∴Kermit is poetical.

ii. All Oreo cookies have a filling.
<u>All Fig Newtons have a filling.</u>
∴All Fig Newtons are Oreo cookies.

iii. Some plants have flowers.
<u>All things that have flowers are beautiful.</u>
∴Some plants are beautiful.

iv. No squares are triangles.
Some triangles are equilateral.
∴No squares are equilateral.

v. No rocker would do the Mariachi.
<u>All baseball fans do the Mariachi.</u>
∴No rocker is a baseball fan.

vi. Nuclear energy is not safe.
Some electric energy is safe.
∴No electric energy is nuclear energy.

vii. Some birds bite.
<u>All things that bite are dangerous.</u>
∴Some birds are dangerous.

viii. All fish can swim.
<u>That barracuda can swim.</u>
∴That barracuda is a fish.

ix. All men behave badly.
<u>Some hockey players behave badly.</u>
∴Some hockey players are men.

x. All grass is green.
<u>That ground cover is not green.</u>
∴That ground cover is not grass.

xi. Most teenagers drink soda.
<u>No CEOs drink soda.</u>
No CEO is a teenager.

xii. Some students like history.
<u>Vern is a student.</u>
∴Vern likes history.

xiii. No mathematics test is fun.
<u>All fun things are worth your time.</u>
∴No mathematics test is worth your time.

xiv. All prudent people shun sharks.
<u>No accountant is imprudent.</u>
∴No accountant fails to shun sharks.

xv. All candidates without a master's degree will not be considered for the position of director.
<u>All candidates who are not considered for the position of director should apply for the position of assistant.</u>
∴All candidates without a master's degree should apply for the position of assistant.

xvi. Some whales make good pets. Some good pets are cute.
<u>Some cute pets bite.</u>
∴Some whales bit

CHAPTER

3

MODELS

LINEAR & NON - LINEAR

All around us, we observe relationships in which some quantities are determined by others. The temperature outside changes with the time of day, different prices of gasoline can be matched to different dates in the past, and the value of a used car may depend on how long ago it was built. In mathematics, many of these relation-ships are considered functions. One of the most common types of functions is a linear function, in which consistent changes in one value cause consistent changes in the related value.

For example, a rule of thumb that is often used to predict the maximum heart rate for women who exercise is to subtract the age from 226. So if A is the age of a woman, her maximum predicted heart rate R is $R = 226 - A$. According to this rule, every time a woman gets one year older, her maximum heart rate decreases by one beat per minute. This consistent change is the defining characteristic of a linear function. We will see many additional examples of linear functions in this chapter.

In this chapter we will learn about;

- Deterministic and population growth models
- Introduction to rectangular coordinates system
- Introduction to functions, their graphs and slopes
- Finding linear and non – linear models
- Exponential and logarithmic functions

Deterministic Models

Deterministic model is a mathematical model that predicate the outcome of the system based on a set of fixed inputs. That is Deterministic model describes systems where the output is uniquely determined by the input, with no randomness or uncertainty. For example Y = a + bx and Area = πr^2 are examples of deterministic models. Solved examples are as follows.

Characteristics

1. Predictable outcomes
2. No randomness or uncertainty
3. Unique solution
4. Cause-and-effect relationships

1. Find the output of the system $y = 2x + 3$ when $x = 4$.
Solution. $y = 11$

2. Solve the differential equation $dy/dx = 2x$, given $y(0) = 1$.
Solution. $y = x^2 + 1$

3. Evaluate the function $f(x) = 3x^2 - 2x + 1$ at $x = 2$.
Solution. $f(2) = 9$

4. Find the equilibrium point of the system $x' = 2x - 3$.
Solution. $x = 3/2$

5. Solve the system of linear equations:
$2x + 3y = 7$; $x - 2y = -3$
Solution. $x = 1, y = 2$

6. Find the derivative of $f(x) = 4x^3 - 2x^2 + x$.
Solution. $f'(x) = 12x^2 - 4x + 1$

7. Find the critical points of $f(x) = x^3 - 2x^2 - 5x + 1$.
Solution. $x \approx 2.53$ or $x \approx -0.87$

Population Growth Models

Population growth models describe the change in population size over time.

Types of Models

- Exponential Growth Model
- Logistic Growth Model
- Malthusian Growth Model
- Verhulst Model

8. **Exponential Growth Model:** Find the population size after 5 years, given an initial population of 1000, growth rate of 0.05, and exponential growth.
   Solution.
   The exponential growth model is given by $P(t) = P_0\, e^{rt}$, where P_0 is the initial population, r is the growth rate, and t is time.
   $P(5) = 1000\ e^{(0.05\times5)} \approx 1276.78$

9. **Logistic Growth Model:** Solve the logistic growth equation $\frac{dP}{dt} = rP(1 - \frac{P}{K})$, with r = 0.2, K = 1000, and P(0) = 500.
   Solution.
   The logistic growth model accounts for carrying capacity (K). The solution involves separating variables and integrating.
   $P(t) = \frac{1000}{(1 + e^{-0.2t})}$

10. **Malthusian Growth Model**
    Evaluate the Malthusian growth model $P(t) = P_0\, e^{rt}$, with $P_0 = 500$, r = 0.03, and t = 10.
    Solution.
    The Malthusian growth model assumes exponential growth without limits.
    $P(10) \approx 674.03$

11. **Verhulst Model:** Solve the Verhulst equation $\frac{dP}{dt} = rP\left(1 - \frac{P}{K}\right)^2$, with r = 0.2, K = 1000, and P(0) = 500.
    Solution.
    The Verhulst model modifies the logistic growth model with a quadratic term.
    $P(t) = \frac{1000}{(1 + (1 - 0.5)e^{-0.2t})}$

12. Population Doubling Time: Find the population doubling time for an exponential growth model with r = 0.04.

Solution.

The population doubling time is calculated using the formula $T = \frac{\ln(2)}{r}$.

$T = \frac{\ln(2)}{r} \approx 17.33$ years

13.Comparative Growth Rates: Compare the growth rates of two populations, one growing exponentially (r = 0.05) and one logistically (r = 0.05, K = 1000).

Solution.

Exponential growth is faster initially, but logistic growth slows due to carrying capacity. Exponential growth initially faster, but logistic growth slows as population approaches carrying capacity.

14.Equilibrium Population Size: Find the equilibrium population size for a logistic growth model with r = 0.1 and K = 500.

Solution.

The equilibrium population size occurs when $\frac{dP}{dt} = 0$.

P = K = 500

15.Differential Equation Solution: Solve the differential equation $\frac{dP}{dt} = 0.02P - 0.001P^2$, with P(0) = 100.

Solution.

The solution involves separating variables and integrating.

$P(t) = \frac{100}{(1 + 0.01e^{-0.02t})}$

16. Model Comparison: Compare the predictions of exponential, logistic, and Malthusian growth models.

Solution.

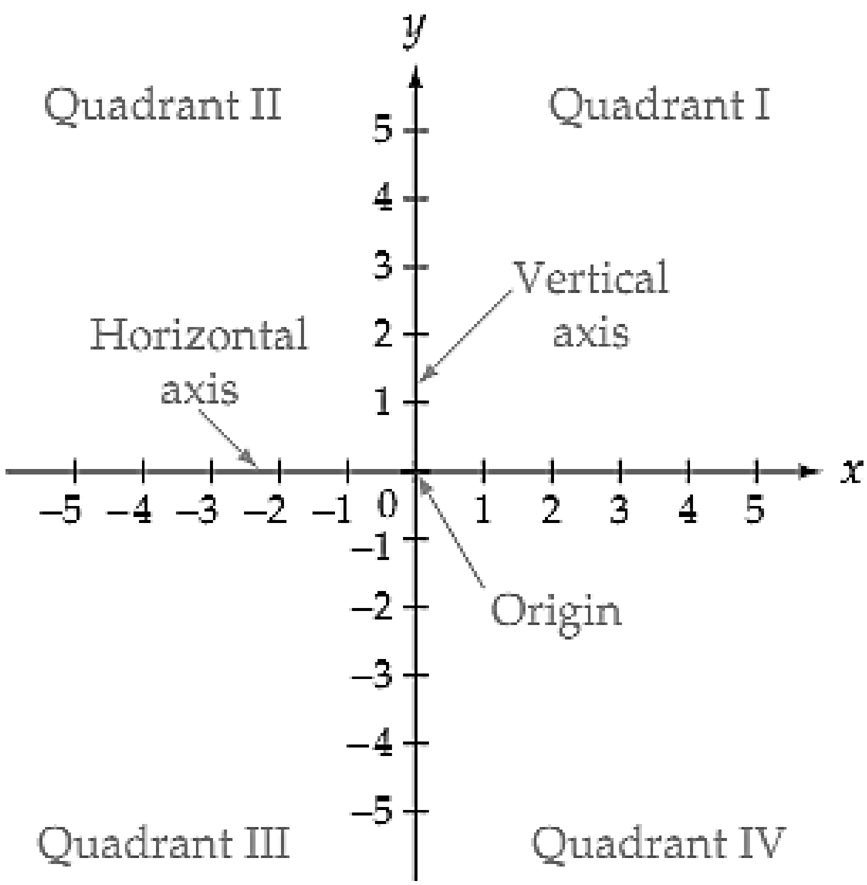
Exponential growth overestimates, logistic growth accounts for carrying capacity, and Malthusian growth underestimates.

Introduction to Rectangular Coordinate Systems

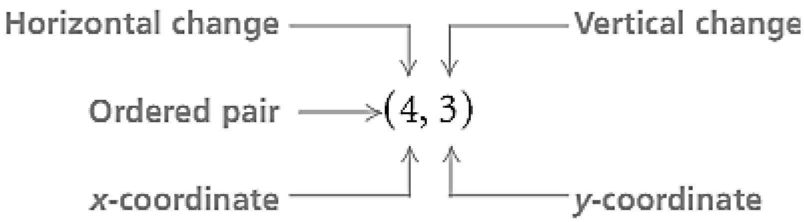
In mathematics we encounter a problem of locating a point in a plane. One way to solve the problem is to use a rectangular coordinate system.

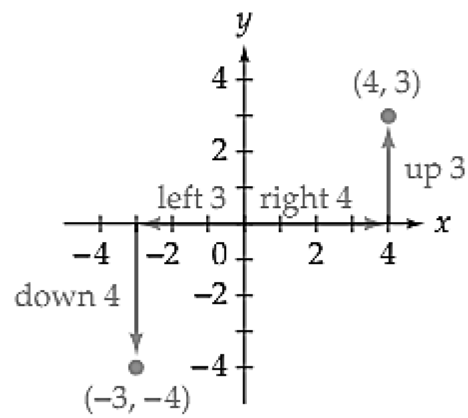
A **rectangular coordinate system** is formed by two number lines, one horizontal and one vertical, that intersect at the zero point of each line. The point of intersection is called the **origin**. The two number lines are called the **coordinate axes**, or simply the **axes**. Frequently, the horizontal axis is labeled the x-axis and the vertical axis is labeled the y-axis. In this case, the axes form what is called the **xy-plane**.

The two axes divide the plane into four regions called **quadrants**, which are numbered counterclockwise, using Roman numerals, from I to IV, starting at the upper right.

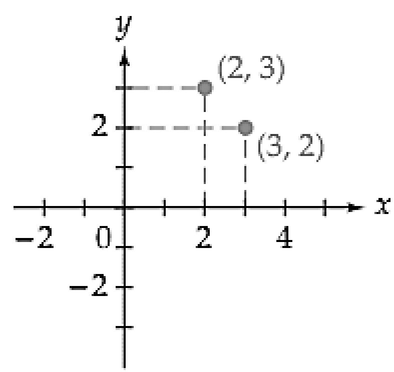


Each point in the plane can be identified by a pair of numbers called an ordered pair. An **ordered pair** is a pair of coordinates, and the order in which the coordinates are listed matters. The first number of the ordered pair measures a horizontal change from the y-axis and is called the **abscissa**, or **x-coordinate**. The second number of the ordered pair measures a vertical change from the x-axis and is called the **ordinate**, or **y-coordinate**. The ordered pair (x, y) associated with a point is also called the **coordinates** of the point.





To **graph**, or **plot**, a point means to place a dot at the coordinates of the point. For example, to graph the ordered pair $(4,3)$ start at the origin. Move 4 units to the right and then 3 units up. Draw a dot. To graph $(-3,-4)$ start at the origin. Move 3 units left and then 4 units down. Draw a dot.
The **graph of an ordered pair** is the dot drawn at the coordinates of the point in the plane. The graphs of the ordered pairs $(4,3)$ and $(-3,-4)$ are shown at the upper right. The graphs of the points whose coordinates Are $(2,3)$ and $(3,2)$ are shown at the right. Note that they are different points. The order in which the numbers in an ordered pair are listed is important.



If the axes are labeled with letters other than x and y, then we refer to the ordered pair using the given labels. For instance, if the horizontal axis is labeled t and the vertical axis is labeled d, then the ordered pairs are written as (t, d). We sometimes refer to the first number in an ordered pair as the **first coordinate** of the ordered pair and to the second number as the **second coordinate** of the ordered pair.

One purpose of a coordinate system is to draw a picture of the solutions of an equation in two variables. Examples of equations in two variables are shown at the right.

$$y = 3x - 2$$
$$x^2 + y^2 = 25$$
$$s = t^2 - 4t + 1$$

A **solution of an equation in two variables** is an ordered pair that makes the equation a true statement. For instance, as shown below, $(2,4)$ is a solution of $y = 3x - 2$ but $(3,-1)$ is not a solution of the equation.

$y = 3x - 2$			$y = 3x - 2$		
4	$3(2) - 2$	• $x = 2, y = 4$	-1	$3(3) - 2$	• $x = 3, y = -1$
4	$6 - 2$		-1	$9 - 2$	
$4 = 4$		• Checks.	$-1 \neq 7$		• Does not check.

1. Is $(-2,1)$ a solution of $y = 3x + 7$?

Answer.

Yes, because $1 = 3(-2) + 7$.

The **graph of an equation in two variables** is a drawing of all the ordered-pair solutions of the equation. Many equations can be graphed by finding some ordered-pair solutions of the equation, plotting the corresponding points, and then connecting the points with a smooth curve.

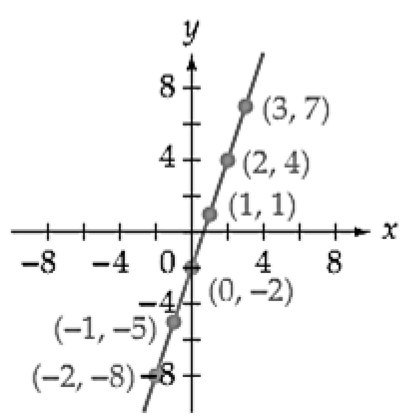
2. Graph $y = 3x - 2$.

Solution.

To find ordered-pair solutions, select various values of x and calculate the corresponding values of y. Plot the ordered pairs. After the ordered pairs have been graphed, draw a smooth curve through the points. It is convenient to keep track of the solutions in a table.

When choosing values of x,we often choose integer values because the resulting ordered pairs are easier to graph.

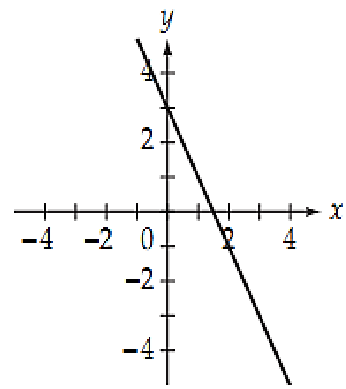
x	$3x - 2 = y$	(x, y)
-2	$3(-2) - 2 = -8$	$(-2, -8)$
-1	$3(-1) - 2 = -5$	$(-1, -5)$
0	$3(0) - 2 = -2$	$(0, -2)$
1	$3(1) - 2 = 1$	$(1, 1)$
2	$3(2) - 2 = 4$	$(2, 4)$
3	$3(3) - 2 = 7$	$(3, 7)$



3. Graph $y = -2x + 3$.

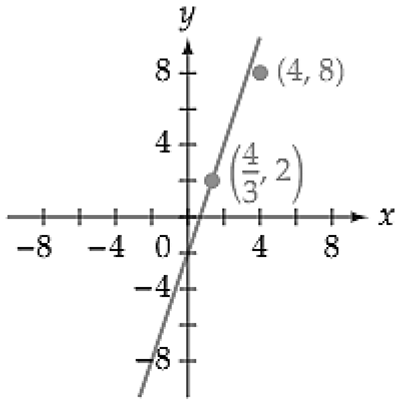
Solution.

x	$-2x + 3 = y$	(x, y)
-2	$-2(-2) + 3 = 7$	$(-2, 7)$
-1	$-2(-1) + 3 = 5$	$(-1, 5)$
0	$-2(0) + 3 = 3$	$(0, 3)$
1	$-2(1) + 3 = 1$	$(1, 1)$
2	$-2(2) + 3 = -1$	$(2, -1)$
3	$-2(3) + 3 = -3$	$(3, -3)$



Note

The graph of $y = 3x - 2$ is shown below. Note that the ordered pair $\left(\frac{4}{3}, 2\right)$ is a solution of the equation and is a point on the graph. The ordered pair $(4,8)$ is not a solution of the equation and is not a point on the graph. **Every ordered-pair solution of the equation is a point on the graph, and every point on the graph is an ordered-pair solution of the equation**.

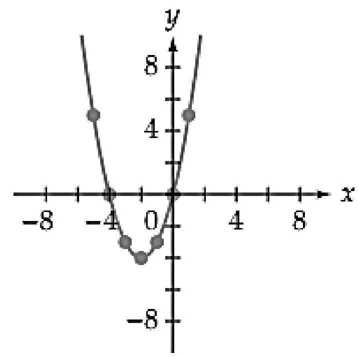


4. Graph $y = x^2 + 4x$.

Solution.

Select various values of x and calculate the corresponding values of y. Plot the ordered pairs. After the ordered pairs have been graphed, draw a smooth curve through the points. The following table shows some ordered pair solutions.

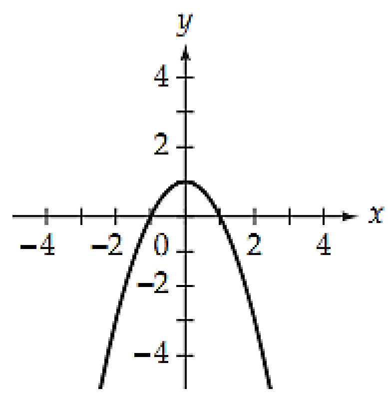
x	$x^2 + 4x = y$	(x, y)
−5	$(-5)^2 + 4(-5) = 5$	(−5, 5)
−4	$(-4)^2 + 4(-4) = 0$	(−4, 0)
−3	$(-3)^2 + 4(-3) = -3$	(−3, −3)
−2	$(-2)^2 + 4(-2) = -4$	(−2, −4)
−1	$(-1)^2 + 4(-1) = -3$	(−1, −3)
0	$(0)^2 + 4(0) = 0$	(0, 0)
1	$(1)^2 + 4(1) = 5$	(1, 5)



5. Graph $y = -x^2 + 1$.

Solution.

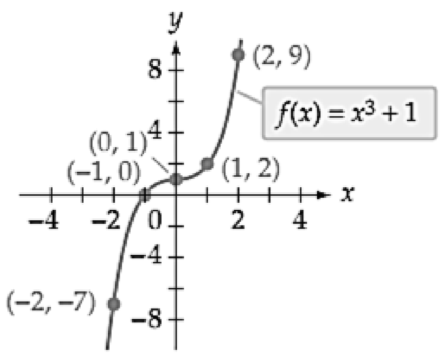
x	$-x^2 + 1 = y$	(x, y)
−3	$-(-3)^2 + 1 = -8$	(−3, −8)
−2	$-(-2)^2 + 1 = -3$	(−2, −3)
−1	$-(-1)^2 + 1 = 0$	(−1, 0)
0	$-(0)^2 + 1 = 1$	(0, 1)
1	$-(1)^2 + 1 = 0$	(1, 0)
2	$-(2)^2 + 1 = -3$	(2, −3)
3	$-(3)^2 + 1 = -8$	(3, −8)



6. Graph $f(x) = x^3 + 1$.

Solution.

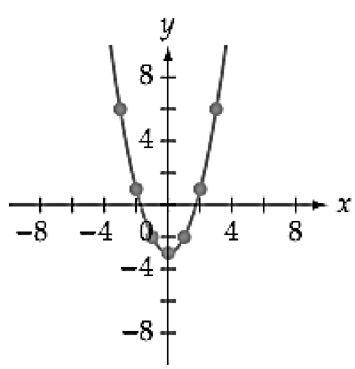
x	$f(x) = x^3 + 1$	(x, y)
-2	$f(-2) = (-2)^3 + 1 = -7$	$(-2, -7)$
-1	$f(-1) = (-1)^3 + 1 = 0$	$(-1, 0)$
0	$f(0) = (0)^3 + 1 = 1$	$(0, 1)$
1	$f(1) = (1)^3 + 1 = 2$	$(1, 2)$
2	$f(2) = (2)^3 + 1 = 9$	$(2, 9)$



7. Graph $f(x) = x^2 - 3$.

Solution.

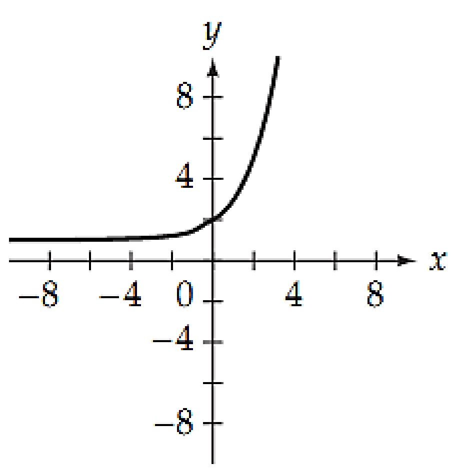
x	$h(x) = x^2 - 3$	(x, y)
-3	$h(-3) = (-3)^2 - 3 = 6$	$(-3, 6)$
-2	$h(-2) = (-2)^2 - 3 = 1$	$(-2, 1)$
-1	$h(-1) = (-1)^2 - 3 = -2$	$(-1, -2)$
0	$h(0) = (0)^2 - 3 = -3$	$(0, -3)$
1	$h(1) = (1)^2 - 3 = -2$	$(1, -2)$
2	$h(2) = (2)^2 - 3 = 1$	$(2, 1)$
3	$h(3) = (3)^2 - 3 = 6$	$(3, 6)$



8. Graph $f(x) = 2 - \frac{3}{4}x$.

Solution.

x	$f(x) = 2 - \frac{3}{4}x$	(x, y)
-3	$f(-3) = 2 - \frac{3}{4}(-3) = 4\frac{1}{4}$	$\left(-3, 4\frac{1}{4}\right)$
-2	$f(-2) = 2 - \frac{3}{4}(-2) = 3\frac{1}{2}$	$\left(-2, 3\frac{1}{2}\right)$
-1	$f(-1) = 2 - \frac{3}{4}(-1) = 2\frac{3}{4}$	$\left(-1, 2\frac{3}{4}\right)$
0	$f(0) = 2 - \frac{3}{4}(0) = 2$	$(0, 2)$
1	$f(1) = 2 - \frac{3}{4}(1) = 1\frac{1}{4}$	$\left(1, 1\frac{1}{4}\right)$
2	$f(2) = 2 - \frac{3}{4}(2) = \frac{1}{2}$	$\left(2, \frac{1}{2}\right)$
3	$f(3) = 2 - \frac{3}{4}(3) = -\frac{1}{4}$	$\left(3, -\frac{1}{4}\right)$



Introduction to Functions

A **function** is a correspondence, or relationship, between two sets called the **domain** and **range** such that for each element of the domain there corresponds exactly one element of the range.

Or Let A and B be two non-empty sets such that:

i) f is a relation from A to B that is , f is a subset of $A \times B$
ii) Dom f = A
iii) First element of no two pairs of f are equal, then f is said to be a function from A to B.

The function f is also written as: $f: A \rightarrow B$

which is read: f is a function from A to B .

If (x, y) in an element of f when regarded as a set of ordered pairs,

we write y = f(x). y is called the **dependent** value of f for x that is **independent** value or image of x under f.

The process of finding f(x) for a given value of x is called **evaluating the function**.

Examples

The ordered pairs, the graph, and the equation are all different ways of expressing the correspondence, or relationship, between the two variables. These are called a function. Here are some additional examples of functions, along with a specific example of each correspondence.

To each real number 5	there corresponds ⟶	its square 25
To each score on an exam 87	there corresponds ⟶	a grade B
To each student Alexander Sterling	there corresponds ⟶	a student identification number S18723519

An important fact about each of these correspondences is that each result is unique. For instance, for the real number 5, there is exactly one square, 25.

9. Evaluate $f(t) = 2t^2 - 3t + 1$ when $t = -2$.

Solution.

$f(t) = 2t^2 - 3t + 1$

$f(-2) = 2(-2)^2 - 3(-2) + 1 = 15$

10. Evaluate $f(z) = z^2 - z$ when $z = -3$.

Solution.

$f(t) = 2t^2 - 3t + 1$

$f(-3) = (-3)^2 - (-3) = 12$

11. The surface area of a cube (the sum of the areas of each of the six faces) is given by $SA(s) = 6s^2$ where $SA(s)$ is the surface area of the cube and s is the length of one side of the cube. Find the surface area of a cube that has a side of length 10 centimeters.

Solution.

$SA(s) = 6s^2$

$SA(10) = 6(10)^2 = 600$

The surface area of the cube is 600 square centimeters.

12. A **diagonal** of a polygon is a line segment from one vertex to a nonadjacent vertex, as shown. The total number of diagonals of a polygon is given by $N(s) = \frac{s^2-3s}{2}$ where $N(s)$ is the total number of diagonals and s is the number of sides of the polygon. Find the total number of diagonals of a polygon with 12 sides.

Solution.

$N(s) = \frac{s^2-3s}{2}$

$N(s) = \frac{(12)^2-3(12)}{2} = 54$

A polygon with 12 sides has 54 diagonals.

Linear Functions and Linear Growth Models

A linear function is one that can be written in the form $f(x) = mx + b$, where m is the coefficient of x and b is a constant.

Here are some other examples of linear functions.

$f(x) = 2x + 5$	• $m = 2, b = 5$
$g(t) = \frac{2}{3}t - 1$	• $m = \frac{2}{3}, b = -1$
$v(s) = -2s$	• $m = -2, b = 0$
$h(x) = 3$	• $m = 0, b = 3$
$f(x) = 2 - 4x$	• $m = -4, b = 2$

13. Are the given functions linear functions?

a. $f(x) = 2x^2 + 5$ ***b.*** $g(x) = 1 - 3x$

Solution.

a. Because $f(x) = 2x^2 + 5$ has an x^2 term, f is not a linear function.

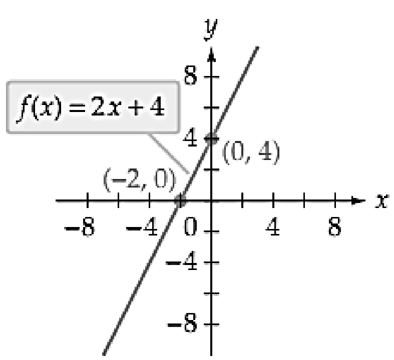
b. Because $g(x) = 1 - 3x$ can be written in the form $f(x) = mx + b$ as $g(x) = -3x + 1$ (m = −3 and b = 1),, g is a linear function.

Note

Note that **the graph of a linear function is a straight**. Observe that when the graph crosses the x-axis, the y-coordinate is 0. When the graph crosses the y-axis, the x-coordinate is 0. The table confirms these observations.

14. Graph $f(x) = 2x + 4$

Solution.



x	$f(x) = 2x + 4$	(x, y)
−3	$f(-3) = 2(-3) + 4 = -2$	(−3, −2)
−2	$f(-2) = 2(-2) + 4 = 0$	(−2, 0)
−1	$f(-1) = 2(-1) + 4 = 2$	(−1, 2)
0	$f(0) = 2(0) + 4 = 4$	(0, 4)
1	$f(1) = 2(1) + 4 = 6$	(1, 6)

15. Find the x- and y-intercepts of the graph $g(\mathrm{x}) = -3\mathrm{x} + 2$

Solution.

For x – intercept: Put $\boldsymbol{y} = \boldsymbol{g}(\mathbf{x}) = \mathbf{0}$

$g(\mathrm{x}) = -3\mathrm{x} + 2 \Rightarrow -3\mathrm{x} + 2 = 0 \Rightarrow \mathrm{x} = \frac{2}{3} \Rightarrow$ The x-intercept is $\left(\frac{2}{3}, 0\right)$

For y – intercept: Put $\boldsymbol{x} = \mathbf{0}$

$g(\mathrm{x}) = -3\mathrm{x} + 2 \Rightarrow g(0) = -3(0) + 2 \Rightarrow \mathrm{y} = 2 \Rightarrow$ The y-intercept is $(0,2)$

16. Find the x- and y-intercepts of the graph $g(\mathrm{x}) = \frac{1}{2}\mathrm{x} + 3$

Solution.

For x – intercept: Put $\boldsymbol{y} = \boldsymbol{g}(\mathbf{x}) = \mathbf{0}$

$g(\mathrm{x}) = \frac{1}{2}\mathrm{x} + 3 \Rightarrow \frac{1}{2}\mathrm{x} + 3 = 0 \Rightarrow \mathrm{x} = -6 \Rightarrow$ The x-intercept is $(-6,0)$

For y – intercept: Put $\boldsymbol{x} = \mathbf{0}$

$g(\mathrm{x}) = \frac{1}{2}\mathrm{x} + 3 \Rightarrow g(0) = \frac{1}{2}(0) + 3 \Rightarrow \mathrm{y} = 3 \Rightarrow$ The y-intercept is $(0,3)$

17. After a parachute is deployed, a function that models the height of the parachutist above the ground is $f(\mathrm{t}) = -10\mathrm{t} + 2800$ where $f(\mathrm{t})$ is the height, in feet, of the parachutist t seconds after the parachute is deployed. Find the intercepts on the vertical and horizontal axes and explain what they mean in the context of the problem.

Solution.

For Horizontal – intercept: Put $\boldsymbol{y} = \boldsymbol{f}(\mathbf{t}) = \mathbf{0}$

$f(\mathrm{t}) = -10\mathrm{t} + 2800 \Rightarrow -10\mathrm{t} + 2800 = 0 \Rightarrow \mathrm{t} = 280$

The intercept on the horizontal axis is $(280,0)$. This means that the parachutist reaches the ground 280 seconds after the parachute is deployed.

For Vertical – intercept: Put $\boldsymbol{t} = \mathbf{0}$

$f(\mathrm{t}) = -10\mathrm{t} + 2800 \Rightarrow f(0) = -10(0) + 2800 \Rightarrow \mathrm{y} = 2800$

The intercept on the vertical axis is $(0,2800)$. This means that the parachutist is 2800 feet above the ground when the parachute is deployed.

Note that the parachutist reaches the ground when $f(\mathrm{t}) = 0$.

18. A function that models the descent of a certain small airplane is given by $g(\mathrm{t}) = -20\mathrm{t} + 8000$ where $g(\mathrm{t})$ is the height, in feet, of the airplane t seconds after it begins its descent. Find the intercepts on the vertical and horizontal axes, and explain what they mean in the context of the problem.

Solution.

For Horizontal – intercept: Put $y = g(\mathrm{t}) = 0$

$g(\mathrm{t}) = -20\mathrm{t} + 8000 \Rightarrow -20\mathrm{t} + 8000 = 0 \Rightarrow \mathrm{t} = 400$

The intercept on the horizontal axis is $(400,0)$. This means that the plane reaches the ground 400 seconds after beginning its descent.

For Vertical – intercept: Put $t = 0$

$g(\mathrm{t}) = -20\mathrm{t} + 8000 \Rightarrow g(0) = -20(0) + 8000 \Rightarrow \mathrm{y} = 8000$

The intercept on the vertical axis is $(0,8000)$. This means that the plane is at an altitude of 8000 feet when it begins its descent.

Slope of Graph of linear Function

For a linear function given by the slope of the graph of the function $f(\mathrm{x}) = \mathrm{mx} + \mathrm{b}$ is m, the coefficient of the variable.

19. What is the slope of each of the following?

a. $y = -2x + 3$ **b.** $f(x) = x + 4$ **c.** $g(x) = 3 - 4x$

d. $y = \frac{1}{2}x - 5$

Solution.

a. -2 **b.** 1 **c.** -4 **d.** $\frac{1}{2}$

Slope of a Line

Let $(\mathrm{x}_1, \mathrm{y}_1)$ and $(\mathrm{x}_2, \mathrm{y}_2)$ be two points on a nonvertical line. Then the **slope** of the line through the two points is the ratio of the change in the y-coordinates to the change in the x-coordinates.

$\mathrm{m} = \frac{\text{change in y}}{\text{change in x}} = \frac{\mathrm{y}_2 - \mathrm{y}_1}{\mathrm{x}_2 - \mathrm{x}_1}$; $\mathrm{x}_1 \neq \mathrm{x}_2$

20. Why is the restriction $\mathrm{x}_1 \neq \mathrm{x}_2$ required in the definition of slope?

Answer.

If $\mathrm{x}_1 = \mathrm{x}_2$ then the difference $\mathrm{x}_2 - \mathrm{x}_1 = 0$. This would make the denominator 0, and division by 0 is not defined.

21. Find the slope of the line between the two points.

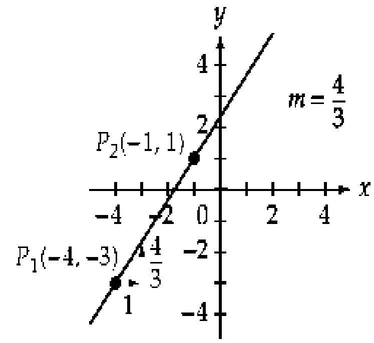
a. $(-4, -3)$ and $(-1, 1)$ **b.** $(-2, 3)$ and $(1, -3)$

c. $(-1, -3)$ and $(4, -3)$ **d.** $(4, 3)$ and $(4, -1)$

Solution.

a. $(x_1, y_1) = (-4, -3), (x_2, y_2) = (-1, 1)$

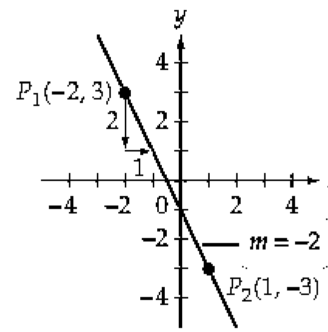
$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{1 - (-3)}{-1 - (-4)} = \frac{4}{3}$$



The slope is $\frac{4}{3}$. A **positive** slope indicates that the line slopes upward to the right. For this particular line, the value of y increases by $\frac{4}{3}$ when x increases by 1.

b. $(x_1, y_1) = (-2, 3), (x_2, y_2) = (1, -3)$

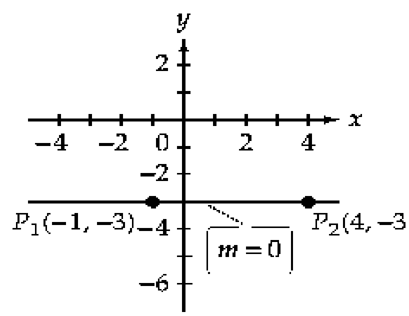
$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-3 - 3}{1 - (-2)} = \frac{-6}{3} = -2$$



The slope is -2. A **negative** slope indicates that the line slopes downward to the right. For this particular line, the value of y decreases by 2 when x increases by 1.

c. $(x_1, y_1) = (-1, -3), (x_2, y_2) = (4, -3)$

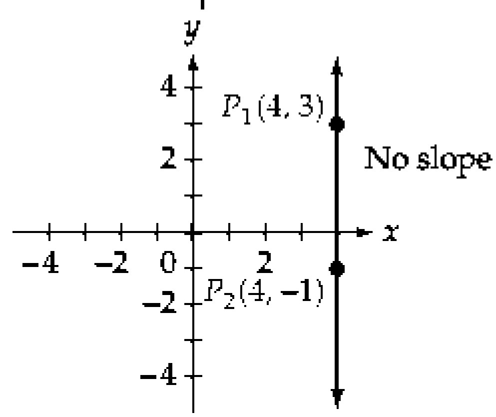
$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-3 - (-3)}{4 - (-1)} = \frac{0}{5} = 0$$



The slope is 0. A zero slope indicates that the line is **horizontal**. For a horizontal line, the value of y stays the same when x increases by any amount.

d. $(x_1, y_1) = (4, 3), (x_2, y_2) = (4, -1)$

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-1 - 3}{4 - 4} = \frac{-4}{0}$$



If the denominator of the slope formula is zero, the line has no slope. Sometimes we say that the slope of a vertical line is **undefined**.

22. Find the slope of the line between the two points.

a. $(-6, 5)$ and $(4, -5)$ **b.** $(-5, 0)$ and $(-5, 7)$

c. $(-7, -2)$ and $(8, 8)$ **d.** $(-6, 7)$ and $(1, 7)$

Solution.

a. $(x_1, y_1) = (-6, 5), (x_2, y_2) = (4, -5)$

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-5 - 5}{4 - (-6)} = \frac{-10}{10} = -1$$

The slope is -1.

b. $(x_1, y_1) = (-5, 0), (x_2, y_2) = (-5, 7)$

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{7 - 0}{-5 - (-5)} = \frac{7}{0}$$

The slope is undefined.

c. $(x_1, y_1) = (-7, -2), (x_2, y_2) = (8, 8)$

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{8 - (-2)}{8 - (-7)} = \frac{10}{15} = \frac{2}{3}$$

The slope is $\frac{2}{3}$.

d. $(x_1, y_1) = (-6, 7), (x_2, y_2) = (1, 7)$

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{7 - 7}{1 - (-6)} = \frac{0}{7} = 0$$

The slope is 0.

23. The function $T(x) = -6.5x + 20$ approximates the temperature $T(x)$, in degrees Celsius, at x kilometers above sea level. What is the slope of this function? Write a sentence that explains the meaning of the slope in the context of this application.

Solution.

For the linear function $T(x) = -6.5x + 20$ the slope is the coefficient of x. Therefore, the slope is -6.5. The slope means that the temperature is decreasing (because the slope is negative) 6.5°C for each 1-kilometer increase in height above sea level.

24. The distance that a homing pigeon can fly can be approximated by $d(t) = 50t$ where $d(t)$ is the distance, in miles, flown by the pigeon in t hours. Find the slope of this function. What is the meaning of the slope in the context of the problem?

Solution.

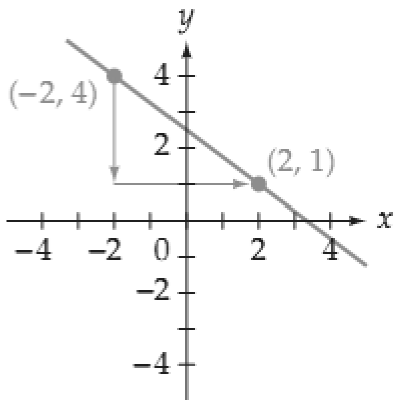
For the linear function $d(t) = 50t$ the slope is the coefficient of t. Therefore, the slope is 50. This means that a homing pigeon can fly 50 miles for each 1 hour of flight time.

Slope Intercept Form of a Straight Line

Let m be the slope and c be the y – intercept of a non – vertical line then the equation of line is $f(x) = y = mx + b$, this equation is called the slope–intercept form of a straight line.

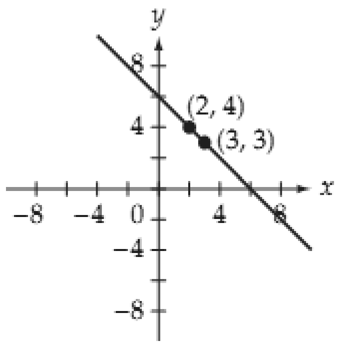
25. Draw the line that passes through $(-2,4)$ and has slope $-\frac{3}{4}$.

Solution.



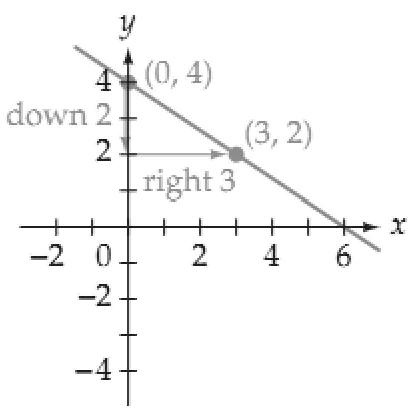
26. Draw the line that passes through $(2,4)$ and has slope -1.

Solution.



27. Graph $f(x) = -\frac{2}{3}x + 4$ by using the slope and y-intercept.

Solution.

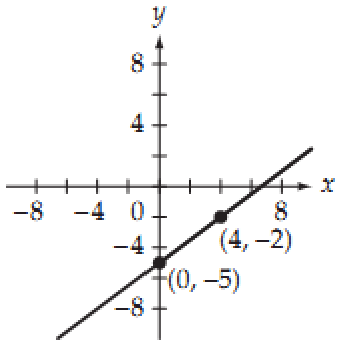


y-intercept = $(0,4)$

slope = $m = -\frac{2}{3} = \frac{-2}{3}$

28. Graph $f(x) = \frac{3}{4}x - 5$ by using the slope and y-intercept.

Solution.



29. Suppose that a car uses 0.04 gallon of gas per mile driven and that the fuel tank, which holds 18 gallons of gas, is full. Using this information, determine a linear model for the amount of fuel remaining in the gas tank after driving x miles.

Solution.

The slope is the rate at which the car is using fuel, 0.04 gallon per mile. Because the car is consuming the fuel, the amount of fuel in the tank is decreasing. Therefore, the slope is negative and we have m = − 0.04.

The amount of fuel in the tank depends on the number of miles x the car has been driven. Before the car starts (that is, when x = 0), there are 18 gallons of gas in the tank.

The y-intercept is $(0,18)$

Using this information, we can create the linear function.

$f(x) = mx + b$

$f(x) = -0.04x + 18$

The linear function that models the amount of fuel remaining in the tank is given by $f(x) = -0.04x + 18$ where $f(x)$ is the amount of fuel, in gallons, remaining after driving x miles. The graph of the function is shown follows.

The x-intercept of a graph is the point at which $f(x) = 0$.

For this application, $f(x) = 0$ when there are 0 gallons of fuel remaining in the tank.

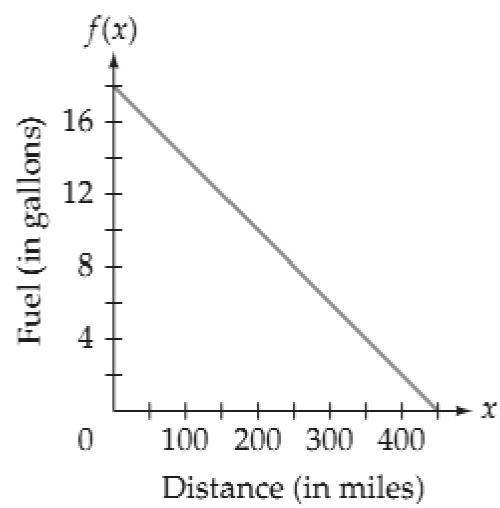
For x – intercept: Put $\boldsymbol{y = f(x) = 0}$

$f(x) = -0.04x + 18$

$\Rightarrow -0.04x + 18 = 0$

$\Rightarrow x = 450$

The car can travel 450 miles before running out of gas.



30. Suppose a 20-gallon gas tank contains 2 gallons when a motorist decides to fill up the tank. If the gas pump fills the tank at a rate of 0.1 gallon per second, find a linear function that models the amount of fuel in the tank t seconds after fueling begins.

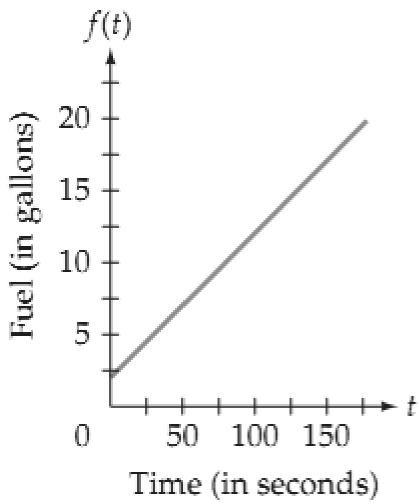
Solution.

When fueling begins, at t = 0 there are 2 gallons of gas in the tank. Therefore, the y-intercept is $(0,2)$. The slope is the rate at which fuel is being added to the tank. Because the amount of fuel in the tank is increasing, the slope is positive and we have m = 0.1. To find the linear function, replace m and b by their respective values.

$f(\mathrm{t}) = \mathrm{mt} + \mathrm{b}$

$f(\mathrm{t}) = 0.1\mathrm{t} + 2$

The linear function is $f(\mathrm{t}) = 0.1\mathrm{t} + 2$ where $f(\mathrm{t})$ is the number of gallons of fuel in the tank t seconds after fueling begins.



31. The boiling point of water at sea level is 100°C.

The boiling point decreases 3.5°C per 1 kilometer increase in altitude. Find a linear function that gives the boiling point of water as a function of altitude.

Solution.

$f(\mathrm{a}) = \mathrm{ma} + \mathrm{b}$

$f(\mathrm{a}) = -3.5\mathrm{a} + 100$

The linear function is $f(\mathrm{a}) = -3.5\mathrm{a} + 100$ where $f(\mathrm{a})$ is the boiling point of water in degrees Celcius at an altitude of ‘a’ kilometers above sea level.

Point–Slope Formula of a Straight Line
Let (x_1, y_1) be a point on a line and let m be the slope of the line. Then the equation of the line can be found using the point–slope formula
$y - y_1 = m(x - x_1)$

32. Find the equation of the line that passes through $(1, -3)$ and has slope -2.
Solution.

$$y - y_1 = m(x - x_1)$$
$$y - (-3) = -2(x - 1)$$
$$y + 3 = -2x + 2$$
$$y = -2x - 1$$

33. Find the equation of the line that passes through $(-2,2)$ and has slope $-\frac{1}{2}$.
Solution.
$y - y_1 = m(x - x_1) \Rightarrow y - 2 = -\frac{1}{2}\left(x - (-2)\right) \Rightarrow y = -\frac{1}{2}x + 1$

34. Based on data from the Kelley Blue Book, the value of a certain car decreases approximately \$250 per month. If the value of the car 2 years after it was purchased was \$14,000, find a linear function that models the value of the car after x months of ownership. Use this function to find the value of the car after 3 years of ownership.
Solution.
Let V represent the value of the car after x months. Then $V = 14000$ when (2 years is 24 months). A solution of the equation is $(24, -14000)$.The car is decreasing in value at a rate of \$250 per month. Therefore, the slope is -250. Now use the point–slope formula to find the linear equation that models the function.

$$V - V_1 = m(x - x_1)$$
$$V - 14{,}000 = -250(x - 24)$$
$$V - 14{,}000 = -250x + 6000$$
$$V = -250x + 20{,}000$$

A linear function that models the value of the car after x months of ownership is
$V(x) = -250x + 20{,}000$
To find the value of car after 3 years (36 months), evaluate the function when x = 36.

$$V(x) = -250x + 20{,}000$$
$$V(36) = -250(36) + 20{,}000 = 11{,}000$$

The value of the car is \$11,000 after 3 years of ownership.

35. During a brisk walk, a person burns about 3.8 calories per minute. If a person has burned 191 calories in 50 minutes, determine a linear function that models the number of calories burned after t minutes.

Solution.

$$C - C_1 = m(t - t_1)$$
$$C - 191 = 3.8(t - 50)$$
$$C - 191 = 3.8t - 190$$
$$C = 3.8t + 1$$

A linear function that models the number of calories burned after minutes is

$$C(t) = 3.8t + 1$$

36. Find the equation of the line that passes through $P_1(6, -4)$ and $P_2(3,2)$.

Solution.

Find the slope of the line between the two points.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{2 - (-4)}{3 - 6} = \frac{6}{-3} = -2$$

Use the point–slope formula to find the equation of the line.

$$y - y_1 = m(x - x_1)$$
$$y - (-4) = -2(x - 6)$$
$$y + 4 = -2x + 12$$
$$y = -2x + 8$$

37. Find the equation of the line that passes through $P_1(-2,3)$ and $P_2(4,1)$.

Solution.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{1 - 3}{4 - (-2)} = \frac{-2}{6} = -\frac{1}{3}$$

$$y - y_1 = m(x - x_1)$$

$$y - 3 = -\frac{1}{3}[x - (-2)]$$

$$y - 3 = -\frac{1}{3}x - \frac{2}{3}$$

$$y = -\frac{1}{3}x + \frac{7}{3}$$

Polynomial Function

A function of the form

$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0\ ;\ a_n \neq 0$

Where **n** is a non – negative integer and the coefficients $a_n, a_{n-1}, \ldots, a_1, a_0$ are real numbers. It can be considered as a polynomial function of x.

Degree of the Polynomial

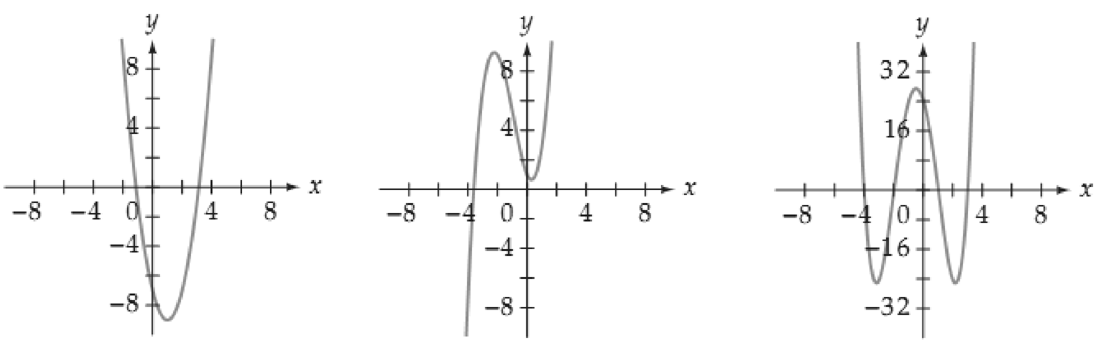
The highest power of x in a polynomial in x is called the **degree of the polynomial**. So the expression is the polynomial of degree n.

Here are some examples of polynomial functions and their degrees.

Polynomial Function	**Degree**
$f(x) = 2x - 3$	1
$g(x) = 2x^2 - 4x - 7$ $f(t) = t^2$	2
$p(x) = 1 - 2x + 3x^2 + x^3$ $s(t) = 5t^3 - t + 8$	3
$z(x) = x^4 + 2x^3 - 13x^2 - 14x + 24$	4

Nonlinear functions

Polynomial functions of degree greater than 1 are part of a class of functions that are called **nonlinear functions** because their graphs are not straight lines. Here are some graphs of polynomial functions of degree greater than 1.



$g(x) = 2x^2 - 4x - 7$
2nd degree

$p(x) = 1 - 2x + 3x^2 + x^3$
3rd degree

$z(x) = x^4 + 2x^3 - 13x^2 - 14x + 24$
4th degree

Just as a first-degree polynomial function is also called a linear function, some of the other polynomial functions have special names.

Polynomial Function	Name
First-degree	Linear function
Second-degree	Quadratic function
Third-degree	Cubic function
Fourth-degree	Quartic function

Fifth-degree polynomials are called quintic polynomials. Generally, however, polynomials of degree 4 and higher are referred to by their degrees. So, for instance, $f(x) = 3x^8 + 4x^5 - 3x^2 + 4$ represents an eighth-degree polynomial function.

38. Let $f(x) = 3x - 2x^2 - 4x^3 + 2$

a. Write the polynomial in standard form.
b. Name the degree of the function.
c. Evaluate the function when x = 3.

Solution.

a. Write the polynomial in decreasing powers of x.
$f(x) = -4x^3 - 2x^2 + 3x + 2$
b. The degree is 3, the largest exponent on the variable.
c. Value of function at x = 3.

$$f(x) = -4x^3 - 2x^2 + 3x + 2$$
$$f(3) = -4(3)^3 - 2(3)^2 + 3(3) + 2$$
$$= -4(27) - 2(9) + 3(3) + 2$$
$$= -108 - 18 + 9 + 2 = -115$$

39. Let $g(t) = -4t + 3t^2 + 5$

a. Write the polynomial in standard form.
b. Name the degree of the function.
c. Evaluate the function when x = 3.

Solution.

a. $g(t) = 3t^2 - 4t + 5$
b. The degree is 2, the largest exponent on the variable.
c. Value of function at x = 3.

$$g(t) = 3t^2 - 4t + 5$$
$$g(3) = 3(3)^2 - 4(3) + 5$$
$$= 27 - 12 + 5$$
$$= 20$$

Things to Remember

- **Constant Polynomial/Equation** has degree **0**.
- **Linear Polynomial/Equation** has degree **1**.
- **Quadratic Polynomial** has degree **2**.
- **Cubic Polynomial/Equation** has degree **3**.
- **Degree 4**, is a **quadratic polynomial**.
- **0** is a **polynomial** of degree **0**.
- A polynomial having **one term** is called **monomial**.
- A polynomial having **two terms** is called **binomial**.
- A polynomial having **three terms** is called **trinomial**.

Quadratic Function

A quadratic function in x is a function that can be written in the form
$f(x) = ax^2 + bx + c$; where a, b and c are real numbers and $a \neq 0$.

Or A quadratic function in x is **2^{nd} Degree Polynomial** in x.

Examples

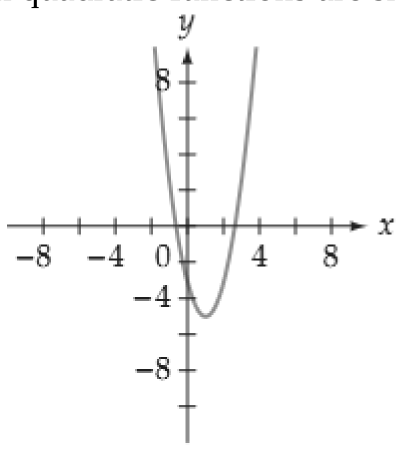
$f(x) = x^2 - 3x + 1$ • $a = 1, b = -3, c = 1$

$g(t) = -2t^2 - 4$ • $a = -2, b = 0, c = -4$

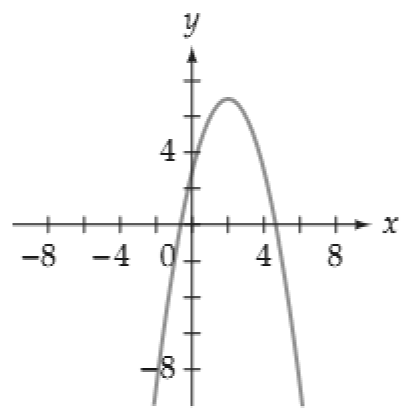
$h(p) = 4 - 2p - p^2$ • $a = -1, b = -2, c = 4$

$f(x) = 2x^2 + 6x$ • $a = 2, b = 6, c = 0$

The graph of a quadratic function in a single variable x is a parabola. The graphs of two such quadratic functions are shown below.



$a > 0$



$a < 0$

The figure on the left is the graph of $f(x) = 2x^2 - 4x - 3$.The value of a is positive $(a = 2)$ and the graph opens up.

The figure on the right is the graph of $f(x) = -x^2 + 4x + 3$. The value of a is negative $(a = -1)$ and the graph opens down.

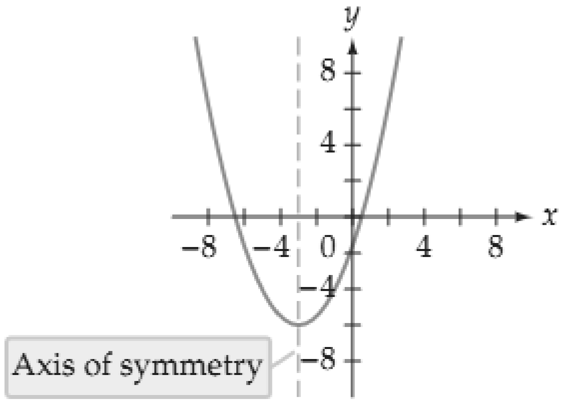
Vertex of the parabola

The point at which the graph of a parabola has a minimum or a maximum is called the vertex of the parabola. The vertex of a parabola is the point with the smallest y-coordinate when a > 0 and the point with the largest y-coordinate when a < 0.

