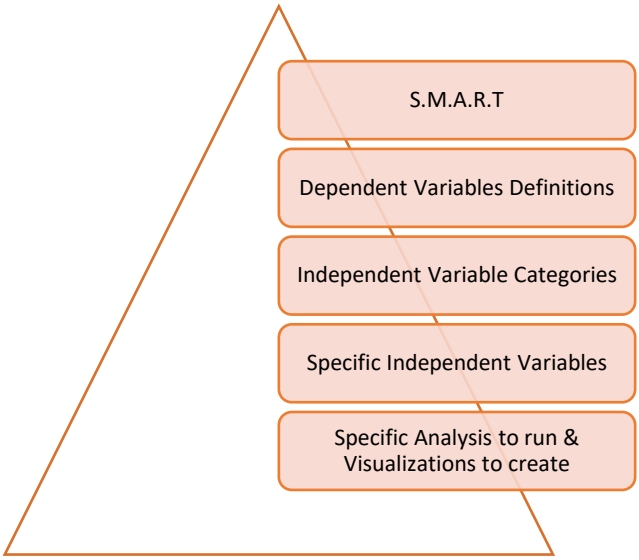
# World Happiness Analysis: Structured Pyramid Analysis Plan




## S.M.A.R.T:

### S: SPECIFIC

•Our objective is to perform Multivariate analysis on World Happiness for which we have obtained dataset from Kaggle.

• We aim to understand the different factors that contribute to the Happiness of a nation, and how these factors interact with one another.

• We will explore how the Happiness score varies across countries in different parts of the world.

### M: MEASURABLE

• Perform operations on the dataset depending upon the nature of the dataset and the questions asked.

• The performance can be measured based on the completion of the assigned task and also on the results from the techniques performed.

### A: ACHIEVABLE

Operations performed on the World Happiness Score dataset:

1. Understood the nature of dataset.

2. Formed analysis questions which can be performed on the dataset.

3. Cleaned the data i.e missing values and unwanted columns.

4.Data Visualization for finding relationship between variables, detecting outliers, correlation etc .

**R: RELEVANT**

- Machine learning can help in identifying the major factors which increases the happiness score and knowing this information can empower the government of different countries to achieve higher happiness score for its residents.
- Using different analytic techniques and algorithms can provide better information to governments in understanding how economy development can impact the happiness score.

**T: TIME-BOUND**

- The long-term goal is to understand how the countries can maximize the happiness score.
- Research and perform the best algorithms which will help us in accurate prediction.

## Dependent Variable Definitions:

- Country/Rank
- Happiness Score

These variables depend on the variation of the other independent variables. Each independent variable has a different impact on the output variable.

## Independent Variable Categories:

The following are the independent variables:

- Economy (GDP per Capita): The extent to which GDP contributes to the calculation of the Happiness Score.
- Family: The extent to which Family contributes to the calculation of the Happiness Score
- Health (Life Expectancy): The extent to which Life expectancy contributed to the calculation of the Happiness Score
- Freedom: The extent to which Freedom contributed to the calculation of the Happiness Score.

• Trust (Government Corruption): The extent to which Perception of Corruption contributes to Happiness Score.

• Generosity: The extent to which Generosity contributed to the calculation of the Happiness Score.

• Dystopia Residual: The extent to which Dystopia Residual contributed to the calculation of the Happiness Score.

## Specific Analysis to run & Visualizations to create

- Questions based on our data set –

1. How are the major various factors related?

a. Economy Gdp per capita

b. Family

c. Life Expectancy

2. Are there any significant differences between the means of Government related factors?
a. Freedom

b. Trust/Government Corruption

3. Are there any significant differences between the standard deviations of Government related factors?

a. Freedom

b. Trust/Government Corruption

4. Can the factors leading to happiness score describe a function?

a. Economy

b. Family

c. Life Expectancy

d. Freedom

e. Generosity

f. Trust

g. Dystopia residual

5. What is the effect of happiness score based on continent?

a. Country categorized to continents [new independent variable]

- Data Visualisations:

1.Correlation matrix

2. Correlation Plot

3. Box Plot

4.Scatter Plot/ 3D Scatter Plot