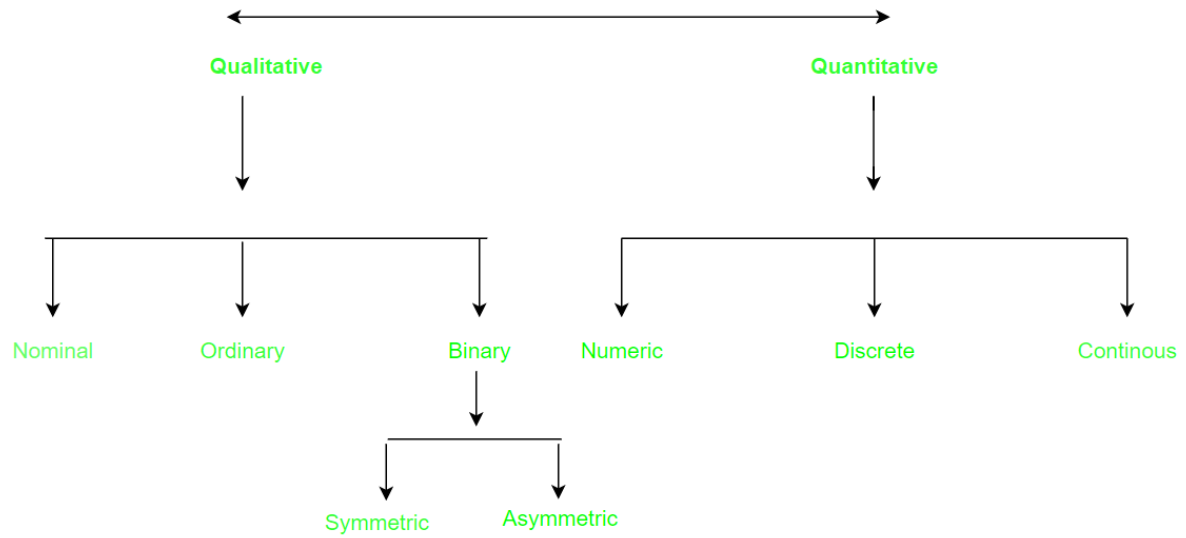
Type of attributes :

Description of attribute types.

1. Qualitative (Nominal (N), Ordinal (O), Binary(B)).
2. Quantitative (Numeric, Discrete, Continuous)



Qualitative Attributes:

1. Nominal Attributes – related to names: The values of a Nominal attribute are names of things, some kind of symbols. Values of Nominal attributes represents some category or state and that's why nominal attribute also referred as **categorical attributes** and there is no order (rank, position) among values of the nominal attribute.
Example :

Attribute	Values
Colours	Black, Brown, White
Categorical Data	Lecturer, Professor, Assistant Professor

2. Binary Attributes: Binary data has only 2 values/states. For Example yes or no, affected or unaffected, true or false.

- **Symmetric:** Both values are equally important (Gender).
- **Asymmetric:** Both values are not equally important (Result).

Attribute	Values
Gender	Male , Female

Attribute	Values
Cancer detected	Yes, No
result	Pass , Fail

3. Ordinal Attributes : The Ordinal Attributes contains values that have a meaningful sequence or ranking(order) between them, but the magnitude between values is not actually known, the order of values that shows what is important but don't indicate how important it is.

Attribute	Value
Grade	A,B,C,D,E,F
Basic pay scale	16,17,18

Quantitative Attributes:

1. Numeric: A numeric attribute is quantitative because, it is a measurable quantity, represented in integer or real values. Numerical attributes are of 2 types, **interval**, and **ratio**.

- An **interval-scaled** attribute has values, whose differences are interpretable, but the numerical attributes do not have the correct reference point, or we can call zero points. Data can be added and subtracted at an interval scale but can not be multiplied or divided. Consider an example of temperature in degrees Centigrade. If a day's temperature of one day is twice of the other day we cannot say that one day is twice as hot as another day.
- A **ratio-scaled** attribute is a numeric attribute with a fix zero-point. If a measurement is ratio-scaled, we can say of a value as being a multiple (or ratio) of another value. The values are ordered, and we can also compute the difference between values, and the mean, median, mode, Quantile-range, and Five number summary can be given.

2. Discrete : Discrete data have finite values it can be numerical and can also be in categorical form. These attributes has finite or countably infinite set of values.

Example:

Attribute	Value
Profession	Teacher, Business man, Peon
ZIP Code	301701, 110040

3. Continuous: Continuous data have an infinite no of states. Continuous data is of float type. There can be many values between 2 and 3.
Example :

Attribute	Value
Height	5.4, 6.2 ...etc
weight	50.33 .....etc

Ordinal data?

Ordinal data is a statistical type of quantitative data in which variables exist in naturally occurring ordered categories. However, it cannot be used to determine the distance between the two categories.

In statistics, a group of ordinal numbers indicates this data, and a group of this data is represented using an ordinal scale.

Uses of ordinal data

It is used for surveys and questionnaires due to its "ordered" nature.

Researchers use this type of data to collect useful information about the subject of their research. Medical researchers

used by businesses to improve overall customer service.

Employers will occasionally use a Likert scale when collecting information from **job applicants** during the application process. For example, when an applicant is applying for a position as a social media manager, a Likert scale may be used to determine how familiar an applicant is with Facebook, Twitter, LinkedIn, and so on.

Two Ordinal Attributes

Any of the methods previously described for bivariate analysis can also be used in the presence of two ordinal attributes. However:

- Spearman's rank correlation should be used instead of the Pearson correla-tion.
- Scatter plots with ordinal attributes usually have the problem that there are many values falling at the same point, making it impossible to evaluate the number of values per point. In order to avoid this problem, some software packages use a jitter effect which add a random deviation to the values, making it possible to evaluate how large the cloud is.
- Contingency tables can be used and mosaic plots too. The values should be in increasing order.