Heart Disease Diagnostic Analysis
By Vinotha

PROJECT DETAIL:

Project Title	Heart Disease Diagnostic Analysis
Technologies	Data Science
Domain	Healthcare
Project Difficulties level	Intermediate

PROBLEM STATEMENT:

Health is real wealth in the pandemic time we all realized the brute effects of covid-19 on all irrespective of any status. You are required to analyze this health and medical data for better future preparation.

Do ETL: Extract- Transform and Load data from the heart disease diagnostic database

You can perform EDA through python. The database extracts various information such as Heart disease rates, Heart disease by gender, by age.

You can even compare attributes of the data set to extract necessary information. Make the necessary dashboard with the best you can extract from the data. Use various visualization and features and make the best dashboard

Find key metrics and factors and show the meaningful relationships between attributes.

STEPS INVOLVED:

Raw data collection
Importing libraries in jupyter notebook
Load dataset
Missing value imputations
Handling outliers
Data cleaning
Exploratory data analysis(EDA)
Modelling
Reporting

DATASET INFORMATION:

age: The person's age in years

sex: The person's sex (1 = male, 0 = female)

cp: The chest pain experienced (Value 1: typical angina, Value 2: atypical angina, Value 3: non-anginal pain, Value 4: asymptomatic)

trestbps: The person's resting blood pressure (mm Hg)

chol: The person's cholesterol measurement in mg/dl

fbs: The person's fasting blood sugar (> 120 mg/dl, 1 = true; 0 = false)

restecg: Resting electrocardiographic measurement (0 = normal, 1 = having ST-T wave abnormality, 2 = showing probable or definite left ventricular hypertrophy by Estes' criteria)

thalach: The person's maximum heart rate achieved

exang: Exercise induced angina (1 = yes; 0 = no)

oldpeak: ST depression induced by exercise relative to rest

slope: the slope of the peak exercise ST segment (Value 1: upsloping, Value 2: flat)

ca: The number of major vessels (0-3)

thal: A blood disorder called thalassemia (1 = normal; 1 = fixed defect; 3 = reversable defect)

target : Heart disease (0 = no, 1 = yes)

Dataset:

Dataset is available in the given link. You can download it at your convenience.

https://drive.google.com/file/d/1U8CHK_ye5jmcuYEeIOYIYcMzK2ooqLUV/view

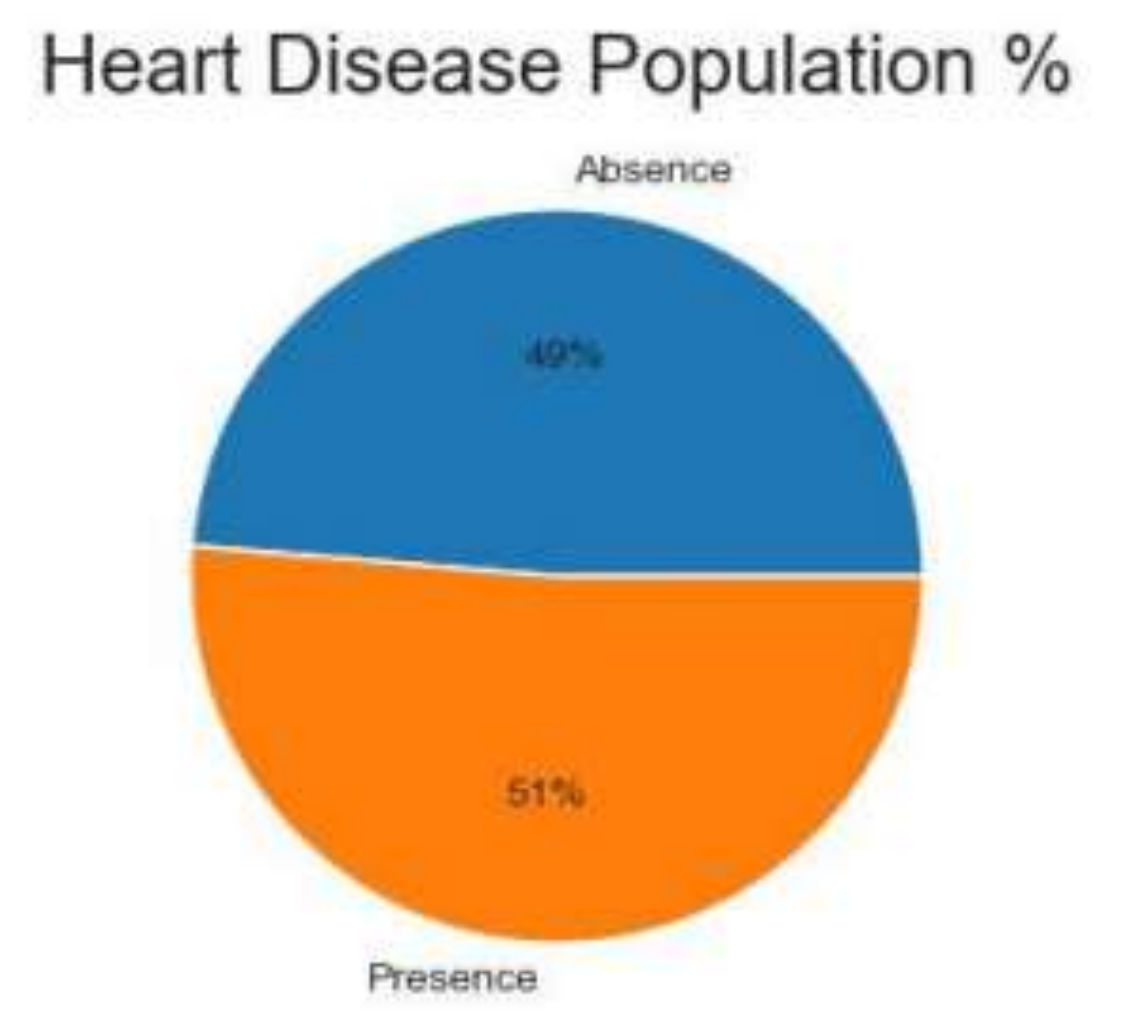
TOOL USED:

JUPYTER NOTEBOOK

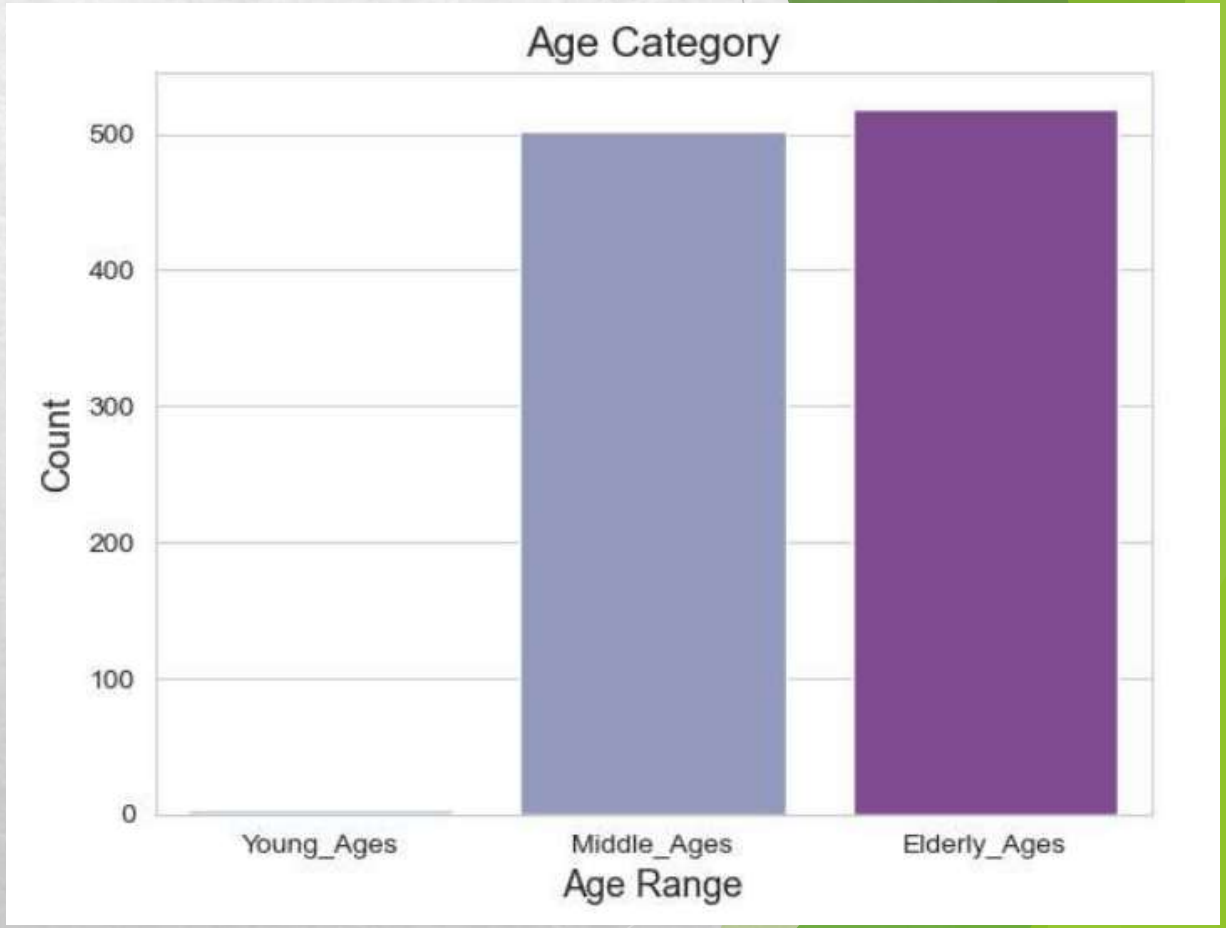
LANGUAGE USED:

PYTHON

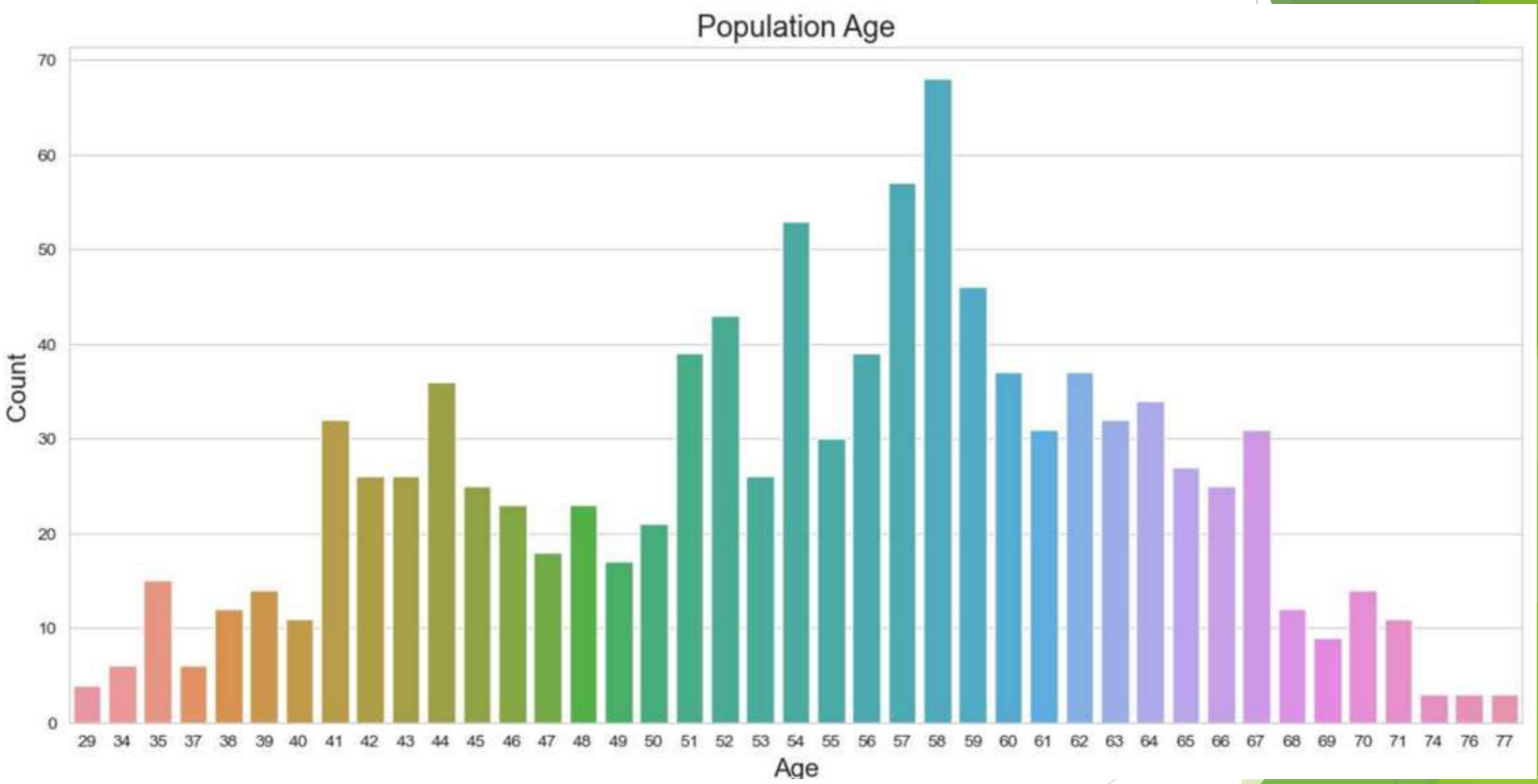
What Kind of Population do we have?

