

**Data Visualization**

**Assignment-2**

**BY**

**Kartheek Prasanna**

**Rushab S.L.**

**Juliana Fernandes**

# INTRODUCTION

#### This database is about the heart diseases and various factors involved in it.

#### LITERATURE REVIEW

Consistently, present day PC based frameworks gather a lot of information utilizing programmed information record frameworks in the medicinal services field where data mining can separate a profitable learning from them. The next section briefly explains heart disease and the application of data mining techniques in treating such diseases.

HEART DISEASE

As the leading cause of death in the world, heart disease, according to WHO, accounts for 3.8 million and 3.4 million deaths in males and females, respectively.

The symptoms and incidence of heart disease differ from one person to another. However, they commonly include chest pain, jaw pain, neck pain, back pain, [stomach disorders](https://www.omicsonline.org/hepatology-gastrointestinal-disorders.php); arms and shoulders pains and shortness of breath. Different heart problems induce different heart diseases including coronary artery disease, heart failure and stroke.

Although heart disease has been identified as the most chronic disease across the world, it is the most preventable one at the same time. A healthy life style (primary prevention) and timely diagnosis (secondary prevention) are two main elements of heart disease control. Conducting regular check-ups (secondary prevention) plays a remarkable role in the diagnosis and early prevention of heart disease complication. Several tests including, chest X-rays, angiography, echocardiography and exercise tolerance test contribute to this important issue. However, these tests are costly and require accurate medical equipment.

All distributed trials allude to utilize a subset of 14 of them**.** The 14 attributes are:



# BUSINESS QUESTION

1.) What is the average cholesterol level of patients depending upon the gender?

2.) What is the maximum heart rate achieved according to the chest pain type?

3.) What is the slope of the ST depression in male and female?

# DISCOVERIES AND INSIGHTS

Here, we have chosen R studio to implement the code and we visualize the interactive graphs through shiny app.

In the Business question we are finding out what is the cholesterol level in the patients depending on the gender. This is a necessary insight because elevated cholesterol itself does not cause any indications, such as significant number of individuals are unconscious that their cholesterol levels are excessively high. **Lowering cholesterol levels that are too high lessens the risk for developing heart disease and reduces the chance of a heart attack or dying of heart disease**.

In the figure 1 we are creating an interactive graph in the shiny app by having the cholesterol levels of men in the ascending order. The X-axis shows the average cholesterol level and the Y-axis depicts the number of patients in the dataset.

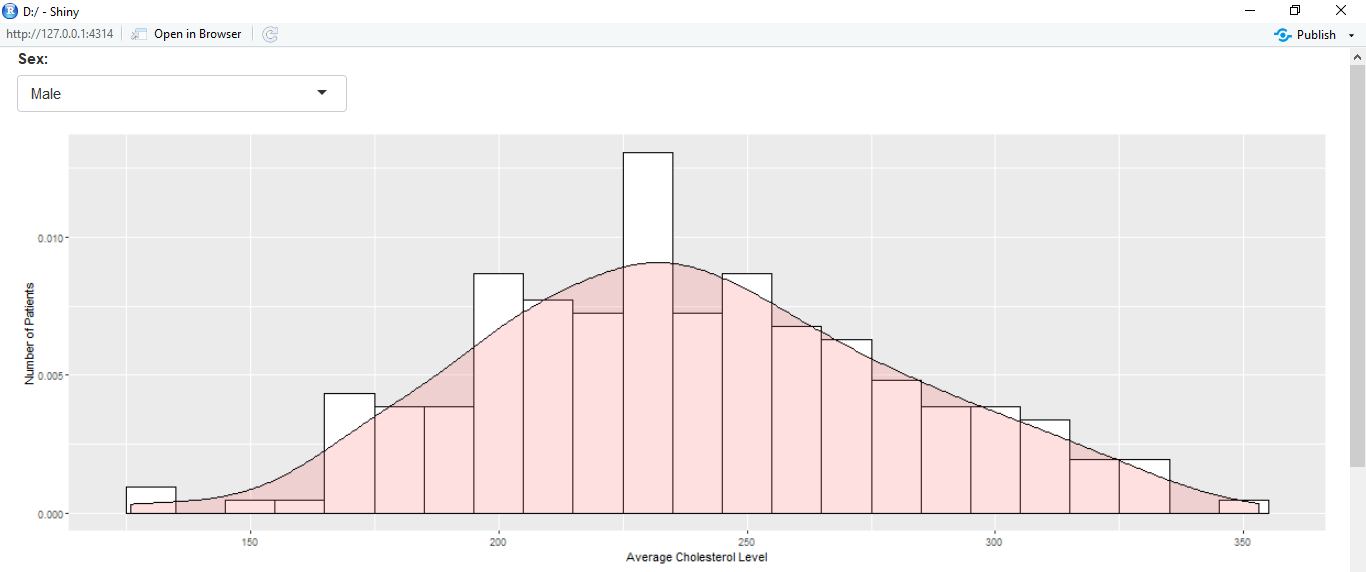


Figure 1.

