SNS COLLEGE OF TECHNOLOGY

An Autonomous Institution
Coimbatore-35

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

23ADT202 – FUNDAMENTALS OF DATA SCIENCE AND ANALYTICS

II YEAR IV SEM

UNIT II – VARIABILITY FOR QUALITATIVE AND RANKED DATA

Qualitative Data

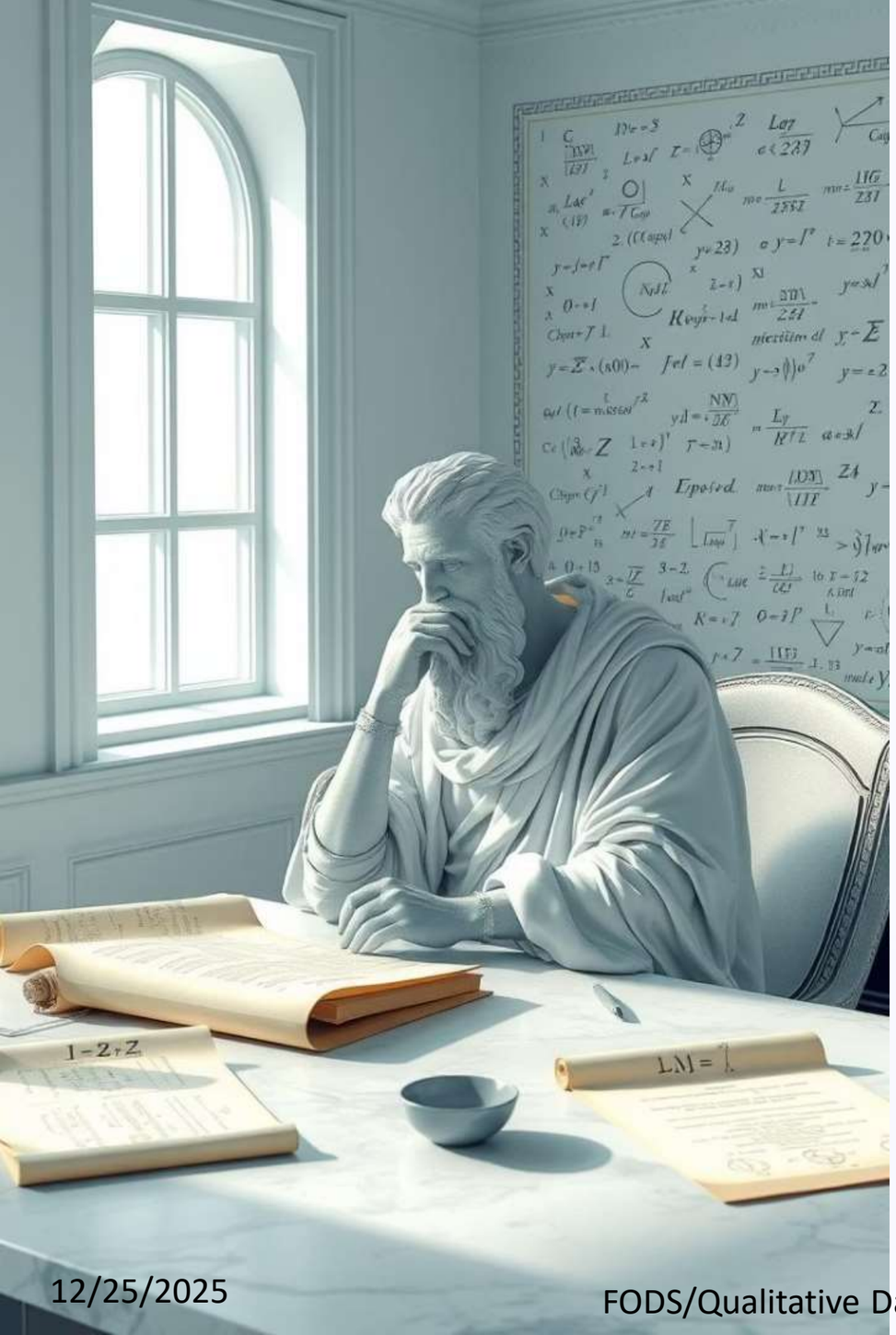
- Qualitative data is defined as data that approximates and characterizes, it can be observed and recorded.
- In the field of analysis, the terms "qualitative data" and "quantitative data" are used frequently.
- Quantitative and Qualitative are the two sides of the coin named "Data in Statistics" but as many people are familiar with quantitative data (i.e., numerical data of various sorts), qualitative data is often less understood.
- Understanding the qualitative data is essential for researchers, analysts, decision-makers, or anyone who wants to gain deep insights into people's behaviors, attitudes, and experiences.

