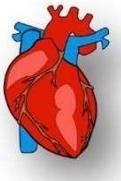
Heart Disease Diagnostic Analysis

DETAILED PROJECT REPORT



Shubham Tembhurne

PROJECT DETAIL

Project Title	Heart Disease Diagnostic Analysis
Technology	**Business Intelligence**
Domain	**Healthcare**
Project Difficulty level	**Advance**
Programming Language Used	**Python**
Tools Used	**Jupyter Notebook, MS-Excel, Tableau**

OBJECTIVE

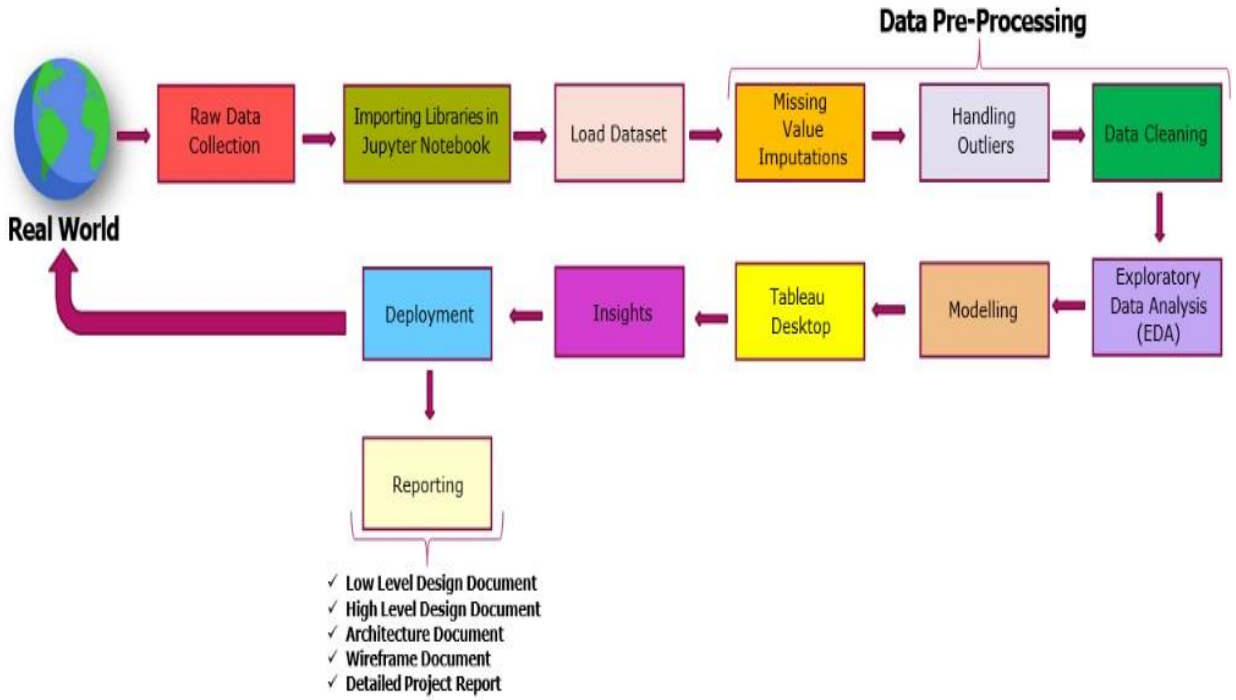
- The goal of this project is to analyze the heart disease

Occurrence based on a combination of features that describes heart disease.

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
- A data set that is formed by taking into consideration some of the information of 303 individuals. The problem is based on the given information about each individual we have to calculate whether that individual will suffer from heart disease or not.

Architecture



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 on admission to the hospital)

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, Value 3: downsloping)

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

thal: A blood disorder called thalassemia (3 = normal; 6 = fixed defect; 7 = reversable defect) num: Heart disease (0 = no, 1 = yes)

Why these Parameter are important?

Age: Age is the most important risk factor in developing cardiovascular or heart diseases, with approximately a tripling of risk with each decade of life. Coronary fatty streaks can begin to form in adolescence. It is estimated that 82 percent of people who die of coronary heart disease are 65 and older. Simultaneously, the risk of stroke doubles every decade after age 55.

