

Heart Health Data Analysis

Objectives

- 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?

The Process

- Data Collection
- Data Cleaning
- Data Analysis
- Insights

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.

Data Cleaning

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

CP – Chest Pain Severity

RPB(Diastolic) – Resting Blood Pressure

SC – Serum Cholesterol in mg/dl

FBS – Is Fasting Blood Glucose greater than 120mg

RER – Resting Electrocardiographic Results

MHRA – Maximum Heart Rate Achieved

EIA – Exercise Induced Angina

Data Cleaning

Oldpeak – ST depression induced by exercise relative to rest

Slope – the slope of the peak exercise ST segment

Flourosopy – Number of major vessels coloured by flourosopy

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

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.

Data Cleaning

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, I **replaced all the values of 3 with 2**.

Data Analysis

1025

Number of Patients

54

Avg Age

312

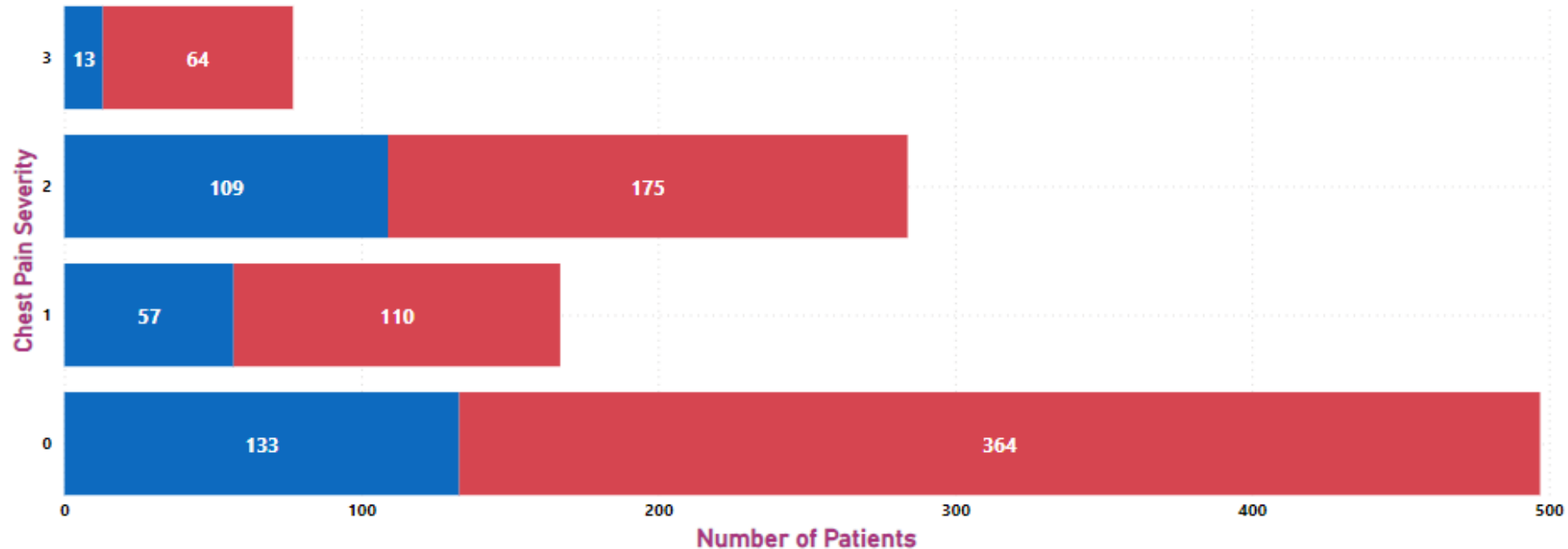
No. of Females

713

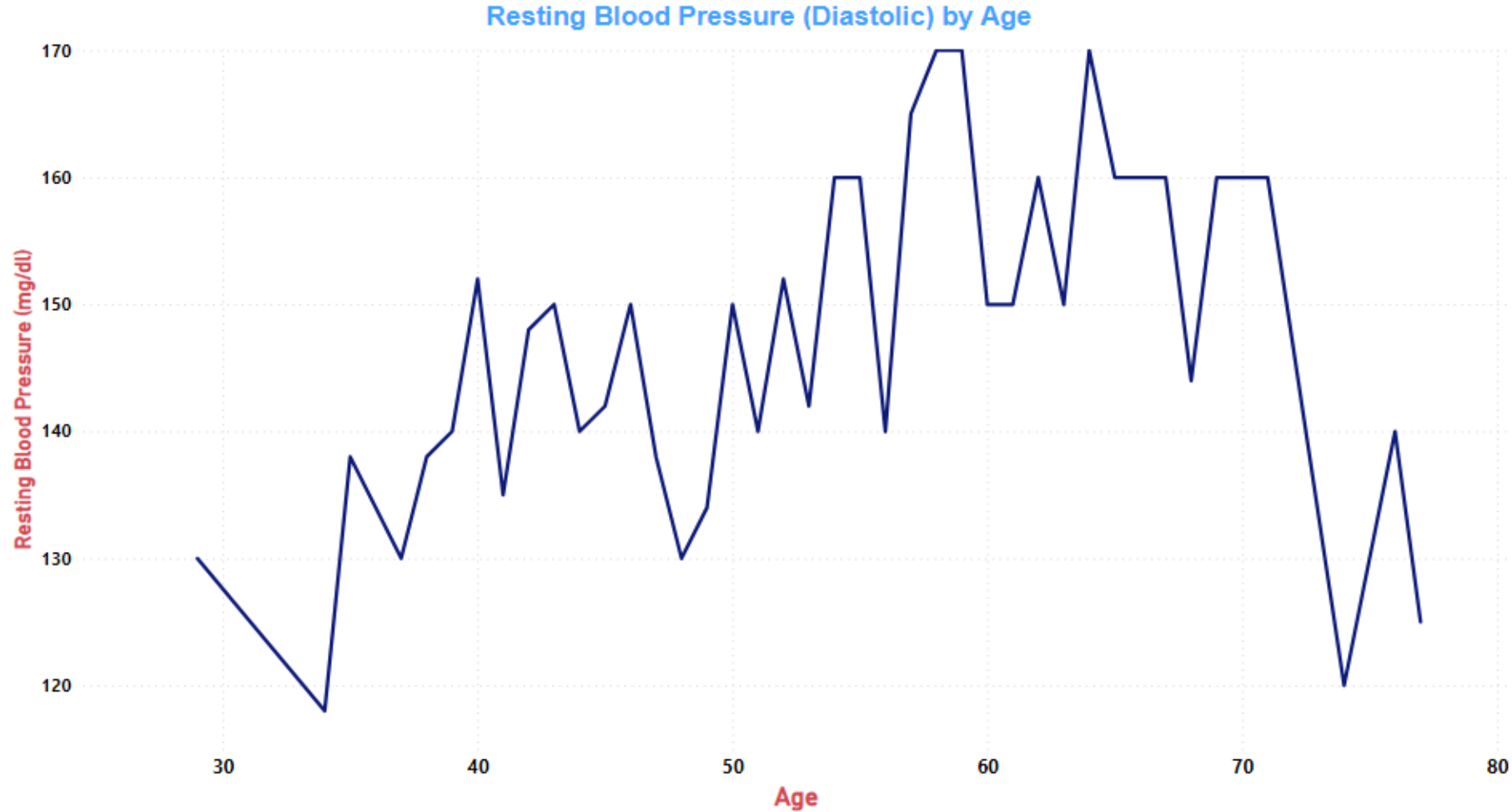
No. of Males

Patients Distribution by Chest Pain Severity

Gender ● Female ● Male



Data Analysis



Data Analysis

853

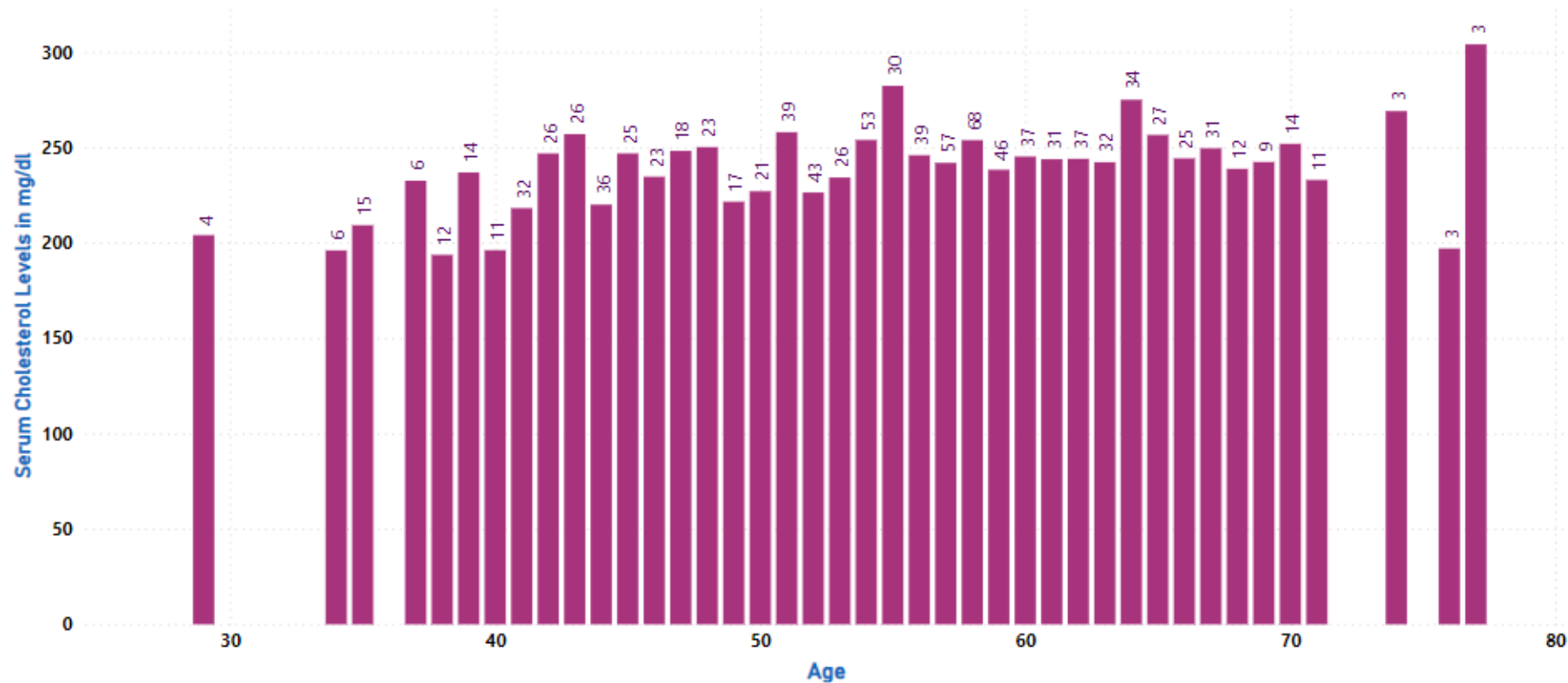
Number of Patients having SC Value > 200mg/dl