In the graph the cholesterol density for the male range between 200-250. Complete cholesterol levels under 200 milligrams for each deciliter (mg/dL) are viewed as normal for grown-ups. A perusing somewhere in the range of 200 and 239 mg/dL is viewed as marginal high and a perusing of 240 mg/dL or more is viewed as high.

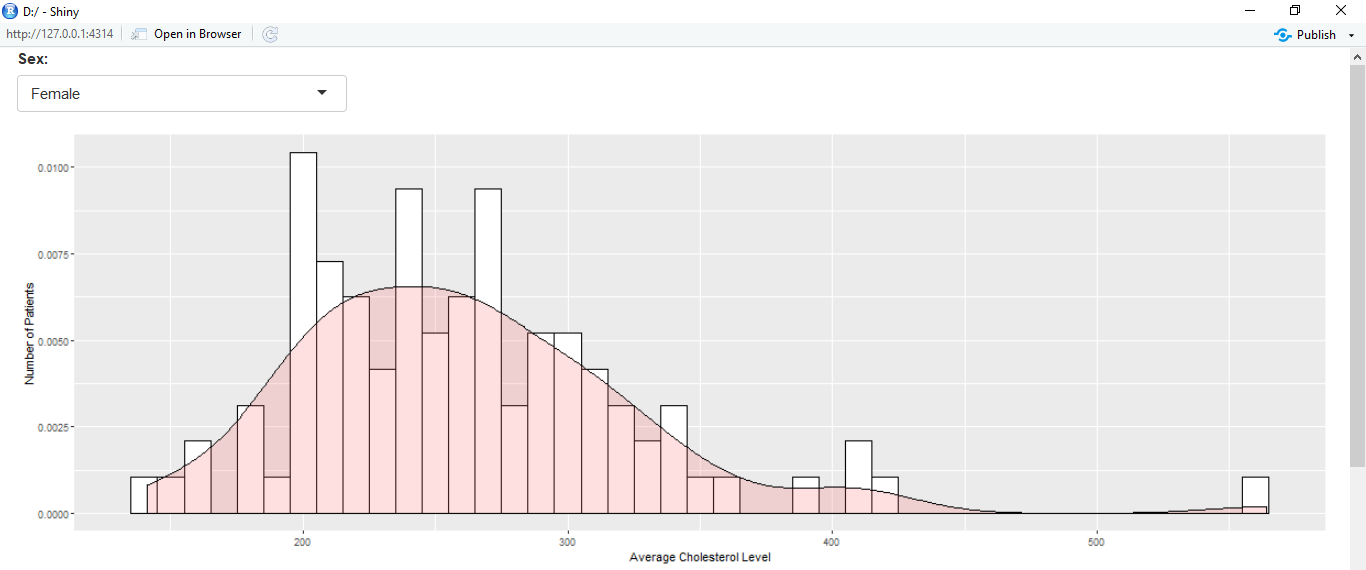


Figure 2

In the figure 2 we are creating an interactive graph in the shiny app by having the cholesterol levels of women in the ascending order. The X-axis shows the average cholesterol level and the Y-axis depicts the number of patients in the dataset.

According to Bell distribution in the histogram, 68.2% of women population have cholesterol nearly 200-300 (mg/dL) and it depicts women are less prone to heart diseases while compared to men.