Sex: Men are at greater risk of heart disease than pre-menopausal women. Once past menopause, it has been argued that a woman's risk is similar to a man's although more recent data from the WHO and UN disputes this. If a female has diabetes, she is more likely to develop heart disease than a male with diabetes.

Fasting Blood Sugar: Not producing enough of a hormone secreted by your pancreas (insulin) or not responding to insulin properly causes your body's blood sugar levels to rise, increasing your risk of a heart attack.

Cholesterol: A high level of low-density lipoprotein (LDL) cholesterol (the "bad" cholesterol) is most likely to narrow arteries. A high level of triglycerides, a type of blood fat related to your diet, also ups your risk of a heart attack. However, a high level of high-density lipoprotein (HDL) cholesterol (the "good" cholesterol) lowers your risk of a heart attack

Resting Blood Pressure: Over time, high blood pressure can damage arteries that feed your heart. High blood pressure that occurs with other conditions, such as obesity, high cholesterol, or diabetes, increases your risk even more.

Resting ECG: For people at low risk of cardiovascular disease, the USPSTF concludes with moderate certainty that the potential harms of screening with resting or exercise ECG equal or exceed the potential benefits. For

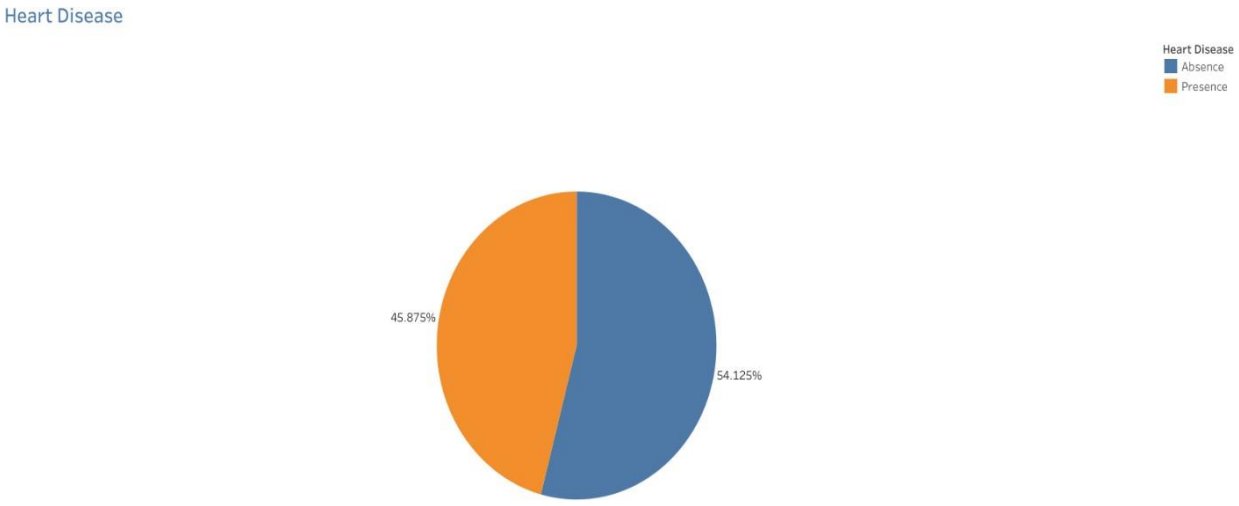
people at intermediate to high risk, current evidence is insufficient assess the balance of benefits and harms of screening.

