# Flow to visualize data and test created models

1. Get the data and create a dataframe

A dataframe in Python offers numerous advantages for data manipulation, analysis, and visualization. After the data is loaded, a dataframe summary can be retrieved using df.info().

1. Identify columns containing missing data.

When possible, values should be generated.

* Missing values in attributes that have continuous data are best replaced using Mean value.
* Missing values in attributes that have categorical data are best replaced using the most frequent value.

1. Normalization of data. This can be achieved by division of all column values by the maximum value of that column. The result will be a value between 0 and 1.
2. Visualization
   1. SLR (regression)
   2. MLR (histplot distribution)
   3. Residuals
   4. Polynomial regression
3. Evaluation
   1. Corr() for correlation
   2. R^2
   3. Mean Squared Error (MSE)