



# HEART DISEASE DIAGNOSIS ANALYSIS

DETAILED PROJECT REPORT

SRIMANTA SHEKHAR MISHRA  
SWADESH RANJAN PANDA

# PROJECT DETAILS

<b>Project Title</b>	Heart Disease Diagnostic – Analysis
<b>Technology</b>	Business Intelligence
<b>Domain</b>	Healthcare
<b>Project Difficulty level</b>	Advanced
<b>Programming Language Used</b>	Python
<b>Tools Used</b>	Jupyter Notebook, MS-Excel, MS-Power BI

# OBJECTIVE

- ▶ The goal of this project is to analyse the heart disease occurrence, based on a combination of features that describes the 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 analyse this health and medical data for better future preparation.
- ▶ A dataset is formed by taking into consideration some of the information of 303 individuals.

# 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)

**restbpps:** 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)

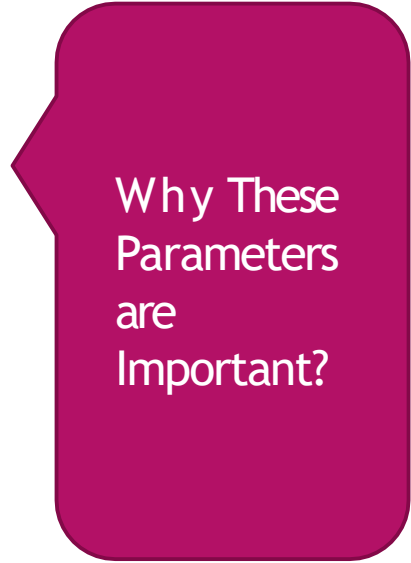
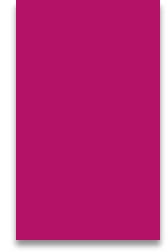
**oldpeak:** ST depression induced by exercise relative to rest

**slope:** the slope of the peak exercise ST segment (Value 1: upsloping, Value 2: flat, Value 3: down sloping)

**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)



**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.

**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.

**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 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 heart attack. However, a high level of high-density lipoprotein (HDL) cholesterol (the "good" cholesterol) lowers your risk of heart attack.

**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 to assess the balance of benefits and harms of screening.



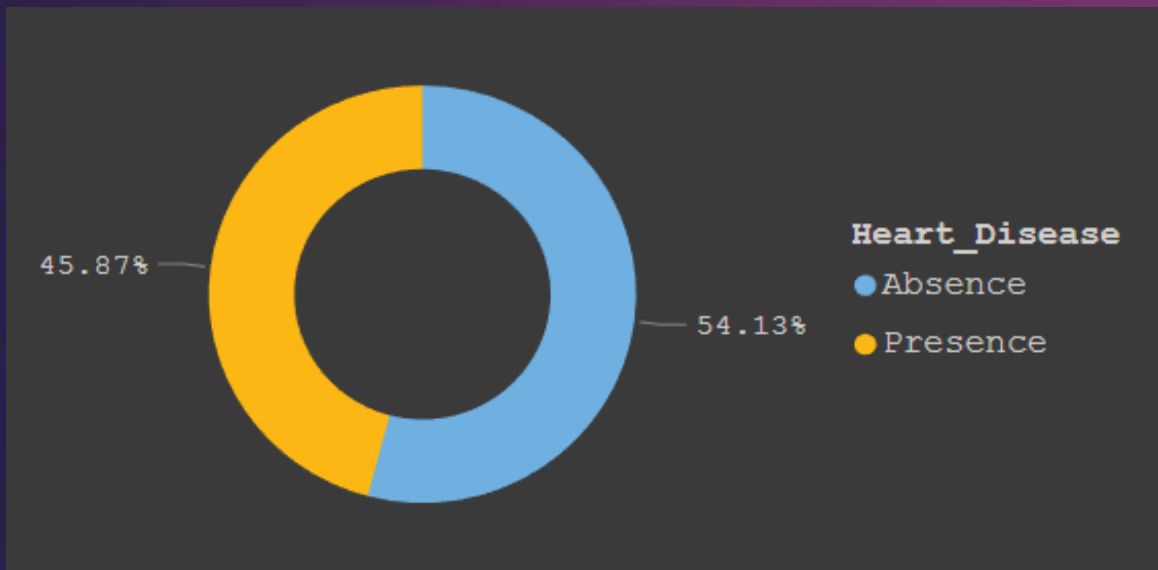
**Max heart rate achieved:** The increase in the cardiovascular risk, associated with the acceleration of heart rate, was comparable to the increase in risk observed with high blood pressure. It has been shown that an increase in heart rate by 10 beats per minute was associated with an increase in the 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 to an increased risk of subsequent cardiac events. In these patients an early invasive strategy substantially decreases death/myocardial infarction.

