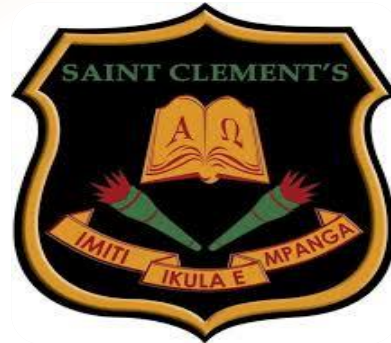


ST CLEMENT'S SECONDARY SCHOOL



DATA REPRESENTATION AND PROCESSING

Senior Secondary Grade 10

Data representation

- **Data** is a collection of raw or unprocessed facts that does not have any meaning on its own.

Example of data

- Text
- Numbers
- Audio
- Video
- Images
- **Information** is processed data that convey meaning and its useful to the user or organization in making decisions or solving problems.

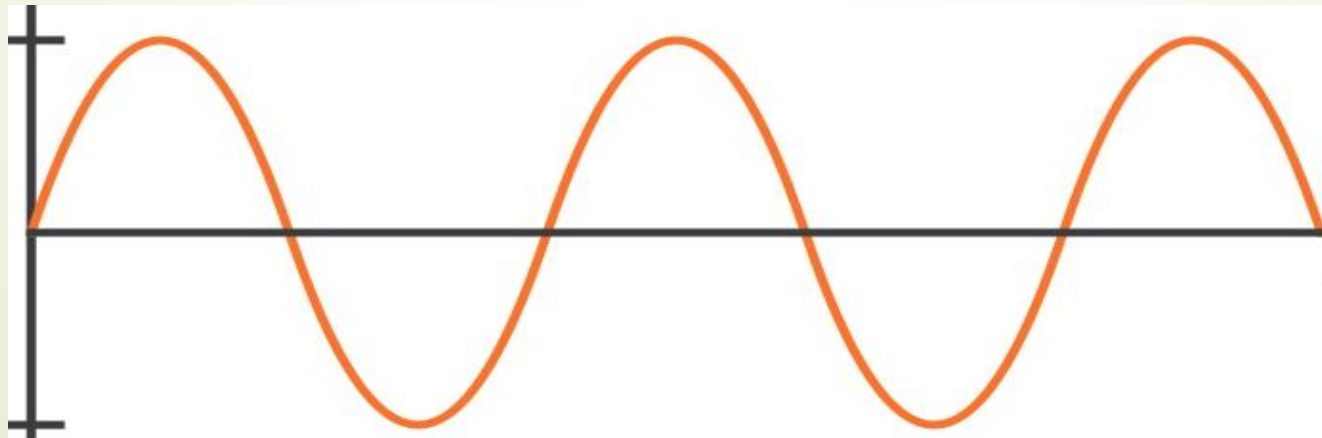


The concept of data representation

- **Data representation** refers to the form in which data is stored, processed and transmitted.
- **Data** can be represented using digital or analogue methods.

Analogue signal

- **An analogue** signal is the one that has a value that varies smoothly.
- Examples of analogue signal
 - Human voice
 - Dimmer light switch



Digital signal

- This is the signal represented in one of distinct states, either 0 or 1.
- Text, music, images, numbers, video and audio data are in **digital** form.
- The computer uses millions of electronic switches to represent data in only one of the two distinct states- off (0) or on (1).
- Representing data as 1s and 0s is known as **binary system**.
- Data is also stored in digital form on CDs DVDs and flash memory.



Coding scheme

- The ASCII (American Standard Code for Information Interchange).
 - It is most widely used common scheme for data representation, it can represent 65 000 codes.
- Unicode
 - It can represent 65 000 characters and symbols.
 - It is capable of representing all world's current languages.

Reasons for data representation

- Makes it possible for humans to communicate with a digital computer that uses 1s and 0s for processing.
- When data is in human-readable form via input devices, the data is converted into machine-readable (binary) form which a computer can process and store in its memory.
- Makes it possible for components of the computer to communicate with each other successfully.
- Standards also enable manufacturers to produce a component and be confident that they will operate correctly in a computer.



