SMA 2102: IT FOR STATISTICS

1. INTRODUCTION TO IT FOR STATISTICS

- Data is raw facts.
- IT (Information Technology) refers to use of technological tools to store, process, transmit and manage information or processed data.
- For Statistics; it is referred to as the application of computer systems and statistical software to make data handling faster, more accurate and easier to interpret (collect, store, process, analyze and present)

Role of IT

- Data analysis
  Convert raw data to information.
  Systematic process of inspecting and interpreting data to come up with meaningful conclusions and support informed decision.
  Programs used; SPSS, EXCEL, R, PYTHON.
- Data collection
- Data storage
- Data processing
- Data presentation

Role of IT for Statistics in Society

- Financial Impact:
  Enhances employability and addressing skills gap.
  Vocational training and apprenticeship.
  Financial literacy and economic resilience.
  Reduced financial vulnerability.
  Economic participation and growth.
- Social Impact:
  Reduced economic dependency.
  Diverse talent pipeline.
  Community Integration.
- Data collection and enhancement.
- Supports planning and policy making.
- Enables storage and management.
- Boosts research and innovation.

2. Fundamentals of Computer Operations.

- Computer is an electronic device that inputs data, process it and give out information.
- Computers are made hardware (physical parts) and software.

Components of a computer

Output devices
Input devices
CPU (Central processed unit) – process instuctions/data:
ALU – performs calculations
Control Unit – direct co-ordinates
Register – small storage……
Motherboard – connects all components.
Memory – RAM (temporary) ROM (permanent)
Secondary: USB, Hard drive, SSD, CD (compact disk), Memory card.

Software – set of instructions that tell a computer what to do.

System – controls computer hardware and provides a platform for application software e.g Windows

Application – programs designed to perform specific tasks e.g Microsoft Word

Basic Operations of a Computer

- Input – receive data/ input data
- Processing – process statistical formular, performing calc/data manipulation.
- Storage – saving data permanently or temporarily.
- Output – displaying; display results using charts and graphs.
- Control – managing entire process; use tools like Excel, SPSS, R, Python.
- IPO Cycle: Input – process – Output

Computer Operation Cycle (Machine Cycle)

- Fetch – get instruction from memory.
- Decode – Interpret the instruction.
- Execute – perform the operation.
- Store – save result in memory or register.

3. Computer Hardware Basics

Computer hardware refers to the physical components of a computer system. It includes input devices like the keyboard and mouse, output devices such as the monitor and printer, and internal components found in the system unit. Key internal parts are the CPU, which processes data, RAM for temporary storage, the motherboard that connects all components, and storage devices like HDDs and SSDs for long-term data storage. Other hardware includes the power supply, which provides electricity, and communication devices like network cards that enable connectivity. Together, these parts allow the computer to input, process, store, and output information.

4. Computer Software Basics

- Software is a collection of programs, instructions and data that tell a computer what to do.
- It is intangible – you cannot touch it like hardware.
- Two types of software:

a) System Software

- This is software that managesand controls the hardware so that other software can run, it includes:

I. Operating Systems (OS) – e.g., Windows, Linux, macOS, Android.

II. Functions: manages files, memory, processes, hardware, and provides the user interface.

III. Utility Programs – perform maintenance tasks.

IV. Examples: antivirus, disk cleanup, file compression, backup tools.

V. Device Drivers – help hardware communicate with the OS.

Examples: printer driver, keyboard driver, graphics driver.

B. Application Software

These are programs used to perform specific tasks for the user.

Categories:

I. General-purpose software – Word processors (Word), Spreadsheets (Excel), Browsers (Chrome).

II. Special-purpose software – Designed for a specific task.

Examples: payroll systems, school management systems, banking software.

III. Customized software – Created to meet the unique needs of an organization.

Differences Between System and Application Software

System Software	Application Software
Controls/manages hardware	Helps users perform tasks
Runs automatically when computer starts	Runs when user opens it
Example: OS	Example: Word, Chrome

Software Acquisition Methods

Ways to get software:

- Purchase – buying a license.
- Freeware – free to use.
- Shareware – free trial then paid.
- Open-source – source code available for modification.
- Custom development – software built for specific needs.

Software Licensing

A license is a legal agreement that states how software can be used.

Types:

- Single-user license
- Multi-user license
- Site license
- Free and open-source license
- Proprietary license

Software Installation and Updates

Installation puts the program onto the computer.

Updates fix bugs, add features, and improve security.

Importance of Software

- Enables communication with hardware.
- Helps users complete tasks.
- Provides a platform for applications to run.
- Enhances productivity and efficiency.