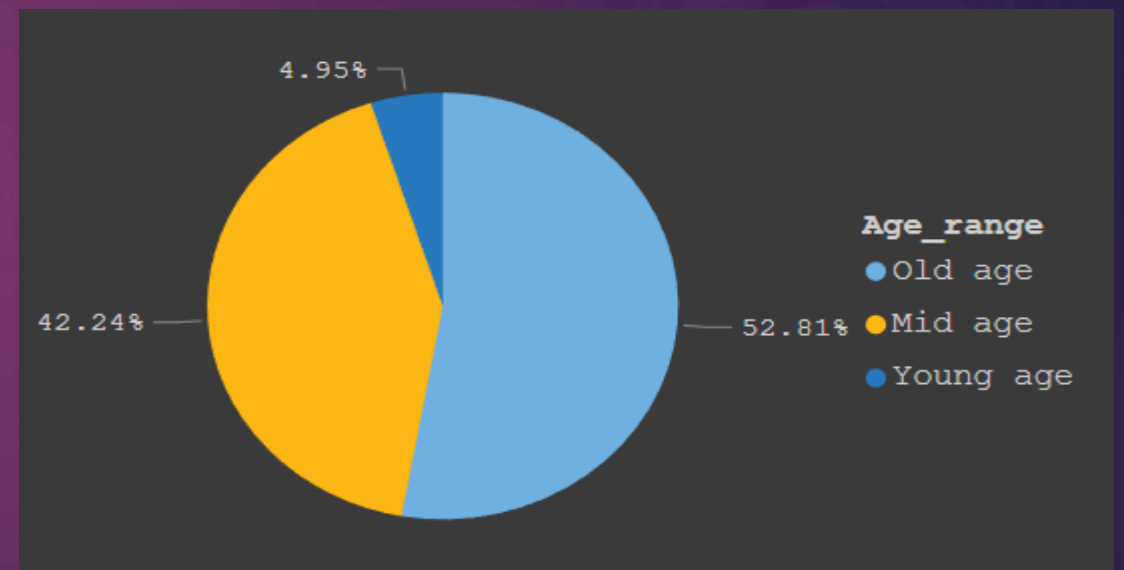


# INSIGHTS

## What Kind of Population do we have?

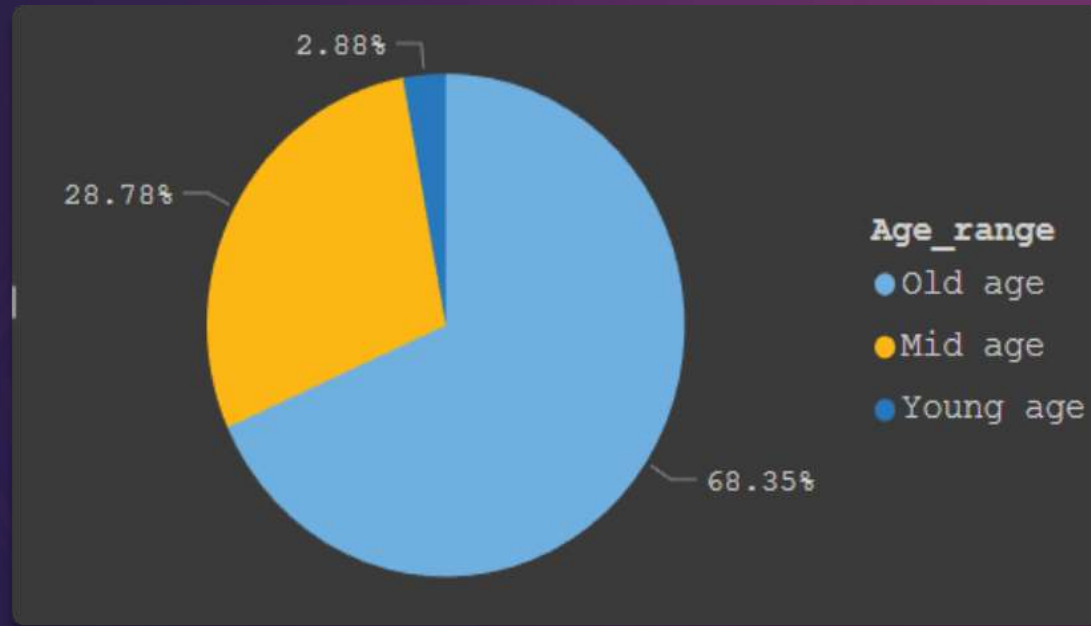


