DATA TYPES

Anand Jha, M.Tech(DS&E) BITS Pilani

What is DATA?

- Data … Data and Data everywhere, but what is data? Data is a collection of facts, such as numbers, words, measurements, observations or even just descriptions of things.
- There are software in market which will help you identify the type of the data, but it is important to have at least a basic understanding of the different types of data.
- In this article, you will have an basic understanding of different types of data.

TYPES OF DATA

Quantitative

- Continuous
- Discrete

Qualitative

CATEGORICAL

- Binomial (Binary)
- Nominal
- Ordinal

Data can be Quantitative or Qualitative

Quantitative

- Numerical information (numbers)
- deals with numbers and things you can measure objectively: dimensions such as height, width, and length. Temperature and humidity. Prices. Area and volume.

Qualitative

- Descriptive information (describing/categorizing something)
- deals with characteristics and descriptors that can't be easily measured, but can be observed subjectively—such as smells, tastes, textures, attractiveness, and color.

This is just a high level classification of data, there are also different types of quantitative and qualitative data.

Quantitative data can be Continuous or Discrete

Continuous

- represent measurements
- For example, you can measure the height at progressively more precise scales: meters, centimeters, millimeters, and beyond; so height is continuous data.

Discrete

- represent items that can be counted
- For example, total number of males in a group of 10, it can be 0 to 10 (finite case), you can't have 2.5 males, or 4.7 females.

NOTE: if I measure height of males in a group of 10, then the data becomes “Continuous data”

Qualitative data can be of three types

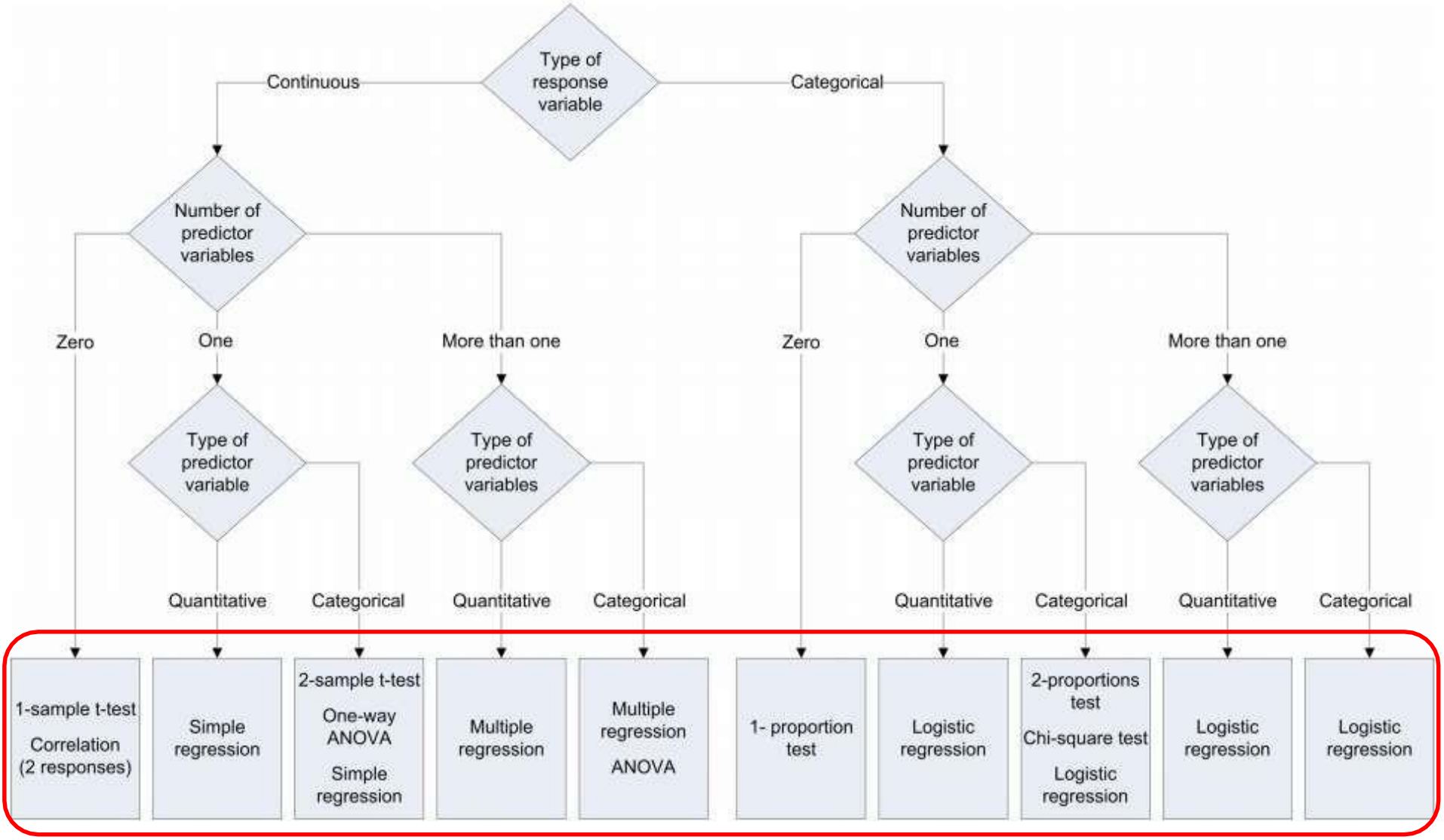
- *Categorical* data represent characteristics such as a person's gender, marital status or hometown. Categorical data can take on numerical values (such as "1" indicating male and "2" indicating female), but those numbers don't have mathematical meaning.

- ***Binary*** data is a data whose unit can take on only two mutually exclusive categories (in accordance with the binary numeral system and Boolean algebra): right/wrong, true/false, or accept/reject.

- ***Nominal*** scales are used only for labeling, without any quantitative value, such as color of each ball in a basket full of colored balls, Name of your school, type of car.

- ***Ordinal or Ordered*** data mixes numerical and categorical data. The data fall into categories, but the numbers placed on the categories have meaning. For example, rating a restaurant on a scale from 0 (lowest) to 4 (highest) stars gives ordinal data. Ordinal data are often treated as categorical, where the groups are ordered when graphs and charts are made. However, unlike categorical data, the numbers do have mathematical meaning. For example, if you survey 100 people and ask them to rate a restaurant on a scale from 0 to 4, taking the average of the 100 responses will have meaning. This would not be the case with categorical data.

Choosing Right test is always a problem?



For more information please contact:

Anand Jha – M.Tech (DS&E) , BITS PILANI

email: anand@ineuron.ai , +918618800846