## INTRODUCTION:

In **Data Science**, we aim to do different experiments with raw data and finds some good insights from the data. To drive any business on the right path, data is very important or we can say that **“Data is the fuel”**. It can at least provide some actionable insights that can help to:

* Strategize current campaigns,
* Easily organize the launch of new products, or
* Try out different experiments.

In all the above-mentioned things, the one common driving component is Data. We are entering into the digital era where we produce a lot of Data every day.

***For Example,******on a daily basis, a company like Flipkart produces more than 2-TB of data.***

Due to so much importance of data in our life, it becomes very crucial to properly store and process this data without any error. While dealing with datasets, the data type or category of the data plays an important role to find the answer to the questions below:

* Which pre-processing strategy would work for a particular set to get the right results?
* Which type of statistical analysis should be applied for the best results?

***Following will be the agenda for studying the data and its types:***

1. Introduction to Data and its Types and their Importance
2. Qualitative vs Quantitative Data
3. Qualitative Data
   1. Nominal Data
   2. Ordinal Data
4. Quantitative Data
   1. Discrete Data
   2. Continuous Data
   3. Interval Data
   4. Ratio Data

**Introduction to Data and its Types and their Importance:**

* Data can be defined as a systematic record of a particular quantity.
* It is the different values of that quantity represented together in a set.
* It is the collection of facts and figures to be used for a specific purpose such as a survey or analysis.
* When arranged in an organized form, can be called information.

In Statistics, Data Types play a very crucial and important role, which needs to be understood, to apply statistical measurements correctly to your data so that we can correctly conclude certain assumptions about the data.

Similarly, we need to know which data analysis and its type you are working on to select the correct perception technique since different data types are considered as an approach to arrange various types of variables.

**Qualitative vs Quantitative Data:**

### ***Qualitative Data***

1. Qualitative data can’t be expressed as a number, so it can’t be measured. It mainly consists of words, pictures, and symbols, but not numbers.
2. It is also known as Categorical Data as the information can be sorted by catego1ry, not by number.
3. These can answer the questions like:

* “How this has happened”, or
* “Why this has happened”.

**Examples of qualitative data:**

* Colours e.g. the colour of the sea.
* Popular holiday destinations such as Switzerland, New Zealand, South Africa, etc.
* Ethnicity such as American Indian, Asian, etc.

**In general, there are 2 types of qualitative data:**

* Nominal data.
* Ordinal data.

### ***Quantitative Data:***

1. These types of data seem to be the easiest to explain. It tries to find the answers to questions such as
   * “how many,
   * “how much” and
   * “how often”
2. It can be expressed as a number, so it can be quantified. In simple words, it can be measured by numerical variables.
3. These are easily open for statistical manipulation and can be represented by a wide variety of statistical types of graphs and charts like line charts, bar graphs, scatter plots, etc.

**Examples of quantitative data:**

* + Scores of tests and exams e.g. 74, 67, 98, etc.
  + The weight of a person.
  + The temperature in a room.

