HEART DISEASE DIAGNOSTIC ANALYSIS

AGENDA

OBJECTIVES

PROBLEM STATEMENT

THE PROCESS

KEY PERFORMANCE INDICATOR (KPI)

INSIGHTS

CONCLUSION

Objectives

The goal of this project is to analyze the heart disease occurrence, based on a combination of features that describes the heart disease.

- Which age group has the highest chances of getting heart disease.

- Relationship between Chest Pain Severity and Heart Disease.

- Do male are more prone to heart disease or female?

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.

The Process
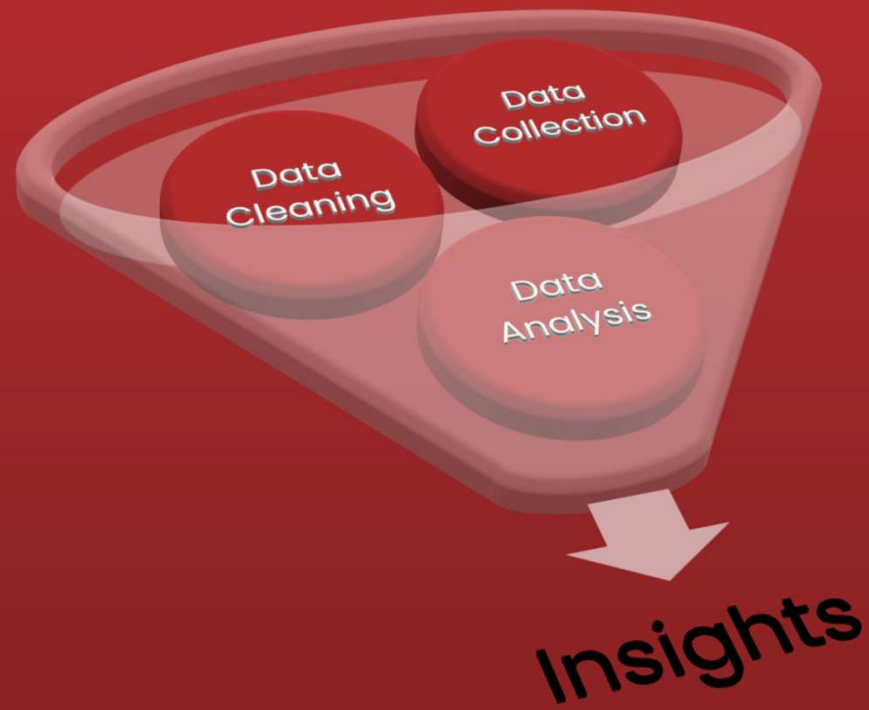
Data Collection

- The Data has been collected in the form of a csv file named "**Heart Disease Data.csv**".

- File contains **Heart Health** data of about 1025 people

- It has a columns: **age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, thal** and **target.**

Data Cleaning

Changed the column names to – Age, Gender, CP, RPB(Diastolic), SC, FBS, RER, MHRA, EIA, Oldpeak, Slope, Flourosopy, Thallium Test, HD.

Age – The person's age in years

Gender – The person's sex (1 – Male, 0 – Female).

CP – Chest Pain Severity (Value 1: typical angina, Value 2: atypical angina, Value 3: non-anginal pain, Value 4: asymptomatic)

RPB(Diastolic) – Resting Blood Pressure

SC – Serum Cholesterol in mg/dl

FBS – Is Fasting Blood Glucose greater than 120mg (1 = true; 0 = false)

Data Cleaning

RER – Resting Electrocardiographic Results (0 = normal, 1 = having ST-T wave abnormality, 2 = showing probable or definite left ventricular hypertrophy by Estes' criteria)

MHRA – Maximum Heart Rate Achieved

EIA – 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)

Flourosopy – Number of major vessels coloured by flourosopy (0-3)

Data Cleaning

Thallium Test – To check how well blood flows to the heart muscles. (0-Normal, 1-fixed defect, 2-reversable defect)

HD – Heart Disease (0 = No, 1 = Yes)

Data Cleaning

Changed value **0** and **1** of **Gender** column to **Male and Female** respectively.