Exercise

1. Define data representation.
2. Explain the meaning of term data.
3. Explain two reasons for data representation .
4. Draw how data is represented in:
 - (a) Digital form
 - (b) analogue



Data collection and preparation

- Data collection refers to the process of bringing data from various sources to one central point.



Methods of data collection

1. Data from digital text repositories.
2. Interviews



Data from digital text repositories.

- The advantage of this method is that you gain access to all the information online without having to leave your home.



Interviews

- There several types of interviews.
 1. Face-to-face interviews
 2. Telephone interview
 3. Computer-Aided Personal Interviewing



Face-to-face interviews

- This is the type of interview where the researcher establishes a rapport with the potential participants and thereby gains their cooperation.
- This method yields the highest response rate
- It allows the researcher to clarify ambiguous responses and when appropriate to seek a follow-up.



Advantages of using face-to-face interviews

- Accurate screening.
- Capture verbal and non-verbal cues.
- Keep focus
- Capture emotions and behaviors



Disadvantages of using face-to-face interviews

- Cost
- Quality of data by interviewer
- Manual data entry
- Limit in sample size



Telephone interview


Advantages of telephone interview

- These are less time consuming and less expensive.
- The researcher can access to any participant with a telephone.



Disadvantages of telephone interview

- The response rate is not as high as face-to-face but still better than a mailed questionnaire.
- The sample maybe biased because people without a telephone are excluded.



Computer-Aided Personal Interviewing (CAPI)

- This is a form of personal interviewing where instead of participants completing a questionnaire, the interviewer captures the information directly in database using a laptop or hand-held computer.



Advantages of c CAPI

- This method saves time.
- Facilitates logic checks
- Validations during interviews.



Disadvantage of CAPI

- The survey may be difficult to conduct in crime-ridden areas.
- The interviewer must have the required typing skills.
-



Questionnaires

- Researchers develop questionnaires to obtain information.
- **1. Paper-pencil questionnaires:** these can be sent to a large number of people

Advantages

1. saves researcher time and money.
2. Participants are more truthful

Disadvantages

1. Many people who may receive questionnaires may not return them.

Department / Medical Unit:**Date:****Hour:**

Please fill in the appropriate answers:

1. How long have you been working in this building?

- Less than 1 year;
- 1-2 years;
- 3-5 years;
- More than 5 years;

2. How long have you been working in this department/operating room?

- Less than 6 months;
- 6-12 months;
- 1 an - 3 years;
- More than 3 years;

3. In a normal week how many hours do you spend in the operating room?

- Less than 10 hours;
- 10-20 hours;
- 20-30 hours;
- More than 30 hours;

4. What age do you have?

- Less than 30 years;
- 30-40 years;
- 40-50 years;
- More than 50 years;

5. Specify your gender:

- Female;
- Male;



1. Web-based questionnaire:

- participants receive an e-mail with a link to secure website to fill in a questionnaire.
- It is quicker with fewer details.
- But it does exclude people without access to a computer



9:41

Stop

Smartphone Survey

Recap

Which smartphone brand do you own?

- Samsung
- Huawei
- Asus
- Apple
- Xiaomi
- Acer

How much are you satisfied about Apple brand?



How likely is it that you would recommend iPhone to a friend?

