**Structures Pyramid Analysis Plan**

**S.M.A.R.T. Goal:**To determine the factors that influence the algae blooms.

**Specific Analysis and Visualization to be created:**

For the river size: Box and Whisker’s plot

For Season: Box and Whisker’s plot

For Speed: Box and Whisker’s Plot

Why Box and Whisker’s plot:

All the above variables are categorical/ordinal in nature and the dependent/Target Variable is numerical in nature, hence a Box and Whisker’s plot would be an ideal way of Visualization.

Chemical River Characteristics: Scatter Plot.

Why Scatter plot:

We wanted to see the range in which the maximum algae are blooming. Also we had to use two variables for doing it hence Scatter plot was the best choice.

**Specific Independent Variables:**

River Size: Large, Medium, Small

Speed: High, Medium, Low

All the Chemical River Characteristic Independent Variables are Numeric in nature.

Season: Autumn, Spring, Summer, Winter

**Independent Variable Categories:** We have divided the data set in 3 categories for the independent variables.

The categories are as follows:

1. Physical River Characteristics:

River Size, River Speed

1. Chemical River Characteristics:

mxPH, mnO2, Cl, NO3, NH4, oPO4, PO4, Chla

1. Season:

**Dependent Variable Definition:**All the 7 different types of Algae’s i.e. a1, a2, a3, a4, a5, a6,a7