DATA PROCESSING.
Data processing deals with how data is organized & processed in the computer.

DATA:
- Data is a collection of facts & figures, which can be processed to produce information.
- Data are the facts relating to an activity in a given environment.
  The activity can be Accounting, Inventory control, etc. Environment can be business, scientific, education, etc.

Examples:
- In an educational environment, when students sit for exams, the grades obtained represent the data to be processed by the computer. In this case, data can be Names of students & Marks obtained.
- In a business environment, data can be the No. of Hours worked, names of employees, Stock.

Data can also be described as Raw data, if they are not yet processed, i.e. if they do not convey particular meaning to a given activity within any given environment.

It therefore means that, Data are unprocessed information consisting of details relating to business transactions. For example, in a Payroll system, data are employee’s names, basic salary, department number, marital status, etc.

DATA PROCESSING:
- The collection, manipulation & distribution of data (i.e.) letters, numbers & graphic symbols, to achieve certain objectives.
  The processing may involve calculations, comparisons, decision-making and/or any other logic to produce the required result.
- The activity of manipulating the raw facts to generate a set of meaningful data (described as Information), which is able to convey some meaning.
- Those activities, which are concerned with the systematic recording, arranging, filing, processing, and dissemination of facts relating to the physical events occurring in a business.

Data processing is a very important activity in any organization of any size or nature because it generates information for decision-making.

If the data processing uses complicated processing tools or aids, e.g. the computer, it is described as Electronic Data Processing (EDP).

INFORMATION.
- Information is data, which is summarized and processed in the way you want it, so that it is useful in your work.
- Information is an assembly of meaningful data items.

The information in Payroll activity includes; Net pay, Total Tax deductions, etc. In Stock Control, the information generated includes; Closing stock, Total cost of the items, Purchases, Sales, etc.

The information is obtained by applying some processing procedures onto the raw data being input. For example, to get the Net pay in a Payroll activity, the procedure would be;

Net pay = (Basic salary + Allowances + Overtime, if any) – Taxes.

Information is the end product of data processing available at the right place, the right time and in the right form.

The information generated by the data processing activities is very important in the working strategies of any organization, because it is used by the organization to make decisions.

Characteristics/ Features of good Information.
Data processing techniques

It should:

(i). Have and serve a purpose.
(ii). Be relevant to its purpose.
(iii). Be complete, accurate, and comprehensive.
(iv). Have been obtained from a reliable source.
(v). Be communicated to the right person and in the right time (i.e. it should be timely).
(vi). Be clear and understandable by the user.
(vii). The user must have confidence in it.

Relationship between Data, Data Processing, and Information.

Data are the facts which relate to any particular activity, and do not have any specific meaning. Information is data with a definite meaning.

Data processing is the process, which transforms data into information.

In a Manufacturing industry, data may be compared to raw materials and Information to finished products. Just as raw materials are transformed into finished products, raw data are transformed into information.

In order to generate information from data items, a set of processing activities have to be performed on the data items in a specific sequence depending on the desired final result. Performing these processes is known as Data processing.

![Data processing cycle diagram]

Exercise.

1. Define the terms:
   (i). Data.
   (ii). Information.
   (iii). Data processing.

2. Distinguish between the following terms:
   (i). Data.
   (ii). Data processing.
   (iii). Information

3. Using examples, explain the difference between ‘Data’ and ‘Information’.

DATA PROCESSING CYCLE.

Data processing cycle refers to the various stages involved in converting data into information.

Basic stages in the Data processing cycle.

There are 5 primary elements/functions of data processing system. They include; Input, Processing, Storage, Output, and Control.

<table>
<thead>
<tr>
<th>ORIGINATION OF DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data originates from Source documents, Time cards, Sales orders, Purchase orders, Invoices, etc</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PREPARATION OF DATA FOR INPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data is Transcribed, sorted, etc.</td>
</tr>
</tbody>
</table>
Data processing techniques

Notes.

Exercise I.
1. (a). What is a Data Processing cycle.
   (b). State and describe the stages involved in data processing cycle.
2. Draw and label a clear flow diagram of the stages involved in a data processing cycle.
3. List the various steps in the data processing cycle and briefly describe what happens at each stage.

DATA COLLECTION.

Data Collection is the process involved in getting the data from the point of its origin to the computer in a form suitable for processing.

Note. Data collection starts at the source of the raw data & ends when valid data is within the computer in a form ready for processing.

DATA CAPTURE & DATA ENTRY.

Data Entry:
Nowadays, most end-users input data to the computer using Keyboards on PCs, Workstations, or Terminals.
Data can originate in many forms, but the computer can only accept it in a machine-sensible form.

Problems of Data Entry.
1. The data to be processed by the computer must be presented in a Machine-sensible form (i.e. in the language of a particular input device).
   Note that most of the data originates in a form that is not machine-sensible. Therefore, the data must undergo the process of Transcription before it is suitable for input to the computer.
2. The process of Data collection involves getting the original data to the “processing center”, transcribing it, sometimes converting it from one medium to another, and finally getting it
Data processing techniques into the computer. This process involves a great number of people, many machines, and much expense.

Data Capture:
Data Capture is the process of obtaining data in a computer-sensible form at the point of origin. Obtaining of data in a computer-sensible form helps to avoid many of the problems of data entry. The captured data may be stored in some intermediate form for later entry into the main computer in the required form. If data is input directly into the computer at its point of origin, the data entry is said to be On-Line. In addition, if the method of direct input is a terminal or workstation, the method of input is known as Direct Data Entry (DDE).

STAGES IN DATA COLLECTION.
The process of data collection may involve any number of the following stages depending on the methods used.

1. Data Creation.
This involves 2 basic alternatives:

(a). Source documents.
Source document is the original document used to record data and/or instructions. Most of the data is in form of a manually scribed or typewritten documents, i.e. the data is on clerically prepared source documents.

(b). Data capture. This involves preparing the source document itself in a machine-sensible form so that it may be used as input to the computer without the need for transcription. The prepared source document is then read directly by a suitable device, e.g. a Bar code reader.
Data capture eliminates the need for transcription.

Note. The method and medium adopted for data creation will depend on factors such as Cost, Type of application, etc.

2. Data Transmission.
This will depend on the method & medium of data collection involved/adopted. If the computer is located at a central point, the documents will be physically “transmitted”, i.e. by the Post office or a Courier to the central point.
The data can also be transmitted by means of Telephone lines to the central computer. In this case, no source documents would be involved in the transmission process.

3. Data Preparation.
Data Preparation is the term given to the transcription of data from the source document to a machine-sensible medium.
There are 2 parts involved in the data preparation:
(a). The original transcription itself, and
(b). The Verification process that follows.

4. Conversion of data from one medium to another.
Data is prepared in a particular medium & converted to another medium for faster input into the computer.
For example; data might be prepared on Diskette, or captured onto Cassette, and then converted to magnetic Tape for input.
The conversion will be done on a computer that is separate from the one for which the data is intended.