Let $f(x) = ax^2 + bx + c$ be the equation of a parabola. The coordinates of the vertex are $\left(-\frac{b}{2a}, f\left(-\frac{b}{2a}\right)\right)$

Axis of symmetry

The axis of symmetry of the graph of a quadratic function is a vertical line that passes through the vertex of the parabola.



40. Find the vertex of the parabola whose equation is $y = -3x^2 + 6x + 1$.

Solution.

$$x = -\frac{b}{2a} = -\frac{6}{2(-3)} = 1$$

$$y = -3x^2 + 6x + 1$$

$$y = -3(1)^2 + 6(1) + 1$$

$$y = 4$$

The vertex is $(1,4)$

41. Find the vertex of the parabola whose equation is $y = x^2 - 2$.

Solution.

$a = 1, b = 0;\ x = -\frac{b}{2a} = -\frac{0}{2(1)} = 0$

$y = (0)^2 - 2 = -2$

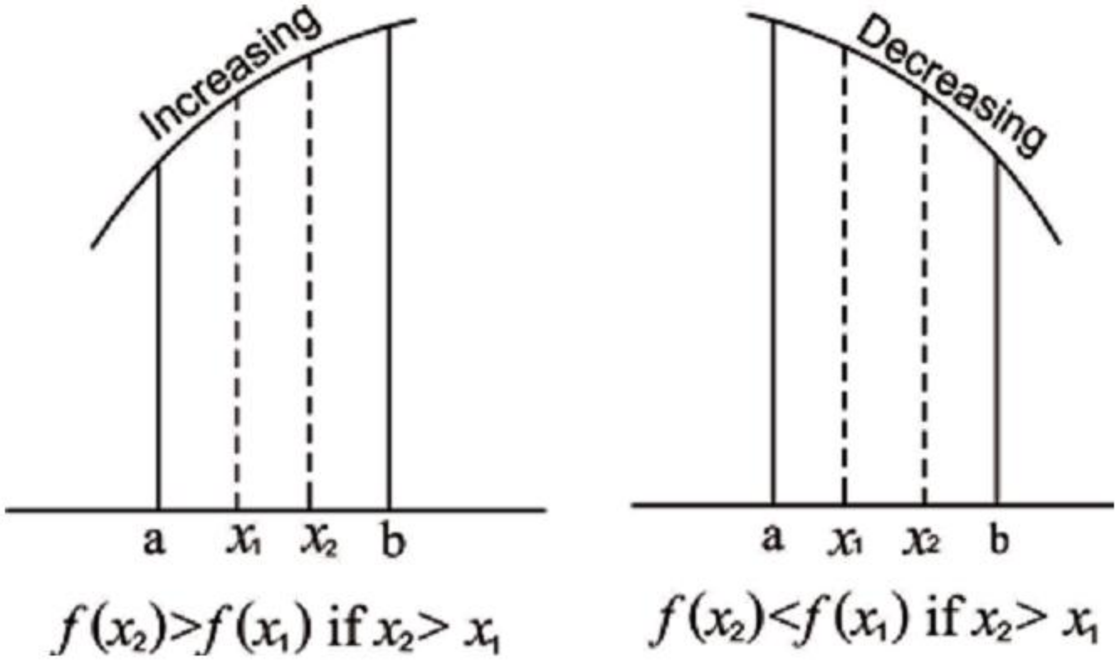
The vertex is $(0, -2)$

Increasing and Decreasing Functions

Let *f* be defined on an interval (a, b) and let $x_1, x_2 \in$ (a, b). Then

(i) *f* is increasing on the interval (a, b) if $f(x_2) > f(x_1)$ whenever $x_2 > x_1$.

(ii) *f* is decreasing on the interval (a, b) if $f(x_2) < f(x_1)$ whenever $x_2 > x_1$.



Fist Derivative Theorem:

Let *f* be a differentiable function on the open interval (a,b). Then

(i) *f* is increasing on (a,b) if $f'(x) > 0$ for each x ∈ (a, b).

(ii) *f* is decreasing on (a,b) if $f'(x) < 0$ for each x ∈ (a, b).

(iii) *f* is neither increasing nor decreasing on (a,b) if $f'(x) = 0$ for each x ∈ (a, b).

Stationary Point

Any point where *f* is neither increasing nor decreasing.

Critical value or Critical Point

If c ∈ Dom$f$ and $f'(\text{c}) = 0$ or $f'(\text{c})$ does not exists then c is called Critical value or Critical Point.

Relative Maxima / Maximum and Relative Minima / Minimum

Function f has relative maxima at c if $f''(\text{c}) < 0$ and function f has relative minima at c if $f''(\text{c}) > 0$.

Point of Inflection

Point where the function f is increasing before x = 0 and also after x = 0, such point is called the point of inflection.

42. Show that $f(\mathrm{x}) = \mathrm{x}^2$ is increasing or decreasing function on the interval $(-\infty, \infty)$.

Solution.

Given that $f(\mathrm{x}) = \mathrm{x}^2$ then

$f'(\mathrm{x}) = 2\mathrm{x}$

If $\mathrm{x} > 0$ then $f'(\mathrm{x}) = 2\mathrm{x} > 0$ so $f(\mathrm{x}) = \mathrm{x}^2$ is increasing in $(0, \infty)$

If $\mathrm{x} < 0$ then $f'(\mathrm{x}) = 2\mathrm{x} < 0$ so $f(\mathrm{x}) = \mathrm{x}^2$ is decreasing in $(-\infty, 0)$

If $\mathrm{x} = 0$ then $f'(\mathrm{x}) = 2\mathrm{x} = 0$ so $f(\mathrm{x}) = \mathrm{x}^2$ is neither increasing nor decreasing.

Hence x = 0 is stationary point. And function has minimum at x = 0.

43. Find the maximum value of $f(\mathrm{x}) = -2\mathrm{x}^2 + 4\mathrm{x} + 3$.

Solution.

$$x = -\frac{b}{2a} = -\frac{4}{2(-2)} = 1$$

$$f(x) = -2x^2 + 4x + 3$$

$$f(1) = -2(1)^2 + 4(1) + 3$$

$f(1) = 5$

The maximum value of the function is 5, the y-coordinate of the vertex.

44. Find the minimum value of $f(\mathrm{x}) = 2\mathrm{x}^2 - 3\mathrm{x} + 1$.

Solution.

$$a = 2,\ b = -3;\ x = -\frac{b}{2a} = -\frac{-3}{2(2)} = \frac{3}{4}$$

$$f(x) = 2x^2 - 3x + 1$$

$$f\left(\frac{3}{4}\right) = 2\left(\frac{3}{4}\right)^2 - 3\left(\frac{3}{4}\right) + 1$$

$$f\left(\frac{3}{4}\right) = -\frac{1}{8}$$

The vertex is $(\frac{3}{4}, -\frac{1}{8})$.

The minimum value of the function is $-\frac{1}{8}$, the y-coordinate of the vertex.

45. Find the extreme values for the function $f(\mathrm{x}) = 1 - \mathrm{x}^3$.

Solution.

Given that $f(\mathrm{x}) = 1 - \mathrm{x}^3$ then

$f'(\mathrm{x}) = -3\mathrm{x}^2$ and $f''(\mathrm{x}) = -6\mathrm{x}$

$f'(\mathrm{x}) = 0 \Rightarrow -3\mathrm{x}^2 = 0 \Rightarrow \mathrm{x} = 0$

$f''(0) = 0$ gives no information.

$f'(0-\in) = -3(0-\in)^2 = -3\in^2 < 0$

$f'(0+\in) = -3(0+\in)^2 = -3\in^2 < 0$

First derivative does not change sign at x = 0.

Hence $(0, f(0) = 1)$ is a point of inflection.

46. Find the extreme values for the function $f(\mathrm{x}) = \mathrm{x}^2 - \mathrm{x} - 2$.

Solution.

Given that $f(\mathrm{x}) = \mathrm{x}^2 - x - 2$ then

$f'(\mathrm{x}) = 2\mathrm{x} - 1$ and $f''(\mathrm{x}) = 2$

$f'(\mathrm{x}) = 2\mathrm{x} - 1 = 0 \Rightarrow 2\mathrm{x} = 1 \Rightarrow \mathrm{x} = \frac{1}{2}$

$f''\left(\frac{1}{2}\right) = 2 > 0$. So $f(\mathrm{x})$ is minimum at $\mathrm{x} = \frac{1}{2}$

$f\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right) - 2 = -\frac{9}{4}$ is required minimum value.

47. Find the extreme values for $f(\mathrm{x}) = 5 + 3\mathrm{x} - \mathrm{x}^3$.

Solution.

Given that $f(\mathrm{x}) = 5 + 3\mathrm{x} - \mathrm{x}^3$ then

$f'(\mathrm{x}) = 3 - 3\mathrm{x}^2$ and $f''(\mathrm{x}) = -6\mathrm{x}$

$f'(\mathrm{x}) = 3 - 3\mathrm{x}^2 = 0 \Rightarrow \mathrm{x} = \pm 1$

$f''(1) = -6 < 0$. So $f(\mathrm{x})$ is maximum at $\mathrm{x} = 1$

$f(1) = 5 + 3(1) - (1)^3 = 7$ is required maximum value.

$f''(-1) = 6 > 0$. So $f(\mathrm{x})$ is minimum at $\mathrm{x} = \frac{1}{2}$

$f(-1) = 5 + 3(-1) - (-1)^3 = 3$ is required minimum value.

48. The vertex of a parabola that opens up is $(-4, 7)$. What is the minimum value of the function?

Answer.

The minimum value of the function is 7, the y-coordinate of the vertex.

49. A mining company has determined that the cost c, in dollars per ton, of mining a mineral is given by $C(x) = 0.2x^2 - 2x + 12$, where x is the number of tons of the mineral that is mined. Find the number of tons of the mineral that should be mined to minimize the cost. What is the minimum cost?

Solution.

To find the number of tons of the mineral that should be mined to minimize the cost and to find the minimum cost, find the x- and y-coordinates of the vertex of the graph of $C(x) = 0.2x^2 - 2x + 12$.

$x = -\frac{b}{2a} = -\frac{-2}{2(0.2)} = 5$

To minimize the cost, 5 tons of the mineral should be mined.

$C(x) = 0.2x^2 - 2x + 12 \Rightarrow C(5) = 0.2(5)^2 - 2(5) + 12 = 7$

The minimum cost per ton is $7.

50. The height s, in feet, of a ball thrown straight up is given by $s(t) = -16t^2 + 64t + 4$, where t is the time in seconds after the ball is released. Find the time it takes the ball to reach its maximum height. What is the maximum height?

Solution.

$x = -\frac{b}{2a} = -\frac{64}{2(-16)} = 2$

The ball reaches its maximum height in 2 seconds.

$s(t) = -16t^2 + 64t + 4$

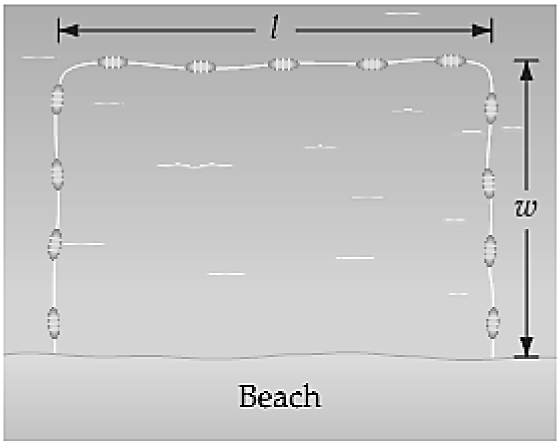
$s(2) = -16(2)^2 + 64(2) + 4 = 68$

The maximum height of the ball is 68 feet.

51. A lifeguard has 600 feet of rope with buoys attached to lay out a rectangular swimming area on a lake. If the beach forms one side of the rectangle, find the dimensions of the rectangle that will enclose the greatest swimming area.

Solution.

Let l represent the length of the rectangle, let w represent the width of the rectangle, and let A (which is unknown) represent the area of the rectangle. See the figure. Use these variables to write expressions for the perimeter and area of the rectangle.



Perimeter: $\mathrm{w} + l + \mathrm{w} = 600$

$2\mathrm{w} + l = 600$

$l = -2\mathrm{w} + 600$

Area: $A = l\mathrm{w}$

$A = (-2\mathrm{w} + 600)\mathrm{w}$

$A = -2\mathrm{w}^2 + 600\mathrm{w}$

Find the w-coordinate of the vertex.

$$\mathrm{w} = -\frac{b}{2a} = -\frac{600}{2(-2)} = 150$$

The width is 150 feet. To find l, replace w by 150 in $l = -2\mathrm{w} + 600$ and solve for l.

$l = -2\mathrm{w} + 600$

$l = -2(150) + 600 = 300$

The dimensions of the rectangle with maximum area are 150 feet by 300 feet.

52. A mason is forming a rectangular floor for a storage shed. The perimeter of the rectangle is 44 feet. What dimensions will give the floor a maximum area?

Solution.

Perimeter: $w + l + w + l = 44 \Rightarrow w + l = 22 \Rightarrow l = -w + 22$

Area: $A = lw \Rightarrow A = (-w + 22)w \Rightarrow A = -w^2 + 22w$

Find the w-coordinate of the vertex: $w = -\frac{b}{2a} = -\frac{22}{2(-1)} = 11$

The width is 11 feet. To find l, replace w by 11 in $l = -w + 22$ and solve for l.

$l = w + 22 \Rightarrow l = -11 + 22 = 11$

The length is 11 feet. The dimensions of the rectangle with maximum area are 11 feet by11 feet.

53. An open box is made from a square piece of cardboard that measures 50 inches on a side. To construct the box, squares inches on a side are cut from each corner of the cardboard. The remaining flaps are folded up to create a box.

a. Express the volume of the box as a polynomial function in x.

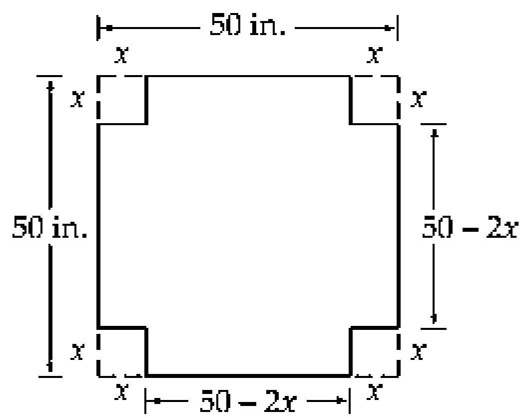
b. What is the volume of the box when squares 5 inches on a side are cut out?

c. Is it possible for the value of to be 30? Explain your answer.

Solution.

a. The volume of a box is a product of its length, width, and height.

From the diagram, the length is 50 – 2x, the width is 50 – 2x, and the height is x. Therefore, the volume is given by



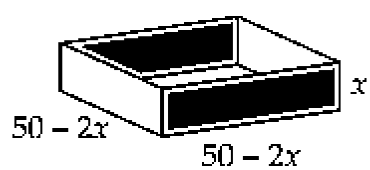
$$V = LWH$$

$$V(x) = (50 - 2x)(50 - 2x)x$$

$$= 4x^3 - 200x^2 + 2500x$$

The volume is given by $V(x) = 4x^3 - 200x^2 + 2500x$.

b. To find the volume when squares 5 inches on a side are cut out, evaluate the volume function when x = 5.



$$V(x) = 4x^3 - 200x^2 + 2500x$$

$$V(5) = 4(5)^3 - 200(5)^2 + 2500(5)$$

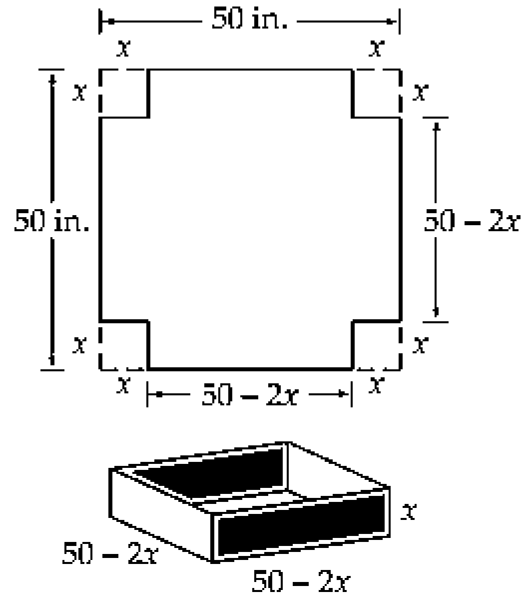
$$= 4(125) - 200(25) + 2500(5) = 8000$$

When squares 5 inches on a side are removed, the volume of the box is 8000 cubic inches.

c. If x = 30, then the value of 50 – 2x would be $50 - 2x = 50 - 2(30) = -10$. Because a length of -10 inches is not possible, the value of x cannot be 30.

54. An open box is made from a square piece of cardboard that measures 50 inches on a side. To construct the box, squares inches on a side are cut from each corner of the cardboard. The remaining flaps are folded up to create a box.

Express the surface area of the box as a function of x. The surface area is the sum of the areas of the four sides of the box and its bottom.



Solution.

The surface area is the sum of the areas of the four sides of the box and the area of its bottom.

$$S(x) = \underbrace{4}_{\text{4 sides}}\overbrace{(50-2x)x}^{\text{Area of each side}} + \overbrace{(50-2x)(50-2x)}^{\text{Area of the base}}$$

$$= 200x - 8x^2 + 2500 - 200x + 4x^2$$

$$= -4x^2 + 2500$$

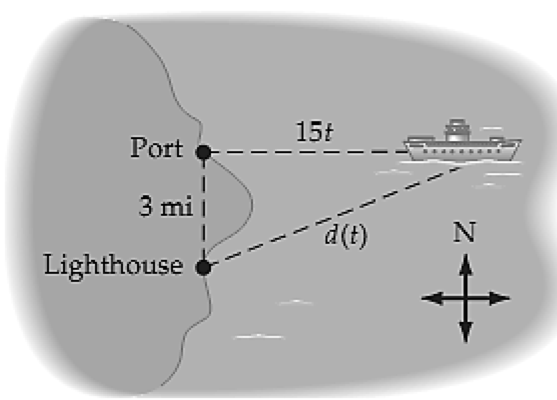
55. A lighthouse is 3 miles south of a port. A ship leaves the port and sails east at 15 mph.

a. Express the distance $d(t)$, in miles, between the ship and the lighthouse in terms of t, the number of hours the ship has been sailing.

b. Find the distance of the ship from the lighthouse after 3 hours. Round to the nearest tenth.

Solution.

Because the ship is sailing at 15 mph, after t hours the ship has traveled 15t miles, as shown in the diagram.



Using the Pythagorean Theorem $c^2 = a^2 + b^2$, where c is the length of the hypotenuse of a right triangle and a and b are the lengths of the legs, we have

$$[d(t)]^2 = (15t)^2 + 3^2$$
$$[d(t)]^2 = 225t^2 + 9$$
$$d(t) = \sqrt{225t^2 + 9}$$

To find the distance after 3 hours, replace by 3 and simplify.

$$d(t) = \sqrt{225t^2 + 9}$$
$$d(3) = \sqrt{225(3)^2 + 9} = \sqrt{2034}$$
$$\approx 45.1$$

After 3 hours, the ship is approximately 45.1 miles from the lighthouse.

56. A plane flies directly over a radar station at an altitude of 2 miles and a speed of 400 miles per hour.

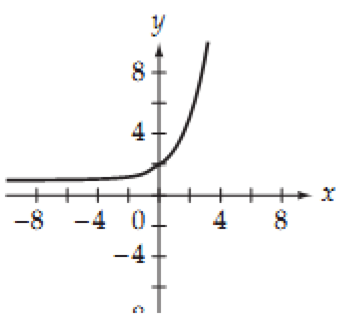
a. Express the distance $d(t)$, in miles, between the plane and the radar station in terms of t, the number of hours after the plane passes over the radar station.

b. Find the distance of the plane from the radar station after 3 hours. Round to the nearest tenth.

Solution.

a. $[d(t)]^2 = 2^2 + (400t)^2$

$$[d(t)]^2 = 4 + 160{,}000t^2$$
$$d(t) = \sqrt{4 + 160{,}000t^2}$$



b. $d(t) = \sqrt{4 + 160{,}000t^2}$

$$d(3) = \sqrt{4 + 160{,}000(3)^2}$$
$$d(3) \approx 1200$$

57. One of the considerations for a retail company is the cost of maintaining its inventory. The annual **inventory cost** is the cost of storing the items plus the cost of reordering the items. A lighting store has determined that the annual cost, in dollars, of storing *x* 25-watt halogen bulbs is 0.10x. The annual cost, in dollars, of reordering *x* 25-watt halogen bulbs is $\frac{0.15x+2}{x}$.

a. Express the inventory cost, $C(x)$, for the halogen bulbs in terms of x.

b. Find the inventory cost if the company wants to maintain an inventory of 150 25-watt halogen bulbs. Round to the nearest cent.

Solution.

a. $$C(x) = 0.10x + \frac{0.15x + 2}{x}$$

b. $$C(x) = 0.10x + \frac{0.15x + 2}{x}$$

$$C(150) = 0.10(150) + \frac{0.15(150) + 2}{150}$$

$$\approx 15.16$$

The inventory cost is $15.16.

58. A manufacturer has determined that the total cost, in dollars, of producing x straight-back wooden chairs is given by $C(x) = 35x + 500$. The average cost per chair, A(x), is the quotient of the total cost and x.

a. Find the function for the average cost per chair.

b. What is the average cost per chair when the manufacturer produces 40 chairs?

Solution.

a. Let $A(x)$ be the average cost per chair.

$$A(x) = \frac{\text{total cost}}{x} = \frac{C(x)}{x}$$

$$= \frac{35x + 500}{x}$$

b. $$A(x) = \frac{35x + 500}{x}$$

$$A(40) = \frac{35(40) + 500}{40} = 47.50$$

The average cost is $47.50 per chair.

Exponential Function

The exponential function is defined by $f(x) = b^x$ where b is called the base, b > 0, $b \neq 1$ and x is any real number.

59. Evaluate $f(x) = 3^x$ at $x = 2, x = -4$ and $x = \pi$. Round approximate results to the nearest hundred thousandth.

Solution.

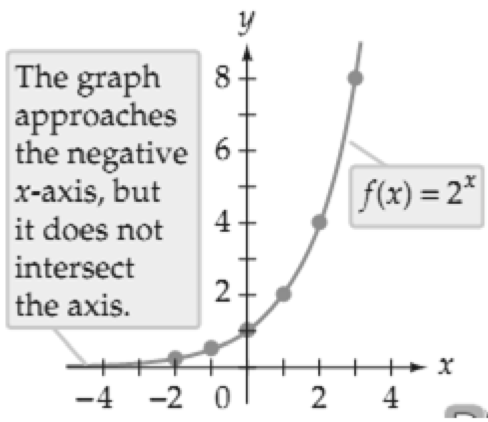
$$f(2) = 3^2 = 9$$

$$f(-4) = 3^{-4} = \frac{1}{3^4} = \frac{1}{81}$$

$$f(\pi) = 3^{\pi} \approx 3^{3.1415927} \approx 31.54428$$

Graphs of Exponential Functions/ Exponential Growth Model

The graph of $f(x) = 2^x$ is shown in Figure. The coordinates of some of the points on the graph are given in the table.



x	$f(x) = 2^x$	(x, y)
-2	$f(-2) = 2^{-2} = \frac{1}{4}$	$\left(-2, \frac{1}{4}\right)$
-1	$f(-1) = 2^{-1} = \frac{1}{2}$	$\left(-1, \frac{1}{2}\right)$
0	$f(0) = 2^0 = 1$	$(0, 1)$
1	$f(1) = 2^1 = 2$	$(1, 2)$
2	$f(2) = 2^2 = 4$	$(2, 4)$
3	$f(3) = 2^3 = 8$	$(3, 8)$

Observe that the values of y increase as x increases. This is an **exponential growth function**. This is typical of the graphs of all exponential functions for which the base is greater than 1. For the function $f(x) = 2^x, b = 2$ which is greater than 1.

60. Evaluate $g(x) = \left(\frac{1}{2}\right)^x$ at $x = 3, x = -1$ and $x = \sqrt{3}$. Round approximate results to the nearest hundred thousandth. Also graph it.

Solution.

$$g(x) = \left(\frac{1}{2}\right)^x$$

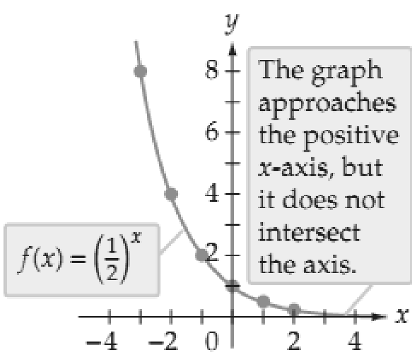
$$g(3) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

$$g(-1) = \left(\frac{1}{2}\right)^{-1} = \frac{1}{\frac{1}{2}} = 2$$

$$g(\sqrt{3}) = \left(\frac{1}{2}\right)^{\sqrt{3}} \approx \left(\frac{1}{2}\right)^{1.732} \approx 0.301$$

61. Graph the function $f(x) = \left(\frac{1}{2}\right)^x$.

Solution.



x	$f(x) = \left(\frac{1}{2}\right)^x$	(x, y)
-3	$f(-3) = \left(\frac{1}{2}\right)^{-3} = 8$	$(-3, 8)$
-2	$f(-2) = \left(\frac{1}{2}\right)^{-2} = 4$	$(-2, 4)$
-1	$f(-1) = \left(\frac{1}{2}\right)^{-1} = 2$	$(-1, 2)$
0	$f(0) = \left(\frac{1}{2}\right)^{0} = 1$	$(0, 1)$
1	$f(1) = \left(\frac{1}{2}\right)^{1} = \frac{1}{2}$	$\left(1, \frac{1}{2}\right)$
2	$f(2) = \left(\frac{1}{2}\right)^{2} = \frac{1}{4}$	$\left(2, \frac{1}{4}\right)$

Observe that the values of y decrease as x increases. This is an **exponential decay function**. This is typical of the graphs of all exponential functions for which the positive base is less than 1. For the function $g(x) = \left(\frac{1}{2}\right)^x$, $b = \frac{1}{2}$ which is less than 1.

62. Is $f(x) = 0.25^x$, an exponential growth function or an exponential decay function?

Answer.

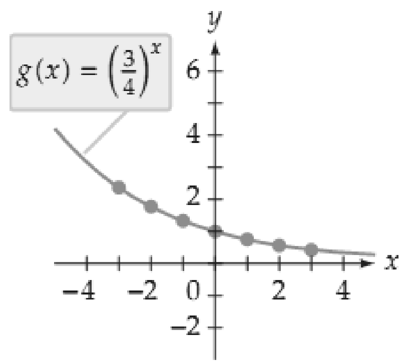
The base is 0.25, which is less than 1. The function is an exponential decay function.

63. State whether $g(x) = \left(\frac{3}{4}\right)^x$ is an exponential growth function or an exponential decay function. Then graph the function.

Solution.

Because the base $\frac{3}{4}$ is less than 1, g is an exponential decay function. Because it is an exponential decay function, the y-values will decrease as x increases. The y-intercept of the graph is the point $(0,1)$ and the graph also passes through $\left(\frac{1,3}{4}\right)$. Plot a few additional points. Then draw a smooth curve through the points, as shown in the figure.

x	$g(x) = \left(\frac{3}{4}\right)^x$	(x, y)
-3	$g(-3) = \left(\frac{3}{4}\right)^{-3} = \frac{64}{27}$	$\left(-3, \frac{64}{27}\right)$
-2	$g(-2) = \left(\frac{3}{4}\right)^{-2} = \frac{16}{9}$	$\left(-2, \frac{16}{9}\right)$
-1	$g(-1) = \left(\frac{3}{4}\right)^{-1} = \frac{4}{3}$	$\left(-1, \frac{4}{3}\right)$
0	$g(0) = \left(\frac{3}{4}\right)^{0} = 1$	$(0, 1)$
1	$g(1) = \left(\frac{3}{4}\right)^{1} = \frac{3}{4}$	$\left(1, \frac{3}{4}\right)$
2	$g(2) = \left(\frac{3}{4}\right)^{2} = \frac{9}{16}$	$\left(2, \frac{9}{16}\right)$
3	$g(3) = \left(\frac{3}{4}\right)^{3} = \frac{27}{64}$	$\left(3, \frac{27}{64}\right)$

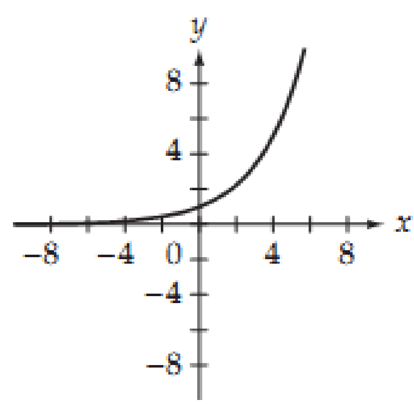


64. State whether $f(x) = \left(\frac{3}{2}\right)^x$ is an exponential growth function or an exponential decay function. Then graph the function.

Solution.

Because the base $\frac{3}{2}$ is greater than 1, f is an exponential growth function.

x	$f(x) = \left(\frac{3}{2}\right)^x$	(x, y)
-3	$f(-3) = \left(\frac{3}{2}\right)^{-3} = \frac{8}{27}$	$\left(-3, \frac{8}{27}\right)$
-2	$f(-2) = \left(\frac{3}{2}\right)^{-2} = \frac{4}{9}$	$\left(-2, \frac{4}{9}\right)$
-1	$f(-1) = \left(\frac{3}{2}\right)^{-1} = \frac{2}{3}$	$\left(-1, \frac{2}{3}\right)$
0	$f(0) = \left(\frac{3}{2}\right)^{0} = 1$	$(0, 1)$
1	$f(1) = \left(\frac{3}{2}\right)^{1} = \frac{3}{2}$	$\left(1, \frac{3}{2}\right)$
2	$f(2) = \left(\frac{3}{2}\right)^{2} = \frac{9}{4}$	$\left(2, \frac{9}{4}\right)$
3	$f(3) = \left(\frac{3}{2}\right)^{3} = \frac{27}{8}$	$\left(3, \frac{27}{8}\right)$



The Number 'e'

The number e is defined as the number that $\left(1 + \frac{1}{n}\right)^n$ approaches as n increases without bound.

The letter e was chosen in honor of the Swiss mathematician Leonhard Euler. The value of e accurate to eight decimal places is 2.71828183.

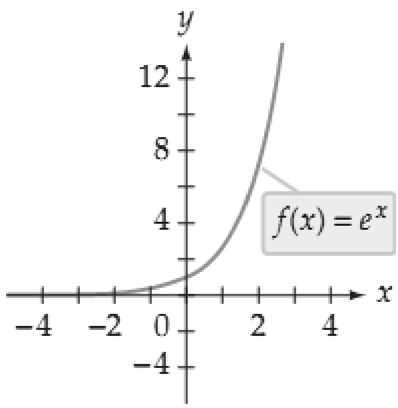
The Natural Exponential Function

For all real numbers x, the function defined by $f(x) = e^x$ is called the natural exponential function.

65. Graph $f(x) = e^x$.

Solution.

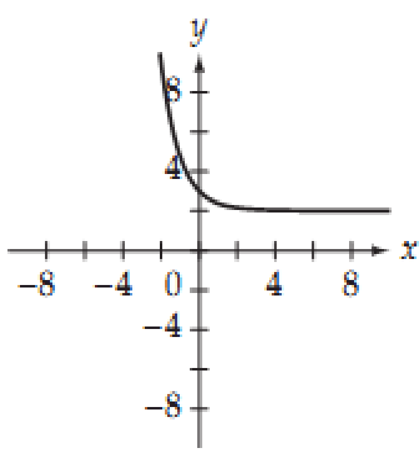
x	−2	−1	0	1	2
$f(x) = e^x$	0.1	0.4	1.0	2.7	7.4



66. Graph $f(x) = e^{-x} + 2$.

Solution.

x	−2	−1	0	1	2
$f(x) = e^{-x} + 2$	9.4	4.7	3	2.4	2.1



67. When an amount of money P is placed in an account that earns compound interest, the value A of the money after t years is given by the compound interest formula $A = P\left(1 + \frac{r}{n}\right)^{nt}$ where r is the annual interest rate as a decimal and n is the number of compounding periods per year. Suppose \$500 is placed in an account that earns 8% interest compounded daily. Find the value of the investment after 5 years.

Solution.

Use the compound interest formula. Because interest is compounded daily, n = 365

$$\begin{aligned} A &= P\left(1 + \frac{r}{n}\right)^{nt} \\ &= 500\left(1 + \frac{0.08}{365}\right)^{365(5)} \\ &\approx 500(1.491759) \approx 745.88 \end{aligned}$$

After 5 years, there is \$745.88 in the account.

68. The radioactive isotope iodine-131 is used to monitor thyroid activity. The number of grams N of iodine-131 in the body t hours after an injection is given by $N(t) = 1.5\left(\frac{1}{2}\right)^{\frac{t}{193.7}}$. Find the number of grams of the isotope in the body 24 hours after an injection. Round to the nearest ten-thousandth.

Solution.

$$\begin{aligned} N(t) &= 1.5\left(\frac{1}{2}\right)^{t/193.7} \\ N(24) &= 1.5\left(\frac{1}{2}\right)^{24/193.7} \\ &\approx 1.5(0.9177) \approx 1.3766 \end{aligned}$$

After 24 hours, there is approximately 1.3766 grams of the isotope in the body.

The next example is based on Newton's Law of Cooling. This exponential function can be used to model the temperature of something that is being cooled.

69. A cup of coffee is heated to 160°F and placed in a room that maintains a temperature of 70°F. The temperature T of the coffee after t minutes is given by $T(t) = 70 + 90e^{-0.0485t}$. Find the temperature of the coffee 20 minutes after it is placed in the room. Round to the nearest degree.

Solution.

Evaluate the function $T(t) = 70 + 90e^{-0.0485t}$ for $t = 20$.

$$T(t) = 70 + 90e^{-0.0485t}$$
$$T(20) = 70 + 90e^{-0.0485(20)}$$
$$\approx 70 + 34.1$$
$$\approx 104.1$$

After 20 minutes the temperature of the coffee is about 104°F.

70. The function $A(t) = 200e^{-0.014t}$ gives the amount of aspirin, in milligrams, in a patient's bloodstream t minutes after the aspirin has been administered. Find the amount of aspirin in the patient's bloodstream after 45 minutes. Round to the nearest milligram.

Solution.

$$A(t) = 200e^{-0.014t}$$
$$A(45) = 200e^{-0.014(45)}$$
$$\approx 107$$

After 45 minutes, there is approximately 107 milligrams of aspirin in the patient's bloodstream.

Logarithm

For $b > 0$, $b \neq 1, y = \log_b x$ is equivalent to $x = b^y$.

71. Which of the following is the logarithmic form of $4^3 = 64$?

a. $\log_4 3 = 64$ **b.** $\log_3 4 = 64$ **c.** $\log_4 64 = 3$

Solution.

$\log_4 64 = 3$ is equivalent to $4^3 = 64$.

Remark
The equation $y = \log_b x$ is the logarithmic form of $b^y = x$ and the equation $b^y = x$ is the exponential form of $y = \log_b x$. These two forms state exactly the same relationship between x and y.

72. Write a Logarithmic Equation in Exponential Form and an Exponential Equation in Logarithmic Form

a. Write $2 = \log_{10}(x + 5)$ in exponential form.
b. Write $2^{3x} = 64$ in logarithmic form.

Solution.

a. $2 = \log_{10}(x + 5)$ if and only if $10^2 = x + 5$
b. $2^{3x} = 64$ if and only if $\log_2 64 = 3x$

73. Write a Logarithmic Equation in Exponential Form and an Exponential Equation in Logarithmic Form

a. Write $\log_2(4x) = 10$ in exponential form.
b. Write $10^3 = 2x$ in logarithmic form.

Solution.

a. $\log_2(4x) = 10$ if and only if $2^{10} = 4x$
b. $10^3 = 2x$ if and only if $\log_{10} 2x = 3$

Equality of Exponents Property
If $b > 0$ and $b^x = b^y$, then $x = y$.

74. Evaluate the logarithms. $\log_8 64$ and $\log_2\left(\frac{1}{8}\right)$.

Solution.

a.
$$\log_8 64 = x$$
$$8^x = 64$$
$$8^x = 8^2$$
$$x = 2$$
$$\log_8 64 = 2$$

b.
$$\log_2\left(\frac{1}{8}\right) = x$$
$$2^x = \frac{1}{8}$$
$$2^x = 2^{-3}$$
$$x = -3$$
$$\log_2\left(\frac{1}{8}\right) = -3$$

75. Evaluate the logarithms. $\log_{10}(0.001)$ and $\log_5 125$.

Solution.

a. $\log_{10} 0.001 = x$

$$10^x = 0.001$$
$$10^x = 10^{-3}$$
$$x = -3$$
$$\log_{10} 0.001 = -3$$

b. $\log_5 125 = x$

$$5^x = 125$$
$$5^x = 5^3$$
$$x = 3$$
$$\log_5 125 = 3$$

76. Solve $\log_3 x = 2$.

Solution.

$$\log_3 x = 2$$
$$3^2 = x$$
$$9 = x$$

77. Solve $\log_2 x = 6$.

Solution.

$$\log_2 x = 6$$
$$2^6 = x$$
$$64 = x$$

Common and Natural Logarithms

The function defined by $f(x) = \log_{10} x$ is called the **common logarithmic function**. It is customarily written without the base as $f(x) = \log x$.

The function defined by $f(x) = \log_e x$ is called the natural logarithmic function. It is customarily written as $f(x) = \ln x$.

78. Solve each of the following equations. Round to the nearest thousandth.

$\log x = -1.5$; $\ln x = 3$

Solution.

$\log x = -1.5 \Rightarrow 10^{-1.5} = x \Rightarrow 0.032 \approx x$

$\ln x = 3 \Rightarrow e^3 = x \Rightarrow 20.086 \approx x$

79. Solve each of the following equations. Round to the nearest thousandth.

$\log x = -2.1$; $\ln x = 2$

Solution.

$\log x = -2.1 \Rightarrow 10^{-2.1} = x \Rightarrow 0.008 \approx x$

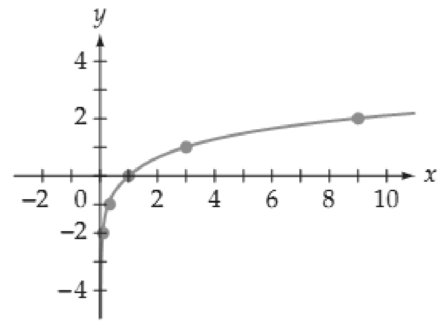
$\ln x = 2 \Rightarrow e^2 = x \Rightarrow 7.389 \approx x$

80. Graph $f(x) = \log_3 x$.

Solution.

$\log_3 x = y \Rightarrow x = 3^y$

$x = 3^y$	$\frac{1}{9}$	$\frac{1}{3}$	1	3	9
y	−2	−1	0	1	2

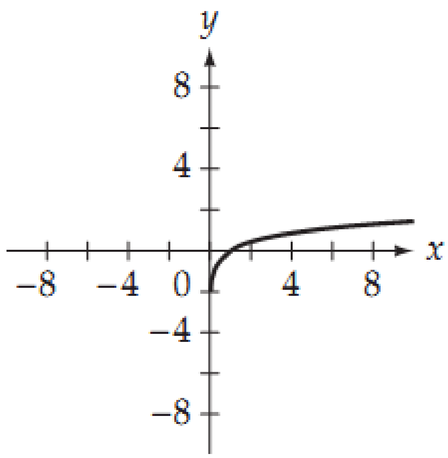


81. Graph $f(x) = \log_5 x$.

Solution.

$\log_5 x = y \Rightarrow x = 5^y$

$x = 5^y$	$\frac{1}{25}$	$\frac{1}{5}$	1	5	25
y	−2	−1	0	1	2



82. During the 1980s and 1990s, the average time T of a major league baseball game tended to increase each year. If the year 1981 is represented by x = 1, then the function $T(x) = 149.57 + 7.63 \ln x$ approximates the average time T, in minutes, of a major league baseball game for the years x = 1 to x = 19.

a. Use the function to determine the average time of a major league baseball game during the 1981 season and during the 1999 season. Round to the nearest hundredth of a minute.

b. By how much did the average time of a major league baseball game increase from 1981 to 1999?

Solution.

a. The year 1981 is represented by x = 1 and the year 1999 by x = 19.

$T(1) = 149.57 + 7.63 \ln(1) = 149.57$

In 1981 the average time of a major league baseball game was about 149.57 minutes.

$T(19) = 149.57 + 7.63 \ln(19) \approx 172.04$

In 1999 the average time of a major league baseball game was about 172.04 minutes.

b. $T(19) - T(1) \approx 172.04 - 149.57 = 22.47$

From 1981 to 1999, the average time of a major league baseball game increased by about 22.47 minutes.

83. The following function models the average typing speed S, in words per minute, of a student who has been typing for t months.

$S(t) = 5 + 29 \ln(t + 1) \quad ; 0 \leq t \leq 9$

a. Use the function to determine the student's average typing speed when the student first started to type and the student's average typing speed after 3 months. Round to the nearest whole word per minute.

b. By how much did the typing speed increase during the 3 months?

Solution.

a. $S(0) = 5 + 29 \ln(0 + 1) = 5$

The average typing speed when the student first started to type was 5 words per minute.

$S(3) = 5 + 29 \ln(3 + 1) \approx 45$

The average typing speed after 3 months was about 45 words per minute.

b. $S(3) - S(0) = 45 - 5 = 40$

The typing speed increased by 40 words per minute during the 3 months.

The Richter Scale Magnitude of an Earthquake

An earthquake with an intensity of I has a Richter scale magnitude of $\mathbf{M = \log\left(\frac{I}{I_0}\right)}$

Where I_0 is the measure of the intensity of a zero-level earthquake.

84. Find the Richter scale magnitude of the 2003 Amazonas, Brazil earthquake, which had an intensity of $I = 12589254I_0$. Round to the nearest tenth.

Solution.

$$M = \log\left(\frac{I}{I_0}\right) = \log\left(\frac{12{,}589{,}254I_0}{I_0}\right) = \log(12{,}589{,}254) \approx 7.1$$

The 2003 Amazonas, Brazil earthquake had a Richter scale magnitude of 7.1.

85. Previous example: What is the Richter scale magnitude of an earthquake whose intensity is twice that of the Amazonas, Brazil earthquake.

That is Find the Richter scale magnitude of the 2003 Amazonas, Brazil earthquake, which had an intensity of $I = 2(12589254I_0)$. Round to the nearest tenth.

Solution.

$$I = 2\cdot(12{,}589{,}254I_0) = 25{,}178{,}508I_0$$

$$M = \log\left(\frac{I}{I_0}\right) = \log\left(\frac{25{,}178{,}508I_0}{I_0}\right) = \log(25{,}178{,}508) \approx 7.4$$

The Richter scale magnitude of an earthquake whose intensity is twice that of the Amazonas, Brazil earthquake is 7.4.

86. Find the intensity of the 2003 Colina, Mexico earthquake, which measured 7.6 on the Richter scale. Round to the nearest thousand.

Solution.

$M = \log\left(\frac{I}{I_0}\right) = 7.6 \Rightarrow \frac{I}{I_0} = 10^{7.6} \Rightarrow I = 10^{7.6}I_0 \Rightarrow I \approx 39810717I_0$

The 2003 Colina, Mexico earthquake had an intensity that was approximately 39,811,000 times the intensity of a zero-level earthquake.

87. On April 29, 2003, an earthquake measuring 4.6 on the Richter scale struck Fort Payne, Alabama. Find the intensity of the quake. Round to the nearest thousand.

Solution.

$M = \log\left(\frac{I}{I_0}\right) = 4.6 \Rightarrow \frac{I}{I_0} = 10^{4.6} \Rightarrow I = 10^{4.6}I_0 \Rightarrow I \approx 39811I_0$

The April 29, 2003 earthquake had an intensity that was approximately 40,000 times the intensity of a zero-level earthquake.

The pH of a Solution

The pH of a solution with a hydronium-ion concentration of H^+ moles per liter is given by $\mathbf{pH = -log[H^+]}$

88. Find the pH of each liquid. Round to the nearest tenth.

a. Orange juice containing an H^+ concentration of 2.8×10^{-4} mole per liter
b. Milk containing an H^+ concentration of 3.97×10^{-7} mole per liter
c. A baking soda solution containing an H^+ concentration of mole 3.98×10^{-9} per liter

Solution.

a. $\text{pH} = -\log[H^+]$

$\text{pH} = -\log(2.8 \times 10^{-4}) \approx 3.6$

The orange juice has a pH of about 3.6.

b. $\text{pH} = -\log[H^+]$

$\text{pH} = -\log(3.97 \times 10^{-7}) \approx 6.4$

The milk has a pH of about 6.4.

c. $\text{pH} = -\log[H^+]$

$\text{pH} = -\log(3.98 \times 10^{-9}) \approx 8.4$

The baking soda solution has a pH of about 8.4.

89. Find the pH of each liquid. Round to the nearest tenth.

a. A cleaning solution containing an H^+ concentration of 2.41×10^{-13} mole per liter
b. A cola soft drink containing an H^+ concentration of 5.07×10^{-4} mole per liter
c. Rainwater containing an H^+ concentration of 6.31×10^{-5} mole per liter

Solution.

a. $\text{pH} = -\log[H^+] = -\log(2.41 \times 10^{-13}) \approx 12.6$

The cleaning solution has a pH of 12.6.

b. $\text{pH} = -\log[H^+] = -\log(5.07 \times 10^{-4}) \approx 3.3$

The cola soft drink has a pH of 3.3.

c. $\text{pH} = -\log[H^+] = -\log(6.31 \times 10^{-5}) \approx 4.2$

The rainwater has a pH of 4.2.

90. A sample of blood has a pH of 7.3. Find the hydronium-ion concentration of the blood.
Solution.

$$pH = -\log[H^+]$$

$$7.3 = -\log[H^+]$$

$$-7.3 = \log[H^+]$$

$$10^{-7.3} = H^-$$

$$5.0 \times 10^{-8} \approx H^-$$

The hydronium-ion concentration of the blood is about 5.01×10^{-8} mole per liter.

91. The water in the Great Salt Lake in Utah has a pH of 10.0. Find the hydronium-ion concentration of the water.
Solution.

$$pH = -\log[H^+]$$

$$10.0 = -\log[H^+]$$

$$-10.0 = \log[H^+]$$

$$10^{-10.0} = H^+$$

$$1.0 \times 10^{-10} = H^+$$

The hydronium-ion concentration of the water in the Great Salt Lake in Utah is 1.0×10^{-10} mole per liter.

Exercise

1) In Exercises i–vi, graph each equation.

i. $y = 2x - 1$
ii. $y = -3x + 2$
iii. $y = \frac{2}{3}x + 1$
iv. $y = -\frac{x}{2} - 3$
v. $y = \frac{1}{2}x^2$
vi. $y = |x - 1|$

2) In Exercises i–iv, evaluate the function for the given value.

i. $f(x) = 2x + 7$; $x = -2$
ii. $f(x) = 1 - 3x$; $x = -4$
iii. $f(t) = t^2 - t - 3$; $t = 3$
iv. $T(p) = \frac{p^2}{p-2}$; $p = 0$

3) In Exercises i–xiv, find the x- and y-intercepts of the graph of the equation.

1. $f(x) = 3x - 6$
2. $f(x) = 2x + 8$
3. $y = \frac{2}{3}x - 4$
4. $y = -\frac{3}{4}x + 6$
5. $y = -x - 4$
6. $y = -\frac{x}{2} + 1$
7. $3x + 4y = 12$
8. $5x - 2y = 10$
9. $2x - 3y = 9$
10. $4x + 3y = 8$
11. $\frac{x}{2} + \frac{y}{3} = 1$
12. $\frac{x}{3} - \frac{y}{2} = 1$
13. $x - \frac{y}{2} = 1$
14. $-\frac{x}{4} + \frac{y}{3} = 1$

4) In Exercises i–vi, find the slope of the line containing the two points.

i $(1, 3), (3, 1)$
ii $(2, 3), (5, 1)$
iii $(-1, 4), (2, 5)$
iv $(3, -2), (1, 4)$
v $(-1, 3), (-4, 5)$
vi $(-1, -2), (-3, 2)$

5) In Exercises 1–8, find the equation of the line that passes through the given point and has the given slope.

1. $(0, 5), m = 2$
2. $(2, 3), m = \frac{1}{2}$
3. $(-1, 7), m = -3$
4. $(0, 0), m = \frac{1}{2}$
5. $(3, 5), m = -\frac{2}{3}$
6. $(0, -3), m = -1$
7. $(-2, -3), m = 0$
8. $(4, -5), m = -2$

6) In Exercises 9–16, find the equation of the line that passes through the given points.

9. $(0, 2), (3, 5)$

10. $(0, -3), (-4, 5)$

11. $(0, 3), (2, 0)$

12. $(-2, -3), (-1, -2)$

13. $(2, 0), (0, -1)$

14. $(3, -4), (-2, -4)$

15. $(-2, 5), (2, -5)$

16. $(2, 1), (-2, -3)$

7) In Exercises i to vi, a. write (if necessary) the polynomial function in standard form, b. give the degree of the polynomial function, and c. evaluate the function for the given values of the variable.

1. $f(x) = 2x^2 + 4x - 10, f(2)$

2. $f(x) = 1 + 2x^2, f(-2)$

3. $g(x) = x^2 + 2x^3 - 3x - 1, g(-1)$

4. $s(t) = 1 - t^2 - t^4, s(3)$

5. $y(z) = 2z^3 - 3z^2 + 4z - z^5 + 6, y(-2)$

6. $p(x) = -2x^3, p(-3)$

8) In Exercises i to vi, find the minimum or maximum value of each quadratic function. State whether the value is a minimum or a maximum.

i $f(x) = x^2 - 2x + 3$

ii $f(x) = -2x^2 + 4x - 5$

iii $f(x) = x^2 + 3x - 1$

iv $f(x) = 2x^2 + 4x$

v $f(x) = 3x^2 + 3x - 2$

vi $f(x) = x^2 - 5x + 3$

9) Find the extreme values for the following functions defined as;

i. $f(x) = 5x^2 - 6x + 2$

ii. $f(x) = 3x^2$

iii. $f(x) = 3x^2 - 4x + 5$

iv. $f(x) = 2x^3 - 2x^2 - 36x + 3$

v. $f(x) = x^4 - 4x^2$

vi. $f(x) = (x-2)^2(x-1)$

10) Show that $y = \frac{\ln x}{x}$ has maximum value at x = e.

11) Show that $y = x^x$ has maximum value at $x = \frac{1}{e}$.

12) Given $f(x) = 3^x$, evaluate

a. $f(2)$ **b.** $f(0)$ **c.** $f(-2)$

13) Given $H(x) = 2^x$, evaluate

a. $H(-3)$ **b.** $H(0)$ **c.** $H(2)$

14) Given $G(r) = \left(\frac{1}{2}\right)^{2r}$, evaluate

a. $G(0)$ **b.** $G\left(\frac{3}{2}\right)$ **c.** $G(-2)$

15) In Exercises i –x, graph the equation.

i $f(x) = 2^x + 1$

ii $f(x) = 3^x - 2$

iii $g(x) = 3^{x/2}$

iv $h(x) = 2^{-x/2}$

v $f(x) = 2^{x+3}$

vi $g(x) = 4^{-x} + 1$

vii $H(x) = 2^{2x}$

viii $F(x) = 2^{-x}$

ix $f(x) = e^{-x}$

x $y(x) = e^{2x}$

16) In Exercises 1–8, write the exponential equation in logarithmic form.

1. $7^2 = 49$

2. $10^3 = 1000$

3. $5^4 = 625$

4. $2^{-3} = \frac{1}{8}$

5. $10^{-4} = 0.0001$

6. $3^5 = 243$

7. $10^y = x$

8. $e^y = x$

17) In Exercises 9–16, write the logarithmic equation in exponential form.

9. $\log_3 81 = 4$

10. $\log_2 16 = 4$

11. $\log_5 125 = 3$

12. $\log_4 64 = 3$

13. $\log_4 \frac{1}{16} = -2$

14. $\log_2 \frac{1}{16} = -4$

15. $\ln x = y$

16. $\log x = y$

18) In Exercises 17–24, evaluate the logarithm.

17. $\log_3 81$

18. $\log_7 49$

19. $\log 100$

20. $\log 0.001$

21. $\log_3 \frac{1}{9}$

22. $\log_7 \frac{1}{7}$

23. $\log_2 64$

24. $\log 0.01$

19) In Exercises 25–32, solve the equation for x.

25. $\log_3 x = 2$

26. $\log_5 x = 1$

27. $\log_7 x = -1$

28. $\log_8 x = -2$

29. $\log_3 x = -2$

30. $\log_5 x = 3$

31. $\log_4 x = 0$

32. $\log_4 x = -1$

20) In Exercises 45–50, graph the function.

i $g(x) = \log_2 x$

ii $g(x) = \log_4 x$

iii $f(x) = \log_3(2x - 1)$

iv $f(x) = -\log_2 x$

v $f(x) = \log_2(x - 1)$

vi $f(x) = \log_3(x - 2)$

CHAPTER

4

LINEAR

EQUATIONS & INEQUALITIES

Systems of linear equations play an important and motivating role in the subject of linear algebra. In fact, many problems in linear algebra reduce to finding the solution of a system of linear equations. Thus, the techniques introduced in this chapter will be applicable to abstract ideas introduced later. On the other hand, some of the abstract results will give us new insights into the structure and properties of systems of linear equations. All our systems of linear equations involve scalars as both coefficients and constants, and such scalars may come from any number field **F**. There is almost no loss in generality if the reader assumes that all our scalars are real numbers — that is, that they come from the real field **R**.

In this chapter we will learn about;

- Analytical approach to solve simultaneous equations
- Inequalities and their application
- Comparing quantities using analytical tools

Linear Equation: $(\mathbf{ax + b = 0}\ ;\ \mathbf{a \neq 0})$

It is an algebraic equation in which each term has an exponent of one and graphing of equation results in a straight line.

Or A linear equation in unknowns $x_1, x_2, \ldots, x_n$ is an equation that can be put in the standard form $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ where $a_1, a_2, \ldots, a_n$ and b are constants. The constant a_k is called the coefficient of x_k, and b is called the constant term of the equation. e.g. $6x_1 + 7x_2 = 5$, $2x + 3y + 4z = -1$

Solutions of Linear Equation:

A **solution** of the linear equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ is a list of values for the unknowns or, equivalently, a vector **u** in $\mathbf{R}^n$, say $x_1 = k_1, x_2 = k_2, \ldots, x_n = k_n$ or $\mathbf{u} = (k_1, k_2, \ldots, k_n)$ such that the following statement (obtained by substituting k_i for x_i in the equation) is true: $a_1k_1 + a_2k_2 + \cdots + a_nk_n = b$

In such a case we say that **u** satisfies the equation.

Examples for Linear and Non – Linear Equations

- $x + 3y = 7$ linear
- $5x + 7y - 8yz = 16$ not linear
- $x + \pi y + ez = \log 5$ linear for constants π, e
- $\frac{1}{2}x - y + 3z = -1$ linear
- $x_1 - 2x_2 - 3x_3 + x_4 = 0$ linear
- $x_1 + x_2 + x_3 + \cdots \ldots + x_n = 1$ linear
- $x + 3y^2 = 4$ not linear
- $3x + 2y - xy = 5$ not linear
- $\mathrm{Sin}x + y = 0$ not linear
- $\sqrt{x_1} + 2x_2 + x_3 = 1$ not linear
- $x_1 + 5x_2 - \sqrt{2}x_3 = 1$ linear
- $x_1 + 3x_2 + x_1x_3 = 2$ not linear
- $x_1 = -7x_2 + 3x_3$ linear
- ${x_1}^{-2} + x_2 + 8x_3 = 5$ not linear
- ${x_1}^{3/5} - 2x_2 + x_3 = 4$ not linear
- $\pi x_1 - \sqrt{2}x_2 = 7^{\frac{1}{3}}$ linear
- $2^{\frac{1}{3}}x + \sqrt{3}y = 1$ linear
- $\mathrm{Cos}\left(\frac{\pi}{7}\right)x - 4y = \log 3$ not linear
- $xy = 1$ not linear
- $y + 7 = x$ linear

TRY OTHERS ALSO!!!!!!!!