0 1 2 3 4 5 6 7 8 9 10

Next



Why questionnaires are important

- They are fast
- Efficient
- Inexpensive
- Reach people quickly
- Flexibility
- Data accuracy



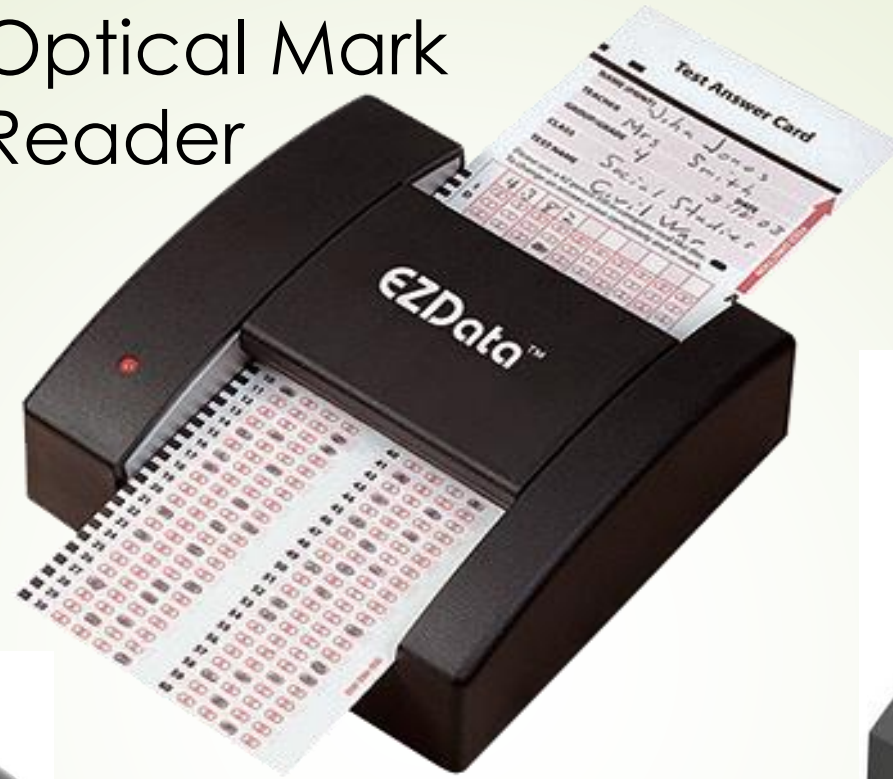
Other sources of data

- Barcodes
- Optical mark recognition (OMR)
- Optical Character Recognition (OCR)
- Magnetic ink Character Recognition (MICR)
- Smart cards
- Magnetic stripes

Barcode reader



Optical Mark Reader



MICR magnetic ink
Character Recognition





Point of sale
machine



OCR (Optical Character Recognition)

Quick Response





Magnetic
Stripe Cards



Magnetic
Stripe Card
Reader



Methods of data preparation and data errors



Data preparation

- **Data preparation** is a process of cleaning and transforming raw data prior to processing and analysis.
- Once data is collected it must be organized correctly to save time and prevent mistakes.
- Most researchers use the following tools to organize and analyse data effectively.
 1. Database
 2. Statistical analysis program.
 - Microsoft excel
 - SPSS



Data processing errors

1. **Routing errors:**

Asking questions in wrong order

2. **Consistence errors:**

Contradictory responses

3. **Range errors**

Responses outside the range of reasonable answers

4. **Coping errors**

5. **Coding error**

Data verification

- Data must always be validated and verified for human errors.
- **Verification** is the process of performing checks to ensure that the data entered exactly matches the original source.

Methods of data verification

1. Double entry
2. Proofreading data

Data validation

- **Data validation** is a process of comparing data against the set of rules to determine whether the data is valid.
- before processing, data validation techniques such as format, range and data type checks are applied to the input to increase accuracy and consistency of data preventing invalid data to be entered.

Methods for checking validity

1. Use of validation rules
2. Required fields
3. Specific file size
4. Input masks
5. Default values



Exercise

1. A group of grade 10 G pupils at St Clements wanted to collect information about malaria outbreak in school by interviewing fellow pupils.
 - (a) State two types of interviews that can be appropriate to use to collect data in school.
 - (i)
 - (ii)
 - (a) Explain the reason why questionnaire is used to collect data.
 - (b) The information captured must be verified. Give the reason for this.
 - (c) Explain the meaning of verification.

