Lecture 01 21/08/2025

Data and Society (DS 1107)

Lecture - Prof. BTGS Kumara (credit 01)

÷ Data

Data are raw facts and figures that on their own have no meaning

These can be, Text, numbers, symbols

÷ Data into Information.

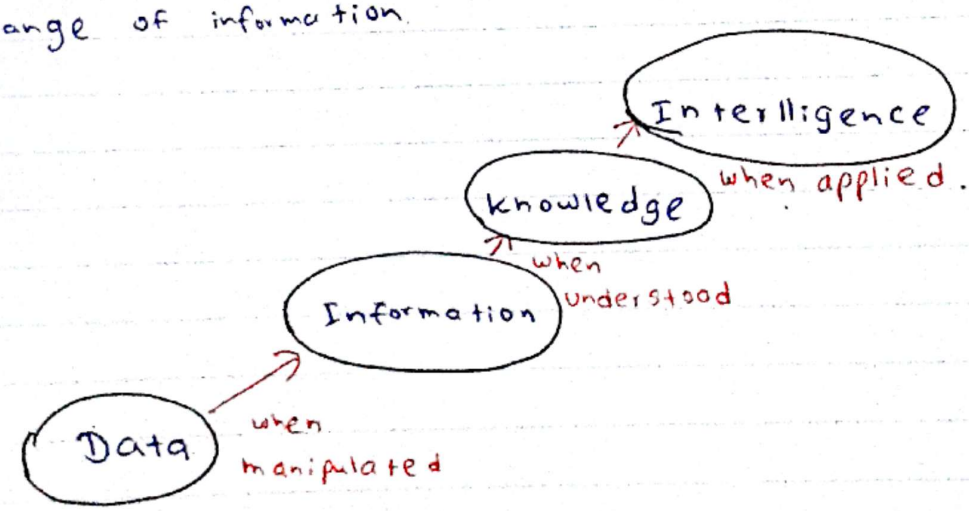
- To achive its aims the organization will need to process data into information.

information - Aggregation of data that makes decision making easier.

÷ knowledge

knowledge get by information

knowledge is derived from information inthe same way information is derived from data. it is person's range of information



Types of Data

- Qualitative Data
  - It's descriptive, expressed in terms of language rather than numerical values.
  - It can help us to understand why, how, or what happened behind certain behaviors.
  - It is subjective and unique.

- Quantitative Data
  - Quantitative data refers to any information that can be quantified.
  - It is numbers-based, countable, or measurable.
  - It tells us how many, how much, or how often in calculations.

Quantitative vs Qualitative Data

Quantitative (Example)	Qualitative
How many people attended last week's webinar?	Why do people prefer using one product over another?
How much revenue did our company make last year?	How do customers feel about their customer service experience?
How often does a customer rage click on this app?	What do people think about a new feature in the app?

Quantitative vs Qualitative Data

Quantitative	Qualitative
Deals with numbers	Deals with descriptions.
Data which can be measured	Data can be observed but not measured
Data is collected by interviewing and observing.	Data is gathered by measuring and counting
Data is analyzed by grouping it in terms of meaningful categories or themes	Data is analyzed using statistical analysis

How is data generated?

Quantitative	Qualitative
Surveys and questionnaires	Surveys and questionnaires
Analytics tools (Google Analytics gathers data such as traffic, number of page views, and average session length)	Interviews
sensors	Observations:

Real-world examples of quantitative data

- Measurements such as height, length, and weight
- Counts, such as the number of website visitors, sales, or email sign-ups
- Calculations, such as revenue
- Projections, such as predicted sales or projected revenue increase expressed as a percentage

Real-world examples of qualitative data?

- The text included in an email or social media post
- Product reviews and customer testimonials
- Observations and descriptions; e.g. “I noticed that the teacher was wearing a red saree.”
- Labels and categories used in surveys and questionnaires, e.g. selecting whether you are satisfied, dissatisfied, or indifferent to a particular product or service.