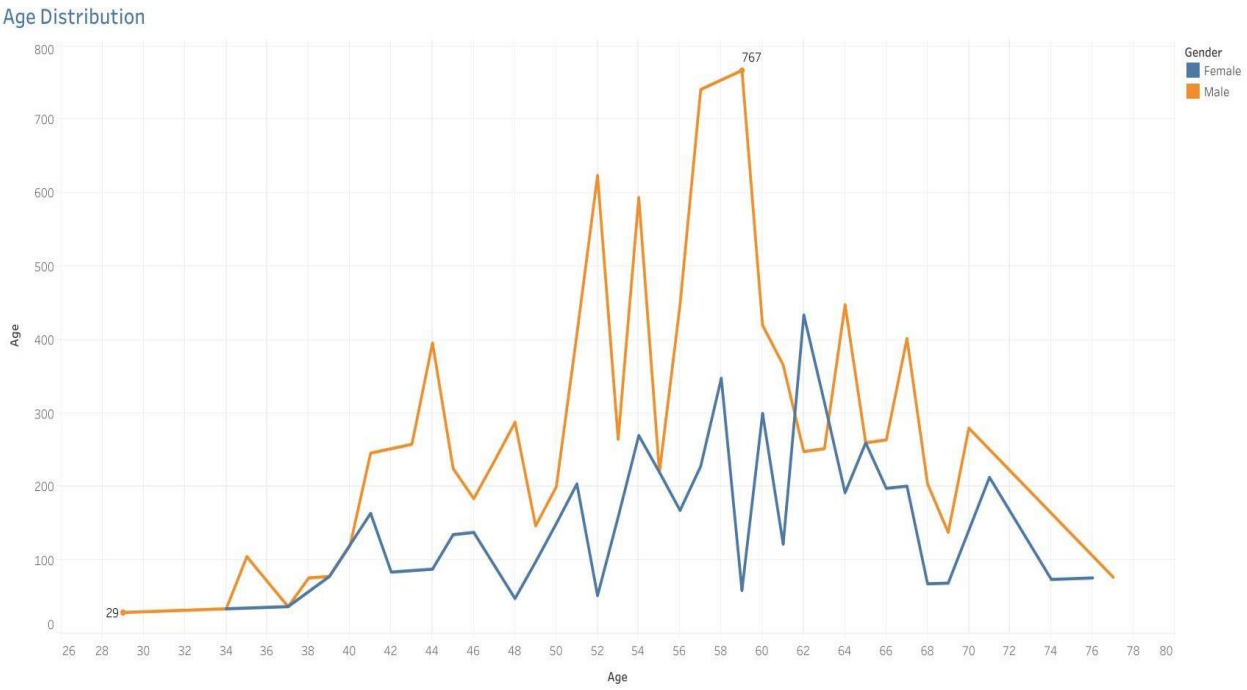
Max heart rate achieved: The increase in the cardiovascular risk, associated with the acceleration of heart rate, was comparable to the increase in risk observed with high blood pressure. It has been shown that an increase in heart rate by 10 beats per minute was associated with an increase in the risk of cardiac death by at least 20%, and this increase in the risk is similar to the one observed with an increase in systolic blood pressure by 10 mm Hg.

ST Depression: In unstable coronary artery disease, ST-segment depression is associated with a 100% increase in the occurrence of three-vessel/left main disease and an increased risk of subsequent cardiac events. In these patients, an early invasive strategy substantially decreases death/myocardial infarction.