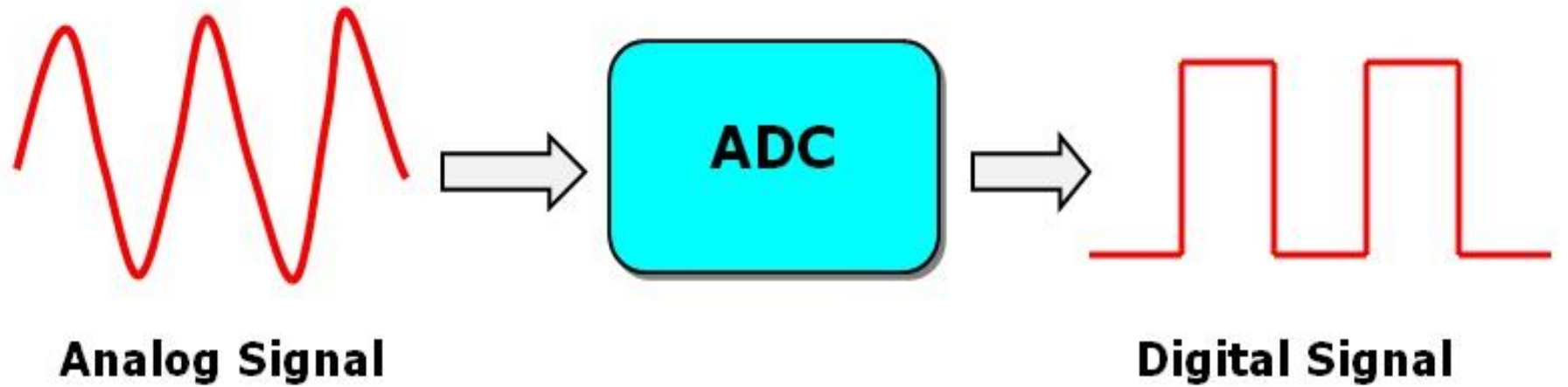


Conversion between analogue and digital

Analogue-to-digital converters

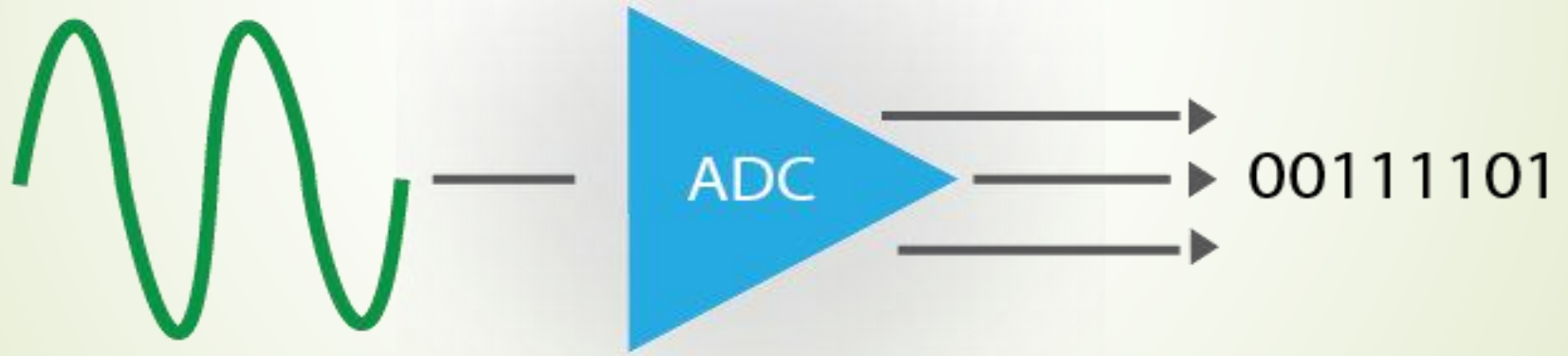
- It is abbreviated as ADC, A/D or A to D.
- It is a device that converts a smooth continuous signal into digital.
- For example when you plug a microphone into a computer, you actually connect it to the input jack of the sound card.
- Sound card in a computer system is circuit board that converts sound from analogue to digital.