5. Input.
Data processing techniques

The data, now in magnetic form, is put into the computer and subjected to validity checks by a computer program before being used for processing.

This stage is required to re-arrange the data into the sequence required for processing.
Sorting is necessary for efficient processing of sequentially organized data in many commercial and financial applications.

7. Control.
In all the stages of data collection, control must be established and applied where necessary. In other words, Control is usually applied throughout the whole process of data collection.

METHODS OF DATA COLLECTION.
The following are alternatives that can be used to collect data:
(i). Use of Data Capture devices such as Scanners, Kimball Tags, Point-of-Sale systems, Barcode readers & Magnetic strip readers.

TYPES OF ERRORS IN DATA PROCESSING.
The System designer must guard against the following types of errors:
(a). Transcription (copying) errors.
(b). Missing source documents.
(c). Source documents whose entries are omitted, illegal and suspicious/doubtful.
(d). Program faults (errors).
(e). Machine hardware faults.

Note. Machine hardware faults are less common because modern computers have self-checking facilities & usually signal any internal failure.

DATA INTEGRITY.
(i). Accuracy:
(ii). Timeliness:
(iii). Relevance:

DATA CONTROL.
The quality of Input data is important to the accuracy of output. Control must be instituted as early as possible in the system & everything possible must be done to ensure that data is complete and accurate before being input to the computer.

Objectives of Data Control.
The objectives of Control are:
(i). To detect, correct and re-process all errors.
(ii). To ensure that all data is processed.
(iii). To preserve the integrity/reliability of maintained data.
(iv). To prevent and detect fraud/deception.

Note. Control must be designed into the system & thoroughly tested. Failure to build in adequate control may cause expensive systems to fail. In addition, all users must be fully consulted to ensure that adequate controls are implemented.

Types of Data Controls.
The following are controls that can be used to ensure data accuracy:
(1). Verification:
Data processing techniques

This is the process of checking & ensuring that data has been transcribed/ written out correctly.

Verification is whereby several computer users are given data to enter into the computer and the results are compared. Or else, a second transcription is compared with the first one. If the results are different, then there is inaccuracy in that data. This method is mostly used to verify password changes.

Note. Verification calls for manual intervention, hence errors are possible. Note that some copying/transcription mistakes that bypass the verification stage are difficult to isolate during verification, e.g. the confusion of l (letter l) and 1 (one). In this case, l might be input instead of 1 and vice versa, hence such mistakes go undetected.

The main types of errors, which might occur are:

(i). Missing data.
(ii). Duplicating of data.
(iii). Use of outdated records.
(iv). Incorrect batches of input data.
(v). Incorrect recording at the source.
(vi). Incorrect data preparation.

(2). Manual controls.

This involves considerable checking of the source documents. Such checks may be:
- Inspecting the source documents to detect missing entries, illegible entries, illogical or unlikely entries.
- Comparing the document against stored data to verify entries.
- Re-calculating to check calculations made on the document.

VALIDATION CHECKS.

A Computer cannot notice errors in the data being processed in the way that a Clerk or Machine operator does.

Data validation is the process of preventing wrong data from being processed. It involves checking whether the results generated by the computer are valid or applicable. During input or data preparation, the data must be checked for transcription errors, through a process known as Verification.

Once the data is brought into the computer memory directly from an input device, immediately before processing, the data is again subjected to checks built in the program described as validation checks, to check the data integrity or the conformity of the data to the processing requirements.

Data validation includes testing for the following:

(a). Test for reasonableness.
    The computer program checks whether the data is reasonable, e.g., number of people should not be represented in decimals, i.e. 9½ children.

(b). Test for numbers.
    E.g., numbers should not be given as alphabets.

(c). Test for alphabets.
    E.g., alphabets should not be represented as numbers.

These checks can be made at 2 stages:

(1). Input stage: When data is first input to the computer, different checks can be applied to prevent errors going forward for processing. For this reason, the first computer run is often referred to as Validation or Data vet.
(2). **Updating stage:** Further checking is possible during data processing (or when the data input are being processed).

The program checks the consistency of the input data with existing stored data. This check is possible during the input run if the stored data is on-line at the time.

**Note.** Validation is an online process (i.e. validation checks are build into the computer programs using the input data, so that incorrect data items are detected and reported). Since the checks are under the influence of the computer, they are not prone to errors.

**Exercise I.**

1. Distinguish between Data verification and Validation as used in the context of data collection.

**METHODS OF DATA PROCESSING.**

(1). **MANUAL SYSTEMS.**
Data processing techniques

In Manual systems, the data processing activities are carried out manually by the human Clerks assisted by some calculating tools such as Slide rule, Logarithms, etc.

In individual business units, the transactions are recorded on the source documents, which are taken to the data processing department for processing. Human beings work on source documents mentally or with the aid of some simple manipulation tools.

The files maintained are updated appropriately to reflect the correct image of the business.

The records are stored in form of Ledger cards, in the filing trays or in cabinets. The Ledger cards contain the sales data (the amount owed by customers) and purchases data (the amounts owed to suppliers).

The Information (in the form of business documents) is generated, e.g., Statements of Accounts, and sent to the customers.

Control is carried out/ monitored by the Supervisor guided by the instructions written down in a Procedure manual.

In Manual systems, the data being used by one individual becomes inaccessible to another individual.

(2). MECHANICAL SYSTEMS.

Mechanical systems are data processing systems whose activities are carried out by Keyboard devices operated by human beings. The devices include; Accounting machines, Cash registers, Calculators, etc.

Data is keyed in by the Machine operator, manipulated by the machine, and the output is obtained in form of printed documents.

Once the machine is switched on & given the relevant instructions, it works on the data input automatically.

Note. The instructions, in this case, may be pressing the relevant Keyboard button, e.g. pressing the button for addition, after a set of values have already been keyed in or as they are being keyed in.

The control activity is carried out automatically by the machine itself or by a human machine-operator guided by the instructions laid down in a Procedure manual. Other control strategies include; Self-experience on the job and Supervision.

(3). ELECTRONIC SYSTEMS.

Electronic Data Processing (E.D.P) systems use electronic machines, such as Computers, to process data. This is because of the volume of data to be processed, and timing of the information expected from such processing activities.

Data that is to be input into an E.D.P system should be first prepared into machine-sensible form. This means that, data cannot be input directly through the terminal or Keyboard, connected online to the computer system. In such a case, the Key-to-disk data preparation method could be used. The contents of the disk are input using the reading/writing unit of the disk. The disk pack is mounted onto its drive and the computer is activated to read or transfer the contents of the disk into its memory, where the data are held temporarily to await processing based on the instructions given.

The processing is done automatically by the computer under the influence of a set of instructions (programs).

The master files are stored in the mass storage media, e.g. Disk. The disk contents are updated accordingly during the processing run.

The type of output generated by the E.D.P system is influenced by the type of output device used, e.g. hardcopy outputs are produced through the Printers, while Softcopies are produced through the Screen displays.