INSIGHTS

1. What Kind of the Population do we have?

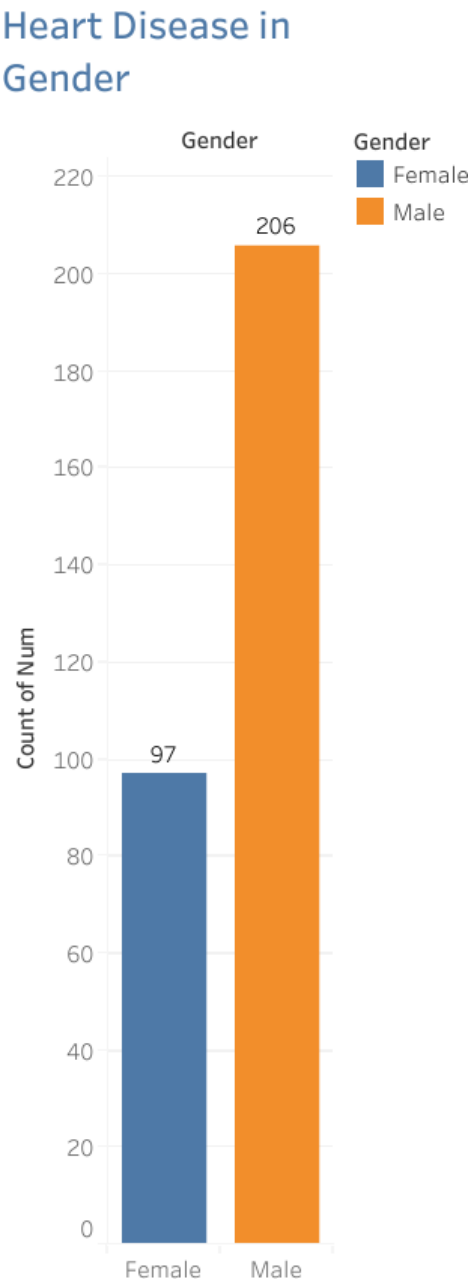


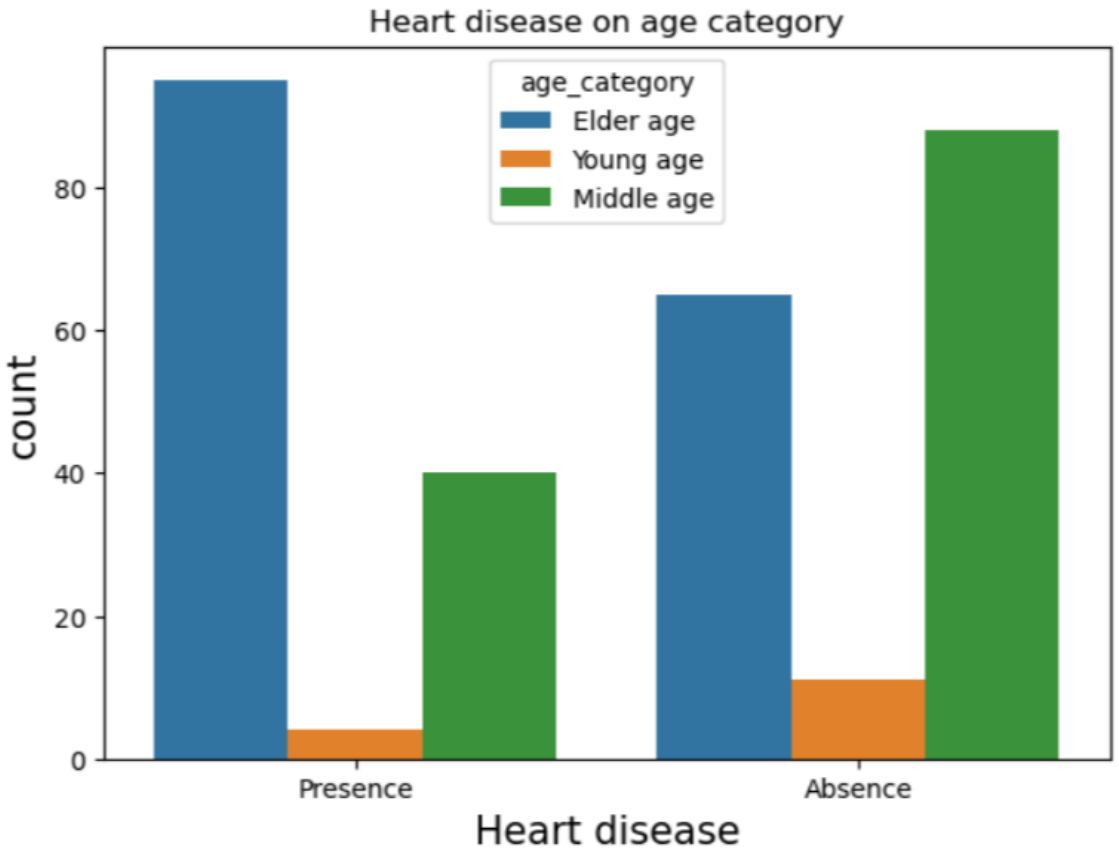


- The People are not suffering with Heart Disease are 54.13%.

- The People are suffering from Heart Disease Presence are 45.87%.
- I found that maximum male is from middle age (50-60) Years and most of the female are Elder age (55-65) Year category.