Select all

Female

Male

Serum Cholesterol Distribution by Age



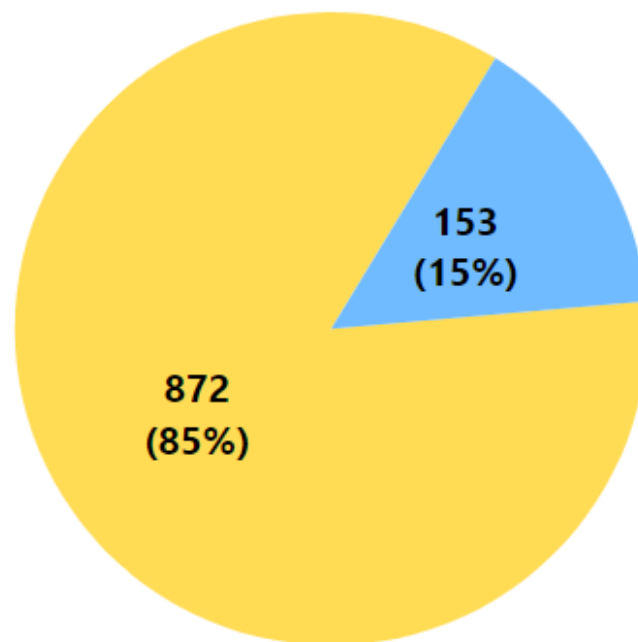
Data Analysis

Select all

Female

Male

Patients Distribution by Fasting Blood Sugar (>120ml/dl)