**There are 2 general types of quantitative data:**

* + Discrete data
  + Continuous data

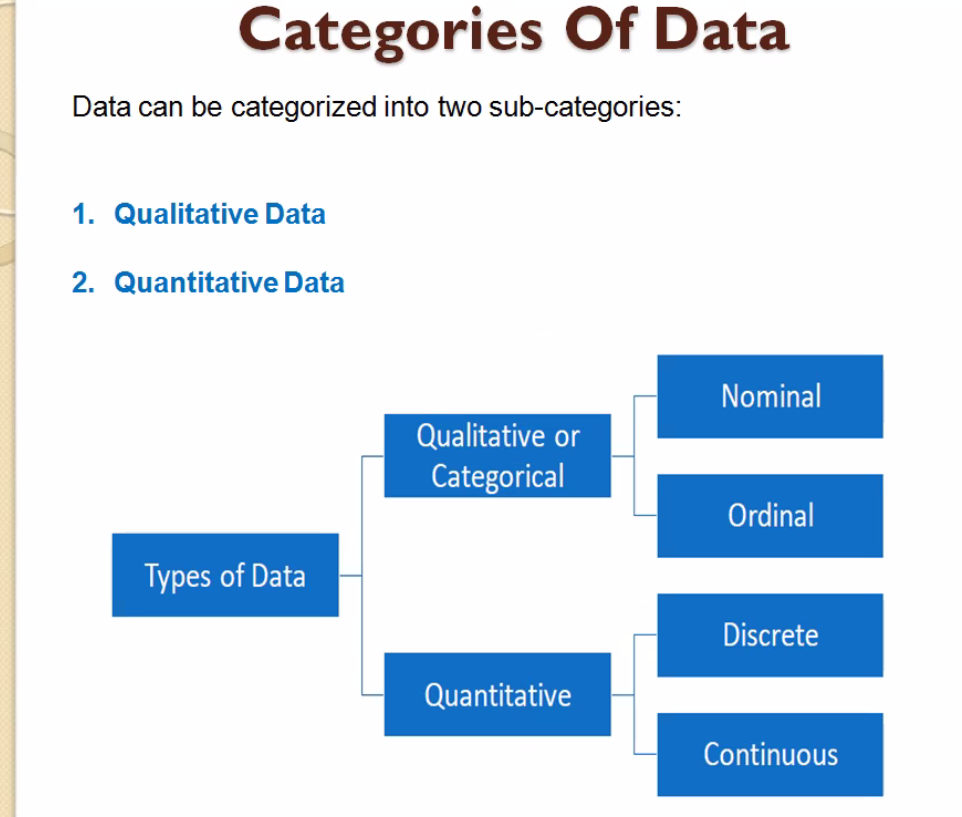


Image Source: Google

## *Qualitative Data*

## *Nominal Data:*

1. This data type is used just for labelling variables, without having any quantitative value. Here, the term ‘nominal’ comes from the Latin word “nomen” which means ‘name’.
2. It just names a thing without applying for any particular order. The nominal data sometimes referred to as “labels”.

**Examples of Nominal Data:**

* Gender (Women, Men)
* Hair colour (Blonde, Brown, Brunette, Red, etc.)
* Marital status (Married, Single, Widowed)

As you can observe from the examples there is no intrinsic ordering to the variables.

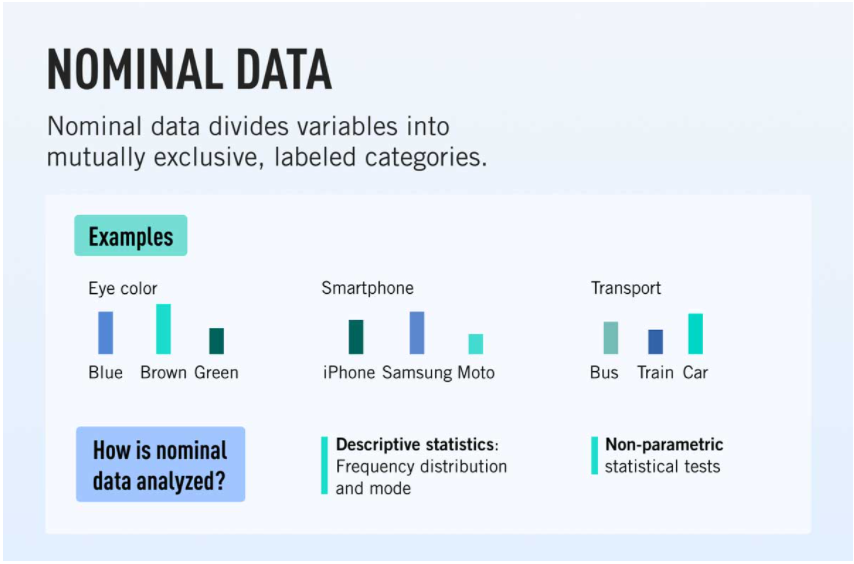
Eye colour is a nominal variable having a few levels or categories such as Blue, Green, Brown, etc and there is no possible way to order these categories in a rank-wise manner i.e, from highest to lowest or vice-versa.

