**CSE523 Machine Learning**

**Heart Attack Prediction**

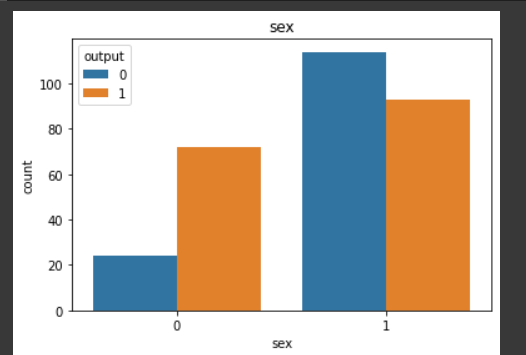
| **Group Members** | **Roll No** |
| --- | --- |
| **Shivam Thakker** | **AU1940193** |
| **Devarsh Sheth** | **AU1940189** |
| **Pranav Gandhi** | **AU1940313** |
| **Meet Jhaveri** | **AU1940284** |

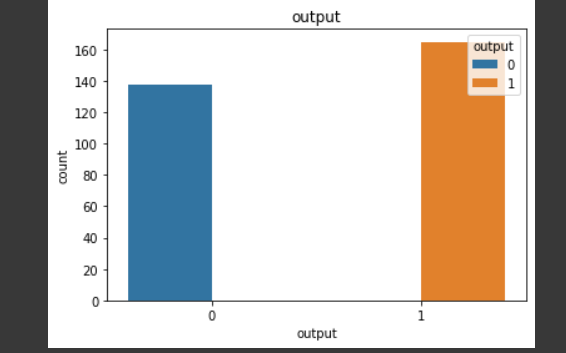
**Group-3 Week-1 Report**

* We first went through our dataset, understood the parameters used in it. As we don’t have a medical background it took us sometime to get some data terms.

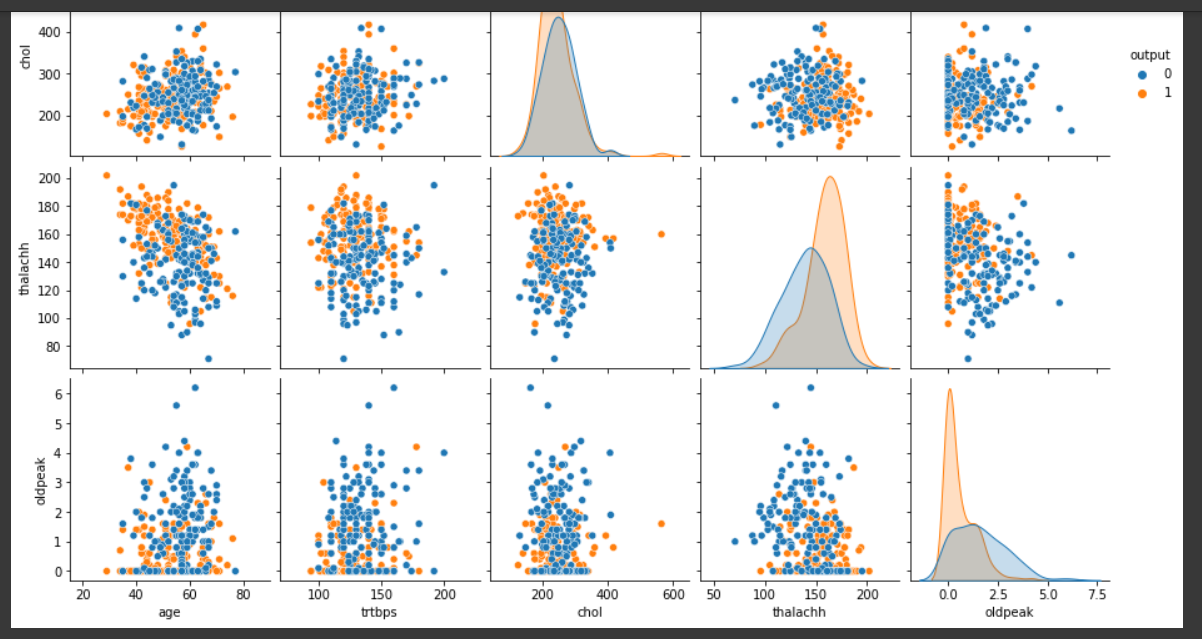
**These are the data terms that we went through:**

* + Age: Age of the patient
  + Sex: Sex of the patient
  + exng: Exercise induced angina (1 = yes, 0 = no)
  + caa: Number of major vessels (0-3)
  + cp: Chest pain type
    - 1: Typical angina
    - 2: Atypical angina
    - 3: Non-anginal pain
    - 4: Asymptomatic
  + trtbps: Resting blood pressure
  + chol: Cholesterol in mg/dl
  + fbs: Fasting blood sugar > 120 mg/dl (1 or 0)
  + rest\_ecg: Resting electrocardiographic results
    - 0: Normal1: Having ST-T Wave Abnormality
    - 2: Showing probable or definite left ventricular hypertrophy by Estes' criteria
  + thalachh: Maximum heart rate achieved
  + target: 0=less chance of herat attack 1 = More chance of heart attack
* The next thing we did was, we read and analysed the data.
* For the cleaning data step, we did the missing value analysis, to find out further we carried a Unique value analysis, so that we know how many unique values each column header has.
* Further we did categorical value analysis by plotting graphs using sns.countplot(). We plotted a graph of sex vs. count , cp vs. count , fbs vs. count, restecg vs. count , exng vs. count , slp vs. count , caa vs. count, thall vs. count, and output vs. count.





* Further we also did numeric value analysis by using numericList = ["age", "trtbps", "chol", "thalachh", "oldpeak", "output"]



* Plotted graph of each element vs. other element with respect to visualizing the output. For eg:- age vs age, age vs. trtbps, age vs. chol and so on with taking hue = output(to visualize the output) in sns.pairplot().
* Then we carried out Standardization using StandardScalar() and .fit\_transform().