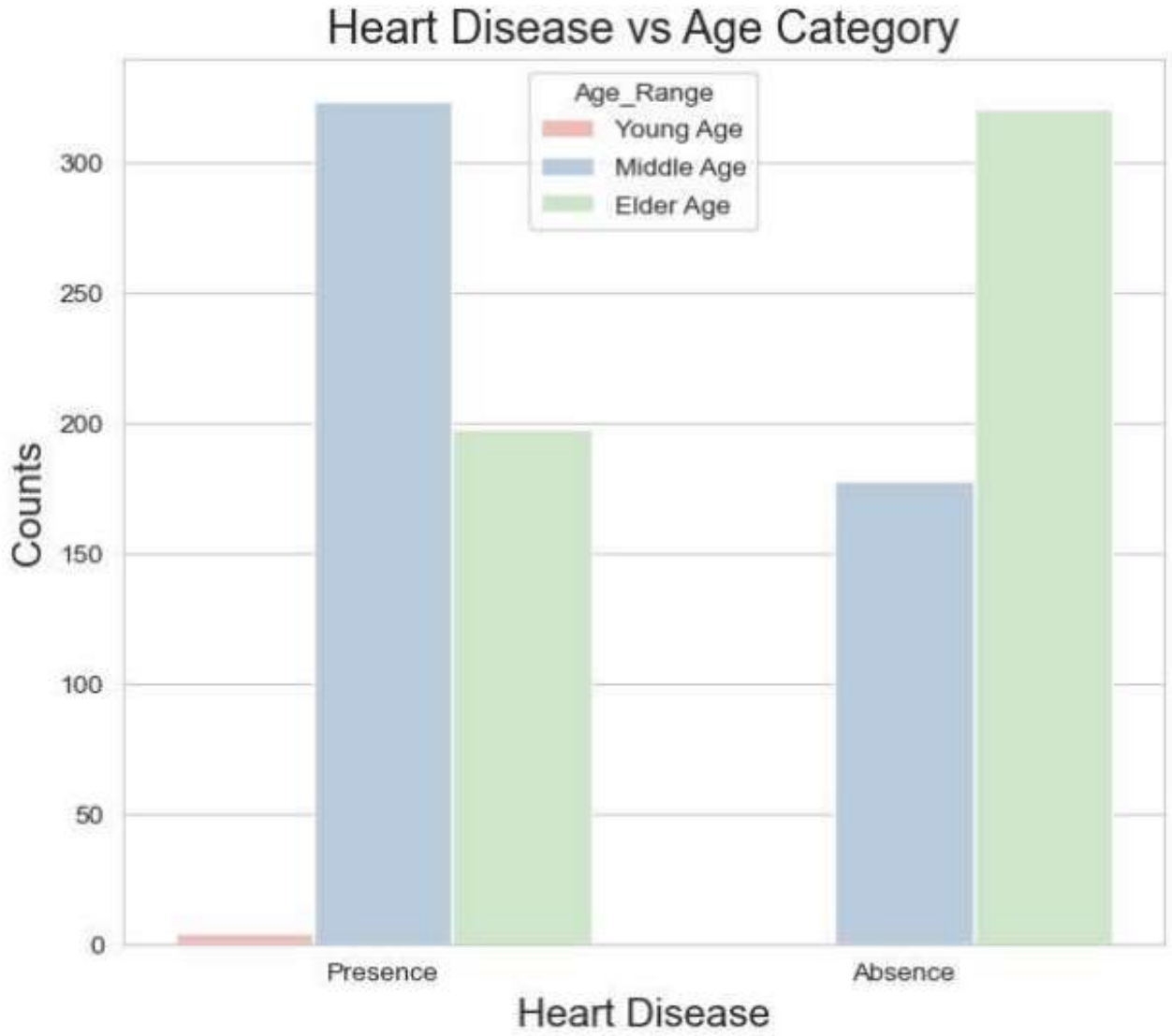


51% suffer from heart disease

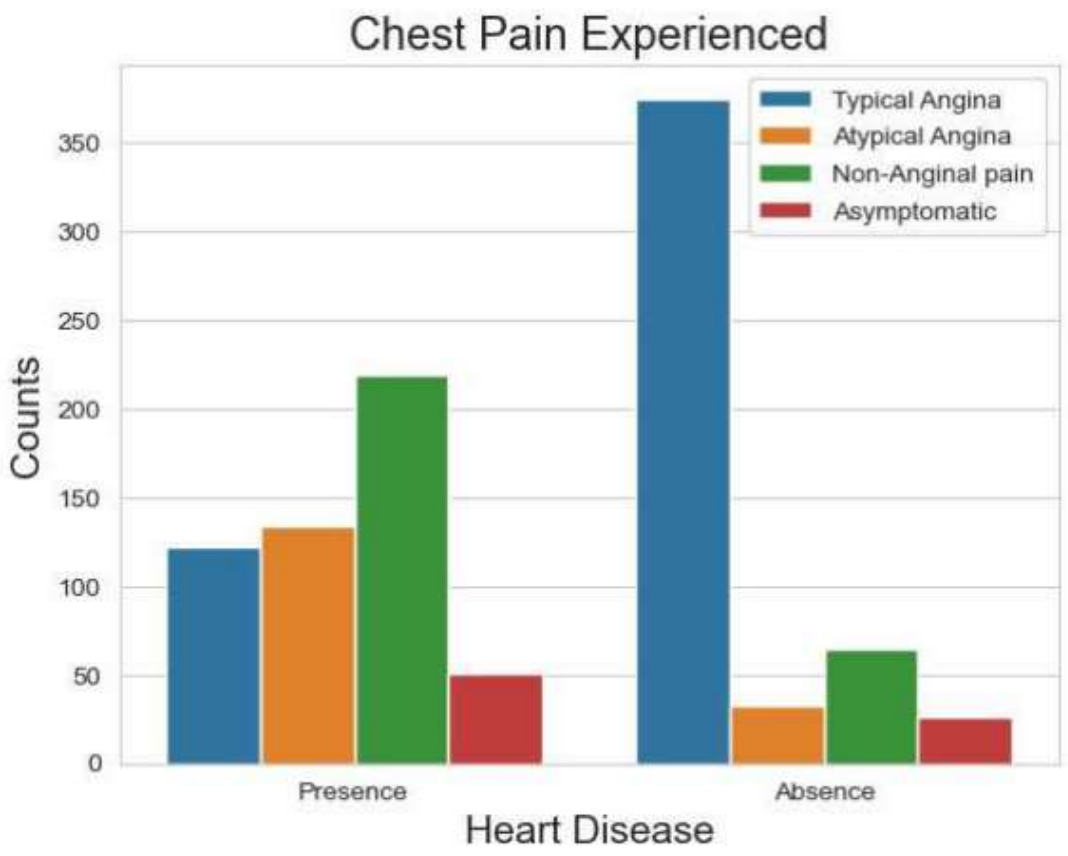


Middle age and elderly age are likely to suffer from heart diseases

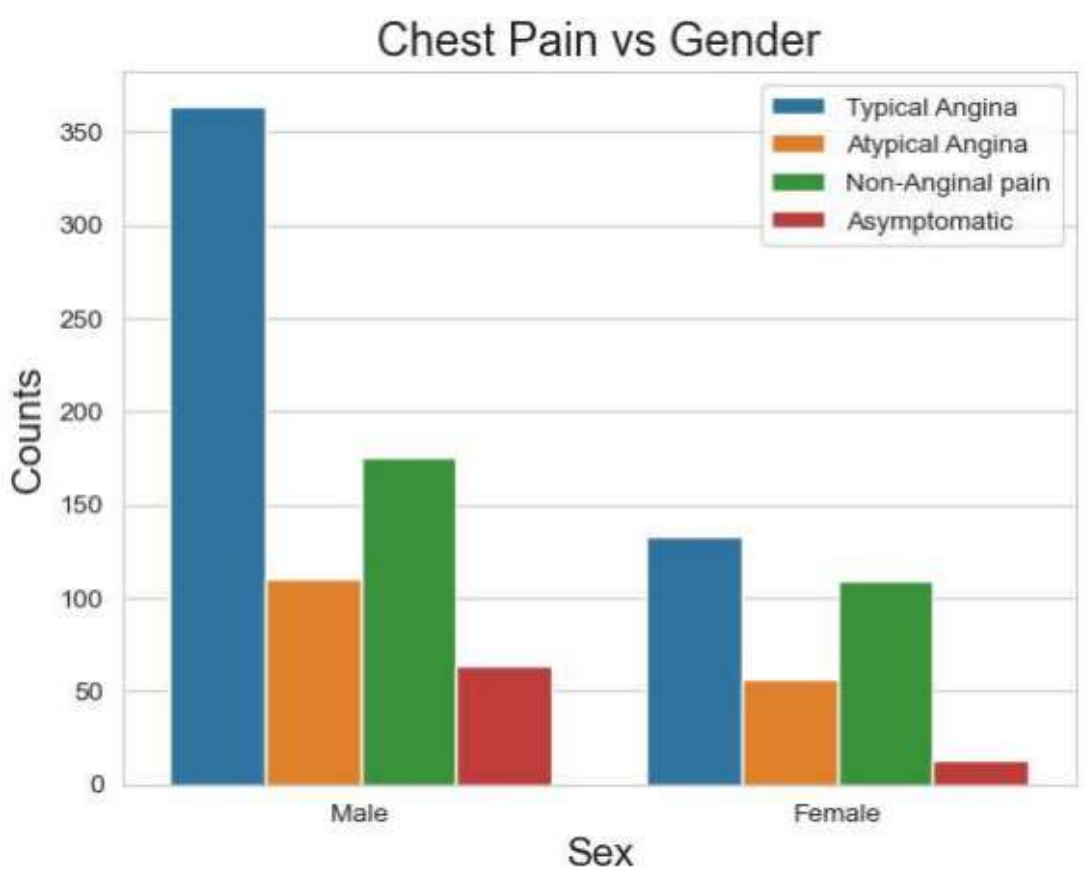

Population Age
Count
70
60
50
40
30
20
10
0
29 34 35 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 74 76 77
Age



Middle Age People are most affected by Heart Disease AND Young Age People are mostly FREE from any kind of Heart Disease

