Business - has transformed business landscape by automating processes, enabling e-comm
and facilitating global trade.

Finance - Finance industry relies on IT for secure transactions, risk management and data
analysis eg online banking, investment services.

Health care - helps in patient management, medical research and treatment delivery.
Electronic health systems, records, telemedicine and health information
systems enhances quality of health care.

Agriculture - farmers use IT for weather updates, market places and modern
farming tools.

Innovation and creativity - supports new ideas like artificial intelligence, robots,
editing and producing digital work.

BASIC COMPUTER TERMINOLOGIES.

1. Central Processing Unit (CPU)

It executes instructions organised into programs that determine computer's actions.
Role in biostatistics - enable simulation ie epidemiology modelling.
execute statistical algorithm. ie survival analysis.

2. Memory (RAM)

Temporarily stores active programs, data currently being processed and intermediate results
Role: holding large patient databases during longitudinal analysis.
NB. Insufficient RAM causes disk swapping which slows computations in both
health and financial simulation.

3. Mass Storage Devices

Slower, long term memory that retains data btwn computational jobs eg
Hard disk devices (HDD)
Solid state devices (SSDs) - offer high read, crucial for large scale risk simulation.
Tape devices (archival) - relevant for long-term archival in hospitals or financial institution
Role: longitudinal surveys
Historical clinical trial data.
Genomic sequences.

3. Mini computer - A multi-user computer capable of supporting up to hundreds of users simultaneously.

4. Mainframe - a powerful multi-user computer capable of supporting many hundred thousand of users simultaneously.

5. Supercomputer - an extremely fast computer that can perform hundreds of million of instruction per second.

B: FUNDAMENTALS OF COMPUTER OPERATIONS

raw data Input - computer receives data through keyboard, mouse or touchscreen.

Processing - the computer system starts executing the instructions given by the user.

processed data Output - the result of the instruction given by a user.

Storing - memory and storage are used to hold data temporarily / permanently.

Controlling - type of process that monitors the instruction given by the user from the time it is executed to the output result.

* INPUT DEVICES - Keyboard   Touchscreen Microphone Game controller.
                  Mouse   scanner   Camera

✓ Processing DEVICES - Central Processing Unit - the brain

  Arithmetic Logic Unit - performs arithmetic & logical operation.
  Central Unit - direct all computer operations.
  Graphic Processing unit - handle graphics and parallel tasks.

Functions involved in Processing.

Running program Instruction.

Performing arithmetic calculation

Decision making through logic operation.

Managing flow of data between components.

✓ MEMORY & STORAGE - Primary memory

  Random Access memory - temporary working.
  Cache - high speed memory inside CPU.
  Registers - fastest storage inside CPU.
    Secondary memory.

  Solid state drive - fast permanent storage   Memory card - used in
  Hard Disk Disk - mechanical permanent storage   phone & cam era
  USB Flash Drive - portable external storage.

* OUTPUT DEVICES - Moniter   Speakers   VR Handset
                   Printer   Projectors

computer - an electronic device that receives information, processes it and gives
           back results
           - It's an electronic device that processes input data to output data.
     characteristics of a computer -
Speed - very fast device capable of performing calculations of very large amount of data
Accuracy - calculations are 100% error free.
storage - have large storage capacity.
Diligence - free from monotony, tiredness and lack concentration. It can work
           continuously without error.
Reliability - made designed to make maintainance easy.
Versatility - flexible in performing the jobs to be done.
Automation - they perform a given task automatically

I CPU
Consists of: a) Arithmetic and Logic unit.
             Is where most arithmetic and logical data computer operations are
             executed.
             NOTE: All operands (numbers) may not reside in the main memory. Processor
                   contains a number of high speed storage elements called Registers which
                   may be used for temporary storage of frequently used operands.
          b) Control unit
             Acts as a nerve center that sends control signal to other unit.
             Generates timing signal that governs the I/O transfer.
             Generates synchronization signals.
             Maintains order and direct the operation of the entire system by
             selecting, interpreting and executing the program instruction.