- 45.87% People suffering from heart disease from our Total population.

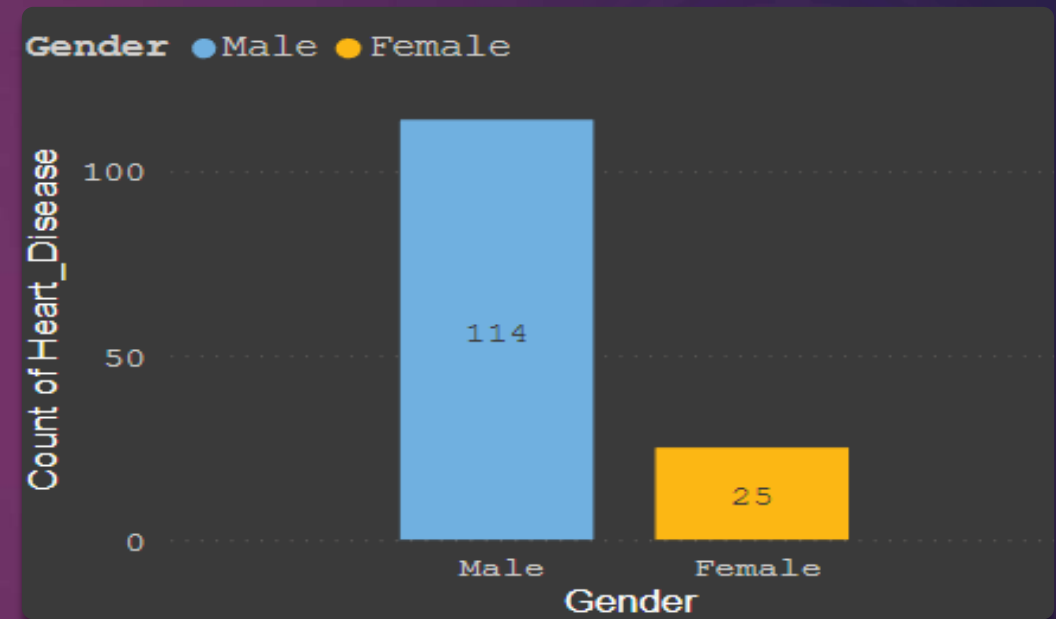


- We have more Old age people (>55) i.e. 52.81% in our dataset.