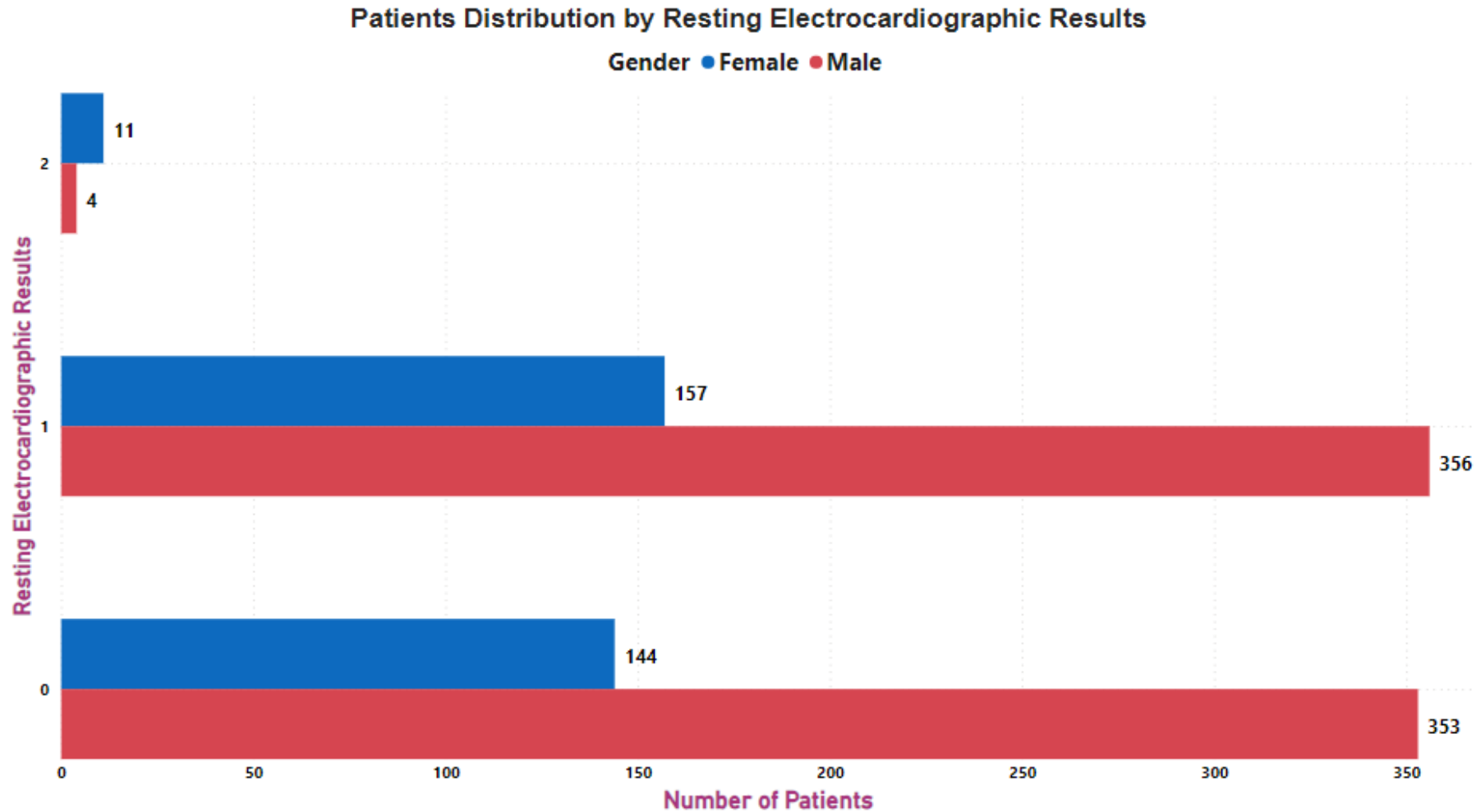


Fasting Blood Sugar > 120 mg/dl (0-No,1-Yes)