The control of the Electronic systems is automatic under the influence of the Control unit (CU) of the computer, whose actions are influenced by stored programs.
Comparison between Manual, Mechanical and Electronic Systems.

<table>
<thead>
<tr>
<th>Functions</th>
<th>Manual Systems</th>
<th>Mechanical Systems</th>
<th>Electronic Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
<td>- Human observation</td>
<td>- Human observation</td>
<td>- Keyboard</td>
</tr>
<tr>
<td></td>
<td>- Keyboard</td>
<td>- Keyboard</td>
<td>- Data entry terminals</td>
</tr>
<tr>
<td></td>
<td>- Written records</td>
<td>- Written records</td>
<td>- Document readers</td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>- Human brain</td>
<td>- Conventional machines, e.g. Calculators, Cash registers, etc</td>
<td>- Computer’s C.P.U.</td>
</tr>
<tr>
<td></td>
<td>- Calculators</td>
<td>- Written procedures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Written procedures</td>
<td>- Calculator display.</td>
<td></td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>- Human voice</td>
<td>- Human voice</td>
<td>- Visual Display unit</td>
</tr>
<tr>
<td></td>
<td>- Written documents</td>
<td>- Written reports</td>
<td>- Voice Output</td>
</tr>
<tr>
<td></td>
<td>- Calculator display.</td>
<td>- Printed reports.</td>
<td>- Printed output.</td>
</tr>
<tr>
<td></td>
<td>- Written records</td>
<td>- Written procedures on manuals.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Personnel experience.</td>
<td>- Personnel experience.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Supervision</td>
<td>- Supervision</td>
<td></td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>- Human brain</td>
<td>- Manual files or folders.</td>
<td>- Computer’s Main memory.</td>
</tr>
<tr>
<td></td>
<td>- Manual files</td>
<td></td>
<td>- Storage media - disks, tapes, etc</td>
</tr>
<tr>
<td></td>
<td>- Written records</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control</strong></td>
<td>- Written procedures on manuals.</td>
<td>- Machine influence</td>
<td>- Computer’s Control unit (CU).</td>
</tr>
<tr>
<td></td>
<td>- Personnel experience.</td>
<td>- Written procedures on manuals.</td>
<td>- Stored programs.</td>
</tr>
<tr>
<td></td>
<td>- Supervision</td>
<td>- Personnel experience.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Supervision</td>
<td></td>
</tr>
</tbody>
</table>


The following are the factors, which may necessitate the change from Manual to Mechanical or to Electronic data processing method:

(i). **Operation Speed.**

The timing aspect of information availability (i.e. when the information is required) is very important.

Electronic & Mechanical systems provide automatic processing of the input data. This quickens the operations on the input data to produce timely information.

For example, a Clerk assisted by mechanical or electronic devices takes shorter time to complete the posting of a transaction.

(ii). **Accuracy of the information.**

The use of mechanical or electronic data processing tools makes information more accurate & neat, by removing the use of illegible hand written entries.

In addition, verification is made easy; hence wrong data are easily prevented from entering the processing stage.

(iii). **Volume of data.**

The data processing method selected should be able to cope with the processing tasks, in respect to the data held. The data (records) of an organization depends on the size & the nature of the business.

Small organizations with low volumes of data, require few personnel with little or no data processing aids.

Large or complex business organizations, with high volumes of data, require the use of sophisticated processing tools, if the information is to be produced on time.

(iv). **Convenient.**
Data processing techniques

Data processing that requires repeated operations may be boring & tedious when carried out manually. In such a case, mechanical or computer machines may be employed to assist in the processing depending on the nature of the business.

(v). Linked Applications.

In a situation where there is a common data pool that supports several applications, and e.g., Manual D.P method is used, then different operations may be required to produce different informations. However, if Electronic D.P method is used, the informations can be easily produced from the same data. This is because, the computer is versatile, and can operate in any desired manner provided the relevant programs are available.

(vi). Better services to customers.

As Data processing systems produce information, the recipient of such information should receive them immediately to enable them take decisions that control their business operations.
Using the sophisticated processing aids, such as Computer as in Electronic D.P systems, improves the quality of information produced, e.g. statistical summaries are produced in good time, enquiries are answered in good time, and orders are dispatched promptly.

Factors that determine the Methods of Data Processing.

The following are the factors that influence the method of data processing selected:

(i). Size and Type of business.
(ii). Timing Aspects of the information produced from the system.
(iii). Link between Applications.

Size and Type of Business.

Simple or small business organizations require relatively fewer personnel and processing methods that are less complicated.
In a very small company, a single person can be used to produce all the information required, but as the volume of business increases, more people and tools/aids in the form of Calculators and small Computers may be employed. Large volumes of data and information will require the use of large computers.

For example;
In some companies, the Payroll may involve paying a member of staff the same amount each month, while in others a complex payment system may be involved.
Similarly, producing an Invoice may be a matter of simply copying from the customer’s order, or it may require complex discount calculations.
Simple calculations indicate the need for fewer people and tools to produce the information, while complex situations indicate the need for more people and aids.

Timing Aspects of the information produced.

Some applications/ jobs require much shorter time between the origination of the transaction and the production of information (e.g. Hotel bookings), while other business applications may require the information to be made available after a relatively longer period, e.g. in Passport application, where information is required periodically.
Some information requirements are less important than others. E.g., the Payroll and Statement of Accounts may only be produced once a month, whereas in certain companies, the Invoices may be produced all the time (i.e. as a customer collects the goods).

Link between Applications.

In some applications, the same data items may be used in producing more than one information; hence, the most suitable data processing system should be used depending on circumstances surrounding these information requirements.
Data processing techniques
E.g. a particular item sold may be needed to produce the Invoice & to amend the recorded Stock position (i.e. to make adjustment of Stock level, and the Bank account or Cash account).

Exercise I.
2. Distinguish between Manual, Mechanical and Electronic systems.
3. Describe the reasons for changing from Manual to Mechanized or Electronic systems.

Exercise II.
1. Write short notes on the following:
   (i). Manual systems.
   (ii). Mechanized systems.
   (iii). Electronic systems.
2. By use of a clear table and brief explanation, show the differences between manual, mechanized and electronic systems touching on the following functions: input, process, output, storage and control. (20 marks)

COMPUTER FILES
A File is a collection of related records (i.e. several records put together) that give a complete set of information about a certain item or a particular business entity.
Files are important in any business because; they provide up-to-date information relating to the entity sets of the business, e.g., the suppliers, employees, customers, etc of the organization.

Entities are things whose facts need to be recorded. Each entity has its attributes (i.e., individual properties), e.g., Employee (which is an entity) has attributes such as; Name, salary, address, etc.

A file can be stored manually in a file cabinet or electronically in a computer’s secondary storage device such as a Floppy disk or hard disk.

**Advantages of computerized filing system over manual filing systems.**

1. Information takes up less space than the manual filing.
2. It is much easier to update or modify information.
3. It offers faster access and retrieval of data.
4. It enhances data integrity.
5. Reduces duplication of data, or of the stored records

**Logical and physical files**

Computer files are classified as either; Logical or Physical.