Freshman Class

Qualitative data:

- friendly
- environmentalists
- positive school spirit

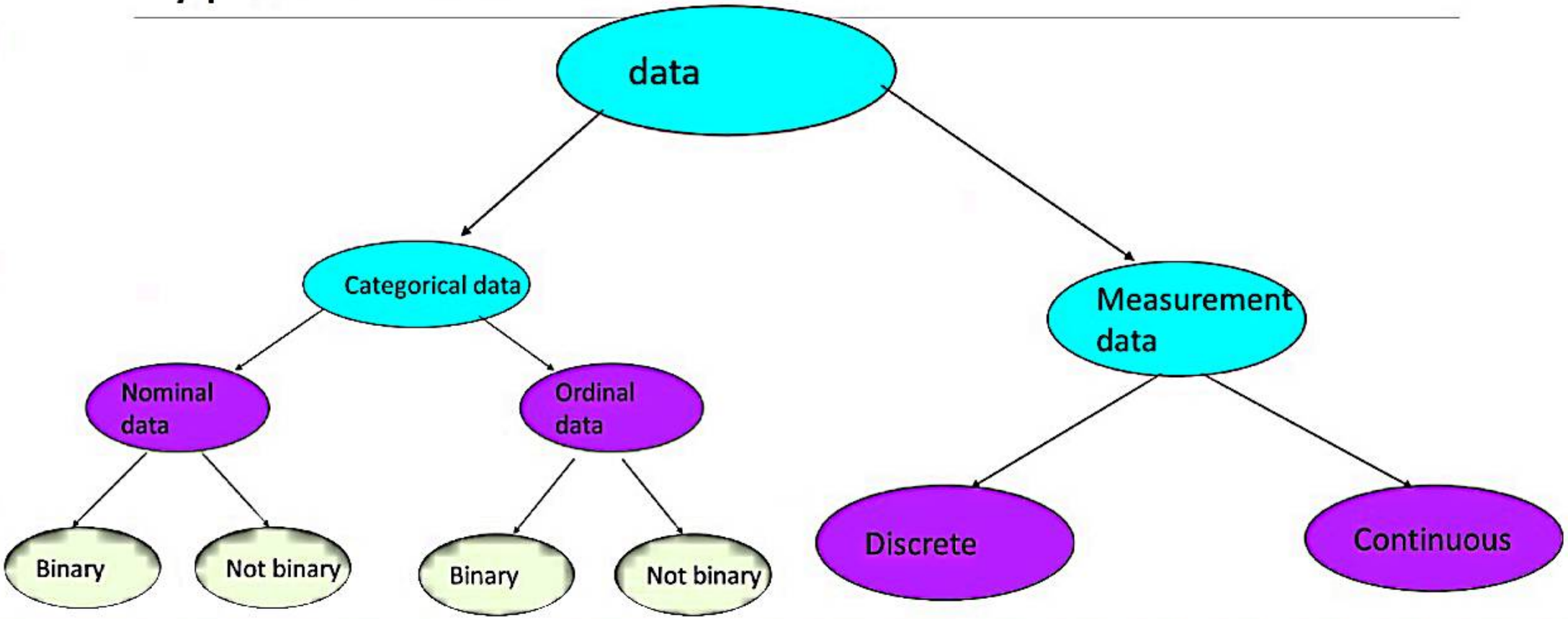
Quantitative data:

- 44 students
- 11 girls, 33 boys
- 20 students accelerated in mathematics

Qualitative or quantitative

1. The customer has clicked on the button 13 times.
2. Kirihami has curly brown hair and blue eyes.
3. The player has played 34 matches in this year.
4. The team has completed 7 upgrades this month.
5. 4 cartons of eggs were purchased by Manika.
6. My friend is funny, loud, and a good listener.
7. The customer has a very friendly face.
8. The eggs were delicious
9. There are five good people

Types of Data:



Categorical Data

- The objects being studied are grouped into categories based on some qualitative trait.