● 0

● 1

Data Analysis



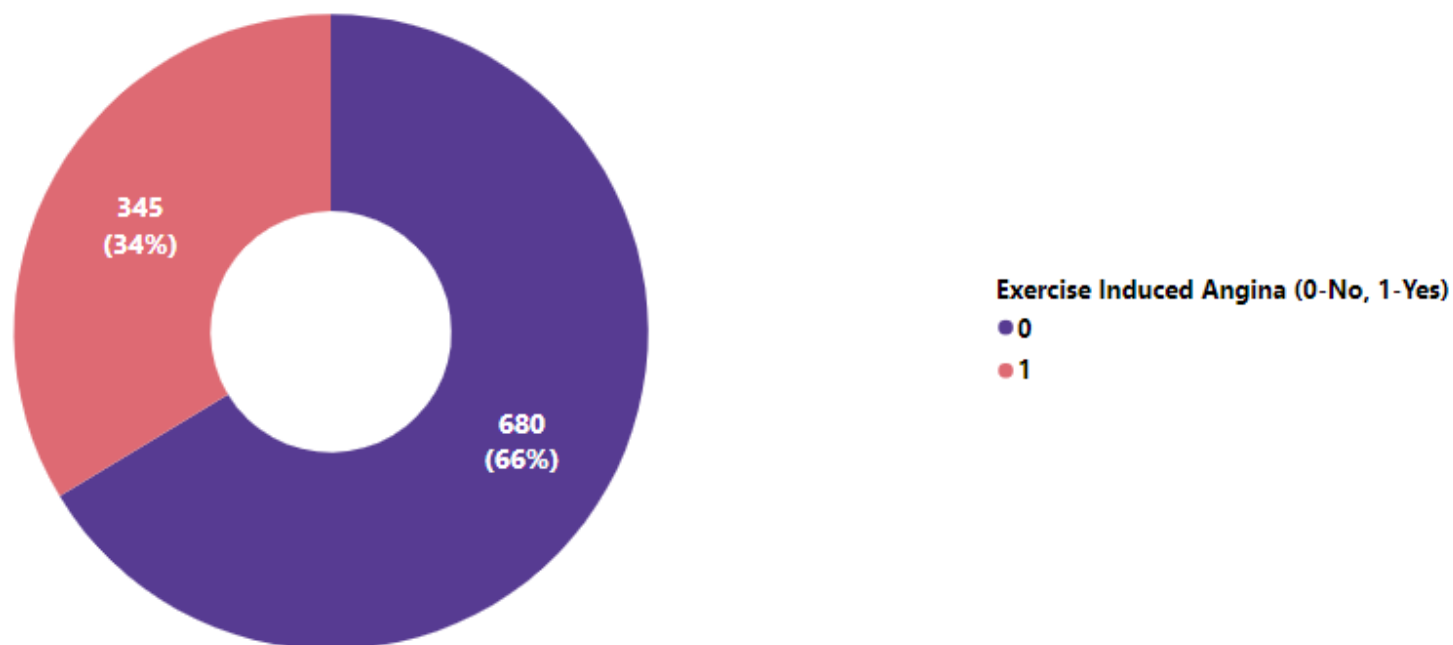
Data Analysis

Select all

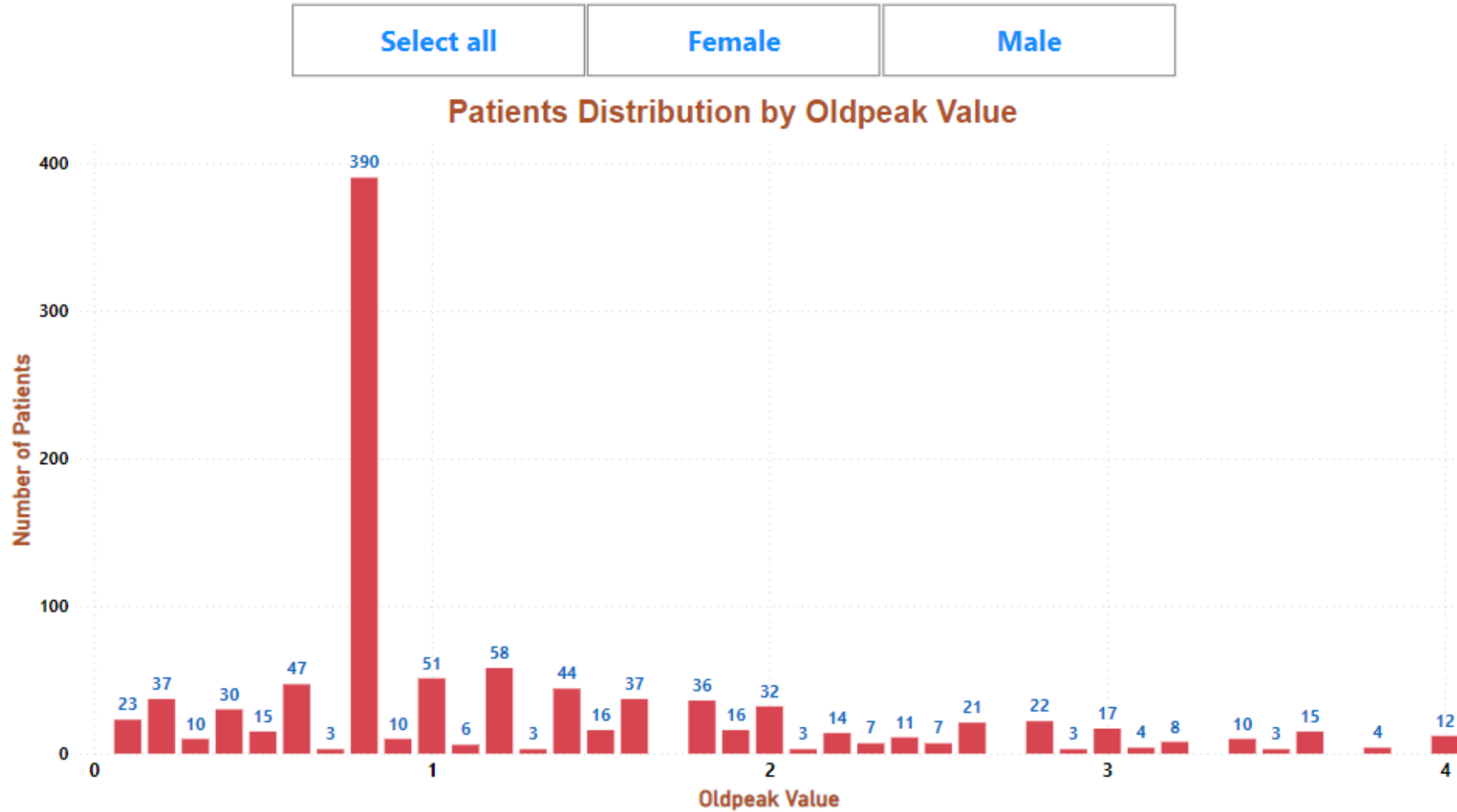
Female

Male

Patients Distribution by Exercise Induced Angina



Data Analysis



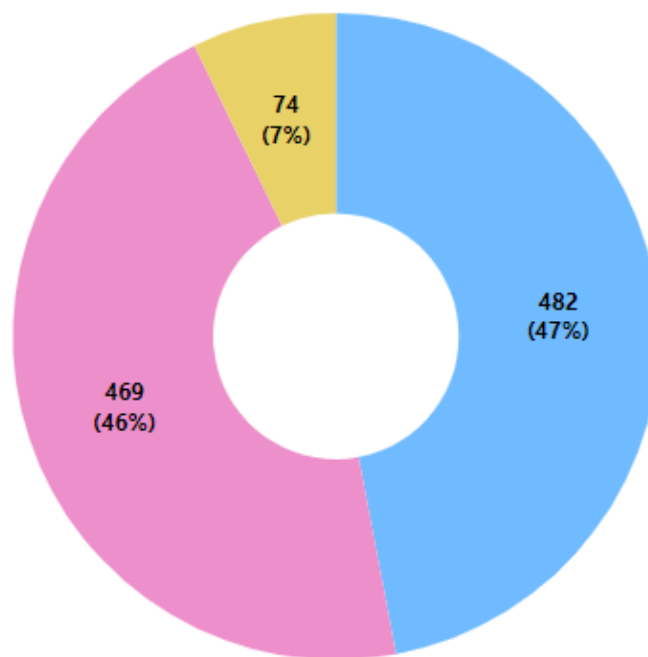