Image Source: Google

## *Ordinal Data:*

1. The crucial difference from nominal types of data is that Ordinal Data shows where a number is present in a particular order.
2. This type of data is placed into some kind of order by their position on a scale. Ordinal data may indicate superiority.
3. We cannot do arithmetic operations with ordinal data because they only show the sequence.
4. Ordinal variables are considered as “in-between” qualitative and quantitative variables.
5. In simple words, we can understand the ordinal data as qualitative data for which the values are ordered.
6. In comparison with nominal data, the second one is qualitative data for which the values cannot be placed in an order.
7. Based on the relative position, we can also assign numbers to ordinal data. But we cannot do math with those numbers. For example, “first, second, third…etc.”

**Examples of Ordinal Data:**

* Ranking of users in a competition: The first, second, and third, etc.
* Rating of a product taken by the company on a scale of 1-10.
* Economic status: low, medium, and high.

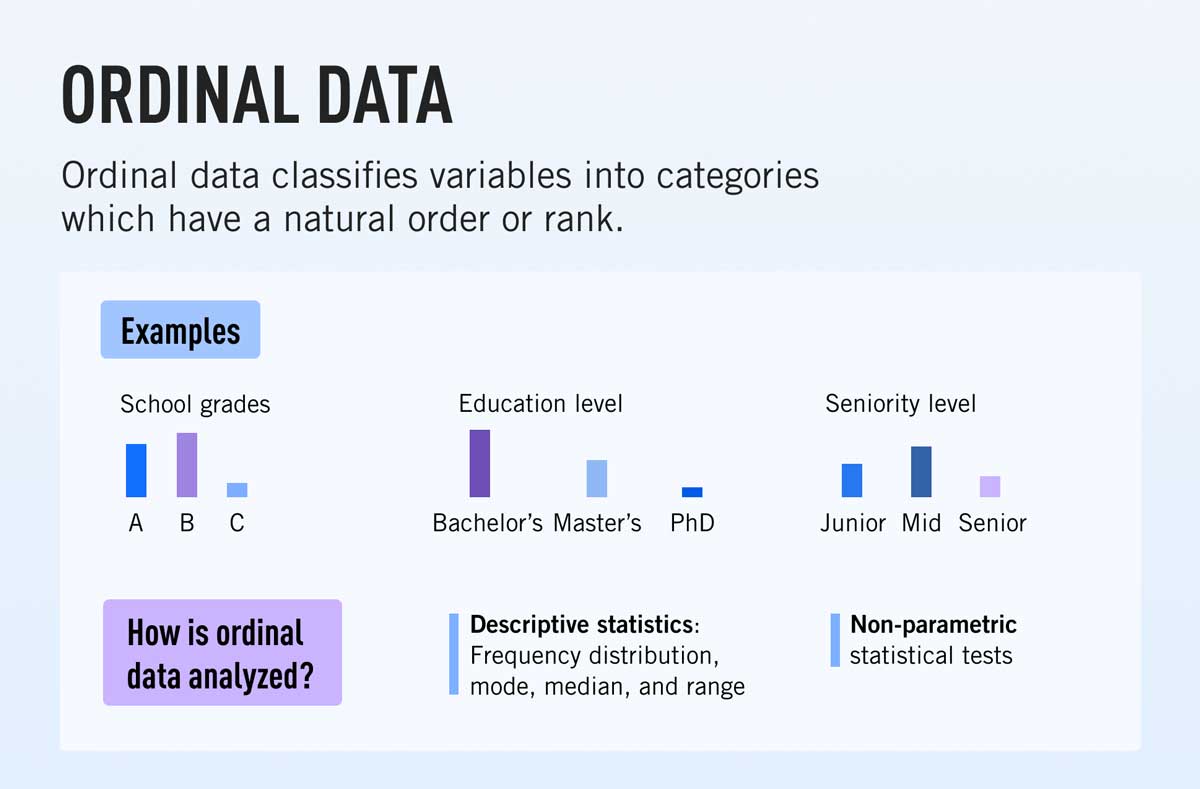


Image Source: Google

## *Quantitative Data:*

## *Discrete Data:*

1. It shows the count that involves only integers and we cannot subdivide the discrete values into parts.

**For Example,** the number of students in a class is an example of discrete data since we can count whole individuals but can’t count like 2.5, 3.75, kids.

1. In simple words, discrete data can take only certain values and the data variables cannot be divided into smaller parts.