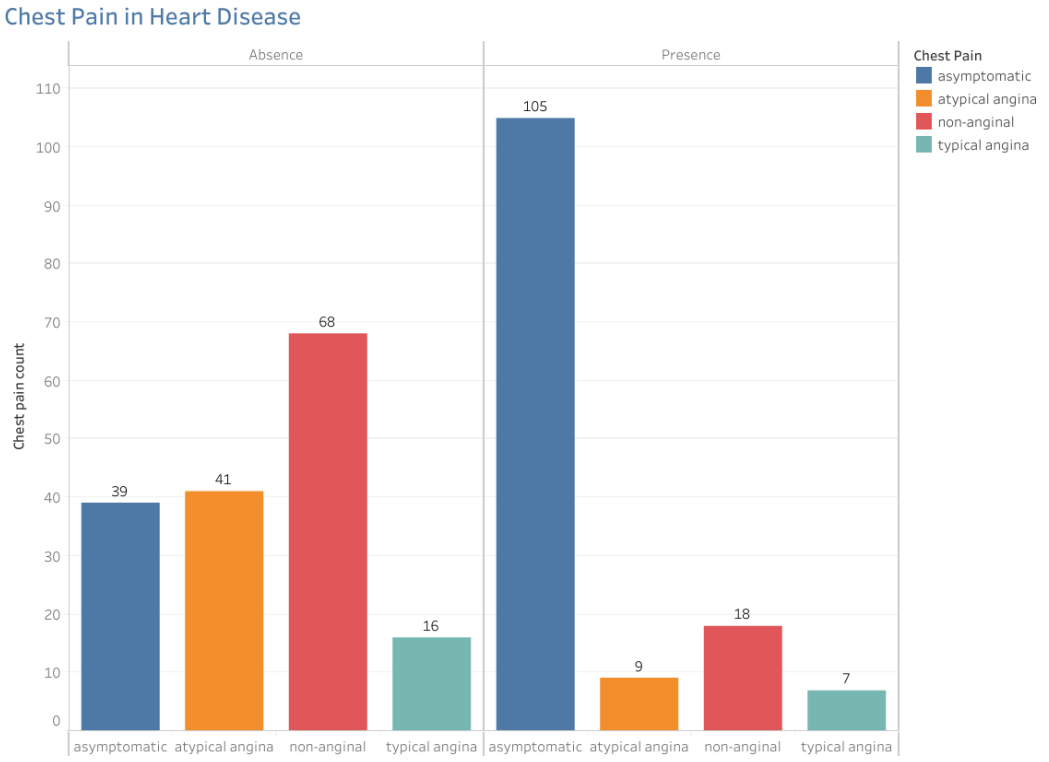
2. Who Suffered from heart Disease?