ADC are also used in devices that capture that capture the output of the video.



Analog Signal


Digital Signal



00111101

Digital to Analogue converters

- DAC converts a digital signal into analogue signal.
- When you connect a loud speaker or headphones to your computer, you require a DAC. This serves as a DAC and the sound card converts the digital data from the computer into analogue signals which loud speaker then converts into sound.
- CD players, audio players, old TVs have DAC.
- DAC uses a technique called **interpolation** to ensure that the analogue signal is as smooth and continuous as possible.

- 
- In a Dual-up connection, you need modem or router that acts as an ADC and a DAC
 - A modem receives the analogue signals and converts it into digital signal for processing.



Exercise

1. State the type of data that can be captured by a microphone.
2. State one application where a microphone is used to input data.
3. Define the following
 - (a) DAC
 - (b) ADC



Types of data representation




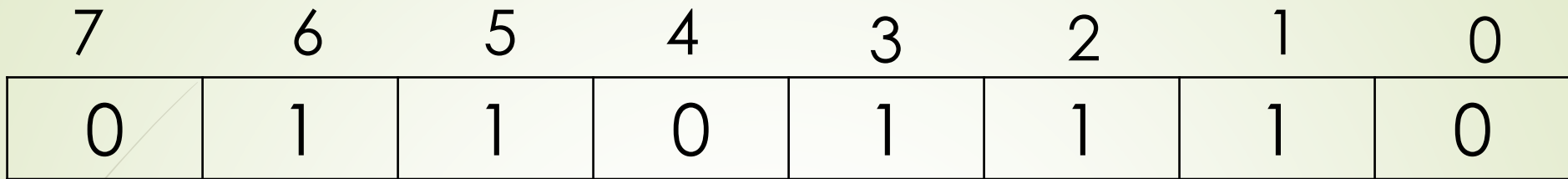
Data storage/ capacity

- Storage hold data, information and instructions.
- Secondary memory in a computer is referred to as storage, external, long-term or backup storage.

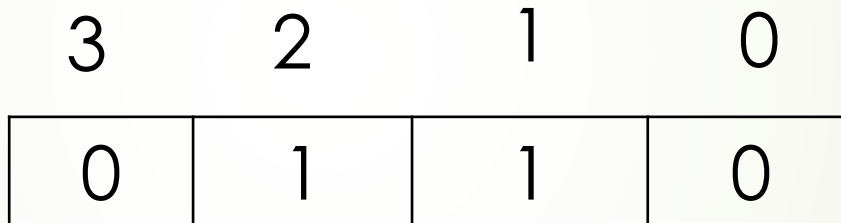
Examples of secondary memory

- CDs
- Hard disks
- DVD
- USB flash drives

- 
- the unit of measurement in a computer system such as Kilobit and megabytes are derived from the fact that computers use a binary number system (1s and 0s) to represent, process or store data.
 - 0s and 1s are referred to as binary digits (bits)
 - A **bit** is the smallest unity of data that a computer can process and is capable of representing only one of two different values.
 - A group of 8 bits is called a byte.
 - A **byte** provides $2^8 = 256$ different combinations of 0s and 1s.
 - A byte has numbers from 0-7 starting from the right to the left.
 - A **nibble** is a collection of 4 bits and provide $2^4 = 16$ combinations of 0s and 1s.