لینئر کی پہچان: ویری ایبل کی پاور ایک سے کم زیادہ نہ ہو۔ ویری ایبل کے ساتھ اس کا ڈیری ویٹو نہ لکھا گیا ہو۔ دو ویری ایبل اکٹھے نہ ہوں۔

Variable not appears in this form $\tan x, \log x, \sin x, \cos x, \sqrt{x}, e^x$ etc.

System of Linear Equations **(System in which more than one linear equations involve)**

A system of linear equations is a list of linear equations with the same unknowns. In particular, a system of 'm' linear equations $L_1, L_2, \dots, L_m$ in 'n' unknowns $x_1, x_2, \dots, x_n$ can be put in the standard form

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

::

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

$$a_{ij}x_j = b_i$$

$m =$ No. of equations

$n =$ No. of unknowns

where the a_{ij} and b_i are constants. The number a_{ij} is the coefficient of the unknown x_j in the equation L_i, and the number b_i is the constant term of the equation L_i.

- The system of linear equations is called an $m \times n$ system. It is called a **square system** if $m = n$ that is, if the number m of equations is equal to the number n of unknowns.
- The system is said to be **homogeneous** if all the constant terms are zero that is, if $b_1 = 0, b_2 = 0, \dots, b_n = 0$. Otherwise the system is said to be **nonhomogeneous (inhomogeneous**.).
- A **solution** (or a particular solution) of the system (above) is a list of values for the unknowns or, equivalently, a vector **u** in $\mathbf{R}^n$, which is a solution of each of the equations in the system. The set of all solutions of the system is called **the solution set or the general solution of the system**.
- A finite set of linear equations is called a **system of linear equations**, or more briefly a **linear system**. The variables are called **unknown**.
- A linear equation does not involve any products or roots of variables. All variables occur only to the first power, and do not appear as arguments of trigonometric, logarithmic or exponential functions.

1. Solve the equation $3x + 7 = 1$.

Solution

$3x + 7 = 1$

$\Rightarrow 3x + 7 - 7 = 1 - 7$

$\Rightarrow 3x = -6$

$\Rightarrow \frac{3x}{3} = -\frac{6}{3} \Rightarrow x = -2$

2. Solve the equation $2x - 3 - 2 = 3$.

Solution

$2x - 3 - 2 = 3$

$\Rightarrow 2x - 5 = 3$

$\Rightarrow 2x - 5 + 5 = 3 + 5$

$\Rightarrow 2x = 8$

$\Rightarrow \frac{2x}{2} = \frac{8}{2}$

$\Rightarrow x = 4$

3. Solve the equation $-7x + 1 + 2x = 9x - 8 + 1$.

Solution

$-7x + 1 + 2x = 9x - 8 + 1$

$\Rightarrow -7x + 2x - 9x = -8 + 1 - 1$

$\Rightarrow -14x = -8 \Rightarrow x = \frac{-8}{-14}$

$\Rightarrow x = \frac{4}{7}$

4. Solve the equation; $2(x - 3) - 17 = 13 - 3(x + 2)$.

Solution

$2(x - 3) - 17 = 13 - 3(x + 2)$

$\Rightarrow 2x - 6 - 17 = 13 - 3x - 6$

$\Rightarrow 2x + 3x = 13 - 6 + 6 + 17$

$\Rightarrow 5x = 30 \Rightarrow x = 6$

5. Solve the equation involving fraction; $\frac{x+2}{4} - \frac{x-1}{3} = 2$.

Solution

$\frac{x+2}{4} - \frac{x-1}{3} = 2$

$\Rightarrow 12 \times \frac{x+2}{4} - 12 \times \frac{x-1}{3} = 12 \times 2$

$\Rightarrow 3(x + 2) - 4(x - 1) = 24$

$\Rightarrow 3x + 6 - 4x + 4 = 24$

$\Rightarrow x = -14$

6. Solve the equation involving fraction; $\frac{3x}{5} - \frac{x-3}{2} = \frac{x+2}{3}$.

Solution

$\frac{3x}{5} - \frac{x-3}{2} = \frac{x+2}{3}$

$\Rightarrow 30 \times \frac{3x}{5} - 30 \times \frac{x-3}{2} = 30 \times \frac{x+2}{3}$

$\Rightarrow 6(3x) - 15(x-3) = 10(x+2)$

$\Rightarrow 18x - 15x + 45 = 10x + 20$

$\Rightarrow x = \frac{25}{7}$

7. Consider the following linear equation in three unknowns x, y, z:

$x + 2y - 3z = 6$

Check $\vec{u} = (5,2,1)$ is solution of equation or not.

Solution

We note that $x = 5;\ y = 2;\ z = 1,$

or, equivalently, the vector $\vec{u} = (5,2,1)$ is a solution of the equation.

That is, $5 + 2(2) - 3(1) = 6$

8. Consider the following linear equation in three unknowns x, y, z:

$x + 2y - 3z = 6$

Check $\vec{v} = (1,2,3)$ is solution of equation or not.

Solution

$\vec{v} = (1,2,3)$ is not a solution, because on substitution,

we do not get a true statement: $1 + 2(2) - 3(3) = -4 \neq 6$

9. Solve the system of equations by addition method;

$2x + 3y = 48 \quad and \quad 9x - 8y = -24$

Solution

$2x + 3y = 48$ …………(i)

$9x - 8y = -24$ …………(ii)

Multiplying (i) with -3

$-9x - 6y = -144$

$9x - 8y = -24$

Then adding we have y = 12 and putting in (i) we have x = 8.

Hence the solution set is {8,12}.

Radical Equations

An equation in which the unknown letter (variable) appears under a radical sign is called a radical equation.

For example: $\sqrt{x+1}=7$, $\sqrt{x}=9$, $\sqrt{2x-3}=\sqrt{x+5}$

10. Solve $\sqrt{x}+3=7$

Solution

$\sqrt{x}+3=7$

$\sqrt{x}=4$

$\left(\sqrt{x}\right)^2=4^2$

$x=16$

11. Solve $4+2\sqrt{3y+1}=3$

Solution

$4+2\sqrt{3y+1}=3$

$2\sqrt{3y+1}=-1$

$\sqrt{3y+1}=-\frac{1}{2}$

$\left(\sqrt{3y+1}\right)^2=\left(-\frac{1}{2}\right)^2$

$3y+1=\frac{1}{4}$

$3y=-\frac{3}{4}$

$y=-\frac{1}{4}$ it is an extraneous root and the solution set is Φ

12. Solve $\sqrt{x-5}=3$

Solution

$\sqrt{x-5}=3$

$\left(\sqrt{x-5}\right)^2=(3)^2$

$x-5=9$

$x=14$

Absolute Value

The absolute value of a real number x is defined as follows;

$$|x| = \begin{cases} x & ; x > 0 \\ 0 & ; x = 0 \\ -x & ; x < 0 \end{cases}$$

Absolute Value Equations

An equation that contains a variable inside the absolute value bars is called an absolute value equation.

For example: $|x| + 1 = 5$, $|x - 3| = 4$

13. Solve $|x| = 8$

Solution

$|x| = 8$

$x = \pm 8$

14. Solve $|x| = -6$

Solution

$|x| = -6$ not possible for real numbers.

So this equation has no solution.

15. Solve $|2x + 5| = 11$

Solution

$|2x + 5| = 11$

$2x + 5 = \pm 11$

$x = 3, -8$

16. Solve $|a - 1| = |2a - 3|$

Solution

$|a - 1| = |2a - 3|$

$a - 1 = \pm(2a - 3)$

$a = 2, \frac{4}{3}$

17. Consider the following system of linear equations:

$$x_1 + x_2 + 4x_3 + 3x_4 = 5$$

$$2x_1 + 3x_2 + x_3 - 2x_4 = 1$$

$$x_1 + 2x_2 - 5x_3 + 4x_4 = 3$$

It is a 3×4 system because it has three equations in four unknowns. Determine whether (a) $u = (-8, 6, 1, 1)$ and (b) $v = (-10, 5, 1, 2)$ are solutions of the system.

Solution:

(a) Substitute the values of **u** in each equation, obtaining

$$-8 + 6 + 4(1) + 3(1) = 5 \Rightarrow 5 = 5$$

$$2(-8) + 3(6) + 1 - 2(1) = 1 \Rightarrow 1 = 1$$

$$-8 + 2(6) - 5(1) + 4(1) = 3 \Rightarrow 3 = 3$$

Yes, **u** is a solution of the system because it is a solution of each equation.

(b) Substitute the values of **v** into each successive equation, obtaining

$$-10 + 5 + 4(1) + 3(2) = 5 \Rightarrow 5 = 5$$

$$2(-10) + 3(5) + 1 - 2(2) = 1 \Rightarrow -8 \neq 1$$

No, **v** is not a solution of the system, because it is not a solution of the second equation. (We do not need to substitute **v** into the third equation.)

Consistent and Inconsistent Solutions:

The system of linear equations is said to be **consistent** if it has one or more solutions, and it is said to be **inconsistent** if it has no solution.

Underdetermined: A system of linear equations is considered underdetermined if there are fewer equations than unknowns. $m < n$

Over determined: A system of linear equations is considered over determined if there are more equations than unknowns. $n < m$

18. Solve the system of equations by substitution; $2x + y = 5 \quad and \quad x - 2y = 15$

Solution

$2x + y = 5 \quad \ldots\ldots\ldots\ldots(i)$

$x - 2y = 15 \quad \ldots\ldots\ldots\ldots(ii)$

$(ii) \Rightarrow x = 2y + 15$

$(i) \Rightarrow 2(2y + 15) + y = 5 \Rightarrow 4y + 30 + y = 5 \Rightarrow 5y = -25 \Rightarrow \mathbf{y = -5}$

$(ii) \Rightarrow x = 2(-5) + 15 \Rightarrow x = -10 + 15 \Rightarrow \mathbf{x = 5}$

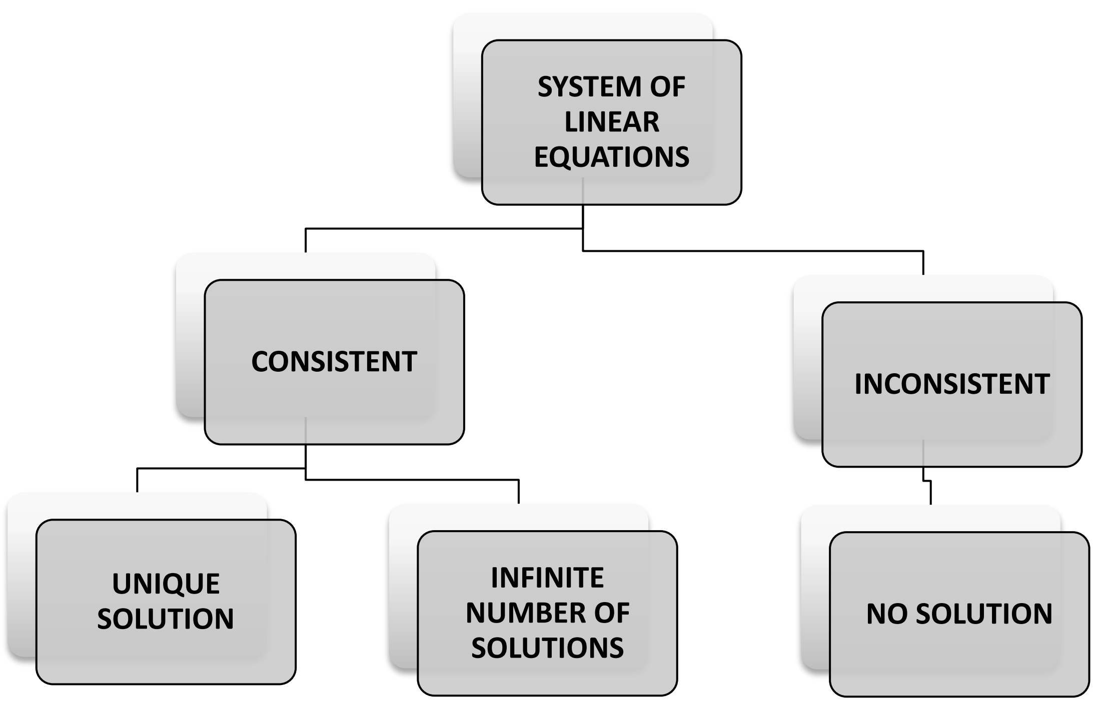
So our final answer is $x = 5, y = -5$

Remember

If the field **F** of scalars is infinite, such as when **F** is the real field **R** or the complex field **C**, then we have the following important result.

Result: Suppose the field **F** is infinite. Then any system of linear equations has

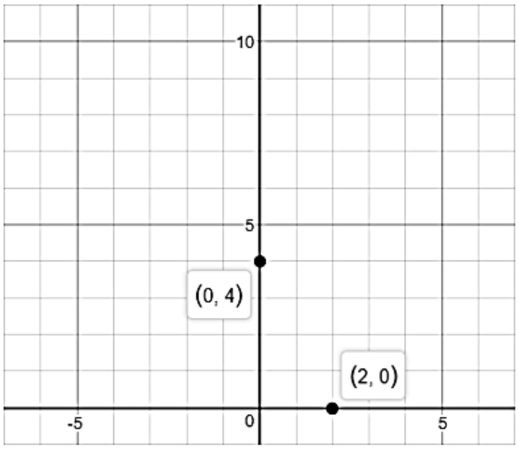
(i) a unique solution, (ii) no solution, or (iii) an infinite number of solutions.

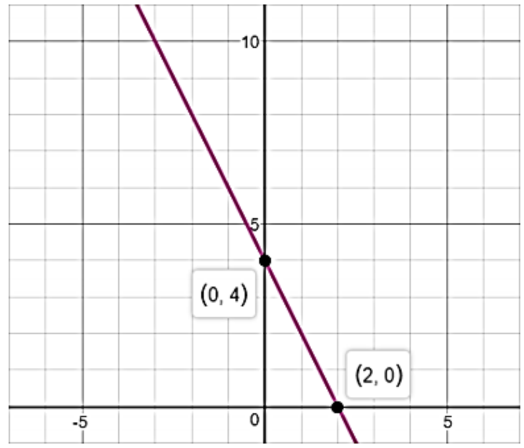


Graphing Linear Equations

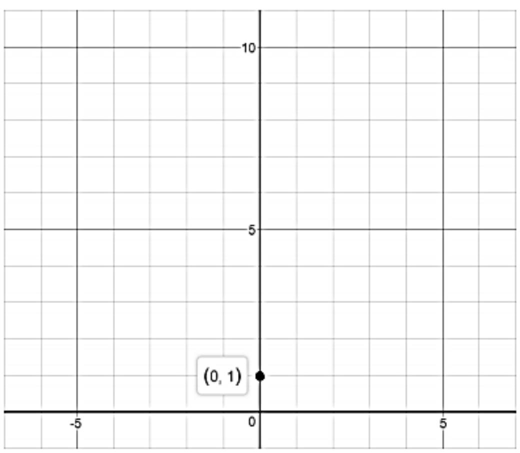
There are two methods to graph a line:

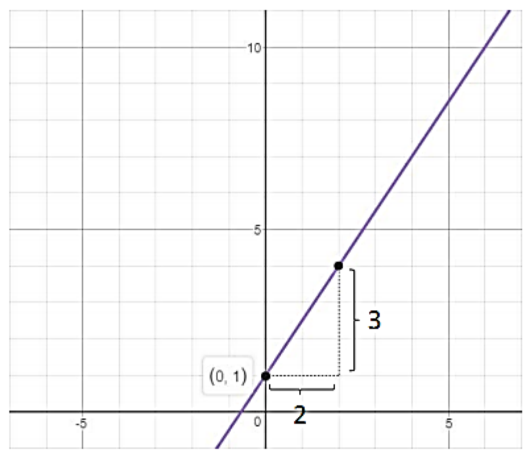
1) Use the x and y intercepts: If you have the coordinates of the x and y intercepts, plot those two points and connect them with a line.





2) Use slope-intercept form: If you have an equation in slope intercept form ($y = nx + b$), plot the y-intercept and then construct the line using the slope. If the equation is not in slope-intercept form, it can be rearranged to be in slope-intercept form.





$$m = \frac{3}{2}$$

19. Graph the equation $6x + 2y = 12$.

Solution

Given that $6x + 2y = 12$

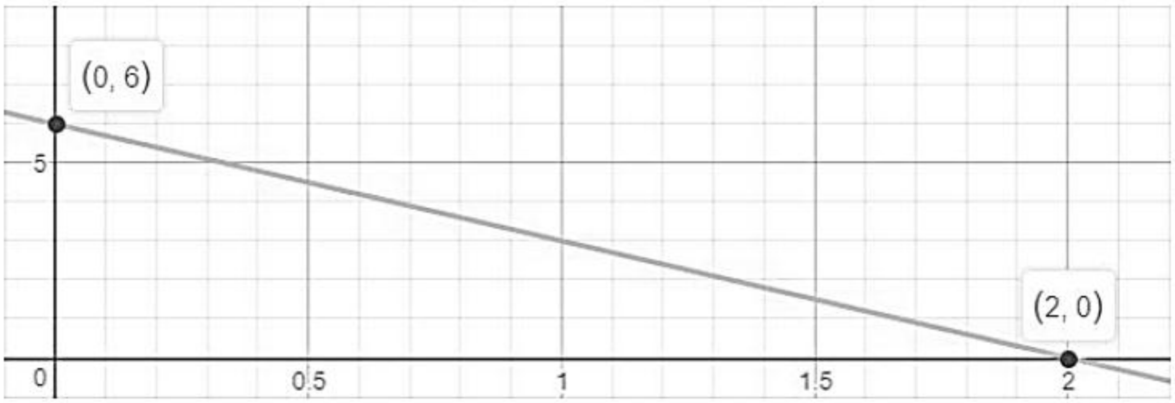
To find the x-intercept, set y equal to zero and solve for x.

$6x + 2y = 12 \Rightarrow 6x + 2(0) = 12 \Rightarrow 6x = 12 \Rightarrow x = 2$

To find the y-intercept, set x equal to zero and solve for y.

$6x + 2y = 12 \Rightarrow 6(0) + 2y = 12 \Rightarrow 2y = 12 \Rightarrow y = 6$

Now, plot the points (2,0) and (0,6) on the graph and connect them with a line:



20. Graph the equation $6x + 2y = 12$.

Solution

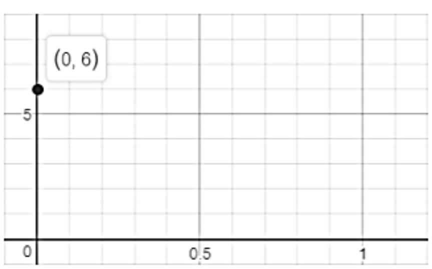
Given that $6x + 2y = 12$

$\Rightarrow y = -3x + 6$

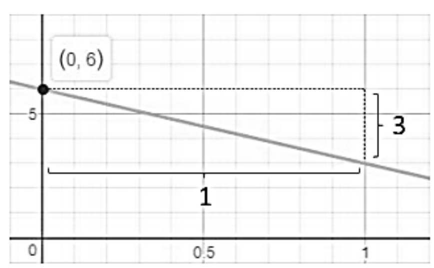
Comparing with $y = mx + c$ we have $m = -3$ and $b = 6$.

In other words, the slope is -3 and y – intercept is at (0,6)

Plot the y-intercept:



Use the slope to draw the rest of the line:



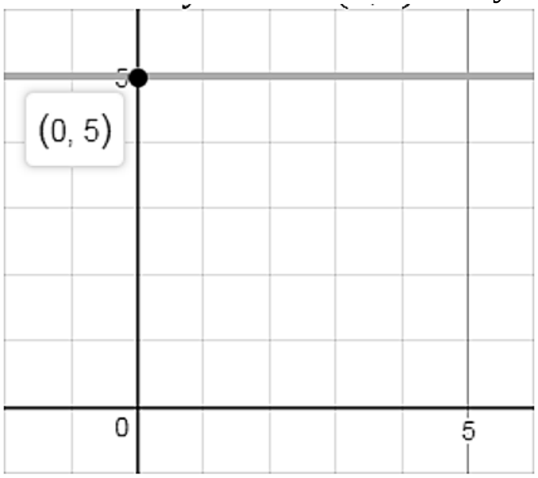
Graphs with one variable

If a graph only involves one variable, it can be graphed as a horizontal or vertical line. Below are two examples.

21. Graph the equation $y = 5$.

Solution

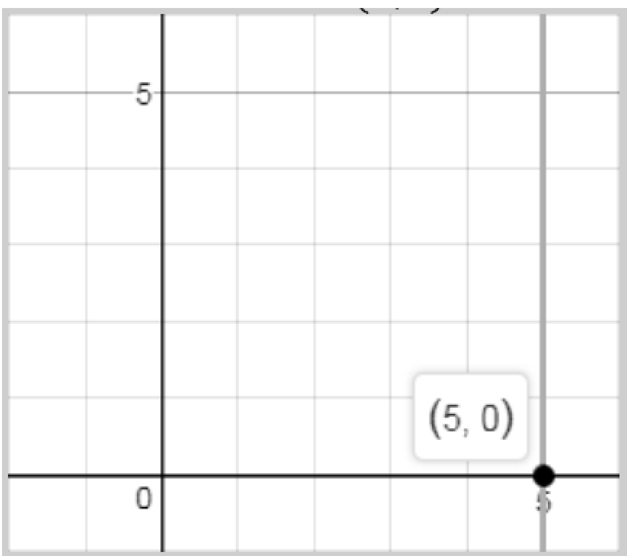
A horizontal line that crosses the y – axis at (0,5)



22. Graph the equation x = 5.

Solution

A vertical line that crosses the x – axis at (5,0)



Remark: (Geometrical Presentation / Graphics)

Linear system in two unknowns arise in connection with intersection of lines.

- The lines may be parallel and distinct, in which case there is no intersection and consequently no solution.
- The lines may be intersect at only one point, in which case the system has exactly one solution.
- The lines may coincide, in which case there are infinitely many points of intersection (the points on the common line) and consequently infinitely many solutions. (in such system, all equations will be same with few common factors)

23. (A Linear System with one Solution): Solve the following system of linear equations

$x - y = 1$ ………..(i)

$2x + y = 6$ ………..(ii)

Solution

(i) $\Rightarrow -2x + 2y = -2$ multiplying with -2

Adding (i) with (ii) $\Rightarrow -2x + 2y + 2x + y = -2 + 6 \Rightarrow y = \frac{4}{3}$

(i) $\Rightarrow x - \frac{4}{3} = 1 \Rightarrow x = \frac{7}{3}$

$S.S = \left(x = \frac{7}{3}, y = \frac{4}{3}\right)$

Geometrically this means that the lines represented by the equations in the system intersect at the single point $\left(x = \frac{7}{3}, y = \frac{4}{3}\right)$

24. (A Linear System with No Solution): Solve the following system of linear equations

$x + y = 4$ ………..(i)

$3x + 3y = 6$ ………..(ii)

Solution

(i) $\Rightarrow -3x - 3y = -12$ multiplying with -3

Adding (i) with (ii) $\Rightarrow -3x - 3y + 3x + 3y = -12 + 6 \Rightarrow 0 = -6$

The result is contradictory, so the given system has no solution. Geometrically this means that the lines may be parallel and distinct, in this case there is no intersection and consequently no solution.

25. (A Linear System with Infinitely many Solutions): Solve the following system of linear equations

$4x - 2y = 1$ ……….(i)

$16x - 8y = 4$ ……….(ii)

Solution:

(i) $\Rightarrow -16x + 8y = -4$ multiplying with -4

Adding (i) with (ii) $\Rightarrow -16x + 8y + 16x - 8y = -4 + 4 \Rightarrow 0 = 0$

Equation $0 = 0$ does not impose any restriction on 'x' and 'y' and hence can be omitted. Thus the solution of the system are those values of 'x' and 'y' that satisfy the single equation $4x - 2y = 1$

Geometrically this means that the lines corresponding to the two equations in the original system coincide. And this system will have infinitely many solutions.

How to Find Few Solutions of Such System?

Find the value of 'x' from Common equation.

Put $y = t$ 't' being **Parameter** (arbitrary value instead of actual value)

Replace $y = t$ in given system.

Use $t = 0,1,2,3,\ldots$ Upon your taste and get different answers.

We may apply same procedure by replacing 'x' and 'y'

26. Find different solutions for problem as follows using **Parametric Equation** (arbitrary equation using Parameter instead of actual value).

$4x - 2y = 1$ ……….(i)

$16x - 8y = 4$ ……….(ii)

Solution

$4x - 2y = 1$

$\Rightarrow x = \frac{1}{4} + \frac{1}{2}y$ and put $y = t$

$\Rightarrow x = \frac{1}{4} + \frac{1}{2}t$

$S.S = \left(x = \frac{1}{4}, y = 0\right)$ $t = 0$,

$S.S = \left(x = \frac{3}{4}, y = 1\right)$ $t = 1$

$S.S = \left(x = -\frac{1}{4}, y = -1\right)$ $t = -1$

> How to find solution of more than two equations?
>
> 1st method: find x,y,z solving equations in pair (Lengthy Process)
>
> 2nd: solve two equations, find x,y and put in 3rd equation to get value of z.
>
> 3rd method: observe given equations and take common if possible and then check all equations are same or not, if same then solution will be infinite.

27. Find different solutions for problem as follows using **Parametric Equation** (arbitrary equation using Parameter instead of actual value).

$x - y + 2z = 5$ ………..(i)

$2x - y + 4z = 10$ ………..(ii)

$3x - 3y + 6z = 15$ ………..(ii)

Solution

Since above all equations have same graphics or formation. Therefore will have infinitely many solutions. We will solve it using parametric equations.

In above all equations we have the parallel form $x - y + 2z = 5$

$\Rightarrow x = 5 + y - 2z$ and put $y = r, z = s$

$\Rightarrow x = 5 + r - 2s$

$S.S = (x = 5, y = 0, z = 0)$ $r = 0, s = 0$

$S.S = (x = 6, y = 1, z = 0)$ $r = 1, s = 0$

$S.S = (x = 4, y = 1, z = 1)$ $r = 1, s = 1$

General Solution $= \{(5,0,0), (6,1,0), (4,1,1)\}$

Try Others Also!!!!!!!!

Modeling with System of Linear Equations and their Solutions

28. A car rental company charges Rs.30 a day and 15 Pesa a mile for renting a car. Ali rents a car for two days and his bill comes to Rs. 108. How many miles did he drives?

Solution

Let x be number of miles driven. Mileage cost is 0.15x and daily cost is Rs. $2(30)$.

Now according to mathematical model

$\text{Mileage cost} + \text{daily cost} = \text{total cost}$

$0.15x + 2(30) = 108$

$x = 320$ miles

29. Usman inherits Rs. 100,000 and inverts it in two certificates of deposit. One certificate pays 6% and the other pays $4\frac{1}{2}\%$ simple interest annually. If usman's total interest is Rs. 5025 per year, how much money is invested at each rate?

Solution

Let x be amount invested at 6%, amount invested at $4\frac{1}{2}\%$ is 100,000 – x, interest earned at 6% is 0.06x and interest earned at $4\frac{1}{2}\%$ is $0.045(100{,}000 - x)$.

Then according to mathematical model

$\text{interest at } 6\% + \text{interest at } 4\frac{1}{2}\% = \text{total interest}$

$0.06x + 0.045(100{,}000 - x) = 5025$

$x = 35000$ rupees

So, usman has invested Rs. 35000 at 6% and the remaining Rs. 65,000 at $4\frac{1}{2}\%$.

30. A square garden has a walkway 3 ft wide around its outer edge. If the area of entire garden including the walkway is 18000 ft^2 what are the dimensions of the planted area?

Solution

Let x be the length of planted area, length of entire garden $(x + 6)$ and area of entire garden $(x + 6)^2$.Then according to mathematical model

$\text{area of entire garden} = 18000\ ft^2$

$(x + 6)^2 = 18000$

$x + 6 = \sqrt{18000}$

$x \approx 128\ ft$

The planted area of garden is about 128 ft by 128 ft.

31. A rectangular building lot is 8ft longer than its width and has an area of $2900\ ft^2$. Find the dimension of the lot.

Solution

Let w be the width of lot, length of lot is $(w + 8)$.

Then according to mathematical model

$\text{width of lot} \times \text{length of lot} = 2900\ ft^2$

$w(w + 8) = 2900\ ft^2$

$w^2 + 8w = 2900$

$w^2 + 8w - 2900 = 0$

$w = 50 \quad \text{or} \quad w = -58$

Since the width of the lot must be a positive numbers, we conclude that $w = 50$ft. And the length of the lot is $w + 8 = 58$ft.

32. A man who is 6ft tall wishes to find the height of a certain four story building. He measures its shadow and find it to be 28ft long, while his own shadow is $3\frac{1}{2}$ft long. How tall is building?

Solution

Let h be the height of building. Then according to mathematical model

$$\frac{\text{height in large triangle}}{\text{base in large triangle}} = \frac{\text{height in small triangle}}{\text{base in small triangle}}$$

$$\frac{h}{28} = \frac{6}{3.5} \Rightarrow h = 48\ ft$$

So the building is 48ft tall.

33. A manufacturer of soft drinks advertises their orange soda as "naturally flavored" although it contains only 5% orange juice. A new federal regulation stipulates that to be called 'natural' a drink must contain at least 10% fruit juice. How much pure orange juice must this manufacturer add to 900 gal of orange soda to conform to the new regulation?

Solution

Let x be the amount of orange juice to be added. Amount of the mixture is $(900 + x)$

Amount of orange juice in the first vat $= 0.05(900) = 45$

Amount of orange juice in the second vat $= 1.x = x$

Amount of orange juice in the mixture $= 0.10(900 + x)$

Then according to mathematical model

Amount of orange juice in the first vat + Amount of orange juice in the second vat = Amount of orange juice in the mixture

$45 + x = 0.10(900 + x)$

$45 + x = 90 + 0.1x$

$x = 50$

The manufacturer should add 50 gal of pure orange juice to the soda.

34. Because of an anticipated heavy rainstorm, the water level in a reservoir must be lowered by 1ft. Opening spillway A lowers the level by this amount in 4 hours, whereas opening the smaller spillway B does the job in 6 hours. How long will it take to lower the water level by 1ft if both spillways are opened?

Solution

Time it takes to lower level 1ft with A and B together $= x$

Distance A lowers level in 1 hour $= \frac{1}{4}$ ft

Distance B lowers level in 1 hour $= \frac{1}{6}$ ft

Distance A and B together lowers level in 1 hour $= \frac{1}{x}$ ft

Then according to mathematical model

Fraction done by A + Fraction done by B = Fraction done by both

$\frac{1}{4} + \frac{1}{6} = \frac{1}{x}$

$x = \frac{12}{5}$

It will takes $2\frac{2}{5}$ hours, or 2h 24 min, to lower the water level by 1ft if both spillways are open.

35. Solve the following systems of linear equations by matrix inversion method.

$2x - 2y = 4\ ;\ \ 3x + 2y = 6$

Solution.

$2x - 2y = 4$

$3x + 2y = 6$

by matrix inversion method

$$\begin{bmatrix} 2 & -2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

Let $A = \begin{bmatrix} 2 & -2 \\ 3 & 2 \end{bmatrix}$

$|A| = \begin{vmatrix} 2 & -2 \\ 3 & 2 \end{vmatrix} = (2)(2) - (-2)(3) = 4 + 6 = 10 \neq 0$

$\text{Adj}A = \begin{bmatrix} 2 & 2 \\ -3 & 2 \end{bmatrix}$

Since $AX = B$ therefore $X = A^{-1}B$

$X = \frac{1}{|A|} \times \text{Adj}A \times B$

$X = \frac{1}{10} \times \begin{bmatrix} 2 & 2 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 20 \\ 0 \end{bmatrix}$

$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$x = 2, y = 0$

36. Solve the following systems of linear equations by Cramer's rule.

$2x - 2y = 4\ ;\ \ 3x + 2y = 6$

Solution.

$2x - 2y = 4$

$3x + 2y = 6$

We have

$$\begin{bmatrix} 2 & -2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

Let $A = \begin{bmatrix} 2 & -2 \\ 3 & 2 \end{bmatrix}$

$|A| = \begin{vmatrix} 2 & -2 \\ 3 & 2 \end{vmatrix} = (2)(2) - (-2)(3) = 4 + 6 = 10 \neq 0$

$|A_x| = \begin{vmatrix} 4 & -2 \\ 6 & 2 \end{vmatrix} = (4)(2) - (-2)(6) = 8 + 12 = 20$

$|A_y| = \begin{vmatrix} 2 & 4 \\ 3 & 6 \end{vmatrix} = (2)(6) - (4)(3) = 12 - 12 = 0$

$x = \frac{|A_x|}{|A|} = \frac{20}{10} = 2, y = \frac{|A_y|}{|A|} = \frac{0}{10} = 0$

$x = 2, y = 0$

37. Use Cramer's rule to solve the system.

$3x_1 + x_2 - x_3 = -4$

$x_1 + x_2 - 2x_3 = -4$

$-x_1 + 2x_2 - x_3 = 1$

Solution.

$3x_1 + x_2 - x_3 = -4$

$x_1 + x_2 - 2x_3 = -4$

$-x_1 + 2x_2 - x_3 = 1$

We have

$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 1 & -2 \\ -1 & 2 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \\ 1 \end{bmatrix}$$

$$|A| = \begin{vmatrix} 3 & 1 & -1 \\ 1 & 1 & -2 \\ -1 & 2 & -1 \end{vmatrix} = 3(-1+4) - 1(-1-2) - 1(2+1) = 9 + 3 - 3 = 9 \neq 0$$

$$|A_{x_1}| = \begin{vmatrix} -4 & 1 & -1 \\ -4 & 1 & -2 \\ 1 & 2 & -1 \end{vmatrix} = -4(-1+4) - 1(4+2) - 1(-8-1) = -12 - 6 + 9$$

$$|A_{x_1}| = -9$$

$$|A_{x_2}| = \begin{vmatrix} 3 & -4 & -1 \\ 1 & -4 & -2 \\ -1 & 1 & -1 \end{vmatrix} = 3(4+2) + 4(-1-2) - 1(1-4) = 18 - 12 + 3$$

$$|A_{x_2}| = 9$$

$$|A_{x_3}| = \begin{vmatrix} 3 & 1 & -4 \\ 1 & 1 & -4 \\ -1 & 2 & 1 \end{vmatrix} = 3(1+8) - 1(1-4) - 1(2+1) = 27 + 3 - 12$$

$$|A_{x_3}| = 18$$

$$x_1 = \frac{|A_{x_1}|}{|A|} = \frac{-9}{9} = -1, x_2 = \frac{|A_{x_2}|}{|A|} = \frac{9}{9} = 1, x_3 = \frac{|A_{x_3}|}{|A|} = \frac{18}{9} = 2$$

$$x_1 = -1, x_2 = 1, x_3 = 2$$

Echelon Form of a Matrix:

A matrix is said to be in echelon form if it has the following structure;

i. All the non – zero rows proceed the zero rows.
ii. The first non – zero element in each row is **1**.
iii. The preceding number of zeros before the first non – zero element **1** in each row should be greater than its previous row.

For example followings are in echelon form.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 7 & 3 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Reduced Echelon Form of a Matrix:

A matrix is said to be in reduced echelon form if it has the following structure;

i. Matrix should be in echelon form.
ii. If the first non – zero element **1** in the i^{th} row of matrix lies in the j^{th} column then all other elements in the j^{th} column are zero.

For example followings are in reduced echelon form.

$$\begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Remark about echelon forms:

i. Every matrix has a unique reduced row echelon form.
ii. Row echelon forms are not unique.
iii. Although row echelon forms are not unique, the reduced row echelon form and all row echelon forms of a matrix A have the same number of zero rows, and the leading 1's always occur in the same positions.

Gaussian Elimination

The main method for solving the general system of linear equations is called Gaussian elimination. It essentially consists of two parts:

Part A. (Forward Elimination) Step-by-step reduction of the system yielding either a degenerate equation with no solution (which indicates the system has no solution) or an equivalent simpler system in triangular or echelon form.

Part B. (Backward Elimination) Step-by-step back-substitution to find the solution of the simpler system.

Gaussian Elimination steps (Procedure):

i. Reduce the augmented matrix into echelon form. In this way, the value of last variable is calculated.
ii. Then by backward substitution, the values of remaining unknown can be calculated.

Example: Solve the matrix using Gauss's Elimination method.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Solution: Firstly we reduce the given matrix in echelon form.

Step – I: locate the left most column that does not consist entirely of zeros.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Step – II: interchange the top row with another row, if necessary, to bring a non – zero entry to the top of the column found in step – I.

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \qquad \sim R_{12}$$

Step – III: if the entry that is now at the top of the column found in step – I is 'a', multiply the first row by '1/a' in order to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \qquad \sim \frac{1}{2}R_1$$

Step – IV: add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zero.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix} \qquad \sim R_3 - 2R_1$$

Step – V: now cover the top row in the matrix and begin again with step – I applied to the submatrix that remains or remaining rows. Continue in this way until the entire matrix is in row echelon form.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix} \qquad \sim -\frac{1}{2}R_2$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} \qquad \sim R_3 - 5R_2$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \qquad \sim 2R_3$$

Hence above matrix is in row echelon form.

Thus corresponding system is

$x_5 = 2$

$x_3 - \frac{7}{2}x_5 = -6$

$x_1 + 2x_2 - 5x_3 + 3x_4 + 6x_5 = 14$

Solving for leading variables we obtain and in next line solution using free variable

$x_5 = 2,\ x_3 = -6 + \frac{7}{2}x_5,\ x_1 = 14 - 2x_2 + 5x_3 - 3x_4 - 6x_5$

$x_5 = 2,\ x_3 = -6 + \frac{7}{2}(2) = 1$

$x_1 = 2 - 2x_2 + 5(1) - 3x_4 - 6(2) = 7 - 2x_2 - 3x_4$

Finally we express the general solution of the system parametrically by assigning the free variables x_2, x_4 arbitrary values 'r' and 's' respectively. This yield

$x_1 = 7 - 2r - 3s\ , x_2 = r\ \ , x_3 = 1\ , x_4 = s\ \ ,\ \ x_5 = 2$

Above is our required solution.

Question

Solve the following system by Gauss's Elimination method

$3x_1 + x_2 - x_3 = -4$

$x_1 + x_2 - 2x_3 = -4$

$-x_1 + 2x_2 - x_3 = 1$

Solution

$$\begin{bmatrix} 3 & 1 & -1 & \vdots -4 \\ 1 & 1 & -2 & \vdots -4 \\ -1 & 2 & -1 & \vdots\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & 1 & -1 & \vdots -4 \\ 0 & 2 & -5 & \vdots -8 \\ 0 & 7 & -4 & \vdots -1 \end{bmatrix} \sim 3R_2 - R_1 \ ; \ \sim 3R_3 + R_1$$

$$\Rightarrow \begin{bmatrix} 1 & \frac{1}{3} & -\frac{1}{3} & \vdots -\frac{4}{3} \\ 0 & 2 & -5 & \vdots -8 \\ 0 & 7 & -4 & \vdots -1 \end{bmatrix} \sim \frac{1}{3}R_1$$

$$\Rightarrow \begin{bmatrix} 1 & \frac{1}{3} & -\frac{1}{3} & \vdots -\frac{4}{3} \\ 0 & 1 & -\frac{5}{2} & \vdots -4 \\ 0 & 7 & -4 & \vdots -1 \end{bmatrix} \sim \frac{1}{2}R_2$$

$$\Rightarrow \begin{bmatrix} 1 & \frac{1}{3} & -\frac{1}{3} & \vdots -\frac{4}{3} \\ 0 & 1 & -\frac{5}{2} & \vdots -4 \\ 0 & 0 & \frac{27}{2} & \vdots\ 27 \end{bmatrix} \sim R_3 - 7R_2$$

$\Rightarrow \frac{27}{2}x_3 = 27 \Rightarrow \boldsymbol{x_3 = 2}$

$\Rightarrow x_2 - \frac{5}{2}x_3 = -4 \Rightarrow x_2 - \frac{5}{2}(2) = -4 \Rightarrow \boldsymbol{x_2 = 1}$

$\Rightarrow x_1 + \frac{1}{3}x_2 - \frac{1}{3}x_3 = -\frac{4}{3} \Rightarrow x_1 + \frac{1}{3}(1) - \frac{1}{3}(2) = -\frac{4}{3} \Rightarrow \boldsymbol{x_1 = -1}$

Hence solution set is $\{-1,1,2\}$

Example

Solve the linear system by Gauss Elimination method. Or show that system has no solution.

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 2 & 4 & -4 & 6 & 10 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix}$$

Solution:

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 2 & 4 & -4 & 6 & 10 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 6 & 7 \end{bmatrix} \quad \sim R_2 - 2R_1 \; also \sim R_3 - 3R_1$$

$$\Rightarrow \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix} \quad \sim R_3 - \frac{3}{2}R_2$$

The matrix is now in echelon form. The third row of the echelon matrix corresponds to the degenerate equation $0x_1 + 0x_2 + 0x_3 + 0x_4 = -2$ which has no solution, thus the system has no solution.

Gauss Jordan Elimination

Procedure:

i. In this method we reduce the augmented matrix into <u>reduced echelon</u> form. In this way, the value of last variable is calculated.
ii. Then by backward substitution, the values of remaining unknown can be calculated.

Example

Solve the matrix using Gauss's Elimination method.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Solution: Firstly we reduce the given matrix in reduced echelon form.

Step – I: locate the left most column that does not consist entirely of zeros.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Step – II: interchange the top row with another row, if necessary, to bring a non – zero entry to the top of the column found in step – I.

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \quad \sim R_{12}$$

Step – III: if the entry that is now at the top of the column found in step – I is 'a', multiply the first row by '1/a' in order to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \quad \sim \frac{1}{2}R_1$$

Step – IV: add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zero.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix} \quad \sim -2R_1 + R_3$$

Step – V: now cover the top row in the matrix and begin again with step – I applied to the submatrix that remains or remaining rows. Continue in this way until the entire matrix is in row echelon form.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix} \quad \sim -\frac{1}{2}R_2$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} \quad \sim R_3 - 5R_2$$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \quad \sim 2R_3$$

Step – VI: Beginning with the last non – zero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1's.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \quad \sim R_2 + \frac{7}{2}R_3 \Rightarrow \begin{bmatrix} 1 & 2 & -5 & 3 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \quad \sim R_1 - 6R_3$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \quad \sim R_1 + 5R_2$$

Hence above matrix is in row reduced echelon form. Thus corresponding system is

$x_5 = 2,\ x_3 = 1,\ x_1 + 2x_2 + 3x_4 = 7$

Solving for leading variables we obtain and in next line solution using free variable

$x_5 = 2,\ x_3 = 1,\ x_1 = -2x_2 - 3x_4 - 7$

Finally we express the general solution of the system parametrically by assigning the free variables x_2, x_4 arbitrary values 'r' and 's' respectively. These yields

$x_1 = -2r - 3s - 7\ , x_2 = r\ \ , x_3 = 1\ , x_4 = s\ \ ,\ \ x_5 = 2$

Above is our required solution.

Example:

Solve the linear system by Gauss's Jordan Elimination method.

$x_1 + 3x_2 - 2x_3 + 2x_5 = 0$

$2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 = -1$

$5x_3 + 10x_4 + 15x_6 = 5$

$2x_1 + 6x_2 + 8x_4 + 4x_5 + 18x_6 = 6$

Solution:

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 2 & 6 & 0 & 8 & 4 & 18 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix} \quad \sim R_2 - 2R_1 \ and \ R_2 - 2R_4$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix} \sim -1R_2 \Rightarrow \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \end{bmatrix} \sim R_3 - 5R_2 \ and \ R_4 - 4R_2$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim R_{34} \Rightarrow \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \frac{1}{6}R_3 \qquad \text{Echelon form.}$$

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim R_2 - 3R_3 \Rightarrow \begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim R_1 + 2R_2 \qquad \text{reduced E. form.}$$

Thus corresponding system is

$x_6 = \frac{1}{3}, \ x_3 + 2x_4 = 0, \ x_1 + 3x_2 + 4x_4 + 2x_5 = 0$

Solving for leading variables we obtain and in next line solution using free variable

$x_6 = \frac{1}{3}, \ x_3 = -2x_4 \, , \ x_1 = -3x_2 - 4x_4 - 2x_5$

Finally we express the general solution of the system parametrically by assigning the free variables x_2, x_4, x_5 arbitrary values 'r', 's' and 't' respectively. These yields

$x_1 = -3r - 4s - 2t \, , x_2 = \mathrm{r} \, , \ x_3 = -2s, \ x_4 = s \, , \ x_5 = t, \ x_6 = \frac{1}{3}$

Above is our required solution.

Inequality/Inequation

An expression involving the signs $<, \leq, >, \geq, \neq$ is called inequality. It is a statement in which two algebraic expression are not equal.

For example: $x < 3, x \leq 5, x > -2, x \geq 17, x \neq 10$

Strict Inequality

An expression involving the signs $<, >$ is called strict inequality.

For example: $x < 3, x > -20$

Solution of an Inequality

Variable that make the inequality true is called solution of that inequality.

38. Solve $3x + 7 \geq 1$

Solution

$$3x + 7 \geq 1 \Rightarrow 3x + 7 - 7 \geq 1 - 7 \Rightarrow 3x \geq -6 \Rightarrow \frac{3x}{3} \geq -\frac{6}{3} \Rightarrow x \geq -2$$

39. Solve $-5x - 8 \geq 2$

Solution

$$-5x - 8 \geq 2 \Rightarrow -5x - 8 + 8 \geq 2 + 8 \Rightarrow -5x \geq 10 \Rightarrow \frac{-5x}{-5} \leq \frac{10}{-5} \Rightarrow x \leq -2$$

40. Solve $2x + 3 > 7$ or $4x - 1 < 3$

Solution

$2x + 3 > 7$ or $4x - 1 < 3$

$2x > 4$ or $4x < 4$

$x > 2$ or $x < 1$

Solution Set = $\{x | x \in R \wedge x > 2 \text{ or } x < 1\}$

Linear Programming

Programming that deals with the optimization (maximization or minimization) of the function is called linear programming.

Boundary of Half Plane

$ax + by < c$ is called half plane region and line $ax + by = c$ is called boundary of half plane.

Left and Right Half Plane

Vertical line divides the plane into left and right half plane.

Upper and Lower Half Plane

Non – Vertical line divides the plane into lower and upper half plane.

Vertex or Corner Point

A point of a solution region where two of its boundary lines interest is called vertex.

Non – Negative Constraints

The variable used in the system of linear inequalities relating to the problem of every day life are non – negative and are called non – negative constraints.

Decision Variables

The non – negative constraints play an important rule for taking decision, so these variables are called decision variables.

Solution Region

We draw graph of each inequality in the system of the same coordinate axes and then take intersection of the graph. The common region so obtained is called the solution region.

Feasible Region

A region which is restricted to the first quadrant is called feasible region.

Feasible Solution

Each point of feasible region is called feasible solution.

Optimal Solution

The feasible solution which maximizes or minimizes the objective function is called the optimal solution.

Objective Function

A function which is to be maximized or minimized is called an objective function.

Problem Constraints

The systems of linear inequalities involved in the problem concerned are called problem constraints.

Convex

If the line segment obtained by joining any two points of a region lies entirely within the region then the region is called convex.

41. Graph the system of linear inequalities $x - y \leq 6$; $2x + y \geq 2$

Solution

$x - y \leq 6$ ………(i)

$2x + y \geq 2$ ………(ii)

Associated equations

$x - y = 6$ ………(iii)

$2x + y = 2$ ………(iv)

To find Points

(iii) $\Rightarrow$ Put $x = 0$, $y = -3$ then point is $(0, -3)$

(iii) $\Rightarrow$ Put $y = 0$, $x = 6$ then point is $(6,0)$

(iv) $\Rightarrow$ Put $x = 0$, $y = 2$ then point is $(0,2)$

(iv) $\Rightarrow$ Put $y = 0$, $x = 1$ then point is $(1,0)$

To check Region put $(0,0)$ in (i) and (ii)

(i) $\Rightarrow 0 < 6$ true

(ii) $\Rightarrow 0 > 2$ false

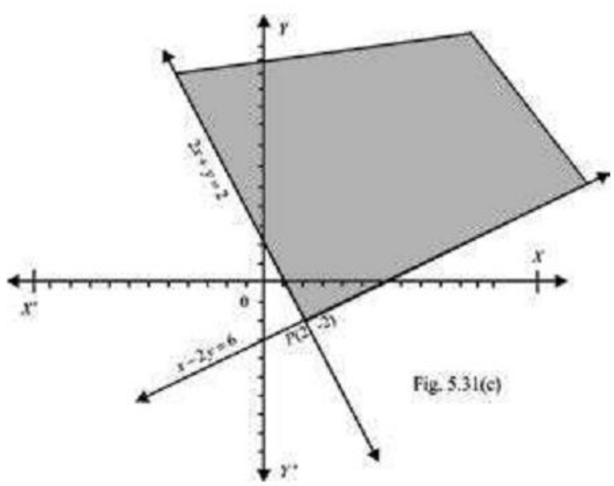


Fig. 5.31(c)

42. Indicate the solution of linear inequalities by shading $2x - 3y \leq 6$; $2x + 3y \leq 12$

Solution

$2x - 3y \leq 6$ ....(i)

$2x + 3y \leq 12$ ....(ii)

Associated equations

$2x - 3y = 6$ ....(iii)

$2x + 3y = 12$ ....(iv)

To find Points

(iii) $\Rightarrow$ Put $x = 0$, $y = -2$ then point is $(0, -2)$

(iii) $\Rightarrow$ Put $y = 0$, $x = 3$ then point is $(3,0)$

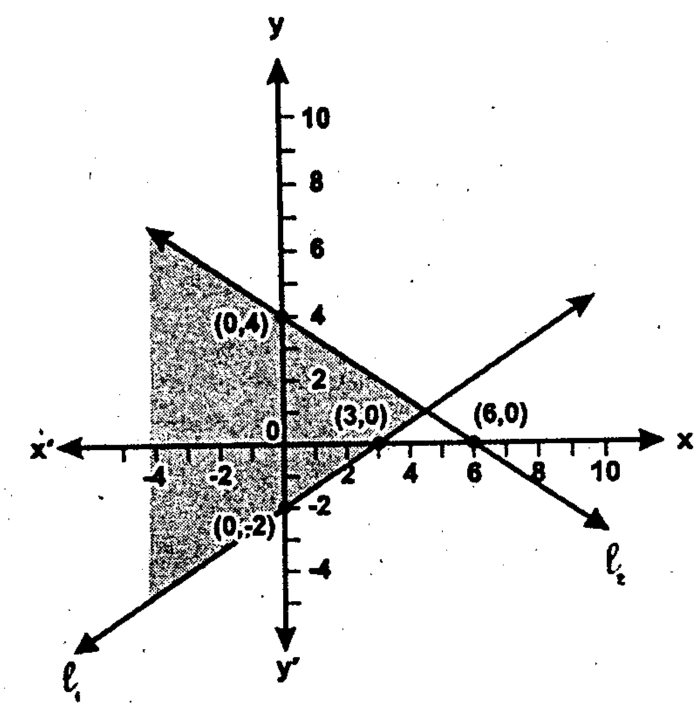
(iv) $\Rightarrow$ Put $x = 0$, $y = 4$ then point is $(0,4)$

(iv) $\Rightarrow$ Put $y = 0$, $x = 6$ then point is $(6,0)$

To check Region put $(0, 0)$ in (i) and (ii)

(i) $\Rightarrow 0 < 6$ true

(ii) $\Rightarrow 0 < 12$ true



43. Graph the solution region also find the corner points in each case

$2x - 3y \leq 6$; $2x + 3y \leq 12$

Solution

$2x - 3y \leq 6$ ....(i)

$2x + 3y \leq 12$ ....(ii)

Associated equations

$2x - 3y = 6$ ....(iii)

$2x + 3y = 12$ ....(iv)

To find Points

(iii) $\Rightarrow$ Put $x = 0$, $y = -2$ then point is $(0, -2)$

(iii) $\Rightarrow$ Put $y = 0$, $x = 3$ then point is $(3,0)$

(iv) $\Rightarrow$ Put $x = 0$, $y = 4$ then point is $(0,4)$

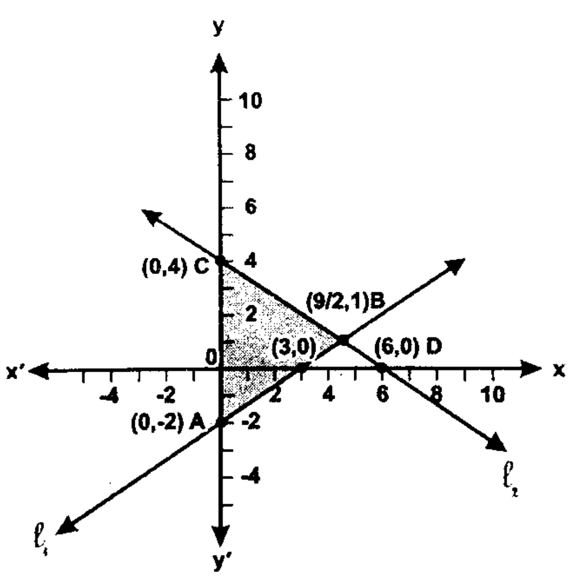
(iv) $\Rightarrow$ Put $y = 0$, $x = 6$ then point is $(6,0)$

To check Region put $(0,0)$ in (i) and (ii)

(i) $\Rightarrow 0 < 6$ true and (ii) $\Rightarrow 0 < 12$ true

Adding (iii) and (iv) we have $x = \frac{9}{2}$ and putting $x = \frac{9}{2}$ in (iii) we have $y = 1$

Corner Points: $A(0, -2), B\left(\frac{9}{2}, 1\right), C(0,4)$



44. Graph the solution region of the following system of linear inequalities by shading

$3x - 4y \leq 12$; $3x + 2y \geq 3$; $x + 2y \leq 9$

Solution

$3x - 4y \leq 12$ ………(i)

$3x + 2y \geq 3$ ………(ii)

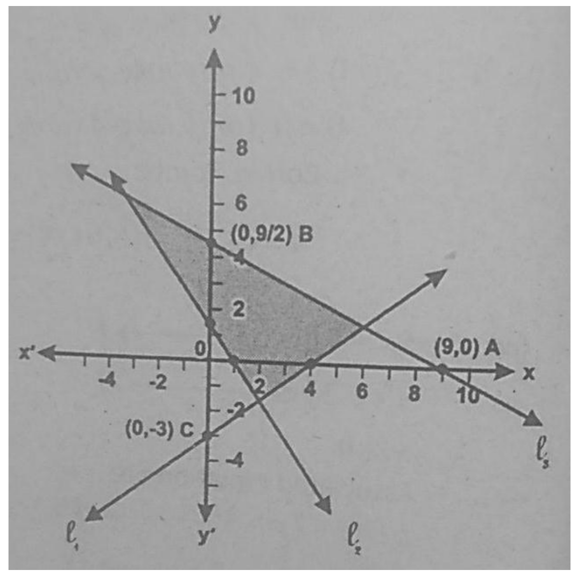
$x + 2y \leq 9$ ………(iii)

Associated equations

$3x - 4y = 12$ ………(iv)

$3x + 2y = 3$ ………(v)

$x + 2y = 9$ ………(vi)



To find Points

(iv) $\Rightarrow$ Put $x = 0$, $y = -3$ then point is $(0, -3)$

(iv) $\Rightarrow$ Put $y = 0$, $x = 4$ then point is $(4,0)$

(v) $\Rightarrow$ Put $x = 0$, $y = \frac{3}{2}$ then point is $\left(0, \frac{3}{2}\right)$

(v) $\Rightarrow$ Put $y = 0$, $x = 1$ then point is $(1,0)$

(vi) $\Rightarrow$ Put $x = 0$, $y = \frac{9}{2}$ then point is $\left(0, \frac{9}{2}\right)$

(vi) $\Rightarrow$ Put $y = 0$, $x = 9$ then point is $(9,0)$

To check Region put $(0, 0)$ in (i), (ii) and (iii)

(i) $\Rightarrow 0 < 12$ true, (ii) $\Rightarrow 0 > 3$ false and (iii) $\Rightarrow 0 < 9$ true

45. Graph the feasible region also find corner points

$5x + 7y \leq 35$; $x - 2y \leq 4$; $x \geq 0, y \geq 0$

Solution

$5x + 7y \leq 35$ ....(i)

$x - 2y \leq 4$ ....(ii)

Associated equations

$5x + 7y = 35$ ....(iii)

$x - 2y = 4$ ....(iv)

To find Points

(iii) $\Rightarrow$ Put $x = 0$, $y = 5$ then point is $(0,5)$

(iii) $\Rightarrow$ Put $y = 0$, $x = 7$ then point is $(7,0)$

(iv) $\Rightarrow$ Put $x = 0$, $y = -2$ then point is $(0, -2)$

(iv) $\Rightarrow$ Put $y = 0$, $x = 4$ then point is $(4,0)$

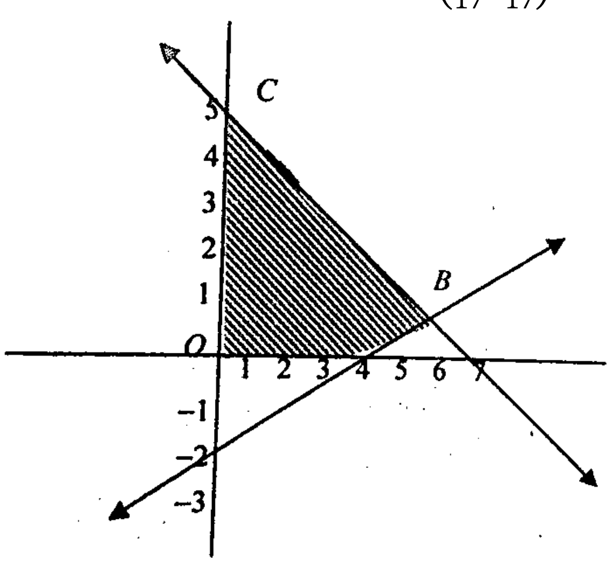
To check Region put $(0, 0)$ in (i) and (ii)

(i) $\Rightarrow 0 < 35$ true and (ii) $\Rightarrow 0 < 4$ true

Multiplying (iv) with 5 and Subtracting from (iii)

We have $y = \frac{15}{17}$ and putting $y = \frac{15}{17}$ in (iv) we have $x = \frac{98}{17}$

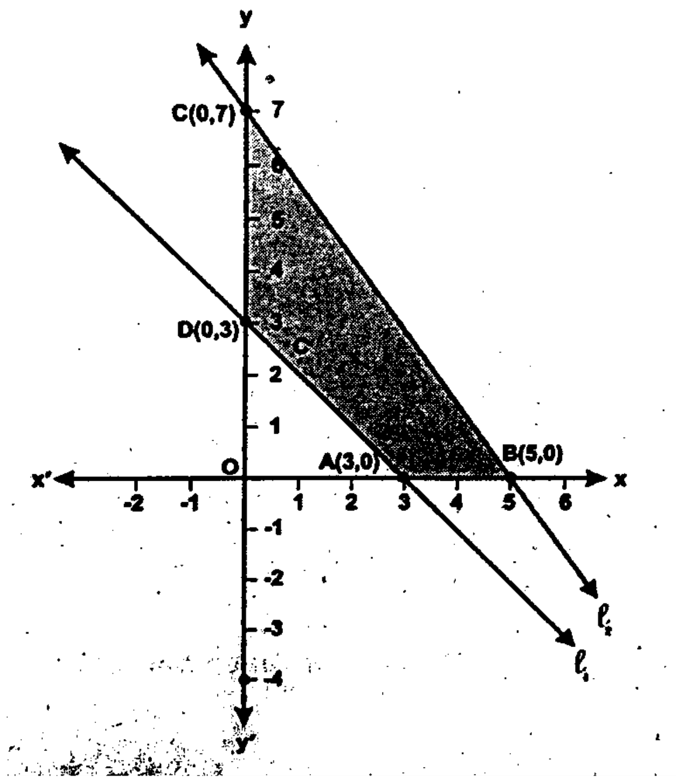
Corner Points: $O(0,0), A(4,0), B\left(\frac{98}{17}, \frac{15}{17}\right), C(0,5)$



46. Maximize as well as minimize $z = 2x + y$ for given linear inequalities

$x + y \geq 3$; $7x + 5y \leq 35$; $x \geq 0, y \geq 0$

Solution



$x + y \geq 3$ ....(i)

$7x + 5y \leq 35$ ....(ii)

Associated equations

$x + y = 3$ ....(iii)

$7x + 5y = 35$ ....(iv)

To find Points

(iii) $\Rightarrow$ Put $x = 0$, $y = 3$ then point is $(0,3)$

(iii) $\Rightarrow$ Put $y = 0$, $x = 3$ then point is $(3,0)$

(iv) $\Rightarrow$ Put $x = 0$, $y = 7$ then point is $(0,7)$

(iv) $\Rightarrow$ Put $y = 0$, $x = 5$ then point is $(5,0)$

To check Region put $(\mathbf{0}, \mathbf{0})$ in $(\mathbf{i})$ and $(\mathbf{ii})$

(i) $\Rightarrow 0 > 3$ false and (ii) $\Rightarrow 0 < 35$ true

Corner Points: $A(3,0), B(5,0), C(0,7), D(0,3)$

At A: $z = f(3,0) = 2(3) + 0 = 6$

At B: $z = f(5,0) = 2(5) + 0 = 10$

At C: $z = f(0,7) = 2(0) + 7 = 7$

At D: $z = f(0,3) = 2(0) + 3 = 3$

So

$z = 2x + y$ is maximum at $(5,0)$

$z = 2x + y$ is minimum at $(0,3)$

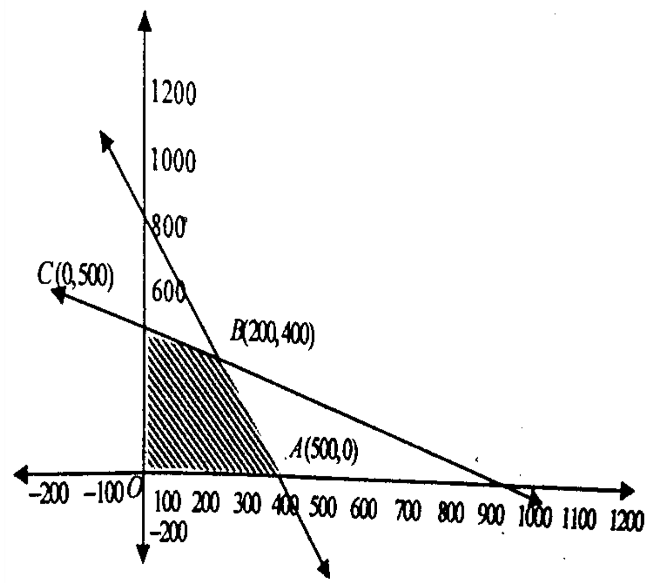
47. A machine can produce product A by using 2 units of chemical and 1 unit of a compound or can produce product B by using 1 unit of chemical and 2 units of the compound. Only 800 units of chemical and 1000 units of the compound are available. The profits per unit of A and B are Rs. 30 and Rs. 20 respectively, maximize the profit function.