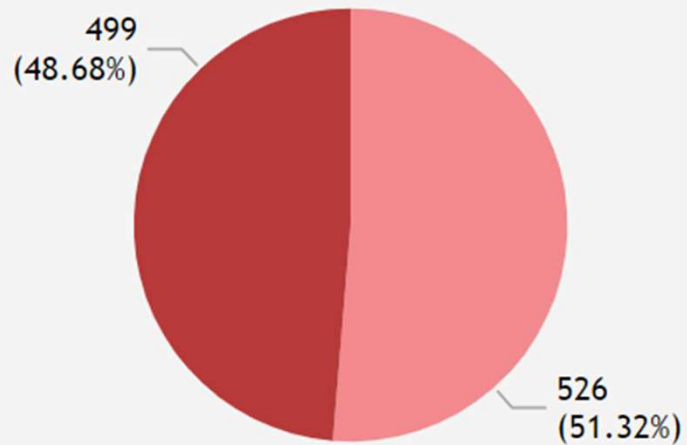Some columns has **outliers**. Replaced those values with the **median of that column** using Python.

Some values of Oldpeak column has value **zero**, which is not possible in practical situation. It is always **greater than zero.** So, replaced those zero values with the **median** of Oldpeak Column.

According to dataset metadata, Thallium Test column should only have values **0,1 and 2.** But some rows has value 3, which is incorrect. So,Its **replaced all the values of 3 with 2**.
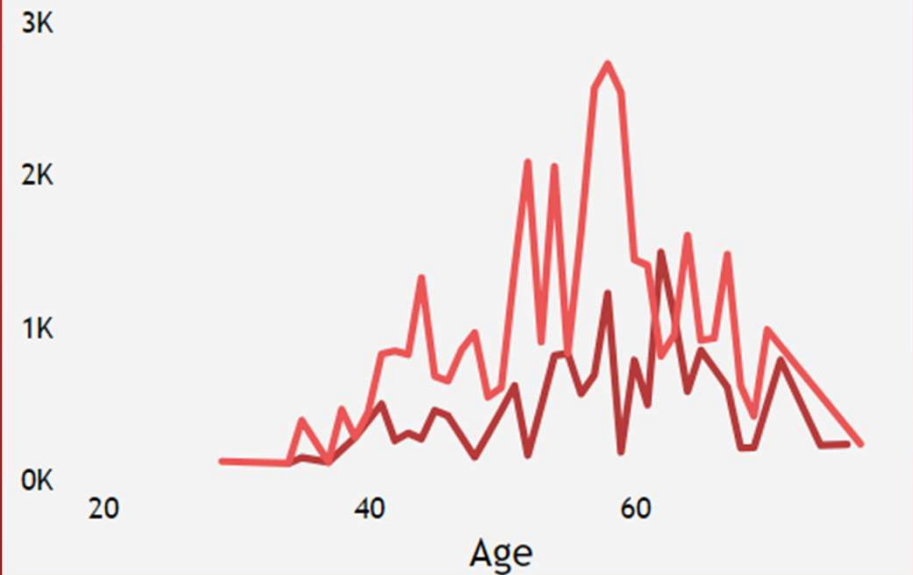
Data Analysis

Heart Disease

HD ● Yes ● No

499
(48.68%)

526
(51.32%)

Age Distribution

Gender ● Female ● Male

3K

2K

1K

0K

20   40   60

Age

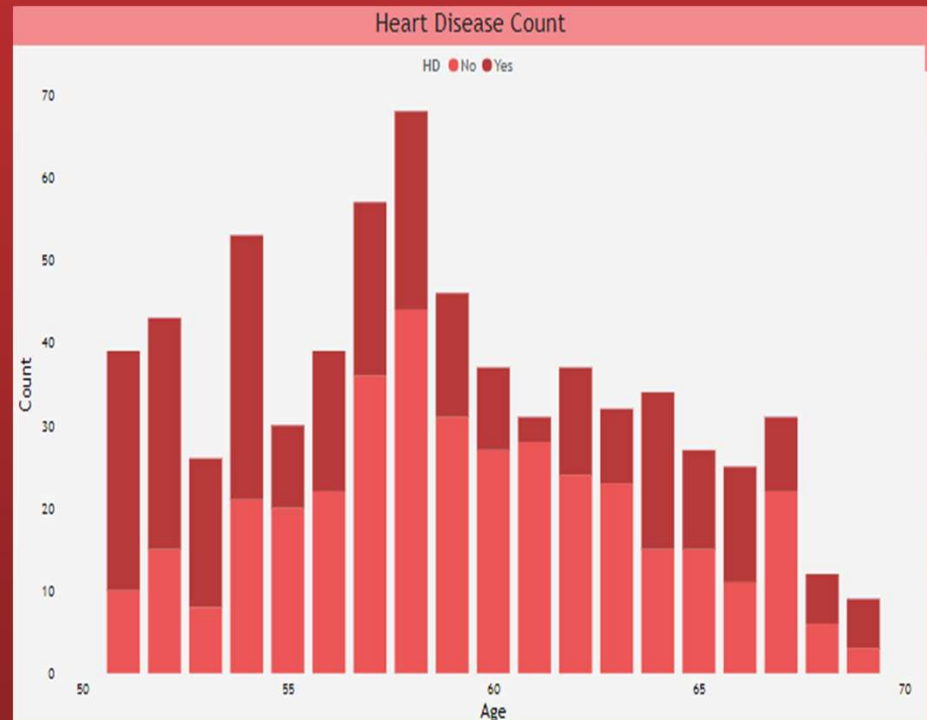✓ **51.32% People suffering from heart disease.**