# Who suffers from heart disease?

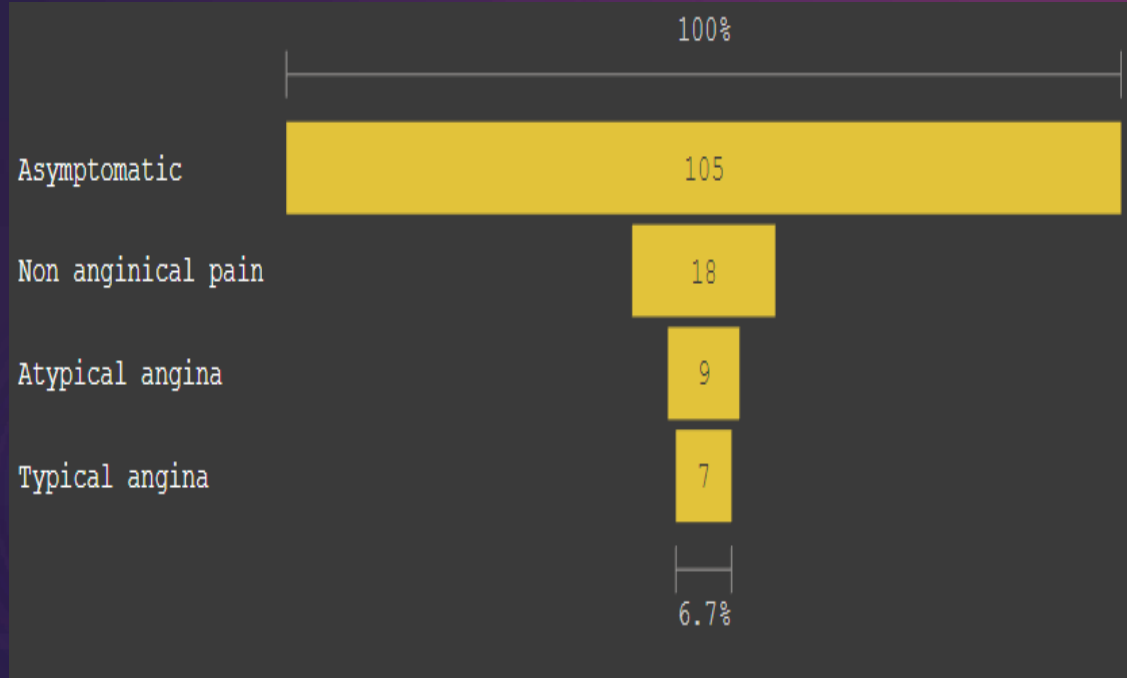


- 68.35 % people who belongs to elder age (>55) are more prone to heart disease.