A byte



A nibble



Measure of storage capacity

1024 bytes = 1 Kb (Kilobyte)

1024 Kb = 1 Mb (Megabyte)

1024 Mb = 1 Gb (Gigabyte)

1024 Gb = 1 Tb (Terabyte)



Number systems



Number systems

1. The decimal number system
2. The binary number system
3. The octal number system
4. The hexadecimal number system



Decimal number system

- Base 10
- Number of elements 10
- Elements 0 1 2 3 4 5 6 7 8 9

Place value system

5432.9

Position	3	2	1	0	-1
Place value has a base	10^3	10^2	10^1	10^0	10^{-1}
Decimal value	1000	100	10	1	0.01
Numbers	5	4	3	2	9



Expanded notation

Example 1

5432.9

Solution

$$= 5 \times 10^3 + 4 \times 10^2 + 3 \times 10^1 + 2 \times 10^0 + 9 \times 10^{-1}$$

$$= 5 \times 1000 + 4 \times 100 + 3 \times 10 + 2 \times 1 + 9 \times 0.1$$

$$= 5000 + 400 + 30 + 2 + 0.9$$

$$= 5432.9$$

Binary number system

- It is used by computers

Base 2

Number of elements 2

Elements 0 1

Position	3	2	1	0	-1
Place value has a base	2^3	2^2	2^1	2^0	2^{-1}
Decimal value	8	4	2	1	0.5
Numbers	1	0	1	1	1

Convert binary to decimal

Example

$$= 1011.1_2$$

$$= 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1}$$

$$= 1 \times 8 + 0 \times 4 + 1 \times 2 + 1 \times 1 + 1 \times 0.5$$

$$= 8 + 0 + 2 + 1 + 0.5$$

$$= 11.5$$

The octal number system

- It is also used in computers
- Base 8
- Number of elements 8
- Elements 0 1 2 3 4 5 6 7

Example: 215.22_8 can be represented as follows

Position	3	2	1	0	-1
Place value has a base	8^2	8^1	8^0	8^{-1}	8^{-2}
Decimal value	64	8	1	0.125	0.015625
Numbers	2	1	5	2	2

Convert an octal to binary

215.22_8

$$= 2 \times 8^2 + 1 \times 8^1 + 5 \times 8^0 + 2 \times 8^{-1} + 2 \times 8^{-2}$$

$$= 2 \times 64 + 1 \times 8 + 5 \times 1 + 2 \times 0.125 + 2 \times 0.015625$$

$$= 128 + 8 + 5 + 0.25 + 0.03125$$

$$= 141.28125_9$$

The hexadecimal number system

Base	16
Number of elements	16
Elements	0 1 2 3 4 5 6 7 8 9 A B C D E F

EXAMPLE $2AF.B_{16}$

Position	2	1	0	-1
Place value has a base	16^2	16^1	16^0	16^{-1}
Decimal value	256	16	1	0.0625
Numbers	2	A	F	B

Expanded notation: $2AF.B_{16}$

$$2 \times 16^2 + A \times 16^1 + F \times 16^0 + B \times 16^{-1}$$



Convert a hexadecimal number to decimal

Example 1

Convert the following

(a) $2AF.B_{16}$

(b) $8E$

Solution 1

$$(a) 2AF.B_{16}$$

$$= 2 \times 16^2 + A \times 16^1 + F \times 16^0 + B \times 16^{-1}$$

$$= 2 \times 256 + 10 \times 16 + 15 \times 1 + 0.0625$$

$$= 512 + 160 + 15 + 0.0625$$

$$= 687.625$$



Solution 2

$$\begin{aligned} & (b) 8E \\ &= 8 \times 16^1 + E \times 16^0 \\ &= 8 \times 16 + 14 \times 1 \\ &= 128 + 14 \\ &= 142 \end{aligned}$$



EXERCISE

Convert the following to decimal numbers

(a) 4B

(b) 2AF

(c) AFB

(d) 8D



Convert from decimal to binary, octal and hexadecimal

- 
- A decimal number has two parts

- (i) Integer part

- (ii) Fractional part

Example 4.15

- 4 is a decimal and .15 is the fractional part.
- The integer part of a decimal number is converted to any base to any base using division operation.