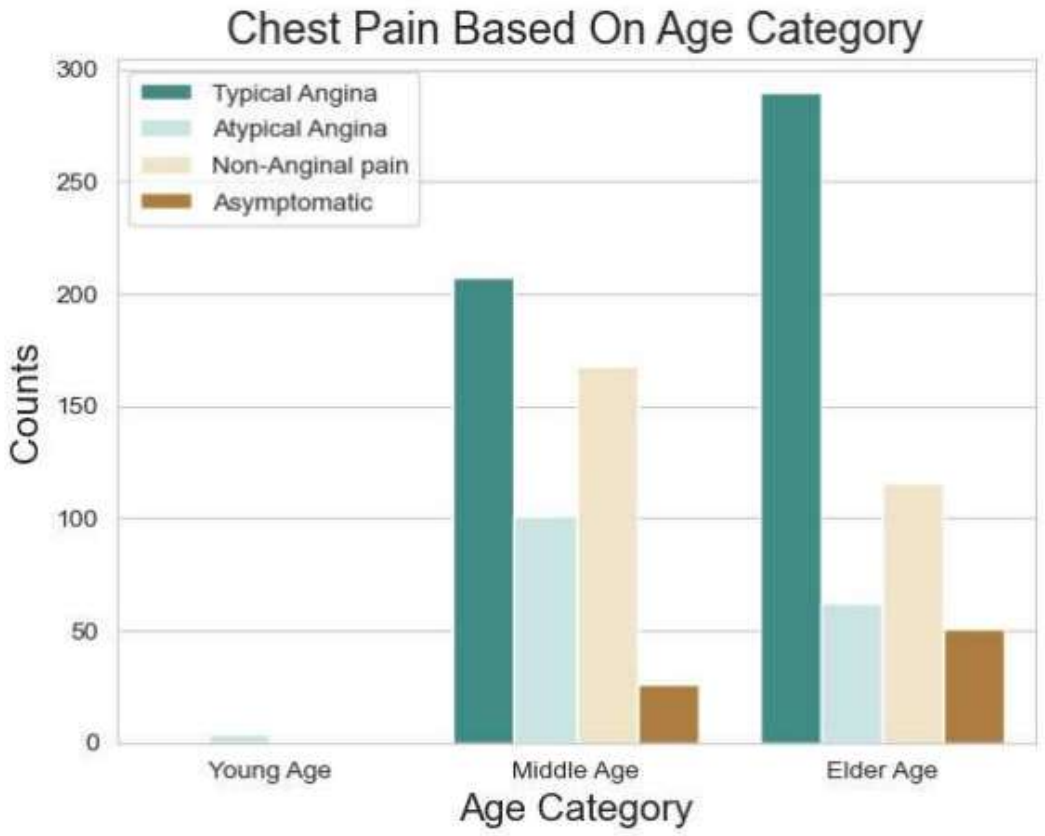


It seems people having non anginal pain have a higher chance of heart disease



Higher number of men are suffering from Typical Angina

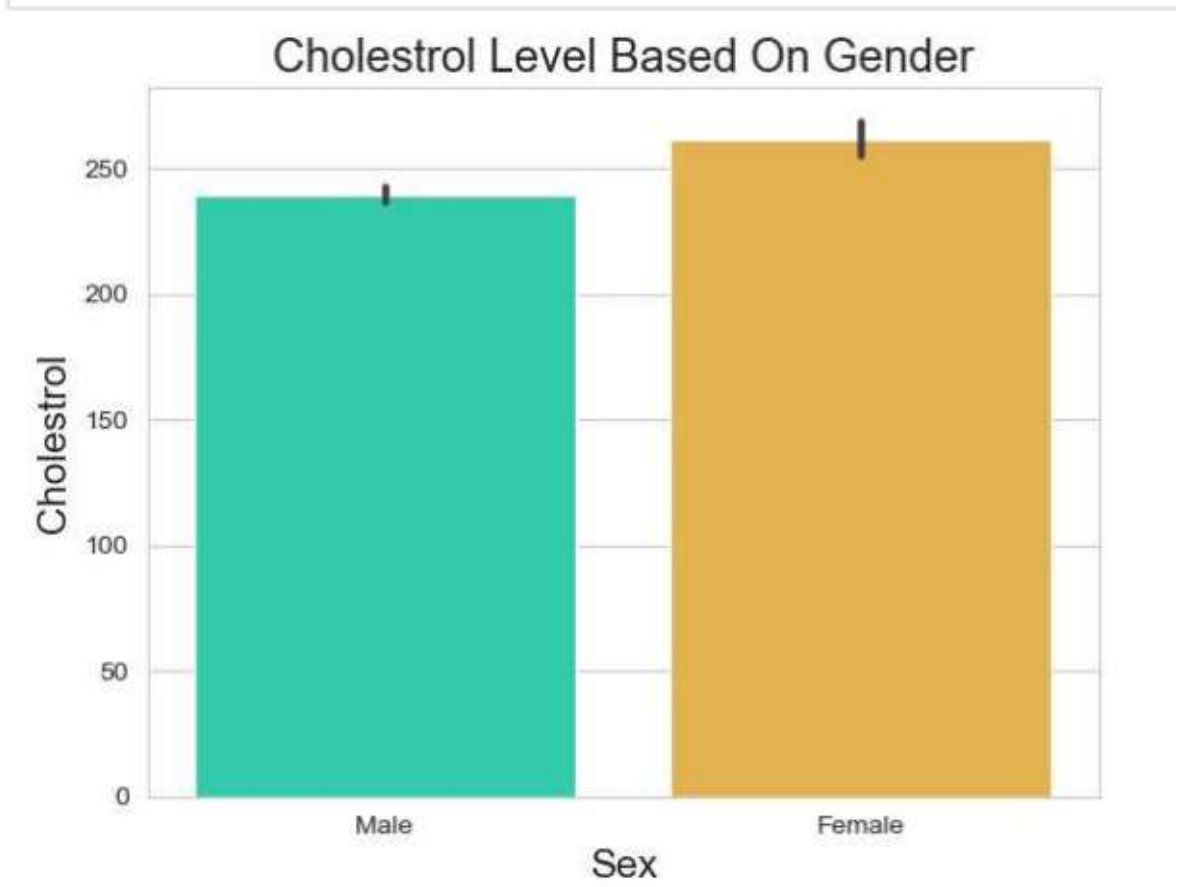
CHEST PAIN EXPERIENCED BY PATIENTS

