



# LECTURE 15

# DATA PROCESSING

## Part 1

# DATA VS. INFORMATION

## Data

Data, which is the plural form of datum are actually representations of facts involving people, objects, ideas, events, etc.

These facts can be collected from careful observations, educated and experienced assumptions or recording from simple straight forward occurrences and events.

Data are represented by symbols such as letters of the alphabets, numerals or special characters.



# DATA VS. INFORMATION

## Information

When data are transformed into meaningful and useful form to serve specific purpose, then we call it information.

When knowledge collected from observation has the capability to convey its meaning and can explain fact to its users according to users' preference, then it is considered information.

Information is the result of processing of data in a way so that it gains the ability to explain its meaning to the users.

# DATA PROCESSING

- Data processing is the act of handling or manipulating data to assign the meaning into it regardless of the steps of activities during its processing. The ultimate goal of all kinds of data processing is to convert data into meaningful information.
- Data processing is a series of actions/operations through which facts and figures are collected, assigned meanings, communicated to others and retained for future use.



## TYPES OF DATA

Data collected from observations and assumptions can have its varieties of nature and characters. For example, sometimes simple numbers can describe the counts of occurrences of events but sometimes just numbers are not enough.

Quantitative observation may be simple but qualitative analysis varies from recorders' perceptions. To solve this calamity, we categorize data according to its type and level of measurement.

# TYPES OF DATA

Type of Data	Level of Measurement	Example
Categorical	Nominal (No inherent order in categories)	Eye color, ethnicity, academic major, location
	Ordinal (Categories have inherent order)	Job grade, educational qualification, age group, rating/ranking
	Binary (Can have only two choices of categories)	Gender, Marital status, Yes/No, True/False

# TYPES OF DATA

Type of Data	Level of Measurement	Example
Quantitative	Discrete (Can be described by whole number)	No. of various products in inventory, sales amount, salary
	Continuous (In theory, it can take any value in a given range, although it is recorded to a predetermined degree of precision and educated assumption)	Percentage of ingredients in a product, height, weight, age, temperature.



# DATA ORGANIZATION

Before processing, data is organized in a hierarchical approach most of the time. This approach involves putting data representing an entity into some predefined fields or attributes of data. Upon fulfilling all the fields corresponding their attributes, a complete record representing the entity (person, object, event, etc.) is achieved.

**Data**



**Field**

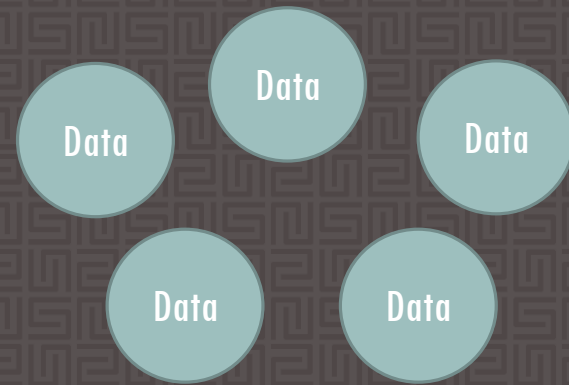
Name	Age	Smoking Status
Jason	35	Smoker

**Record**





# DATA ORGANIZATION



## Data Item

A data item is the smallest unit of information stored in a computer file. It is a single element used to represent a fact such as an employee name in an organization, item price in a retail shop or the ID number of a student in an educational institute.

# DATA ORGANIZATION



## Field

Database fields are the containers that store pieces of record or data in database tables. Data items are arranged in these fields following their attributes. For example, a record representing an entity may contain name, ID, age, contact, etc. data item where each are different in their attributes and convey different aspects of the entity. That is why all these data will be listed under their corresponding field to build the complete record.

# DATA ORGANIZATION



## Record

A record is a collection of data items belonging to their respective fields and we can consider each record to be a specific unit of information provided that the fields are well defined.

## Field

Name	Age	Smoking Status
Jason	35	Smoker

**Record**

## DATA PROCESSING ACTIVITIES

All the activities that are required to transform data into information, presentation of information and utilization of information are listed as data processing activities.

In past, these activities required pen/pencil, papers, filing cabinets, calculators, typing tools, etc. but now just one computer system can complete all these activities in fastest possible time in a lucrative manner requiring minimum manpower.