2. It has a limited number of possible values **e.g. days of the month.**

**Examples of discrete data:**

* The number of students in a class.
* The number of workers in a company.
* The number of test questions you answered correctly.

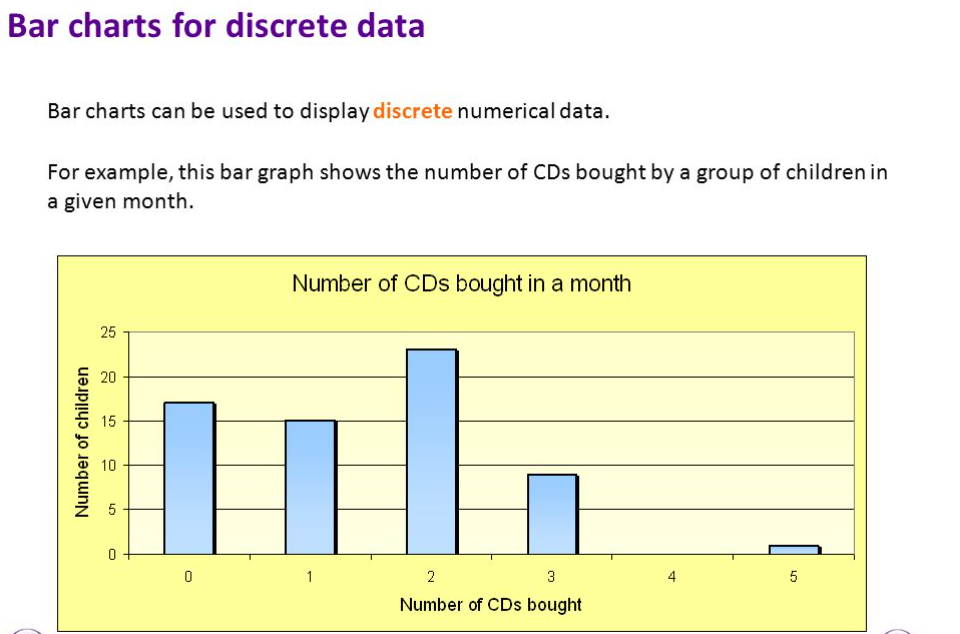
****

Image Source: Google

## *Continuous Data:*

1. It represents the information that could be meaningfully divided into its finer levels. It can be measured on a scale or continuum and can have almost any numeric value.

**For Example**, We can measure our height at very precise scales in different units such as meters, centimetres, millimetres, etc.

1. The key difference between continuous and discrete types of data is that in the former, we can record continuous data at so many different measurements such as width, temperature, time, etc.
2. The continuous variables can take any value between two numbers. For **Example,** between the range of 60 and 82 inches, there are millions of possible heights such as 62.04762 inches, 79.948376 inches, etc.
3. A good great rule for defining if data is continuous or discrete is that if the point of measurement can be reduced in half and still make sense, the data is continuous.

**Examples of continuous data:**

* The amount of time required to complete a project.
* The height of children.
* The speed of cars.

