Data Science

**Type of Variables**

* Continuous Variables: are continuous numerical values
* Categorical variables: are variables with categories like sex, grades.

We have two types of categorical variables

* Ordinal: all variables aren’t equal. There’s an order in the variables. Example, Grades: Good, Bad, Excellent
* Nominal: all variables are equal. No order or rank. Male, female.

**Outliers**: any value which falls outside the normal range

Causes of Outliers:

* Typos during data collection e.g. adding a zero mistakenly
* Measurement Error due to faulty device or operator
* Intentional Error caused by giving false information
* Legit Outliers are real values but very different from others. E.g. CEO salary compared to regular employees