# DATA PROCESSING ACTIVITIES

Although computers have outperformed humans for certain tasks in a lot of sectors but there are still some sectors where human intervene is a must. We consider data processing a sector where all the activities require minimum human intervention.

## Collection

- Originating
- Measuring
- Recording
- Comparing

## Conversion

- Coding
- Classifying
- Verifying
- Transforming

## Manipulation

- Sorting
- Calculating
- Summarizing
- Comparing

## Storage

- Storing
- Retrieving

## Presentation

- Communication
- Reproduction

Input

Process

Output

# DATA PROCESSING ACTIVITIES

## 1. Data Collection

This is where the data processing starts and in this phase data items are collected for input. As we know, data items originate from event observation, entity evaluation and measurement, recording transaction and comparing status and features.

The method involved in data collection is set as per owner's preference and may vary from field to field.

# DATA PROCESSING ACTIVITIES

## 2. Data Conversion

Data conversion phase may start with data coding. For example, a retail shop selling Pepsi can have different packaging and size. They may code 250ml bottle as PepsiS, 500ml bottle as PepsiM, 1000ml bottle as PepsiL and 2000ml (Two liter) bottle as PepsiXL. They can be using code PepsiC for their metal can packaging.

Before transforming data into its electronic form, data classification (if required) and verification reduces the possibility of faulty data being input.

# DATA PROCESSING ACTIVITIES

## **3. Data Manipulation**

Data manipulation is the phase where the main processing is done. This may not necessarily be the third phase activity as at any stage, data manipulation can be required to be done on stored data to generate report or to use processed data in other task.

Data manipulation may involve:

- Sorting
- Calculating
- Summarizing
- Comparing



# DATA PROCESSING ACTIVITIES

## 3. Data Manipulation >>Sorting

Sorting is arranging data or record in a desired sequence. Because of logical sorting, it becomes easier to review records, find meaningful use and make decision based upon the sorted data. Sometime only logical sorting transform data into information.

For example, sales records are easier to interpret if they are sorted according to product type. Telephone database works great when it has alphabetic order.

# DATA PROCESSING ACTIVITIES

## **3. Data Manipulation >>Calculating**

Arithmetic manipulation of data is called calculating.

During inventory management or keeping shipping record, regular addition/subtraction of item quantity is performed.

More advanced calculations are performed in manipulation phase. For example, in a business organization, if payment is calculated on the basis of pay grade and hour spent, then payroll calculation can be done just by multiplying grade by hour and deducting AIT, PF and ESI.

A separate database keeping records of Tax, Insurance and Provident Fund can be auto processed and updated.

# DATA PROCESSING ACTIVITIES

## 3. Data Manipulation >>Summarizing

Summarizing condenses or reduces the volume of data and transform the database into a more usable and concise form. For example, summarizing the total sales of a shop can be done making product wise sales summary, or can have sales person wise sales record. Or the summary can have sorted product wise sales record for individual sales person so that the manager can decide to reassign his/her manpower according to skill and specialty.



# DATA PROCESSING ACTIVITIES

## **3. Data Manipulation >>Comparing**

To compare data is to perform an evaluation in relation to some known measure. For example, manager compare data to discover how well their companies are doing. They may compare current sales figures with those of last year to analyze the performance of the company in the current month.



# DATA PROCESSING ACTIVITIES

## 4. Storage

Data storage is required for storing manipulated data for future use. Storage of data is not just having a disk and writing all the files into it. Data storage must have:

- Secured storage
- Fast retrieval of data for user with authentication
- Strong backup system
- RAID storage for uninterrupted access