Solution

Suppose units of A are x and units of B are y then $f(x,y) = 2x + y$ then by condition

$2x + y \leq 800$ ....(i)

$x + 2y \leq 1000$ ....(ii)



Associated equations

$2x + y = 800$ ....(iii)

$x + 2y = 1000$ ....(iv)

To find Points

(iii) $\Rightarrow$ Put $x = 0$, $y = 800$ then point is $(0,800)$

(iii) $\Rightarrow$ Put $y = 0$, $x = 400$ then point is $(400,0)$

(iv) $\Rightarrow$ Put $x = 0$, $y = 500$ then point is $(0,500)$

(iv) $\Rightarrow$ Put $y = 0$, $x = 1000$ then point is $(1000,0)$

To check Region put $(\mathbf{0},\mathbf{0})$ in $(\mathbf{i})$ and $(\mathbf{ii})$

(i) $\Rightarrow 0 < 800$ true and (ii) $\Rightarrow 0 < 1000$ true

Multiplying (iii) with 2 and subtracting (iii) from (iv)

We have $x = 200$ and putting $x = 200$ in (iv) we have $y = 400$

Corner Points: $O(0,0), A(400,0), B(200,400), C(0,500)$

At O: $f(0,0) = 30(0) + 20(0) = 0$

At A: $f(400,0) = 30(400) + 20(0) = 12000$

At B: $f(200,400) = 30(200) + 20(400) = 14000$

At C: $f(0,500) = 30(0) + 20(500) = 1000$

So profit is maximum at corner $B(200,400)$.

Exercise

1) Solve the following linear equations in one variable.

i. $5x - 2 - x = 4 - 3x - 27$

ii. $7(2 - 5x) + 27 = 18x - 3(8 - 4x)$

iii. $\frac{5x}{4} + \frac{1}{2} = 0$

iv. $\frac{x}{2} + \frac{x}{3} - \frac{x}{4} + \frac{x}{5} = 7\frac{5}{6}$

v. $\frac{y+1}{3} + \frac{y+1}{2} = 2 - \frac{y+3}{2}$

vi. $0.5x = 6.3 - 0.2x$

2) Solve the following radical equations.

i. $\sqrt{2x} = 4$

ii. $\sqrt{x - 3} = 2$

iii. $\sqrt{5x - 4} = 14$

iv. $5 - \sqrt{2x - 1} = 0$

v. $\sqrt{9 - 2x} = \sqrt{5x - 12}$

3) Solve the following absolute value equations.

i. $\frac{|10-x|}{5} = \frac{|2x-5|}{2} \quad ; x \in R$

ii. $|x + 2| = 6$

iii. $|5x| + 10 = 5$

iv. $\frac{|1-2y|}{4} = 3$

v. $|z + 3| - 3 = 5 - |z + 3|$

4) Solve the following inequalities.

i. $2 \le x \le 5$

ii. $-4 < x < -\frac{3}{2}$

iii. $x \le 4 \quad ; x \in W$

iv. $3x + 21 < 1 - x$ or $3x + 8 < 3 - 2x$

v. $1 - 5x > 16$ and $3 - \frac{3x}{2} \le 9$

5) Consider the following system of linear equations:

$$2x_1 - 4x_2 - x_3 = 1$$

$$x_1 - 3x_2 + x_3 = 1$$

$$3x_1 - 5x_2 - 3x_3 = 1$$

Determine whether given 3 – tuples are solutions of the system?

a) $(3,1,1)$
b) $(3,-1,1)$
c) $(13,5,2)$
d) $\left(\frac{13}{2},\frac{5}{2},2\right)$
e) $(17,7,5)$

6) Consider the following system of linear equations:
$x + 2y - 2z = 3;\ 3x - y + z = 1;\ -x + 5y - 5z = 5$
Determine whether given 3 – tuples are solutions of the system?

a) $\left(\frac{5}{7},\frac{8}{7},1\right)$
b) $\left(\frac{5}{7},\frac{8}{7},0\right)$
c) $(5,8,1)$
d) $\left(\frac{5}{7},\frac{10}{7},\frac{2}{7}\right)$

7) Solve the linear system by Gauss's Elimination method.

i. $x_1 + x_2 + 2x_3 = 8$
$-x_1 - 2x_2 + 3x_3 = 1$
$3x_1 - 7x_2 + 4x_3 = 10$

ii. $2x_1 + 2x_2 + 2x_3 = 0$
$-2x_1 + 5x_2 + 2x_3 = 1$
$8x_1 + x_2 + 4x_3 = -1$

iii. $x - y + 2z - w = -1$
$2x + y - 2z - 2w = -2$
$-x + 2y - 4z + w = 1$
$3x - 3w = -3$

iv. $-2b + 3c = 1$
$3a + 6b - 3c = -2$
$6a + 6b + 3c = 5$

8) Solve the linear system by Gauss's Jordan Elimination method.

i. $3x + y = 5$; $2x - y + z = 4$; $3x - z = 3$

ii. $-2x_2 + 3x_3 = 1$; $3x_1 + 5x_2 + 3x_3 = -2$; $6x_1 - 6x_2 + 3x_3 = 5$

iii. $2x_1 - x_2 - x_3 = 4$; $3x_1 + 4x_2 - 2x_3 = 11$; $3x_1 - 2x_2 + 4x_3 = 11$

iv. $x_1 + 3x_2 + 2x_5 = 0$; $2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 = 1$

v. $5x_3 + 10x_4 + 15x_6 = 5$; $2x_1 + 6x_2 + 8x_4 + 18x_6 = 6$

9) Solve the following systems of linear equations by matrix inversion method.

a) $2x + y = 3$; $6x + 5y = 1$

b) $4x + 2y = 8$; $3x - y = -1$

c) $x - 2y + z = -1$; $3x + y - 2z = 4$; $y - z = 1$

10) Solve the following systems of linear equations by Cramer's rule.

a) $2x + y = 3$; $6x + 5y = 1$

b) $4x + 2y = 8$; $3x - y = -1$

c) $2x + 2y + z = 3$; $3x - 2y - 2z = 1$ $5x + y - 3z = 2$

11) In each part, solve the linear system, if possible, and use the result to determine whether the lines represented by the equations in the system have zero, one, or infinitely many points of intersection. If there is a single point of intersection, give its coordinates, and if there are infinitely many, find parametric equations for them.

a) $3x - 2y = 4$ and $6x - 4y = 9$

b) $2x - 4y = 1$ and $4x - 8y = 2$

c) $x - 2y = 0$ and $x - 4y = 8$

12) In each part use parametric equations to describe the solution set of linear equations.

a) $7x - 5y = 3$

b) $x + 10y = 2$

c) $3x_1 - 5x_2 + 4x_3 = 7$

d) $-8x_1 + 2x_2 - 5x_3 + 6x_4 = 1$

e) $3v - 8w + 2x - y + 4z = 0$

f) $4x_1 + 2x_2 + 3x_3 + x_4 = 20$

g) $x_1 + 3x_2 - 12x_3 = 3$

h) $v + w + x - 5y + 7z = 0$

13) In each part use parametric equations to describe the infinitely many solutions of linear equations.

a) $2x - 3y = 1$ and $6x - 9y = 3$
b) $6x_1 + 2x_2 = -8$ and $3x_1 + x_2 = -4$
c) $x_1 + 3x_2 - x_3 = -4$, $3x_1 + 9x_2 - 3x_3 = -12$ and $-x_1 - 3x_2 + x_3 = 4$
d) $2x - y + 2z = -4$, $6x - 3y + 6z = -12$ and $-4x + 2y - 4z = 8$

14) Graph the solution region of the following system of linear inequalities by shading also find the corner points in each case.

a) $x + y \leq 5$; $-2x + y \leq 2$; $y \geq 0$
b) $5x + 7y \leq 35$; $x - 2y \leq 4$; $x \geq 0$
c) $3x + 2y \geq 6$; $x + 3y \leq 6$; $y \geq 0$
d) $2x + y \leq 4$; $2x - 3y \geq 12$; $x + 2y \leq 6$
e) $3x - 2y \geq 3$; $x + 4y \leq 12$; $3x + y \leq 12$

15) Graph the feasible region of the following system of linear inequalities by shading also find the corner points in each case.

a) $x + y \leq 5$; $-2x + y \leq 2$; $x \geq 0, y \geq 0$
b) $2x - 3y \leq 6$; $2x + 3y \leq 12$; $x \geq 0, y \geq 0$
c) $3x + 2y \geq 6$; $x + y \leq 4$; $x \geq 0, y \geq 0$
d) $2x + y \leq 20$; $8x + 15y \geq 120$; $x + y \leq 11$; $x \geq 0, y \geq 0$
e) $x + 2y \leq 14$; $3x + 4y \leq 36$; $2x + y \leq 10$; $x \geq 0, y \geq 0$

16) Maximize $z = 2x + 5y$ for given linear inequalities
$2y - x \leq 8$; $x - y \leq 4$; $x \geq 0, y \geq 0$

17) Maximize $z = 2x + 3y$ for given linear inequalities
$2x + y \leq 8$; $x + 2y \leq 14$; $x \geq 0, y \geq 0$

18) Each unit of food X costs Rs. 25 and contains 2 units of protein and 4 units of iron while each unit of food Y costs Rs. 30 and contains 3 units of protein and 2 unit of iron. Each animal must receive at least 12 units of protein and 16 units of iron each day. How many units of each food should be fed to each animal at the smallest possible cost?

19) A dealer wishes to purchase a number of fans and sewing machines. He has only Rs. 5760 to invest and has space atmost for 20 items. A fan costs him Rs. 360 and a sewing machine costs Rs. 240. His expectation is that the can sell a fan at a profit of Rs. 22 and a sewing machine at a profit of Rs. 18. Assuming that he can sell all the items that he can buy, how should he invest his money in order to maximize his profit?

CHAPTER

5

STATISTICS

SAMPLING & DATA ANALYSIS

Statistics is the science of collecting, analyzing, and interpreting data to extract meaningful insights and inform decision-making. At the heart of statistical analysis lies sampling and data analysis, essential components of statistical research. **Sampling** involves selecting a subset of individuals, objects, or observations from a larger population to estimate population characteristics. Effective sampling methods ensure that the sample accurately represents the population, enabling reliable conclusions. **Data analysis** involves processing, transforming, and examining data to uncover patterns, trends, and relationships. Statistical techniques, such as descriptive statistics, inferential statistics, and graphical methods, are used to extract insights from data.

In this chapter we will learn about;

- Introduction, Probabilistic models, Statistical models.
- Data and distribution of data, tabulation.
- Graph, Charts, Misleading Graphs, Bivariate analysis
- Data tendencies via measure of location and spread of data
- Variability and Measure of dispersion
- Measuring relationships via Regression analysis and correlation.
  i.e. Linear Regression and Correlation
- **Statistical inference:** sampling techniques, estimation techniques and hypothesis testing for decision and policy making

The word "Statistics" is derived from the Latin word **Status** or the Italian word **Statista** or the German word **Statistik** or the French word **Statistique** meaning "a political state" or "the state – man's art". It is the discipline that includes procedures or techniques used to collect process, analyzes numerical data to inferences and to reach decision in the face of uncertainty. It is the science of collection, presentation, analysis and interpretation of numerical data.

According to this definition, there are four stages.

- Collection of data
- Presentation of data
- Analysis of data
- Interpretation

Use/Importance of Statistical Information

- To inform general public
- To explain things that have happened
- To justify a claim
- To provide general comparison
- To estimate the unknown quantities

Limitations of Statistical Information

- Deals with aggregates and not with individuals.
- Deals with numerically specified characteristics.

Types of Statistics

- **Descriptive/Deductive Statistics:** Branch which deals with concepts and methods concerned with summarization and description of the important aspects of numerical data. In it no conclusion is drawn about the population.
- **Inferential/Inductive Statistics:** Branch which deals with drawing inferences about the population on the basis of sample information. Its two most types are; Estimation and Testing of Hypothesis.

Introduction to Probabilistic Models

Probabilistic models are mathematical frameworks used to represent and analyze uncertain relationships between variables, providing a powerful tool for understanding complex phenomena in various fields, including engineering, economics, biology, and social sciences. These models quantify uncertainty using probability theory, enabling predictions, decision-making, and risk assessment under uncertainty. By capturing stochastic relationships, probabilistic models facilitate:

1. Uncertainty quantification
2. Predictive modeling
3. Decision-making under uncertainty
4. Risk analysis

Applications

1. Machine learning (deep learning, natural language processing)
2. Signal processing and communication
3. Finance (risk management, portfolio optimization)
4. Reliability engineering and maintenance
5. Biostatistics and epidemiology

Types of Probabilistic Models

Bayesian networks**,** Markov chain Monte Carlo (MCMC)**,** Probabilistic graphical models**,** Stochastic differential equations**,** Time-series analysis

Benefits

Handling uncertainty and ambiguity**,** Incorporating prior knowledge**,** Flexibility and adaptability**,** Interpretable results**,** Scalability

Deterministic Models/Relations

If there is an extract relationship between dependent and independent variables, and there is no chance of error, the regression line developed for such variables is called deterministic models. The following equation is an example of deterministic model;

Y = a + bX

Probabilistic Models/Relations

If there is no extract relationship between dependent and independent variables, and there is a chance of error, the regression line developed for such variables is called Probabilistic models. The following equation is an example of Probabilistic model;

Y = a + bX + error

Statistical Modeling

Statistical modeling is an elaborate method of generating sample data and making real world predictions using numerous statistical models and explicit assumptions. It helps data scientists visualize the relationship between random variables and strategically interpret dataset.

Types of Probabilistic Models

There are several statistical models, each designed to solve a specific research issue or data format. Here are a few common types of statistical models and their applications.

- **Linear Regression Models:** These models are used to represent the connection between a continuous result variable and one or more predictor variables. For example, depending on a person's height, age, and gender, a linear regression model may be used to estimate their weight.
- **Logistic Regression Models:** Logistic regression models are used to represent the connection between a binary outcome variable (for example, yes/no) and one or more predictor variables. For example, depending on age, blood pressure and cholesterol levels, a regression logistic model may be used to predict if a patient would have a heart attack.
- **Time Series Models:** Time series models are used to model data that change over time, such as stock prices, weather trends, or monthly sales numbers. These type of models may be applied to data to find trends, sessional patterns and other forms of temporal correlations.
- **Multilevel Models:** These models are used to model data having a hierarchical structure, such as pupil in school are patients in hospitals. Multilevel models can be used to investigate how individual – level and group – level factors impact outcomes, as well as to account for the fact that people in the same group maybe more similar to each other than those in different groups.
- **Structural Equation Models:** These types of models are used to represent complicated interactions between several variables. Structural equation model can be used to evaluate ideas regarding casual links between variables and to quantify their strength and direction.
- **Clustering Models:** Clustering models are used to bring together comparable observations based on their similarities in terms of features. Clustering algorithm can be used to uncover patterns in data that would be difficult to detect using other approaches.

Distributions

- A listing of all classes of the data and their frequencies is called a **Frequency distribution**. It is a tabular arrangement for classifying data into different groups and the number of observations falling in each group corresponds to the respective group. On the basis of type of variables, it has two types;
  Discrete frequency distribution
  Continuous frequency distribution
- The data presented in the form of frequency distribution is called **Grouped Data**.
- A listing of all classes and their relative frequencies is called a **Relative Frequency distribution**. Most distributions show frequencies as well as relative frequencies.

1. Discrete Frequency Table by using a Tally Column:

20 coins are tossed 5 times and the number of heads recorded at each toss are given below; 3,4,2,3,3,5,2,2,2,1,1,2,1,4,2,2,3,3,4,2.

Make frequency distribution of number of heads observed.

Solution: Let X = number of heads. The frequency distribution is given below;

X	Tally Marks	frequency (f)
1	|||	3
2	卌 |||	8
3	卌	5
4	|||	3
5	|	1

2. Continuous Frequency Table by listing Actual Values:

For data given below;

51,55,32,41,22,30,35,53,30,60,59,15,7,18,40,49,40,25,14,18,19,2,43,22,39,26,34,19,10, 17, 47,38,13,30,34,54,10,21,51,52.

Make frequency distribution with a class interval of size 10.

Solution:

Class/Groups	Observations	frequency (f)
0 – 9	2,7	2
10 – 19	10,10,13,14,15,17,18,18,19,19	10
20 – 29	21,22,22,25,26	5
30 – 39	30,30,30,32,34,34,35,38,39	9
40 – 49	40,40,41,43,47,49	6
50 – 59	51,51,52,53,54,55,59	7
60 – 69	60	1

3. **Continuous Frequency Table:**

Bradley worked a summer job to earn money for college. His weekly hours over a 12 week period were 25, 32, 36, 32, 18, 28, 30, 36, 12, 16, 35, 36. Find Distribution table.

Solution

The Distribution would be as follows:

Hours	Frequency	Relative Frequency
10-19	3	$\frac{3}{12} = 0.25 = 25\%$
20-29	2	$\frac{2}{12} = 0.167 = 17\%$
30-39	7	$\frac{7}{12} = 0.583 = 58\%$
Total	12	100%

Remember;

- **Class Limits:** The minimum and the maximum values defined for a class or group are called Class Limits. The minimum value is called the **lower class limit** and maximum value is called the **upper class limit** of the class.
- **Class Boundaries:** The real class limits of a class are called **class boundaries**. A class boundary is obtained by adding two successive class limits and dividing the sum by 2. The value so obtained is taken as **upper class boundary** for the previous class and **lower class boundary** for the next class.
- **Mid – Point/ Class Mark:** For a given class the average of that class obtained by dividing the sum of upper and lower class by 2, is called the mid – point of class mark of that class.
- **Interval/ Class Width:** Difference between the class boundaries.
- **Cumulative Frequency:** The total of frequency up to an upper class limit or boundary is called the cumulative frequency.

Classes	Frequency (f)	Class Boundaries	Mid Point	Cumulative Frequency
10 – 14	5	9.5 – 14.5	12	5
15 – 19	12	14.5 – 19.5	17	5 + 12 = 17
20 – 24	30	19.5 – 24.5	22	17 + 30 = 47
25 – 29	25	24.5 – 29.5	27	47 + 25 = 72
30 – 34	6	29.5 – 34.5	32	72 + 6 = 78

Some Facts about Data

- **Observation:** It is a fact or figure; we collect about a given variable. It can be expressed as a number or as a quality.
- **Data:** The collection of raw fact and figures is called data.
- **Data Set:** The collection of observations on one or more variables.
- **Cross Section Data:** Data collected on different elements at the same point in time or for the same period of time are called cross section data.
- **Time Series Data:** Data collected on the same element for the same variable at different points in time or for different periods of time are called time series data.
- **Discrete Data:** A data which is generated by a discrete variable is called discrete data.
- **Continuous Data:** A data which is generated by a continuous variable is called continuous data.
- **Datum:** A single numerical fact is datum.
- There are two types of data: Primary data and Secondary data.
- **Primary Data:** The data that have been initially collected and have not undergone any statistical treatment are called primary data.
- **Source of Primary Data:**
  Direct personal investigation.
  Indirect investigation or interviews.
  Collection through questionnaires.
  Collection through local sources.
  Through internet.
  Experimental research.
- **Secondary Data:** The data which has undergone any statistical treatment at least once is called primary data.
- **Source of Secondary Data:**
  - **Official:** Using the publications of statistical divisions, ministry of finance, the federal and provincial bureau of statistics, ministries of food, agriculture and industry etc.
  - **Semi Official:** State Bank of Pakistan, Railway Board, Central Cotton Committee, Board of Economic Inquiry, District Councils, Municipalities etc.
  - Publications of Trade Association, Chamber of Commerce etc.
  - Technical and Trade Journals and newspapers.
  - Research organizations such as universities and other institutions.

Presentation of Data

The device of gathering data often results in a massive volume of statistical data which are in the form of individual measurement of counts. These are as follows;

- **Classification:** The process of dividing a set of observations or objects into classes or groups. It is the sorting of data into homogeneous classes or groups according to their being alike or not.
- **Tabulation:** A systematic presentation of data classified under suitable heads and subheads and placed in columns and rows. It is an orderly arrangement of data in columns and rows.
- **Graphical Display:** The visual display of statistical data in the form of point lines, areas and other geometrical forms and symbols is in the most general term called graphical display. Such graphical representation divided into **graphs** and **diagrams**.

Data handling

Data handling is the process of securing the research data is gathered, archived or disposed of in a protected and safe way during and after the completion of the analysis process. Data handling means collecting the set of data and presenting in a different form.

Data Handling Steps

The steps involved in the data handling process are as follows:

- Problem Identification
- Data Collection
- Data Presentation
- Graphical Representation
- Data Analysis
- Conclusion
  From the analysis of the data, we can derive the solution to our problem statement. The data can be usually represented in any one of the following ways. They are:
  Bar Graph, Line Graphs, Histograms, Stem and Leaf Plot, Dot Plots, Frequency Distribution, Cumulative Tables and Graphs

Tabulation
An orderly arrangement of data in columns and rows.

Graphs of Data
As the old saying goes, a picture is worth a thousand words. Data summaries can come in pictures or graphs. Here are some of the typical types of graphs to display distributions. They can give us a quick overview of the big picture and the characteristics of the data.

Histogram / Frequency Histogram
A **histogram** is a bar graph where the data is represented in equal intervals. A Frequency Histogram is a graph that displays the classes on the horizontal axis and the frequencies on the vertical axis. It consists of vertical bars, whose height is equal to the frequency of the class (interval). The bars are drawn next to each other (without gaps), since they encompass the range of the data in numerical order. The left side of each bar starts at the lower limit of the class interval. The right side goes up to the lower limit of the next interval. A Histogram is only for quantitative data, not qualitative.

Relative Frequency Histogram
A Relative Frequency Histogram is the same as a frequency histogram, except it uses relative frequencies for the vertical axis and the bar heights.

Cumulative Frequency Polygon or Ogive
It is graphical representation of frequency distribution taking upper class boundary along with x – axis and cumulative frequency along y – axis.

Frequency Polygon
It is graphical representation of frequency distribution taking mid point along x – axis and frequency along y – axis.

Relative Frequency
It is a frequency which is derived by dividing frequency of any particular class by total frequency.

Historiogram
A Historiogram is a graph showing changes in the values of one period of time to the next is known as graph of a time series or histogram. Graph of frequency distribution is called **Histogram** and the graph of time series is called **Historiogram**.

Remember
One advantage of using a relative frequency distribution instead of a grouped frequency distribution is that there is a direct correspondence between the percent values of the relative frequency distribution and probabilities. For instance, in the relative frequency distribution in Table 8.5, the percent of the data that lie between 35 and 40 seconds is 14.9%. Thus, if a subscriber is chosen at random, the probability that the subscriber will require at least 35 seconds but less than 40 seconds to download the music file is 0.149.

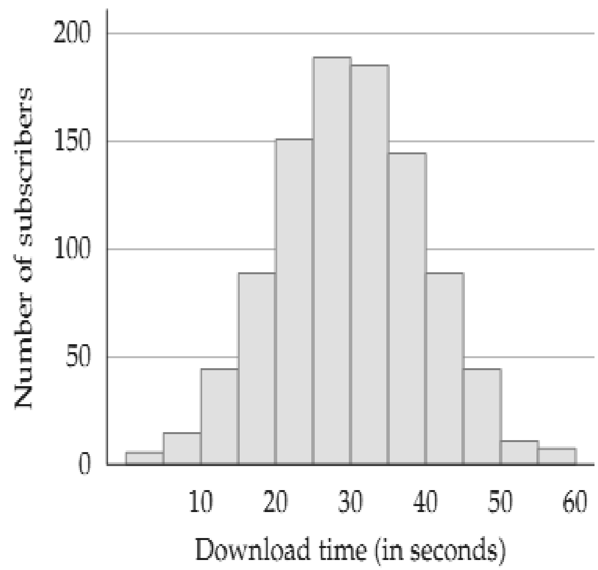
4. An Internet service provider (ISP) has installed new computers. To estimate the new download times its subscribers will experience, the ISP surveyed 1000 of its subscribers to determine the time required for each subscriber to download a particular file from the Internet site music.net. Summarized the survey by frequency distribution and histogram.

Solution

The results of that survey are summarized in Table.

Download Time (in seconds)	Number of Subscribers
0–5	6
5–10	17
10–15	43
15–20	92
20–25	151
25–30	192
30–35	190
35–40	149
40–45	90
45–50	45
50–55	15
55–60	10

A Histogram for the Frequency Distribution



A Grouped Frequency Distribution with 12 Classes

Table 1 is called a **grouped frequency distribution**. It shows how often (frequently) certain events occurred. Each interval, 0–5, 5–10, and so on, is called a **class**. This distribution has 12 classes. For the 10–15 class, 10 is the **lower class boundary** and 15 is the **upper class boundary**. Any data value that lies on a common boundary is assigned to the higher class. <u>The graph of a frequency distribution is called a **histogram**</u>. A histogram provides a pictorial view of how the data are distributed. In Figure, the height of each bar of the histogram indicates how many subscribers experienced the download times shown by the class at the base of the bar.

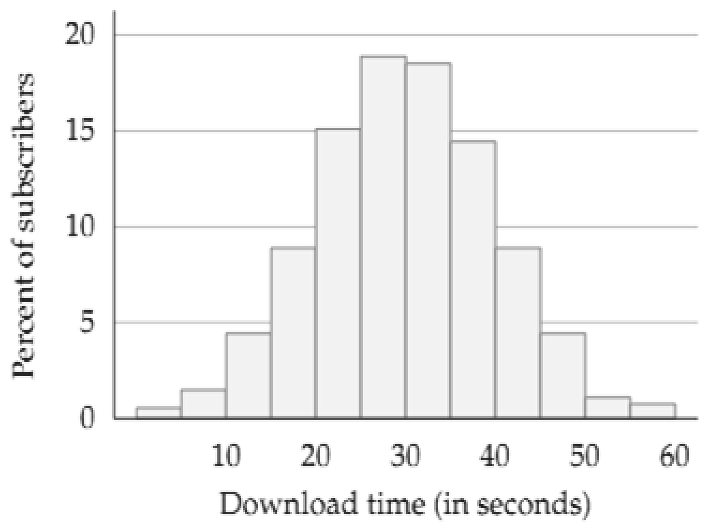
5. An Internet service provider (ISP) has installed new computers. To estimate the new download times its subscribers will experience, the ISP surveyed 1000 of its subscribers to determine the time required for each subscriber to download a particular file from the Internet site music.net. Summarized the survey by relative frequency distribution and relative frequency histogram.

Solution

The results of that survey are summarized in Table.

Download Time (in seconds)	Percent of Subscribers
0–5	0.6
5–10	1.7
10–15	4.3
15–20	9.2
20–25	15.1
25–30	19.2
30–35	19.0
35–40	14.9
40–45	9.0
45–50	4.5
50–55	1.5
55–60	1.0

A Relative Frequency Histogram



A Relative Frequency Distribution

Examine the distribution in Table. It shows the percent of subscribers that are in each class, as opposed to the frequency distribution in Table on the preceding page, which shows the number of customers in each class. The type of frequency distribution that lists the percent of data in each class is called a **relative frequency distribution**. The **relative frequency histogram** in Figure was drawn by using the data in the relative frequency distribution. It shows the percent of subscribers along its vertical axis.

6. Use the relative frequency distribution in Table to determine

a. the percent of subscribers who required at least 25 seconds to download the file.

b. the probability that a subscriber chosen at random will require at least 5 but less than 20 seconds to download the file.

Download Time (in seconds)	Percent of Subscribers	
0–5	0.6	
5–10	1.7	Sum is 15.2%
10–15	4.3	
15–20	9.2	
20–25	15.1	
25–30	19.2	Sum is 69.1%
30–35	19.0	
35–40	14.9	
40–45	9.0	
45–50	4.5	
50–55	1.5	
55–60	1.0	

Solution

a. The percent of data in all the classes with a lower boundary of 25 seconds or more is the sum of the percents for all of the classes highlighted in red in the distribution below. Thus the percent of subscribers who required at least 25 seconds to download the file is 69.1%. See Table.

b. The percent of data in all the classes with a lower boundary of at least 5 seconds and an upper boundary of 20 seconds or less is the sum of the percents in all of the classes highlighted in blue in the distribution above. Thus the percent of subscribers who required at least 5 but less than 20 seconds to download the file is 15.2%. The probability that a subscriber chosen at random will require at least 5 but less than 20 seconds to download the file is 0.152. See Table.

7. Use the relative frequency distribution in Table to determine

a. the percent of subscribers who required less than 25 seconds to download the file.

b. the probability that a subscriber chosen at random will require at least 10 seconds but less than 30 seconds to download the file.

Download Time (in seconds)	Percent of Subscribers
0–5	0.6
5–10	1.7
10–15	4.3
15–20	9.2
20–25	15.1
25–30	19.2
30–35	19.0
35–40	14.9
40–45	9.0
45–50	4.5
50–55	1.5
55–60	1.0

Solution

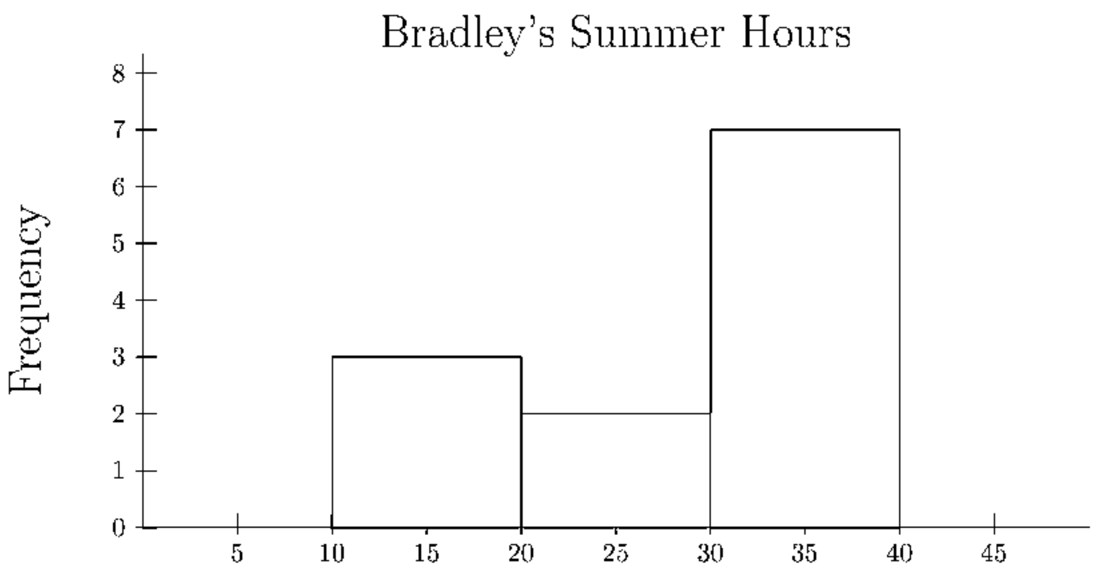
a. The percent of data in all classes with an upper bound of 25 seconds or less is the sum of the percents for the first five classes in Table. Thus the percent of subscribers who required less than 25 seconds to download the file is 30.9%.

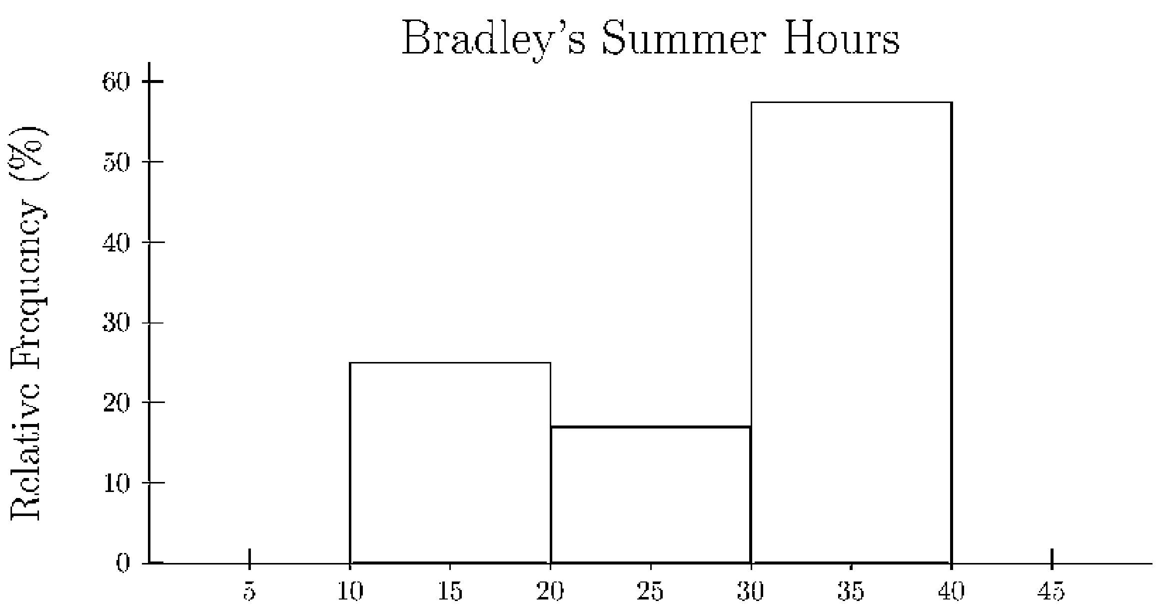
b. The percent of data in all the classes with a lower bound of at least 10 seconds and an upper bound of 30 seconds or less is the sum of the percents in the third through sixth classes in Table. Thus the percent of subscribers who required from 10 to 30 seconds to download the file is 47.8%. The probability that a subscriber chosen at random will require from 10 to 30 seconds to download the file is 0.478.

8. For Bradley's weekly hours at a summer job: 25, 32, 36, 32, 18, 28, 30, 36, 12, 16, 35, 36. Find the frequency and relative frequency histograms hours he worked in a week.

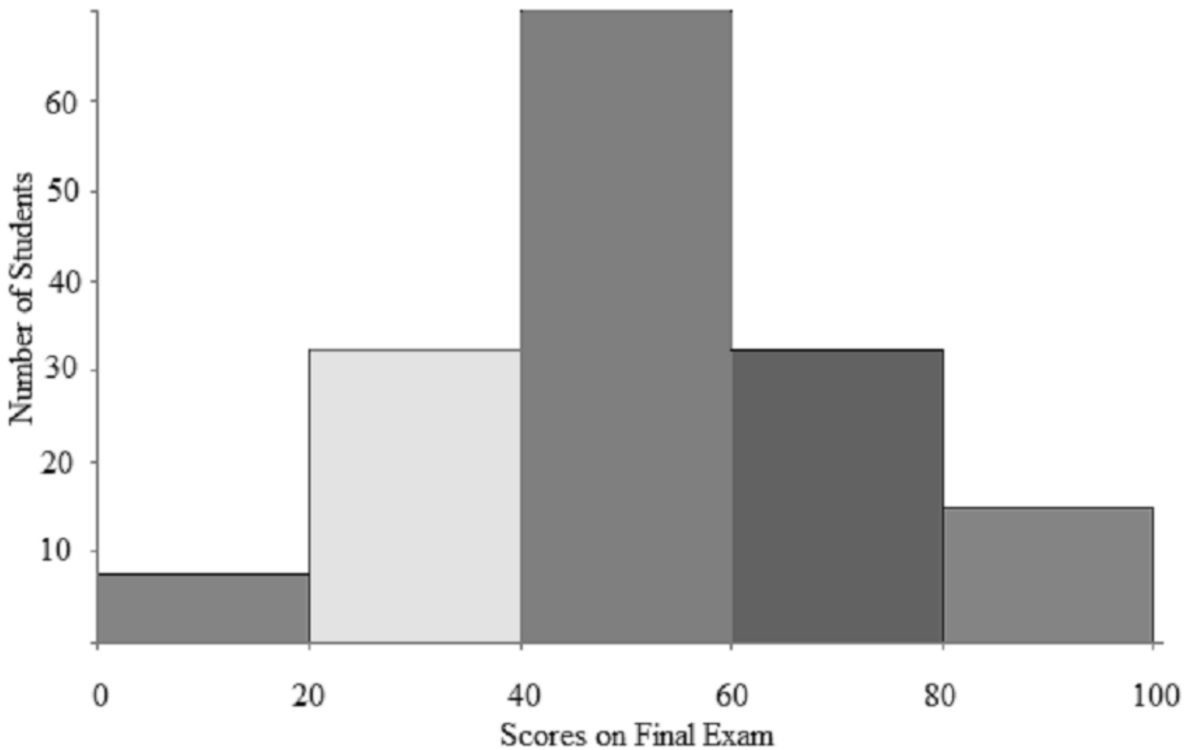
Solution

The frequency and relative frequency histograms for Bradley's summer job data are shown below.



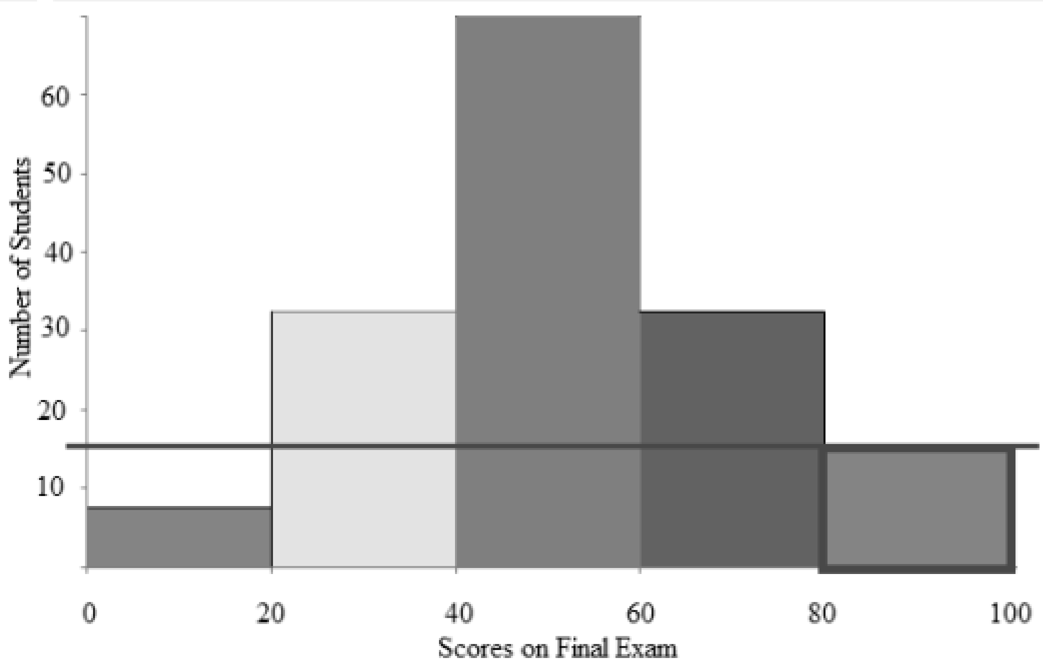


9. In the following graph, how many students got the highest score?



Solution

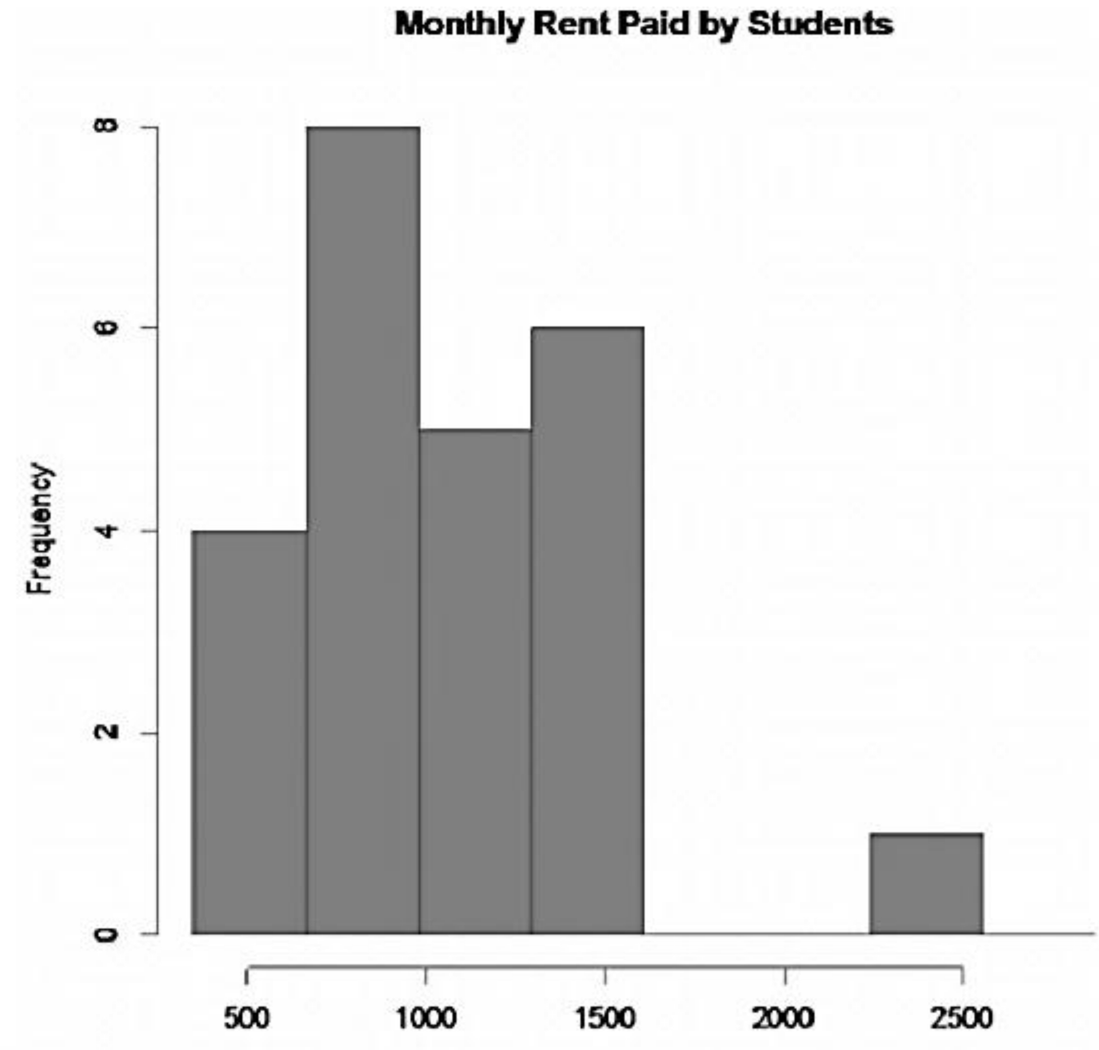
The highest score ranges from 80 to 100. Based on the histogram chart, there were about 15 students. Therefore, the answer is 15.



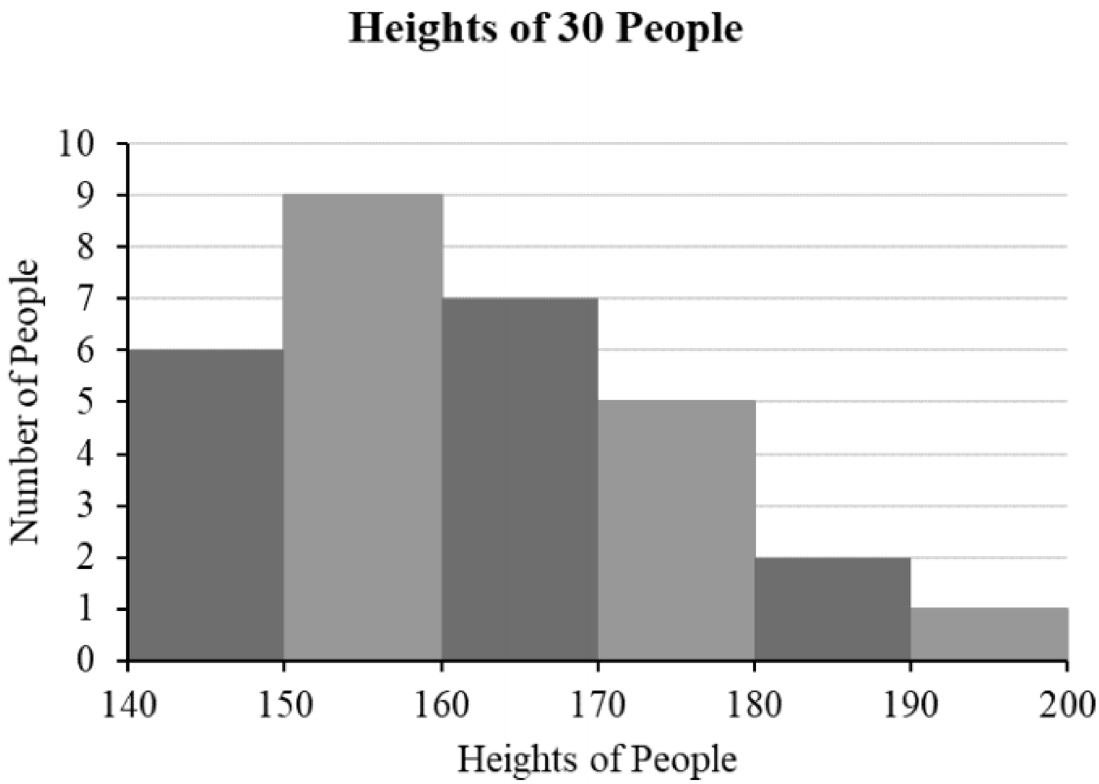
10. Draw a histogram for the distribution for the following data

1500	1350	350	1200	850	900
1500	1150	1500	900	1400	1100
1250	600	610	960	890	1325
900	800	2550	495	1200	690

Solution



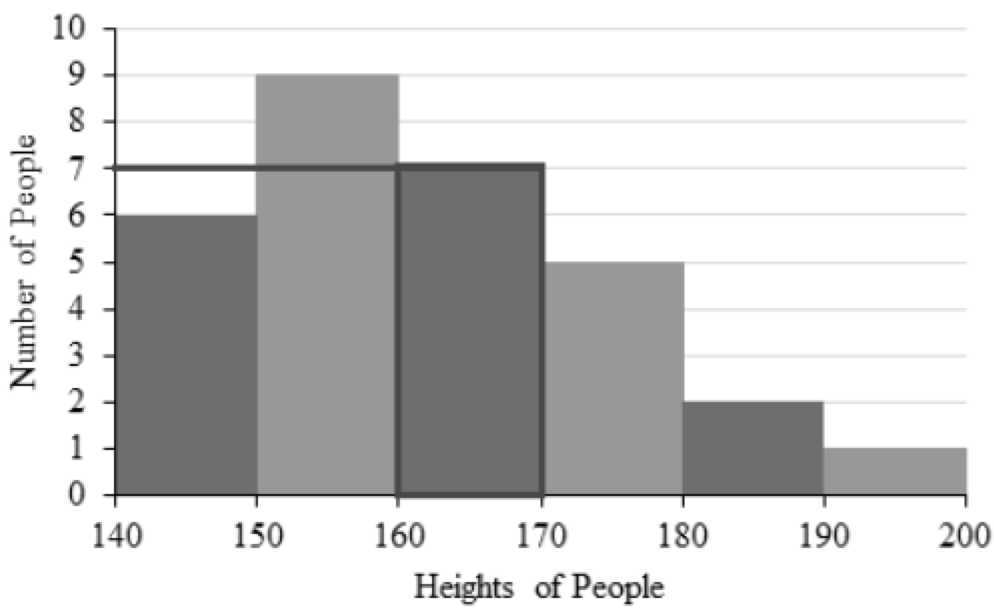
11. The histogram below shows the height (in cm) distribution of 30 people. How many people have heights between 160 and 170 cm? How many people have heights less than 160 cm?