✓ **More men are from age category >50 and females are from category >55**
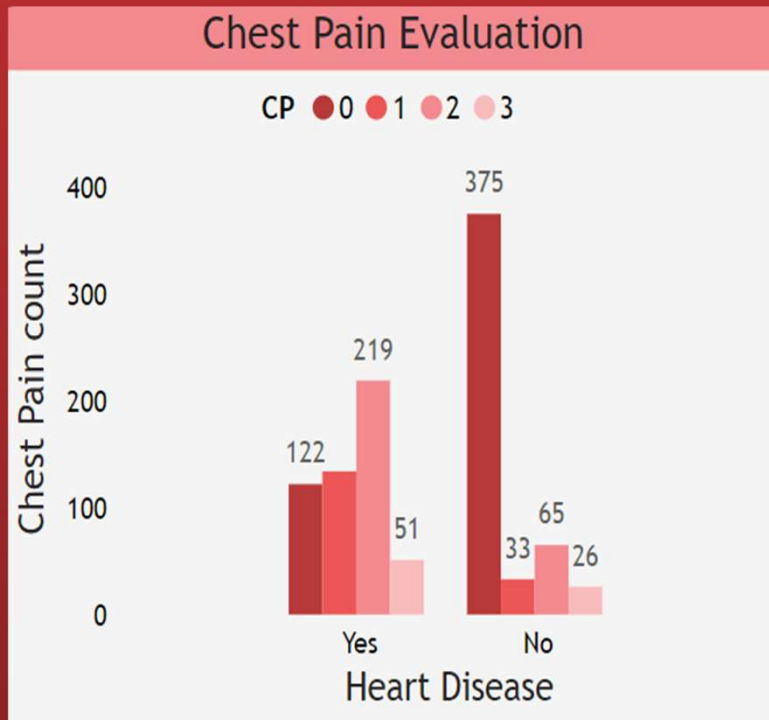
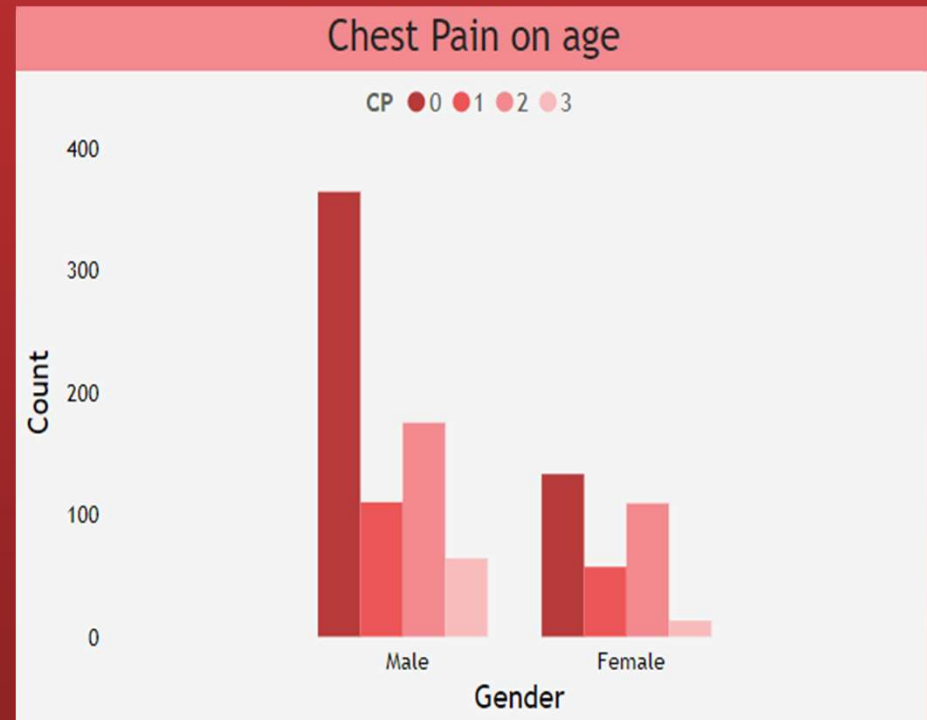# Data Analysis
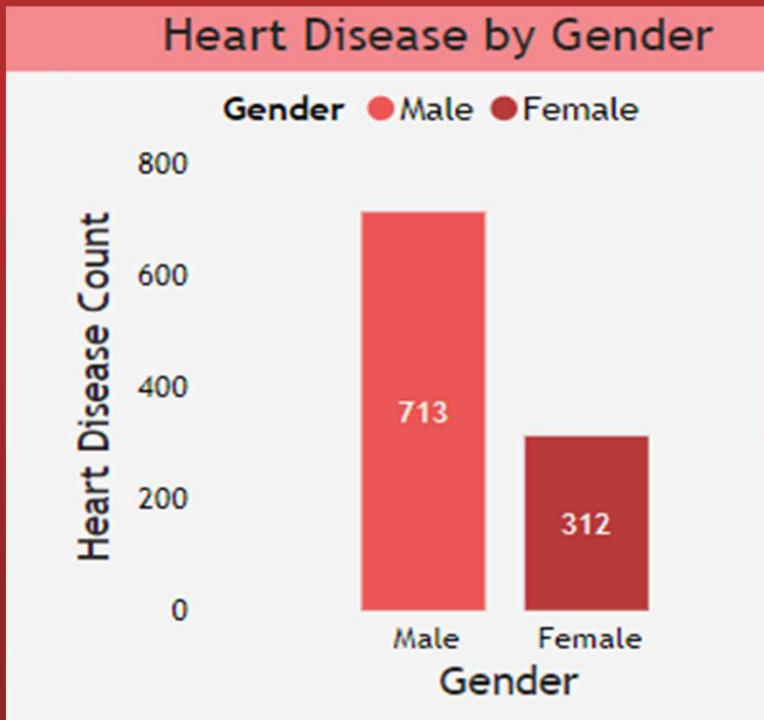
✓ **Males are more prone to Heart Disease**