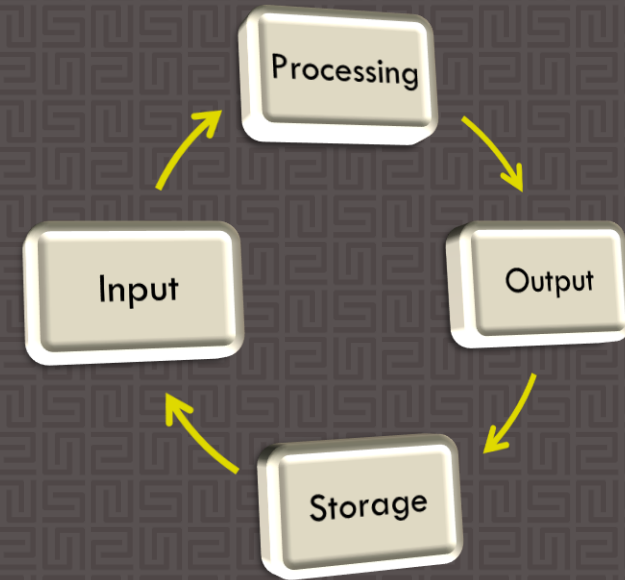
# DATA PROCESSING ACTIVITIES

## 5. Presentation

Data presentation involves communication and reproduction of data. Communication is the process of sharing information with the person or department which requests the data with proper authentication.

Reproduction of data provides us with the copy or duplicate of the data in a more intangible or presentable form.

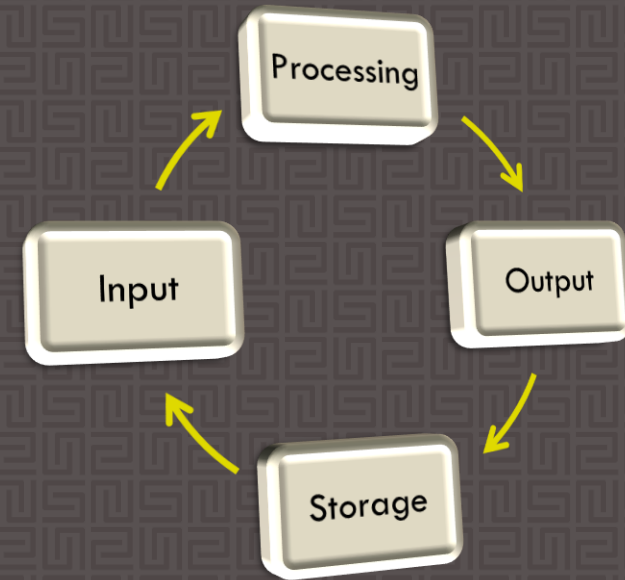
# DATA PROCESSING CYCLE



Complete data processing activities can be described revolving in a four phase cycle and we call it Data Processing Cycle. We summarize them in following four phases:

- Input
- Processing
- Output
- Storage

# DATA PROCESSING CYCLE

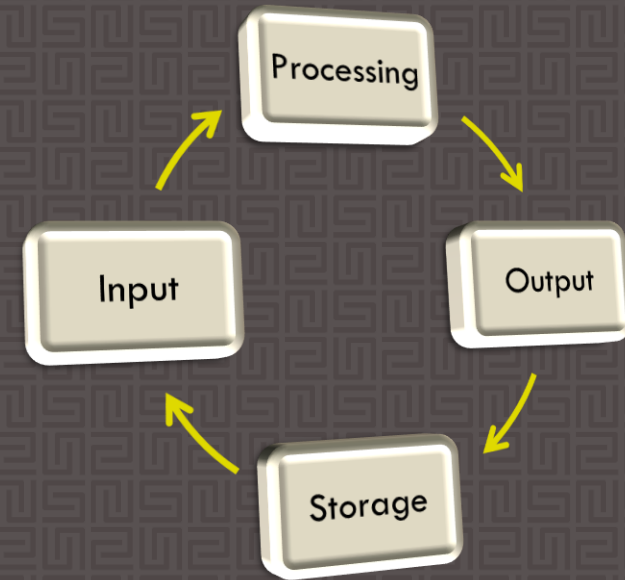


**Input:** All the activities required to record data and to make it available for manipulation and processing phase. So the whole collection, verifying and conversion activities are regarded to be the part of input phase

**Processing:** Data sorting, categorization, calculative processing, summarizing, comparing, etc. are the activities of processing phase.



# DATA PROCESSING CYCLE



**Output:** After processing, transferring the information to the requesting user falls in this phase. This phase often requires conversion of data from electronic to hard copy form and may involve some decoding and presentation.

**Storage:** Data preservation for future usage maintaining security and also facilitating authenticated retrieval are basic requirements of data storage phase.

Stay Home, Stay Safe  
Always put on a mask  
when you are in public!