EMPATHY:

Types of Data in Statistics

The grouping of data can be based on the quantitative and qualitative aspects of the gathered information, and data can be classified into the following types:

- Qualitative Data
- Quantitative Data



Qualitative Data Examples

There are several examples of Qualitative Data in the real world, some of these examples are:

Interview transcripts: Data collected from survey forms after the interviews can provide rich qualitative data that describes the opinions, attitudes, and experiences of participants.

Observation notes: When observing a behavior or phenomenon, recorded data of that phenomenon is also an example of qualitative data as it can tell us about the characteristics, context, and nuances of the observed phenomenon.

Open-ended survey responses: In a survey, there are some open-ended questions sometimes to know about the participant's experiences, perceptions, and opinions on a given topic. This data is also an example of qualitative data

DEFINE:

Types of Qualitative Data

Qualitative data can be further categorized into the following types:

Nominal Data

- Nominal data is represented using names, as indicated by their Latin origin.
- It includes named or labeled data and does not consider numerical values.
- For example, different movie or series genres, such as horror, sci-fi, and rom-com, are nominal categorical data. They are labeled in different forms.

Ordinal Data

- Ordinal qualitative data uses a certain scale or measure to group data into categories or groups.
- The data is generally ordered or measured, but the scale used to represent the data may not be standard or specific.
- This type of data includes numerical values and displays properties of both categorical data and numerical data.

IDEATE:

Advantages of Qualitative Data

Some advantages of Qualitative Data are as follows:

Richness and depth of data: Qualitative data provides a rich and in-depth understanding of the phenomenon being studied and can also reveal complex relationships, social norms, and cultural practices.

Flexibility: Qualitative research is flexible, which means that researchers can adapt their methods of collection, amount and type of data collected, and analysis to the specific needs of the study.

Disadvantages of Qualitative Data

Some disadvantages of Qualitative Data are as follows:

Subjectivity: As qualitative data is often collected through interviews or observations, that's why the researcher's own biases and beliefs can influence the data and that can lead to subjective interpretations of the data.
Small sample sizes: Qualitative research typically involves small sample sizes, which can limit the generalizability of the findings.
Time-consuming: Qualitative research can be time-consuming, particularly when compared to quantitative research as all the steps in this type of data are time-consuming from collection to analysis.

TESTING:

Ranked Data

- A ranked variable is one that has an ordinal value (i.e. 1st, 2nd, 3rd, etc.).
- While the exact value of the variable may not be known, its place relative to the other variables is.
- Ranked data is data that has been compared to the other pieces of data and given a "place" relative to these other pieces of data.
- For example, to rank the numbers 7.6, 2.4, 1.5, and 5.9 from least to greatest, 1.5 is first, 2.4 is second, 5.9 is third, and 7.6 is fourth.
- The numbers within this data set (7.6, 2.4, 1.5, 5.9) are ranked data, and the ordinal numbers used to rank them (1st, 2nd, 3rd, 4th) are ranked variables.

PROTOTYPING:
Types of Ranking

Standard Competition Ranking

Standard Competition Ranking is a ranking system where people or teams on a leaderboard are given positions, taking the possibility of ties occurring into account.

Oftentimes, the number "1224" appears in parentheses with the title of this type of ranking, indicating that data items that are equal in value receive the same ranking.

In other words, when two people or team tie, they tie for the same rank. For example, say the following divers have these scores:

Diver	Score
A	9.8
B	7.9
C	7.9
D	6.2

Diver: A, B, C, D; Score: 9.8, 7.9, 7.9, 6.2

The divers' rankings would be 1, 2, 2, 4, where the two divers with 7.9 are both ranked in second place.

Ordinal Ranking

- Ordinal Ranking is a ranking system where items are ranked/ordered, and are only able to be classified as higher or lower than the other items in the set.

- There are no ties in ordinal ranking, and there is no indication in the final ranking of the degree/amount of distance between items. "1234" usually accompanies the title of this type of ranking, meaning each data item gets a unique or different rank, even in the event of a tie.

Suppose four students are ranked based on their exam performance:

Student	Rank
A	1st
B	2nd
C	3rd
D	4th