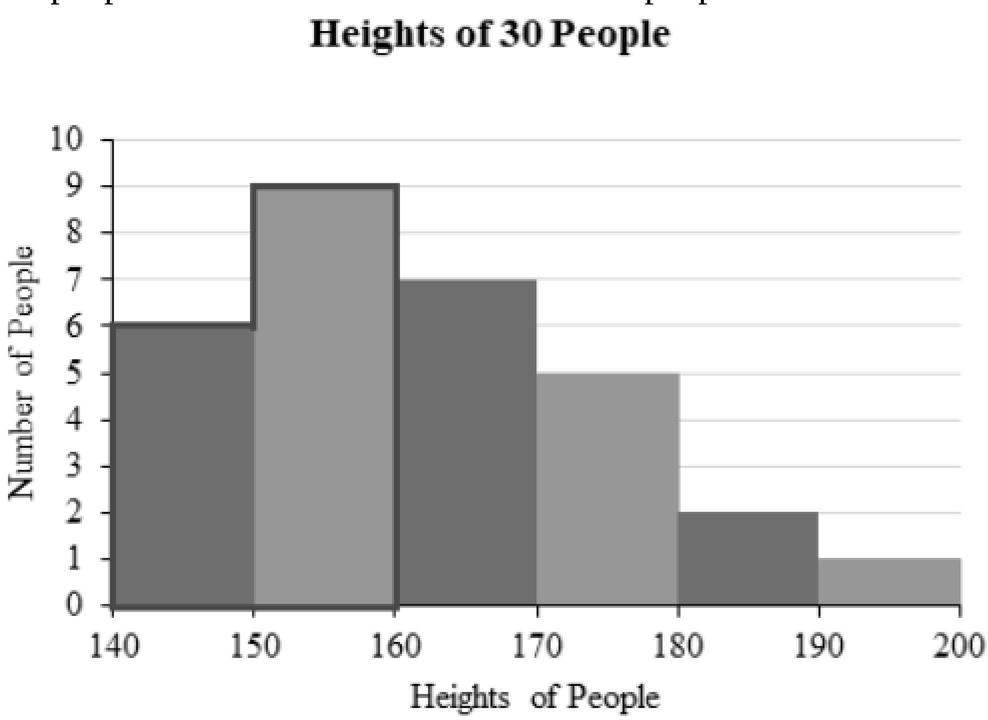
Solution

Solution for heights between 160 and 170 cm:

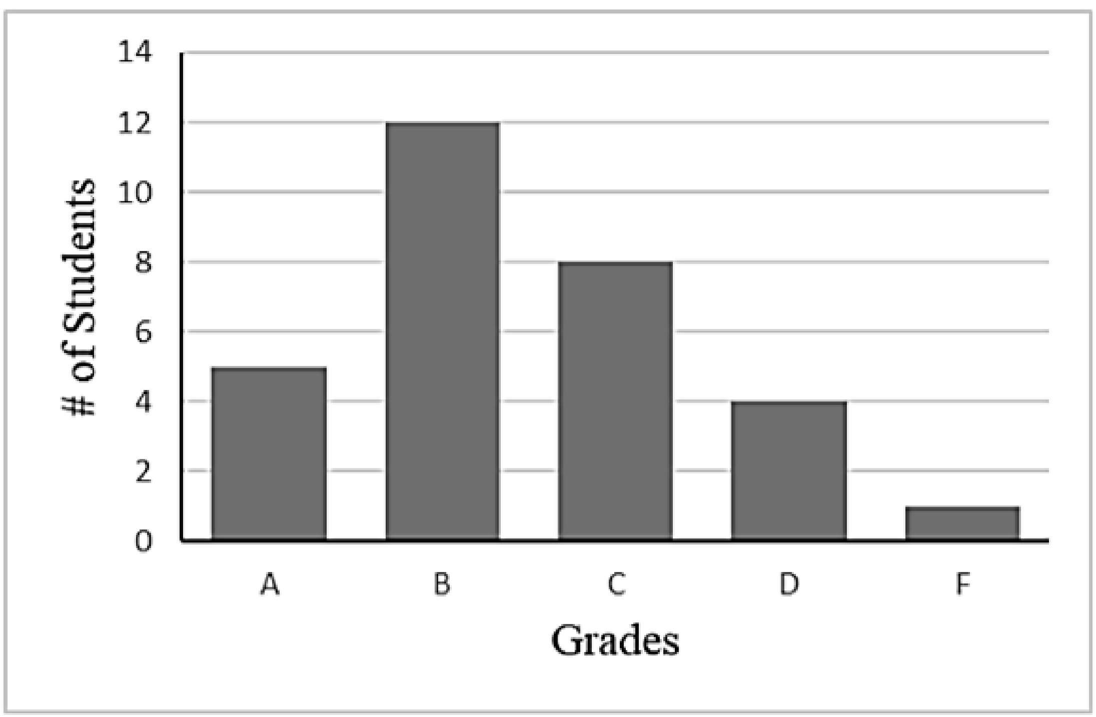
We need to look at the third bar because it ranges from 160 cm to 170 cm. The bar indicates that there are 7 people. Therefore, the answer is 7.



Solution of number of people less than 160 cm:
For the people less than 160 cm height, we have to look at the bars from two categories – 140 cm to 150 cm and 150 cm to 160 cm. Therefore,
140-150: 6 people
150-160: 9 people
Total people less than 160 cm is 6 + 9 = 15 people

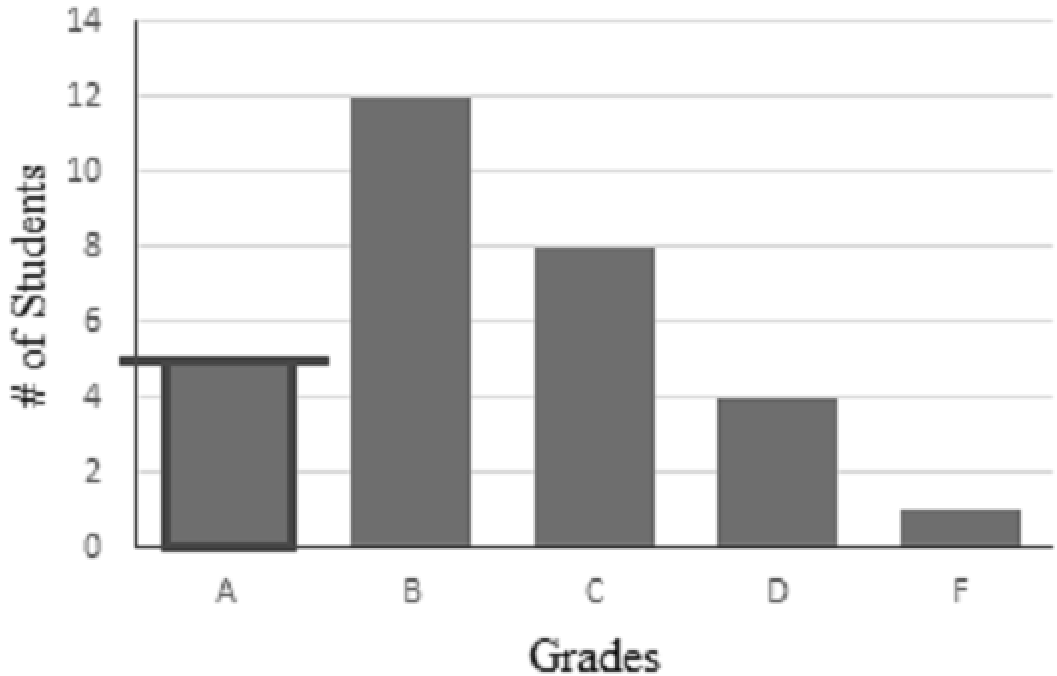


12. How many students got a grade of 'A' based on the following chart.

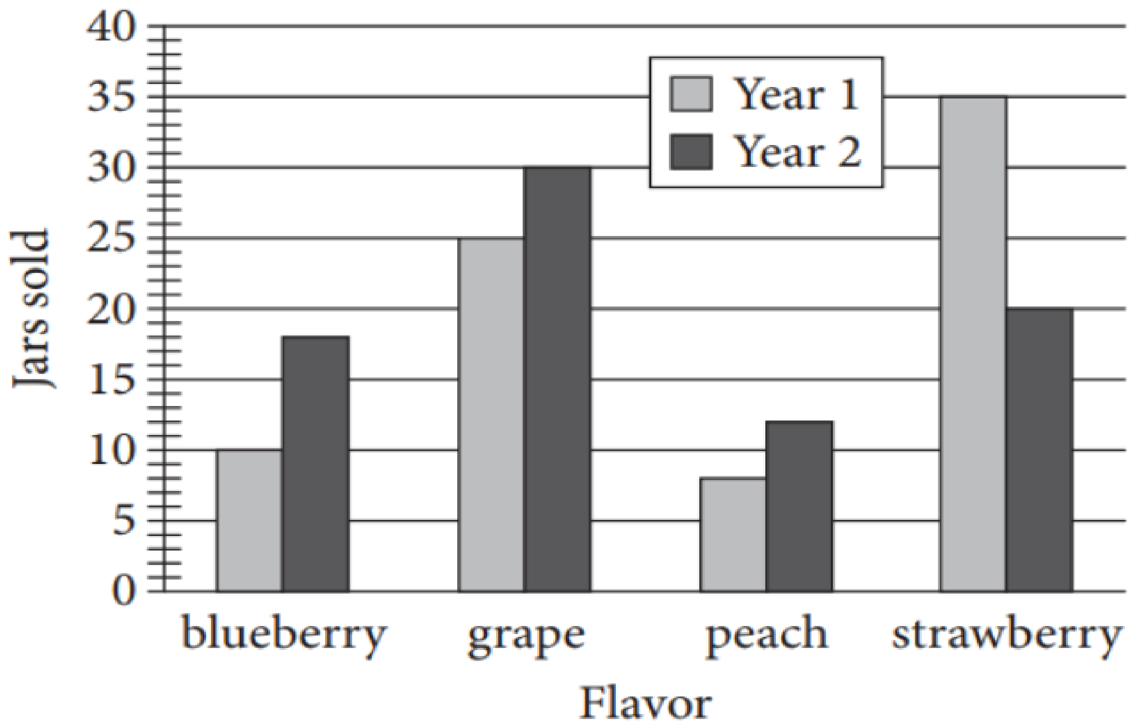


Solution

When we observe the bar A, the bar lines up with 5. Therefore, the answer is 5 students.



13. Logan sells four different flavors of jam at an annual farmers market. The graph below shows the number of jars of each type of jam they sold at the market during the first two years. Which flavor of jam had the greatest increase in number of jars sold from Year 1 to Year 2?



Solution

For this problem, we have to calculate the difference between year 1 and year 2 of all the jars.

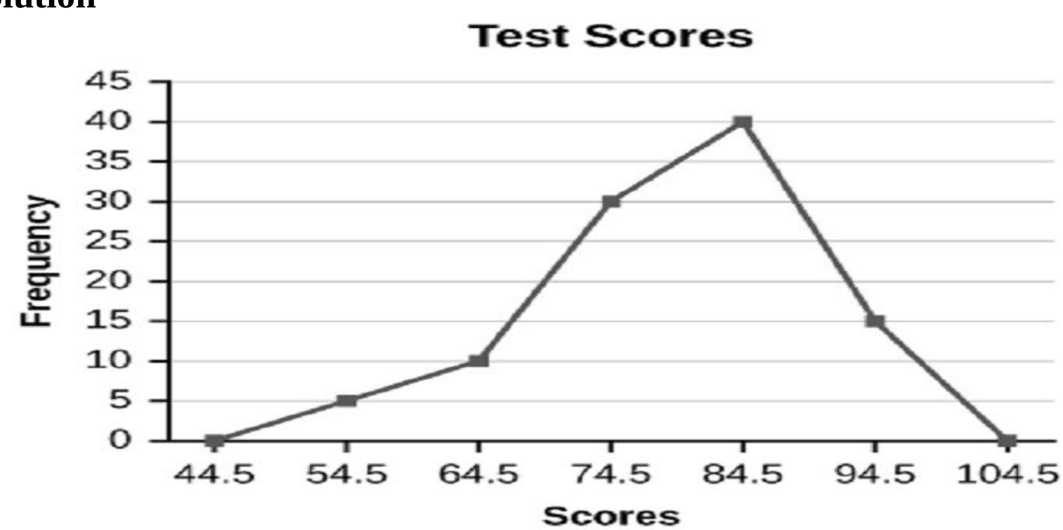
Jars	Year 1	Year 2	Difference (Year 2 – Year 1)
Blueberry	10	18	8
Grape	25	30	5
Peach	8	12	4
Strawberry	35	20	-15

Although strawberry has the highest difference, it is not the answer because the number of jars decreased. Blueberry jars increased from 10 to 18 with an increase of 8 jars. Therefore, blueberry is the answer.

14. Construct a frequency polygon from the following frequency table.

Frequency Distribution for Calculus Final Test Scores			
Lower Bound	**Upper Bound**	**Frequency**	**Cumulative Frequency**
49.5	59.5	5	5
59.5	69.5	10	15
69.5	79.5	30	45
79.5	89.5	40	85
89.5	99.5	15	100

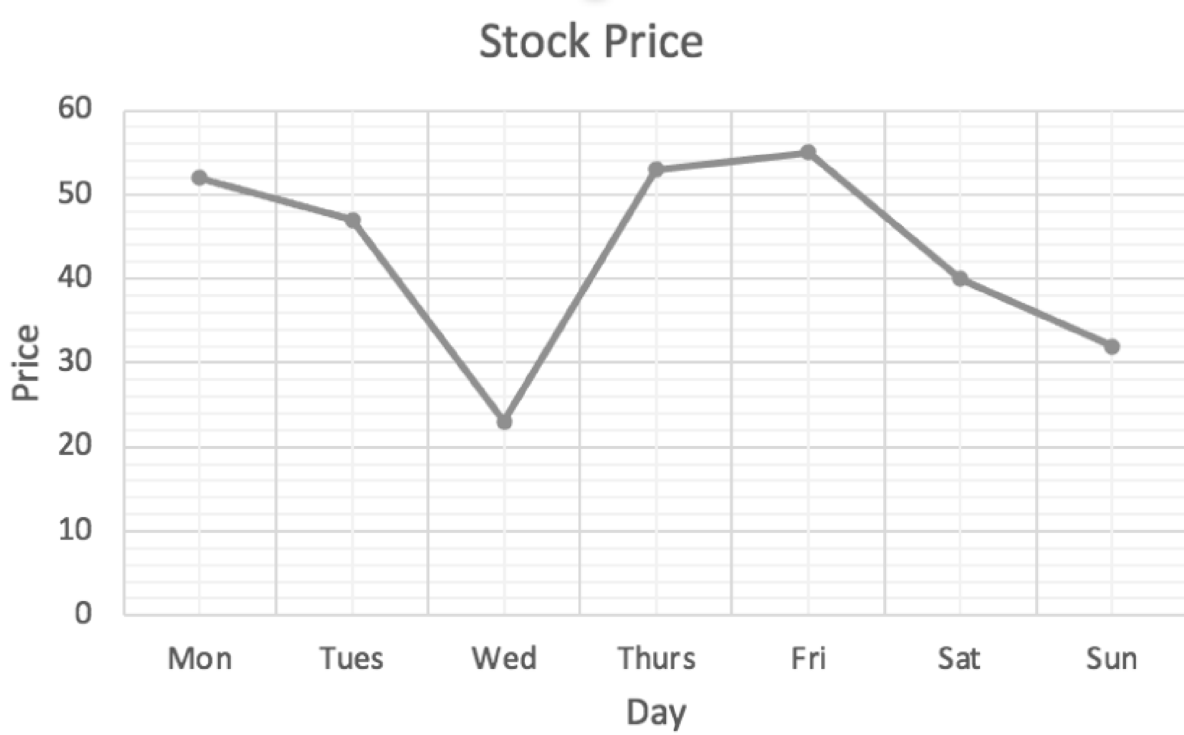
Solution



Time Series Graphs/ Line Graphs

Line Graphs show trends over a period of time. Time is located on the horizontal axis and amounts will be located on the vertical axis. Information will be organized into ordered pairs. To draw the graph, you plot these points and then connect them with straight lines. In line graph, the x- axis (horizontal axis) consists of data values and the y-axis (vertical axis) consists of frequency points. The frequency points are connected using line segments. The graph of time series is called **Historiogram**.

15. Veronica is a stock trader. She followed the value of a stock and recorded the following graph for this past week. Use it to answer the following questions.



a) What was the value of the stock on Tuesday?

b) What day of the week was the stock price equal to $40?

c) Between which two consecutive days was the biggest change in stock prices?

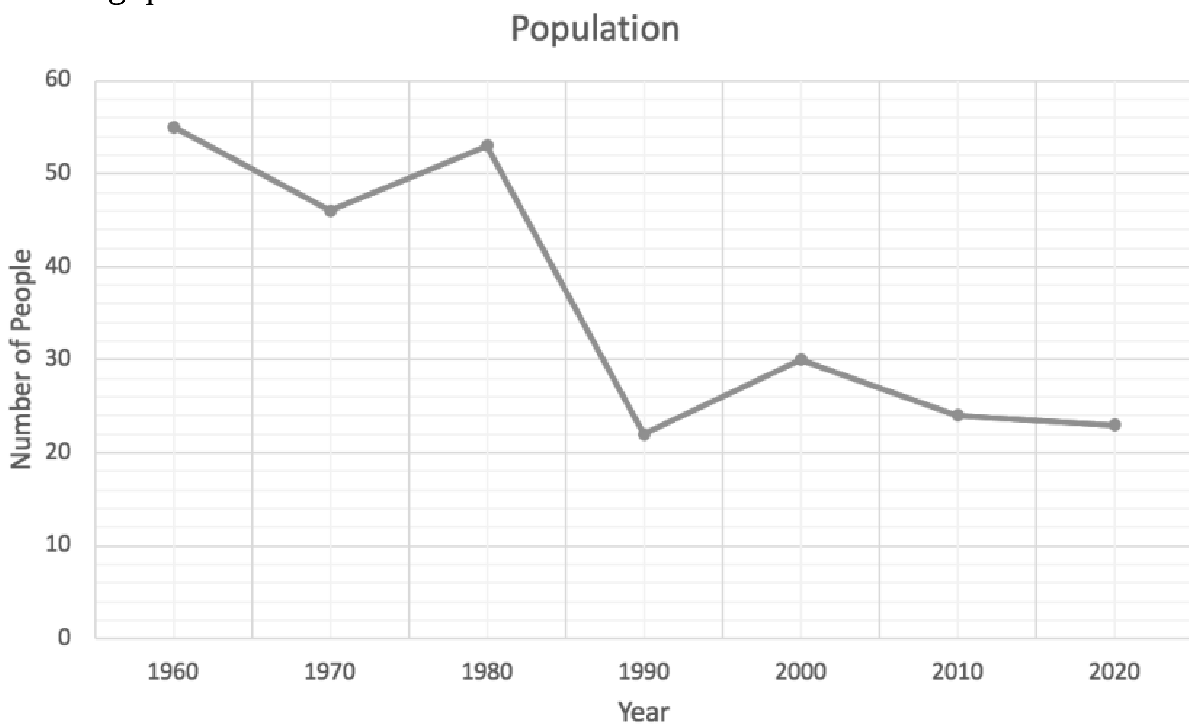
Solution

a) When reading the graph the horizontal axis is the day and the vertical axis tells us the value of the stock that day. For Tuesday the value of the stock was $47.

b) Start by reading the vertical axis and find where the stock price is $40 and then find the corresponding day on the horizontal axis which is Saturday.

c) We need to look at the graph and find two consecutive points who have the biggest difference between stock prices. Looking at the graph this occurs between Wednesday and Thursday.

16. Edge Hill is the smallest incorporated city in the state of Georgia. The line graph below shows its population every ten years starting in 1960. Use this graph to answer the following questions.



a) What was the population in 2000?
b) What year was the population equal to 46 people?
c) At approximately what rate did the population decrease from the year 1980 to 1990?

Solution

a) When reading this graph the horizontal axis represents the year and the vertical axis represents the population size. To answer this problem look for the year 2000 on the horizontal axis and then read the population size from the vertical axis which is 30 people.
b) For this problem look for the point on the line graph where the population is 46 people. This happens for the year 1970.
c) To answer this question we need to compute a rate. Start by identifying two points from the graph at the years 1980 and 1990: (1980, 53) and (1990, 22). Now compute the rate. To do this we will use the slope formula.

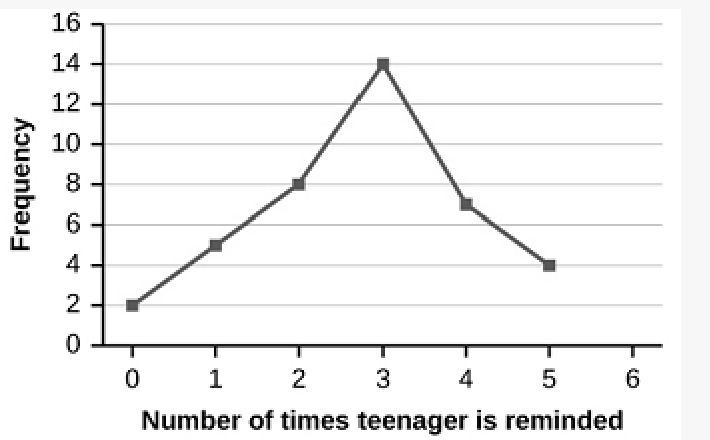
$$\text{Rate} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{22 - 53}{1990 - 1980} = \frac{-31}{10} = -3.1$$

What this tells us is that on average the population decreased by approximately 3 people each year between 1980 and 1990.

17. In a survey, 40 mothers were asked how many times per week a teenager must be reminded to do his or her chores. The results are shown in Table and in Figure.

Number of times teenager is reminded	Frequency
0	2
1	5
2	8
3	14
4	7
5	4

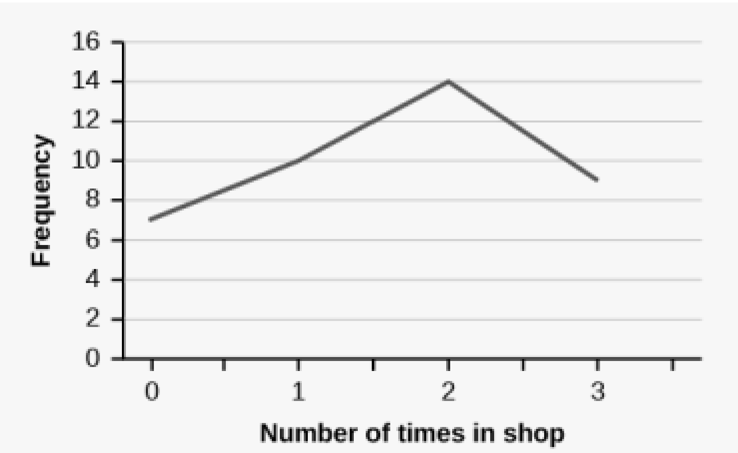
Solution



18. In a survey, 40 people were asked how many times per year they had their car in the shop for repairs. The results are shown in Table. Construct a line graph.

Number of times in shop	Frequency
0	7
1	10
2	14
3	9

Solution

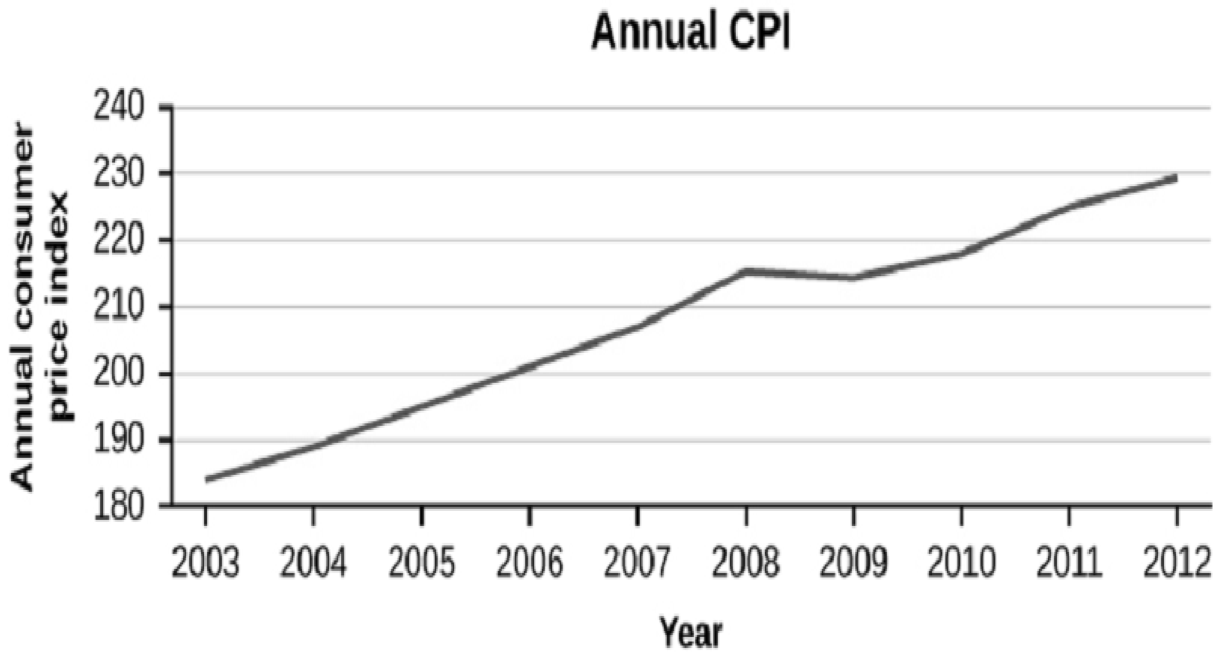


19. The following data shows the Annual Consumer Price Index, each month, for ten years. Construct a time series graph for the Annual Consumer Price Index data only.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul
2003	181.7	183.1	184.2	183.8	183.5	183.7	183.9
2004	185.2	186.2	187.4	188.0	189.1	189.7	189.4
2005	190.7	191.8	193.3	194.6	194.4	194.5	195.4
2006	198.3	198.7	199.8	201.5	202.5	202.9	203.5
2007	202.416	203.499	205.352	206.686	207.949	208.352	208.299
2008	211.080	211.693	213.528	214.823	216.632	218.815	219.964
2009	211.143	212.193	212.709	213.240	213.856	215.693	215.351
2010	216.687	216.741	217.631	218.009	218.178	217.965	218.011
2011	220.223	221.309	223.467	224.906	225.964	225.722	225.922
2012	226.665	227.663	229.392	230.085	229.815	229.478	229.104

Year	Aug	Sep	Oct	Nov	Dec	Annual
2003	184.6	185.2	185.0	184.5	184.3	184.0
2004	189.5	189.9	190.9	191.0	190.3	188.9
2005	196.4	198.8	199.2	197.6	196.8	195.3
2006	203.9	202.9	201.8	201.5	201.8	201.6
2007	207.917	208.490	208.936	210.177	210.036	207.342
2008	219.086	218.783	216.573	212.425	210.228	215.303
2009	215.834	215.969	216.177	216.330	215.949	214.537
2010	218.312	218.439	218.711	218.803	219.179	218.056
2011	226.545	226.889	226.421	226.230	225.672	224.939
2012	230.379	231.407	231.317	230.221	229.601	229.594

Solution



Uses of a Time Series Graph

Time series graphs are important tools in various applications of statistics. When recording values of the same variable over an extended period of time, sometimes it is difficult to discern any trend or pattern. However, once the same data points are displayed graphically, some features jump out. Time series graphs make trends easy to spot.

Stem-and-Leaf Diagram

A clear disadvantage of using a frequency table is that the identity of individual observations is lost in grouping process. To overcome this draw back, **John Turkey** introduced a technique known as *the Stem and Leaf Display*. This technique offers a quick and novel way for simultaneously sorting and simplifying data sets, where each number in data set is divided into two parts, a stem and a leaf. A **stem** is the leading digit(s) of each number and used in sorting, while a **leaf** is the rest of the number or the trailing digit(s) and shown in display. A vertical line separates the leaf (or leaves) from the stem. For example the number 243 could be split two ways;

Leading Digit	Trailing Digit
2 Stem	43 Leaf

Or

Leading Digit	Trailing Digit
24 Stem	3 Leaf

All possible stems are arranged in order from the smallest to the largest and placed on the left hand side of the line.

The stem and leaf display is a useful step for listing the data in an array, leaves are associated with the stem to know the numbers. The stem and leaf table provide a useful description of the data set and can easily be converted to a frequency table. It is a common practice to arrange the trailing digits in each case from smallest to highest.

Steps in the Construction of a Stem-and-Leaf Diagram

1. Determine the stems and list them in a column from smallest to largest or largest to smallest.
2. List the remaining digit of each stem as a leaf to the right of the stem.
3. Include a legend that explains the meaning of the stems and the leaves. Include a title for the diagram.

20. Construct a stem-and-leaf display from the data and list the data in an array.

48,31,54,37,18,64,61,43,40,71,51,12,52,65,53,
42,39,62,74,48,29,67,30,49,68,35,57,26,27,58

Solution

A scan of the data indicates that the observation range is 12 to 74. We use the first observation is 48, which has a stem of 4 and a leaf of 8, the second a stem of 3 and a leaf of 1, etc. Then our required stem and leaf display is as follows;

Stem (Leading Digit)	**Leaf (Trailing Digit)**
1	8 2
2	9 6 7
3	1 7 9 0 5
4	8 3 0 2 8 9
5	4 1 2 3 7 8
6	4 1 5 2 7 8
7	1 4

21. Construct stem-and-leaf diagram for the following history test scores:

65, 72, 96, 86, 43, 61, 75, 86, 49, 68, 98, 74, 84, 78, 85, 75, 86, 73

Solution

In the stem-and-leaf diagram on the following page, we have organized the history test scores by placing all of the scores that are in the 40s in the top row, the scores that are in the 50s in the second row, the scores that are in the 60s in the third row, and so on. The tens digits of the scores have been placed to the left of the vertical line. In this diagram they are referred to as stems. The ones digits of the test scores have been placed in the proper row to the right of the vertical line. In this diagram they are the leaves. It is now easy to make observations about the distribution of the scores. Only two of the scores are in the 90s. Six of the scores are in the 70s, and none of the scores are in the 50s. The lowest score is 43 and the highest is 98.

Stems	Leaves
4	3 9
5	
6	1 5 8
7	2 3 4 5 5 8
8	4 5 6 6 6
9	6 8

Legend: 8|6 represents 86

Remark
The choice of how many leading digits to use as the stem will depend on the particular data set.

22. Construct stem-and-leaf diagram for the following data set, in which a travel agent has recorded the amounts spent by customers for a cruise.

\$3600	\$4700	\$7200	\$2100	\$5700	\$4400	\$9400
\$6200	\$5900	\$2100	\$4100	\$5200	\$7300	\$6200
\$3800	\$4900	\$5400	\$5400	\$3100	\$3100	\$4500
\$4500	\$2900	\$3700	\$3700	\$4800	\$4800	\$2400

Solution
One method of choosing the stems is to let each thousands digit be a stem and each hundreds digit be a leaf. If the stems and leaves are assigned in this manner, then the notation 2|1 with a stem of 2 and a leaf of 1, represents a cost of \$2100, and 5|4 represents a cost of \$5400. A stem-and-leaf diagram can now be constructed by writing all of the stems in a column from smallest to largest to the left of a vertical line, and writing the corresponding leaves to the right of the line.

Stems	Leaves
2	1 1 4 9
3	1 1 6 7 7 8
4	1 4 5 5 7 8 8 9
5	2 4 4 7 9
6	2 2
7	2 3
8	
9	4

Legend:
7|3 represents \$7300

Remark

Sometimes two sets of data can be compared by using a back-to-back stem-and- leaf diagram, in which common stems are listed in the middle column of the diagram. Leaves from one data set are displayed to the right of the stems, and leaves from the other data set are displayed to the left.

23. Construct stem-and-leaf diagram for the following data set below shows the test scores for two classes that took the same test.

Scores in the range of 70s, 80s, 90s, 40s, 50s, and 60s.

Solution

The back-to-back stem-and-leaf diagram below shows the test scores for two classes that took the same test. It is easy to see that the 8 A.M. class did better on the test because it had more scores in the 80s and 90s and fewer scores in the 40s, 50s, and 60s. The number of scores in the 70s was the same for both classes.

Biology Test Scores

8 A.M. Class		10 A.M. Class
2	4	5 8
7	5	6 7 9 9
5 8	6	2 3 4 8
1 2 3 3 3 7 8	7	1 3 3 5 5 6 8
4 4 5 5 6 8 8 9	8	2 3 6 6 6
2 4 5 5 8	9	4 5

Legend: 3 | 7 represents 73

Legend: 8 | 2 represents 82

24. Draw Stem and Leaf graph for Susan Dean's spring pre-calculus class, scores for the first exam were as follows (smallest to largest):

33; 42; 49; 49; 53; 55; 55; 61; 63; 67; 68; 68; 69; 69; 72; 73; 74; 78; 80; 83; 88; 88; 88; 90; 92; 94; 94; 94; 94; 96; 100

Solution

Stem-and-Leaf Graph

Stem	Leaf
3	3
4	2 9 9
5	3 5 5
6	1 3 7 8 8 9 9
7	2 3 4 8
8	0 3 8 8 8
9	0 2 4 4 4 4 6
10	0

The stemplot shows that most scores fell in the 60s, 70s, 80s, and 90s. Eight out of the 31 scores or approximately 26% $\left(\frac{8}{31}\right)$ were in the 90s or 100, a fairly high number of As.

25. For the Park City basketball team, scores for the last 30 games were as follows (smallest to largest):

32; 32; 33; 34; 38; 40; 42; 42; 43; 44; 46; 47; 47; 48; 48; 48; 49; 50; 50; 51; 52; 52; 52; 53; 54; 56; 57; 57; 60; 61

Construct a stem plot for the data.

Solution

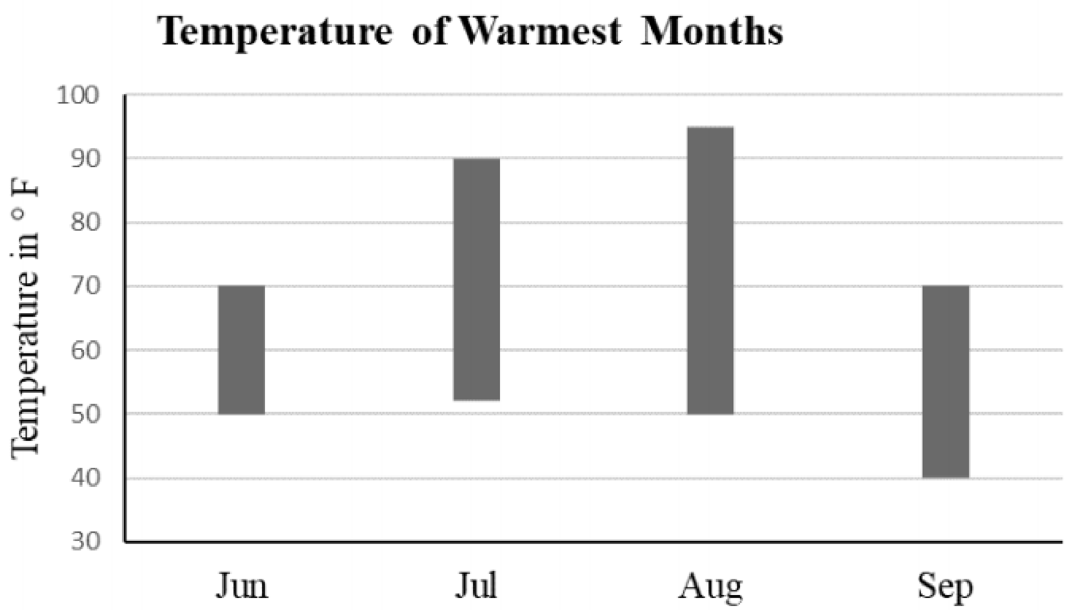
Stem	Leaf
3	2 2 3 4 8
4	0 2 2 3 4 6 7 7 8 8 8 9
5	0 0 1 2 2 2 3 4 6 7 7
6	0 1

The stemplot is a quick way to graph data and gives an exact picture of the data. You want to look for an overall pattern and any outliers. An **outlier** is an observation of data that does not fit the rest of the data. It is sometimes called an **extreme value**. When you graph an outlier, it will appear not to fit the pattern of the graph. Some outliers are due to mistakes (for example, writing down 50 instead of 500) while others may indicate that something unusual is happening.

Range Bar

A range bar graph represents a range of data for each independent variable.

26. In the following chart, which month is the warmest month?



Solution

In June, temperatures range from 50° F to 70° F.

In July, temperatures range from 52° F to 90° F.

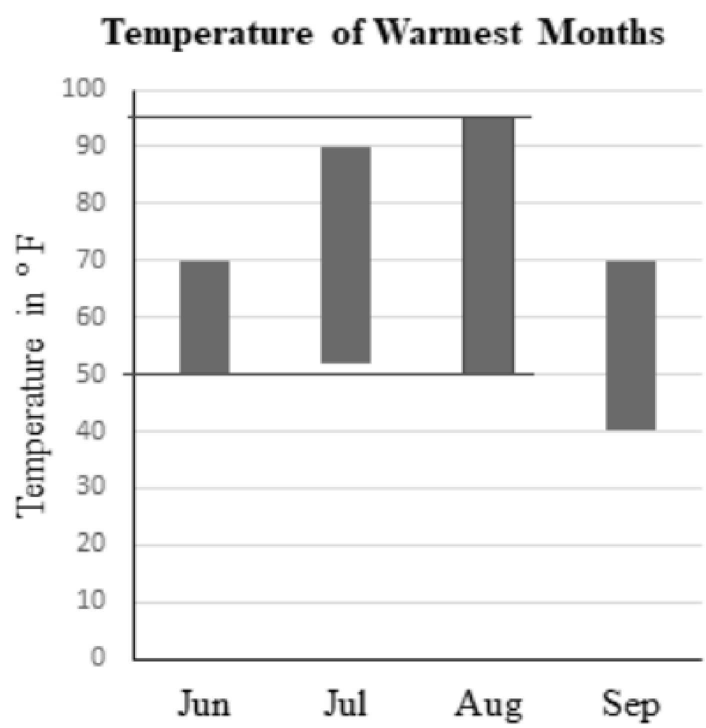
In August, temperatures range from 50° F to 95° F, 95 being the highest.

In September, temperatures range from 40° F to 70° F.

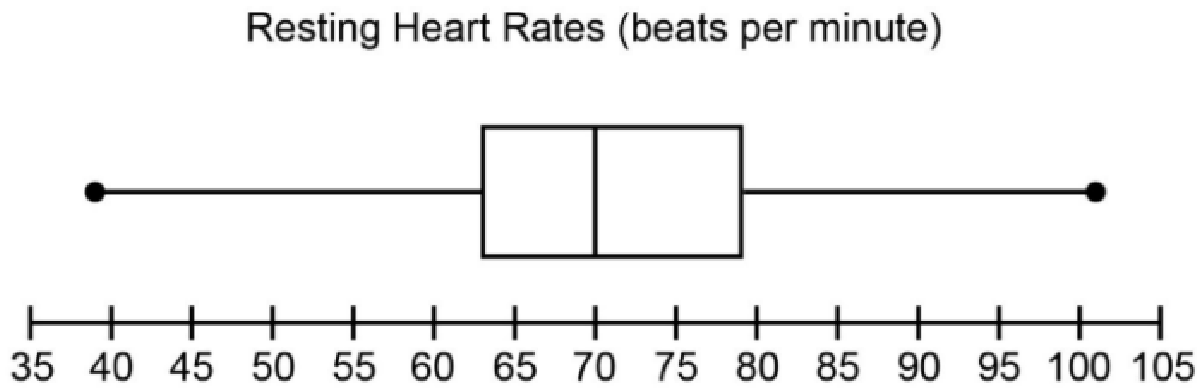
Therefore,

The warmest month is **August**.

The highest temperature is 95° F.

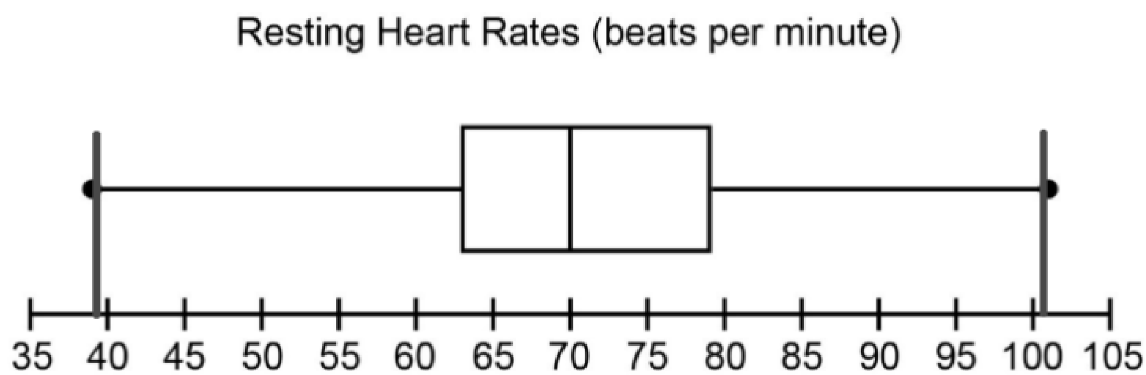


27. The box plot below summarizes the resting heart rates, in beats per minute, of the members at a gym. Which of the following could be the range of resting heart rates, in beats per minute?



Solution

The chart shows resting heart rates (beats/minute), and we are asked to calculate range of resting heart rates.



We know that

Range = Maximum – Minimum

Range = 101 – 139 = 62

Therefore, the answer is 62 beats per minute.

Chart

A chart is the diagrammatic representation of a spatial series where the data is split into different categories.

Different types of Charts

Histograms are for quantitative data. There is a similar graph for qualitative data (categories), called a **Bar Graph**. In a bar graph, the width of the bars is arbitrary and the bars are not connected. A Bar Graph is a graph that consists of bars for each category with the length/height of the bars specifying the frequency for each category. One axis will indicate the categories and the other axis will indicate the frequency. Bar graphs are often used for comparing different characteristics of items. Bar graphs can have horizontal or vertical bars. Can show frequency or relative frequency. Few types are as follows;

Bar Graph/ Bar Chart/ Simple Bar Chart

A simple bar chart is used when the data consists of a single component and also do not involve much variation.

Multiple Bar Charts

When we represent two or more sets of inter related data and also we want to compare different phenomena diagrammatically we use multiple bar chart.

Component Bar Charts

Component bar charts are used when the data consists of more than one homogeneous component, they represents the commutations of the various components of data.

Percentage Component Bar Charts

Percentage component bar charts are used if the comparison of relative values are concerned. These charts are drawn on percentage basis.

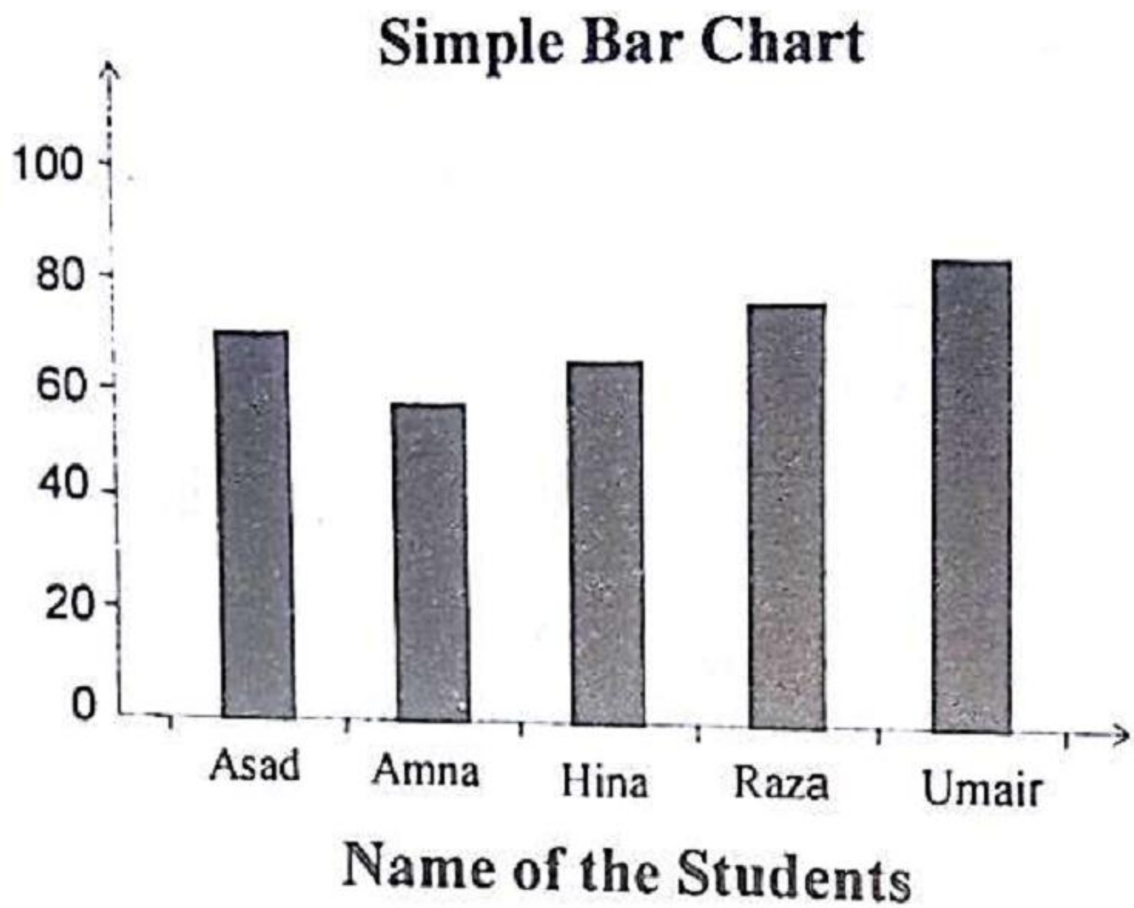
Pareto Chart

A **Pareto chart** is a specific type of bar graph where the classes are reordered so that the bars are in size order.

28. Draw the simple bar chart for the following

Name	Asad	Amna	Hina	Raza	Umair
Marks	70	57	65	75	82

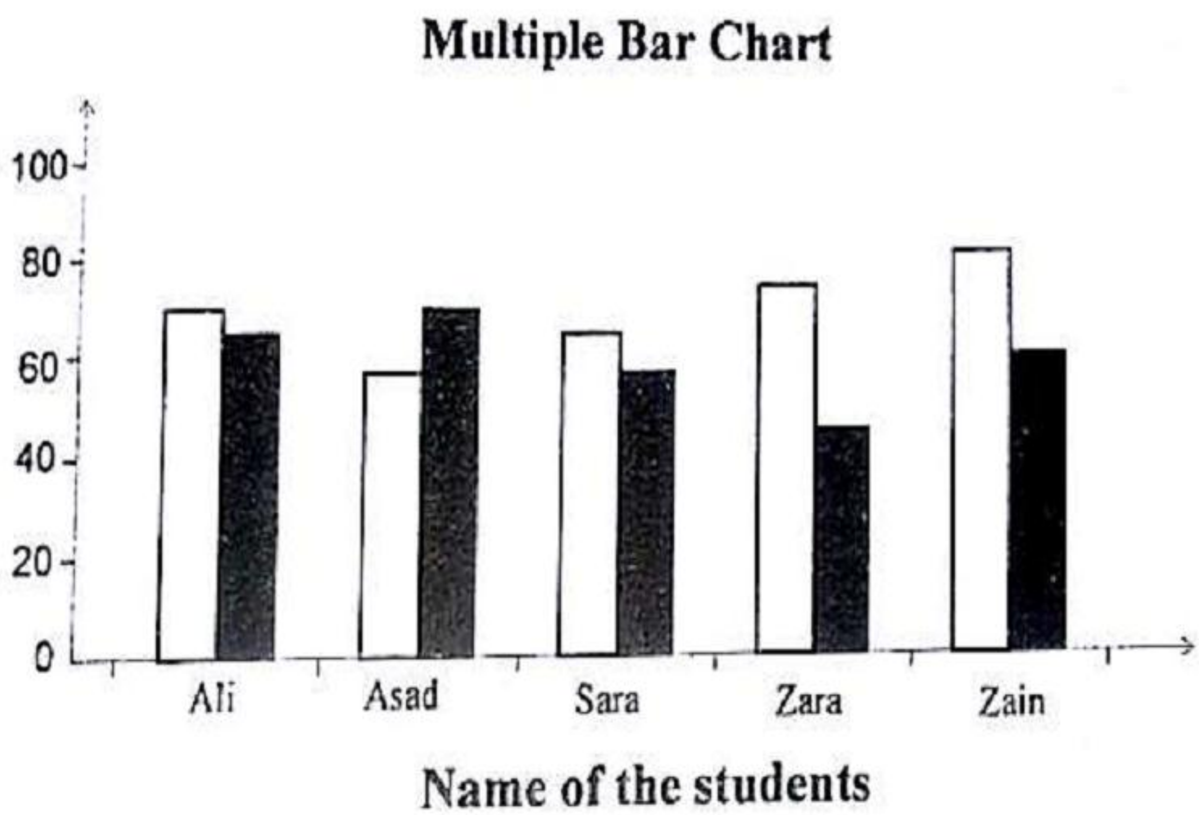
Solution



29. Draw the multiple bar chart for the following

Name	Asad	Amna	Hina	Raza	Umair
MBF	70	57	65	75	82
I to B	65	70	57	45	60

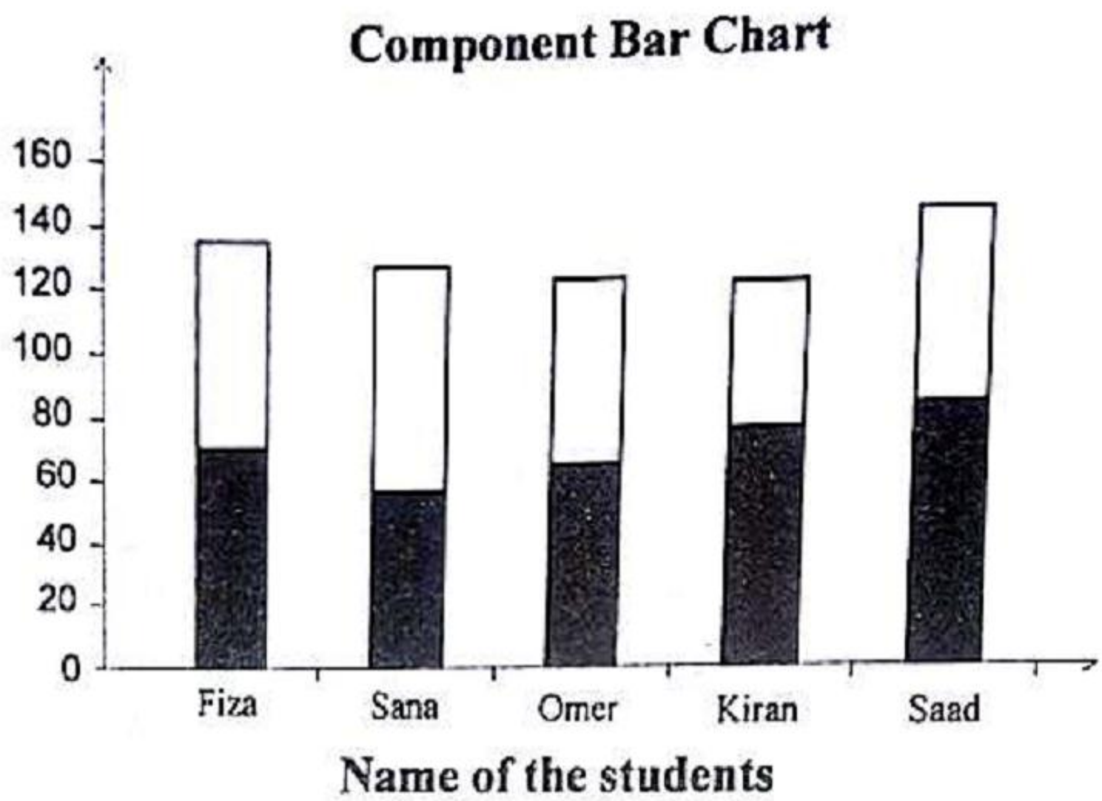
Solution



30. Draw the component bar chart for the following

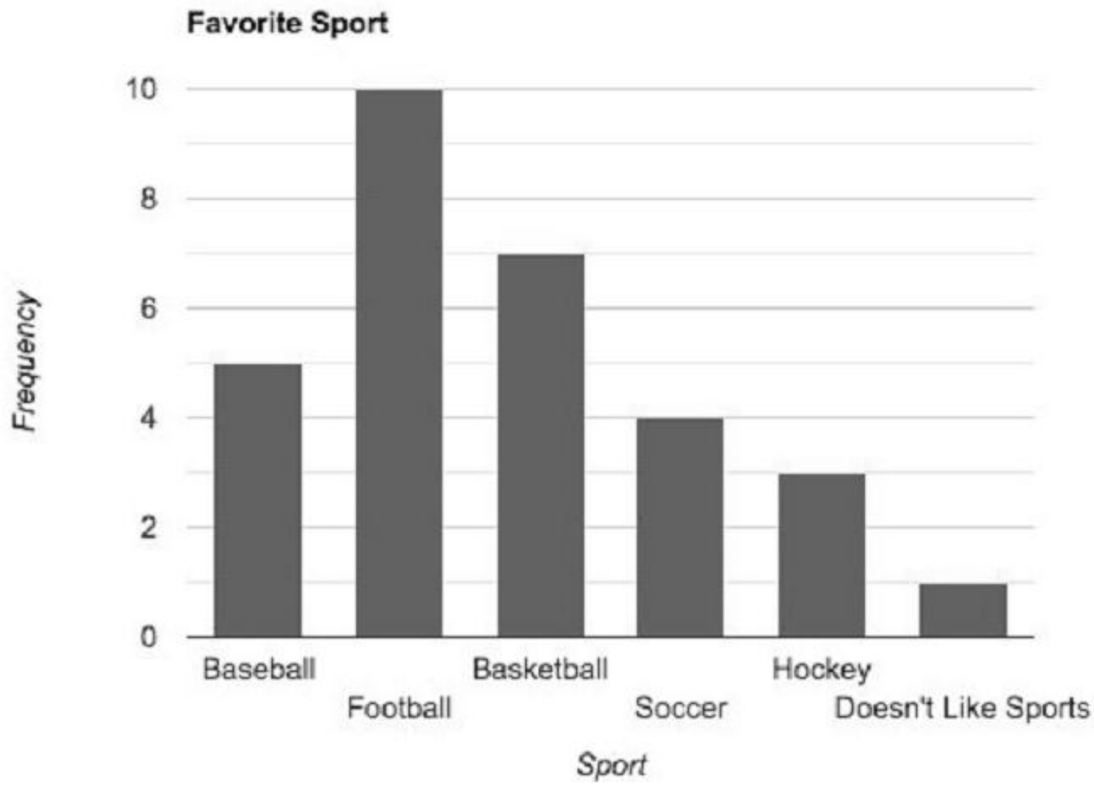
Name	Asad	Amna	Hina	Raza	Umair
MBF	70	57	65	75	82
I to B	65	70	57	45	60

Solution



31. A group of students surveyed their class about what sport was their favorite.

The results are given below.



a) How many students said basketball was their favorite?

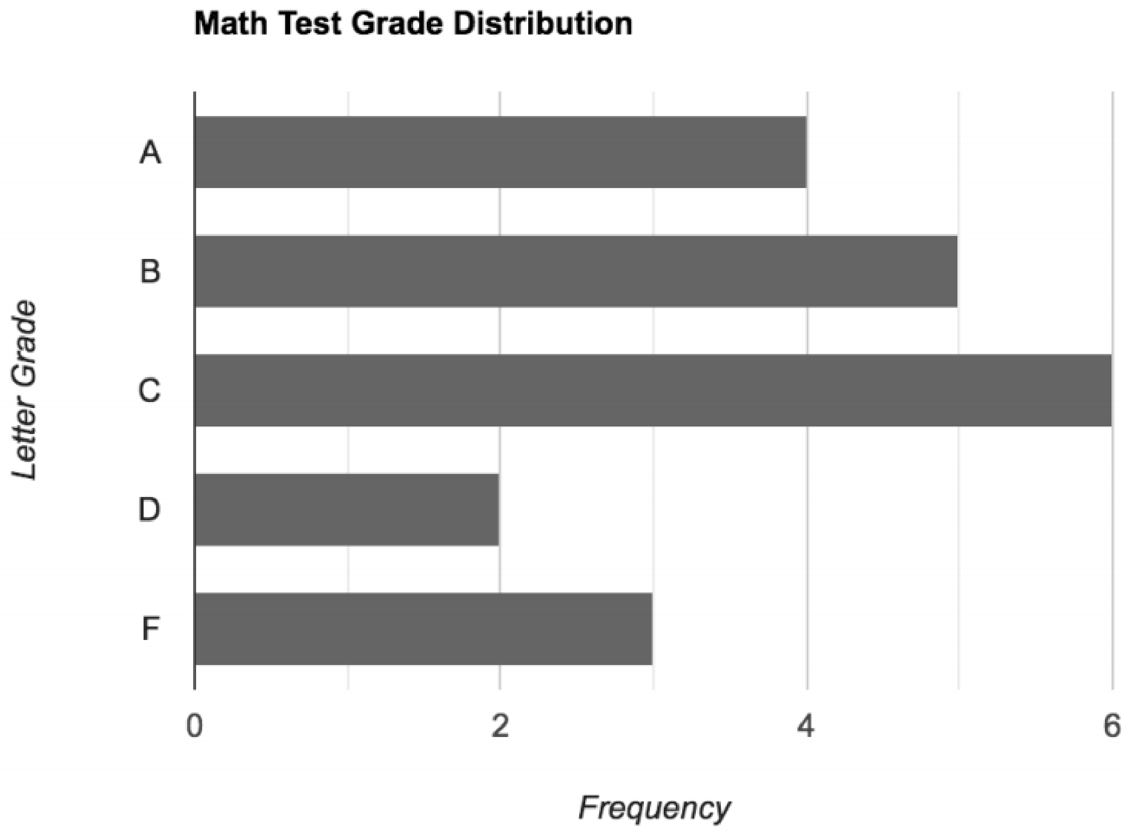
b) How many more students liked Football than Soccer?

