<u>**IDS-FA22-Assignment**</u>   <u>**Due Date: 21-10-2022**</u>

Download the Hello dataset from the following link:
https://drive.google.com/file/d/1xpPrW3ivqQPm029cr0KccSbWUAPiPF6S/view?usp=sharing

Provide detailed responses to the following, as discussed with you in class:

- Manually view the dataset, and see if it requires any cleaning/normalization
  (*if you clean/normalize any values, record them)
- Load the dataset (csv file) into a Pandas DataFrame.
- Print the list of all students whose first name starts with letter the 'H'.
- Print the total number of students who have a three words name (first-middle-surname).
- Print the percentage of students who have a CGPA of 3.0 or above.
- Plot a pie chart to show the ratio of male and female students.
- Plot the CGPA of all male students on a histogram with intervals 2.0-2.5, 2.6-3.0, 3.1-3.5, 3.6-4.0.
- Plot the HSSC-1 marks of all male vs female students on a scatter plot.
- Plot the favorite colors of male vs female students on a bar chart.
- Plot line chart of students and their birth months.
- Create a correlation matrix between HSSC-1 and HSSC-2 marks and then plot on a heatmap.
- Bonus points: What other things (insights) you can get from the dataset?