Your team needs to do the following:

1. Find one or several publicly available datasets containing data of interest to you. Report the origin of the data and give a short explanation why are you interested in this data. Describe the dataset: number of rows and columns, meaning of rows (observations) and column names (attributes). (20%)
2. Import the dataset into Python pandas and display the first 10 lines. Clean and filter the data (if needed). (20%)
3. Conduct quantitative and graphical EDA (include several graphs/charts and statistical summaries) (20%)
4. Conduct some of the statistical analysis techniques (linear regression, confidence intervals, hypothesis testing) covered in HW8 (10%)
5. Suggest and conduct one of the machine learning tasks covered in HW10 (classification, prediction, clustering) (10%)
6. Summarize your results – what were the interesting conclusions from the data analysis (10%)
7. Submit the 2 files (see below) in the correct format and on time. (10%)

The deliverables for the project:

1. A documentaddressing all the items above, explaining the steps done, and showing the obtained results.**(**Word or PDF document).
2. The commented Python code that you wrote to work with the data. (.py/ipynb/.pdf document).