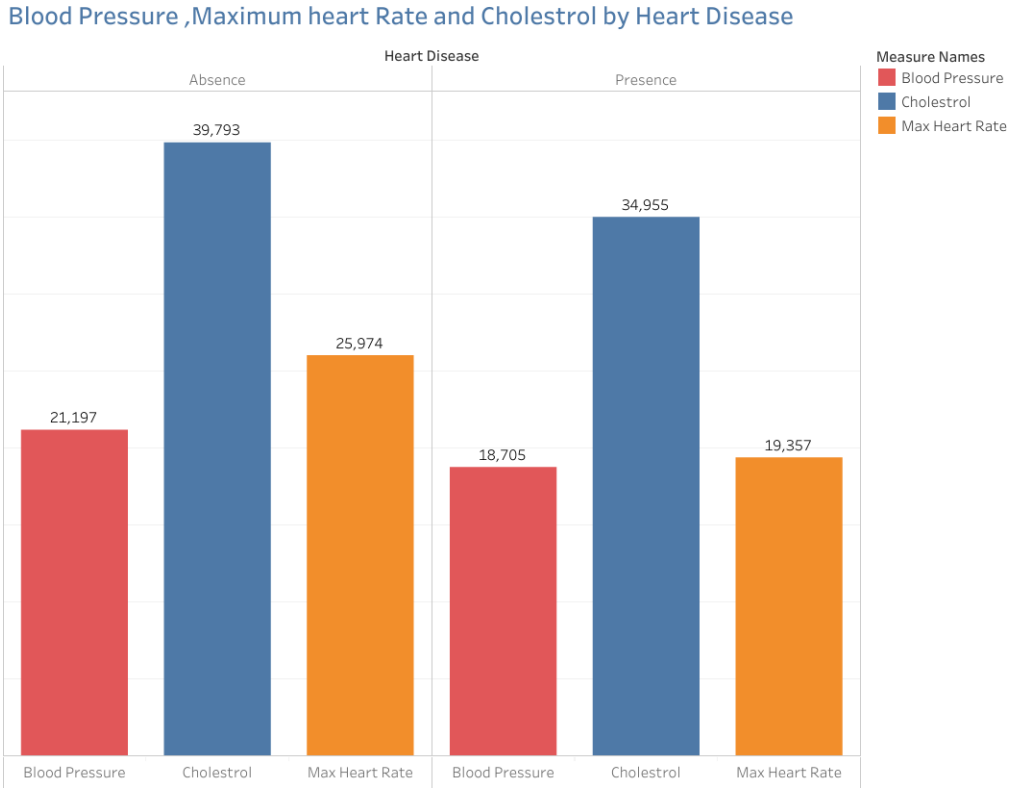




- Males are more prone to heart disease.
- The most of the heart disease present in the elder age and the middle age people are mostly free from the Heart disease.

