About Data

Types of data

In a dataset, data can be either **Categorical** or **Numerical**.

Categorical data describes groups or categories such as car brands, gender, age groups, names, etc. On the other hand, numerical data just as the name reveals represents numbers. Within this category, you can
have **Discrete** and **Continuous** numbers.

- **Discrete -** data which can only take certain values. You only have a fixed set of values you have access to. For example, age, number of cars in a street, number of fingers.
- **Continuous -** data which can take any real or fractional value between a certain range, without any restrictions (e.g. weight, Balance in a bank account, value spent on thepurchase, Grade on Exam,Foot Size)

Levels of Measurement

Data can have two levels of measurement : Qualitative and Quantitative.

Qualitative Data is information that characterizes attributes in data but does not measure them. It can be divided into two types: Nominal or Ordinal.

- **Nominal** : They are not numbers and cannot be put in any order;

Example : names

- **Ordinal** : Consists of groups and categories that follow a strict order.

Example : Grades (e.g. Bad, Satisfy, Good)

Quantitative Data measures attributes in the data. It can be divided into two groups : Interval and Ratio

- **Interval** : Represented by numbers, without having a true zero. In this case, the zero value is meaningless.
- **Ratio**: Represented by numbers and has a true zero.

For quantitative data to be regarded as an interval or ratio, it depends on the context we are using them in. For example, think about temperature. Saying it is 0º Celsius or 0º Fahrenheit has no meaning, since that is not the true zero. The absolute zero temperature in Celsius is -273.15 ºC whereas in Fahrenheit is -459.67º F. Therefore, in this case, the temperature has to be considered as **Interval data, since the zero value is meaningless**.

However, if you analyse temperature in Kelvins, the absolute zero temperature is 0^{o} Kelvin, thus you can say now the temperature value is a Ratio since it has a true zero.