Data Analysis

Select all

Female

Male

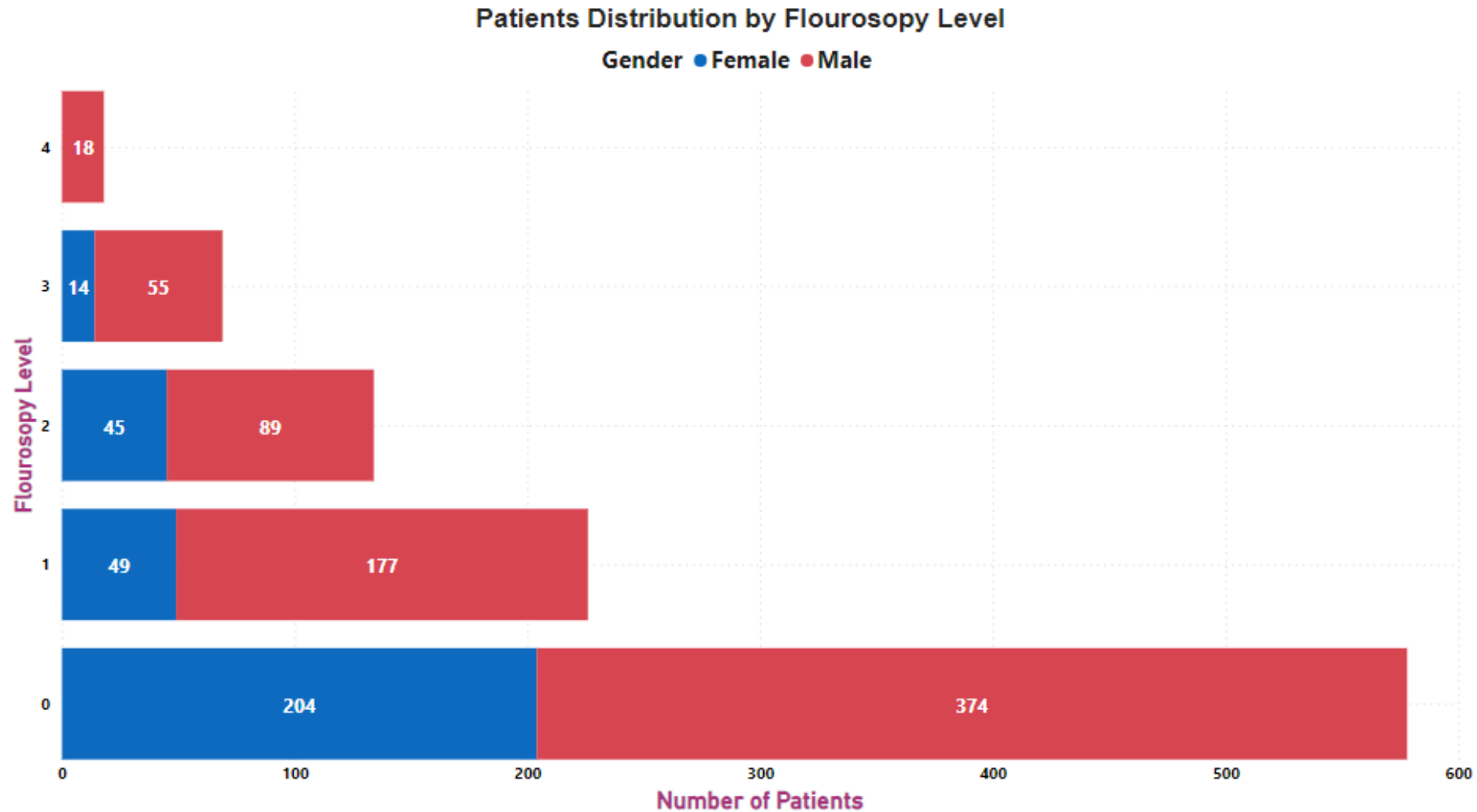
Patients Distribution by Slope



Slope Values

- 1
- 2
- 0

Data Analysis



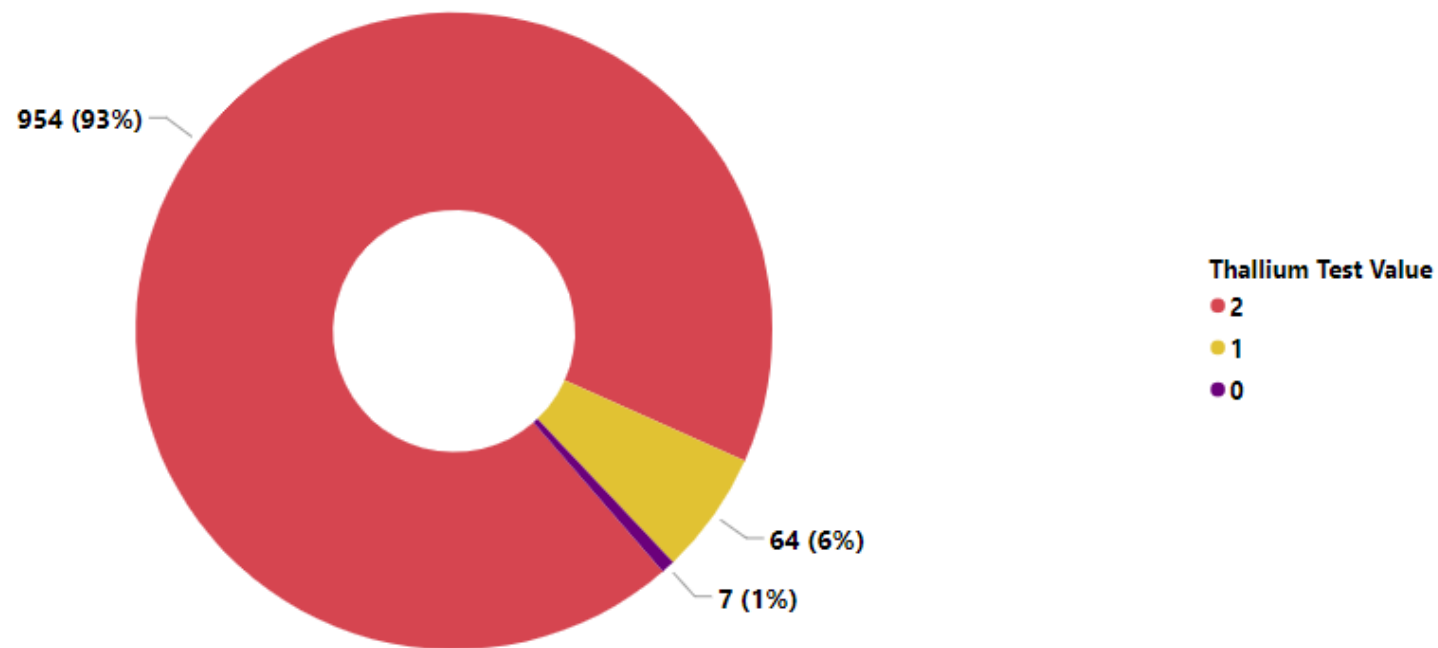
Data Analysis

Select all

Female

Male

Patients Distribution by Thallium Test Value