3. What symptoms do people experience in heart disease?

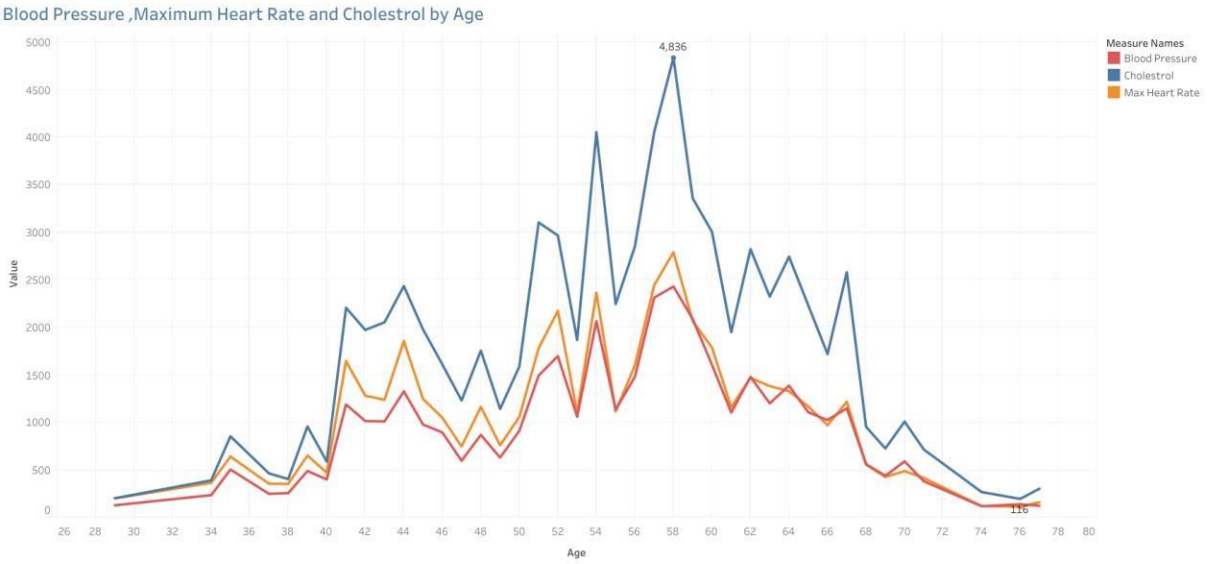


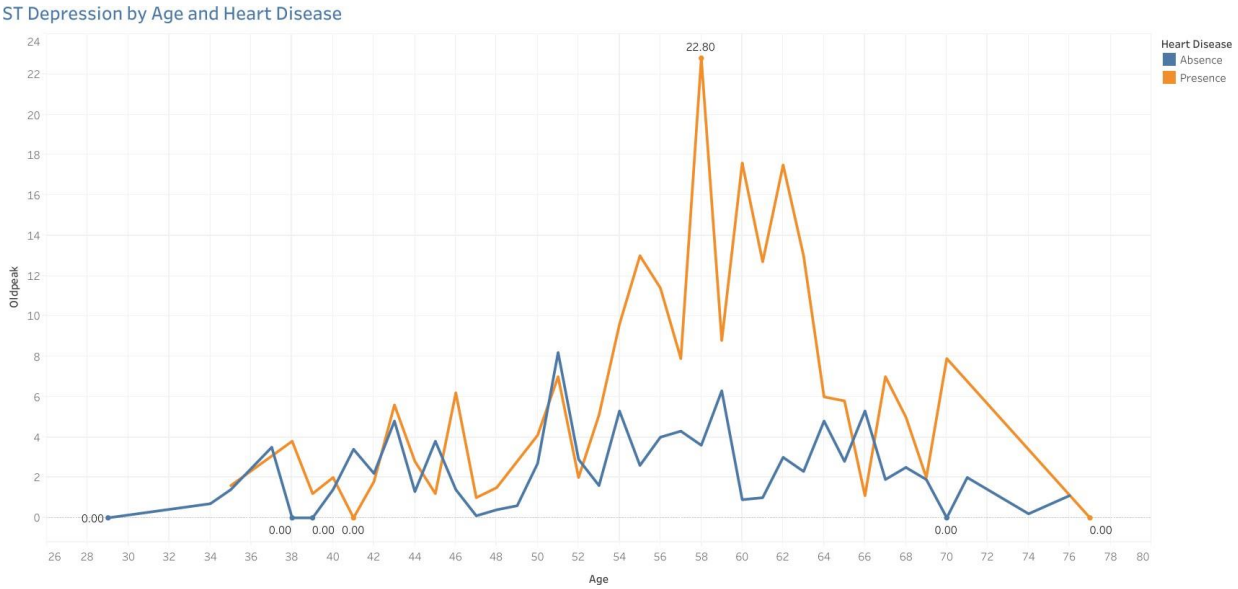


- The chest pain experience in the heart disease is most likely asymptomatic. That have higher chances of heart disease.

- Asymptomatic means there are no symptoms. You are considered asymptomatic if you: Have recovered from an illness or condition and no longer have symptoms of that illness or condition.

- The Cholesterol is highly risk factor which can lead to heart disease.





➢ The cholesterol level is increase at the age of 50-60 and then continue the same pattern.

➢ The blood pressure is increase at the age of 50-65 and then continue the same pattern.

➢ We can see that ST depression is increase suddenly at the age of 30-40. An ST depression can be an outcome of an electrocardiogram (ECG) test. It can indicate health conditions like hypokalemia, myocardial ischemia, or a side effect of medication.

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

- 45.87% of people suffering from heart disease.
- Males are more prone to heart disease.
- The chest pain experience in the heart disease is most likely asymptomatic. That have higher chances of heart disease.
- The Cholesterol is highly risk factor which can lead to heart disease.
- The cholesterol level is increase at the age of 50-60 and then continue the same pattern.
- The blood pressure is increase at the age of 50-65 and then continue the same pattern.
- We can see that ST depression is increase suddenly at the age of 30-40.
- Maximum male is from middle age (50-60) Years and most of the female are Elder age (55-65) Year category.

Q & A

Q1) What's the source of data?

Ans) The Dataset was taken from

https://drive.google.com/drive/folders/165Pjmfb9W9PGy0rZjHEA22LW0Lt3Y-Q8?usp=sharing

Q2) What was the type of data?

Ans) The data was a combination of numerical and Categorical values.

Q3) What's the complete flow you followed in this Project?

Ans) Refer slide 5th for better Understanding

Q4) What techniques were you using for data?

Ans) -Removing unwanted attributes

-Visualizing relation of independent variables with each other and output variables

-Removing outliers

-Cleaning data and imputing if null values are present.

-Converting Numerical data into Categorical values.

Q6) What were the libraries that you used in Python?

Ans) I used Pandas, NumPy, Matplotlib, **and Seaborn libraries in Pandas.**

THANK YOU