5. DATA AND DATA FILES

1. Data

Raw, unprocessed facts (numbers, text, images, sound).

Has no meaning until processed into information.

2. Types of Data

Numeric: numbers used in calculations.

Text/Alphanumeric: letters, digits, symbols.

Boolean: TRUE/FALSE values.

Multimedia: pictures, audio, video.

3. Data Representation

Computers use binary (0s and 1s) to represent all data.

Units: bit> byte> KB> MB> GB >TB.

Bit (short for binary digit)

The smallest unit of digital data.

It can have only one of two values: 0 or 1.

Example: A single light switch (on = 1, off = 0) represents 1 bit.

Byte

A group of 8 bits.
One byte can represent 256 different values ($2^8 = 256$), which is enough to store one text character (like the letter "A") using ASCII/Unicode.
Example: The letter "A" in ASCII is 01000001 in binary → that's 1 byte (8 bits).
KB (Kilobyte)
Approximately 1,000 bytes in everyday/decimals use (used by storage manufacturers),
or exactly 1,024 bytes in technical/binary use (used by computers).
Most of the time in school and exams:
1 KB = 1,024 bytes
MB (Megabyte)
Approximately 1,000,000 bytes (decimal),
or exactly 1,048,576 bytes in binary (1,024 × 1,024).
Again, in most school/technical contexts:
1 MB = 1,024 KB = 1,048,576 bytesText uses ASCII/Unicode; images use pixels; audio uses sampled values.
4. Files
A file is a named collection of related data stored on a device.
Has a name, type, size, and location.
5. Types of Files
Data files: documents, images, audio, video, spreadsheets.
Program files: executable instructions (.exe, drivers, system files).
6. File Management
Creating, saving, renaming, moving, copying, deleting, backing up files.
Organizing files into folders helps easy access and safety
7. Data Security & Integrity
Security: protect data using passwords, antivirus, permissions.
Integrity: ensure data stays accurate and uncorrupted.
Use backups to prevent loss.
8. Relationship
Data > processed by a program >becomes information > stored as file
6. Disk Storage Fundamentals
1. Types of Disk Storage
a) Magnetic Storage
Hard Disk Drive (HDD), Floppy disks
Characteristics: Large capacity, cheaper per GB, slower, mechanical parts (prone to physical damage)
b) Solid-State Storage
SSD, USB flash drives, memory cards
Characteristics: Very fast, no moving parts, durable/shock-resistant, more expensive per GB
c) Optical Storage
CDs, DVDs, Blu-ray
Characteristics: Cheap, portable, good for backups/distribution, low capacity, easily scratched
2. How Disk Storage Works
HDD: Spinning platters coated with magnetic material, read/write heads move over tracks and sectors
SSD: Uses NAND flash memory cells; no moving parts, extremely fast access
Optical: Laser reads/reflects off pits (0) and lands (1) burned into disc surface

3. Key Terms
Track: Circular path on a platter
Sector: Smallest addressable storage unit (usually 512 bytes or 4 KB)
Cylinder: Same track number across multiple platters
Access Time: Time taken to locate and retrieve data
Capacity: Total storage size (e.g., GB, TB)
4. Why Disk Storage Matters
Permanent (non-volatile) data retention
Essential for storing OS, applications, databases, and user files
Directly affects system performance (SSD >> HDD speed)
5. Disk Performance Factors
Seek Time: Time for read/write head to move to correct track (HDD only)
Rotational Latency: Time for platter to rotate to correct sector (HDD only)
Transfer Rate: Speed of actual data read/write once positioned
6. Advantages of Disk Storage
Large capacity (especially HDD/SSD)
Permanent/non-volatile storage
Faster than tape storage
Supports all file types
7. Disadvantages
HDD: Slow, mechanical failure risk, noise/vibration
SSD: Higher cost per GB, limited write cycles
Optical: Very slow, low capacity, easily damaged by scratches
8. Common Uses
Operating systems and applications
Databases and large files
Backups (external drives, optical discs, cloud)
Media storage (photos, videos, music)
9. File Systems (organize & manage data on disks)
FAT32: Old, compatible, max 4 GB file size
NTFS: Windows standard, secure, journaling, large files
exFAT: Good for flash drives, supports large files, cross-platform
ext4: Linux standard, reliable, journaling
APFS: Modern macOS/iOS file system, fast, encrypted
10. Disk Maintenance Best Practices
Disk Cleanup: Remove temporary files, junk, recycle bin
Defragmentation: Reorganize fragmented files (only needed for HDDs; SSDs should NOT be defragged)
Regular Backups: Use external drives, optical media, or cloud storage.