*Logical files.*

A Logical file is a type of file viewed by the user in terms of what data items it contains & what processing operations may be performed on the stored data items.

*Physical files.*

A Physical file is viewed in terms of how the data items found in a file are arranged on the surface of the storage media (e.g., disk, tape), and how the stored data items can be processed.

**DATA HIERARCHY.**

In data processing, data is organized from the smallest element to the most comprehensive.

Bits

Characters

Fields (a group of characters)

Records (a collection of fields)

File (a combination of related data records maintained in some pre-arranged order)

Database (consists of several related and integrated data files)

*Bit:*  
A Bit is the smallest item that can be stored in a physical file. The bit can either be a ‘0’ or a ‘1’; the two states that define the storage cells of a computer memory & a storage media.

Bits combine together to form the Byte (which is the unit of measuring the computer storage). A Byte is the collection of several bits that represent a Character.

**ELEMENTS OF A COMPUTER FILE.**

A computer file is made up of three elements:

1. Characters.
2. Fields.
3. Records.

Character
A character is the smallest element in a computer file, and can refer to a letter, number, & symbol that can be entered, stored and output by a computer. A character is formed by several bits combined together, depending on the character coding system used, e.g., in a 6-bit character coding system, a character is represented by a combination of 6 bits. Characters are normally used to represent data items such as Names, Prices, etc.

Fields
A field is an item of data or information consisting of one or more characters. A Field is made up of a combination of characters, and forms the attribute of a given entity, e.g., in a student’s record, the students Admission number is a field.

There are 2 types of fields;
- **Fixed length fields** – these are fields with the same numbers of characters.
- **Variable length fields** - fields within a record that are made up of different numbers of characters (i.e., fields with different spaces allocated for their characters).

Records
A record is a collection of related fields, which together form or represents a single entity. In any particular file, there is a separate record for each entity, e.g., in a class score sheet, the details of each individual student in a row such as name, admission number, total marks, and position form a record.

There are 2 types of fields;
- **Fixed length records** – records in a file that are made up of the same number of fields.
- **Variable length records** - records that have different number of fields making them. If the records have different spaces preserved for them, then it implies that, all the records in the file will not have the same size.

Note. Variable length records normally utilize the storage efficiently. However, processing or updating them in a computer is difficult because; the programmer is dealing with unknown quantities.

On the other hand, fixed length records do not utilize the storage efficiently, but they are easy to process because; the programmer is dealing with known character quantities.

**TYPES OF COMPUTER PROCESSING FILES.**
There are various types of files used to store data needed for processing. Data processing files are classified according to:
- Their uses within the overall data processing activities.
- The kind of data/information they store.

The main types of data processing files include:
1). Master files.
2). Transaction files.
3). Reference files.
4). Sort files.
5). Backup files.
6). Scratch files.
7). History files.
8). Report files.
Master files.

A Master file is the main file that contains relatively permanent records about particular items or entries against which transactions are processed.

Master files contain records, which have long-term significance, and are very important for the running of the organization.

Master files normally contain 2 types of data: Static data and Dynamic data.

(a). Reference (Static) data:

Static data is relatively permanent, and contain details which do not change, e.g., Name, Sex, Date of birth, Date of hiring, etc.

Static data is processed by amending (i.e. making occasional changes to) the existing records, e.g., inserting new records, deleting outdated records, etc.

(b). Dynamic data:

Dynamic data is temporary and is likely to change frequently, e.g., Salary, Tax rates, hours worked, Rate of pay, etc.

Dynamic data is processed by updating (i.e. changing the values of the various fields).

The accuracy of data within the operational files is achieved by Updating the Master file (i.e., changing the contents in the master files regularly in order to reflect the current state of affairs). This involves adding, removing or adjusting the data in the Master file.

Transaction (Movement) files.

A Transaction file contains individual data about the transactions (activities) that occurred in a business during a particular period of time.

The file contains relatively temporary information such as all incoming or outgoing records resulting from a transaction.

Transaction files are usually created from the source documents, which contain data from the point of their origin.

The contents in a Transaction file are used to update the dynamic data on Master files. For example, in a busy supermarket, daily sales are recorded on a transaction file, and later used to update the stock file. The file is also used by the management to check on the daily or periodic transactions.

Transaction files have a short life span. This is because, once the contents of the file have been used to update the master file, its contents are no longer required, and can be replaced by the next business transactions.

Examples of transaction files:

Files that contain Earnings & deductions of an Employee, or payments received from customers.

Reference files.

A Reference file is used for reference or look-up purposes.

Lookup information is that information which is stored in a separate file, but is required during processing. E.g., the item code entered either manually or using a bar-code reader in a point-of-sale terminal is used to look-up the item description & price from a reference file stored on a storage device.

Reference files contain records that are fairly permanent or semi-permanent such as tax deductions, Wage rates, Customer address, etc, and therefore, they need to be revised occasionally.

Backup files.

A Backup file is used to hold duplicate copies (backups) of data or information from the computer’s fixed storage (hard disk). These files are kept for security purposes.
Data processing techniques

This is because; the operational files held on the hard disk may be corrupted, lost or changed accidentally leading to loss or damage of existing information. It therefore important to keep copies of the recently updated files so that, in case the original file is corrupted or deleted, the backup file can be used in its place or to reconstruct the original file.

Note. The backup file & the operational file should be kept at separate places so that in case of loss or damage, both are not affected.

Sort files.
Sort files are created from existing files, such as Master or Transaction files, and are used mainly for sorting data (i.e., they are used to alter the sequence of the existing files).
A sort file is mainly used where data is to be processed sequentially. In sequential processing, data or records are first sorted and held on a magnetic tape before updating the master file.

Report files.
A Report file contains a set of relatively permanent records extracted from the data in a Master file or generated after processing.
Report files are used to prepare reports, which can be printed at a later date.
Example of Report files:
Report on Overtime, report on Taxes, report on student’s class performance in the term, etc.

Scratch file.
A Scratch file is a temporary file used to hold data during processing. It contains temporary data, which can be erased when the task is finished.

History (Archive) files.
History files are usually old files retained for historical use or for reference purposes, e.g., it can contain Employee details for the last 10 yrs.

Key field.
A Key field is one or more fields in a record that uniquely identifies the record or a group of records.
E.g., an Employees Serial number may be used to identify the employee records in a Payroll file.
Note. Any field in the record can be used as the key field. However, it should display unique identification characteristics.

Review questions
1. Define a computer file.
2. State four advantages of storing data in computer files over the manual filing system.
3. Differentiate between Logical file structure and physical file structure.
4. With the help of a figure, illustrate the information system Data hierarchy.
5. Define the following terms:
   (i). Character.
   (ii). Field.
   (iii). Record.
   (iv). Key field.
6. List 5 types of files used in data processing and their purposes.