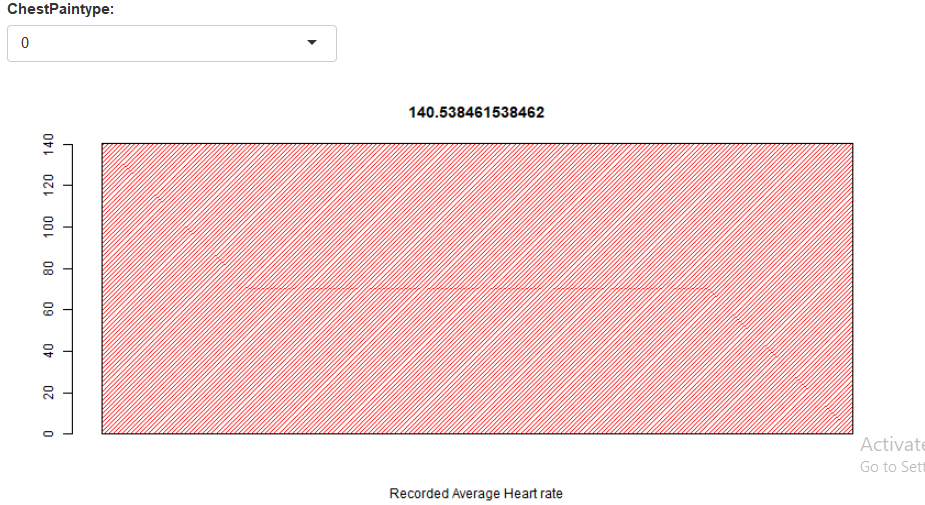
**The second insight in the interactive graph is about the average heart beat recorded according to the chest pain type**. As in the data set we have chest pain type determined as