✓ **More people are from age category >50 and females are from category >60 are prone to HD**

Data Analysis
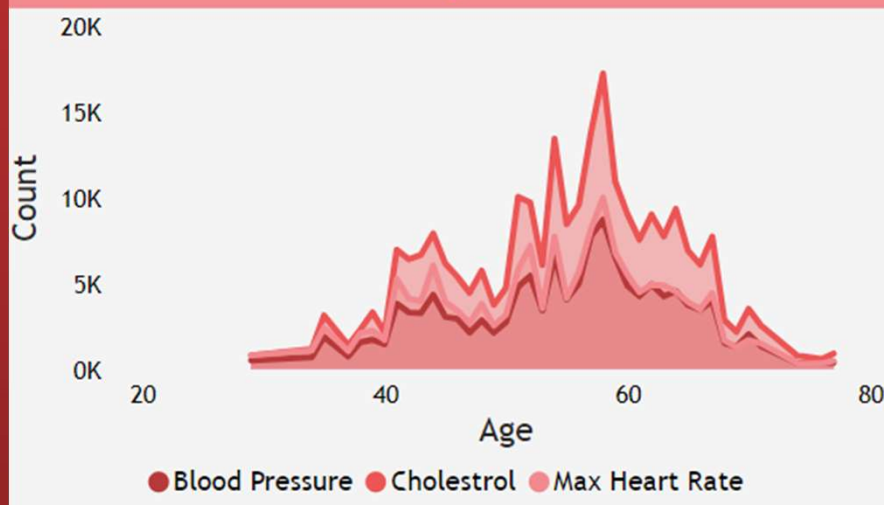

Chest Pain Evaluation

CP ●0 ●1 ●2 ●3

✓ people having asymptomatic chest pain have a higher chance of heart disease.