FILE ORGANIZATION

File organization refers to the way records are arranged (laid out) within a particular file.
Data processing techniques

The term file organization can also refer to the relationship of the Key of a record to the physical location of that record in the computer file.

File organization is very important because; it determines the method of access, efficiency, flexibility, and storage devices to be used.

Methods of file organization.

There are 4 methods by which records of a file can be arranged and accessed. These include:

1. Random.
2. Serial.
3. Sequential.
4. Indexed sequential.

Random (Direct) file organization.

In Random or direct file organization, the records are stored in the file randomly, and in no particular order. This implies that, there is no relationship between two adjacent records.

An Algorithm (mathematical procedure) is applied onto the record key to generate the address of the location where the record would be stored.

<table>
<thead>
<tr>
<th>Record 2</th>
<th>Record 3</th>
<th>Record 8</th>
<th>Record 92</th>
<th>Record 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>K₂</td>
<td>K₃</td>
<td>K₈</td>
<td>K₉₂</td>
<td>K₁</td>
</tr>
</tbody>
</table>

K’s – Record keys.

Random files are usually accessed directly. To access the file, the record key is used to determine where a record is stored on the storage media. Once the record is located, it is then read into the computer memory.

This method is used by Magnetic disks and Optical disks.

Advantages of Random file organization.

1. Records are quickly accessed (i.e. there is fast access to records).
2. Files are easily updated (i.e. adding, deleting, and amending the records is easily achieved).
3. The method does not require the use of indexes, hence saving space.
4. Transactions do not need to be sorted before being updated.
5. New records can be easily inserted into a random file.

Disadvantages of Random file organization

1. Data may be accidentally erased or overwritten unless special precautions are taken.
2. Random files are less efficient in the use of storage space compared to sequentially organized files.
3. Expensive hardware and software resources are required.
4. Relatively complex when programming.
5. System design based on random file organization is complex and costly.

Serial file organization.

In Serial file organization, records in a file are stored one after the other in the order they come into the file without any particular sequence. The records are not sorted in any way on the storage medium, and there is no relationship that exists between adjacent records.

This type of organization is mostly used on Magnetic tapes.

<table>
<thead>
<tr>
<th>Record 1</th>
<th>Record 2</th>
<th>Record 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>File ‘head’</td>
<td>File ‘tail’</td>
<td></td>
</tr>
</tbody>
</table>

Serial files can be accessed serially. This involves searching through the entire file record by record starting from the ‘head’ of the file towards the ‘tail’ of the file.
Data processing techniques

Note. Serial access is suitable where all the records in the file are to be read. This is because; even the records that are not required must be passed over before locating the record of interest. E.g., to access the 10th record in the file, then the computer reads the first 9 records before reading the 10th record. Therefore,

**Sequential file organization.**

In *Sequential file organization*, the records are arranged within the file serially one after the other. However, in sequential file organization, the records are stored in a particular order sorted using a key field; hence, there is a relationship that exists between adjacent records and the key fields.

<table>
<thead>
<tr>
<th>Record 1</th>
<th>Record 2</th>
<th>Record 3</th>
<th>Record 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>K1</td>
<td>K2</td>
<td>K3</td>
<td>K4</td>
</tr>
</tbody>
</table>

K₁ – K₄ – Record keys.

Sequential files are accessed *sequentially*, i.e. the key field is used to search for the particular record required. Searching starts at the beginning of the file and proceeds sequentially towards the ‘tail’ of the file, until the required record is located.

**Advantages of Sequential organization.**

1. The method is simple & easy to understand.
2. Sequential files are easy to organize and maintain.
3. Loading or reading a record requires only the Record Key.
4. It is efficient & economical if the number of file records to be processed is high.
5. Relatively inexpensive Input/Output media and devices may be used.
6. Errors in the files remain localized.

**Disadvantages of Sequential organization.**

1. The entire file must be processed even when the no. of file records to be processed is low.
2. Transactions must be sorted in the sequence of the Master file before they can be processed or updated.
3. Data redundancy/idleness is high since the same data may be stored in several files sequenced in different keys.
4. Random enquiries are almost impossible to handle.

**Indexed Sequential file organization.**

The records are arranged sequentially as in sequential files. However, indexed sequential files have an Index that enables the computer to locate individual records on the storage media. An Index is the address of a particular cylinder or track. The indexes are used to point at the portions where the records are stored in groups. This allows a group of records that are not required in a particular processing run to be bypassed.

<table>
<thead>
<tr>
<th>Record 1</th>
<th>Record 2</th>
<th>Record 3</th>
<th>Record 4</th>
<th>Record 5</th>
<th>Record 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>K₁</td>
<td>K₂</td>
<td>K₃</td>
<td>K₄</td>
<td>K₅</td>
<td>K₆</td>
</tr>
</tbody>
</table>

a, b, c – indexes. K₁-K₆ – Record keys.

To access a record in an indexed sequential file, the Index and the record’s key field are used by the computer to search for the required record before it is read into the computer memory.

**Methods of accessing Indexed sequential files.**

Indexed sequential files may be accessed using 3 methods;
(a). Sequential access.
(b). Selective sequential access.
(c). Random access.

**Sequential access:**
In *sequential access*, the computer reads the records in sequential order (i.e., one record after the other) using the index until the record matching the search key is found. The record is then read into the Main memory.

Sequential access is suitable for high activity files.

**Selective Sequential access:**
In this *selective sequential access*, the transaction file must first be sorted into the same key sequence as the master file. The access mechanism then goes forward in an ordered progression (sequence), and only those records needed are read/processed.

The method is suitable for low activity files.

**Random (direct) access:**
The records in a Random file are not stored in any particular sequence of the key field. This means that, the records can be processed in any sequence, i.e., by moving access mechanism forward and backwards along the file in a non-orderly manner to access the records required.

The method is suitable for low activity files.

**Advantages of Indexed sequential file organization**
1. Records can be accessed sequentially or randomly.
2. Accessing of records can be fast, if done randomly.
3. Records are not duplicated.

**Disadvantages of Indexed sequential file organization.**
1. Accessing of records sequentially is time consuming.
2. Processing of records sequentially may introduce redundancy/idleness.
3. Required expensive storage medium.

**Comparison of File organization methods.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Method of access</th>
<th>Medium used</th>
<th>Example of file</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Serial</td>
<td>Serial</td>
<td>Magnetic tape. Magnetic disk</td>
<td>Unsorted transaction file</td>
</tr>
<tr>
<td>3. Sequential</td>
<td>Serial (Sequential)</td>
<td>Tape, Disk</td>
<td>Sorted transaction file, or Sequential Master file.</td>
</tr>
<tr>
<td>4. Index Sequential</td>
<td>(a). Sequential (b). Selective sequential (c). Random</td>
<td>Magnetic disk</td>
<td>Master files requiring various processing activities.</td>
</tr>
</tbody>
</table>

**File organization & access on a Magnetic Tape.**
In a Magnetic tape, the file records are placed one after the other onto the tape.
There are 2 ways in which files are arranged on tapes:

1). **Serial:**
   - In *serial organization*, the records are written onto the tape without having any relationship between the record keys.
   - ✓ Serial files on a tape are accessed *serially*, i.e., each record is read from the tape into main storage one after the other in the order they occur on the tape.

