**Checklist of Activities performed on the dataset till now**

1. Load the dataset and required libraries
2. Understand the data, what are the features? What are the datatypes?
3. Ask the right question, which we can answer from its analysis

* Analyze the dataset to check for summary and structure
* Analyze what class of datatypes apply to our dataset

1. Data cleaning - dealing with missing data

* Removing Nulls or
* Imputation with mean value or 0 depending on data

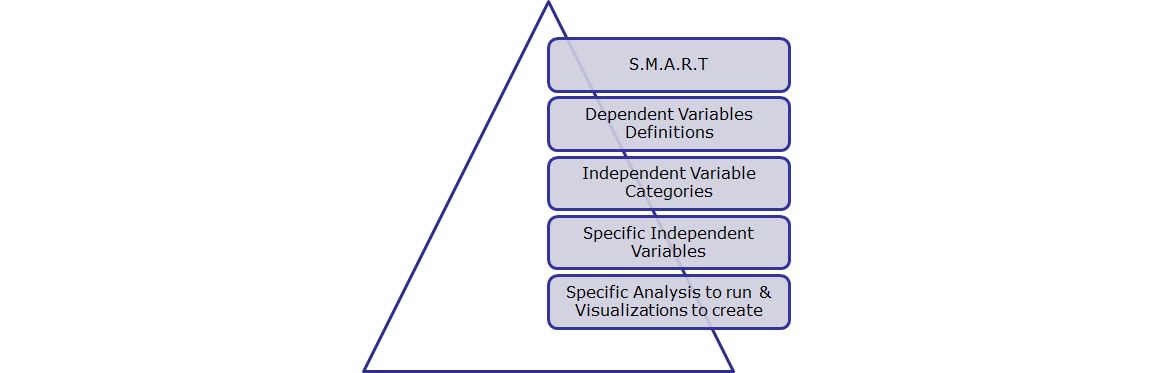
1. Data exploration and visualization

* Boxplot
* Bivariate boxplot – Bvplot
* Histograms
* Scatterplot
* 3D Scatterplot
* Chullplot
* Chiplot
* Scatterplot with linear fit
* Pairs for checking variance

1. Data variance - univariate and multivariate

* T test
* Hotelling T squared Test
* F test
* Leven’s test

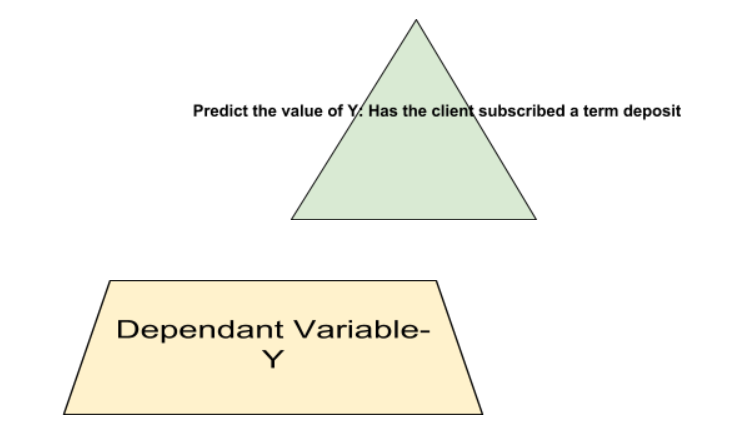
**Structural Pyramid Analysis Plan**



Structured Analytical plan or Structured design plan are architectural methods  that are formulated in order to analyze the data provided in accordance to the actions to be taken on it, and the information that is to be extracted from that data in a precise and efficient manner.

Here, we present some SMART ( SPECIFIC. MEASURABLE . ATTAINABLE. RELEVANT. TIMELY)  goals to display the most relevant entities of our dataset.

DEPENDENT VARIABLE:

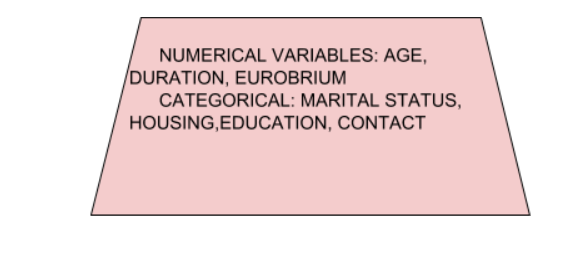


Here, the dependent variable is “y” that has two levels mainly ie. yes , no.

Yes = states the no of people that have subscribed to the term deposit

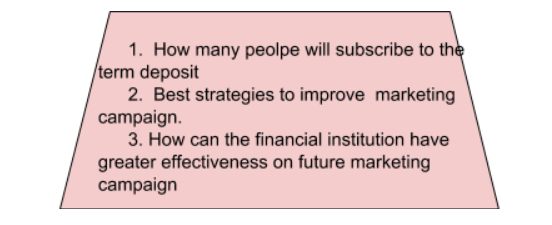
No = states the no of people that declined to the term deposit.

INDEPENDENT VARIABLE:



Next Step is to analyze the independent variables categories to be considered to predict the final value of Y. Below are a few variables we will be considering for our analysis.

QUESTIONS RAISED:



The next stage involves raising specific questions about possible interactions between the types of data.

We have raised the above three questions.