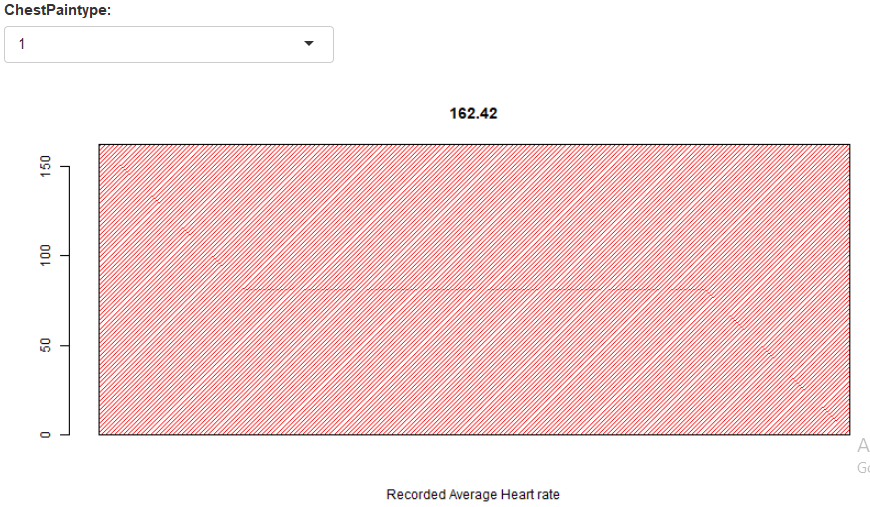
Zero - Typical angina pain,

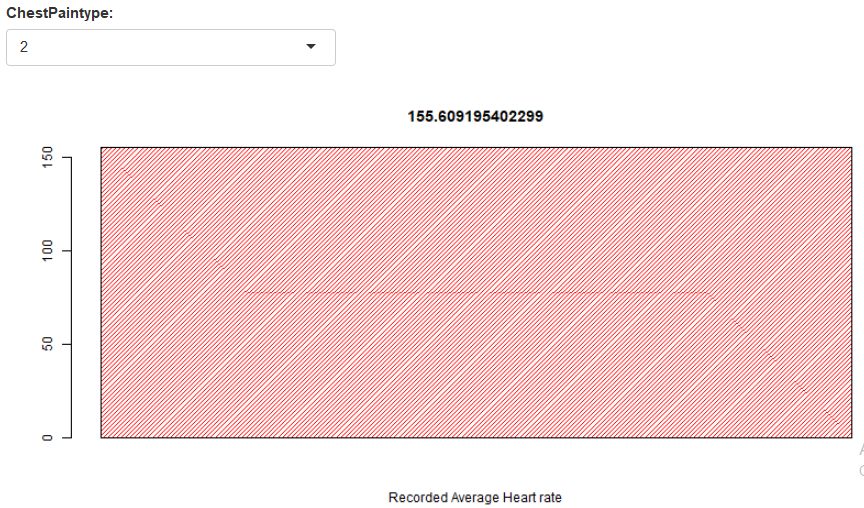
One - Atypical angina pain,

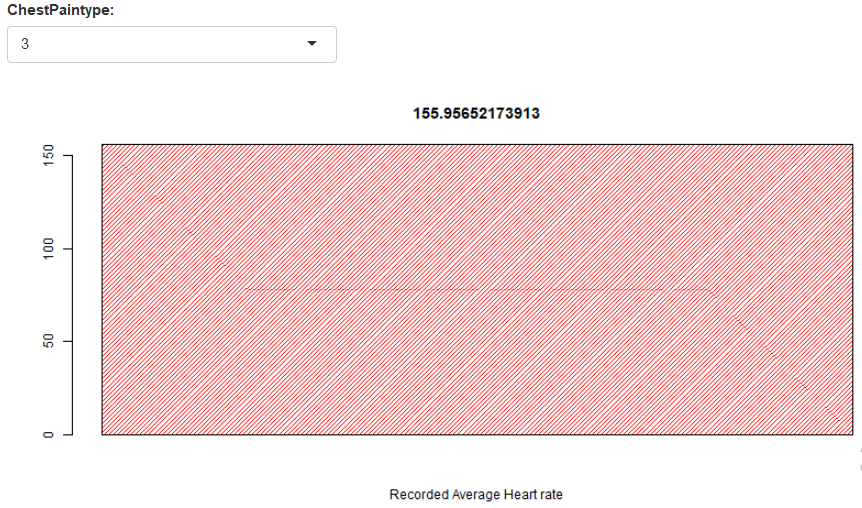
Two - Non angina pain

Three- Asymptomatic. We have the recording as follows.









The above interactive graphs shows the average heart beat rate in the different chest pain types with the highest being in the chest pain type 1 i.e. Atypical angina.

**The third insight in the interactive graph is about slope of the ST depression in male and female.**

ST segment depression may be determined by measuring the vertical distance between the patient's trace and the [isoelectric line](https://en.wikipedia.org/wiki/Isoelectric_line) at a location 2-3 millimeters from the [QRS complex](https://en.wikipedia.org/wiki/QRS_complex" \o "QRS complex).

It is significant if it is more than 1 mm in V5-V6, or 1.5 mm in AVF or III.

In a [cardiac stress test](https://en.wikipedia.org/wiki/Cardiac_stress_test), an ST depression of at least 1 mm after [adenosine](https://en.wikipedia.org/wiki/Adenosine) administration indicates a reversible ischemia, while an exercise stress test requires an ST depression of at least 2 mm to significantly indicate reversible ischemia