B
raw data
processed data

ASSIGNMENT 1

INTRODUCTION TO IT FOR STATISTICS

Information technology - Is the study, design, development, application, implementation support or management of computer-based information systems.

Is the application or use of computers and electronic systems to pro... store and share information.

A. **Role of IT in the society**

Provide productivity tools that enhance the speed and transparency of statistical oper...

Automate institutional work flows eg in hospitals.

Support high stake decision making through real time infor- access

Intergrate data from heterogeneous sources ie stock exchanges.

In academic/Professional fields

Selecting appropriate hardware and software for an organisation.

Maintaining, upgrading and securing the computational infrastructure.

Managing data capture tools for cohort study.

Configuring servers for R, Python and SAS computing

Ensuring security of trading algorithms and risk models.

Managing distributed databases for tick-level or sentimental analysis data.

Deploying low latency architectures for real time market feeds.

Store large data set eg transitional.

Transmit data securely across networks.

Protect sensitive information.

Retrieve information rapidly for decision making.

Transport - Through traffic lights, GPS navigation, online hailing apps like bolt, u... eases transportation.

Security - helps to protect data through passwords, encryption and monitoring systems. Also through CCTVs and alarms.

Education - Has enabled institutions provide online learning platforms, digital reso... and interactive tools ie zeraki, studoco.

Government - has helped improve public services, enhance transparency and engage with citizens eg E-citizen services.

4 Input Devices

They deliver data and instructions into the system eg clinical trial data, lab measurements and survey inputs in biostatistics.

5 Output devices.

Used to visualize computational results eg monitors, printers, dashboard.

6 Supporting components: The bus and interconnects.

The bus - Is a communication backbone linking CPUs, memory, storage and to ensu efficient computational performance.

Role in biostatistics: Influences large matrix operations & distributed dat base queries.

Hard

Hardware - physical component of a computer that are tangible ie keyboard, mouse.

Software - programs that run on a computer.

Network - a group of connected computers that share resources.

Internet - a global network that connects computers.

Operating systems (os) - main software the manages computer hardware and allows d programs to run ie Windows, Android.

IP Address - a unique no. that identifies a device on a network.

File - A digital document or piece of information stored in a computer.

Backup - A copy of data stored to prevent loss.

DATA VS INFORMATION.

Data - plain facts.

Information - data that is processed, organized, structured or presented in a given cont so as to make them meaningful & useful.

Data is the computers language while information is our translation of this language.

computer sizes and power.

1 Personal Computer = a small single-user computer based on a microproce

2 Workstation - a powerful, single-user computer and has a powerful microro ie high quality monitor.

2. REGISTERS

Internal memory organization of Processor

a. Instruction Register (IR)

Holds the instruction that is currently being executed.

Its output is available to the central circuits which operate, generates the timing signals that control the various processing element involved in executing the instruction

- Generates clock pulses
- Determines CPU speed
- Synchronizes computer components.

b. Program Counter (PC)

Contains the address of the instruction currently being executed.

c. General Purpose Register (R_0 to R_{n-1})

Facilitates communication with the main memory.

d. Memory Address Register (MAR)

Holds the address of the location to or from which data are to be transferred

e. Memory Data Register (MDR)

Contains the data to be written into or read out of the address location

3. MACHINE CYCLES

Is the complete set of steps the CPU goes through to process a single instruction or a single piece of data.

Steps: a) fetch - The CPU gets (fetches) the instruction from memory

b) Decode - The CPU interprets (decodes) the instruction to understand what action is required.

c) Execute - The CPU performs the operation eg addition, comparison, reading, writing data etc.

d) Store - The result of the operation is written back to memory or a register

Impacts

They determine how fast a CPU processes instructions.

CPUs with fewer machine cycles per Instruction are faster

A track - is a physical division of data in a disk drive, as used in the cylinder - Head-Record (CCHHRR)

Tracks are divided into blocks or sectors, pages

A sector - is a subdivision of a track on a magnetic disk or optical disk. Each sector stores a fixed amount of user data.

disk sector - intersection of a track and mathematical sector.