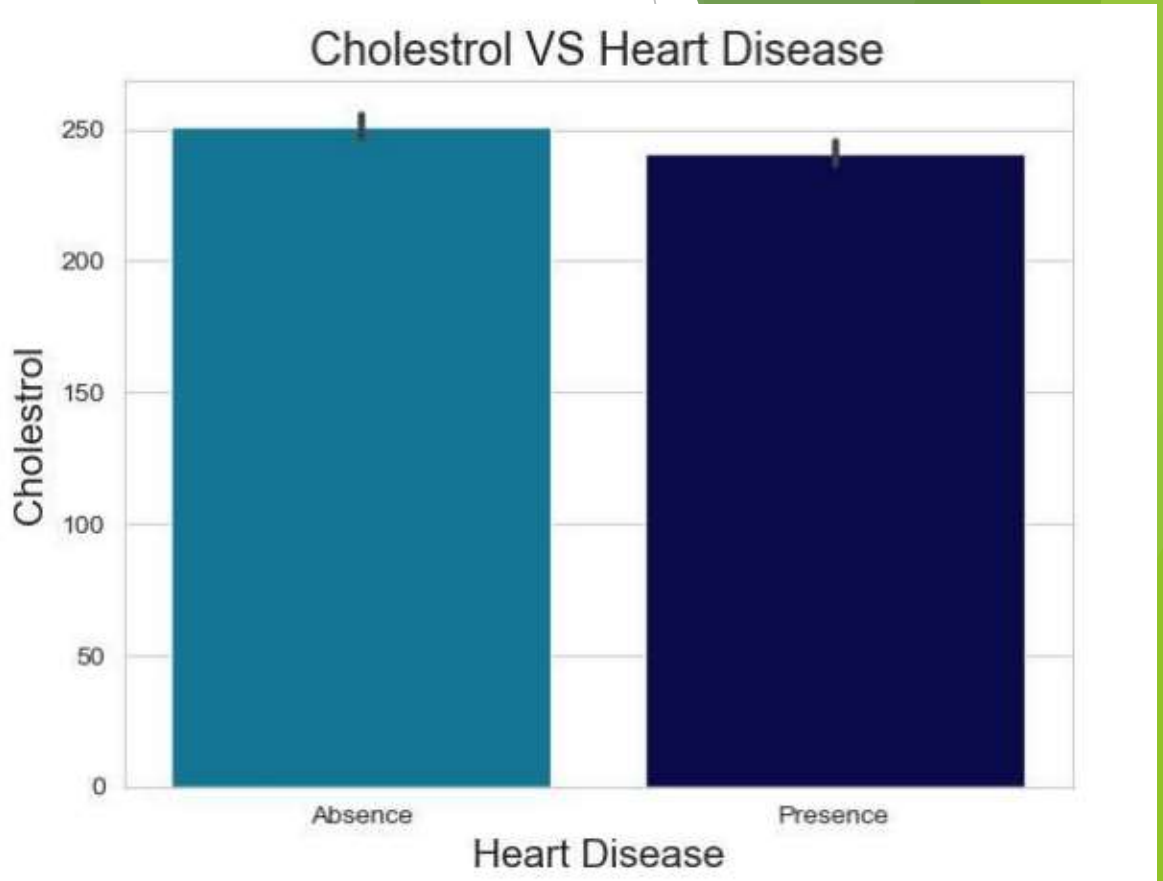


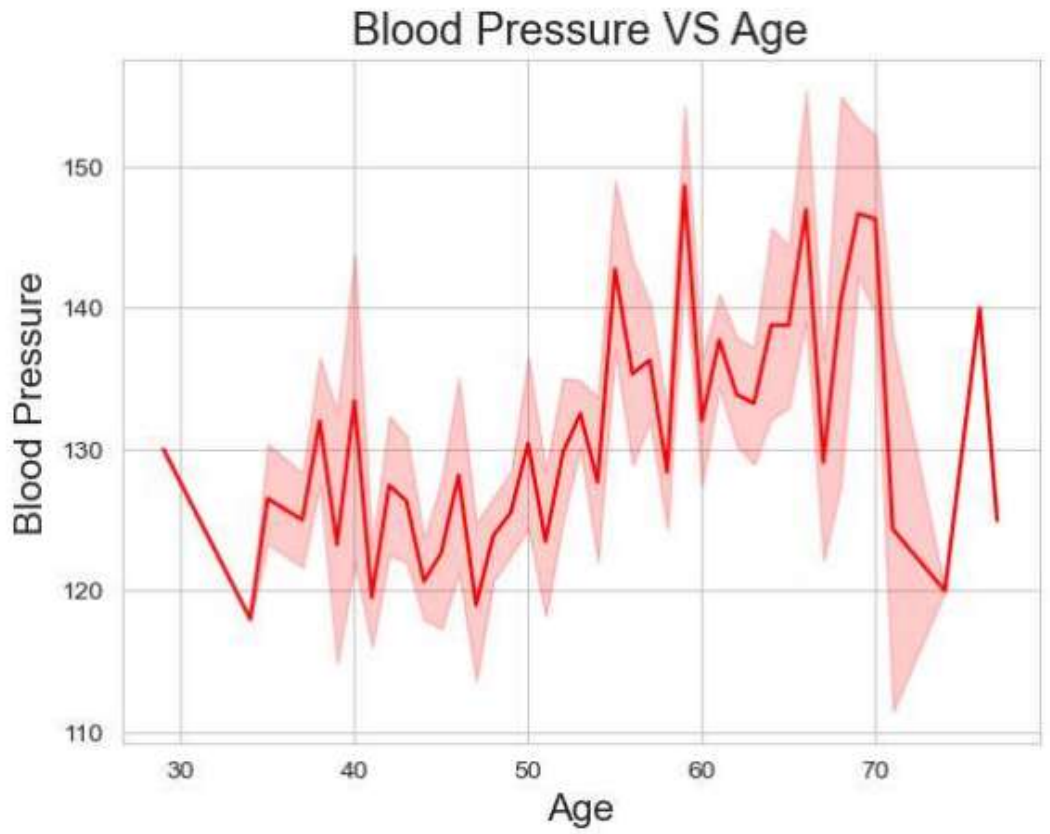
High number of Typical Angina in Elderly age

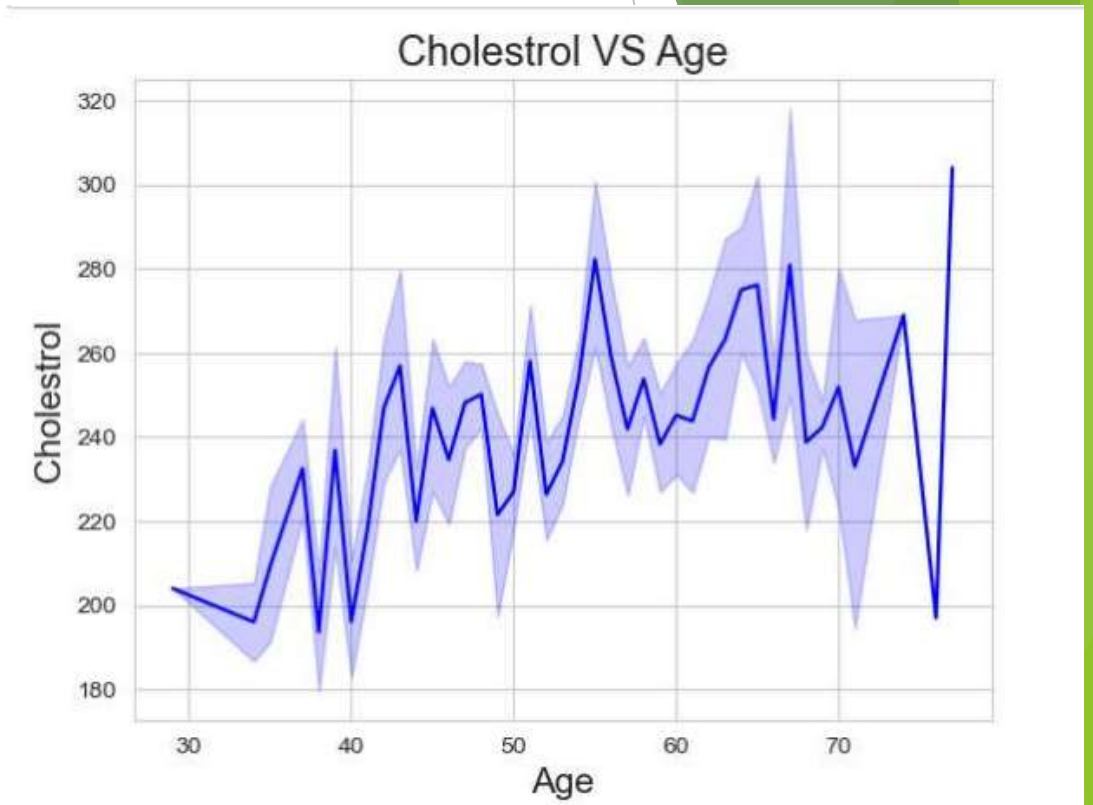
Blood Pressure is almost equal in Male and Female