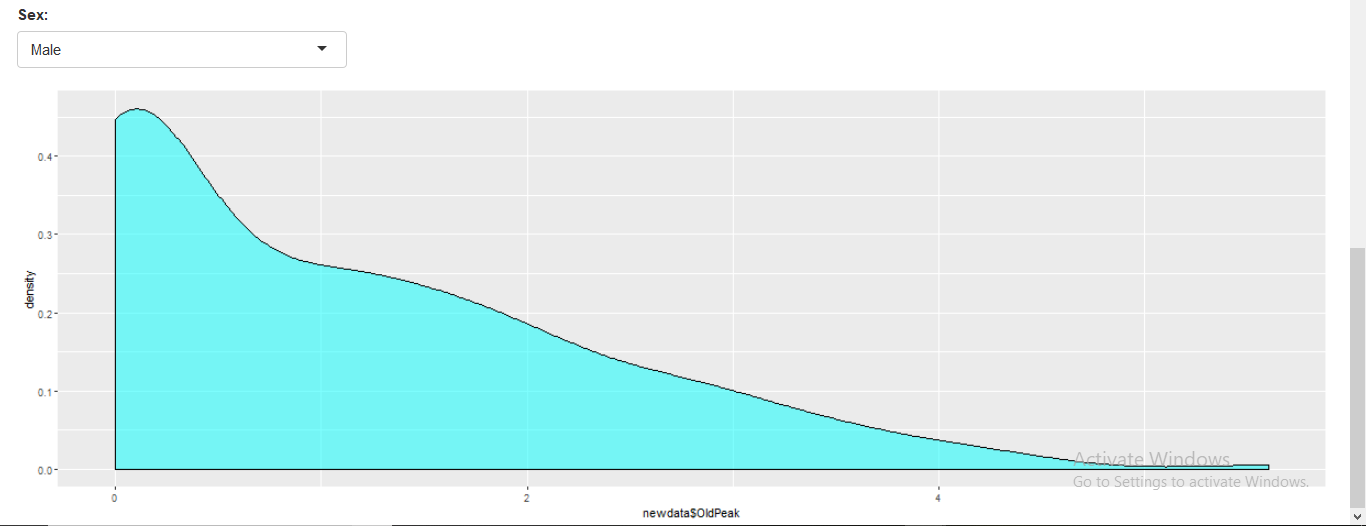
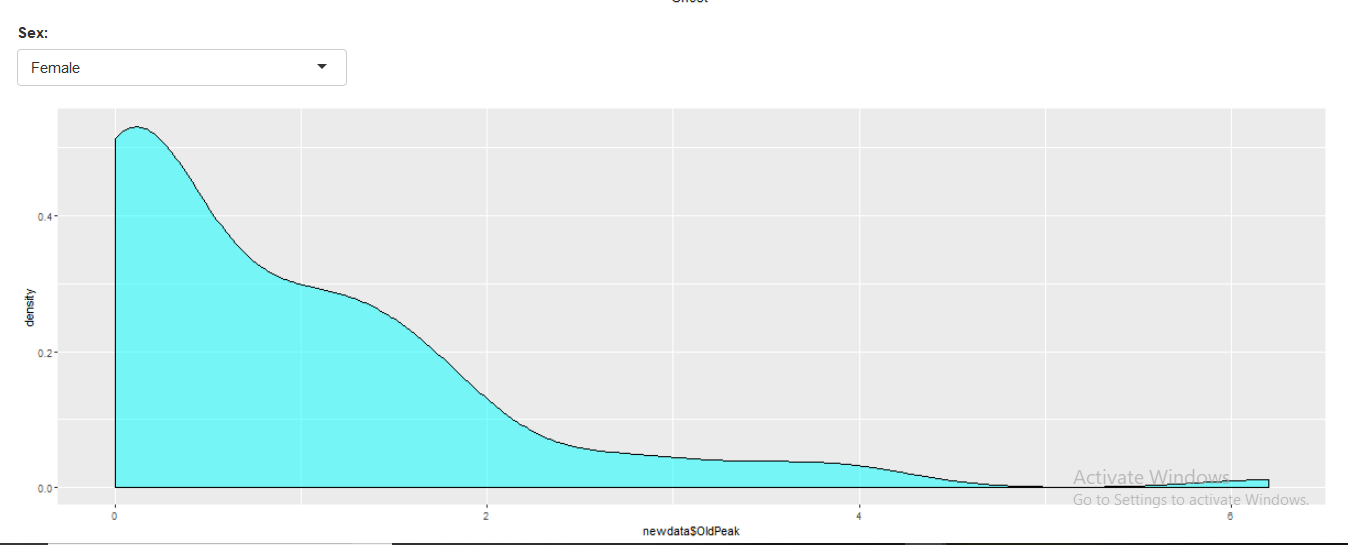


Figure 7

The above graph shows the slope density in the ST depression in the Male patient. The slope density is more in the region of 0-2 mm. Therefore, chances of ischemia i.e. inadequate blood supply to the heart muscles is more.

Figure 8

The above graph shows the slope density in the ST depression in the female patient. The slope density is less in the region of 0-2 mm. Therefore, chances of ischemia i.e. inadequate blood supply to the heart muscles is less.

FUTURE WORKS

Apart from the business questions, we could analyze many other features and compare them between them in the future works to see if there are any patterns. One example is the chart below where we compare the cholesterol and the RestingBPS. It’s a known fact that a high level of cholesterol in your blood, together with other risk factors, can increase your risk of developing coronary heart disease.

We could compare the two features by plotting this chart below, using the library GoogleVis in R.

A screenshot of a social media post

Description automatically generated

Cholesterol levels less than 200 are considered desirable for adults. A reading between 200 and 239 is considered borderline high and a reading of 240 and above is considered high.

We added an option of editing the chart if the users want to. They can do it just by pressing the bottom “Edit this chart if you want!” where they can see the same information in other types of graphs, like below:

A screenshot of a computer

Description automatically generated

CONCLUSION

In the above dataset we saw the various insights about the heart diseases and its types and the various factors affecting it such as cholesterol levels in patients, the average heart beat in different types of chest pain and the ST depression in male and female patients.

By having this data, the doctors can tackle the various causes of heart attack and the health problems associated with it by providing efficient solutions to the patients.