Clusters (Grid storage) - Is a new technology paradigm that pushes the scalability and efficiency boundaries of storage area network to (SAN's) new levels.

Made of network connected storage with an administrative function that manages a collection of physical disks

application for large scale Biostatistical data sets -

- COMPUTER SOFTWARE

① SYSTEM SOFTWARE

Refers to the program, instruction, and data that tells the hardware what to do

→ Functions as a bridge between computer system hardware and the application software.

→ Made up of many control programs including operating system, communication software and database manager.

→ Consists of 3 kinds of Programs: a) System Management Programs
  b) System Support Programs
  c) System Development Program

a) SYSTEM MANAGEMENT PROGRAMS:

Programs that manage the application software, computer hardware and data resources of the computer system.

b) System ~~Support~~ Development Program:

Helps users develop information system programs and prepare user programs for computer.

c) System Support Programs

They help the operations and management of a computer system -

- For accessing data, a distinct address is associated with each word location.
- Data and programs must be in the primary memory for execution.
    - ~~expensive~~
    - ~~execution~~
    - fast memory

Memory access - time required to access one word.

Primary memory allows the computer to store data for immediate manipulation and to keep track of what is currently being processed.

b. secondary storage.
    Are used when large amount of data have to be stored.
    Are less expensive compared to primary storage devices.

NB memory is used when referring to primary memory while storage is used while referring to secondary memory.

        Units used to measure computer memory ~~are~~

    Bit - smallest unit of data on a machine and a single bit can have only one of
            two values; 0 or 1.

    Byte - A unit of eight bits and it's able to contain any binary number between
            00000000 and 11111111. It is represented by an upper case B

    Kilobyte: - kilo refers to 1024.
            ↳ its equal to 1024 bytes - usually represented as KB

    Megabyte: - comprises 1024 kilobyte or 1,048,576 bytes.
            Its the standard unit of measurement for RAM and is represented as MB

    Gigabyte: - consists of 1024 MB.
            It is standard unit of measurement for hard disks and is rep as GB

    Terabyte: - Refers to 1024 GB. Represented as TB

D.   Memory Hierachy
    1. INTERNAL PROCESSOR MEMORY
        Placed within the CPU or it is attached to a special fast bus.
        Includes; a) Cache memory
                    b) special registers.

Cache Memory - a buffer, smaller and faster than main storage, used to hold a copy of instructions and data in main storage that are likely to be needed by the process and that has been obtained automatically from main storage.

2. PRIMARY MEMORY - (RAM, ROM)

a) RAM - is the place where the computer temporarily stores its operating system, application programs and current data so that the computer processor can reach them quickly and easily.

Types of RAM -

i) Static RAM - The info remains stable as long as power supply is supplied but data are lost as soon as the power goes down hence RAM is volatile

ii) Dynamic RAM - Its unstable. The data continues to move in and out of memory as long as power is available - Information needs to be refreshed after every few milliseconds to avoid it being erased.

b) ROM - storage of program and data is permanent.

Types of ROM

i) Programmable ROM (PROM) - Its programmable. However once the chip has been programmed, the recorded info cannot be changed.

ii) Erasable PROM (EPROM) - Info can be erased and the chip reprogrammed to record diff - information.

3. SECONDARY MEMORY / STORAGE DEVICES -

Provides back-up storage for instructions and data -

Computers use disks for storage:

Hard disks - located inside the PC and those that are attached to the PC externally. They are non volatile -

Can be easily removed and attached to other devices ie memory stick, floppy disk

CDs

Cloud storage

A disk drive track - is a circular path on the surface of a disk or diskette on which information is magnetically recorded and from which recorded information is read -

BASICS OF COMPUTER HARDWARE & SOFTWARE

COMPUTER HARDWARE (INPUT & OUTPUT DEVICES)

1 INPUT DEVICES

Hardware components that allow you to enter data, programs, commands and user responses into a computer

Keyboard - enters texts and commands.

Mouse - pointer device.

Scanner - converts physical documents to digital.

