Table of contents

[STATISTIC 2](#_Toc75891088)

[Variables 2](#_Toc75891089)

[DATA SCIENCE 2](#_Toc75891090)

[MACHINE LEARNING 2](#_Toc75891091)

[PYTHON 2](#_Toc75891092)

[OTHERS 2](#_Toc75891093)

# STATISTIC

## Variables

**Numeric variables** have values that describe a measurable quantity as a number, like 'how many' or 'how much'. Therefore numeric variables are **quantitative** variables.

* A **continuous variable** is a numeric variable. Observations can take any value between a certain set of real numbers.
* A **discrete variable** is a numeric variable. Observations can take a value based on a count from a set of distinct whole values

**Categorical variables** have values that describe a 'quality' or 'characteristic' of a data unit, like 'what type' or 'which category'. Categorical variables fall into mutually exclusive (in one category or in another) and exhaustive (include all possible options) categories. Therefore, categorical variables are **qualitative** variables and tend to be represented by a non-numeric value.

* An **ordinal variable** is a categorical variable. Observations can take a value that can be logically ordered or ranked
* A **nominal variable** is a categorical variable. Observations can take a value that is not able to be organised in a logical sequence

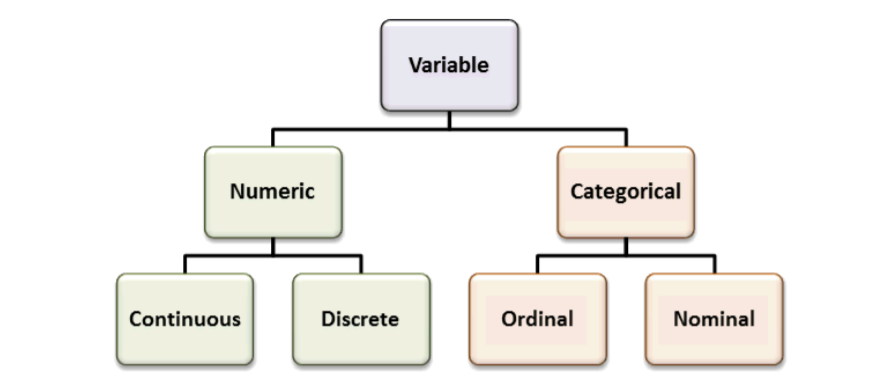
• Nominal: ID numbers, Names of people

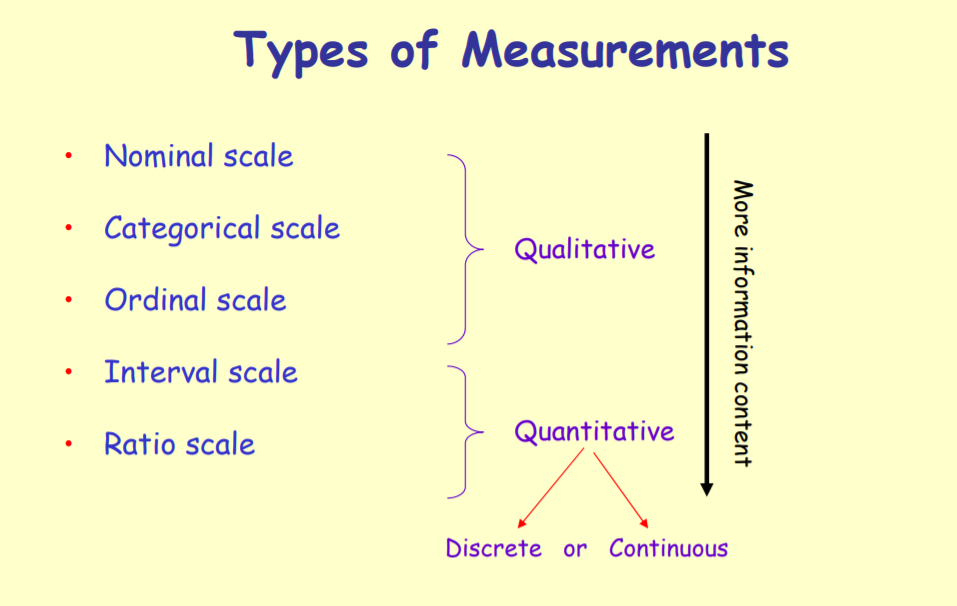
• Categorical: eye color, zip codes

• Ordinal: rankings (e.g., taste of potato chips on a scale from 1-10), grades, height in {tall, medium, short}

• Interval: calendar dates, temperatures in Celsius or Fahrenheit, GRE (Graduate Record Examination) and IQ scores

• Ratio: temperature in Kelvin, length, time, counts





## Discretization

• Divide the range of a continuous attribute into intervals

• Some methods require discrete values, e.g. most versions of Naïve Bayes, CHAID

• Reduce data size by discretization

• Prepare for further analysis

• Discretization is very useful for generating a summary of data

• Also called “binning”

# DATA SCIENCE

# MACHINE LEARNING

# PYTHON

# OTHERS