CHOLESTEROL LEVEL



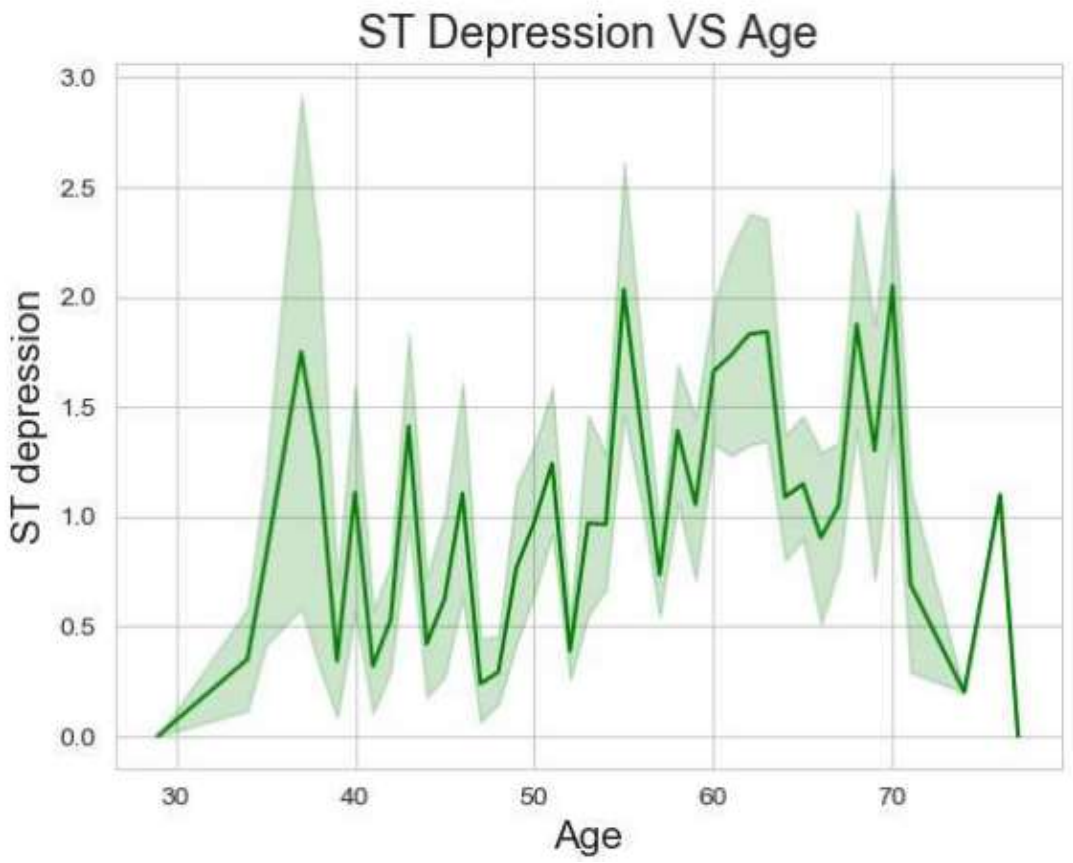




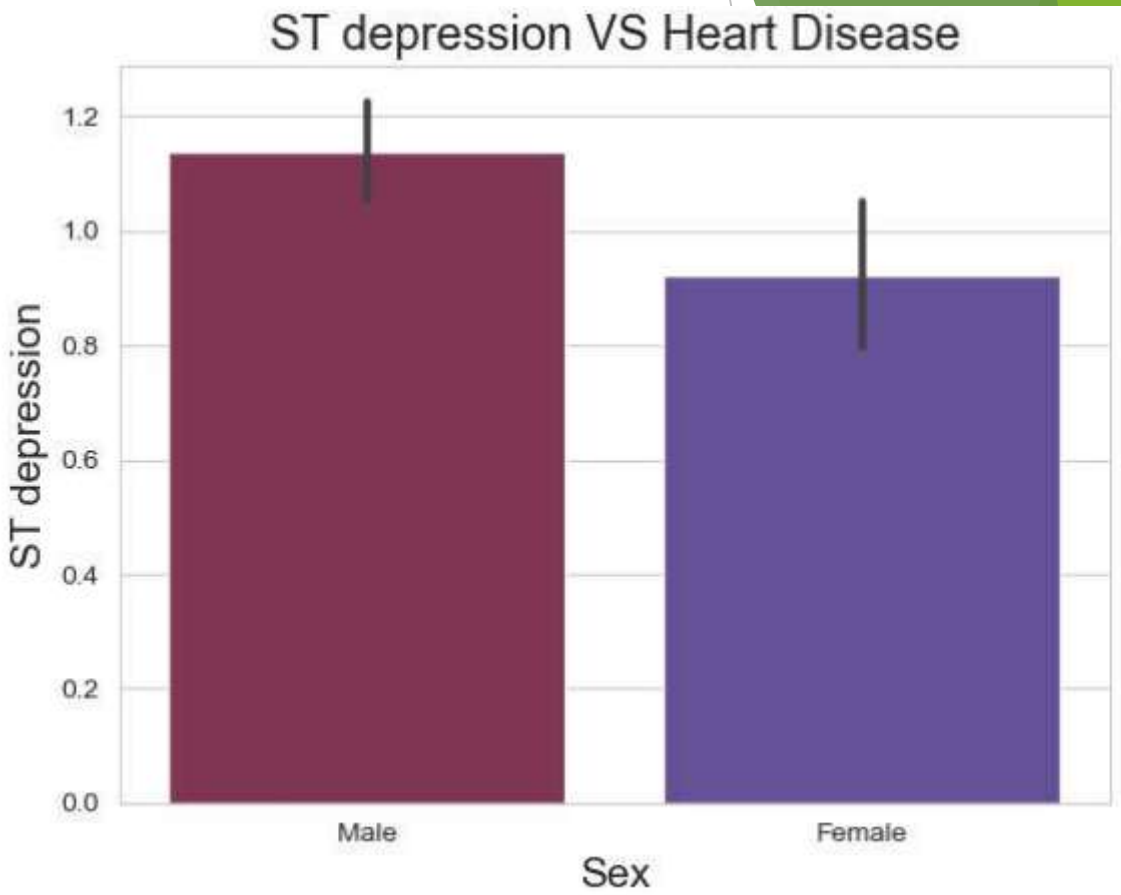


Blood Pressure increases between age of 50 to 60 and reaches peak in elderly people.

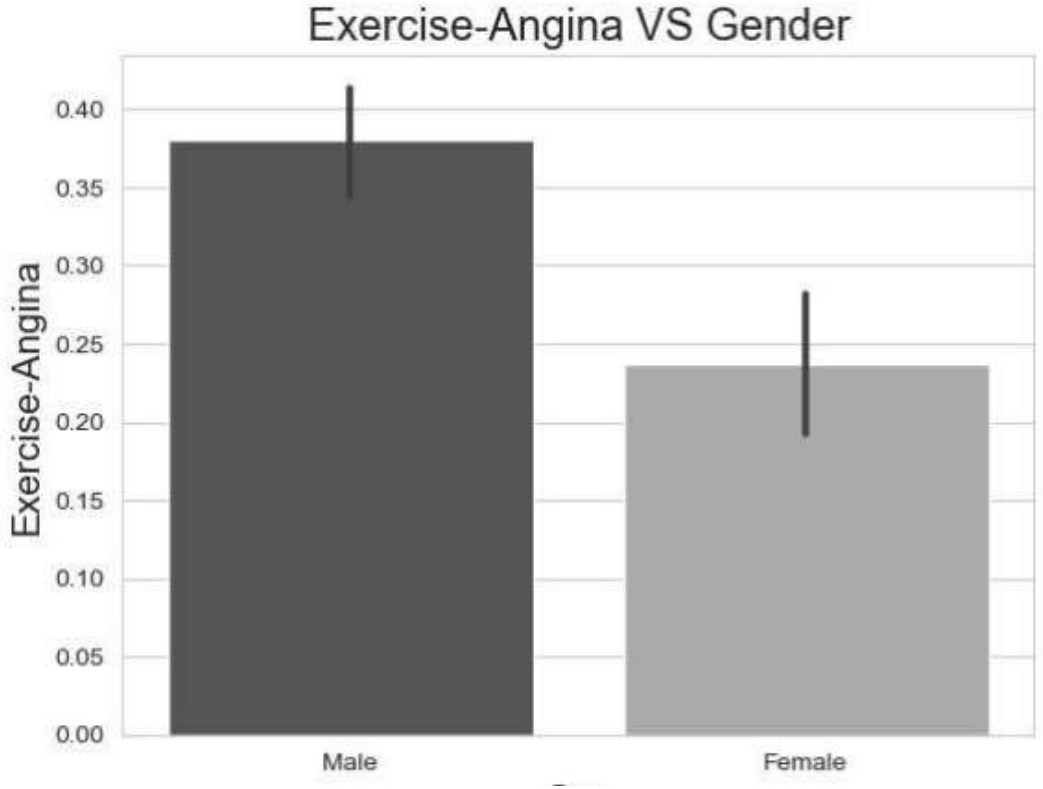
Similarly Cholestrol Increases in the age group of 50-60 and reaches peak in elderly age.



