

# Maths-Assignment1

## Task 1 - Hands On Assignment

- Data Analysis in python
- Web Scraping using Beautiful Soup
- Random Variate Generation

Run the notebooks discussed in class and upload it in the respective folders in the HandsOn link below.

## Task 2 - Case Study Assignment

Perform data analysis on the suicides dataset and answer the following question:

1. Data Preprocessing
  - a. Name the categorical and numerical columns in the dataset.
  - b. Find the number of unique values in the following columns - country, year, sex, age, generation
  - c. Plot a heatmap of the missing values in the dataset.
  - d. Find the number of outliers in each numeric column and remove them.
  - e. Find the attributes with the highest correlation coefficient.
2. Data Visualization
  - a. Plot a pie chart of the 'suicides/100k pop' per generation.
  - b. Plot a line plot of population vs suicides\_no.
  - c. Plot a barplot of the 'gdp\_per\_capita (\$)' per year.
  - d. Plot a distplot for the 'suicides/100k pop'

Link to dataset:

[https://drive.google.com/file/d/1YAmDi9LxISaNVLO5cM2U8bDiAbq1uPwO/view?usp=drive\\_link](https://drive.google.com/file/d/1YAmDi9LxISaNVLO5cM2U8bDiAbq1uPwO/view?usp=drive_link)

## Submission Guidelines

### Task 1

Upload your notebook to the below drive with the following naming format:

SRN\_HandsOn1\_Section.ipynb in each folder.

Drive link for submission:

[https://drive.google.com/drive/folders/1A864j3QNYuaNJD0F\\_s9lgFoF-NfPzCnR?usp=sharing](https://drive.google.com/drive/folders/1A864j3QNYuaNJD0F_s9lgFoF-NfPzCnR?usp=sharing)

## Task 2

Upload your notebook to the below drive with the following naming format:

SRN\_Assignment1\_Section.ipynb

Drive link for submission:

[https://drive.google.com/drive/folders/1bu9ZjUNoOU68rLA9NCM\\_RBivP-09bKWN?usp=sharing](https://drive.google.com/drive/folders/1bu9ZjUNoOU68rLA9NCM_RBivP-09bKWN?usp=sharing)

Deadline for submission - 12 Sept 2024,23:59