Example:

- Hair color: White, brown, red, black
- Opinion of students about online learning: not happy, neutral, happy
- Smoking status: smoker, non-smoker

Categorical Data: Nominal Data

- A type of categorical data in which objects fall into unordered categories (*without* any intrinsic order.)

Example:

- Hair color (White, brown, red, black)
- Gender (male, female)
- Nationality (Sri Lankan, American, Mexican, French)

Categorical Data: Ordinal Data

- A type of categorical data in which order is important (with some intrinsic order or numeric value).

Example:

- Agreement (strongly disagree, disagree, neutral, agree, strongly agree)
- Rating (excellent, good, fair, poor)
- Frequency (always, often, sometimes, never)

Binary Data

- A type of categorical data in which there are only two categories.
- Binary data can either be nominal or ordinal.

Example:

- Smoking status (smoker, non-smoker)
- Attendance (present, absent)
- Did you eat cake at the Damro?" Y/N

Numerical Data (Measurement Data)

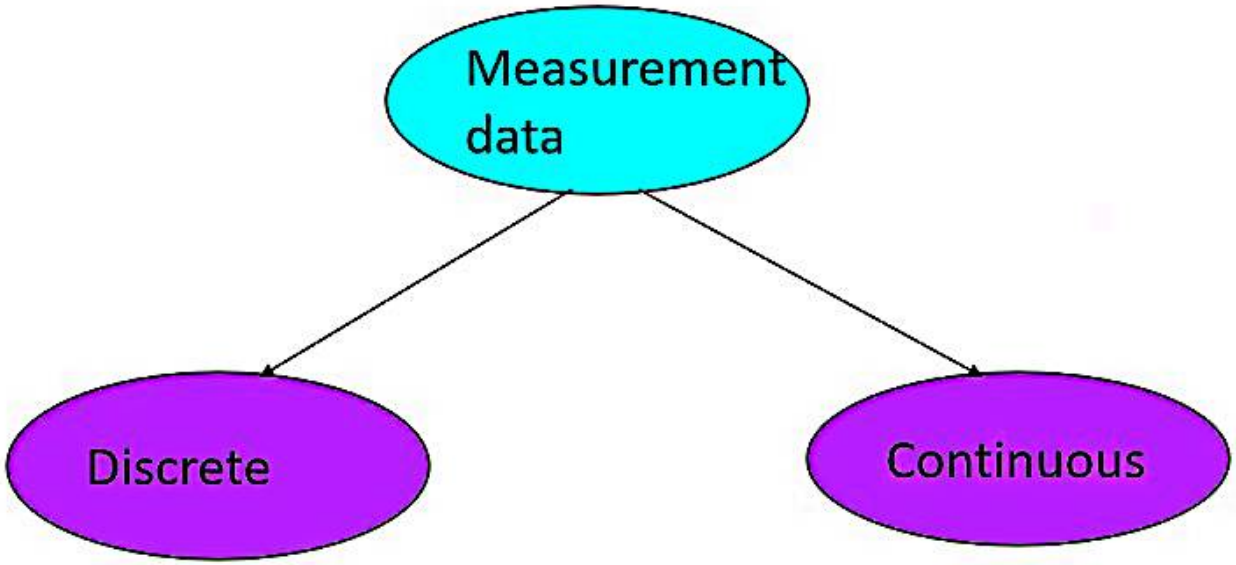
- The objects being studied are “measured” based on some quantitative trait.
- The resulting data are set of numbers.