Chest Pain on age

CP ●0 ●1 ●2 ●3

✓ Higher number of men are suffering from Asymptomatic type of Chest Pain

Data Analysis

Blood Pressure, Cholestrol and Max Heart Rate by Age

✓ Here we can observe that Blood Pressure increases between age of 50 to 60 and somehow continue the pattern till 70. Similarly, Cholesterol and maximum heart rate Increasing in the age group of 50-60.

ST Dipression by OLdpeak and Heart Disease

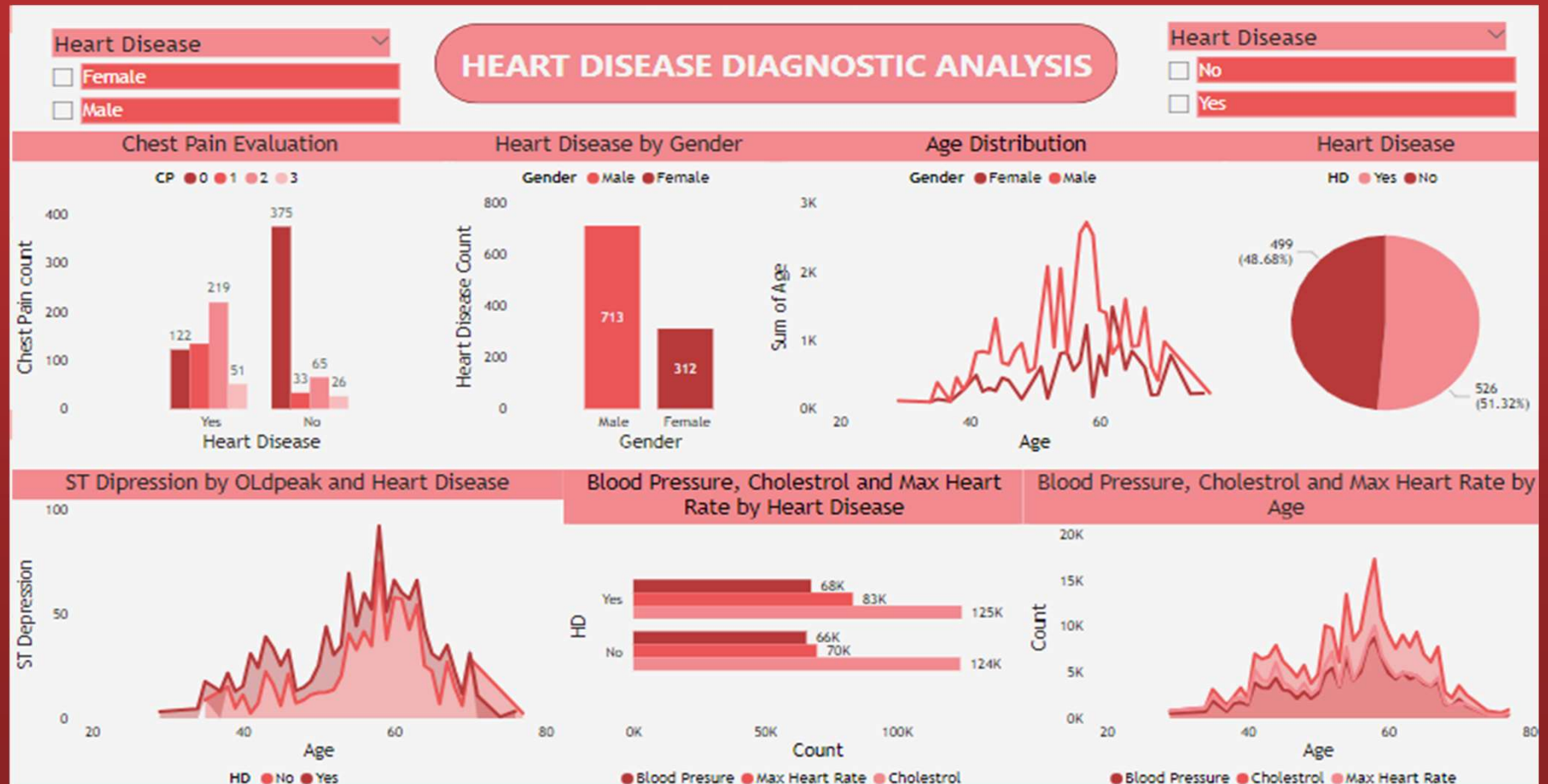✓ we can observe from here that ST depression mostly increases between the age group of 30-40.
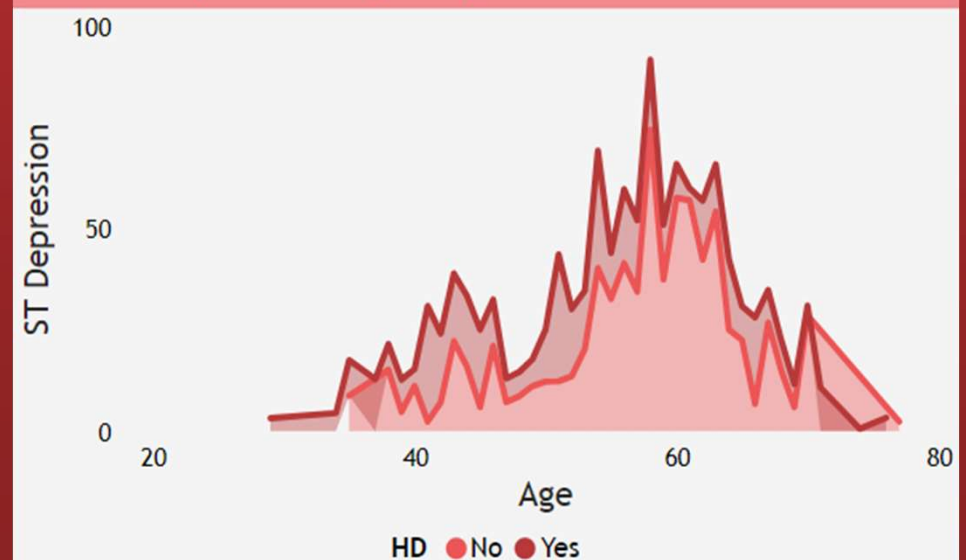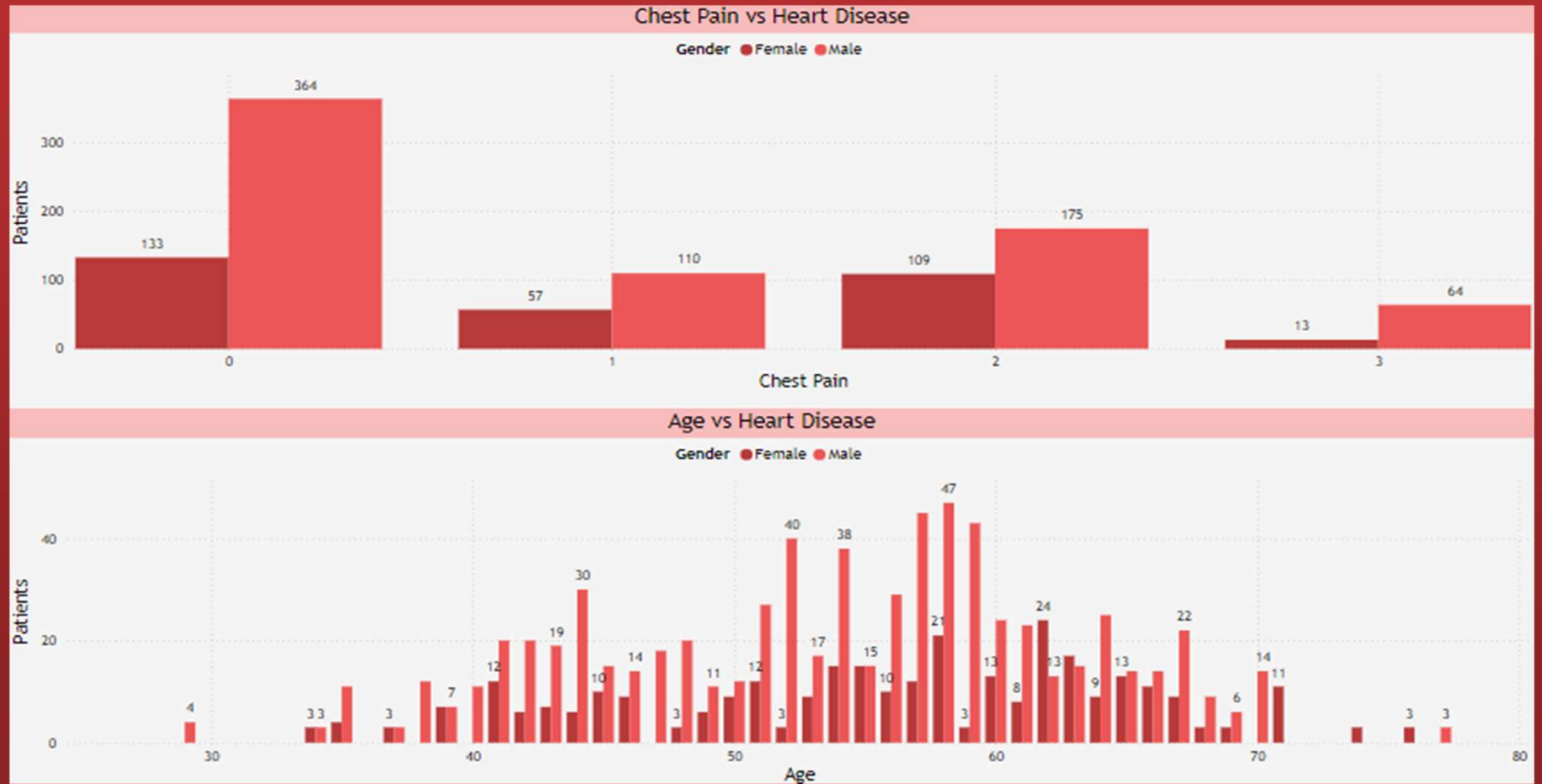
Data Analysis