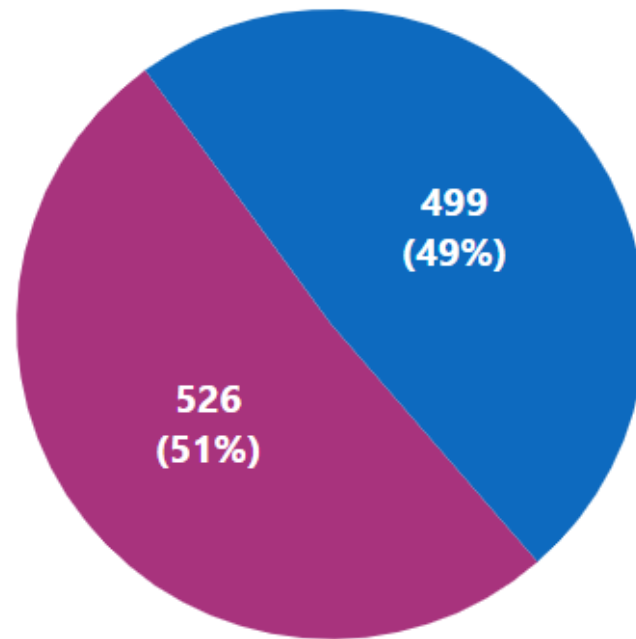
Data Analysis

Select all

Female

Male

Patients Having Heart Disease

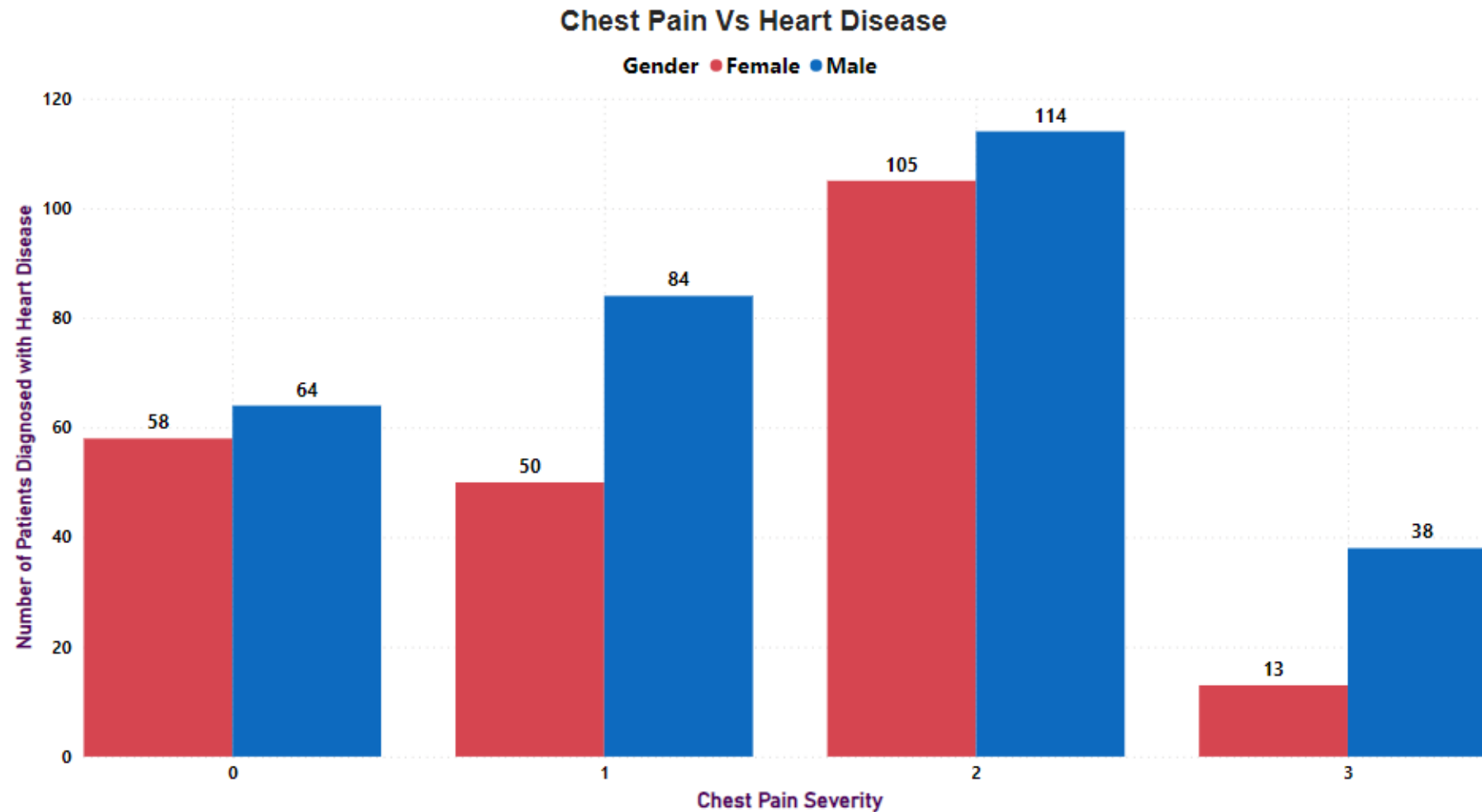


Having Heart Disease (0-No, 1-Yes)

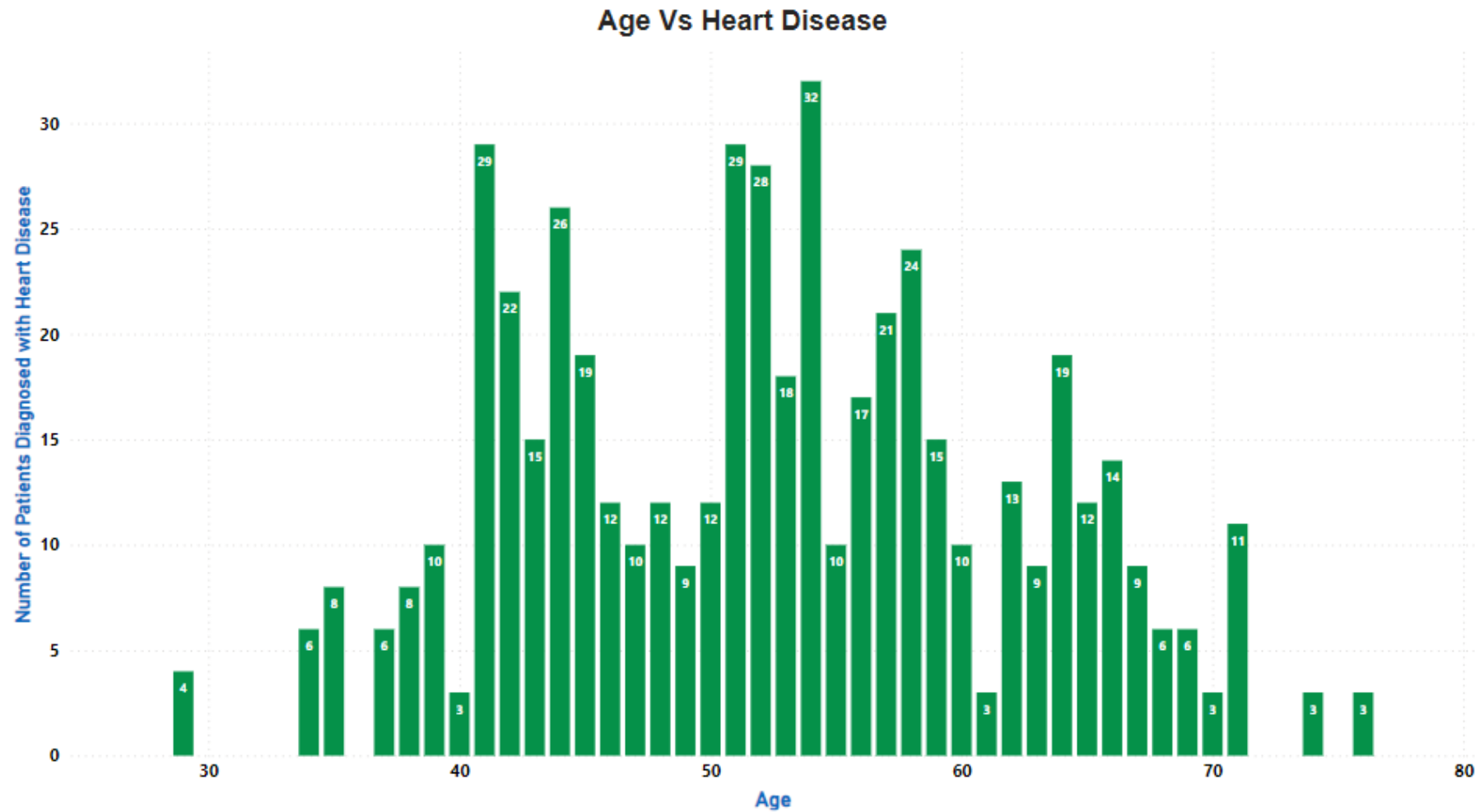
1

0

Data Analysis



Data Analysis



Insights

- Total number of patients are **1025** out of which **312** are females and **713** are males.
- The average age of patients is **54**.
- **48.5%** of patients (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.



Thank you