Example:

- Height
- Age
- blood pressure
- Number of students late for class

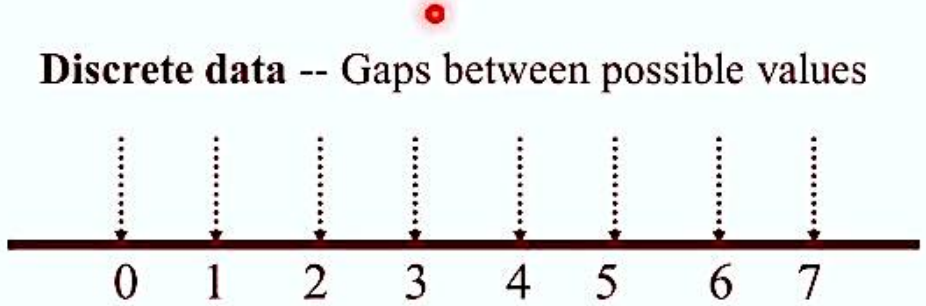
Talking:

Measurement Data



Discrete Data

- Only certain values are possible (there are gaps between the possible values).
- Generally, discrete data are counts.

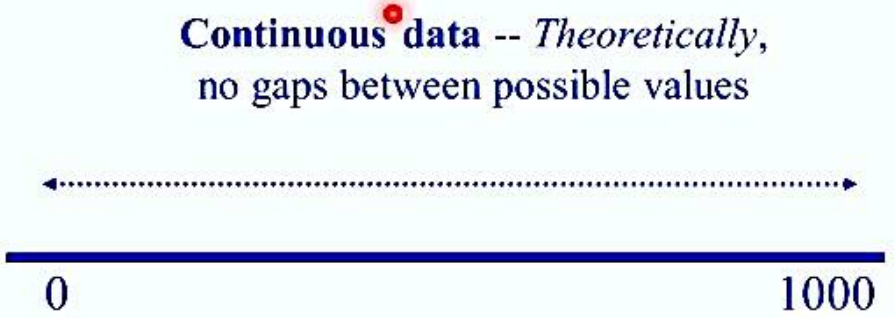


Example:

- Number of students late for class
- Number of crimes reported
- Number of times the word number is used
- Number of students late for class

Continuous Data

- Theoretically, any value within an interval is possible with a fine enough measuring device.



Example:

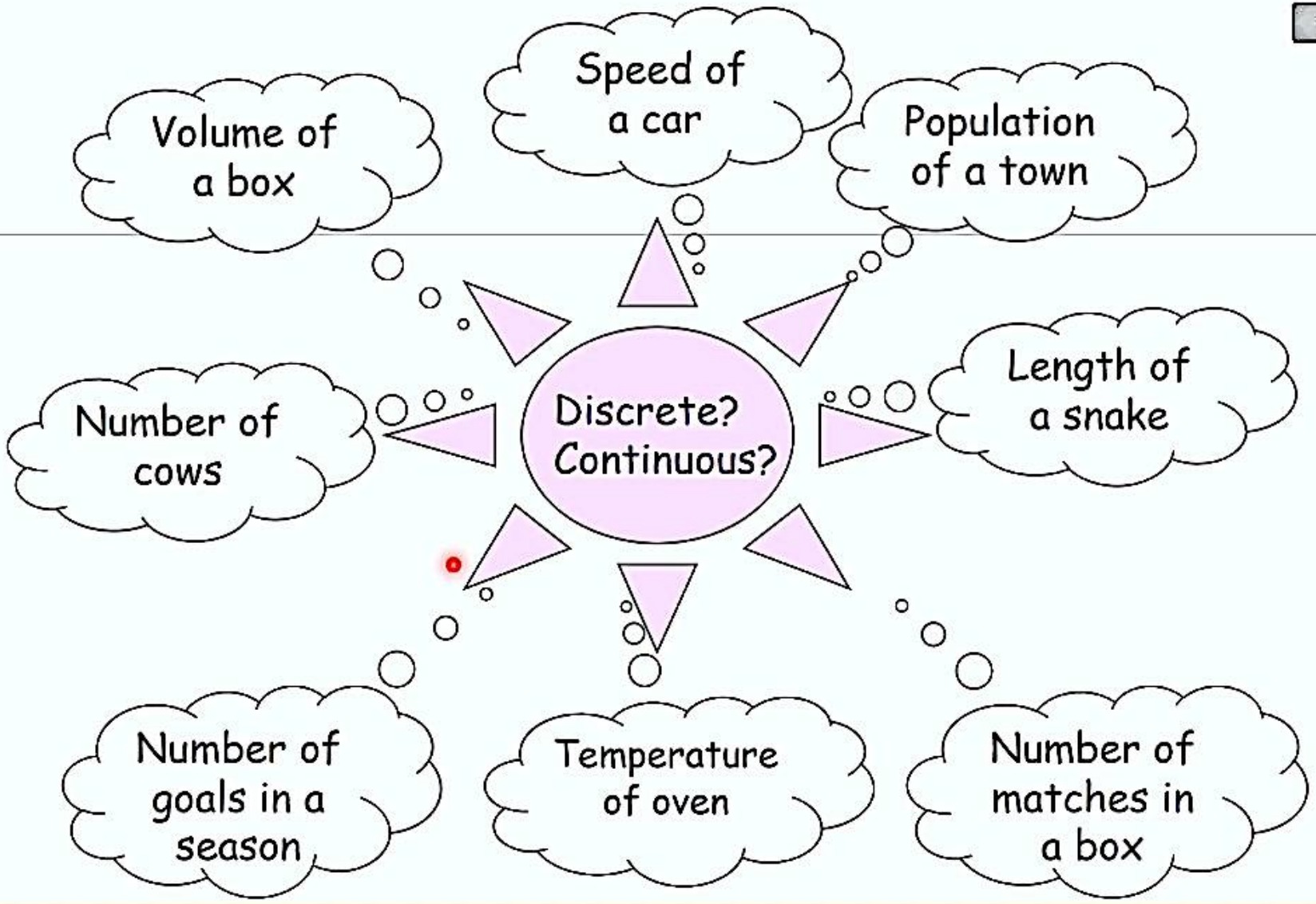
- Cholesterol level
- Height
- Age
- Time to complete a Quiz
- Concentrations of pollutants

Continuous Data

- Continuous data can be classified as interval data or ratio data.
- Interval scales can represent values below zero. Ratio variables, on the other hand, never fall below zero.

Example:

- Interval: A certain temperature can fall below 0 degrees.
- Ratio: Your height and weight cannot equal zero or be below zero.



Nature of data:

- The data can be collected in connection with time or geographical location or in connection with time and location. The following are the three types of data:

Example:

- Time series data.
- Spatial data
- Spacio-temporal data..

Nature of Data: Time Series Data

- It is a collection of a set of numerical values, collected over a period of time.
- The data might have been collected either at regular intervals of time or irregular intervals of time

Year	Food	Education	Others	Total
2001	3000	2000	3000	8000
2002	3500	3000	4000	10500
2003	4000	3500	5000	12500
2004	5000	5000	6000	16000

Nature of Data: Spatial Data

- If the data collected is connected with that of a place, then it is termed as spatial data.

Country	Population (2023)
India	1,428,627,663
China	1,425,671,352
United States	339,996,563
Sri Lanka	21,893,579

Nature of Data: Spacio Temporal Data

- If the data collected is connected to the time as well as place then it is known as spacio temporal data.

- Population:

1950		2020		2100	
China	554	China	1,439	India	1,450
India	376	India	1,380	China	1,065
U.S.	159	U.S.	331	Nigeria	733
Russia	103	Indonesia	274	U.S.	434
Japan	83	Pakistan	221	Pakistan	403
Germany	70	Brazil	213	D.R. Congo	362
Indonesia	70	Nigeria	206	Indonesia	321
Brazil	54	Bangladesh	165	Ethiopia	294
UK	51	Russia	146	Tanzania	286
Italy	47	Mexico	129	Egypt	225