Solution

a) Find the bar above that is labelled "Basketball" on the horizontal axis. The number of students whose favorite is basketball is equivalent to the height of the bar. Read the height of the bar from the vertical axis. The number of students who said basketball was their favorite sport was 7.

b) Begin by observing the number of students who liked Football and Soccer from the bar graph. The number of students who liked Football is 10 and the number of students who liked Soccer is 4. There are $10 - 4 = 6$ more students that liked Football than Soccer.

32. The bar graph below shows scores on a Math test.



a) How many B's were there?
b) How many test grades are there?
c) How many students scored a C or higher?

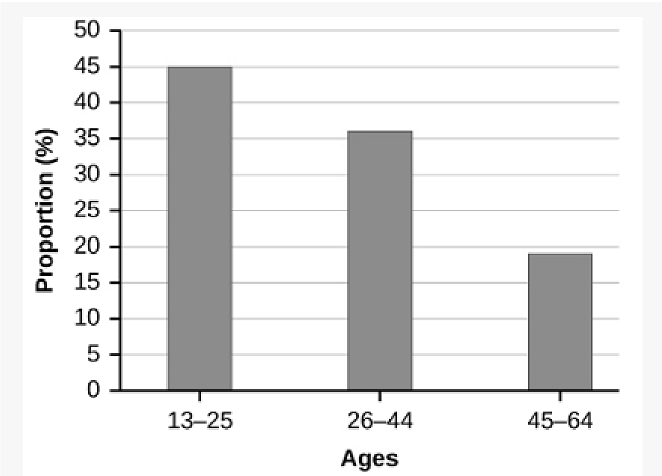
Solution

a) This is a horizontal bar graph. The grade categories are on the vertical axis. Begin by looking for "B". Then look at the horizontal axis for the length of the bar for B's. This gives 5 B's.
b) Add all the lengths of each bar for each letter grade together to find the number of test grades. This gives a total of 4 + 5 + 6 + 2 + 3 = 20
c) Begin by finding the number of test grades for A's, B's, and C's, which are 4, 5, and 6 respectively. Number of grades of C or better are 4 + 5 + 6 = 15.

33. By the end of 2011, Facebook had over 146 million users in the United States. Table shows three age groups, the number of users in each age group, and the proportion (%) of users in each age group. Construct a bar graph using this data.

Age groups	Number of Facebook users	Proportion (%) of Facebook users
13–25	65,082,280	45%
26–44	53,300,200	36%
45–64	27,885,100	19%

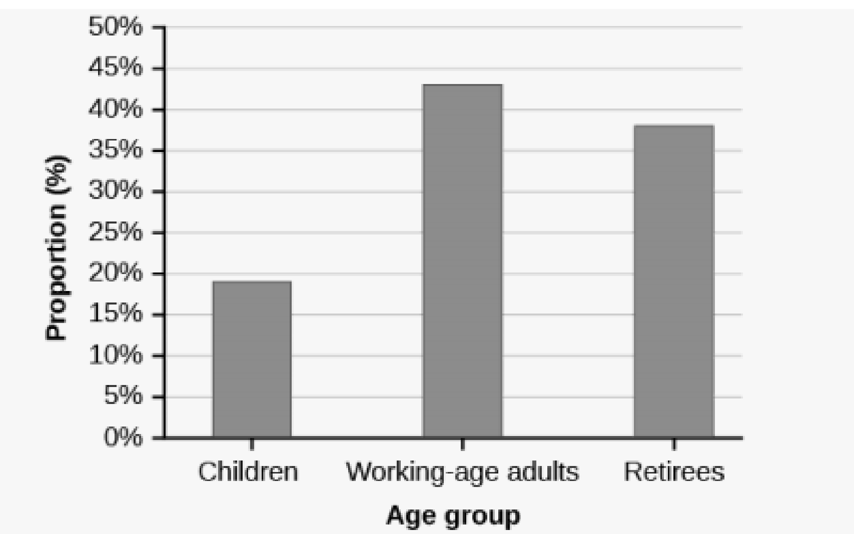
Solution



34. The population in Park City is made up of children, working-age adults, and retirees. Table shows the three age groups, the number of people in the town from each age group, and the proportion (%) of people in each age group. Construct a bar graph showing the proportions.

Age groups	Number of people	Proportion of population
Children	67,059	19%
Working-age adults	152,198	43%
Retirees	131,662	38%

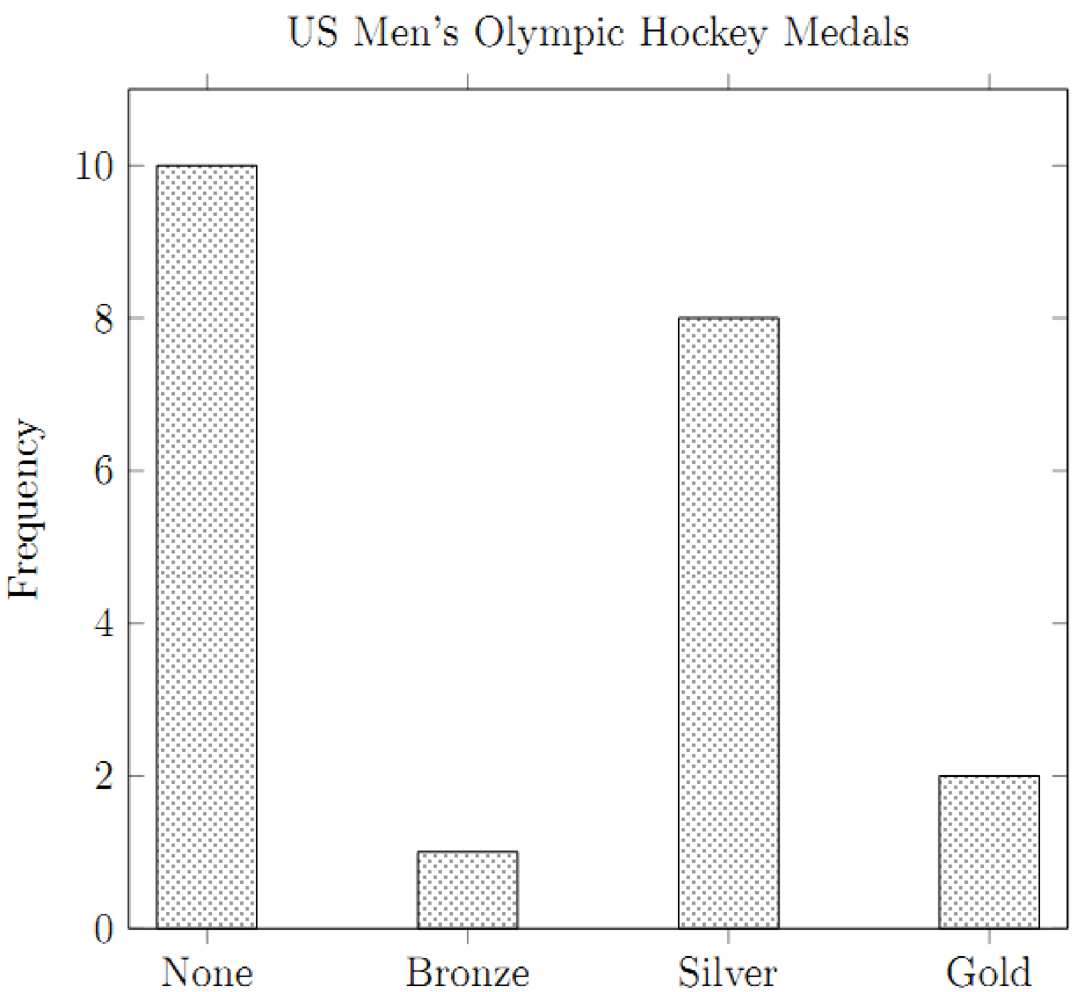
Solution

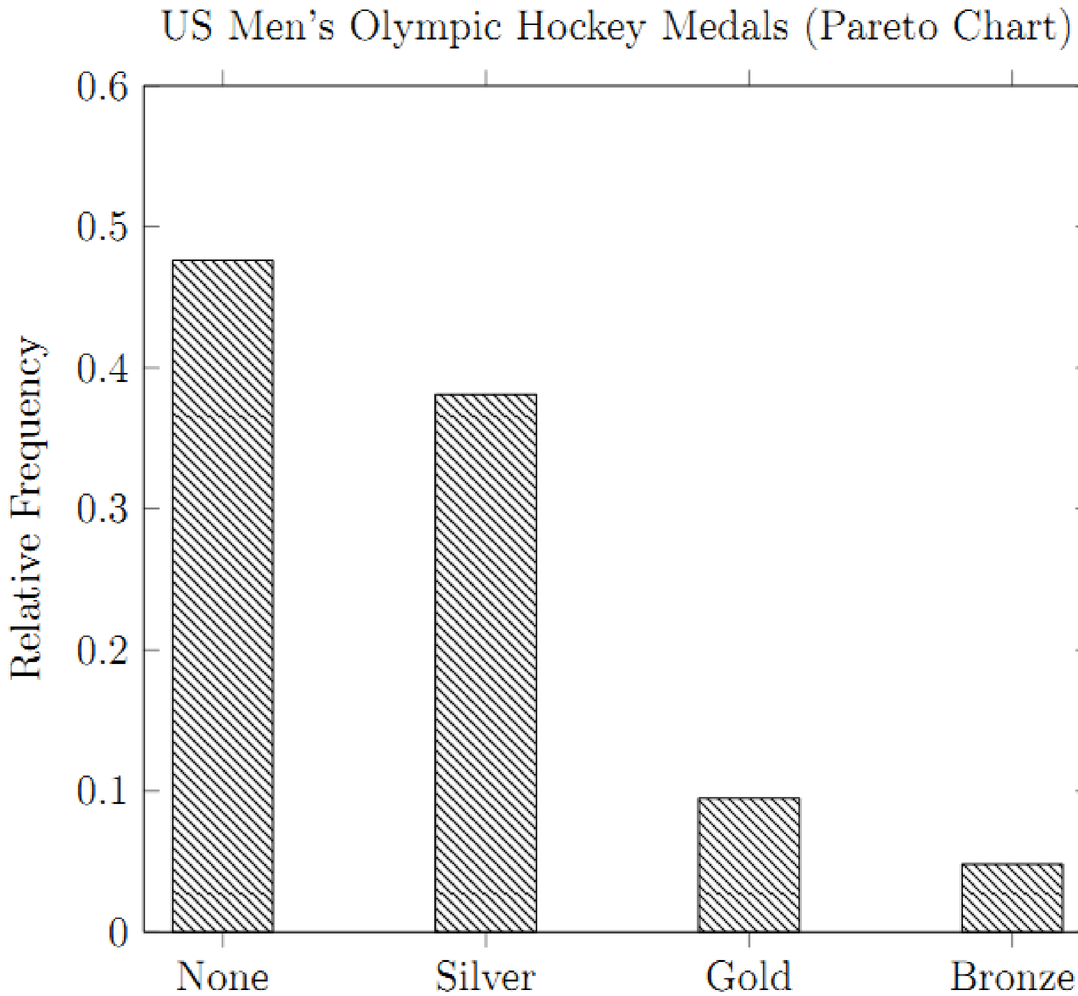


35. The US Men's olympic hockey teams have played in 21 olympic games, winning 11 medals (2 gold, 8 silver, 1 bronze). Make a frequency bar graph and a relative frequency Pareto chart.

Solution

Below are a frequency bar graph (ordered from worst to best finishes) and a relative frequency Pareto chart (ordered from highest to lowest).



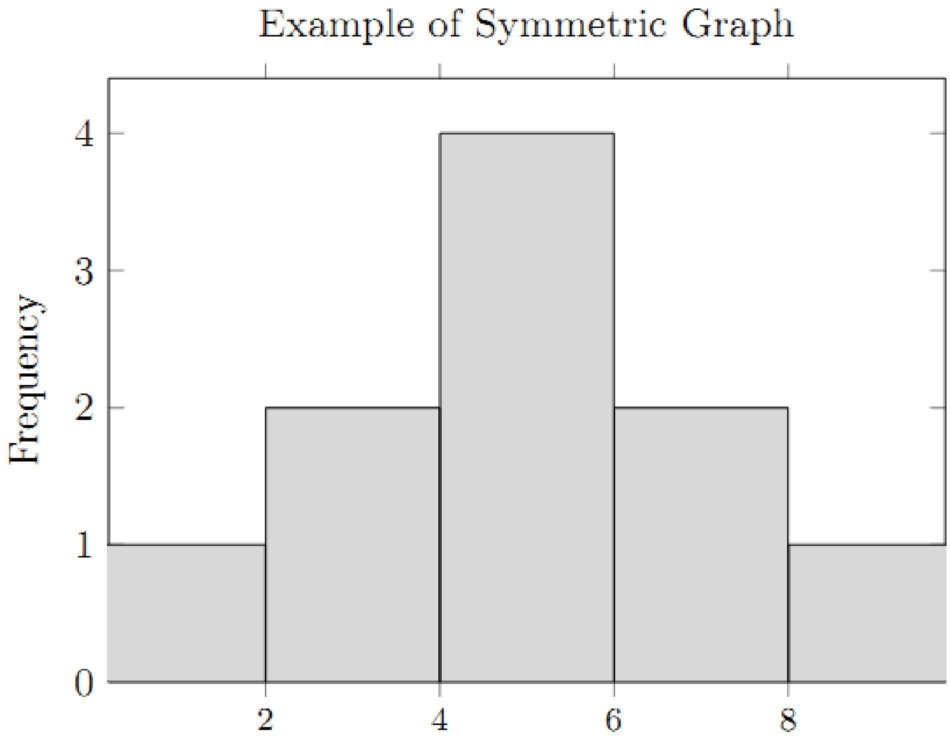


Graphs can help us get an overall view of the data set. When looking at a graph, pay attention to the following:

- **Center:** where is the middle of the graph, and the highest point.
- **Spread:** how are the parts of the graph spread out from each other?
- **Shape:** what shape does the graph have? Bell shape, straight across, repeatedly up and down, random?
- **Symmetry:** graph can be split in half with two mirror image parts, almost equal amount on both sides. Graph that extends more out to left is Left-skewed. Graph that extends more out to right is Right-skewed.
- **Outliers:** are data values (small parts of the graph) that are far from the other data (parts).

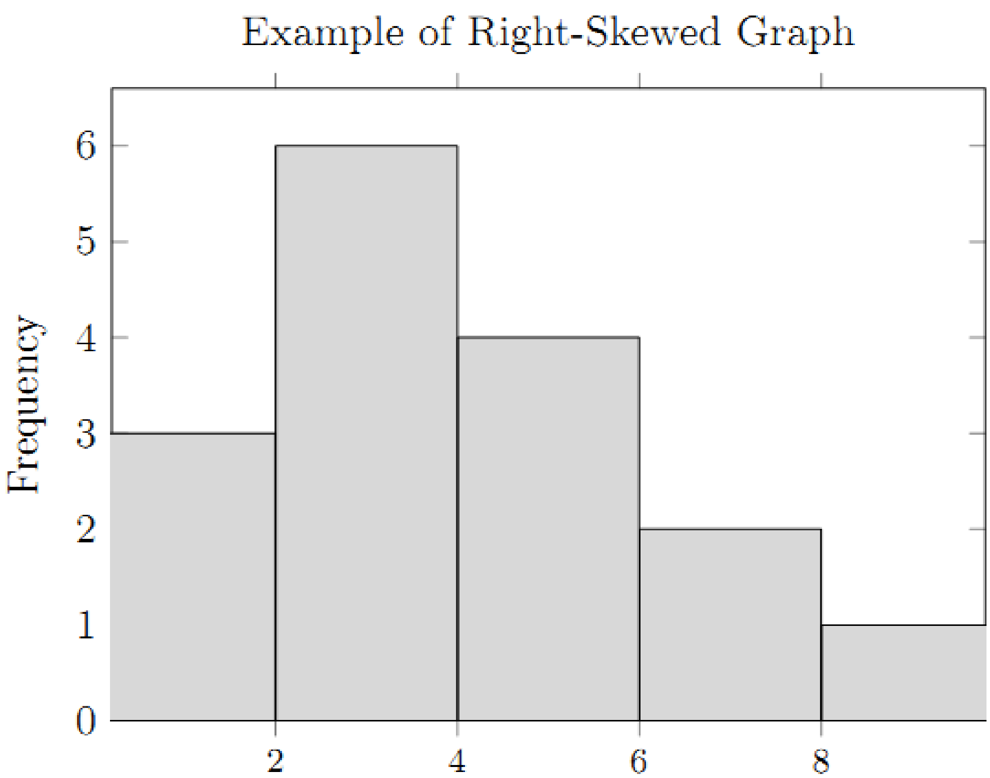
36. Give an example of symmetric graph.

Answer



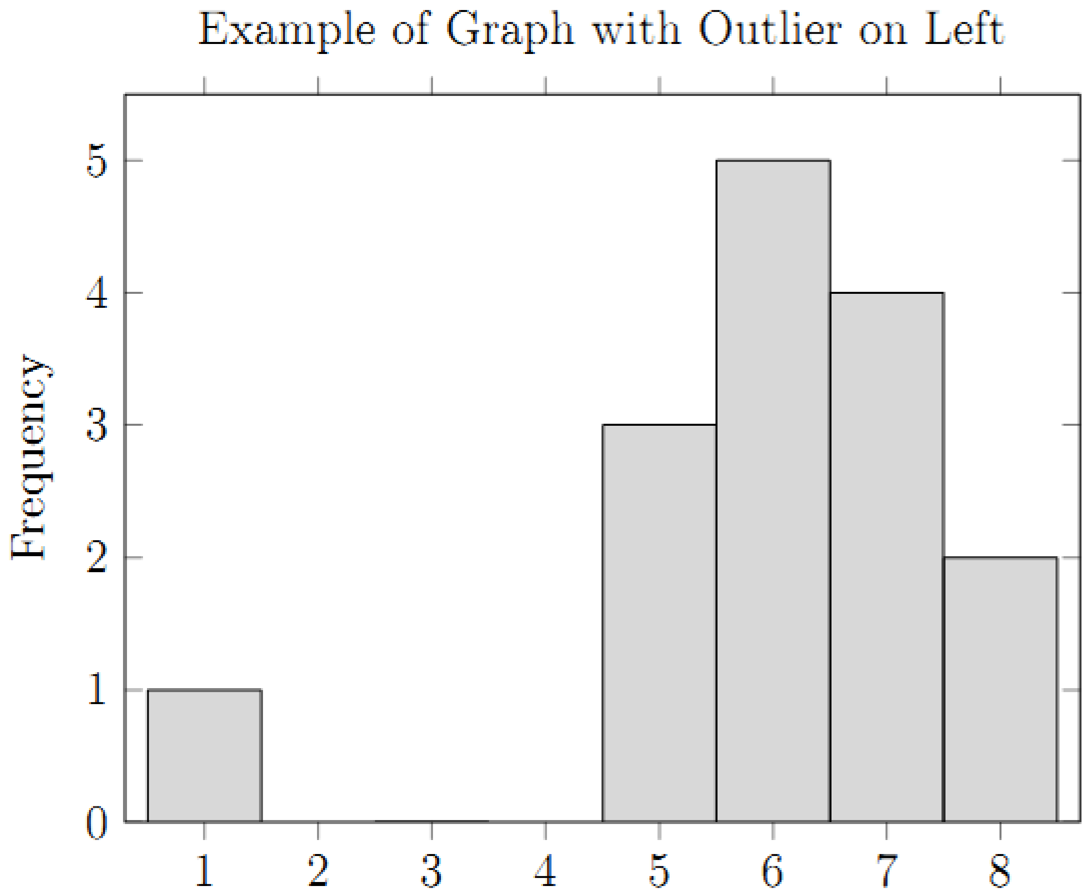
37. Give an example of right skewed graph.

Answer



38. Give an example of graph for outlier.

Answer



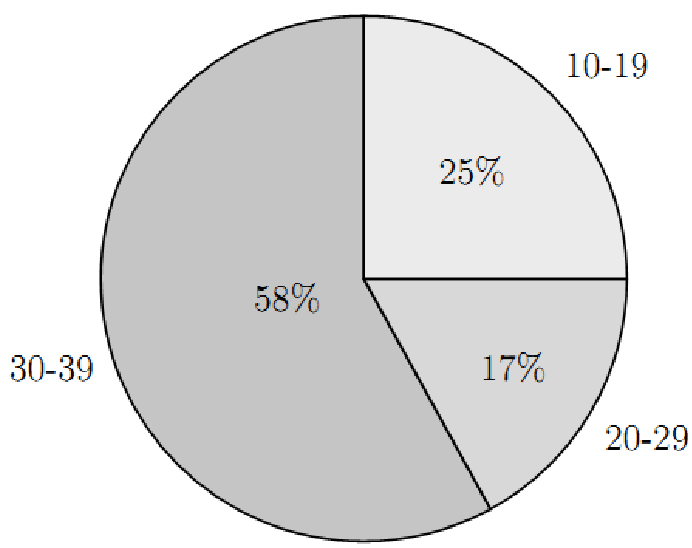
Circle Graph / Pie Chart

A Circle Graph (or Pie Chart) is a circle cut into sections with varying sizes shaped like slices of a pizza. The sizes of the sections are based on the relative frequencies of the categories. The percent or frequency for each category can be specified on the sections of the pie chart.

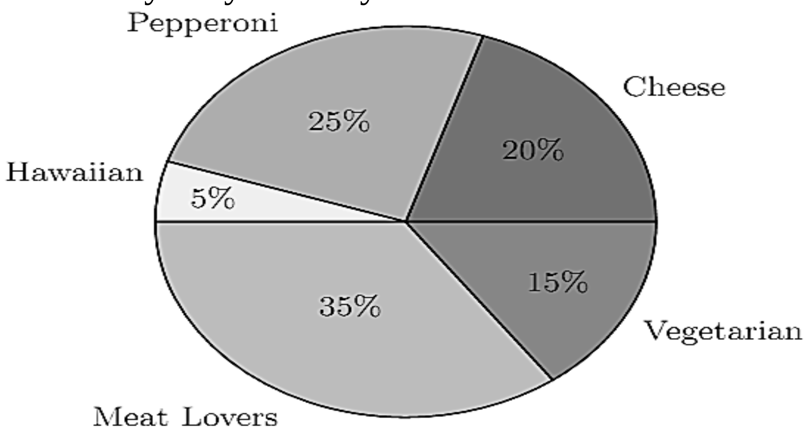
It is a graphical device for presenting qualitative data summaries based on subdivision of a circle into sectors that corresponds to the relative frequency for each class. It is a disk (circle) divided into pie-shaped pieces proportional to the relative frequencies. A pie chart should be labeled well, with class and the relative frequency for each slice. If a slice is very small, then the labels can go outside with an arrow pointing to the corresponding slice. The preferred way to sketch a pie chart is to start slices at 12 o'clock and rotate clockwise.

39. For Bradley's weekly hours at a summer job: 25, 32, 36, 32, 18, 28, 30, 36, 12, 16, 35, 36. Construct a Pie Chart.

Solution



40. A pizza chef at Mario's Pizza makes a list of all the types of pizza he made on a particularly busy Tuesday. He then used this list to create a pie chart for the pizza types.

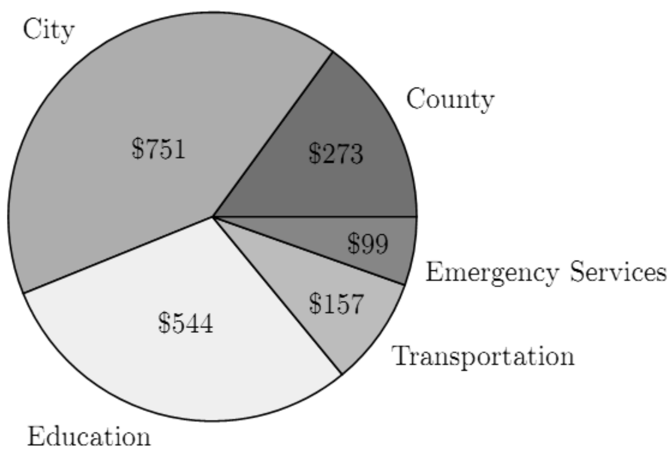


a) What was the least ordered type of pizza?
b) What was the most ordered type of pizza?

Solution

a) Start by identifying the smallest piece on the pie chart. Looking at this graph shows that this category is that of Hawaiian at 5%.
b) Identifying the largest piece on the pie chart. Looking at this graph shows that this category is that of Meat Lovers at 35%.

41. Using the Circle Graph answer the following questions.



a) How much in taxes did Tyrone pay in total?
b) How much of Tyrone's Taxes went to the City?
c) Which of the categories on the pie chart received the smallest amount of money?

Solution

a) Sum of the money amount for each category represented on the chart. The total amount is $99 + $157 + $273 + $544 + $751 = $1824.
b) Look on the chart for the piece labeled City and read the value there which is $751.
c) Look at the chart and look for the category with the lowest amount of money. In this case the category is Emergency Services at $99.

Misleading Graphs

Misleading graphs are visual representations of data that intentionally or unintentionally deceive or mislead the viewer, often by manipulating the presentation, scale, or data selection to support a particular agenda or conclusion.

Examples of Misleading Graphs in Statistics

Truncated Axis: Omitting parts of the axis to exaggerate changes.

Example: A graph showing a 10% increase in sales, but the y-axis starts at 90% instead of 0%.

Logarithmic Scale: Using a logarithmic scale to make small changes appear large.

Example: A graph showing stock prices, using a logarithmic scale to exaggerate fluctuations.

Biased Labeling: Using emotive or misleading labels.

Example: A graph labeled "Huge Increase in Crime Rate" when the actual increase is small.

Cherry-Picked Data: Selectively presenting data to support a claim.

Example: Showing only data that supports a trend, while ignoring contradictory data.

Omitting Data: Leaving out important data to change the narrative.

Example: Omitting data points that contradict a trend.

Dual Scales: Using two different scales on the same graph.

Example: A graph with two different scales, making it difficult to compare data.

Cumulative Graphs: Using cumulative data to make a trend appear more impressive.

Example: Showing cumulative sales over time, rather than monthly sales.

Average vs. Median: Using averages instead of medians to misrepresent data.

Example: Using average income instead of median income to hide income inequality.

Misleading Colors: Using colors to influence interpretation.

Example: Using red to indicate a "bad" trend and green to indicate a "good" trend.

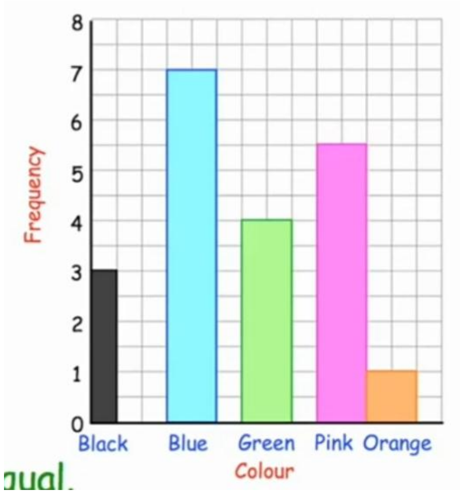
3D Graphs: Using 3D graphs to make data appear more impressive.

Example: A 3D pie chart that exaggerates the size of a particular slice.

42. Using an example show that bar chart is a misleading graph.

Solution.

Colour	Frequency
Black	3
Blue	7
Green	4
Pink	6
Orange	1

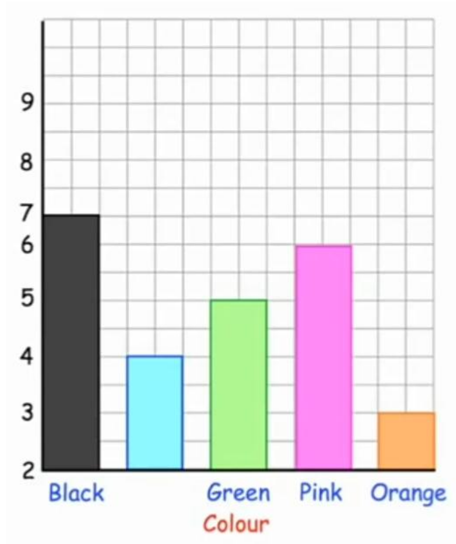


- Bars have not equal width.
- Spacing between bars not equal.
- Bars drawn at the wrong height.

43. Using an example show that bar chart is a misleading graph.

Solution.

Colour	Frequency
Black	7
Blue	4
Green	5
Pink	6
Orange	3

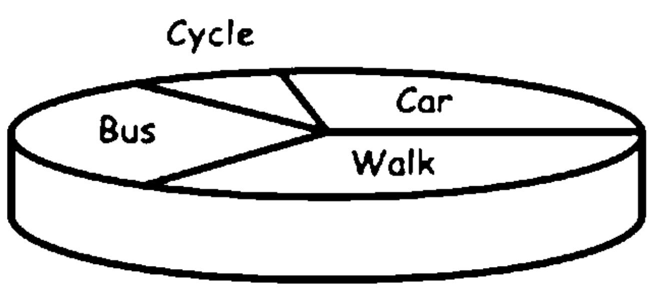


- Frequency axis not starting at zero.
- Frequency axis not labelled.
- One of the bars not labelled.
- Spacing between numbers on frequency axis (or missed).

44. Using an example show that Pie chart is a misleading graph.

Solution.

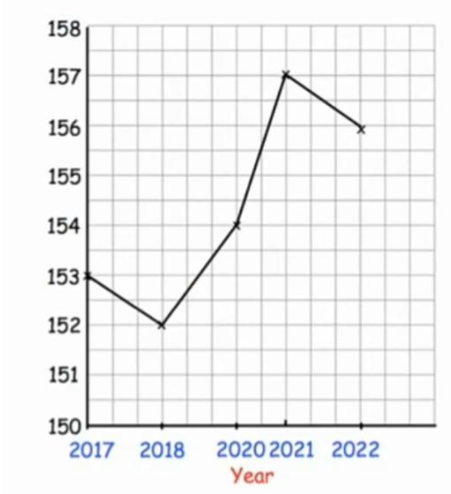
- 3D pie charts can be misleading.
- Sector not labelled.
- Points plotted in the wrong place.
- Wrong angle drawn.
- Wrong angle calculated.



45. Using an example show that line graph is a misleading graph.

Solution.

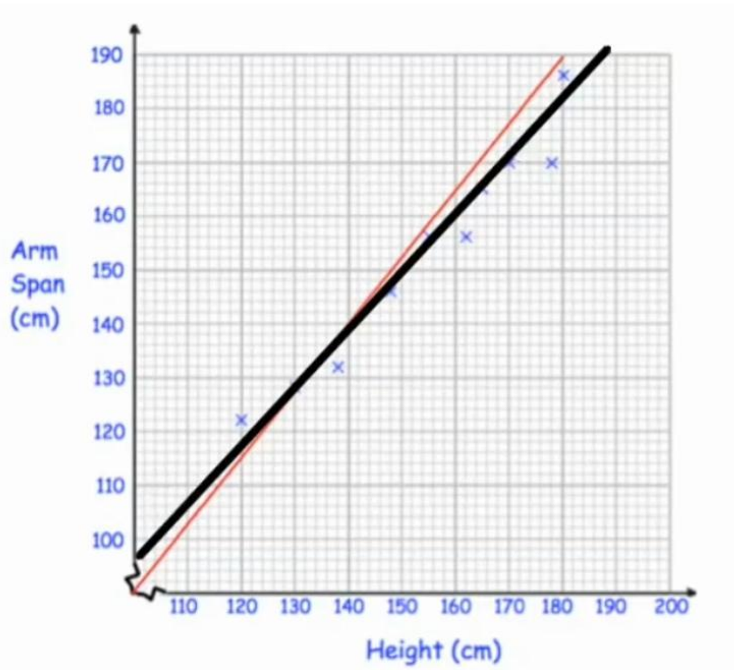
Year	**Share Price**
2017	154
2018	152
2019	154
2020	157
2021	156



- Vertical axis not starting at zero.
- Vertical axis not labelled.
- Points plotted in the wrong place.
- Number missing on horizontal axis.
- Spacing on the axes.
- Curves rather than lines (not shown).

46. Using an example show that scattor graph is a misleading graph.

Solution.



- Points plotted in the wrong place.
- Problems with axes – scale, label, suitable etc.
- Line of best fit not drawn in a suitable location.

Bivariate Analysis

Bivariate analysis is a statistical technique used to examine the relationship between two continuous or categorical variables. One variable is dependent while other is independent.

Purpose

- Identify relationships between variables
- Determine correlation or association strength
- Visualize data distribution

Types of Bivariate Analysis

- **Scatter Plots:** Visualize relationship between two continuous variables
- **Correlation Coefficient:** Measure strength and direction of linear relationship
- **Contingency Tables:** Examine relationship between two categorical variables
- **T-Tests:** Compare means of two groups
- **Regression Analysis:** Model relationship between independent and dependent variables

Key Concepts

- **Correlation:** Measure of linear relationship strength
- **Causation:** Relationship implying cause-and-effect
- **Association:** Relationship between variables
- **Independence:** No relationship between variables

Applications

- **Market Research:** Analyze customer behavior
- **Medical Research:** Investigate disease relationships
- **Social Sciences:** Examine relationships between demographic variables
- **Business:** Identify relationships between sales and marketing strategies

Regression line or Least-Squares Line

The regression line or least-squares line is the line that minimizes the sum of the squares of the vertical distances between the data points and the line. **Or** it is the line for which the sum of the squares of the vertical distances between the data points and the line is a minimum.

Linear Regression
The process of finding the regression line is called **linear regression**.

Regression
The dependence of variable upon one or more other variables is called regression.

Simple Regression
The dependence of one variable upon single independent variable is called simple regression.

Multiple Regressions
The dependence of one variable upon two or more independent variables is called simple regression.

Regressor
The variable that forms the basis of estimation or prediction is called regressor. It is also called as the predictor variable or independent variable or controlled variable or explanatory variable or non – random variable.

Regressand
The variable whose resulting value depends upon the selected value of the independent variable is called the regressand. It is also called the response variable or the predicted variable or dependent variable or explained variable or random variable.

Regression Analysis
Regression analysis refers to the methods of describing the functional (dependent) on the basis of the other (independent) variable.

The Principle of Least Square
This principle says that the sum of the squares of the residual of the observed values from their corresponding estimated values should be the least or minimum.

Mathematically; $\text{Least} = S = \sum\left(Y_i - \hat{Y}\right)^2$

Estimated Regression Line

The estimated regression line of Y on X; $Y = a_{YX} + a_{YX}X$	The estimated regression line of X on Y; $X = a_{XY} + a_{XY}Y$
$b_{YX} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sum(X-\bar{X})^2}$	$b_{XY} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sum(Y-\bar{Y})^2}$
$b_{YX} = \frac{\sum(XY - n\bar{X}\,\bar{Y})}{\sum(X-\bar{X})^2}$	$b_{XY} = \frac{\sum(XY - n\bar{X}\,\bar{Y})}{\sum(Y-\bar{Y})^2}$
$b_{YX} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}}$	$b_{XY} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sum Y^2 - \frac{(\sum Y)^2}{n}}$
$b_{YX} = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2}$	$b_{XY} = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum Y^2 - (\sum Y)^2}$
$b_{YX} = \frac{S_{XY}}{S_X^2} = r\frac{S_Y}{S_X}$	$b_{XY} = \frac{S_{XY}}{S_Y^2} = r\frac{S_X}{S_Y}$
$a_{YX} = \frac{\sum Y - b_{YX}\sum X}{n}$	$a_{XY} = \frac{\sum X - b_{XY}\sum Y}{n}$
$a_{YX} = \bar{Y} - b_{YX}\bar{X}$	$a_{XY} = \bar{X} - b_{XY}\bar{Y}$

47. Fit a regression line Y on X from percentage of marks scored by 12 students in statistics X and economics Y.

x	30	34	26	49	60	62	65	51	44
y	27	18	34	28	26	30	32	30	28

Solution

The estimated regression line is $\hat{y} = a + bx$

x	y	x^2	xy
30	27	900	810
34	18	1156	612
26	34	676	884
49	28	2401	1372
60	26	3600	1560
62	30	3844	1860
65	32	4225	2080
51	30	2601	1530
44	28	1936	1232
$\sum x = 421$	$\sum y = 253$	$\sum x^2 = 21339$	$\sum xy = 11940$

The least square estimate for n = 9 are

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2} = 0.064$$

$$a = \frac{\sum y - b\sum x}{n} = 25.12$$

The best fitted line is $\hat{y} = 25.12 + 0.064x$

48. Fit a regression line X on Y from percentage of marks scored by 12 students in statistics X and economics Y.

x	30	34	26	49	60	62	65	51	44
y	27	18	34	28	26	30	32	30	28

Solution

The estimated regression line is $\hat{y} = a + by$

x	y	y^2	xy
30	27	729	810
34	18	324	612
26	34	1156	884
49	28	784	1372
60	26	676	1560
62	30	900	1860
65	32	1024	2080
51	30	900	1530
44	28	784	1232
$\sum x = 421$	$\sum y = 253$	$\sum x^2 = 7277$	$\sum xy = 11940$

The least square estimate for n = 9 are

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum y^2 - (\sum y)^2} = 0.64$$

$$a = \frac{\sum x - b\sum y}{n} = 28.79$$

The best fitted line is $\hat{y} = 28.79 + 0.64y$

Method of Least Square

A procedure in which regression equation is obtained by minimizing the sum of squares of residuals (errors). The parameter values obtained are called least square estimates.

Normal Equations of Least Square Regression Line

- $\sum Y = na + b\sum X$ this is the normal equation for 'a'
- $\sum XY = a\sum X + b\sum X^2$ this is the normal equation for 'b'

Residual

Residual is the difference between actual value Y_i from predicted value $\widehat{Y}_i$. This error term is denoted by e_i. i.e. Residual $= e_i = Y_i - \widehat{Y}_i = Y_i - (a + bX_i) = Y_i - a - bX_i$

49. Fit a straight line by method of least squares to the following data and estimate Y for X = 30 where X = Supply and Y = Demand.

Score(x)	0	5	10	15	20	25
GPA(y)	12	15	17	22	24	30

Solution

The estimated equation of straight line is $\hat{y} = a + bx$

x	y	x^2	xy
0	12	0	0
5	15	25	75
10	17	100	170
15	22	225	330
20	24	400	480
25	30	625	750
$\sum x = 75$	$\sum y = 120$	$\sum x^2 = 1375$	$\sum xy = 1805$

The normal equations are

$\sum Y = na + b\sum X \Rightarrow 6a + 75b = 120$ ……….. (i)

$\sum XY = a\sum X + b\sum X^2 \Rightarrow 75a + 1375b = 1805$ ……….. (ii)

By $25(\text{i}) - 2(\text{ii})$ we have $a = 11.25, b = 0.7$

The required fitted straight line is $\hat{y} = 11.25 + 0.7x$

To estimate the value of Y put X = 30; $\hat{y} = 11.25 + 0.7(30) = 32.25$

50. Find a straight line by the method of least squares and show that sum of errors is always zero.

Score(x)	0	1	2	3	4	5	6
GPA(y)	12	10	14	11	13	15	16

Solution

The estimated equation of straight line is $\hat{y} = a + bx$

x	y	x^2	xy
0	12	0	0
1	10	1	10
2	14	4	28
3	11	9	33
4	13	16	25
5	15	25	75
6	16	36	96
$\sum x = 21$	$\sum y = 91$	$\sum x^2 = 91$	$\sum xy = 294$

The normal equations are

$\sum Y = na + b\sum X \Rightarrow 7a + 21b = 91$ ……….. (i)

$\sum XY = a\sum X + b\sum X^2 \Rightarrow 21a + 91b = 294$ ……….. (ii)

By (ii) $-$ 3(i) we have $a = 10.75, b = 0.75$

The required fitted straight line is $\hat{y} = 10.75 + 0.75x$

x	y	$\hat{y}$	$y - \hat{y}$
0	12	$10.75 + 0.75(0) = 10.75$	1.25
1	10	$10.75 + 0.75(1) = 10.75$	-1.5
2	14	$10.75 + 0.75(2) = 10.75$	1.75
3	11	$10.75 + 0.75(3) = 10.75$	-2
4	13	$10.75 + 0.75(4) = 10.75$	-0.75
5	15	$10.75 + 0.75(5) = 10.75$	0.50
6	16	$10.75 + 0.75(6) = 10.75$	0.75
$\sum x = 21$	91	91	0

Hence

Sum of errors is always zero. i.e. $\sum(Y - \hat{Y}) = 0$

51. The following sample of 8 grade point averages and marks in matriculation was observed for students from a college.

Score(x)	480	490	510	510
GPA(y)	2.7	2.9	3.3	2.9
Score(x)	530	550	610	640
GPA(y)	3.1	3.0	3.2	3.7

Find the least square line. Estimate the mean GPA of student scoring 600 marks.

Solution

The estimated regression line is $\hat{y} = a + bx$

x	y	x^2	xy
480	2.7	230400	1296
490	2.9	240100	1421
510	3.3	260100	1683
510	2.9	260100	1479
530	3.1	280900	1643
550	3.0	302500	1650
610	3.2	372100	1952
640	3.7	409600	2368
$\sum x = 4320$	$\sum y = 24.8$	$\sum x^2 = 2355800$	$\sum xy = 13492$

The least square estimate for n = 8 are

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2} = 0.00435$$

$$a = \frac{\sum y - b\sum x}{n} = 0.751$$

The best fitted line is $\hat{y} = 0.751 + 0.00435x$

For x = 600 we have $\hat{y} = 0.751 + 0.00435(600) = 3.361$

52. Given the following data

x	0	1	2	3	4
y	1.0	1.8	3.3	4.5	6.3

Determine the least square line taking x as independent variable. Find the estimated values for the given value of x and show that; $\sum y = \sum \hat{y}$; $\sum e = 0$. Also calculate the sum of squares of the residual. And verify that $\sum e^2 = \sum y^2 - a\sum y - b\sum xy$

Solution

The estimated regression line is $\hat{y} = a + bx$

x	y	x^2	xy
0	1.0	0	0
1	1.8	1	1.8
2	3.3	4	6.6
3	4.5	9	13.5
4	6.3	16	25.2
$\sum x = 10$	$\sum y = 16.9$	$\sum x^2 = 30$	$\sum xy = 47.1$

The least square estimate for n = 5 are

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2} = 1.33 \text{ and } a = \frac{\sum y - b\sum x}{n} = 0.72$$

The best fitted line is $\hat{y} = 0.72 + 1.33x$

The estimated values $\hat{y}$ for the given values of x and the residuals $e = y - \hat{y}$ are obtained as shown in the following table.

x	y	$\hat{y}$	e	e^2	y^2
0	1.0	$0.72 + 1.33(0) = 0.72$	0.28	0.0784	1.00
1	1.8	$0.72 + 1.33(1) = 2.05$	-0.25	0.0625	3.24
2	3.3	$0.72 + 1.33(2) = 3.38$	-0.08	0.0064	10.89
3	4.5	$0.72 + 1.33(3) = 4.71$	-0.21	0.0441	20.25
4	6.3	$0.72 + 1.33(4) = 6.04$	0.26	0.0676	39.69
Sum	16.9	16.90	0	0.2590	75.07

It is verified that

$$\sum y = 16.90 = \sum \hat{y} \text{ and } \sum e = \sum(y - \hat{y}) = 0$$

The sum of squares of residual is $\sum e^2 = 0.2590$

Clearly it is verified that $\sum e^2 = \sum y^2 - a\sum y - b\sum xy$
As $\sum e^2 = 75.07 - 0.72(16.9) - 1.33(47.1) = 0.2590$

53. Sodium thiosulfate is used by photographers to develop some types of film. The amount of this chemical that will dissolve in water depends on the temperature of the water. The table below gives the numbers of grams of sodium thiosulfate that will dissolve in 100 milliliters of water at various temperatures.

Temperature, x, in degrees Celsius	20	35	50	60	75	90	100
Sodium thiosulfate dissolved, y, in grams	50	80	120	145	175	205	230

a. Find the linear regression equation for these data.
b. How many grams of sodium thiosulfate does the model predict will dissolve in 100 milliliters of water when the temperature of the water is 70°C?
Round to the nearest tenth of a gram.

Solution

a. the regression equation is $y = 2.2517731x + 5.2482270$

b. Evaluate the regression equation when x = 70

$$y = 2.2517731x + 5.2482270$$
$$= 2.2517731(70) + 5.2482270$$
$$= 162.872344$$

Approximately 162.9 grams of sodium thiosulfate will dissolve when the temperature of the water is 70°C.

54. The heights and weights of women swimmers on a college swim team are given in the table below.

Height, x, in inches	68	64	65	67	62	67	65
Weight, y, in pounds	132	108	108	125	102	130	105

a. Find the linear regression equation for these data.
b. Use your regression equation to estimate the weight of a woman swimmer who is 63 inches tall. Round to the nearest pound.

Solution

a. The regression equation is approximately $y = 5.6333x - 252.8667$.

b. When x = 63, we have $y = 5.6333(63) - 252.8667 \approx 102$.

The estimated weight of a woman swimmer who is 63 inches tall is approximately 102 pounds.

55. Earlier we compared speed and stride length for dogs. The table below lists data from a similar experiment performed with several adult men.

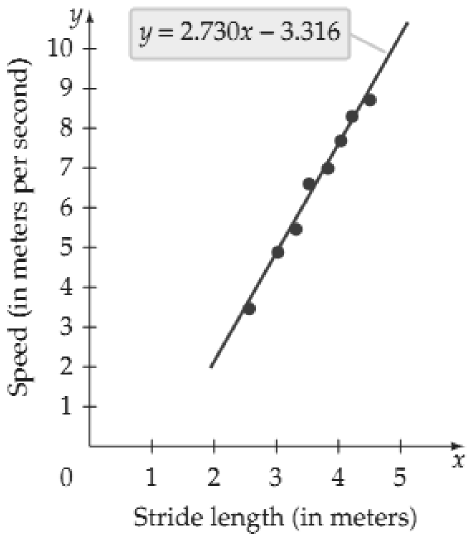
Stride length (meters)	2.5	3.0	3.3	3.5	3.8	4.0	4.2	4.5
Speed (meters per second)	3.4	4.9	5.5	6.6	7.0	7.7	8.3	8.7

a. Find the equation of the regression line for these data.
b. Use your regression equation to predict the average speeds of adult men with stride lengths of 2.8 meters and 4.8 meters. Round your results to the nearest tenth of a meter per second.

Solution

a. the regression equation is approximately $y = 2.730x - 3.316$

The graph is shown as; you can see that the line fits the data well.



b. Evaluate the regression equation when x = 2.8.

$$y = 2.730(2.8) - 3.316$$
$$= 4.328$$

Rounded to the nearest tenth, the predicted average speed for an adult man with a stride length of 2.8 meters is 4.3 meters per second.

Similarly, substituting 4.8 for gives $y = 2.730(4.8) - 3.316 = 9.788$, so 9.8 meters per second is the predicted average speed for an adult man with a stride length of 4.8 meters.

56. The table below lists data from an experiment comparing speed and stride length for several camels.

Stride length (meters)	2.5	3.0	3.2	3.4	3.5	3.8	4.0	4.2
Speed (meters per second)	2.3	3.9	4.4	5.0	5.5	6.2	7.1	7.6

a. Find the equation of the regression line for these data.

b. Use your regression equation to predict the average speeds of camels with stride lengths of 2.7 meters and 4.5 meters. Round your results to the nearest tenth of a meter per second.

Solution

a. The equation of the regression line is approximately

$$y = 3.130x - 5.55.$$

b. When x = 2.7, we have $y = 3.130(2.7) - 5.55 \approx 2.9$,

and when x = 4.5, we have $y = 3.130(4.5) - 5.55 \approx 8.5$.

The predicted average speed of a camel with a stride length of 2.7 meters is about 2.9 meters per second, and the predicted average speed for a camel with a stride length of 4.5 meters is approximately 8.5 meters per second.

Correlation

The interdependence between two or more variables is called correlation. It measures the strength or closeness of relationships between two variables.

Linear Correlation Coefficient (developed by Karl Pearson in the early 1900s)

The **linear correlation coefficient** r is a measurement of the interdependence between the variables. It measures the numerical strength or closeness of linear relationships between two variables. It is a measure of how well the regression line fits the given data. If r is positive, then the closer r is to 1, the stronger the linear relationship between the domain and range values and the better the fit of the regression line to the data. If r is negative, then the closer r is to -1, the stronger the linear relationship between the domain and range values and the better the fit of the regression line to the data.

Types of Correlation

Positive or direct correlation: If r is positive, the relationship between the domain and range values has a **positive or direct correlation**. In this case, if the domain value increases, the range value also tends to increase and vice versa. In this case both variables move in the same direction. The value of correlation coefficient for positive correlation is between 0 and 1. i.e. $0 < r < 1$.

Examples

- Relationship between lung cancer and smoking habits.
- Increase in temperature in summer increase the sale of room coolers.
- Here are the two variables moves in the same direction

x	12	15	20	27	30
y	8	10	12	19	25

Negative or inverse correlation: If r is negative, the linear relationship between the domain and range values has a **negative or inverse correlation**. In this case, if the domain value increases, the range value tends to decrease. In this case both variables move in the opposite direction. The value of correlation coefficient for negative correlation is between -1 and 0. i.e. $-1 < r < 0$.

Examples

- The volume of gas will decrease as the pressure increases.
- Increase in supply of a commodity decreases its price.
- Here are the two variables moves in the opposite direction

x	15	18	25	30	33
y	40	35	30	25	20

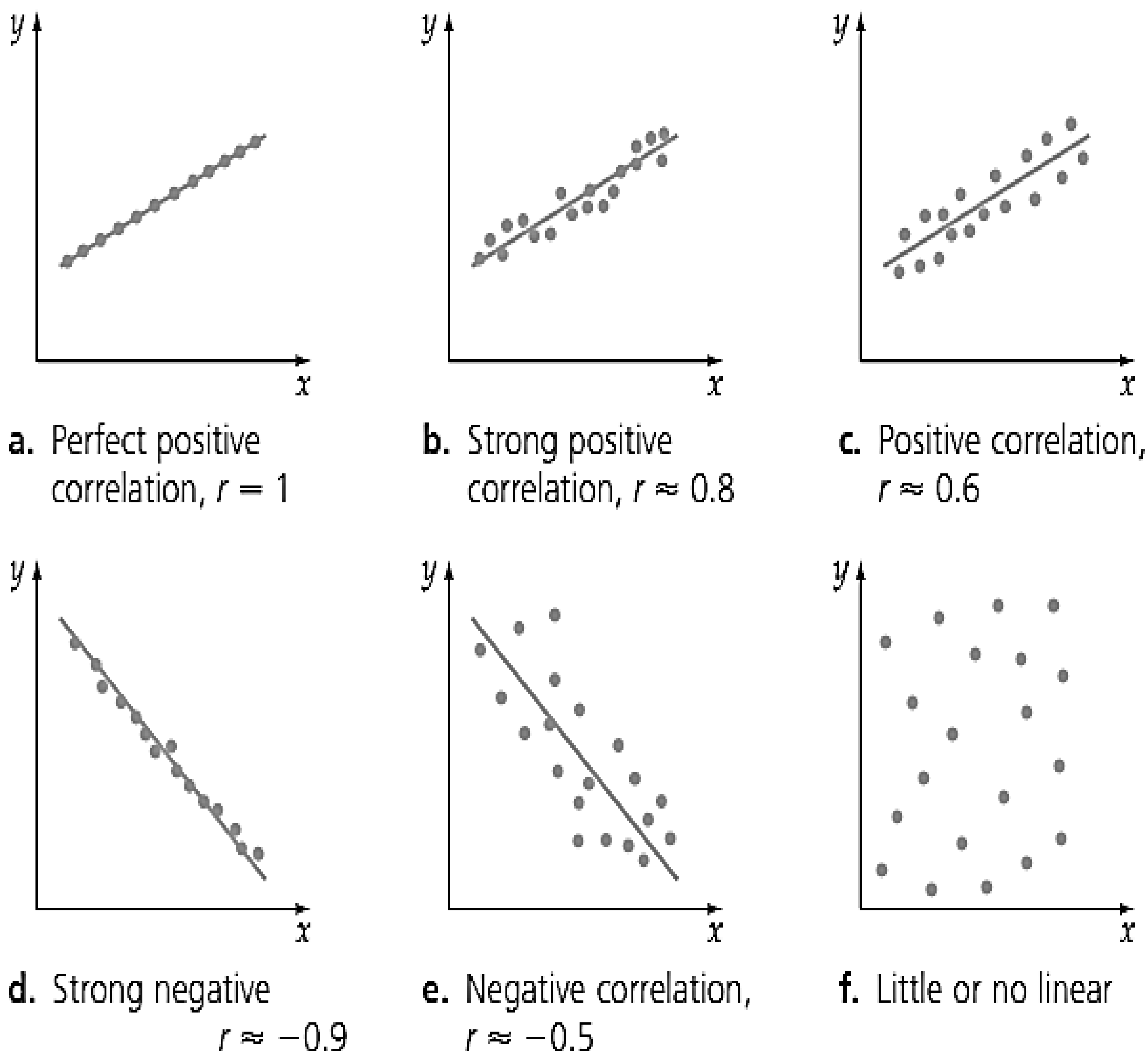
Zero or null correlation: The absence of any relation between the variables is called zero correlation. In this case variables are independent to each other. i.e. $r = 0$.

Examples

- Amount of rainfall and the head sizes.
- Here are the two variables with no effect to each other.

x	1	2	3	4	5
y	7	7	7	7	7

Some Graph of Correlation Types



Correlation Coefficient Formulae

- $r = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$
- $r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$
- $r = \frac{n\sum XY - nXY}{\sqrt{[\sum X^2 - nX^2][\sum Y^2 - nY^2]}}$
- $r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}}$
- $r = \frac{\sum XY - nXY}{\sqrt{[\sum X^2 - nX^2][\sum Y^2 - nY^2]}}$
- $r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{n}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{n}\right]}}$

Properties of Correlation Coefficient

- The Correlation Coefficient always lies between -1 and $+1$. i.e. $-1 \leq r \leq +1$
- The Correlation Coefficient is symmetric with respect to the variables X and Y. i.e. $r_{XY} = r_{YX}$
- The Correlation Coefficient is the geometric mean of the two regression coefficients. i.e. $r_{XY} = \pm\sqrt{b_{XY}.\, b_{YX}}$
- The Correlation Coefficient is a pure number and it has no unit.
- For two independent random variables Correlation Coefficient is zero.
- The Correlation Coefficient is independent of the origin and unit of measurement. i.e. $r_{XY} = r_{UV}$

57. Calculate the correlation coefficient between percentage of marks scored by 12 students in statistics X and economics Y.

x	30	34	26	49	60	62	65	51	44
y	27	18	34	28	26	30	32	30	28

Solution

The estimated regression line is $\hat{y} = a + bx$

x	y	x^2	y^2	xy
30	27	900	729	810
34	18	1156	324	612
26	34	676	1156	884
49	28	2401	784	1372
60	26	3600	676	1560
62	30	3844	900	1860
65	32	4225	1024	2080
51	30	2601	900	1530
44	28	1936	784	1232
421	253	21339	7277	11940

Correlation coefficient between X and Y is

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

$$r = 0.202$$

58. Find the linear correlation coefficient for the data on stride length versus speed of an adult man.

Stride length (meters)	2.5	3.0	3.3	3.5	3.8	4.0	4.2	4.5
Speed (meters per second)	3.4	4.9	5.5	6.6	7.0	7.7	8.3	8.7

Then find the linear correlation coefficient for the speed data for dogs.