2). **Sequential.**
   - In *Sequential organization*, the records are written onto tape in sequence according to the record keys. Sequential files are accessed *sequentially*.

   **Explanation:**
   To process a sequential Master file on a tape, the transaction file must be in the sequence of the Master file. The transaction file is read first, followed by the Master file until the matching file record is found. E.g., if the record required is the 20th record of the file, the computer must first read all the 19 preceding records.

**File organization & access on a Magnetic Disk.**

There are 4 basic methods of organizing files on a Magnetic disk:

1). **Serial:**
   - The records are placed onto the disk one after the other with no regard for sequence.
   - ✓ Serial files on a disk are accessed *Serially*, i.e. each record is read from the disk into main storage one after the other in the order they occur on the disk.

2). **Sequential:**
   - In *sequential organization*, the records are written onto the disk but in a defined sequence according to the record keys.
   - ✓ The Sequential method of access is used to read a sequential disk file.

3). **Random:**
   - In *random organization*, the records are placed onto the disk “randomly”, (i.e. there is no obvious relationship between the records).
   - A mathematical formula is used to generate the address of the location where the record is placed on the disk. During processing, the same record key is used to generate the address which shows the location from which the record is read.
   - ✓ The method of access to random files is **Random (direct) access**.

4). **Indexed Sequential:**
   - In *Indexed Sequential organization*, the records are stored in sequence, but an Index (key field/guide) is provided to enable individual records to be located. In this case, the index will always enable the sequence of the records to be determined.
   - Indexed sequential files can be accessed using *sequential access, selective sequential access,* or *random access* method.

**Factors to consider when choosing the type of file organization to use.**

1). *Frequency of update.*
The file designer should determine how often the file is going to require updating. For periodic updates (e.g., monthly update), the transactions are used to update the master files in one run. For the non-periodic systems, the transactions may be updated anytime as required. The file design selected should therefore be able to meet the update strategies, and at the required time.

2). **File activity.**

The type of file organization adopted should be based on the expected number of records to be processes/accessed in a particular run.

3). **Method of file access.**

This refers to the method the computer shall use to transfer the contents of the file from the storage media into the computer.

4). **Nature of the system.**

Before designing the file(s) to be maintained by a computer system, you have to consider whether the system runs periodically or is an event-driven system. In periodically run systems, all transactions relating to particular business are accumulated over a period of time, after which they are applied to the relevant master files in a single run. Such systems produce periodic reports from the maintained files. On the other hand, event-driven systems allow file enquiries and instant update so long as the transactions are available from the maintained master files for the production of instant information.

5). **Medium for storing the Master file.**

Computer files are stored in the storage media. The type of file organization adopted depends on the medium that will be used to store the computer file. E.g., Serial access devices, such as Magnetic Tapes cannot be used to store Random files or Indexed-sequential files. This is because; searching for the particular record required proceeds serially regardless of the file organization method used.

**Review Questions.**

1. What do you mean by File Organization?
2. State and explain four types of file organization.
3. Distinguish between:
   - (a) Sequential and serial file organization methods.
   - (b) Random and indexed-sequential file organization methods.
4. (a) Describe how files are organized and accessed on tape.
   (b) What are the disadvantages of storing files on tape?
5. Differentiate between Sequential and Indexed Sequential methods of file organization on disk.
6. (a) What is random file organization? State its advantages.
   (b) How are Random files accessed on disk?
7. Identify four file processing methods.
8. Discuss four considerations for choosing a file organization method.

**ELECTRONIC DATA PROCESSING MODES**
Data processing techniques

Data processing modes describe the ways in which a computer, under the influence of an operating system, is designed to handle data or transactions during processing.

Types of electronic data processing modes

1. Batch processing (also referred to as Sequential or Offline processing)
2. Online processing
3. Real-time processing
4. Time-sharing
5. Multi-programming (also referred to as Multi-tasking)
7. Distributed processing
8. Interactive processing

Review questions

1. Define the term “Data processing modes”.
2. Mention five types of electronic data processing modes.

Batch processing

In batch processing, data or transactions are collected & accumulated together over a specified period of time, e.g., daily, weekly, or monthly. The data is then input & processed at once (or as a single unit) to produce a batch of output.

For example:

In a payroll processing system, details of employees such as number of hours worked, rate of pay, may be collected for a period of 1 month, after which they are used to process the payment for the duration worked.

Data collection is usually done off-line (i.e. away from the CPU) on special machines known as Data entry terminals. The data is entered & stored on a disk in a batch queue for a while. It is then input & processed one or more at a time under the control of the Batch operating system, and the result obtained.

Batches of transactions are scheduled for processing by assigning them priorities. The priorities are assigned in terms of percentage ratio, e.g. 95%, 60%, etc. The most priority jobs are processed first, while the less priority jobs are processed once the computer resources (i.e., CPU time, Memory & I/O devices) are released by the most priority jobs.

Once the processing of a given batch starts, there is no interaction between the operator & the CPU. Therefore, the user cannot intervene to perform amendments to the program.

A job is not processed until it is fully input. In addition, a program must wait its turn before processing the data. This means that, there will be a delay in obtaining results. For instance, a job may wait in the batch queue for minutes or hours depending on the workload. Hence, Batch processing cannot be used when the results are needed immediately.

Characteristics/ Features of Batch processing system.

✓ The input device does not necessarily need to be connected to the computer.
   If the device used for data entry is not connected to the computer, it is said to be Off-Line (away from the computer).
✓ The data is not immediately input into the computer, and it is not even immediately recorded in a machine-readable form.
✓ The speed of processing is not important. This implies that, processing of the data is done at whatever time is most convenient.

Application areas for Batch processing systems.
Batch processing is commonly applied in:

1. **Payroll systems.**
   The attendance data of each employee is collected regularly. It is then input weekly or monthly as per the demands of the system, processed, & then the pay figures for each employee is obtained.

2. **Printing systems** (to print documents)

**Advantages of Batch processing.**
1. Batch systems are easy to develop.
2. Processing of data in batches is efficient & economical.
3. The cost of processing per unit is low.
4. Batching provides manageable units for control purposes.
5. Timing of the information (reports) is not a necessity.

**Disadvantages of Batch processing.**
1. There are delays in obtaining information.
2. It leads to overloading of the processing facilities.
3. Late information is not suitable in situations where instant decisions are required.
4. It is difficult to provide the scheduling of the desired priority.

**Review questions**
1. Briefly explain Batch processing.
2. Describe the application, advantages and disadvantages of batch processing.

**Online processing**
In *online processing*, data or the input transactions are processed immediately they are received to produce the information required.

Online processing occurs when the transactions are processed to update (or make any change in) a computer file immediately after the transactions occur.

In online processing, all the Input/Output facilities, and communication equipments are under direct influence of the central Processor.