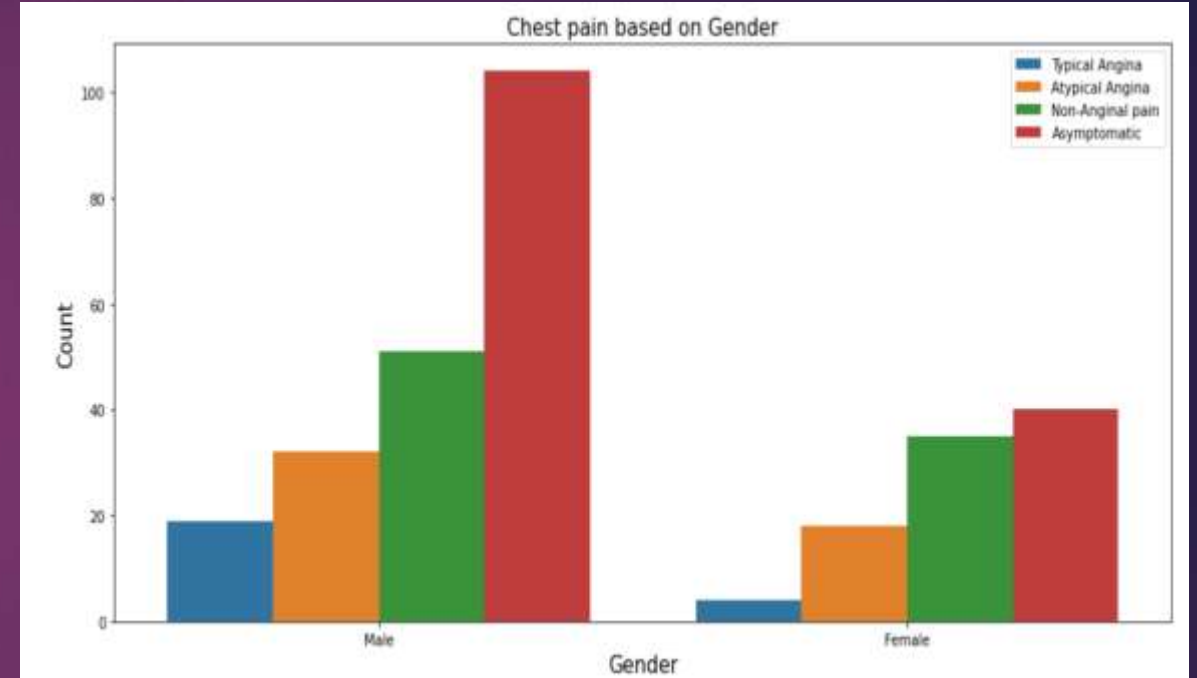


- Males are more affected by heart disease than the females

# Chest pain experienced by Heart patients



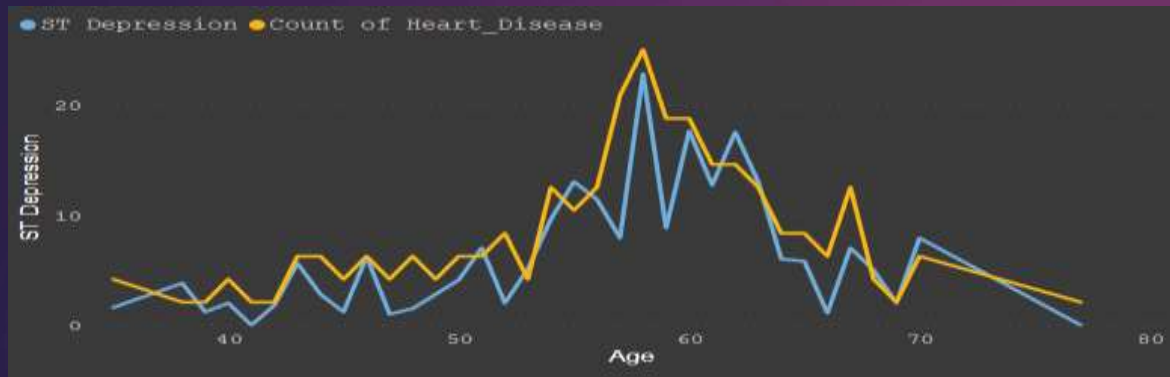
- It is very interesting to see that more heart patients don't have any symptoms of chest pain (asymptomatic).



- Both males and females have more asymptomatic chest pain followed by Atypical angina chest pain.

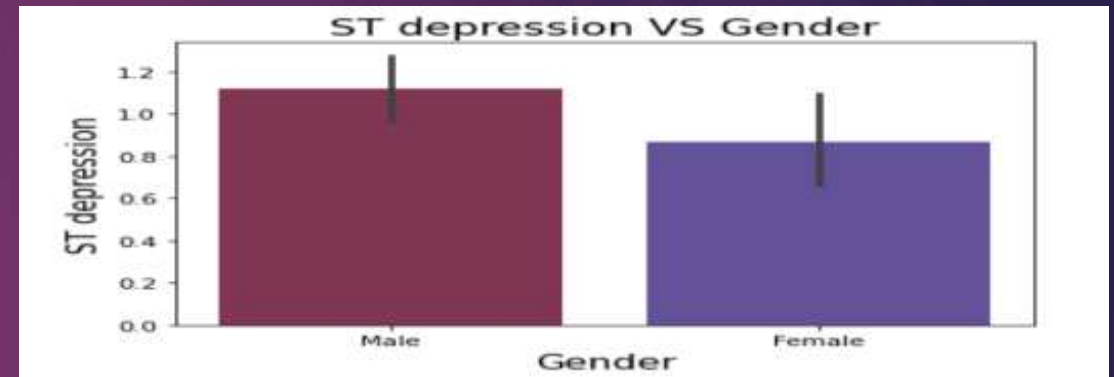
# ST Depression and how it is related to Heart disease

ST depression refers to a finding on an electrocardiogram, wherein the trace in the ST segment is abnormally low below the baseline.



We could see that ST Depression increases after the age 52-55 and reaches it's highest at the age of 57-60 then decreases.

There is a positive correlation between ST Depression and Heart disease as ST depression increases then chances of Heart disease also increase.



We could see that Males are more getting ST depression than the Females.

## Other Observations:

- Avg of Max Heart rate of the patients who is having heart disease (139.26mm/Hg) is less than those who don't have Heart disease.
- ECG Measurement is higher (163) in patients who have heart disease than who don't have heart disease.
- Average cholesterol is higher (251.47) in Heart patients those who have heart disease than those who don't have.

