**Intro to Data Science**

[**https://youtu.be/N6BghzuFLIg**](https://youtu.be/N6BghzuFLIg)

**Types of Data**

1. **Numerical**

Quantitative data.

1. Discrete (limited numbers).
2. Continuous (functions).
3. **Categorical**

Qualitative data.

1. **Ordinal**

Mixing numerical and categorical (subjective data).

**Types of Averages**

1. **Mean**

(Sum / number of values) Is good to use when the values are similar.

1. **Median**

(The middle value) or (if the data are even, it’s the mean of the middle two values) Is good to use when the data have some outliers.

1. **Mode**

(Most common value) Can be used in a categorical data and discrete numerical data. Is good to use when the data reoccurs.

**Types of Spread**

1. **Range + Domain**

Range: (maximum - minimum value).

Domain: The value that your data points can take on.

1. **Variance + Standard Deviation**

Variance: How much the values differ from the mean.

Standard Deviation: Square root of the variance.

1. **Covariance + Correlation**

Covariance: How much one value varies when the other varies.

Correlation: Covariance / the standard deviation of each variable.

**Quantiles + Percentiles**

1. **Quantiles**

Splitting the data into a certain number of regions that each have the same probability or equal number of points.

1. **Percentiles**

Splitting the data into 100 equal segments.

**Types of Graphs**

1. **One Variable Graphs**
   1. **Histogram**

How often each value occurs.

* 1. **Bar Plot**

Compare different groups.

* 1. **Pie Chart**

Most and least prominent.

1. **Two Variable Graphs**
   1. **Scatter plot**

Show the spread of the data between two variables.

* 1. **line plot**

the points must be connected (Numerical Continuous data).

* 1. **2-D Histogram**

Allows us to see where both variables clump, and how one falls of relative to the other at specific point.

* 1. **Box and Whisker Plot**

Shows us the spread of data, with quartiles and median.

1. **Three and Higher Variable Graphs**
   1. **Heat Map**

Let us plot two variables and also show amount/height/intensity.

* 1. **Multi-Variable Bar Plot**

Plotting several variables for one group.