Example 1:

Convert the following

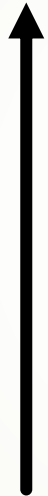
(a) 25_{10} to binary

(b) 453_{10} to octal

(c) 3456_{10} to hexadecimal

Solution (a)

Base	quotient	R
2	25	
2	12	1
2	6	0
2	3	0
2	1	1
	0	1



11001_2

Solution (b)

Base	quotient	R
8	453	
8	56	5
8	7	0
	0	7



$$453_{10} = 705_8$$

Solution (c)

Base	quotient	R
16	3456	
16	216	0
16	13	8
	0	13 (D)



$$3456_{10} = D80_{16}$$

Converting from decimal to hexadecimal


1. Example
2. Convert 3456_{10} to hexadecimal

base	Number (quotient)	Remainder
16	3456	
16	216	0
16	13	8
	0	13



$$3456_{10} = D80_{16}$$

THE DATA PROCESSING CYCLE



The Data Processing Cycle, Errors And Data Integrity

DATA PROCESSING

- **Data** is a single item or fact that has no meaning or news value.
- **Information** is processed data that does have meaning or news value to the users.
- **Data processing cycle** consists of series of steps where raw data (input) is fed into a process (CPU) to produce output.

input

processing

output

INPUT

Acquiring or gathering data and entering it into the computer system.

- The data is validated by checking for completeness and accuracy.
- Data refers to unprocessed text, images, video or audio.
- **Examples**
 - Capturing the prices of items in a supermarket using a bar code reader.
 - Gathering information using a form on a website.
 - Typing using word processing program.



Input devices

- Keyboard
- Mouse
- Scanners
- Barcode readers
- Touchpads
- Digital cameras
- Video cameras
- Microphones
- Voice recognition
- Biometric devices



Processing

- The operation performed on the data to produce information.
- The CPU is responsible for data processing.
- All data that is currently being processed by the CPU is stored temporarily in a random access memory (RAM).



Storage

- Saving data for future use.
- Data and information stored in a secondary storage is not lost when the computer is switched off.

Storage devices

- CDs
- DVDs
- Blu-ray
- External hard drive
- USB drives



Output

- Presenting the in the required format for the user.
- Output can be in form of hardcopy, softcopy, audio, or video.

Output Devices

- Printers
- Monitor
- Fax
- Headset
- Speakers
- Multifunction devices
- Data projectors



Communication

- Computers are able to communicate with other computers and mobile devices.



Exercise

1. Explain the meaning of the term data.
2. State two input devices which automatically capture data.
3. Give an application that uses the devices you have mentioned in question 2.
4. State two examples of output devices.
5. What is data integrity?



Data integrity

- **Data integrity** refers to the accuracy and consistency of data stored in data base.
- if data is inaccurate then the results will be inaccurate, this is called **GIGO** (Garbage in, Garbage Out)
- Data integrity is vital because people make decisions and take actions based on processed data.



Types of computer processing file

- Master and transaction file
- Report files
- Backup files
- Reference files



Master and transaction file

- A **file** is a collection of organized data.
- The **master file** contains all the permanent or semi-permanent data relating to a particular application.
- **Transaction files** contains all the transactions that are captured as they occur over the period of time and it is used to update the master file.



Report files

- Are derived from records within a master file or transaction file.

Backup files

- Backups are duplicate copies of important files that are stored in a safe place used to recreate the master file in case in case of loss of loss.
- Backups can take place in different ways:
 1. In **batch processing**; three generations of the master and transaction files are usually kept. Sometimes refers to as **grandfather-son** method of backup.
 2. **Cloud storage/ online update** – data is stored in an external server on the internet.