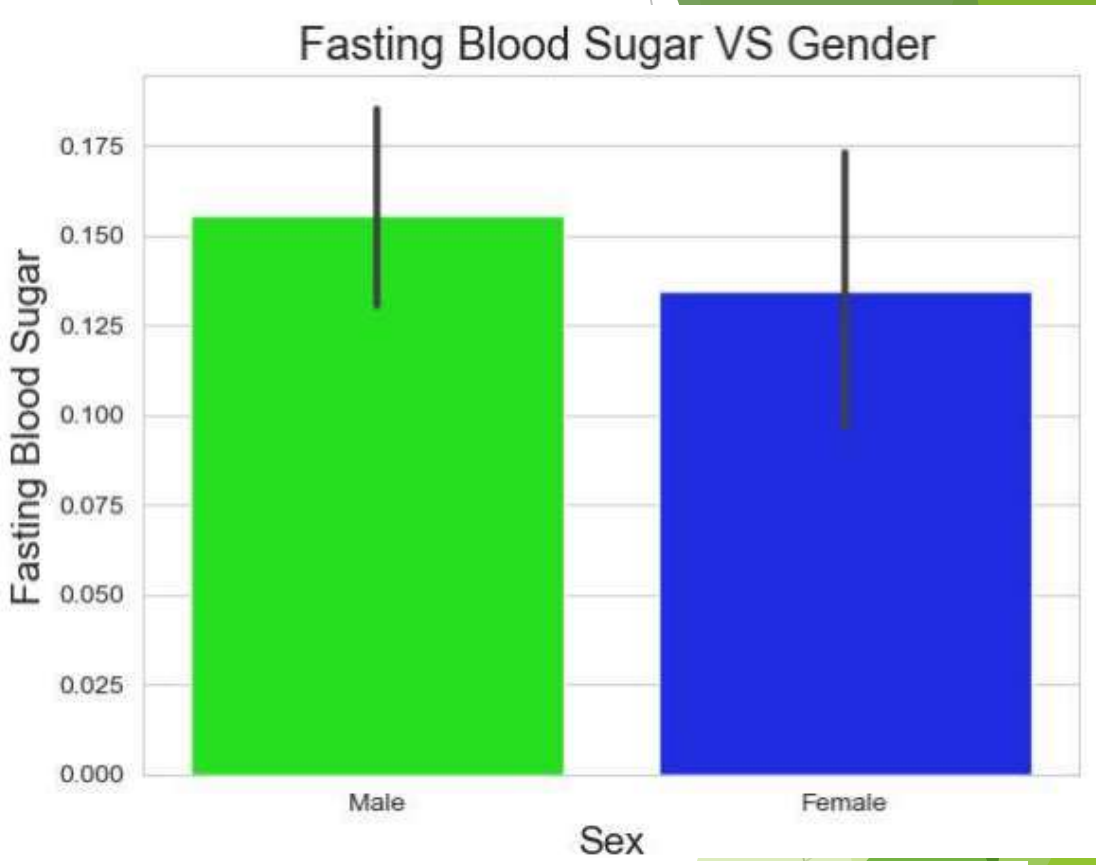
ST depression is high in age group: 30-40



Comparatively male are prone to ST depression than females



-> Male suffer from Angina during exercise



Male have high no. of Fasting Blood Sugar over 120

HEAT MAP


age 1 -0.1 -0.072 0.27 0.22 0.12 -0.13 -0.39 0.088 0.21 -0.17 0.27 0.071 -0.23
sex -0.1 1 -0.041 -0.079 -0.2 0.027 -0.055 -0.049 0.14 0.085 -0.027 0.11 0.2 -0.28
cp -0.072 -0.041 1 0.038 -0.082 0.079 0.044 0.31 -0.4 -0.17 0.13 -0.18 -0.17 0.43
trestbps 0.27 -0.079 0.038 1 0.13 0.18 -0.12 -0.039 0.061 0.19 -0.12 0.1 0.059 -0.14
chol 0.22 -0.2 -0.082 0.13 1 0.027 -0.15 -0.022 0.067 0.065 -0.014 0.074 0.097 -0.1
fbs 0.12 0.027 0.079 0.18 0.027 1 -0.1 -0.0089 0.049 0.011 -0.062 0.14 -0.031 -0.041
restecg -0.13 -0.055 0.044 -0.12 -0.15 -0.1 1 0.048 -0.066 -0.05 0.086 -0.078 -0.016 0.13
thalach -0.39 -0.049 0.31 -0.039 -0.022 -0.0089 0.048 1 -0.38 -0.35 0.4 -0.21 -0.097 0.42
exang 0.088 0.14 -0.4 0.061 0.067 0.049 -0.066 -0.38 1 0.31 -0.27 0.11 0.21 -0.44
oldpeak 0.21 0.085 -0.17 0.19 0.065 0.011 -0.05 -0.35 0.31 1 -0.58 0.22 0.2 -0.44
slope -0.17 -0.027 0.13 -0.12 -0.014 -0.062 0.086 0.4 -0.27 -0.58 1 -0.073 -0.098 0.35
ca 0.27 0.11 -0.18 0.1 0.074 0.14 -0.078 -0.21 0.11 0.22 -0.073 1 0.15 -0.38
thal 0.071 0.2 -0.17 0.059 0.097 -0.031 -0.016 -0.097 0.21 0.2 -0.098 0.15 1 -0.34
target -0.23 -0.28 0.43 -0.14 -0.1 -0.041 0.13 0.42 -0.44 -0.44 0.35 -0.38 -0.34 1
age sex cp trestbps chol fbs restecg thalach exang oldpeak slope ca thal target
1.0 0.8 0.6 0.4 0.2 0.0 −0.2 −0.4

KEY PERFORMANCE INDICATOR (KPI)

1. Percentage of People Having Heart Disease
2. Age Distribution including Gender
3. Gender Distribution Based on Heart Disease
4. Chest Pain Experienced by People Suffering from Heart Disease
5. Blood Pressure, Cholesterol Level and Maximum Heart Rate of People According to their Age and Heart Disease Patients.
6. ST Depression Experienced by People According to their age and heart disease.

CONCLUSION:

- 51% people are suffering from heart disease.
- Middle aged people and Elderly aged People are more prone to heart disease.
- Male are more prone to heart disease.
- People having Typical Angina have a higher chance of heart disease.
- Female have higher cholesterol than male.
- Blood Pressure increases between age of 55 to 70 and
- Cholesterol increases in the age group of 60-70.
- ST depression mostly increases between the age group of 30-40.
- Fasting blood sugar is higher in male.

THANK YOU