In online processing, the operator communicates directly to the computer’s operating system using commands, which are then interpreted by the supervisor. This means that, the operator can interact with the system at any point of processing using the Input/Output facilities.

**Note.** In online processing, the data input units (terminals) are connected directly to the central computer using communication links.

In such a configuration, the data (input transactions) are communicated from the workstations to the central computer for processing, & the results communicated back to the workstations through the telecommunication links.

**Characteristics of Online processing system.**
✓ The input device is connected directly to the computer.
✓ The input data is processed immediately. Processing is completed within a short time (usually 1 or 2 minutes), depending on the speed of the system.

**Application areas for online processing systems.**
1. **Banking:**
   A bank customer can make an inquiry using an online terminal. The system would then respond immediately by accessing the relevant file, and inform the customer on the status of his/her account.

2. **Stock exchanges:**
Data processing techniques

Terminals located in major stock exchanges throughout the country enables quick processing of shares dealings.

3. **Stock control:**
Terminals located in warehouses enable stock records to be re-ordered automatically, make reservations, follow-up of outstanding orders, & print picking lists.

4. **Manufacturing plants:** - to control the progress of work.
5. **Inventory status:** - i.e., ordering & reporting of geographically dispersed distributors.

**Advantages of online processing.**
1. Files are held online; therefore the information generated can be used to update the master files directly.
2. The Information is readily available for immediate decision-making.
3. File enquiries are possible at any given time through the terminals (workstations).

**Disadvantages of online processing.**
1. Online systems are complex to develop.
2. They are costly in terms of hardware, software, storage media, operating system, communication facilities, etc.

**Review questions**
1. (a). Discuss Online processing.
   (b). Mention and explain the Application, Advantages and disadvantages of Online processing mode.

**Real-time systems.**

A **Real-time system** is capable of processing data so quickly such that the results (output) produced are able to influence, control, or affect the outcome of the activity or process currently taking place.

In a **Real-time data processing system**, the computer receives & processes the incoming data as soon as it occurs, updates the transaction file, and gives an immediate response that would affect the events as they happen.

The input-originating workstations may be connected directly to the central processor by appropriate communication equipments. In this case, a transaction is processed & completed immediately or at the same time it occurs. It also ensures quick update to the affected files (records).

The main purpose of a real-time processing is to provide accurate, up-to-date information, hence; better services based on a real situation.

**Requirements of a real-time processing system.**
1). There must be a direct connection between Input/Output devices & the central Processor.
2). The Response time should be fairly fast, to allow a 2-way communication (interaction) between the user & central processor.

**Characteristics/features of real-time data processing.**

✓ Provides immediate response on receipt of a message, e.g., enquiries may be answered immediately.
✓ The system can respond very quickly to a change of circumstances and initiate a feedback.
✓ They have complex & expensive Operating systems, and usually involve **Multi-processing** (use of 2 or more processors) to share the workload.
✓ Real-time systems must be fast and reliable.
✓ Real-time systems are dedicated to one type of application.
Application areas for real-time systems.

Real-time systems can be used in applications that require:

- Quick response.
- A much short time cycle before the information is available to effect the functioning of its environment.

Examples:

1. Airline reservation systems:
   An individual cannot be booked before enquiring whether the airline seat is available. The customer may request for an airline booking information through a remote terminal and the requested information will be given out immediately by the reservation system. If a booking is made, the system immediately updates the reservations file to avoid double-booking, and sends the response back to the customer immediately. This implies that, before the next transaction can be processed, the files must have been updated by the previous transactions.

2. Theatre (cinema) booking.
3. Hotel reservations.
4. Bank systems for making enquiries about customers’ accounts.
5. Police inquiry systems.
6. Chemical processing plants:
   If the temperature is to be regulated, then a rise in temperature requires immediate reaction for correction to restore it to an equilibrium state.

7. In Missile control systems
8. In Hospitals - to monitor the progress of a patient.

Advantages of real time systems.

1. The information is readily available for instant decision-making.
2. Provides better services to users or customers.
3. It is a fast & reliable mode of data processing.
4. Validation of data can be done instantly.
5. Reduces circulation of hardcopies.
6. Provides up-to-date information.

Disadvantages of real-time systems.

1. Real-time systems are very expensive & require complex Operating systems.
2. The systems are not easy to develop.
3. They require large communication equipments, e.g., they require a Front End Processor (FEP), which is used to relieve the central computer by handling some of the limited processing activities, and also link the terminals to the central computer.
4. Real-time systems use 2 or more computer processors sharing up the workloads, which are expensive.

Review questions

1. (a). Name two industries that extensively use Real-time processing.
   (b). Name 3 advantages of a Real-time system.

Time-sharing systems
Data processing techniques

In *time-sharing processing*, the central processor allows 2 or more users, who have different processing requirements, to use one computer at the same time.

The terminal users are usually connected to the central computer using communication links. The CPU time is divided out equally among the users, and each user is allowed a “*Time slice*” – a brief period when he/she is allowed to access the CPU.

The amount of time allocated to each user & the switching from one job to another is controlled by a multi-user operating system. The OS normally assigns priorities to the various jobs entering the system.

*Illustration;*

The OS may give each terminal user 5 seconds to submit a job. The user sits at the terminal & issues commands to the OS.

After every 5 seconds, the central computer checks all the terminals to see if there is any user who needs assistance. If a particular terminal does not need service, the computer goes onto the next terminal. But if a new command has been issued, the computer will allocate a time-slice to the user. During this time, the computer devotes its full attention to this user. When the time-slice is over or the user’s requests have been satisfied, the computer goes on to the next terminal. The user must now wait until he/she is allocated her next time-slice.

*Note.* The switching of control from one user to another during assigning of the time slices happens so fast that an individual user may think that he/she is the only one using the system. E.g., for 50 users each allocated 10 milliseconds; it takes only 500 milliseconds (½ a second) to service them all.

*Question.* What happens to a user’s job if her time-slice is up and the job is not completed?

The job is interrupted and allocated some space on the disk where the job together with all relevant status information is moved into. When the time comes to resume the job (or during the next allocated time-slice), the job is *rolled-in* from the disk, and processing continues at the point at which the interruption occurred.

*Characteristics of a time-sharing processing mode.*

- Each user has one or more Input/Output devices connected to the central computer by communication lines.
- Each user is independently of the others who are connected to the system.
- Each user has its own private set of programs plus access to a set of public programs.
- The central computer accepts the data & instructions arriving simultaneously from many users, and gives each user a small but frequently repeated segment of computer time.
- Data files, program files, and Input/Output devices are all directly connected to the computer, so that processing can be performed at random as requests for transactions are made.

*Application areas for time-sharing.*

1. In *Bureaus* that serve individuals or small companies who cannot afford the computer facilities.
2. In *learning institutions* where there are many users.