Stride length (meters)	1.5	1.7	2.0	2.4	2.7	3.0	3.2	3.5
Speed (meters per second)	3.7	4.4	4.8	7.1	7.7	9.1	8.8	9.9

Which regression line is a better fit for the corresponding data?

Solution

After entering the data for adult men and finding the linear regression equation, we have the linear correlation coefficient as approximately $r = 0.9937$.

Now clear the data for adult men and enter the data for dogs. The regression equation as approximately $y = 3.212x - 1.092$. and the correlation coefficient as approximately $r = 0.9864$. Both correlation coefficients are positive, but because the value of for the adult men data is closer to 1 than the value of for the dog data, the regression line for the adult men fits better than the one for the dogs.

59. Find the linear correlation coefficient for stride length versus speed of a camel as given.

Stride length (meters)	2.5	3.0	3.2	3.4	3.5	3.8	4.0	4.2
Speed (meters per second)	2.3	3.9	4.4	5.0	5.5	6.2	7.1	7.6

Round your result to the nearest hundredth.

Solution

The table below lists data from an experiment comparing speed and stride length for several camels.

Stride length (meters)	2.5	3.0	3.2	3.4	3.5	3.8	4.0	4.2
Speed (meters per second)	2.3	3.9	4.4	5.0	5.5	6.2	7.1	7.6

We have r = 0.998497842, so the linear correlation coefficient is approximately 1.00.

60. Find the correlation coefficient for the data giving the number of persons employed and cloth manufactured in a textile mill.

Person Employed (x)	137	209	113	189	176	200	219
Cloth manufactured (y)	23	47	22	40	39	51	49

Solution

x	y	x^2	y^2	xy
137	23	18769	529	3151
209	47	43681	2209	9823
113	22	12769	484	2486
189	40	35721	1600	7560
176	39	30976	1521	6864
200	51	40000	2601	10200
219	49	47961	2401	10731
1243	271	229877	11345	50815

Correlation Coefficient Formula for n = 7

$$r_{xy} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

$$r_{xy} = \frac{7(50815) - (1243)(271)}{\sqrt{[7(229877) - (1243)^2][7(11345) - (271)^2]}}$$

$$r_{xy} = \frac{355705 - 336853}{\sqrt{[64090][5974]}}$$

$$r_{xy} = \frac{18852}{19567.16}$$

$$r_{xy} = 0.96$$

61. Find the correlation coefficient by using the deviation from their mean for the data giving the height and weight of 8 men.

Height (x)	78	89	97	69	59	79	68	61
Weight (y)	125	137	156	112	107	136	123	106

Solution

$\bar{x} = \frac{\sum x}{n} = \frac{600}{8} = 75$ and $\bar{y} = \frac{\sum y}{n} = \frac{1000}{8} = 1255$

x	y	$x - \bar{x}$	$y - \bar{y}$
78	125	3	0
89	137	11	12
97	156	22	31
69	112	6	−13
59	107	16	−18
79	136	4	11
68	123	−7	−2
61	106	14	−21
Sum 600	1000	0	0

$(x - \bar{x})^2$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
9	0	0
196	144	168
484	961	682
36	169	78
256	324	288
16	121	44
49	4	14
196	441	294
Sum 1242	2164	1568

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} = 0.956$$

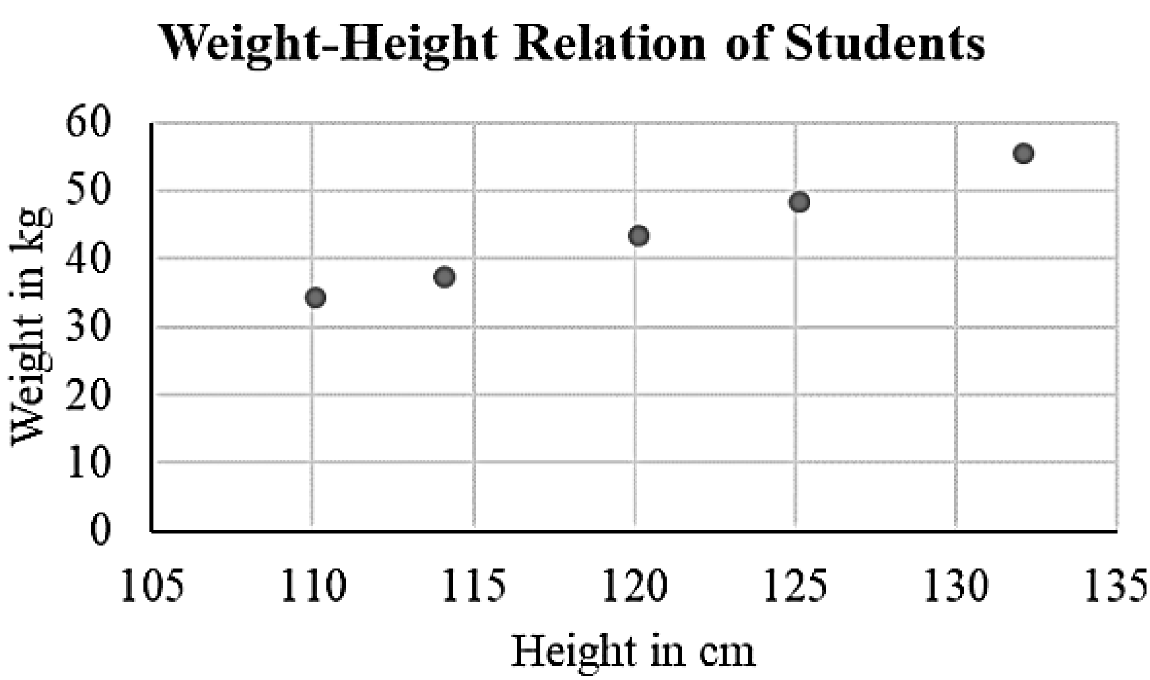
Scatter Plot

A scatter plot is a graph which represents a set of points on the xy - axes. It is a chart type that is normally used to observe and visually display the relationship between variables. Also used for identification of correlational relationships and identification of data patterns.

62. A sample of 5 students have the following body weights and heights. The data is represented in a scatter plot graph. Based on the graph, if a student has a weight of 44 kg, how tall is the student?

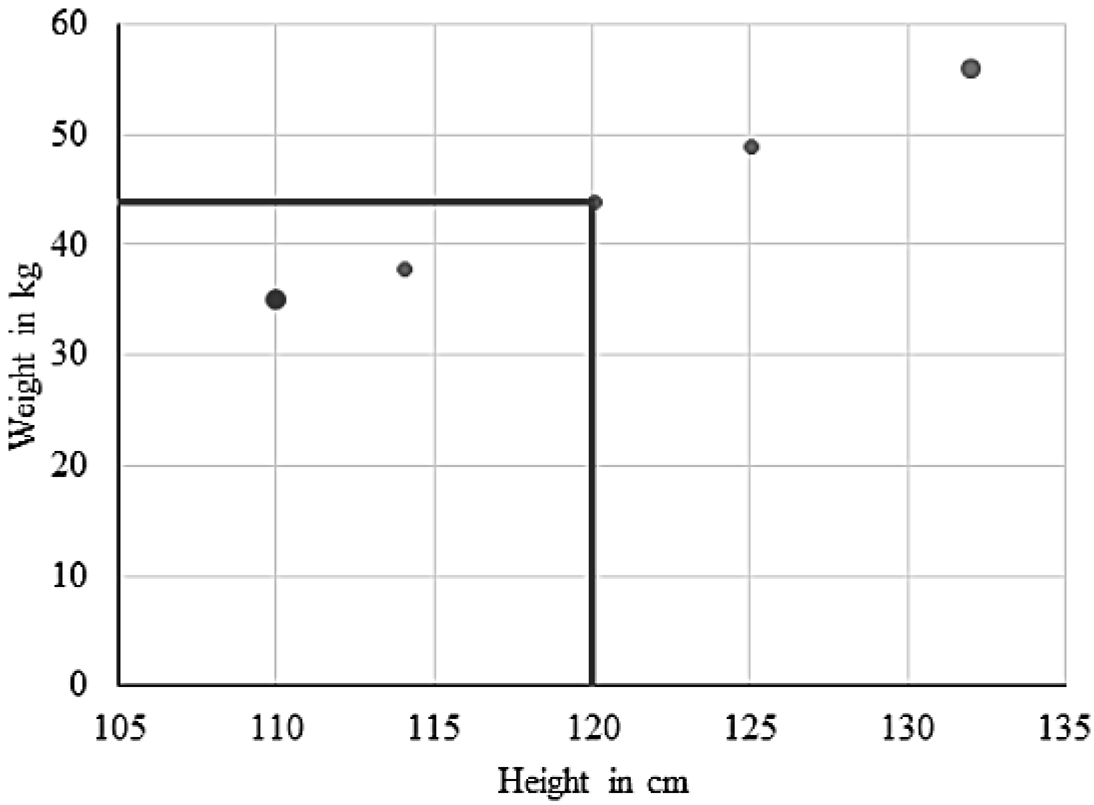
Height in cm	110	114	120	125	132
Weight in kg	35	38	44	49	56

Solution

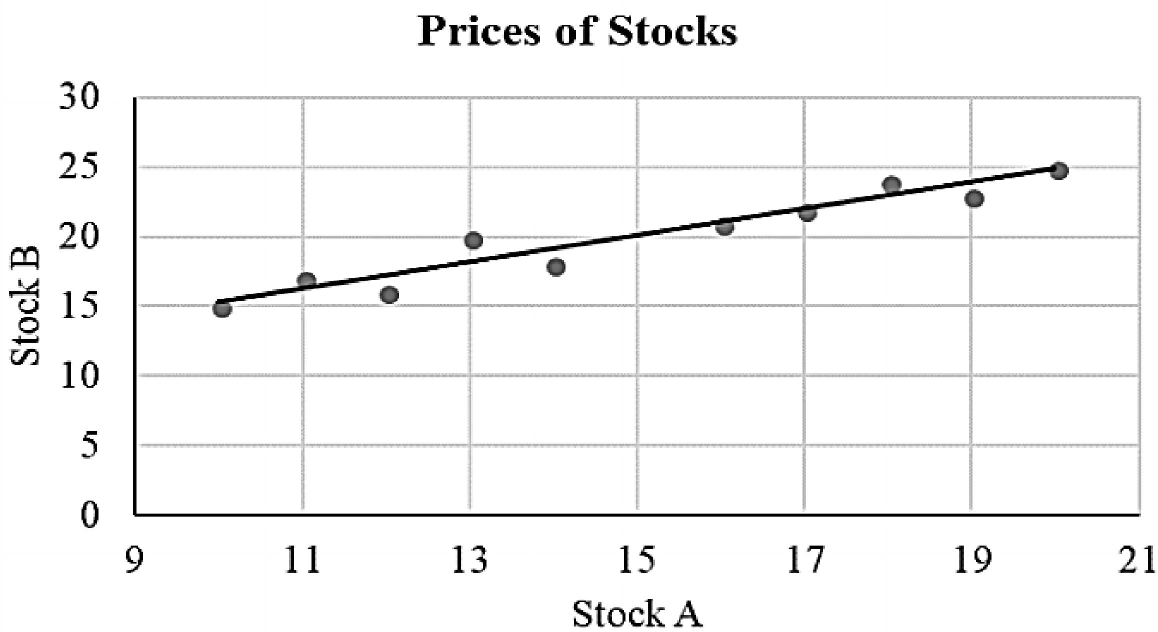


As previously stated, the points are x- y points of the given data. The point (110, 35) is the first point, and the (132, 56) is the last point.

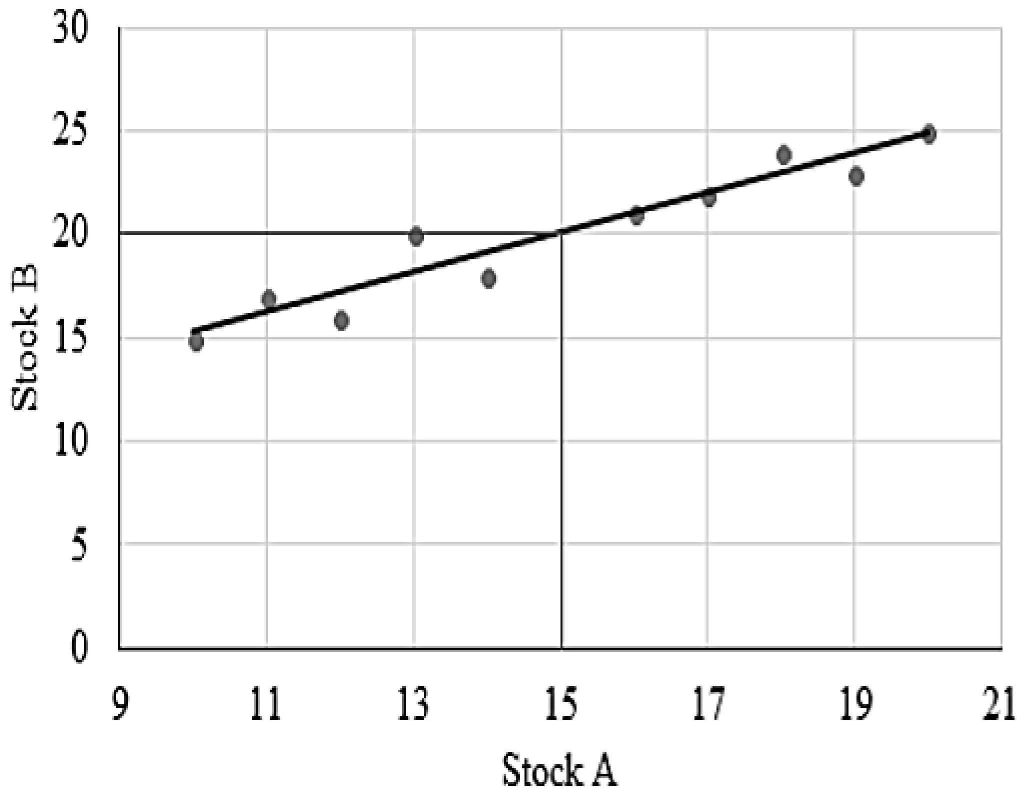
For the height of a student who weighs 44 kg (y-coordinate), we need to find the x-coordinate. The point lines up with **120 cm** of height.



63. A research on two stock companies reveals that the closing prices of stocks were positively correlated to each other. The following chart shows the stock prices of the companies and a line of best fit. Based on the chart, if the price of stock A is $15, what would the price of stock B be?



Solution
If the price of stock A is $15, the price of stock B is about **$20**.

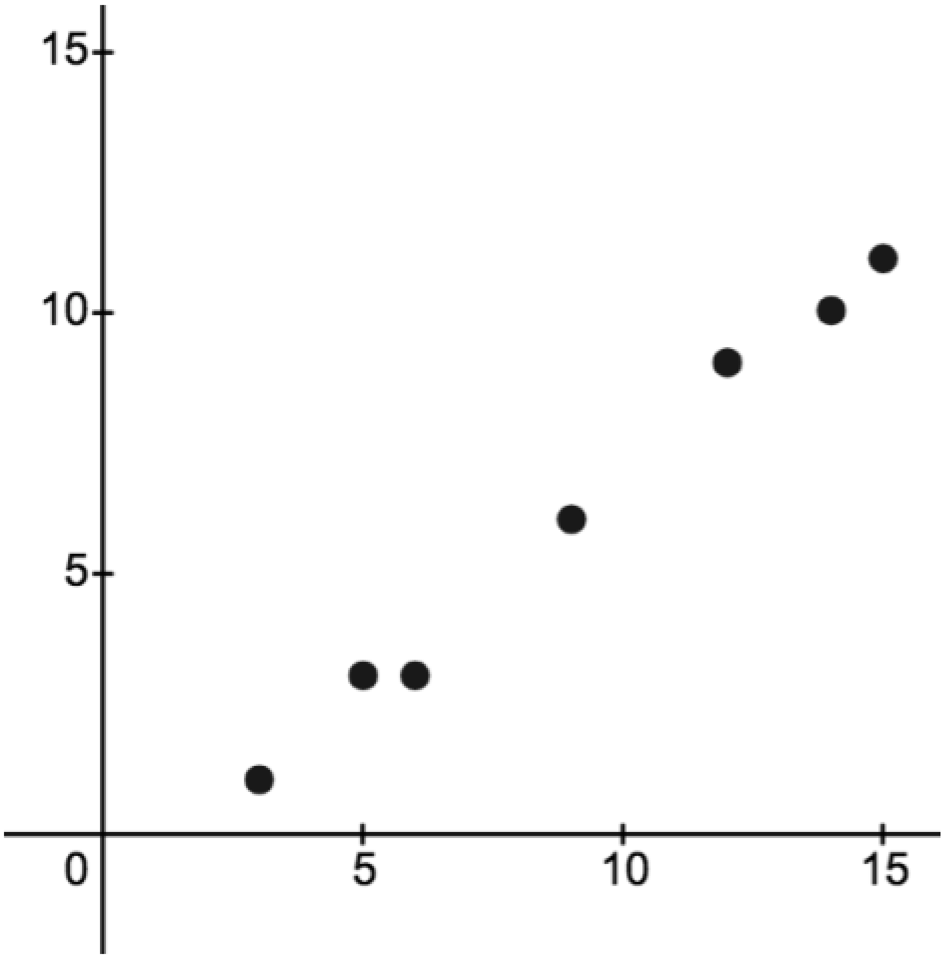


64. Ben started a new job as a car salesman. His supervisor gives him the advice that the more test drives per day he gets his customers to take the more sales he will make per day. He records the following data over the past week.

x (Number of Test Drives Per Day)	y (Number of Sales Per Day)
3	1
5	3
6	3
9	6
12	9
14	10
15	11

Solution

When we plot this set of data as a scatter plot we get the following graph.



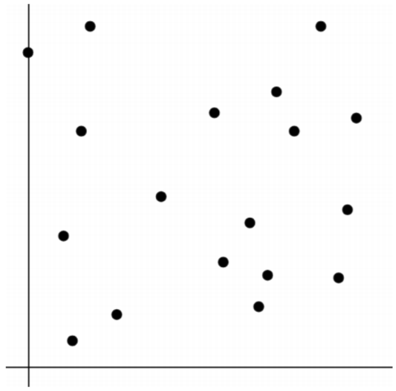
This clearly shows there is a relationship between the two variables. As x increases we see that y also increases. This shows there is what's called a positive linear correlation between the two variables.

Correlation and Scatter Plots

Specifically, linear correlation, is where there appears to be a linear relationship between the two variables. There are three types of correlation:

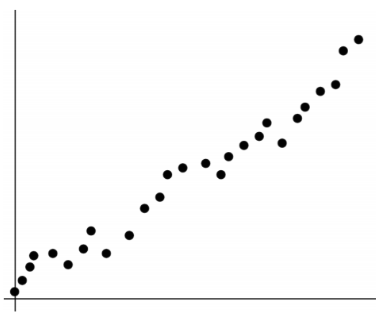
- Positive Correlation. As x increases y increases.
- Negative Correlation. As x increases y decreases.
- Zero Correlation. There is no relationship.

65. Determine if the following scatter plot exhibit positive linear correlation, negative linear correlation, or no linear correlation.



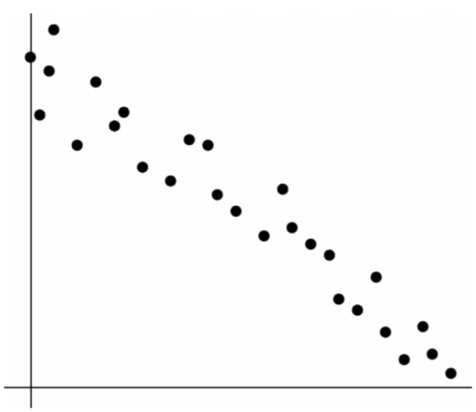
Solution: The graph shows no pattern. In this case there is no correlation.

66. Determine if the following scatter plot exhibit positive linear correlation, negative linear correlation, or no linear correlation.



Solution: The graph shows the pattern that as x increases, y increases. This is positive linear correlation.

67. Determine if the following scatter plot exhibit positive linear correlation, negative linear correlation, or no linear correlation.



Solution

The graph shows the pattern that as x increases, y decreases. This is negative linear correlation.

68. The table below shows the maximum exercise heart rate for specific individuals of various ages who exercise regularly.

Age, x, in years	20	25	30	32	43	55	28	42	50	55	62
Heart rate, y, in maximum beats per minute	160	150	148	145	140	130	155	140	132	125	125

a. Sketch a scatter diagram of the data.

b. Find a linear function that models the data.

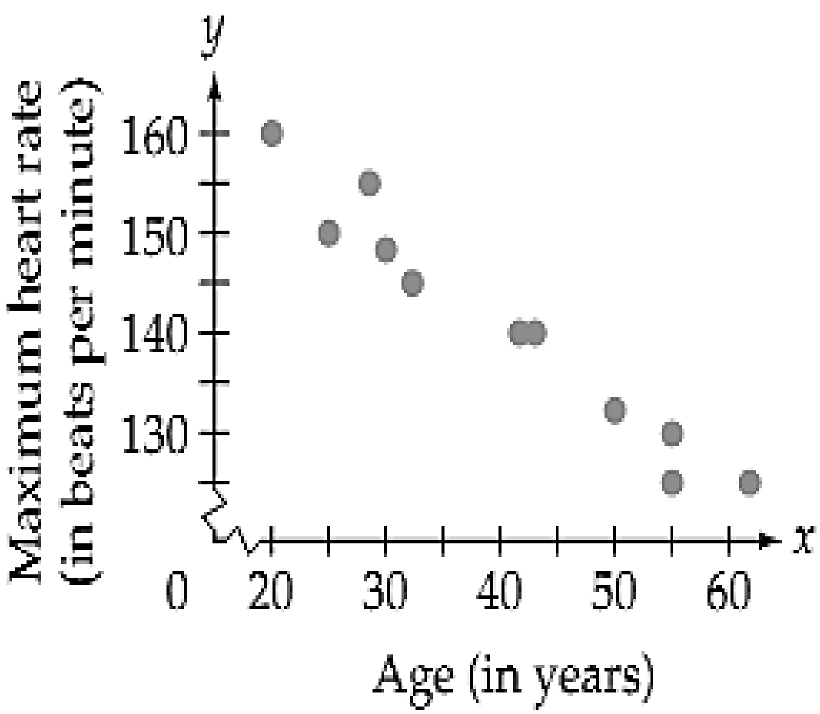
c. Use this function to predict the maximum exercise heart rate recommended for a 28-year-old person.

Solution

The table below shows the maximum exercise heart rate for specific individuals of various ages who exercise regularly.

Age, x, in years	20	25	30	32	43	55	28	42	50	55	62
Heart rate, y, in maximum beats per minute	160	150	148	145	140	130	155	140	132	125	125

The graph as follows, called a **scatter diagram**, is a graph of the ordered pairs of the table. These ordered pairs suggest that the maximum exercise heart rate for an individual decreases as the person's age increases.



Although these points do not lie on one line, it is possible to find a line that approximately fits the data. One way to do this is to select two data points and then find the equation of the line that passes through the two points. To do this, we first find the slope of the line between the two points and then use the point–slope formula to find the equation of the line. Suppose we choose $(20,160)$ as P_1 and $(62,125)$ as P_2.Then the slope of the line between P_1 and P_2 is

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{125 - 160}{62 - 20} = -\frac{35}{42} = -\frac{5}{6}$$

Now use the point–slope formula.

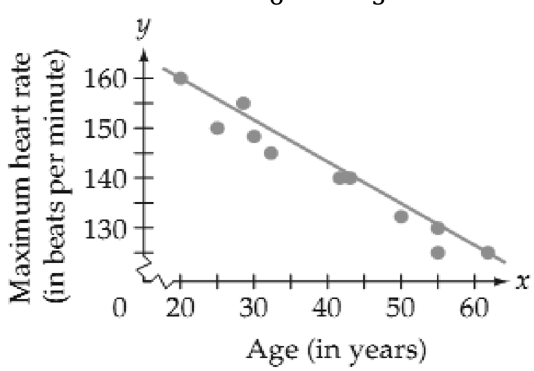
$$y - y_1 = m(x - x_1)$$

$$y - 160 = -\frac{5}{6}(x - 20)$$

$$y - 160 = -\frac{5}{6}x + \frac{50}{3}$$

$$y = -\frac{5}{6}x + \frac{530}{3}$$

The graph of $y = -\frac{5}{6}x + \frac{530}{3}$ is shown as follows.



This line approximates the data and can be used to estimate maximum exercise heart rates for different ages. For example, an exercise physiologist could determine the recommended maximum exercise heart rate for a 28-year-old individual by replacing in the equation by 28 and determining the value of y.

$$y = -\frac{5}{6}x + \frac{530}{3} \Rightarrow y = -\frac{5}{6}(28) + \frac{530}{3} \approx 153.3$$

The maximum exercise heart rate recommended for a 28-year-old person is approximately 153 beats per minute.

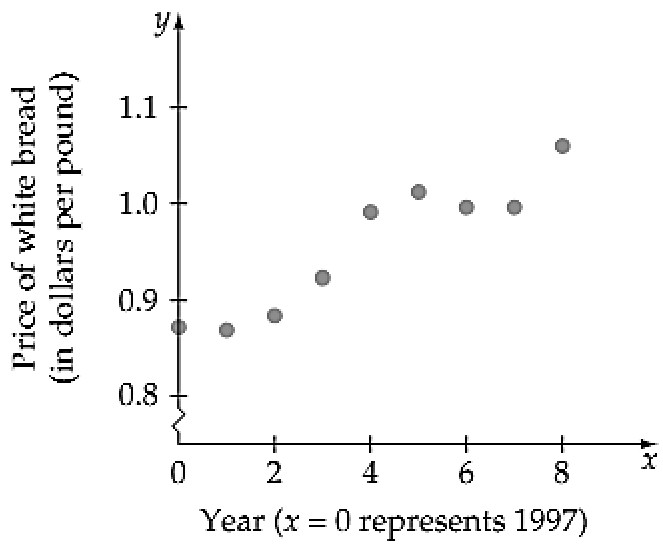
69. The average prices per pound of white bread in U.S. cities, as recorded in August of various years, are listed in the table below.

Year	1997	1998	1999	2000	2001	2002	2003	2004	2005
Price per pound	0.872	0.869	0.884	0.923	0.991	1.012	0.996	0.996	1.060

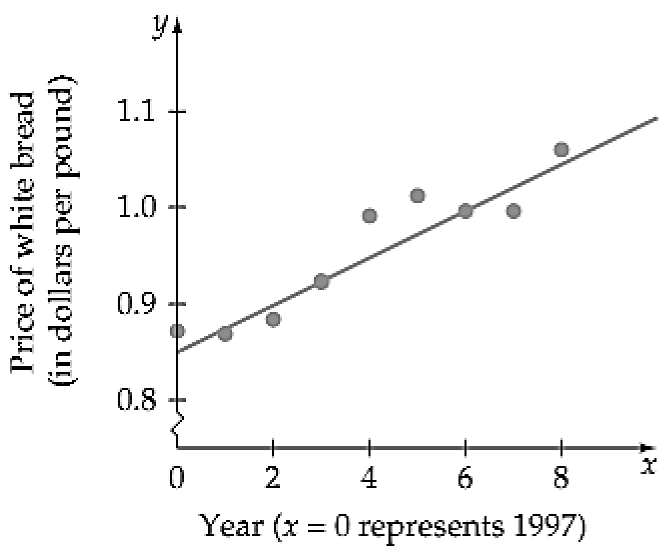
a. Sketch a scatter diagram of the data.
b. Find a linear function that models the data.
c. Use this function to predict the price per pound of white bread in August of 2011.

Solution

a. We can simplify the data by using 0 for 1997, 1 for 1998, etc.



Looking at the scatter diagram, it appears that a line through the points (3, 0.923) and (6, 0.996) will fit the data points reasonably well.



The slope of the line through these points is

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{0.996 - 0.923}{6 - 3} = \frac{0.073}{3}$$

and the equation of the line is

$$y - y_1 = m(x - x_1)$$

$$y - 0.923 = \frac{0.073}{3}(x - 3)$$

$$y - 0.923 = \frac{0.073}{3}x - 0.073$$

$$y = \frac{0.073}{3}x + 0.85$$

or approximately $y = 0.0243x + 0.85$. Thus the price per pound of white bread is approximately $y = 0.0243x + 0.85$, where is the number of years after August 1997. The year 2011 corresponds to x = 14, so we evaluate the function when x = 14.

$f(14) = 0.0243(14) + 0.85 = 1.1902$

The estimated price of white bread in August 2011 is about $1.19 per pound.

70. The populations of a city for various years are given in the table below.

Year	1992	1994	1996	1998	2000	2002	2004	2006
Population (thousands)	20.28	26.31	32.16	37.38	40.11	46.62	49.87	52.91

a. Sketch a scatter diagram of the data.

b. Write an equation for a linear function that models the data.

c. Use the function from part b to predict the city's population in the year 2025.

Solution

a. The regression equation is approximately

$y = 5.6333x - 252.8667$

b. When x = 63, we have $y = 5.6333(63) - 252.8667 \approx 102.$

The estimated weight of a woman swimmer who is 63 inches tall is approximately 102 pounds.

Variability and Measure of dispersion (already discussed in Reasoning – I)

Dispersion

The degree to which numerical data tend to spread about an average value is called the dispersion or variation of data.

Or Dispersion measures the spread or variability of data points within a dataset.

Types of Dispersion

There are two main types of the dispersion;

- **Absolute Dispersion:** It measures the variation among the values in the same unit of measurement in which the original data are given such as rupees, kg, inches etc. commonly used absolute measures are range, quartiles, mean, standard deviation and variance.
- **Relative Dispersion:** If we compare the dispersion of two dissimilar distributions we need relative term is called relative dispersion. **Or** The ratio between the measure of dispersion and corresponding measure of location is called relative dispersion. These measures are free of unit in which the original data is measure. Commonly used relative measures are coefficient of range, coefficient of quartiles, coefficient of mean and coefficient of variation.

Measure of Central Tendencies (Maximum discussed in Reasoning – I)

Central Tendency and Spread of Data

Central tendency is a statistical measure that identifies the middle or typical value of a dataset or distribution. It aims to provide a single value that best represents the entire dataset. It include mean, median, mode and mid ranges. Remaining topic or discussed here.

Percentiles/ p^{th} Percentiles

Percentiles are the values of the variate that divide a set of data into one hundred equal parts after arranging the observations in ascending order of magnitude.

Or

A value x is called the p^{th} percentile of a data set provided p% of the data values are less than x.

For ungrouped data

Position of $P_1 = \left(\frac{n+1}{100}\right)^{th}$ valu and Position of $P_i = \left[\frac{i(n+1)}{100}\right]^{th}$ value

P_i = Positional value + Decimal Part (difference)

For grouped data

$P_1 = l + \frac{h}{f}\left(\frac{n}{100} - C\right)$ and $P_i = l + \frac{h}{f}\left(\frac{in}{100} - C\right)$

Where

l = Lower class boundary of the model class.

h = Size of class interval of the model class.

f = Frequency of the model class.

n = Sum of the frequencies

C = Cumulative frequency of the preceding class of the model class.

71. Obtain D_{65} and D_{94} from the following data;

Class	Frequency	Class	Frequency
110 – 119	2	160 – 169	18
120 – 129	4	170 – 179	13
130 – 139	17	180 – 189	6
140 – 149	28	190 – 199	5
150 – 159	25	200 – 209	2

Solution

Class	Frequency	C.B	C.F
110 – 119	2	109.5 – 119.5	2
120 – 129	4	119.5 – 119.5	6
130 – 139	17	129.5 – 119.5	23
140 – 149	28	139.5 – 119.5	51
150 – 159	25	149.5 – 119.5	76
160 – 169	18	159.5 – 119.5	94
170 – 179	13	169.5 – 119.5	107
180 – 189	6	179.5 – 119.5	113
190 – 199	5	189.5 – 119.5	118
200 – 209	2	199.5 – 119.5	120
Sum	120		

65^{th} Percentile (P_{65})

$l = 159.5, f = 18, h = 10, C = 76$

$P_{65} = l + \frac{h}{f}\left(\frac{65n}{100} - C\right) = 160.61$

94^{th} Percentile (P_{94})

$l = 179.5, f = 6, h = 10, C = 107$

$P_{94} = l + \frac{h}{f}\left(\frac{94n}{100} - C\right) = 189.17$

72. According to the U.S. Department of Labor, the median annual salary in 2003 for a physical therapist was $57,720. If the 85th percentile for the annual salary of a physical therapist was $71,500, find the percent of physical therapists whose annual salaries were

a. more than $57,720.

b. less than $71,500.

c. between $57,720 and $71,500.

Solution

a. By definition, the median is the 50th percentile. Therefore, 50% of the physical therapists earned more than $57,720 per year.

b. Because $71,500 is the 85th percentile, 85% of all physical therapists made less than $71,500.

c. From parts a and b, $85\% - 50\% = 35\%$ of the physical therapists earned between $57,720 and $71,500.

73. According to the U.S. Department of Labor, the median annual salary in 2003 for a police dispatcher was $28,288. If the 30th percentile for the annual salary of a police dispatcher was $25,640, find the percent of police dispatchers whose annual salaries were

a. less than $28,288.

b. more than $25,640.

c. between $25,640 and $28,288.

Solution

a. By definition, the median is the 50th percentile. Therefore, 50% of the police dispatchers earned less than $28,288 per year.

b. Because $25,640 is the 30th percentile, $100\% - 30\% = 70\%$ of all police dispatchers made more than $25,640.

c. From parts a and b, $50\% - 30\% = 20\%$ of the police dispatchers earned between $25,640 and $28,288.

Percentile for a Given Data Value

Given a set of data and a data value x,

Percentile of score x = $\frac{\text{number of data values less than x}}{\text{total number of data values}} . 100$

74. On a reading examination given to 900 students, Elaine's score of 602 was higher than the scores of 576 of the students who took the examination. What is the percentile for Elaine's score?

Solution

$$\begin{aligned}\text{Percentile} &= \frac{\text{number of data values less than 602}}{\text{total number of data values}} \cdot 100 \\ &= \frac{576}{900} \cdot 100 \\ &= 64\end{aligned}$$

Elaine's score of 602 places her at the 64th percentile.

75. On an examination given to 8600 students, Hal's score of 405 was higher than the scores of 3952 of the students who took the examination. What is the percentile for Hal's score?

Solution

$$\begin{aligned}\text{Percentile} &= \frac{\text{number of data values less than 405}}{\text{total number of data values}} \cdot 100 \\ &= \frac{3952}{8600} \cdot 100 \\ &\approx 46\end{aligned}$$

Hal's score of 405 places him at the 46th percentile.

Deciles

These are the values, which divide the set of observations into ten equal parts after arranging the observations in ascending order of magnitude.

Formulae for deciles

For ungrouped data

Position of $D_1 = \left(\frac{n+1}{10}\right)^{th}$ value

Position of $D_2 = \left[\frac{2(n+1)}{10}\right]^{th}$ value

Position of $D_3 = \left[\frac{3(n+1)}{10}\right]^{th}$ value

Position of $D_5 = \left[\frac{5(n+1)}{10}\right]^{th}$ value

Position of $D_9 = \left[\frac{9(n+1)}{10}\right]^{th}$ value

$D_i = \text{Positional value} + \text{Decimal Part (difference)}$

76. Find 4^{th} and 7^{th} deciles from the following data;

17,22,27,29,38,40,42,45,50,54,56,57,60

Solution

4^{th} decile (D_4)

Position of $D_4 = \left[\frac{4(n+1)}{10}\right]^{th}$ value $= \left[\frac{4(13+1)}{10}\right]^{th}$ value $= (5.6)^{th}$value

$D_4 = \text{Positional value} + \text{Decimal Part (difference)}$

$D_4 = 5^{th} + 0.6\left(6^{th} \text{ value} - 5^{th} \text{ value}\right)$

$D_4 = 38 + 0.6\,(40 - 38) = 39.2$

7^{th} decile (D_7)

Position of $D_7 = \left[\frac{7(n+1)}{10}\right]^{th}$ value $= \left[\frac{7(13+1)}{10}\right]^{th}$ value $= (9.8)^{th}$value

$D_7 = \text{Positional value} + \text{Decimal Part (difference)}$

$D_7 = 9^{th} + 0.8\left(10^{th} \text{ value} - 9^{th} \text{ value}\right)$

$D_4 = 50 + 0.8\,(54 - 50) = 53.2$

For grouped data

$$D_1 = l + \frac{h}{f}\left(\frac{n}{10} - C\right)$$

$$D_i = l + \frac{h}{f}\left(\frac{in}{10} - C\right)$$

Where

l = Lower class boundary of the model class.

h = Size of class interval of the model class.

f = Frequency of the model class.

n = Sum of the frequencies

C = Cumulative frequency of the preceding class of the model class.

77. Obtain D_3 and D_7 from the following data;

Class	Frequency	Class	Frequency
110 – 119	2	160 – 169	18
120 – 129	4	170 – 179	13
130 – 139	17	180 – 189	6
140 – 149	28	190 – 199	5
150 – 159	25	200 – 209	2

Solution

Class	Frequency	C.B	C.F
110 – 119	2	109.5 – 119.5	2
120 – 129	4	119.5 – 119.5	6
130 – 139	17	129.5 – 119.5	23
140 – 149	28	139.5 – 119.5	51
150 – 159	25	149.5 – 119.5	76
160 – 169	18	159.5 – 119.5	94
170 – 179	13	169.5 – 119.5	107
180 – 189	6	179.5 – 119.5	113
190 – 199	5	189.5 – 119.5	118
200 – 209	2	199.5 – 119.5	120
Sum	120		

3rd Decile

$$l = 139.5, f = 28, h = 10, C = 23$$

$$D_3 = l + \frac{h}{f}\left(\frac{3n}{10} - C\right) = 144.14$$

7rd Decile

$$l = 159.5, f = 18, h = 10, C = 76$$

$$D_7 = l + \frac{h}{f}\left(\frac{7n}{10} - C\right) = 163.94$$

Quartiles

Quartiles are the values of the variate that divide a set of data into four equal parts after arranging the observations in ascending order of magnitude.

Quartile Deviation/ Semi Inter – Quartile Range

It is also called semi inter – quartile range. The SIQR is the measure of dispersion defined by the difference between third quartile and the first quartile and half of the range is called quartile deviation. The QD is also an absolute measure of dispersion. Its relative measure called coefficient of quartile deviation.

Formulae

$Q.D = \frac{Q_3 - Q_1}{2}$

Where $Q_3 = \left(\frac{3(n+1)}{4}\right)^{th}$ and $Q_1 = \left(\frac{n+1}{4}\right)^{th}$

Coefficient of Q.D $= \frac{Q_3 - Q_1}{Q_3 + Q_1}$

78. Find QD of 45,32,21,65,36,53,48,76,27.

Solution

Arrange: 21,27,32,36,45,48,53,65,76

$Q_1 = \left(\frac{n+1}{4}\right)^{th} = \left(\frac{9+1}{4}\right)^{th} = \left(\frac{10}{4}\right)^{th} = (2.5)^{th} = 27.5$

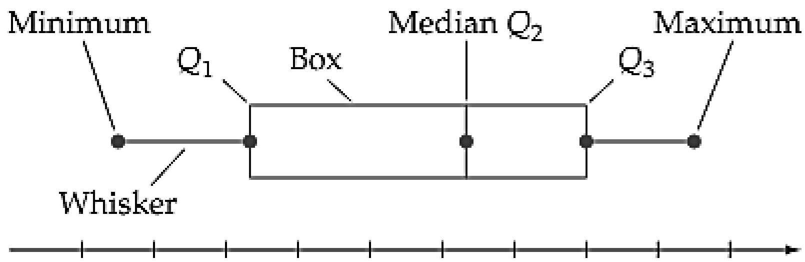
$Q_3 = \left(\frac{3(n+1)}{4}\right)^{th} = \left(\frac{3(9+1)}{4}\right)^{th} \left(\frac{3(10)}{4}\right)^{th} = \left(\frac{30}{4}\right)^{th} = (7.5)^{th} = 53.5$

$Q.D = \frac{Q_3 - Q_1}{2} = \frac{53.5 - 27.5}{2} = \frac{26}{2} = 13$

Coefficient of Q.D $= \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{53.5 - 27.5}{53.5 + 27.5} = \frac{26}{81} = 0.32$

Box-and-Whisker Plots

A box-and-whisker plot (sometimes called a **box plot**) is often used to provide a visual summary of a set of data. A box-and-whisker plot shows the median, the first and third quartiles, and the minimum and maximum values of a data set. See the figure below.



Construction of a Box-and-Whisker Plot

1. Draw a horizontal scale that extends from the minimum data value to the maximum data value.

Above the scale, draw a rectangle (box) with its left side at Q_1 and its right side at Q_3

2. Draw a vertical line segment across the rectangle at the median, Q_2
3. Draw a horizontal line segment, called a whisker, that extends from Q_1 to the minimum and another whisker that extends from Q_3 to the maximum.

Importance of a Box-and-Whisker Plot

Box plots have become popular because they are easy to construct and they illustrate several important features of a data set in a simple diagram. That is we can easily estimate

- the quartiles of the data.
- the range of the data.
- the position of the middle half of the data as shown by the length of the box.

79. The following table lists the calories per 100 milliliters of 25 popular beers. Find the quartiles of the data and Construct a box-and-whisker plot for the data set.

Calories, per 100 Milliliters, of Selected Beer

43	37	42	40	53	62	36	32	50	49
26	53	73	48	45	39	45	48	40	56
41	36	58	42	39					

NOTE: It is important to start a box plot with a scaled number line. Otherwise the box plot may not be useful.

Solution

Step 1: Rank the data, as shown in the following table.

1) 26	**2)** 32	**3)** 36	**4)** 36	**5)** 37	**6)** 39	**7)** 39	**8)** 40	**9)** 40
10) 41	**11)** 42	**12)** 42	**13)** 43	**14)** 45	**15)** 45	**16)** 48	**17)** 48	**18)** 49
19) 50	**20)** 53	**21)** 53	**22)** 56	**23)** 58	**24)** 62	**25)** 73		

Step 2: The median of these 25 data values has a rank of 13. Thus the median is 43. The second quartile Q_2 is the median of the data, so $Q_2 = 43$

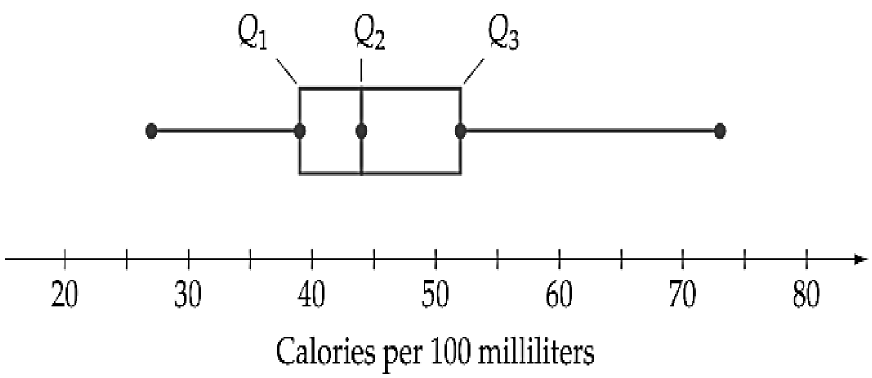
Step 3: There are 12 data values less than the median and 12 data values greater than the median. The first quartile is the median of the data values less than the median. Thus Q_1 is the mean of the data values with ranks of 6 and 7.

$Q_1 = \frac{39+39}{2} = 39$

The third quartile is the median of the data values greater than the median. Thus Q_3 is the mean of the data values with ranks of 19 and 20.

$Q_3 = \frac{50+53}{2} = 51.5$

For the data set, we determined that $Q_1 = 39, Q_2 = 43$ and $Q_3 = 51.5$. The minimum data value for the data set is 26 and the maximum data value is 73. Thus the box-and-whisker plot is as shown in Figure.



80. Graph a box-and-whisker plot for the data values shown.

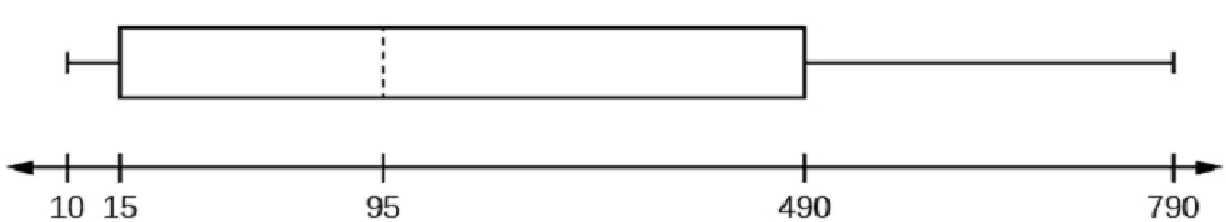
10; 10; 10; 15; 35; 75; 90; 95; 100; 175; 420; 490; 515; 515; 790

The five numbers used to create a box-and-whisker plot are:

Min: 10, Q_1: 15, Med: 95, Q_3: 490, Max: 790

Solution

The following graph shows the box-and-whisker plot.



81. Construct a box-and-whisker plot for the following data.

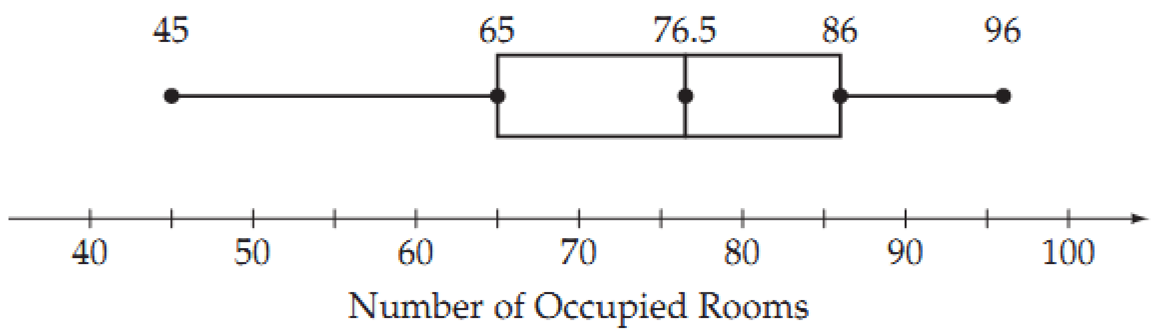
The Numbers of Occupied Rooms in a Resort during an 18-Day Period

86	77	58	45	94	96	83	76	75
65	68	72	78	85	87	92	55	61

Solution

For the data set, we determined that

$Q_1 = 65, Q_2 = 76.5$ and $Q_3 = 86$.



82. Construct box plot for the dataset.

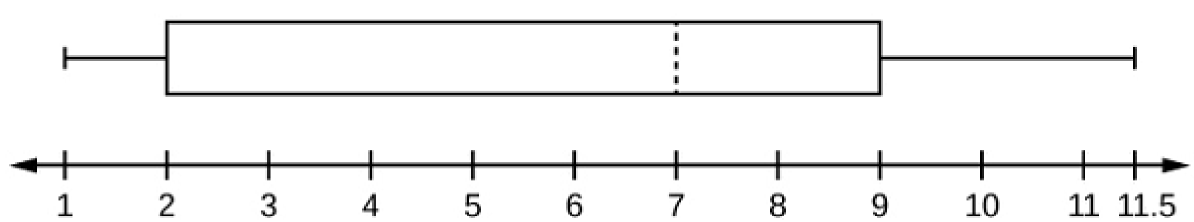
1; 1; 2; 2; 4; 6; 6.8; 7.2; 8; 8.3; 9; 10; 10; 11.5

Solution

Consider the dataset.

1; 1; 2; 2; 4; 6; 6.8; 7.2; 8; 8.3; 9; 10; 10; 11.5

The first quartile is two, the median is seven, and the third quartile is nine. The smallest value is one, and the largest value is 11.5. The following image shows the constructed box plot.



The two whiskers extend from the first quartile to the smallest value and from the third quartile to the largest value. The median is shown with a dashed line.

83. Test scores for a college statistics class held during the day are:

99; 56; 78; 55.5; 32; 90; 80; 81; 56; 59; 45; 77; 84.5; 84; 70; 72; 68; 32; 79; 90

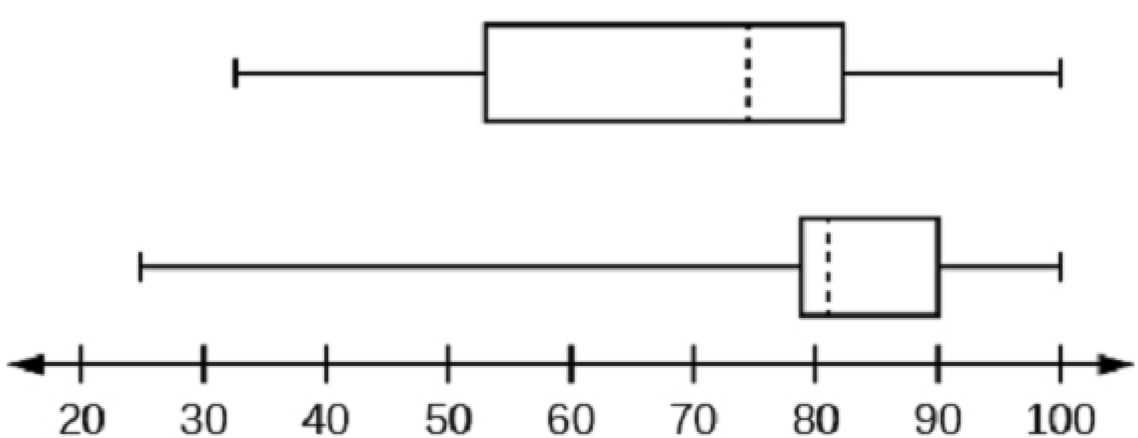
Test scores for a college statistics class held during the evening are:

98; 78; 68; 83; 81; 89; 88; 76; 65; 45; 98; 90; 80; 84.5; 85; 79; 78; 98; 90; 79; 81; 25.5

a. Find the smallest and largest values, the median, and the first and third quartile for the day class.
b. Find the smallest and largest values, the median, and the first and third quartile for the night class.
c. For each data set, what percentage of the data is between the smallest value and the first quartile? the first quartile and the median? the median and the third quartile? The third quartile and the largest value? What percentage of the data is between the first quartile and the largest value?
d. Create a box plot for each set of data. Use one number line for both box plots.
e. Which box plot has the widest spread for the middle 50% of the data (the data between the first and third quartiles)? What does this mean for that set of data in comparison to the other set of data?

Solution

a. Min = 32, Q_1 = 56, M = 74.5, Q_3 = 82.5, Max = 99
b. Min = 25.5, Q_1 = 78, M = 81, Q_3 = 89, Max = 98
c. Day class: There are six data values ranging from 32 to 56: 30%. There are six data values ranging from 56 to 74.5: 30%. There are five data values ranging from 74.5 to 82.5: 25%. There are five data values ranging from 82.5 to 99: 25%. There are 16 data values between the first quartile, 56, and the largest value, 99: 75%.
Night class:



d.
e. The first data set has the wider spread for the middle 50% of the data. The IQR for the first data set is greater than the IQR for the second set. This means that there is more variability in the middle 50% of the first data set.

Variance and Standard Deviation

Variance: Variance is defined as the mean of the squared deviation of $x_i; (i = 1,2,3, \dots, n)$ observations from their arithmetic mean.

Standard Deviation: Standard Deviation is defined as positive square root of the mean of the squared deviation of $x_i; (i = 1,2,3, \dots, n)$ observations from their arithmetic mean

Standard Deviation is the measure of how far, on average, the data is from the mean.

Another related measure, is the **Variance** which is standard deviation squared.

The standard deviation and variance for a **SAMPLE** are calculated by the following symbols and formulas:

- **Variance** $= s^2 = \frac{\sum(X-\bar{X})^2}{n}$ for ungroup data
- **Variance** $= s^2 = \frac{\sum f(X-\bar{X})^2}{\sum f}$ for group data
- **Standard Deviation** $= s = \sqrt{\frac{\sum(X-\bar{X})^2}{n}}$ for ungroup data
- **Standard Deviation** $= s = \sqrt{\frac{\sum f(X-\bar{X})^2}{\sum f}}$ for group data

The standard deviation and variance for a **POPULATION** are calculated by the following symbols and formulas:

- **Variance** $= \sigma^2 = \frac{\sum(X-\mu)^2}{n}$
- **Standard Deviation** $= \sigma = \sqrt{\frac{\sum(X-\mu)^2}{n}}$

Computational formulae of Variance and Standard Deviation:

- **Variance** $= s^2 = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2$ for ungroup data
- **Variance** $= s^2 = \frac{\sum fX^2}{\sum f} - \left(\frac{\sum fX}{\sum f}\right)^2$ for group data
- **Standard Deviation** $= s = \sqrt{\frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2}$ for ungroup data
- **Standard Deviation** $= s = \sqrt{\frac{\sum fX^2}{\sum f} - \left(\frac{\sum fX}{\sum f}\right)^2}$ for group data

84. The following numbers were obtained by sampling a population.

2, 4, 7, 12, 15

Find the standard deviation of the sample.

Solution

$$\bar{X} = \frac{\sum X}{n} = \frac{32 + 4 + 7 + 12 + 15}{5} = \frac{40}{5} = 8$$

x	$x - \bar{x}$	$(x - \bar{x})^2$
2	$2 - 8 = -6$	$(-6)^2 = 36$
4	$4 - 8 = -4$	$(-4)^2 = 16$
7	$7 - 8 = -1$	$(-1)^2 = 1$
12	$12 - 8 = 4$	$4^2 = 16$
15	$15 - 8 = 7$	$7^2 = 49$
		118 ← The sum of the squared deviations

Variance $= s^2 = \frac{\sum(X - \bar{X})^2}{n} = \frac{118}{5} = 23.6$

Standard Deviation $= s = \sqrt{\frac{\sum(X - \bar{X})^2}{n}} = \sqrt{23.6} = 4.858$

85. The marks of six students in Mathematics are as follows. Determine variance and standard deviation.

Student No.	1	2	3	4	5	6
Marks	60	70	30	90	80	42

Solution

Let X = marks of a student.

$\bar{X} = \frac{\sum X}{n} = \frac{372}{6} = 62$

X	X^2	$X - \bar{X}$	$(X - \bar{X})^2$
60	3600	– 2	4
70	4900	8	64
30	900	– 32	1024
90	8100	28	784
80	6400	18	324
42	1764	– 20	400
$\sum X = 372$	$\sum X^2 = 25664$	$\sum(X - \bar{X}) = 0$	$\sum(X - \bar{X})^2 = 2600$

$\textbf{Variance} = s^2 = \frac{\sum(X-\bar{X})^2}{n} \approx 433.3333$

$\textbf{computational Variance} = s^2 = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2 \approx 433.3333$

$\textbf{Standard Deviation} = s = \sqrt{\frac{\sum(X-\bar{X})^2}{n}} \approx 20.81666$

$\textbf{computational Standard Deviation} = s = \sqrt{\frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2} \approx 20.81666$

86. For the following data showing weights of toffee boxes in gm. Determine variance and standard deviation.

X	4.5	14.5	24.5	34.5	44.5	54.5	64.5
f	2	10	5	9	6	7	1

Solution

Let X = marks of a student.