Microphone - voice input.

Camera/webcam - image and video capture.

Touchscreen - input and display.

Joystick/Game ped - gaming input.

Barcode reader - reads product codes.

Biometric devices - made up fingerprints/face scanners.

Functions of input devices
Accept user commands.
Capture raw data.
Enable interaction with Programs.

2 OUTPUT DEVICES

They make the information resulting from processing available for use.

Displays processed information.

Monitor - main visual output.

Projector - large screen display.

Plotter - prints large diagrams and architectural drawings.

Speaker/Headphones - audio output.

Printer - produces hard copies; laser printer - fast high quality.
Ink Jet - cheaper, good for photos.
Dot matrix - impact printing.

Functions
Provide visual/audio display
Present result after processing.

3 STORAGE DEVICES & MEMORY

Memory unit - stores program and data; i) Primary memory
ii) Secondary memory.

a) Primary Memory (Main memory)

Contains large number of semiconductor cells each capable of storing one bit of information.

@ OPERATING SYSTEMS

It's a collection of intergrated computer program that provide recurring services to other programs or to the use of a computer.

These services consist of a disk and file management, memory management and device management.

The operating system executes many functions to operate computer system efficiently ~~effote~~ ie.

a. Resource Management.

Manages a collection of computer hardware resources by using a variety of programs.

It manages computer systems resources: CPU, Primary memory, virtual memory, secondary storage devices, input/output peripherals.

b. Task Management.

It manages one program or many within a computer system simultaneously.

Task management program in an OS provide each task and interupts the CPU operation to manage multitasking capability.

c. File Management

They produce reports on a file.

Manages data files as it contains file management programs that provide ability to create, delete, enter, change, ask and access of files.

d. User Interface.

Allows ~~the~~ users to interact with a computer ie screen design, keyboard command.

@ UTILITY SOFTWARE

Is a system software designed to help analyse, configure, optimize, or maintain a computer.

A single piece of utility software is usually called a utility la tool.

↳ Usually focuses on how the computer infrastructure operates.

E -

DATA PROCESSING.

Steps involved in converting data into information:

1. Collection - gathering raw data.
2. Input - ~~gathering raw data~~ entering data into a system.
3. Processing - sorting, calculating, summarizing.
4. Output - presenting results (reports, charts)
5. Storage - saving for future use.
6. Distribution - sharing information with users.

Data representation in computers.

Computers store everything using binary digits (0 and 1).

Units:

Bit - smallest unit of data.
Byte - 8 bits.
Kilobyte (KB) - 1024 bytes.
Megabyte (MB) - 1024 KB.
Gigabyte (GB) - 1024 MB.
Terabyte (TB) - 1024 GB.

DATA FILES

Is a collection of related data stored together under one name.
It has a file name and an extension (eg report.docx, image)
Files help organize and retrieve information quickly.

Types of data files:

a. Text files - contain readable characters eg text, docx.
b. Binary files - data stored in binary form eg exe, bin, dat.
c. Image files - jpeg, png, gif.
d. Audio files - mp3, wav, aac.
e. Video files - mp4, mkv, avi.
f. Program files - executable or system files; exe, bat.
g. Database files - db, mdb, sql.
h. Compressed files - zip, rar, 7z.

f. **DATA FILES & DATA MANAGEMENT.**

Data - refers to facts, symbols, figures or observations that have not yet been processed.

It has no meaning until it is interpreted or organized - eg

Numbers   images, and audio, video

words or letters   sensor reading.

X-tics of good data -

Accurate - free from error

Complete - contains all necessary details

Reliable - can be trusted

Timely - available when needed

Consistent - does not contradict itself.

Relevant - relates to the subject matter -

Types of Data.

• According to nature   • After

1. Numeric data - numbers

2. Textual - letters

3. Booloon - TRUE, FALSE, YES, NO

4. Auto - music, recording

5. Image data - photo, scans.

6. Video data - moving picture -

7. Graphical data - charts, diagrams.

• According to structure.

