



# **SNS COLLEGE OF ENGINEERING**



**Kurumbapalayam(Po), Coimbatore – 641 107**

**Accredited by NAAC-UGC with 'A' Grade**

**Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai**

## **Department of Information Technology**

**19IT601– Data Science and Analytics**

**III Year / VI Semester**

### **Unit 2 – DESCRIPTIVE ANALYTICS USING STATISTICS**

**Topic 1: Types of Data, Mean, Median, Mode**





# Types of Data



There are several flavors of data, and there are three specific types of data that we will primarily focus on.

- Numerical data
- Categorical data
- Ordinal data

## Numerical data

- It's probably the most common data type.
- Basically, it represents some quantifiable data that you can measure.
- Numerical data refers to the data that is in the form of numbers, and not in any language or descriptive form. Often referred to as quantitative data, numerical data is collected in number form.
- Some examples are
  - heights of people, page load times, stock prices, and so on.
- Things that vary, things that you can measure, things that have a wide range of possibilities.

## Types of Numerical Data

- Discrete data
  - Continuous data
- 19IT601/Types of Data / DSA/ Ashok Kumar / IT /SNSCE

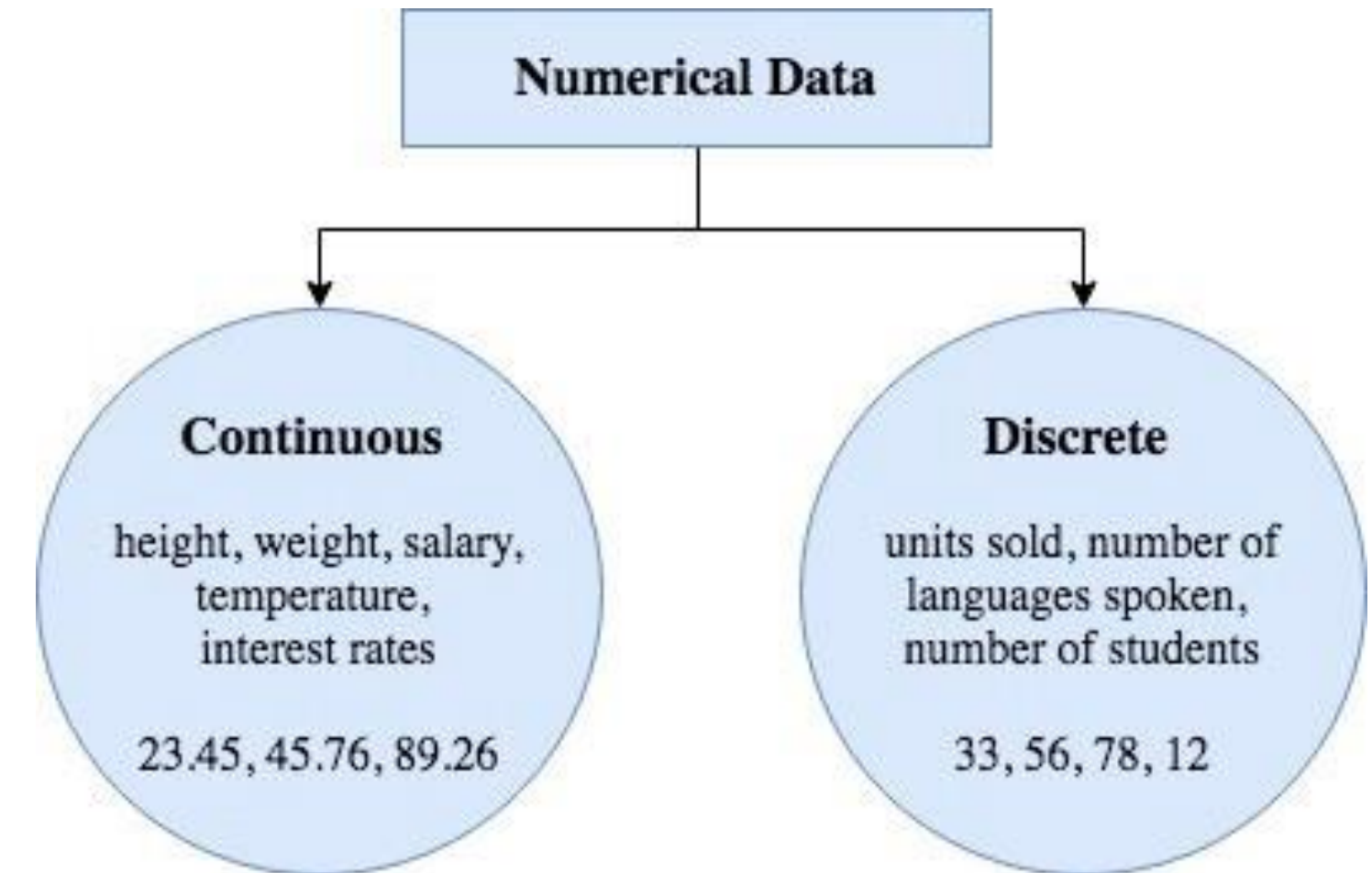
# Numerical Data

## Discrete data

- There's discrete data, which is integer-based and, for example, can be counts of some sort of event.
- Discrete data is used to represent countable items. It can take both numerical and categorical forms and groups them into a list. This list can be finite or infinite too

Some examples are

- how many purchases did a customer make in a year.
- They bought one thing, or they bought two things, or they bought three things. They couldn't have bought, 2.25 things or three and three-quarters things.
- It's a discrete value that has an integer restriction to it.





# Numerical Data

## Continuous data

- Continuous data, is that has an infinite range of possibilities where you can go into fractions.
- So, for example, going back to the height of people, there is an infinite number of possible heights for people. You could be five feet and 10.37625 inches tall, or the time it takes to do something like check out on a website could be any huge range of possibilities, 10.7625 seconds for all you know, or how much rainfall in a given day.
- Again, there's an infinite amount of precision there. So that's an example of continuous data.
- To recap, numerical data is something you can measure quantitatively with a number, and it can be either discrete, where it's integer-based like an event count, or continuous, where you can have an infinite range of precision available to that data.



# Categorical data

- Qualitative data that has no inherent numeric meaning. Categorical data is a collection of information that is divided into groups.

Examples are

- Gender,
- yes/no questions, True or False
- race,
- state of residence,
- product category,
- political party;

you can assign numbers to these categories, and often you will, but those numbers have no inherent meaning.

I might have some sort of numerical assignation to each state. I mean, I could say that Florida is state number 3 and Texas state number 4, but there's no real relationship between 3 and 4 there, right, it's just a shorthand to more compactly represent these categories.

Categorical data does not have any intrinsic numerical meaning; it's just a way that you're choosing to split up a set of data based on categories.



# Ordinal Data

- Ordinal data is one that its a mixture of numerical and categorical data.
- It is categorical data that has mathematical meaning.
- A common example is star ratings for a movie or music, or what have you.
- In this case, we have categorical data in that could be 1 through 5 stars, where 1 might represent poor and 5 might represent excellent, but they do have mathematical meaning.





**THANK YOU**