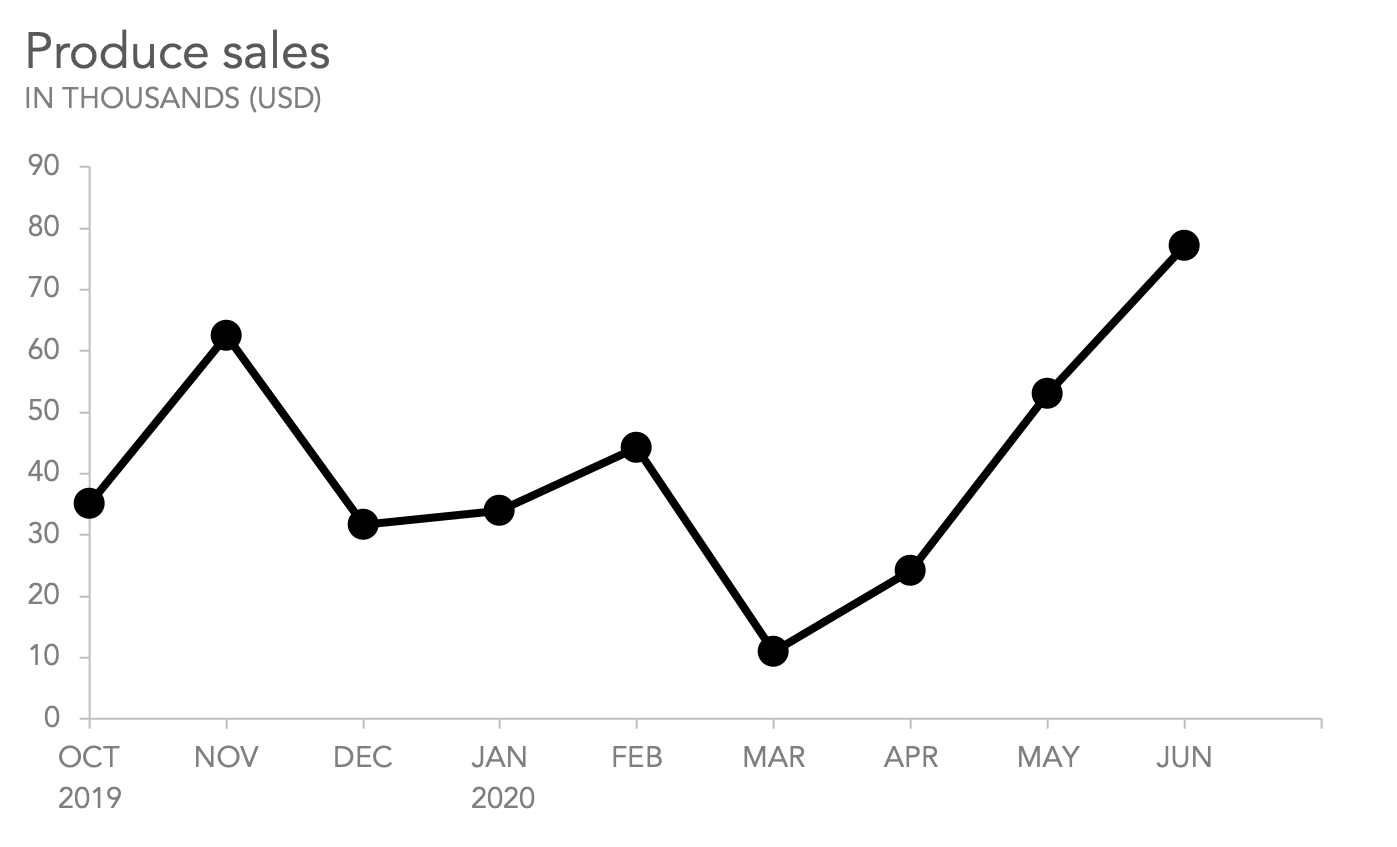


Image Source: Google

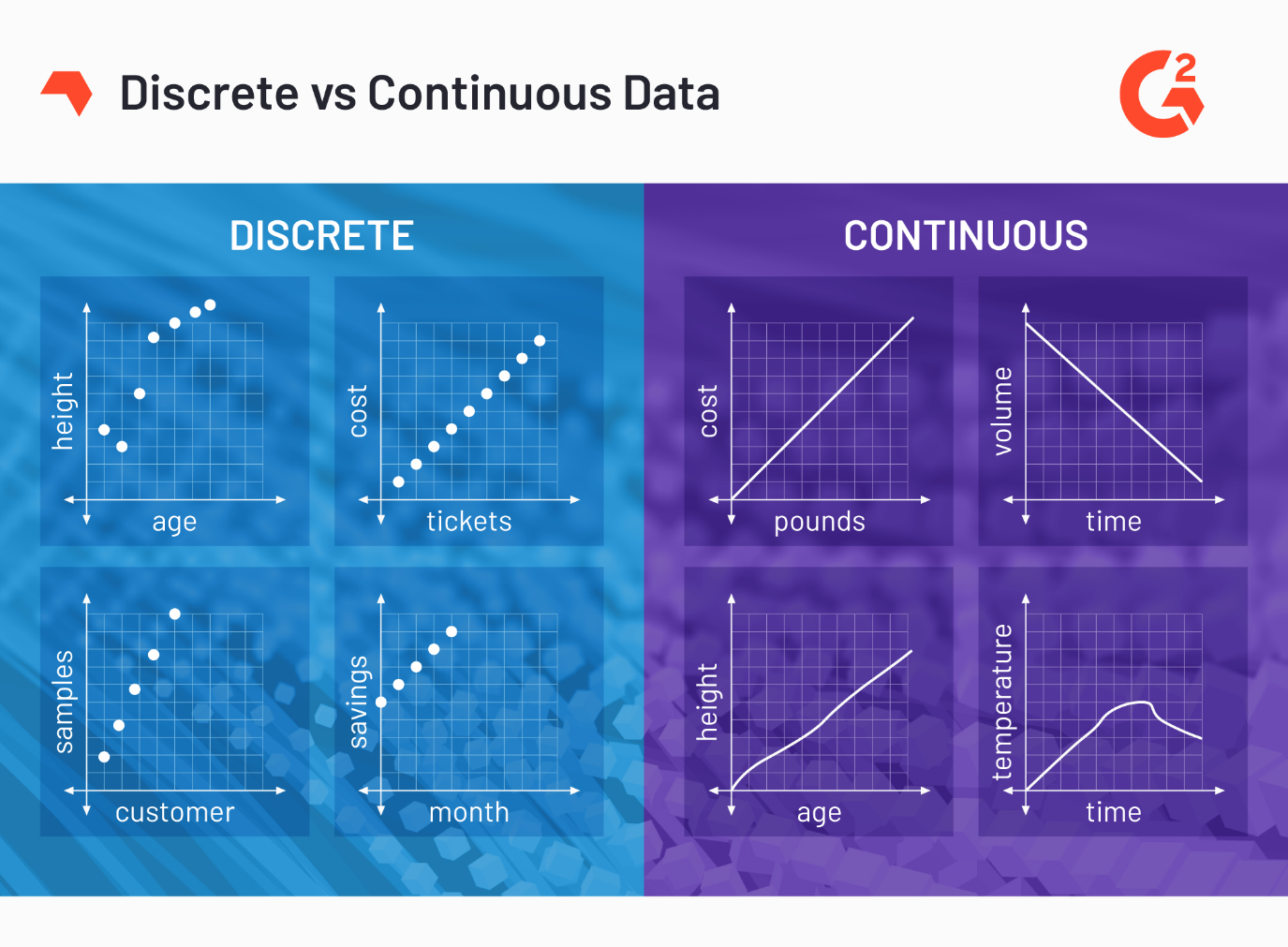


Image Source: Google

## *Interval Data:*

1. These data types are measurable and ordered with the nearest items but have no meaningful zero.
2. In the Interval scale, the term ‘Interval’ signifies space in between, which is a significant thing to recall as interval scales not only educate us about the order but in addition, give information about the value between every item.
3. Fundamentally, we can show interval data in the same way as ratio data, but the thing that we have to note is their characterized zero points.
4. Hence, with the help of interval data, we can easily correlate the degrees of the data and also add or subtract the values.
5. There are some descriptive statistics that we can calculate for interval data such as :

* Central measures of tendency (mean, median, mode)
* Range (minimum, maximum)
* Spread (percentiles, interquartile range, and standard deviation).