## Gender

Female

Male

Avg Cholesterol

Avg Heartbeat

246.69 149.61

## Heart Disease Diagnosis Analysis



## Heart Disease

Absence

Presence

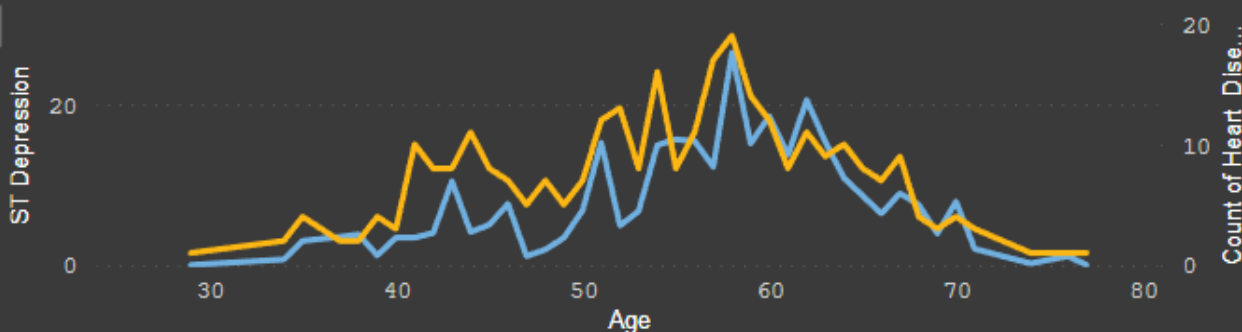
Total Individuals

Avg BP

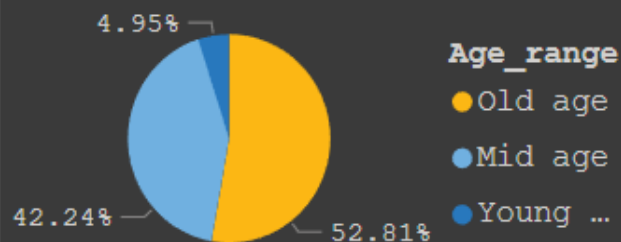
303 131.69

## ST Depression vs Age vs Heart Disease

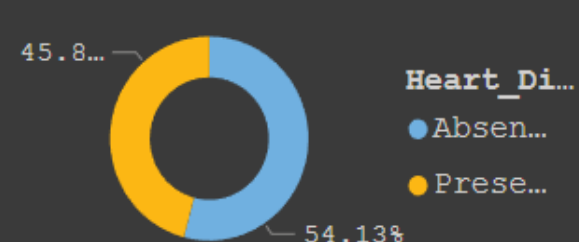
ST Depression Count of Heart\_Disease



## Age category vs Total population

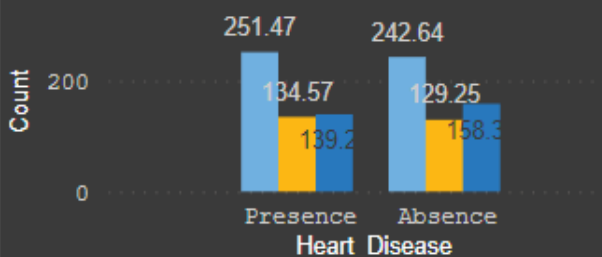


## Heart Disease vs Total population



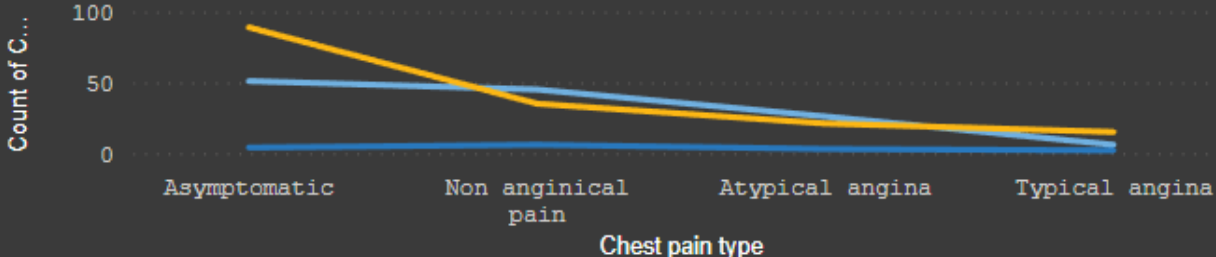
## Chol, BP, Heart Rate vs Heart Disease

