Data Types:

1. Continues

* Mean/Median/Mode Min/Max
* Std Deviation

1. Categarical

* Proportions
* Tabulations

Type of data:

1. Cross-sectional – Measurments taken at one time period

Eg: Students Course evaluaion in a course

* Cross sectional Panel:

Same student’sevaluation of different courses in particular year or subsequent years

1. Time series – Data collected over time

Eg: Un employment rate, monthly retail Sales

* Time series panel:

Eg : Student’s annual satisfaction rating of Ryerson university over 4 years

Univariate and Multivariate

Variable types:

1. Categorical(Nominal)

* Binomial
* Multinomial

Properties :

* No quantitative relationships amoung categories
* Statistics such as averages are usually meaningless

1. Ordinal data : (special type of categorical data)

* Data ordered or ranked according to some relationship to one another

Eg : Number of cars owned by a household

Properties:

* Categories can be compared with one another
* Statistics usually meaningless because of no fixed units of meassument

i.e; differences are meaningless

Eg : Number of cars

1. Ratio data (Continues): data that have a natural zero

Sales dollars, length, weight, time

Properties:

Strongest form of measurement , both ratios and differences are meaningful

1. Interval Data: data that are ordered and characterized by a specified measure of

Specified measure of distance between observations, but with no natural zero

* Temperature scales, time, survey scales that are assumed to be interval

Properties:

* Ratios are meaningless(50 degrees in not twice as hot as 25 degress)
* Differences are meaningul , so statistics such as averages may be compared

Meassure of Central Dependancy:

Population:

µ = i / N

Sample:

µ = i / n

Mean

Median

Mode

Measure of Dispersion (variability, spread, deviation)

Dispersion

Range

Variance:

1. Population : σ2 = i - µ)2 / N
2. Sample : s2 = i - )2 / n – 1

Std Deviation (σ) = sqrt of variance

* Look for low stddev for good data

**Visualization :**

Chart types should be based on type of variables being depicted

* Categorica Variable:

Eg: Counts

* Continues variables:
* Scatter plots
* Statistiscal properties(Averages)

Reference: https://extremepresentation.com/