

SUMMARY OF THE FIRST 6 TOPICS

REG NO: SCM224-0959/2025

NAME: SUMEYA MARIAM ABDI

TOPIC 1: INTRODUCTION TO IT AND COMPUTING

ROLE OF IT IN THE SOCIETY

- Helps in processing information in different fields
- Helps in storing information
- Helps in sharing information
- Protects data and systems through cybersecurity measures
- Supports digital content creation and sharing in entertainment field
- Enhances public service through digital governance and online portals in the government and services
- Supports medical research and data analysis in healthcare
- Provides access to online courses, resources and remote learning in education field
- Creates new job opportunities in technical industries
- Enables global connectivity through internet, email and social media

DATA AND INFORMATION

DATA – refers to raw fact

INFORMATION – refers to processed raw fact (data)

COMPUTING APPLICATION

In Biostatistics:

- Disease modelling and outbreak detection in Epidemiology and public health
- Statistical analysis and data management
- Cloud computing
- Visualization
- Genomics and bioinformatics
- Data management

In financial engineering

- Risk management
- Algorithmic trading

HIGH PERFORMANCE AND CLOUD COMPUTING

HIGH PERFORMANCE – refers to using computers to solve complex calculation quickly

CLOUD COMPUTING – Accessing storage, software and computing power over the internet

Example: Google Drive, AWS

TOPIC 2: FUNDAMENTALS OF COMPUTER OPERATIONS

It covers how a computer works i.e inputting data processing, storing and producing the output

C.P.U – refers to the brain of the computer

It is responsible for executing instructions and processing data from input devices

A.L.U – (ARITHMETIC LOGIC UNIT)

- It does all the arithmetic and logical functions in a computer

CONTROL UNIT – Manages the flow of data transfer between the components i.e registers, ALU, and memory

REGISTERS- A storage space that is used to transfer data immediately to the CPU for data processing

SYSYEM CLOCK – Generates a timing signal and sets the pace

MACHINE CYCLES – refers to the sequence of operations a CPU performs to execute instruction

PARALLEL PROCESSING BASICS – Is a technique where multiple processor or cores execute multiple instructions simultaneously, improving overall system performance and throughout.

(a) Types of parallelism: Data parallelism and Task parallelism

(b) Parallel processing architectures: Multiprocessor systems

Distributed systems

Multicore processors

PERFORMANCE EVALUATION FOR STATISTICS COMPUTATION-

It basically depends on various factors, data size, computational and algorithm efficiency

TOPIC 4:COMPUTER HARDWARE (II)- STORAGE DEVICES AND MEMORY

PRIMARY AND SECONDARY MEMORY

Primary memory – it is responsible for running applications and operating system

Secondary memory – it is responsible for storing data and programs

Example of storage technologies:

- Cache – fast, small memory
- Solid state drive (SSD) – fast storage and reliable
- RAID (Redundant Array of Independent Disks) – improves data security
- Cloud storage- data is stored online

DISK STORAGE:

- Tracks -this is the circular path where data is written on the disk surface
- Sectors – are the divisions of tracks
- Clusters – are group of sectors
- Surface – data is written on the surface

IMPLICATION FOR LARGE SCALE BIOSTATISTICAL AND FINANICIAL DATA SET

- (a) High capacity storage are required hence use of distributed storage system for scalability
- (b) Fast retrieval is essential for timely analysis thus usage of high performance storage solutions like SSDs
- (c) Efficient data management and organization is important hence implementation of data compression techniques

TOPIC 5: COMPUTER SOFTWARE

SYSTEM SOFTWARE:

Responsible for managing and controlling computer systems.

OPERATING SYSTEM:

- It manages computer hardware resources and provides a platform for running application
- It also allows the user to interact with the device

COMPLIERS:

Are software that translate source code written in high level programming languages into machine code that computers can execute

UTILITIES:

Are software tools that perform specific tasks often related to system maintenance and management.

APPLICATION SOFTWARE OVERVIEWS:

- Word – Word processing software for creating documents, letters, reports, essays e.t.c
- Excel – Spreadsheets software for data analysis, budgeting and charting
- Access – Database management software for storing and managing data
- PowerPoint- Presentation software for creating slideshows and presentations

WORKFLOW OPTIMIZATION FOR STATISTICAL ANALYSIS

- Optimization helps in minimizing waste and errors hence streamlining processes to maximize efficiency, productivity and output quality.

-It involves analyzing and improving workflows to:

- Enhance collaboration
- Increase automation
- Reduce unnecessary steps
- Improves task prioritization
- Improves resource allocation

TOPIC 6:DATA FILES AND FILE MANAGEMENT

DATA FILES:

Data are stored in various formats and accessed differently

Data access methods:

- Random Access – Data can be accessed directly without reading the entire file
- Sequential Access – Data must be read in a specific order

Data types:

- Structured Data – Organized, formatted and easily searchable
- Unstructured Data – Not organized, lacks format and it's hard to access

KEY CONCEPTS IN DATA MANAGEMENT:

- 1) File creating – process of making a new file
- 2) Indexing – way of organizing data
- 3) Retrieval – getting the data you want from a file

OPTIMIZATION; RELATIONAL DATABASES AND NORMALIZATION

- Optimization refers to making database run faster and more efficiently
- Relational database refers to organizing data into tables with defined relationships
- Normalization refers to structuring data to minimize redundancy and improve data integrity

FILE MANAGEMENT – A file is a collection of information and the OS carries out the following activities with respect to file management