Data Analysis

Key Performance Indicator (KPI)

➢ Percentage of People Having **Heart Disease**

➢ Age **Distribution** including Gender

➢ Gender Distribution Based on **Heart Disease**

➢ Chest Pain **Experienced** by People Suffering from Heart Disease

➢ Blood Pressure, Cholesterol Level and Maximum Heart Rate of People According to their **Age and Heart Disease Patients**

➢ ST Depression Experienced by People According to their age and heart disease.

Insights

- Total number of patients are **1025** out of which **312** are females and **713** are males.

- The average age of patients is **54**.

- Patients of 48.5% (13% females and 35.5% males) don't have any kind of Chest Pain.

- Only **7.5%** patients (77 patients – 13 females and 64 males) have very severe Chest Pain.

- Most of the patients between the **age of 55 and 65** have the highest Resting Blood Pressure (Diastolic value) which is **greater than 150mg/dl**. This shows they are more prone to High Blood Pressure.

Insights

➤ A total of **853** patients (267 females and 586 males), approx. **83%** patients have Serum Cholesterol(SC) Value **greater than 200mg/dl**, which increases the risk of Cardiovascular Diseases.

➤ Only **15%** patients (42 females and 111 males) have Fasting Blood Sugar levels **greater than 120mg/dl.** This indicates that they are either prediabetics or have Type-2 Diabetes.

➤ Only **1.5%** patients (11 females and 4 males) have Resting Electrocardiographic value of 2, which indicates **signs of Heart Attack.**

➤ **48.5%** patients don't have any kind of cardiac abnormalities.

➤ **Male patients** have good Resting Electrocardiographic Results as compared to female patients.

Insights

➢**34%** of the patients (74 females and 271 males) have Exercise Induced Angina which means **they feel chest pain due to the insufficient blood flow** to the heart muscle during physical exercise.

➢**390** patients have Oldpeak value of **0.8**, which is not severe but have moderate level of **ischemia** (reduced blood flow to the heart).

➢**47%** patients have normal slope value of 1 as compared to 46% patients who have slope value of 2. But **7%** patients have a slope value of 0, which shows more **severe ischemia or coronary artery diseases.**

➢**18** patients (**all male**) have flourosopy level of 4, which is very **critical** and requires immediate action. **56%** of the patients have flourosopy value of 0, which indicate their heart works fine.

Insights

➢Around **93%** patients have Thallium Test value of **2**, which shows **decreased blood flow to the heart or even a scar tissue** in that region.

➢Approximately **21%** patients (105 females and 114 males) having Chest pain Severity of **2**, got diagnosed with Heart Disease. From the chart also we can see that **males are more prone to heart diseases.**

➢**51%** (**almost half**) of the patients are diagnosed with Heart Disease.

➢Patients within the age range of **40-45 and 50-60** have shown the signs of Heart Disease.

Conclusion

➢ The **51.32%** People suffering from heart disease.

➢ Men are more in **50 to 60 Years** and Females are more in **55 to 65 Years** Category.

➢ The **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.

Conclusion

➢ 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.**