These are not the only statistical things to be calculated, but we can calculate more things also.

**Examples of Interval data:**

* Temperature (°C or F, but not Kelvin)
* Dates (1055, 1297, 1976, etc.)
* Time Gap on a 12-hour clock (6 am, 6 pm)

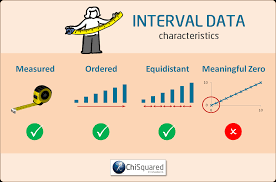


Image Source: Google

## *Ratio Data:*

1. These data are also in the ordered units that have the same difference.
2. Ratio values are the same as interval values, but the only difference is that Ratio data do have an absolute zero. For Example, height, weight, length, etc.
3. These are measured and ordered with equidistant items with a meaningful zero and never be negative like interval data.
4. Let’s understand this with an outstanding example- Measurement of heights.
5. Height can be measured in units like centimeters, inches, meters, or feet and it is not possible to have a negative value of height.
6. It enlightens us regarding the order for variables, the contrasts among them, and they have absolutely zero.
7. Ratio data is fundamentally the same as interval data, aside from zero means none.
8. The descriptive statistics which we can calculate for ratio data are the same as interval data such as :

* Central measures of tendency (mean, median, mode)
* Range (minimum, maximum)
* Spread (percentiles, interquartile range, and standard deviation).

**Example of Ratio data:**

* Age (from 0 years to 100+)
* Temperature (in Kelvin, but not °C or F)
* Time interval (measured with a stop-watch or similar)

For the above examples of ratio data, we see that there is an actual and meaningful zero-point like the age of a person, absolute zero, distance calculated from a specified point or time all have real zeros.

***Conclusion on the types of Data:***

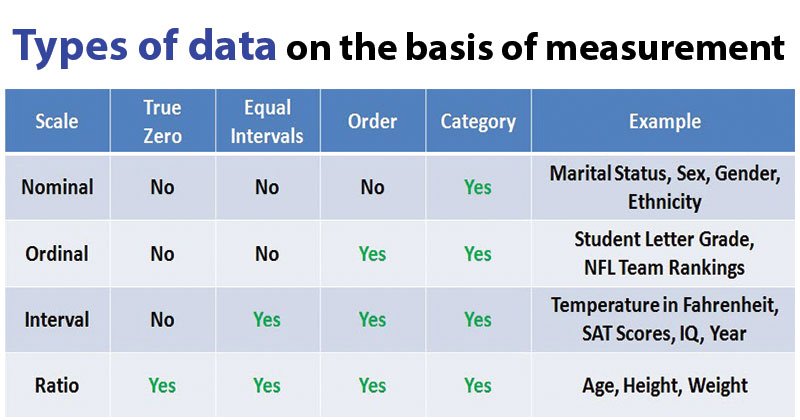


Image Source: Google

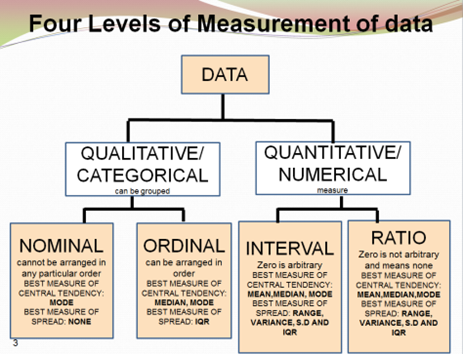


Image Source: Google