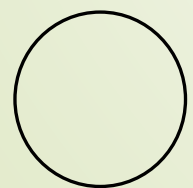
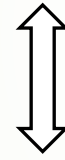
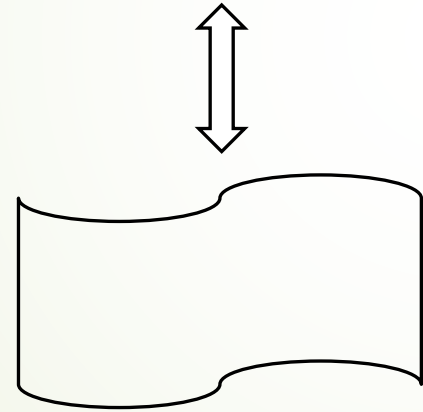
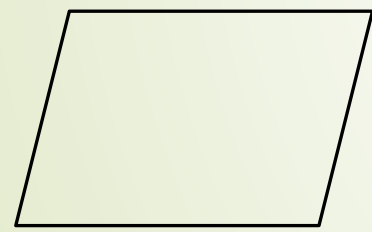
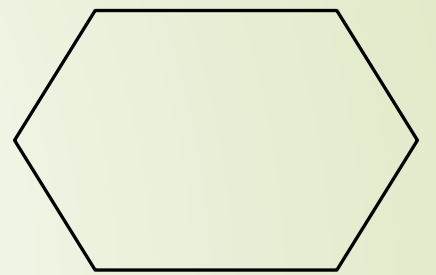
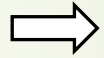
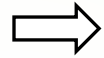
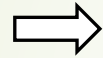
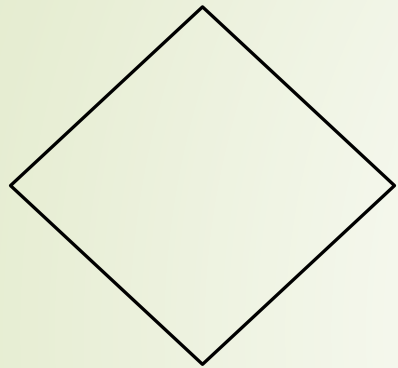
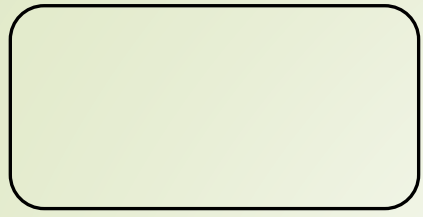
Reference files

- These files contain referential data, such as tables and lists which are necessary to support the data processing of an organization when performing calculations or checking the accuracy of input data. e.g. charts, tables of inventory codes



Electronic data processing modes

1. Online processing
2. Interactive processing
3. Distributed processing
4. Time- sharing processing
5. Batch processing
6. Multiprocessing
7. Multiprogramming
8. Multi-tasking
9. Real-time processing



MARIO@2022



Online processing

- This is an automated way to enter and process data as long as the document is available.
- Each transaction is processed as soon as it is entered.
- It is a real-time method
- It is referred to as an **online transaction processing system** (OLTP)
- Examples of applications that use online processing **bar code scanning**.

Interactive processing

- The user provides the computer with instructions/input during processing and observe the results.

Distributed processing

- Geographically distributed resources such storage devices, data sources, software and multiple independent are connected in a single network to perform a particular

Time- sharing processing

- A computer system is shared among multiple users, giving each other user illusion that they have exclusive control of the system.



Batch processing

- Transaction are accumulated over a period of time in a transaction file.
- The transaction file is used to update the master file in on processing run without human intervention.

Multiprocessing

- The operating system supports multi-core processor multi processing of programs by more than one processor.

Multiprogramming

- It was designed to optimize CPU usage.



Multi-tasking

- The operating system allows a single user to work on two or more programs at the same time.
- Users can run two or more programs concurrently.
- The program in the foreground, which is currently in use is called **active program**.
- The other programs that are running, but not in use are in the background.

Real-time processing

- Transactions are processed as soon as they take place and the relevant master files are updated.
- Example; online banking and ATM transaction.



Exercise

1. Explain the meaning of Real time.
2. List two basic actions that are performed on a current file to create new master file.
3. Name a file used to update a master file.
4. State one application which uses
 - (a) Real time processing
 - (b) On-line processing

Fortune
favors the
ready mind



MARIO CHONGO MAKHANGA
BICTed