Assessment #04 - Python SciPy and Pandas

Dead Line for Upload of the Assignment Records to VTOP is 29-Mar-2020

Dear students,

Please download the following three data files

- DOData.csv (https://bsrvp.github.io/data/DOData.csv)
- NutAverage.xlsx (https://bsrvp.github.io/data/NutAverage.xlsx)
- PhytoBiomass.xlsx (https://bsrvp.github.io/data/PythoBiomass.xlsx)

The first data file **DOData.csv** contains the monthly mean values of Dissolved Oxygen in water for a year.

The second data file **NutAverage.xlsx** lists the monthly mean values of NH4-N (Ammonia), NO2-N (Nitrite), NO3-N (Nitrate) and TN (Total Nitrogen) content in the water for a year.

The final data file **PythoBiomass.xlsx** lists the monthly mean biomass of the two Phytoplanktons namely Cyanophycean and Chlorophycean along with the Total Biomass .

You need to perform the following tasks on these three data sets.

Q1) Read the data file **NutAverage.xlsx** into a DataFrame and perform the following tasks:

- Add a column **DIN** (stands for Dissolved Inorganic Nitrogen) to this DataFrame, where **DIN = NH4-N+NO2-N+NO3-N**.
- Add another column **DON** (stands for Dissolved Organic Nitrogen) to this DataFrame, where **DON = TN – DIN**.
- Describe characteristics of the DataFrame.
- Plot the all the data (except the Day Count column) using area plot and box plot of DataFrame.
- Compare the DIN vs DON composition graphically.

Q2) Read the data file **PhytoBiomass.xlsx** into a DataFrame and perform the following tasks:

- Add a column **Others** which list the biomass of other phytoplankton groups obtained by substracting Total Biomass with sum of the biomass of Cynophyceans and Chlorophyceans.
- Describe the characteristics of the DataFrame.
- Plot the biomass composition of each group using bar plot.

Q3) Read the data file DOData.csv into a DataFrame and perform the following tasks:

- Plot the monthly data using bar plot.
- Construct a Spline interpolation polynomial and use this to estimate the DO for the complete year starting from day 1 to day 364 and visualize the both the data sets.
- Perform a comparative analysis of DO with that of TN by using regression analysis or curve fit routine of SciPy. Plot DO vs TN as well the Regression line.