1. structured data - neatly organised (spread sheets)

2. Unstructured " - no fixed format (videos, PDFs)

3. semi-structured - mixture of both (emails, user files)

Information vs Data -

Information	Data
Processed data	Raw fact
Has meaning	No meaning.

Information = Data + Processing.

File Organization Methods:-

1. Sequential file Organisation
   Records stored data one after the another in order
   Best for batch processing

2. Direct / Random file Organization
   Records assessed using a key field.
   fast for large files.

3. Indexed - sequential Organization
   Combines sequential and direct access.
   Uses an index for quick searches.

File Management - (activities involved in handling files)
   * Creating file
   * naming files
   * Opening and closing file
   * saving and retrieving files.
   * Renaming file
   * Moving files between folders.
   * Deleting file.
   * Backing up important data.

File extension.
txt - text	mp3 - sound
docx - document	mp4 - video
pdf - portable document.	exe - executable
jpg png - image	xlsx - spreadsheet.

File security
Passwords - restrict access.
Encryption - scrambles data so only authorised can read it.
Access rights permission - control who can read, edit or delete file.
Firewalls - block both unauthorized network access.
Antivirus software - protects files from malware.

System monitors - for monitoring resources and performance in a computer system.
System profiles - provides detailed info about the software installed and hardware attached to the computer.

4 COMPILING PROGRAM.

A compiler is a computer program that transforms source code written in a programming language into another computer language.

Decompiler - program that translates from a low level language to a high level language.

Main components.

Compiler - does the actual translation of source code into machine code
Lexical analyser - breaks the code into tokens (words/symbols)
Syntax " - checks if the code follows language rule
Code generator - produces the final machine code
Error handler - shows mistakes in the code.

How it works
You write a code in a programming language.
The compiler reads the code and checks for errors
If everything is correct, the compiler converts the code into machine language.
The computer can now execute the program.

Importance.
They turn human readable programmes into executable files.
They help detect errors before the program runs.
They make programmes run faster because machine code is optimized.

Types of Utilities

Anti-virus - utilities scan for computer viruses.

Backup - utilities can make a copy of all information stored on a disk and store either the entire disk or selected files.

Data compression - utilities output a shorter stream or a smaller file when provided with a stream or a file.

Disk checkers - can scan operating hard drive.

Disk cleaners - can find files that are unnecessary to computer operation, or to up considerable amounts of space.

Disk compression - utilities can transparently compress/uncompress the contents of disk, increasing the capacity of a disk.

Disk fragmenters - can detect computer files whose contents are broken across several locations on the hard disk, and move the fragments to one location to increase efficiency.

Disk partitions - can divide an individual drive into multiple logical drives, each with its own file system which can be mounted by the operating system and treated as an individual drive.

Disk space analyzers - for the visualisation of disk space usage by getting the size for each folder and files in folder or drive showing the distribution of the used space.

Archieve - utilities output a stream or a single file when provided with a directory or a set of files.

File managers - provides a convenient method of performing routine data mgt. ie moving, copying, merging, renaming, deleting, generating and modifying data sets.

Cryptographic - utilities encrypt and decrypt streams and files.

Hex editors - directly modify the text or data of a file.

Memory testers - check for memory failures.

Network utilities - analyze the computer's network connectivity, configure network system, check data transfer or log events.

Registry cleaners - clean and optimize the windows registry by removing old registry keys that are no longer in use.

File storage Devices-
HDD (Hard Disk Drive) - large capacity, cheaper.
SDD (Solid State Drive) - faster, more durable.
USB flash drive - portable and convenient.
Memory cards - used in phones/cameras.
Optical Discs - CDs, DVDs.
Cloud storage - Google Drive, One drive, icloud.

Data Back-up.
Prevent loss of data.
   Methods
Full backup - copies everything.
Incremental backup - copies only changes since last backup
Differential backup - copies changes since last full backup.
Cloud backup - remote servers.
External drive - offline backup.

Importance of Data files and file Management.
Prevents data loss
Ensures quick access to information
Increases efficiency.
Improves decision - making.
Enhances security and organization
Reduces error in duplication.