*Advantages of time-sharing systems.*

1. Provides better services to users since the output is fast.
2. The processor’s idle time is utilized. This is because; the processor doesn’t have to wait for the slower communicating peripherals.
3. Files are held online, hence enquiries or file interrogation is possible.
4. The user and the computer can communicate through the terminals connected to the central computer.
5. It avoids duplication of software.
6. Terminal users can benefit from the facilities of the central computer.
Data processing techniques

7. It is helpful to small companies, which cannot afford to purchase a computer & the related facilities. The company can be served at a fee through time-sharing processing mode.

Disadvantages of time-sharing systems.
1. Users have no control over the central computer.
2. The response is slow especially when there are many tasks.
3. There is no security of data.

Multi-programming systems

Multi-programming (also referred to as Multi-tasking) refers to a type processing where more than one programs residing in the computer memory are executed concurrently by a single Processor.

A multi-programming system allows the user to run 2 or more programs, all of which are in the computer’s Main memory, at the same time.

The jobs are scheduled to run automatically by the Processor under the influence of a Multi-programming or Multi-tasking operating system).

The schedule is such that; the Processor bound jobs (i.e., jobs that require much of the C.P.U time as compared to the peripheral time) are assigned low priorities for them not to tie up the C.P.U time. The Peripheral or Print bound jobs (i.e., jobs that require much of the peripheral time as compared to the C.P.U time) are allocated the C.P.U time whenever it is available.

The OS allocates each program a time-slice, and decides the order in which they will be executed. In this case, the programs take turns at short intervals of processing time. The programs to be run are loaded into the memory and the CPU begins execution of the first one. When the request is satisfied, the second program is brought into memory and its execution starts, and so on.

Note. A Multi-programming system is able to work on several programs at the same time. It works on the programs one after the other, and at any given time it executes instructions from one program only. However, the computer works so quickly that it appears to be executing the programs at the same time.

Advantages of multi-programming.
1. Increases productivity of a computer.
2. Reduces the CPU’s idle time.
3. Reduces the incidence of peripheral bound operation.

Disadvantages of multi-programming.
1. Requires more expensive CPU.
2. A Multi-tasking operating system is complex & difficult to operate.
3. Requires more expensive Input/Output facilities.

Review questions
1. (a). Define Multi-programming.
   (b). What are the factors which make multiprogramming possible?
   (c). State the benefits to be derived from Multi-programming?
   (d). Discuss the hardware and software facilities necessary to facilitate Multi-programming.

Distributed processing.
Data processing techniques

**Distributed data processing** refers to dividing of processing tasks among 2 or more computers that are located on physically separate sites, but connected by data transmission media.

*For example:*

An organization may have various computers that are located at various departments or business sites, but linked together by communication lines. In such a case, each individual department or business site is being served individually by its own computer resources. The computers at different departments are usually of limited processing power (e.g., microcomputers), and only serve as terminals from the various departments. They are then connected to a central computer of enhanced processing ability such as a Mini or a Mainframe computer.

**Salient features of distributed processing mode.**

- Each department or business site is served individually by the computer resources employed.
- The Information generated in each department is used to influence the decisions of individual departments appropriately.

**Computer configurations that can be used for Distributed processing systems.**

Configurations where computers communicate to each other are described as **Networks.**

The following are computer arrangements that can be used for Distributed processing systems:

1. **Local Area Networks (LAN).**
   - This is whereby the network is within the same locality, and does not require the use telecommunication links, e.g., same building.

2. **Wide Area Networks (WAN) or Long Haul Networks (LHN).**
   - These are networks that involve computers separated by long distance; hence, they communicate through telecommunication links.

**Note.** Networks within the same city may be linked through **Telephone lines** or **special Coaxial cables,** while far distant places may be linked through **Satellite transmission channels** or **ground Microwave systems.**

**Application areas of distributed processing.**

1. **In Banks:**
   - All the branches have **Intelligent terminals** (usually microcomputers) linked to a big computer at the Head Office. The customers’ accounts are operated on the servers in the branches, while data from the branches is sent to the main server where it is processed.

**Advantages of distributing processing.**

1. The load on the Host computer is greatly reduced.
2. The distribution of processing power increases efficiency, and also the speed of processing.
3. The use of low cost minicomputers minimizes the cost in data processing.
4. Delays in data processing are reduced.
5. Provides better service to the customers.
6. There is less risk of system breakdown.
7. The design & implementation of the system is less complex due to decentralization.
8. The level of expertise required is less.

**Disadvantages of distributing processing system.**

1. It is expensive due to the extra cost of communication equipments.
2. Data duplication is very common.
3. Programming problems occur with Micro and Minicomputers.
4. More training is needed for the users involved.

**Review questions**
Data processing techniques

1. Most companies are now shifting from the use of centralized mainframe computers to the use of geographically distributed personal computers. This method of data processing is known as Distributed data processing (DDP).

Required:
(i). Name any three computing resources that can be distributed.
(ii). Name four examples of industries and business organizations that extensively use distributed processing systems.
(iii). List and explain three ways of networking microcomputers/personal computers to form a distributed data processing system.
(iv). Name three risks that might be associated with the distributed data processing system.

Interactive processing

Interactive processing occurs if the computer & the terminal user can communicate with each other. It allows a 2-way communication between the user & the computer.

As the program executes, it keeps on prompting the user to provide input or respond to prompts displayed on the screen. In other words, the user makes the requests and the computer gives the responses.

In Interactive processing, the data is processed individually and continuously as transactions take place and output is generated instantly.

Interactive processing is mostly applied in Ticket reservation systems.

Multi-processing systems

Multiprocessing refers to the processing of more than one task at the same time on different processors of the same computer.

In a multi-processing system, a computer may have 2 or more independent processors, which work together in a coordinated manner, and are sharing the same computer memory.

This means that, at any given time, the processors could execute instructions from two or more different programs, or different parts of one program simultaneously. In such systems, each CPU is dedicated to one type of application, e.g., one CPU may handle all terminal users, while another may process only the batch jobs.

The activities of the system are coordinated by the Multi-processing operating system.

Advantage: - if one CPU fails, the other(s) can take over the workload until repairs are made.

Review questions
1. Explain the difference between Multi-Programming and Multi-Processing.

Remote job entry (RJE) systems.

This refers to batch processing where jobs are entered at a terminal remote from the computer and transmitted into the computer, e.g., by means of telecommunication links.

Conversational mode.

This is interactive computer operation where the response to the user’s message is immediate.

Factors to consider when selecting the Data processing mode.
Data processing techniques

1. Type of computer configuration to install.
2. Suitability of accumulating data (transactions) into batches for processing.
3. How the processing time would be optimized.
4. How fast the information should be produced if it is to be used for decision-making.
5. Cost of acquiring the relevant hardware, software, storage media & the cost of maintenance.
6. Ease of development, use, and subsequent maintenance.
7. How the resources (e.g., files, I/O devices) would be controlled.
8. The data communication equipment needed, the cost involved & their convenience.
9. The need for sharing resources among several users.

Review questions

1. Mention 5 factors to be considered when selecting the data processing mode suitable for use in organization.