$\bar{X} = \frac{\sum X}{\sum f} = \frac{241.5}{40} = 6.0376$

X	f	$X - \bar{X}$	$(X - \bar{X})^2$	$f(X - \bar{X})^2$	fX	fX^2
4.5	2	– 28	784	1568	9	40.5
14.5	10	-18	324	3240	145	2102.5
24.5	5	– 8	64	320	122.5	3001.5
34.5	9	2	4	36	310.5	10712.25
44.5	6	12	144	864	267	11881.5
54.5	7	22	484	3388	381.5	20791.75
64.5	1	32	1024	1024	64.5	4160.25

Variance $= s^2 = \frac{\sum f(X-\bar{X})^2}{\sum f} = \frac{10600}{40} = 265$

computational Variance $= s^2 = \frac{\sum fX^2}{\sum f} - \left(\frac{\sum fX}{\sum f}\right)^2 = \frac{52690}{40} - \left(\frac{1300}{40}\right)^2 = 261$

Standard Deviation $= s = \sqrt{261} = 16.155$

computational Standard Deviation $= s = \sqrt{261} = 16.155$

87. A consumer group has tested a sample of eight size D batteries from each of three companies. The results of the tests are shown in the following table. According to these tests, which company produces batteries for which the values representing hours of constant use have the smallest standard deviation?

Company	Hours of Constant Use per Battery
EverSoBright	6.2, 6.4, 7.1, 5.9, 8.3, 5.3, 7.5, 9.3
Dependable	6.8, 6.2, 7.2, 5.9, 7.0, 7.4, 7.3, 8.2
Beacon	6.1, 6.6, 7.3, 5.7, 7.1, 7.6, 7.1, 8.5

Solution

The mean for each sample of batteries is 7 hours.

The batteries from Ever So Bright have a standard deviation of

$$\textbf{Standard Deviation} = s = \sqrt{\frac{\sum(X-\bar{X})^2}{n}} = \sqrt{\frac{12.34}{7}} \approx 1.328 \text{ hours}$$

The batteries from Dependable have a standard deviation of

$$\textbf{Standard Deviation} = s = \sqrt{\frac{\sum(X-\bar{X})^2}{n}} = \sqrt{\frac{3.62}{7}} \approx 0.719 \text{ hours}$$

The batteries from Beacon have a standard deviation of

$$\textbf{Standard Deviation} = s = \sqrt{\frac{\sum(X-\bar{X})^2}{n}} = \sqrt{\frac{5.38}{7}} \approx 0.877 \text{ hours}$$

The batteries from Dependable have the smallest standard deviation. According to these results, the Dependable company produces the most consistent batteries with regard to life expectancy under constant use.

88. Can the variance of a data set be smaller than the standard deviation of the data set?

Answer

Yes. The variance is smaller than the standard deviation whenever the standard deviation is less than 1.

Sampling Techniques

Sampling techniques are methods used to select a representative subset of data from a larger population. It is used in almost every field of life. Some examples are as follows;

- A cook taste a bit of cooked food to find whether it has been properly cooked or not.
- A food inspector takes a sample of food or items like milk, flour etc. to find whether they are pure or not.
- Cement, steel and bricks are examined before using them in different places.

Below are some important terminologies we will be using in sampling.

- **Data** is the collection of all observations for a particular variable or variables, from one more people or things.
- A **Population** is the collection of all individuals or items under consideration in a study **or** the totality of individuals is called population. **e.g.** Total number of absent students, number of colour TV sets, Monthly salaries of all employees, number of computers sold out. Total number of objects in a population is called **population size**.
- A **Sample** is the part of a population from which information is actually collected. **e.g.** wheat yield per acre for 5 pieces of land. Total number of objects in a sample is called **sample size**.
- **Sampling** is the process of selecting a small portion of the population which represent all the characteristics of the population.
- **Sample Survey** is the collection of information from a representative part of the population. It is carried out by an experimental design.
- A **Census** is information (data) obtained from the entire population.
- A **Parameter** is a numerical measurement describing some characteristic of a population. Such as mean, median or standard deviation calculated from the population.
  Examples: The average starting salary of elementary school teachers in Georgia is $33,673.
  The average for the whole United States is $35,763.
- A **Statistic** is a numerical measurement describing some characteristic of a sample. Example: A survey of ten job postings for elementary school teachers in the Atlanta area, had an average starting salary of $38,541.
- **Sampling Error** is the difference between the sample static and the population parameter is called sampling error. i.e. $\in = t - \theta$, where t is sample static and θ is corresponding population parameter.
- **Standard Error** is the standard deviation of sampling distribution of any static.
- **Bias** is the difference between the expected value of the sample static and the true value of the population parameter. i.e. $B = \in(t) - \theta$, where t is sample static used to estimate the population static θ.

Some useful Formulae

- **Formulae of Parameters;**

  Population Mean $= \mu = \frac{\sum x}{N}$

  Population Variance $= \sigma^2 = \frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2$

  Population Proportion $= \pi = \frac{x}{N}$
- **Formulae of Static;**

  Sample Mean $= \bar{x} = \frac{\sum x}{n}$

  Sample Variance $= s^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2$

  Sample Proportion $= p = \frac{x}{n}$

Sampling with replacement

Sampling is said to be with replacement when from a finite population, a sampling unit is drawn, observed and then returned to the population before another unit is drawn. The population in this case remains the same and a sampling unit might be selected more than once.

Formula to find number of samples, when sampling is done with replacement is $\mathbf{m = N^n}$ where m is possible number of samples, N is population size and n is sample size.

Results for Sampling Distribution of Means (with replacement)

- **Formulae of Parameters;**

  Mean of Means or Mean of Sampling Distribution of Means $= \mu_{\bar{x}} = \sum \bar{x} f(\bar{x})$

  Also we may use in need: $\mu_{\bar{x}} = \mu = \frac{\sum x}{N}$

  Standard Error of Means $= \sigma_{\bar{x}} = \sqrt{\sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2}$

  Standard Error of Sampling Distribution of Means $= \sigma_{\bar{x}} = \sqrt{\sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2}$

  Also we may use in need: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2}}{\sqrt{n}}$

  Variance of Means $= \sigma_{\bar{x}}^2 = \sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2$

  Variance of Sampling Distribution of Means $= \sigma_{\bar{x}}^2 = \sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2$

  Also we may use in need: $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2}{n}$

89. A population consists of 3,7,11,15,19. Take all possible sample of size 2 with replacement. Form the sampling distribution of sample mean $\bar{x}$. Find its means and variance. Compare it with population mean and variance.

Solution

$x = 3,7,11,15,19$ $N = 5$ $n = 2$ with replacement

Population Mean $= \mu = \frac{\sum x}{N} = \frac{55}{5} = 11$ ………… (i)

x	3	7	11	15	19
x^2	9	49	121	225	361

Population Variance $= \sigma^2 = \frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2 = 32$

$\frac{\sigma^2}{n} = \frac{32}{2} = 16$ ………… (ii)

Possible samples of size 2 WR $= m = N^n = 5^2 = 25$

Samples	$\bar{x}$	**Samples**	$\bar{x}$
3,3	3	7,3	5
3,7	5	7,7	7
3,11	7	7,11	9
3,15	9	7,15	11
3,19	11	7,19	13
Samples	$\bar{x}$	**Samples**	$\bar{x}$
11,3	7	19,3	11
11,7	9	19,7	13
11,11	11	19,11	15
11,15	13	19,15	17
11,19	15	19,19	19

Sampling distribution of sample mean $\bar{x}$ is

$\bar{x}$	f	$f(\bar{x})$	$\bar{x}f(\bar{x})$	$\bar{x}^2 f(\bar{x})$
3	1	1/25	3/25	9/25
5	2	2/25	10/25	50/25
7	3	3/25	21/25	147/25
9	4	4/25	36/25	324/25
11	5	5/25	55/25	605/25
13	4	4/25	52/25	676/25
15	3	3/25	45/25	675/25
17	2	2/25	34/25	578/25
19	1	1/25	19/25	361/25
total	25	1	275/25	3425/25

Mean of Means $= \mu_{\bar{x}} = \sum \bar{x} f(\bar{x}) = \frac{275}{25} = 11$ ………… (iii)

From (i) and (iii) $\boldsymbol{\mu_{\bar{x}} = \mu}$

Variance of Means $= \sigma_{\bar{x}}^2 = \sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2 = \frac{3425}{25} - (11)^2$

$\sigma_{\bar{x}}^2 = 16$ ………… (iv)

From (ii) and (iv) $\boldsymbol{\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}}$

90. A population consists of 4,6,8. Take all possible sample of size 3 with replacement. Form the sampling distribution of sample mean $\bar{x}$. Calculate mean and standard error of mean. Verify the results with population mean and standard deviation.

Solution

x = 4,6,8 N = 3 n = 3 with replacement

Population Mean $= \mu = \frac{\sum x}{N} = \frac{18}{3} = 6$ ………… (i)

x	4	6	8	**18**
x^2	16	36	64	**116**

Population Variance $= \sigma = \sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2} = 1.63$

$\frac{\sigma}{\sqrt{n}} = \frac{1.63}{\sqrt{3}} = 0.94$ ………… (ii)

Possible samples of size 3 WR = $m = N^n = 3^3 = 27$

Samples	$\bar{x}$	**Samples**	$\bar{x}$
4,4,4	4	6,4,4	4.67
4,4,6	4.67	6,4,6	5.33
4,4,8	5.33	6,4,8	6
4,6,4	4.67	6,6,4	5.33
4,6,6	5.33	6,6,6	6
4,6,8	6	6,6,8	6.67
4,8,4	5.33	6,8,4	6
4,8,6	6	6,8,6	6.67
4,8,8	4.67	6,8,8	7.33

Samples	$\bar{x}$
8,4,4	5.33
8,4,6	6
8,4,8	6.67
8,6,4	6
8,6,6	6.67
8,6,8	7.33
8,8,4	6.67
8,8,6	7.33
8,8,8	8

Sampling distribution of sample mean $\bar{x}$ is

$\bar{x}$	f	$f(\bar{x})$	$\bar{x}f(\bar{x})$	$\bar{x}^2 f(\bar{x})$
4	1	1/27	4/27	16/27
4.67	3	3/27	14.01/27	65.4267/27
5.33	6	6/27	31.98/27	170.4534/27
6	7	7/27	42/27	252/27
6.67	6	6/27	40.02/27	266.9334/27
7.33	3	3/27	21.99/27	161.1867/27
8	1	1/27	8/27	64/27
total	27	1	162/27	996.0002/27

Mean of Means $= \mu_{\bar{x}} = \sum \bar{x} f(\bar{x}) = \frac{162}{27} = 6$ ………… (iii)

From (i) and (iii) $\boldsymbol{\mu_{\bar{x}} = \mu}$

Standard Error of Means $= \sigma_{\bar{x}} = \sqrt{\sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2} = \sqrt{\frac{996.0002}{27} - (6)^2}$

$\sigma_{\bar{x}} = 0.94$ ………… (iv)

From (ii) and (iv) $\boldsymbol{\sigma_{\bar{x}} = \frac{\sigma^2}{n}}$

Sampling without replacement

If the sample is taken without replacement from a finite population, the selected element is not returned to the population before drawing the next element. In without replacement sampling an element can be selected only once.

Formula to find number of samples, when sampling is done without replacement is

m= $^{N}C_{n}$ where m is possible number of samples, N is population size and n is sample size.

Results for Sampling Distribution of Means (with replacement)

- **Formulae of Parameters;**

Mean of Means or Mean of Sampling Distribution of Means $= \mu_{\bar{x}} = \sum \bar{x} f(\bar{x})$

Also we may use in need: $\mu_{\bar{x}} = \mu = \frac{\sum x}{N}$

Standard Error of Means $= \sigma_{\bar{x}} = \sqrt{\sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2}$

Standard Error of Sampling Distribution of Means $= \sigma_{\bar{x}} = \sqrt{\sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2}$

Also we may use in need: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{\sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

Variance of Means $= \sigma_{\bar{x}}^2 = \sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2$

Variance of Sampling Distribution of Means $= \sigma_{\bar{x}}^2 = \sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2$

Also we may use in need: $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2}{n}\left(\frac{N-n}{N-1}\right)$

91. A population consists of 6 members 2,4,6,8,10 and 12. Take all possible sample of size 2 without replacement. Form the sampling distribution of means. Find its means and variance. Compare it with population mean and variance.

Solution

x = 2,4,6,8,10,12 N = 6 n = 2 WOR

Population Mean $= \mu = \frac{\sum x}{N} = \frac{42}{6} = 7$ ………… (i)

x	2	4	6	8	10	12
x^2	4	16	36	64	100	144

Population Variance $= \sigma^2 = \frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2 = 11.67$

$\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{11.67}{2}\left(\frac{6-2}{6-1}\right) = 4.67$ ………… (ii)

Possible samples of size 2 WOR = $m = {}^{N}C_n = {}^{6}C_2 = 15$

Samples	$\bar{x}$	**Samples**	$\bar{x}$
2,4	3	4,6	5
2,6	4	4,8	6
2,8	5	4,10	7
2,10	6	4,12	8
2,12	7	6,8	7

Samples	$\bar{x}$
6,10	8
8	9
8,10	9
8,12	10
10,12	11

Sampling distribution of sample mean $\bar{x}$ is

$\bar{x}$	f	$f(\bar{x})$	$\bar{x}f(\bar{x})$	$\bar{x}^2 f(\bar{x})$
3	1	1/15	3/15	9/15
4	1	1/15	4/15	16/15
5	2	2/15	10/15	50/15
6	2	2/15	12/15	72/15
7	3	3/15	21/15	147/15
8	2	2/15	16/15	128/15
9	2	2/15	18/15	162/15
10	1	1/15	10/15	100/15
11	1	1/15	11/15	121/15
total	15	1	105/15	805/15

Mean of Means $= \mu_{\bar{x}} = \sum \bar{x} f(\bar{x}) = \frac{105}{15} = 7$ ………… (iii)

From (i) and (iii) $\boldsymbol{\mu_{\bar{x}} = \mu}$

Variance of Means $= \sigma_{\bar{x}}^2 = \sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2 = \frac{805}{15} - (7)^2$

$\sigma_{\bar{x}}^2 = 46.7$ ………… (iv)

From (ii) and (iv) $\boldsymbol{\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)}$

92. A population consists of 6 members 3,6,9,12,15 and 18. Take all possible sample of size 3 without replacement. Form the sampling distribution of means. Find its means and standard deviation. Also find the standard error.

Solution

x = 3,6,9,12,15,18 N = 6 n = 3 WOR

Population Mean $= \mu = \frac{\sum x}{N} = \frac{63}{6} = 10.5$

x	3	6	9	12	15	18
x^2	9	36	81	144	225	324

Population S.D $= \sigma = \sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2} = 5.12$

Possible samples of size 3 WOR = $m = {}^{N}C_3 = {}^{6}C_3 = 20$

Samples	$\bar{x}$	**Samples**	$\bar{x}$
3,6,9	6	6,9,12	9
3,6,12	7	6,9,15	10
3,6,15	8	6,9,18	11
3,6,18	9	6,12,15	11
3,9,12	8	6,12,18	12
3,9,15	9	6,15,18	13
3,9,18	10	9,12,15	12
3,12,15	10	9,12,18	13
3,12,18	11	9,15,18	14
3,15,18	12	12, 15,18	15

Sampling distribution of sample mean $\bar{x}$ is

$\bar{x}$	f	$f(\bar{x})$	$\bar{x}f(\bar{x})$	$\bar{x}^2 f(\bar{x})$
6	1	1/20	6/20	36/20
7	1	1/20	7/20	49/20
8	2	2/20	16/20	128/20
9	3	3/20	27/20	243/20
10	3	3/20	30/20	300/20
11	3	3/20	33/20	363/20
12	3	3/20	36/20	432/20
13	2	2/20	26/20	338/20
14	1	1/20	14/20	196/20
15	1	1/20	15/20	225/20
total	20	1	210/20	2310/20

Mean of Means $= \mu_{\bar{x}} = \sum \bar{x} f(\bar{x}) = \frac{210}{20} = 10.5$

Standard Error of Means $= \sigma_{\bar{x}} = \sqrt{\sum \bar{x}^2 f(\bar{x}) - (\mu_{\bar{x}})^2}$

$\sigma_{\bar{x}} = \sqrt{\frac{2310}{20} - (10.5)^2} = 2.29$

Estimation

Estimation is a procedure by which we obtain the value of unknown population parameters by using the sample information. It is the process of approximating a population parameter (e.g., mean, proportion, variance) using sample data. It involves making an educated guess about a population characteristic based on a representative subset of data.

Types of Estimation

Estimation is divided into two types;

- **Point Estimation:** The process of finding a single value from the sample.
- **Interval Estimation:** The process of finding a range of values within which the population parameter is expected to lie with a certain degree of confidence.

Point Estimate

A single numerical value calculated from the sample.

Point Estimator

The rule or formula that is used to estimate a population parameter. Its important characteristics are unbiasedness, consistency, efficiency and sufficiency.

Estimate

An estimate is defined as numerical values of the unknown population parameter obtained by apply an estimator. Estimate is divided into two types;

- **Point Estimate:** It is a single numerical value from the sample.
- **Interval Estimate:** It is a range of values within which the population parameter is expected to lie with a certain degree of confidence.

Degree of Freedom

The term degree of freedom is defined as the number of independent or freely chosen variables.

Central Limit Theorem

The distribution of the means of a large number of samples of size taken from a population is approximately a normal distribution.

Confidence Intervals

A **confidence interval** is a type of estimate but, instead of being just one number, it is an interval of numbers. It provides a range of reasonable values in which we expect the population parameter to fall. There is no guarantee that a given confidence interval does capture the parameter, but there is a predictable probability of success. Confidence intervals are based on the Central Limit Theorem and the hypothesis testing equations.

Let $(1 - \alpha)$ be a specified high probability and L and U be the functions of sample observations $X_1, X_2, X_3, \dots, X_n$ such that:

$$P(L < \theta < U) = 1 - \alpha \quad ; 0 < \alpha < 1$$

Then the interval (L, U) is called a $100(1 - \alpha)\%$ confidence interval for the parameter θ.

Level of Confidence

The probability of accepting a true null hypothesis is called level of hypothesis. It is denoted by $(1 - \alpha)$.

Level of Significance

The probability of rejecting a null hypothesis when it is actually true is called level of significance. It is denoted by α.

Test Statistic

A test statistics is a function or formula of sample observations that provides a basis for testing a null hypothesis. The most commonly used test statistics are Z and T – Tests.

Test Significance

Test of significance is a procedure which enables us, on the basis of sampling distribution, whether to accept or reject a hypothesis.

T-Test/ Small Sample Test

A t-test is a statistical hypothesis test used to determine if there's a significant difference between the means of two groups. It's commonly used for:

1. Comparing two independent samples (e.g., control vs. treatment).

2. Comparing a sample mean to a known population mean.

3. Testing the significance of regression coefficients.

Types of T-Tests

1. One – sample T-test.

2. Independent samples T-test (two-sample t-test).

3. Paired samples T-test.

Z –Test/ Z – Score Test

A Z-test is a statistical hypothesis test used to determine if a sample mean is significantly different from a known population mean. It's commonly used when:

1. Sample size is large (n > 30).

2. Population standard deviation is known.

3. Data is normally distributed.

Key differences between T-Test and Z-Test

1. Sample size: T-test is used for smaller samples (n < 30), while Z-test is used for larger samples.

2. Population standard deviation: T-test estimates standard deviation from sample data, while Z-test requires known population standard deviation.

3. Distribution: T-test assumes normal distribution, while Z-test assumes normal distribution and large sample size.

Student T-Test vs. T-Test

Student's T-test and T-test are often used interchangeably. However, technically:

1. Student's T-test refers specifically to the test developed by William Sealy Gosset (under the pseudonym "Student") for small samples.

2. T-test is a broader term encompassing various types of T-tests.

In practice, the terms are used synonymously, and the distinction is often ignored.

93. A random sample selected from a normal population with mean μ and variance σ^2 gave the values 25,31,23,33,28,36,22,26. Give the point estimator for μ and σ^2 and find their point estimates.

Solution

x	x^2
25	625
31	961
23	529
33	1089
28	784
39	1296
22	484
26	676
224	6444

Point estimator of population mean $\mu = \bar{x} = \frac{\sum x}{n} = \frac{224}{8} = 28$

Point estimate of population mean μ is 28.

Point estimator of population variance $\sigma^2 = \hat{s}^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \frac{6444-8(28)^2}{8-1} = 24.57$

Point estimate of population variance σ^2 is 24.57.

Z – values for Commonly used Confidence Levels

Confidence Level	Area	Z value
90%	0.0500 and 0.9500	1.64 or 1.65
95%	0.0250 and 0.9750	1.96
96%	0.0200 and 0.9800	2.05
97%	0.0150 and 0.9850	2.17
98%	0.0100 and 0.9900	2.33
99%	0.0050 and 0.9950	2.57 or 2.58

94. Confidence Interval for Mean (Z – test): A normal population has a variance of 100. A random sample of size 16 selected from the population has mean of 52.50. Construct the 90% confidence interval estimate of the population mean μ. Interpret the result.

Solution

$n = 16, \sigma^2 = 100, \sigma = 10, \bar{x} = 52.50, 90\%CI = ?$

$1 - \alpha = 90\% \Rightarrow \alpha = 1 - 0.90 \Rightarrow \alpha = 0.10 \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow Z_{\frac{\alpha}{2}} = Z_{0.05} = 1.645$

$90\%CI = \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 52.50 \pm (1.645)\frac{10}{\sqrt{16}} = 52.50 \pm 4.11$

$90\%CI = 52.50 - 4.11 = 48.39$

$90\%CI = 52.50 + 4.11 = 56.64$

Hence 90% confidence interval for population mean μ obtained from the observed sample is $(48.39, 56.64)$.

95. Confidence Interval for Mean (Z – test): A particular component in a transistor circuit has a lifetime which is known to follow a skew distribution. A random sample of 250 components from a week's production given an average lifetime of 840 hours, and the variance of lifetime is 483(Hours^2). Find approximately 95% confidence limits to the true mean lifetime in the whole population of the product.

Solution

$$n = 250, \hat{s}^2 = 483, \hat{s} = \sqrt{483} = 21.98, \bar{x} = 840, 95\%CI = ?$$

$$1 - \alpha = 95\% \Rightarrow \alpha = 1 - 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$$

$$95\%CI = \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\hat{s}}{\sqrt{n}} = 840 \pm (1.96)\frac{21.98}{\sqrt{250}} = 840 \pm 2.72$$

$$95\%CI = 840 - 2.72 = 837.28$$

$$95\%CI = 840 + 2.72 = 842.72$$

Hence 95% confidence interval for population mean μ obtained from the observed sample is (837.28,842.72).

96. Confidence Interval for Mean (Z – test): A random sample of size n = 200 selected without replacement from a finite population of size N = 1000 with σ = 1.28 showed that $\bar{x} = 68.60$. Construct a 97% confidence interval for the mean of the population.

Solution

$$n = 200, N = 1000, \sigma = 1.28, \bar{x} = 68.60, 97\%CI = ?$$

$$1 - \alpha = 97\% \Rightarrow \alpha = 1 - 0.97 \Rightarrow \alpha = 0.03 \Rightarrow \frac{\alpha}{2} = 0.015 \Rightarrow Z_{\frac{\alpha}{2}} = Z_{0.015} = 2.17$$

$$97\%CI = \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = 68.60 \pm (2.17)\frac{1.28}{\sqrt{220}}\sqrt{\frac{1000-200}{1000-1}} = 68.60 \pm 0.17$$

$$97\%CI = 68.60 - 0.17 = 68.43$$

$$97\%CI = 68.60 + 0.17 = 68.77$$

Hence 97% confidence interval for population mean μ obtained from the observed sample is (68.43,68.77).

97. Confidence Interval for Mean (t – test): Ten packets of a particular brand of biscuits are chosen at random and their mass measured in grams. The results are;
$n = 10, x = 3978.70, \sum x^2 = 1583098.30$
assuming that the sample is taken from a normal population with mean mass μ calculate the 98% confidence interval for μ.

Solution

$$n = 10, x = 3978.70, \sum x^2 = 1583098.30, v = n - 1 = 10 - 1 = 9, 98\%CI = ?$$

$$\bar{x} = \frac{\sum x}{n} = \frac{3978.70}{10} = 397.87, \hat{s} = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n-1}} = 3.21$$

$$1 - \alpha = 98\% \Rightarrow \alpha = 1 - 0.98 \Rightarrow \alpha = 0.02 \Rightarrow \frac{\alpha}{2} = 0.01 \Rightarrow t_{\frac{\alpha}{2}(v)} = t_{0.01(9)} = 2.821$$

$$98\%CI = \bar{x} \pm t_{\frac{\alpha}{2}(v)} \frac{\hat{s}}{\sqrt{n}} = 397.87 \pm (2.821)\frac{3.21}{\sqrt{10}} = 397.87 \pm 2.86$$

$$98\%CI = 397.87 - 2.86 = 395.01$$

$$98\%CI = 397.87 + 2.86 = 400.73$$

Hence 98% confidence interval for population mean μ obtained from the observed sample is (395.01,400.73).

98. Confidence Interval for Mean (t – test): A random sample of eight observations of a normal variable gave $\sum x = 261.20, \sum(x - \bar{x})^2 = 3.22$. Calculate a 95% confidence interval for the population mean μ.

Solution

$$n = 8, \hat{s} = ?, \bar{x} = ?, v = n - 1 = 8 - 1 = 7, \ 95\%CI = ?$$

$$\bar{x} = \frac{\sum x}{n} = \frac{261.20}{8} = 32.65, \hat{s} = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = 0.68$$

$$1 - \alpha = 95\% \Rightarrow \alpha = 1 - 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow t_{\frac{\alpha}{2}(v)} = t_{0.025(7)} = 2.365$$

$$95\%CI = \bar{x} \pm t_{\frac{\alpha}{2}(v)} \frac{\hat{s}}{\sqrt{n}} = 32.65 \pm (2.365)\frac{0.68}{\sqrt{8}} = 32.65 \pm 0.57$$

$$95\%CI = 32.65 - 0.57 = 32.08$$

$$95\%CI = 32.65 + 0.57 = 33.22$$

Hence 95% confidence interval for population mean μ obtained from the observed sample is (32.08,33.22).

Hypothesis

Any statement which may or may not be true is called hypothesis.

Hypothesis Testing

It is a procedure which enables us to decide on the basis of information obtained from sample data whether to accept or reject any specified statement or hypothesis or assumption about the value of population parameter.

Statistical Hypothesis

A statistical hypothesis is a statement about one or more parameter of a population. This statement may or may not be true. Its validity is tested on the basis of sample obtained from the population.

Null Hypothesis

Any hypothesis which is tested for possible rejection under the assumption that it is true is called null hypothesis. It is generally denoted by $\mathbf{H_0}$.

Alternative Hypothesis

Any hypothesis which is different from the null hypothesis. It is accepted when null hypothesis is rejected. It is generally denoted by $\mathbf{H_A}$ or $\mathbf{H_1}$.

Types of Null Hypothesis and Alternative Hypothesis

Null Hypothesis $(\mathbf{H_0})$	Alternative Hypothesis $(\mathbf{H_1})$
$\theta = \theta_0$	$\theta \neq \theta_0$
$\theta \leq \theta_0$	$\theta > \theta_0$
$\theta \geq \theta_0$	$\theta < \theta_0$

Simple Hypothesis

A hypothesis in which all parameter of the distribution are specified is called simple hypothesis. For example, if the average age of ICS students is 16 year. i.e. $H_0 = \mu = 16$ is a simple hypothesis.

Composite Hypothesis

A hypothesis in which all parameter of the distribution are not specified is called composite hypothesis. For example, $H_1 = \mu < 16$ or $H_1 = \mu > 16$ years are composite hypothesis.

Power of a Test

If we reject a false null hypothesis, it is called power of a test. It is denoted by $(1-\beta)$.

True Situation	Decision	
	Accept H_0	Reject H_0
H_0 is true	Correct decision Or level of confidence $= 1-\alpha$	Wrong decision= α
H_0 is false	Wrong decision= β	Correct decision $= 1-\beta$

Critical Values

The values of test statistic which separates the rejection and acceptance region are called as critical values.

Critical Region/ Rejection Region

Critical region is the part of sampling distribution of a statistics which leads to the rejection of the null hypothesis.

Acceptance Region

Acceptance region is the part of sampling distribution of a statistics which leads to the acceptance of the null hypothesis.

Some useful Formulae

- Z – test for Hypothesis Testing: $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$
- T – test for Hypothesis Testing: $t = \frac{\bar{x}-\mu}{\frac{\hat{s}}{\sqrt{n}}}$

Choice of Test for Testing of Hypothesis

Variance	Sample Size	Test	Use Variance in Formula
σ^2 known	$n < 30$ or $n > 30$	Z – test	σ^2
σ^2 unknown	$n > 30$	Z – test	$\hat{s}^2$
σ^2 unknown	$n < 30$	t – test	$\hat{s}^2$

General Procedure of testing of Hypothesis (or Null Hypothesis)

The procedure for testing a hypothesis about population parameter involves the following steps;

- State your problem and formulate appropriate null hypothesis H_0 with an alternative hypothesis H_1.
- Decide upon a level of significance α of the test, which is the probability of rejecting the null hypothesis when H_0 is true.
- Choose an appropriate test static.
- Calculate the value of test static from sample data.
- Determine the critical region depends upon the alternative hypothesis H_1.
- Make a conclusion. That is, if the value of test static falls in the critical region then reject H_0 and if the value of test static falls in the accepting region then accept H_0.

Two Tailed Test

If the critical region is located equally in both tails of the sampling distribution of test statistic, the test is called two tailed test. It is also called two sided test.

One Tailed Test

If the critical region is located equally in only one tail of the sampling distribution of test statistic, the test is called one tailed test. It is also called one sided test.

Testing of Hypothesis using Z – test

- Formulate a hypothesis as
  $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ (two sided or two tailed test)
  $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$ (one sided or one tailed test)
  $H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$ (one sided or one tailed test)
- Decide level of significance $\alpha = 0.05$ (generally)
- Use formula $z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$
- Calculate the value of test static from sample data.
- **Critical Regions:**

Variance	Sample Size
$H_1: \mu \neq \mu_0$	$\lvert Z \rvert \geq Z_{\frac{\alpha}{2}}$
$H_1: \mu > \mu_0$	$Z > +Z_{\alpha}$
$H_1: \mu < \mu_0$	$Z < -Z_{\alpha}$

- **Conclusions:**
  If value of test static falls in the critical region then reject H_0
  If value of test static falls in the acceptance region then accept H_0

Testing of Hypothesis using t – test

- Formulate a hypothesis as
  $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ (two sided or two tailed test)
  $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$ (one sided or one tailed test)
  $H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$ (one sided or one tailed test)
- Decide level of significance $\alpha = 0.05$ (generally)
- Use formula $t = \frac{\bar{x}-\mu_0}{\frac{\hat{s}}{\sqrt{n}}}$
- Calculate the value of test static from sample data.
- **Critical Regions:**

Variance	Sample Size
$H_1: \mu \neq \mu_0$	$\lvert t \rvert \geq t_{\frac{\alpha}{2}(v)}$
$H_1: \mu > \mu_0$	$t > +t_{\alpha(v)}$
$H_1: \mu < \mu_0$	$t < -t_{\alpha(v)}$

- **Conclusions:**
  If value of test static falls in the critical region then reject H_0
  If value of test static falls in the acceptance region then accept H_0

Z – scores

The z-score for a given data value x is the number of standard deviations that x is above or below the mean of the data. The following formulas show how to calculate the z-score for a data value x in a population and in a sample.

Population: $z_x = \frac{x-\mu}{\sigma}$ Sample: $z_x = \frac{x-\bar{x}}{s}$

99. Must the z-score for a data value be a positive number?

Answer

No. The z-score for a data value x is positive if x is greater than the mean, it is 0 if x is equal to the mean, and it is negative if x is less than the mean.

100. Raul has taken two tests in his chemistry class. He scored 72 on the first test, for which the mean of all scores was 65 and the standard deviation was 8. He received a 60 on the second test, for which the mean of all scores was 45 and the standard deviation was 12. In comparison with the other students, did Raul do better on the first test or the second test?

Solution

Find the z-score for each test.

$$z_{72} = \frac{72 - 65}{8} = 0.875 \qquad z_{60} = \frac{60 - 45}{12} = 1.25$$

Raul scored 0.875 standard deviation above the mean on the first test and 1.25 standard deviations above the mean on the second test. These z-scores indicate that in comparison with his classmates, Raul scored better on the second test than he did on the first test.

101. Cheryl has taken two quizzes in her history class. She scored 15 on the first quiz, for which the mean of all scores was 12 and the standard deviation was 2.4. Her score on the second quiz, for which the mean of all scores was 11 and the standard deviation was 2.0, was 14. In comparison with her classmates, did Cheryl do better on the first quiz or the second quiz?

Solution

$$z_{15} = \frac{15 - 12}{2.4} = 1.25 \qquad z_{14} = \frac{14 - 11}{2.0} = 1.5$$

These z-scores indicate that in comparison with her classmates, Cheryl did better on the second quiz than she did on the first quiz.

102. A consumer group tested a sample of 100 light bulbs. It found that the mean life expectancy of the bulbs was 842 hours, with a standard deviation of 90. One particular light bulb from the Dura Bright Company had a z-score of 1.2.What was the life span of this light bulb?

Solution

Substitute the given values into the z-score equation and solve for x.

$$z_x = \frac{x - \bar{x}}{s}$$

$$1.2 = \frac{x - 842}{90}$$

$$108 = x - 842$$

$$950 = x$$

The light bulb had a life span of 950 hours.

103. Roland received a score of 70 on a test for which the mean score was 65.5. Roland has learned that the z-score for his test is 0.6.What is the standard deviation for this set of test scores?

Solution

$$z_x = \frac{x - \mu}{\sigma}$$

$$0.6 = \frac{70 - 65.5}{\sigma}$$

$$\sigma = \frac{4.5}{0.6} = 7.5$$

The standard deviation for this set of test scores is 7.5.

104. The mean life time of electric bulbs produce by a company has in the past been 1120 hours with a standard deviation of 125 hours. A sample of 100 electric bulbs recently chosen from a supply of newly produced bulbs showed a mean life time of 1070 hours. Test the hypothesis that the mean life time of bulbs has not changed, using 5 % level of significance.

Solution

$H_0 = \mu = 1120$ and $H_1 = \mu \neq 1120$

Level of significance: $\alpha = 5\% = 0.05$

$\mu = 1120, \sigma = 125, n = 100, \bar{x} = 1070$

Test Statistic: $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = -4$ implies $|Z| = 4$

Critical Region: Reject H_0 if $|Z| > Z_{\frac{\alpha}{2}} = 1.96$

Since the calculated value of Z lies in critical region, so we reject H_0 and conclude that the mean life time of bulbs has changed.

105. It has been found from experience that the mean breaking strength of thread is 9.63N with a standard deviation of 1.40N. Recently a sample of 36 pieces of thread showed a mean breaking strength of 8.93N. Can we conclude at 1 % level of significance that thread has become inferior?

Solution

$H_0 = \mu = 9.63$ and $H_1 = \mu < 9.63$

Level of significance: $\alpha = 1\% = 0.01$

$\mu = 9.63, \sigma = 1.40, n = 36, \bar{x} = 8.93$

Test Statistic: $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = -3$ implies $|Z| = 3$

Critical Region: Reject H_0 if $|Z| < -Z_{\alpha} = -2.33$

Since the calculated value of Z lies in critical region, so we reject H_0 and conclude that the thread has become inferior.

106. The breaking strength of cables produced by a company has a mean of 1800 pounds and standard deviation of 100 pounds. By a new technique in production process, it is claimed that breaking strength can be increased. To test this claim, a random sample of 50 cables is tested and it is found that the mean breaking strength is 1850 pounds. Can we support claim at 0.01 significance level?

Solution

$H_0 = \mu = 1800$ and $H_1 = \mu \neq 1800$

Level of significance: $\alpha = 1\% = 0.01$

$\mu = 1800, \sigma = 100, n = 50, \bar{x} = 1850$

Test Statistic: $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = 3.54$

Critical Region: Reject H_0 if $|Z| > Z_{\frac{\alpha}{2}} = 2.33$

Since the calculated value of Z lies in critical region, so we reject H_0 and conclude that the claim should be supported.

107. A company claims that the average amount of coffee it supplies in jars is 6.0 oz with a standard deviation of 0.2 oz. a random sample of 100 jars is selected and average is found to be 5.9. Is the company cheating the customers? Use 5% level of significance.

Solution

$H_0: \mu \geq 6.0$ and $H_1: \mu < 6.0$

Level of significance: $\alpha = 5\% = 0.05$

$\mu = 6.0, \sigma = 0.2, n = 100, \bar{x} = 5.9$

Test Statistic: $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = -5$

Critical Region: $Z < -Z_\alpha = Z_{0.05} = -1.645$

Since the calculated value of Z lies in critical region, so we reject H_0.

108. A timber company is interested in seeing if the number of board feet per tree has decreased since moving to a new location of timber. In the past, the company has an average of 93 board feet per tree. The company believes that the production has decreased since changing locations, a random sample of 25 trees yields $\bar{x} = 89$ and $\hat{s} = 20$. Assuming the normality of the data, test the hypothesis at a 10% level of significance.

Solution

$H_0: \mu \geq 93$ and $H_1: \mu < 93$

Level of significance: $\alpha = 5\% = 0.05$

$\mu = 93, \hat{s} = 20, n = 25, \bar{x} = 89, v = n - 1 = 25 - 1 = 24$

Test Statistic: $t = \frac{\bar{x}-\mu}{\hat{s}/\sqrt{n}} = -1$

Critical Region: $t < -t_{\alpha(v)} = -t_{0.10(24)} = -1.318$

Since the calculated value of t lies in acceptance region, so we accept H_0.

109. Ten cartons are taken at random from an automatic filling machine. The mean net weight of the 10 cartons is 15.90 oz and the sum of squared deviation from this mean is 0.276 $(oz)^2$. Does the sample mean differ significantly from intended weight of 16 oz?

Solution

$H_0: \mu = 16$ and $H_1: \mu \neq 16$

Level of significance: $\alpha = 5\% = 0.05$

$\mu = 16, n = 10, \bar{x} = 15.90, v = n - 1 = 10 - 1 = 9$

$\sum(x - \bar{x})^2 = 0.276,\ \hat{s} = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = 0.18$

Test Statistic: $t = \frac{\bar{x}-\mu}{\hat{s}/\sqrt{n}} = -1$

Critical Region: $|t| \geq t_{\frac{\alpha}{2}(v)} = t_{0.025(9)} = 2.262$

Since the calculated value of t lies in acceptance region, so we accept H_0.

Exercise

1) Marks obtained by 60 students of a class are given below;
   60,50,46,28,58,64,36,20,50,18,42,56,20,38,40,34,24,64,64,42,46,52,50,44,36,0, 24,30,46,40,64,40,36,14,36,8,56,40,30,36,24,22,36,5058,16,40,34,0,42,42,0,36, 18,18,68,30,46,38,16.
   Make frequency distribution using appropriate class interval.
2) For data given below;
   1.36,1.46,1.50,1.32,1.45,1.24,1.49,1.64,1.47,1.59,1.41,1.48,1.36,1.48,1.51,1.45, 1.26,1.38,1.76,1.63,1.19,1.56,1.65,1.54,1.61,1.73,1.60,1.50,1.45,1.76,1.67,1.35, 1.55,1.68,1.46,1.40,1.32,1.47,1.64,1.45
   Make frequency distribution taking 0.05 as class interval and 1.19 as the lowest class limit.
3) Draw the histogram for the following frequency distribution;

Classes	frequency	Classes	frequency
110 – 119	2	160 – 119	18
120 – 129	4	170 – 179	13
130 – 139	17	180 – 189	6
140 – 149	28	190 – 199	5
150 – 159	25	200 – 209	2

4) Draw the histogram and a polygon for the following frequency distribution;

X	14	16	18	20	22	24
f	20	22	30	25	13	4

5) The heights of college students are given below;

Heights	14	16	18	20	22	24
Students	20	22	30	25	13	4

   Draw a histogram and ogive.
6) The weights of 50 football players are listed below;

193	240	217	283	268	212	251	263	275	208
230	288	259	225	252	230	243	247	280	234
250	236	277	218	245	268	231	269	224	259
258	231	255	228	202	245	246	271	249	255
265	235	243	219	255	245	238	257	254	284

   Make a stem and leaf display for the data and convert it to a frequency table with 10 class beginning with 190. Also make its frequency distribution.

7) The data are the distances (in kilometers) from a home to local supermarkets. Create a stemplot using the data: 1.1; 1.5; 2.3; 2.5; 2.7; 3.2; 3.3; 3.3; 3.5; 3.8; 4.0; 4.2; 4.5; 4.5; 4.7; 4.8; 5.5; 5.6; 6.5; 6.7; 12.3
Do the data seem to have any concentration of values?
HINT: The leaves are to the right of the decimal.

8) The following data show the distances (in miles) from the homes of off-campus statistics students to the college. Create a stem plot using the data and identify any outliers:
0.5; 0.7; 1.1; 1.2; 1.2; 1.3; 1.3; 1.5; 1.5; 1.7; 1.7; 1.8; 1.9; 2.0; 2.2; 2.5; 2.6; 2.8; 2.8; 2.8; 3.5; 3.8; 4.4; 4.8; 4.9; 5.2; 5.5; 5.7; 5.8; 8.0

9) Draw simple bar chart to represent the production of commodity "A" during the years 2000 to 2008:

Years	**Product**	**Years**	**Product**
2000	115	2005	145
2001	113	2006	190
2002	110	2007	210
2003	135	2008	258
2004	100		

10) The following table shows disability in simple population;

Type of disability	**Number of Persons**
Blind	13
Deaf and dumb	26
Crippled	41
Other handicapped	33

Draw a simple bar chart.

11) The following table gives the birth and death rates per thousand of few countries. Represent this data by multiple bar chart;

Country	**Birth rate**	**Death rate**
India	33	24
Japan	32	19
Germany	16	10
Egypt	44	24
Australia	20	9
New Zealand	18	8
France	21	16
Russia	38	16

12) The table shows that quantities in hundreds of Kg of wheat, barley and oats produced on certain farm during the year 1971 – 75;

Years	**Wheat**	**Barley**	**Oats**
1971	34	18	27
1972	43	14	24
1973	43	16	27
1974	45	16	27
1975	50	13	34

Construct the percentage component bar chart to illustrate the data.
Also draw a multiple bar chart.

13) Represent the data by a Pie Chart;

Districts	LHR	MTN	RWP	DGK
Area	50	115	135	165

14) Represent the data by a Pie Chart;

Items	Food	Clothing	House Rent	Edu	Misc.
Expenditure	75	50	30	25	20

15) Obtain the equation of regression line Y on X between the given values;

X	78	77	85	88	83	83	82
Y	84	80	82	83	88	90	88
X	78	76	83	97	98		
Y	91	83	89	78	96		

16) Obtain the equation of regression line X on Y between the given values;

X	78	77	85	88	83	83	82
Y	84	80	82	83	88	90	88
X	78	76	83	97	98		
Y	91	83	89	78	96		

17) Price indices of cotton X and wool Y are given below for the 12 months of a year. Obtain the correlation coefficient between X and Y and obtain the equation of the lines of regression between indices.

X	78	77	85	88	83	83	82
Y	84	80	82	83	88	90	88
X	78	76	83	97	98		
Y	91	83	89	78	96		

18) Calculate the correlation coefficient between percentage of marks scored by 12 students in statistics X and economics Y.

x	50	54	56	59	60	61
y	22	25	34	28	26	30
x	62	65	67	71	71	74
y	32	30	28	36	36	60

19) Calculate the correlation coefficient between supply and demand from the following data;

x	400	200	700	100	500	300	600
y	60	30	70	10	40	20	52

20) Obtain D_3, D_7 and D_8 from the following data;

i. 127,113,132,128,125,130,119,117,121

ii. 121,115,79,52,102,126,81,65,109,119,115,121,103,75,59,110

21) Find2nd, 3rd, 4th , 5th, 6th, 7th and 8th deciles from the following data;

Class	Frequency	Class	Frequency
10 – 20	7	50 – 60	18
20 – 30	10	60 – 70	10
30 – 40	16	70 – 80	5
40 – 50	24	80 – 90	5

22) Obtain P_{38}, P_{45}, P_{67} and P_{86} from the following data;

i. 127,113,132,128,125,130,119,117,121

ii. 121,115,79,52,102,126,81,65,109,119,115,121,103,75,59,110

23) Find 47th and 83th percentiles from the following data;

Class	Frequency	Class	Frequency
15 – 30	2	111 – 126	15
31 – 46	5	127 – 142	11
47 – 62	9	143 – 158	8
63 – 78	13	159 – 174	6
79 – 94	18	175 – 190	3
95 – 110	25		

24) The following data are the number of pages in 40 books on a shelf. Construct a box plot

136; 140; 178; 190; 205; 215; 217; 218; 232; 234; 240; 255; 270; 275; 290; 301; 303; 315; 317; 318; 326; 333; 343; 349; 360; 369; 377; 388; 391; 392; 398; 400; 402; 405; 408; 422; 429; 450; 475; 512

25) Graph a box-and-whisker plot for the data values shown.

0; 5; 5; 15; 30; 30; 45; 50; 50; 60; 75; 110; 140; 240; 330

26) The following data are the heights of 40 students in a statistics class.

59; 60; 61; 62; 62; 63; 63; 64; 64; 64; 65; 65; 65; 65; 65; 65; 65; 65; 65; 66; 66; 67; 67; 68; 68; 69; 70; 70; 70; 70; 70; 71; 71; 72; 72; 73; 74; 74; 75; 77

Construct a box plot with the following properties; the calculator instructions for the minimum and maximum values as well as the quartiles follow the example.

- Minimum value = 59
- Maximum value = 77
- Q_1: First quartile = 64.5
- Q_2: Second quartile or median= 66
- Q_3: Third quartile = 70

27) A population consists of 2,3,6,8. Take all possible sample of size 2 with replacement. Form the sampling distribution of sample mean $\bar{x}$. Also obtain mean and standard deviation. Verify the results with population mean and standard deviation.

28) For the population consisting of 4,6,8,10,12. Take all possible sample of size 2 with replacement. Find the means of these samples and make frequency distribution of the sample means. Calculate the mean and variance of this frequency distribution and compare it with the mean and variance of the population.

29) A population consists of 1000 students has a height distribution with $\sigma = 3$. Find the standard error of mean height for a random sample of 50 students selected without replacement.

30) A population consists of 1000 students has a height distribution with $\sigma = 3$. Find the standard error of mean height for a random sample of 50 students selected with replacement.

31) A population consists of 6 members 2,4,6,8,10 and 12. Take all possible sample of size 2 with replacement and without replacement. Form the sampling distribution of means. Find its means and variance. Compare it with population mean and variance.

32) A random sample of n = 25 values given $\bar{x} = 83$. Can this sample be regarded as drawn from a normal population with mean $\mu = 80$ and $\sigma = 7$ at 5 % level of significance.

33) **Confidence Interval for Mean (Z – test):** An auditor has selected a simple random sample of 100 accounts from the 8042 accounts receivable of a freight company to estimate the total audit amount of the receivable in the population. The sample mean is 33.19 and the sample standard deviation is $\hat{s}$ =34.48. Obtain the 95.44 percent confidence interval for the mean audit amount in the population. Hint: use $\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\hat{s}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$ with $Z_{\frac{\alpha}{2}} = 2.00$.

34) **Confidence Interval for Mean (Z – test):** Find a 90 percent confidence interval for the mean of a normal distribution with $\sigma = 3$ given the sample as $2.3, -0.2, -0.4, -0.9$. Hint: use $\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ with $Z_{\frac{\alpha}{2}} = 1.645$.

35) **Confidence Interval for Mean (t – test):** A random sample of seven observations of a normal variable gave $\sum x = 35.90, \sum x^2 = 186.19$. Calculate a 90% confidence interval for the population mean μ. Hint: use $\bar{x} \pm t_{\frac{\alpha}{2}(v)} \frac{\hat{s}}{\sqrt{n}}$.

36) **Confidence Interval for Mean (t – test):** A random sample of twelve measurements of the breaking strength of cotton threads gave a mean $\bar{x} = 209$ grams and a standard deviation $\hat{s} = 35$ grams. Calculate 95% and 99% confidence limits for the actual mean breaking strength. Hint: use $\bar{x} \pm t_{\frac{\alpha}{2}(v)} \frac{\hat{s}}{\sqrt{n}}$.

37) We wish to test the hypothesis that the mean weight of a population of people is 140 **lb**, using $\sigma = 15$ **lb**, $\alpha = 0.05$ **lb** and a sample of 36 people, find the values of $\bar{x}$ which would lead to rejection of the hypothesis.

38) A sample of 400 males students is found to have a mean height of 67.47 inches. Can it be regarded as a simple random sample from a large population with mean height 67.39 with standard deviation of 1.3 inches?

39) Injection of certain type of hormone into hens is said to increase the mean weight of eggs by 0.3 oz. A sample of 30 eggs has an arithmetic mean 0.4 oz above the pre injection mean and a value of $\hat{s}$ equal to 0.20. Is this enough reason to accept the statement that the mean increase is more than 0.3 oz?

40) A random sample of 25 hens from a normal population showed that the average laying is 272 eggs per year with a variance of 625 eggs. The company claimed that the average laying is at least 285 eggs per year. Test the claim of the company at $\alpha = 0.05$.

41) A random sample of size n is drawn from normal population with mean 5 and variance σ^2. If n = 9, $\bar{x} = 2$ and $t = -2$ what is $\hat{s}$?

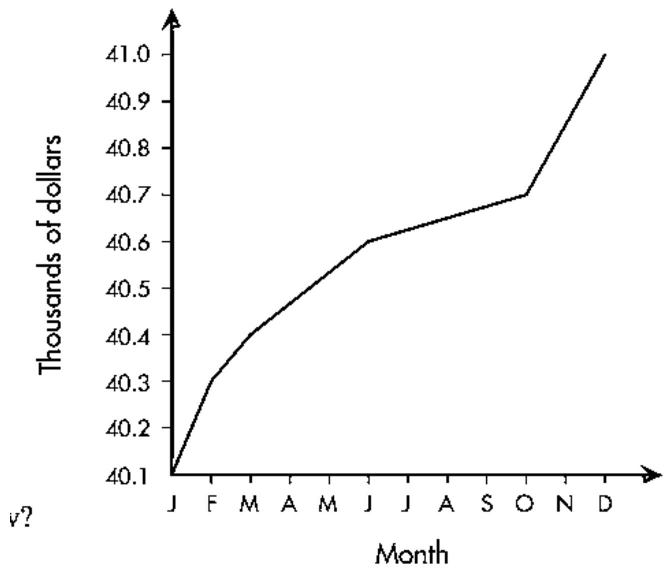
42) A random sample of size n is drawn from normal population with mean 5 and variance σ^2. If n = 25, $\hat{s} = 10$ and $t = 2$ what is $\bar{x}$?

43) Blood Pressure A blood pressure test was given to 450 women ages 20 to 36. It showed that their mean systolic blood pressure was 119.4 mm Hg, with a standard deviation of 13.2 mm Hg.

a. Determine the z-score, to the nearest hundredth, for a woman who had a systolic blood pressure reading of 110.5 mm Hg.

b. The z-score for one woman was 2.15.What was her systolic blood pressure reading?

44) The Brown Disc company presented the following graph to show its profits over the year.



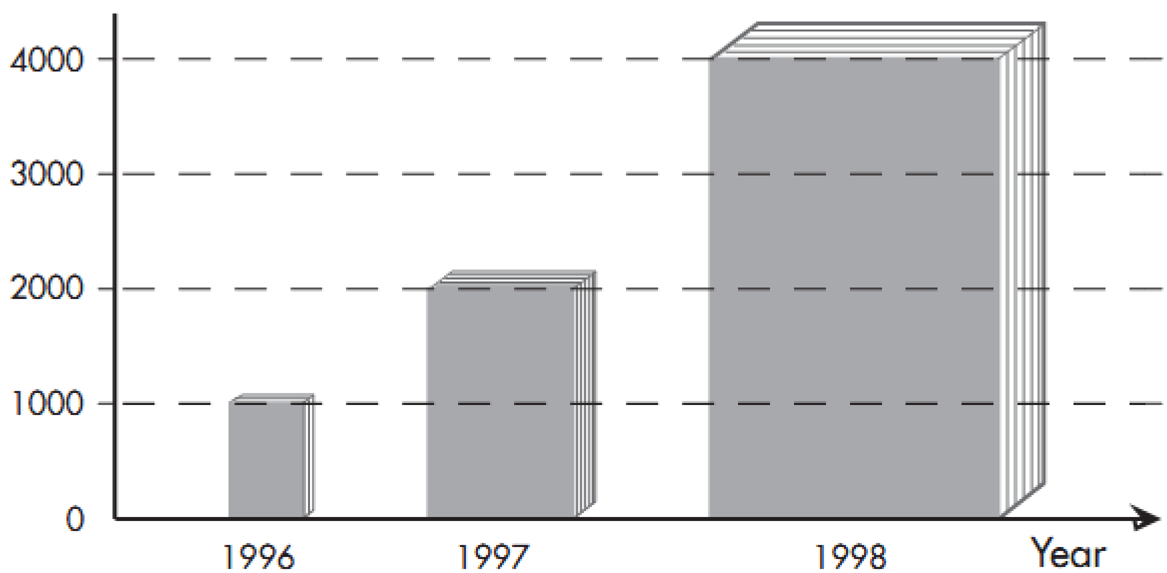
a) What impression does this graph give?

b) What is misleading about this graph?

c) Why do you think the company would present this graph?

d) What does your version of the graph show?

45) This graph shows the sale of books in a store over three years.



a) In what way is this graph misleading?

b) Give the number of books sold in:1996, 1997, 1998

46) The heights of all of the players on a basketball team are shown in the table below. Calculate the standard deviation of the population.

Player	Height
Laura	183
Jamie	165
Deepa	148
Colleen	146
Ingrid	181
Justiss	178
Sheila	154

47) Felix and Melanie have a job laying patio stones. Their boss in interested in who the better worker is so randomly throughout the week he chooses a few hours to record how many stones each of the workers lays. The data is recorded in the table below:

Felix	34	41	40	38	38	45
Melanie	51	28	36	44	41	46

Calculate the mean and standard deviation of each sample and compare use them to compare the two workers.

Recommended Texts

- Akar, G. K., Zembat, İ. Ö., Arslan, S., & Thompson, P. W. Quantitative Reasoning in Mathematics and Science Education.
- Sharma, A. K. (2005). Text book of elementary statistics.
- Blitzer, R. (2014). Pre – calculus, Pearson Education, Limited.
- Gupta, S. C., & Kapoor, V. K. (2020). Fundamentals of mathematical statistics.
- Aufmann, R. N., Lockwood, J., Nation, R. D., & Clegg, D. K. (2007). Mathematical thinking and quantitative reasoning. Cengage Learning
- Blitzer, R., & White, J. (2005). Thinking mathematically.
- Elementary statistics: A step by step approach, by Allan Bluman.
- Introductory Statistics by Barbara Illowsky.
- Statistics and Probability by Las Positas College Publications.
- Introduction to Statistical Theory by Prof Sher Muhammad Choudhry and Prof. Dr. Shahid Kamal.
- Applied Mathematics for Business, Economics and Social Sciences by Frank S Budnick.
- Discrete Mathematics with Applications by Susanna S. Epp.
- Quantitative Reasoning Algebra and Statistics by MSU Denver Tutoring Center.
- Quantitative Skills & Reasoning by the University of West Georgia.
- "Using and Understanding Mathematics: A Quantitative reasoning Appraoch, by Benitt, J.O. Briggs, W.L., Badalamenti, A.
- Applied Statistical Modeling by Salvatore Babones.
- Barrons SAT, by Sharvon Weiner Green, M.A. and Ira K. Wolf.
- Discrete Mathematics by K. H. Rosen.
- Introductory Statistics by Prem S. Mann.

حرفِ آخر (08-10-2024)

خوش رہیں خوشیاں بانٹیں اور جہاں تک ہو سکے دوسروں کے لیے آسانیاں پیدا کریں۔

اللہ تعالیٰ آپ کو زندگی کے ہر موڑ پر کامیابیوں اور خوشیوں سے نوازے۔(آمین)

محمد عثمان حامد

چک نمبر 105 شمالی (گودھے والا) سرگودھا

UNIVERSITY OF SARGODHA

PUNJAB, PAKISTAN