Avg Cholesterol Avg Blood pressure Avg Heart Rate



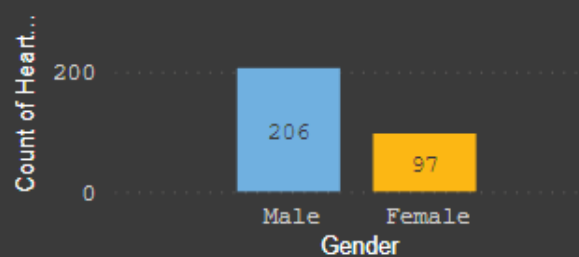
## Age Category vs Chest pain type

Age\_range Mid age Old age Young age

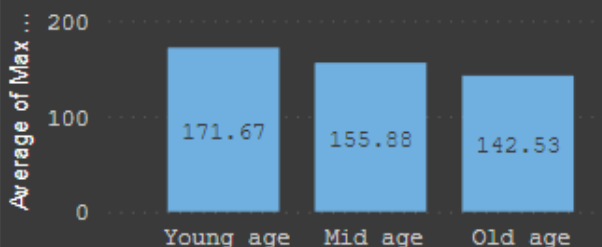


## Gender vs Total population

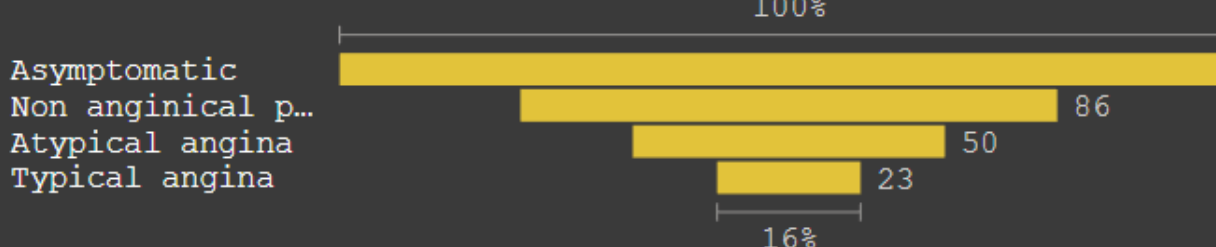
Gender Male Female



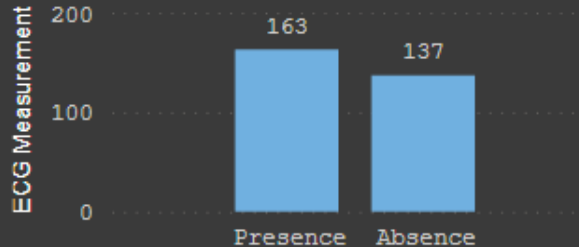
## Average Heart Rate vs Age Range



## Heart Disease vs Chest pain



## ECG vs Heart Disease



# KEY PERFORMANCE INDICATOR (KPI)

1. Age distribution based on Heart disease
2. Gender distribution based on Heart disease
3. ST depression and age based on Heart disease
4. Age category vs Chest pain type
5. Heart disease (Presence/Absence) in Total population
6. Chest pain type vs Heart disease
7. Cholesterol, Blood pressure, Heart rate based on Heart disease patients
8. ECG Measurement vs Heart disease



# CONCLUSION

- 45.87% People suffering from heart disease.
- Elderly Aged Men are more (50 to 60 Years) and Females are more in 55 to 65 Years Category
- Males are more prone to heart disease.
- Elderly Aged People are more prone to heart disease.
- People having asymptomatic chest pain have a higher chance of heart disease.
- High number of cholesterol level in people having heart disease.
- Blood Pressure increases between age of 50 to 60 and somehow continue till 70.
- Cholesterol and maximum heart rate Increasing in the age group of 50-60.
- ST depression mostly increases between the age group of 30-40.



# Q & A

**Q1) What's the source of data?**

**Ans)** The Dataset was taken from iNeuron's Provided Project Description Document.  
<https://drive.google.com/drive/folders/165Pjmf9W9PGy0rZjHEA22LW0Lt3Y-Q8>

**Q2) What was the type of data?**

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

**Q 3) What's the complete flow you followed in this Project?**

**Ans)** Refer slide 5<sup>th</sup> 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.

**Q 6) What were the libraries that you used in Python?**

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



THANK YOU!