WIREFRAME DOCUMENT

Heart Disease Data Analysis

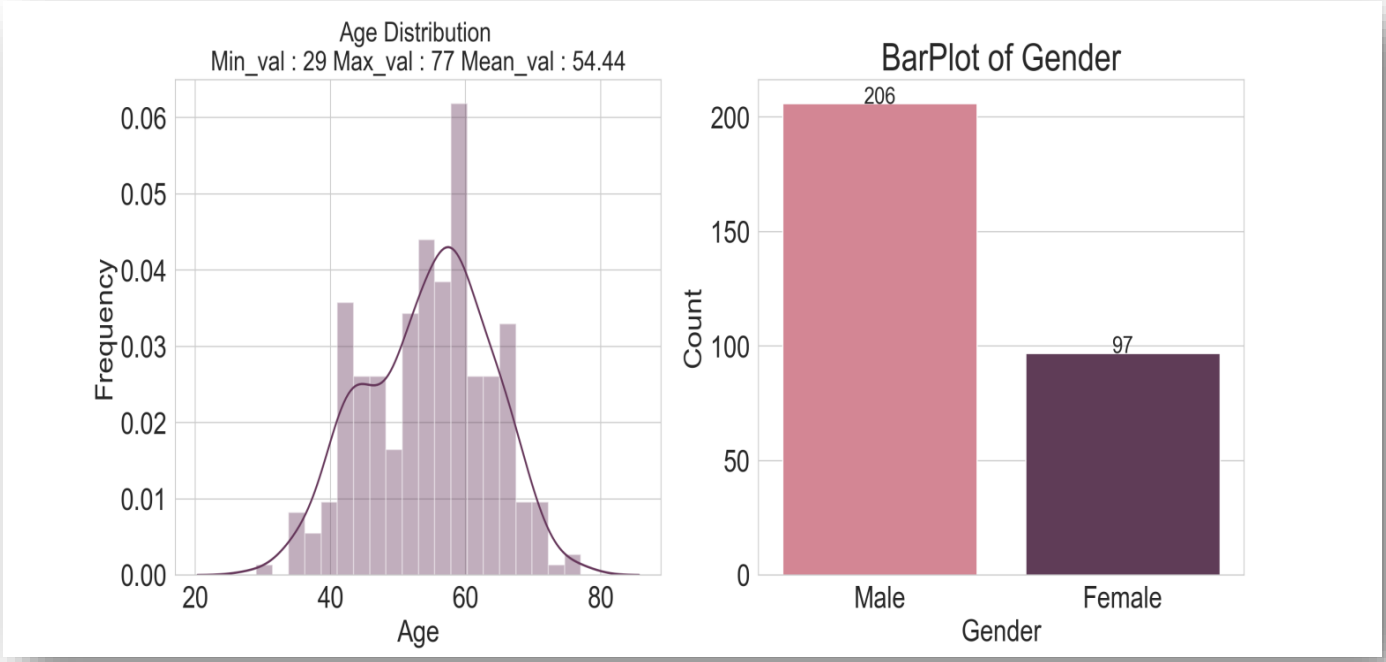
REVISION NUMBER – 1.2
LAST DATE OF REVISION – 16/05/2022

SHIVANSH KAUSHAL, CHIRAG SHARMA

Observations from Explanatory Data Analysis of Heart Disease Diagnostics Dataset

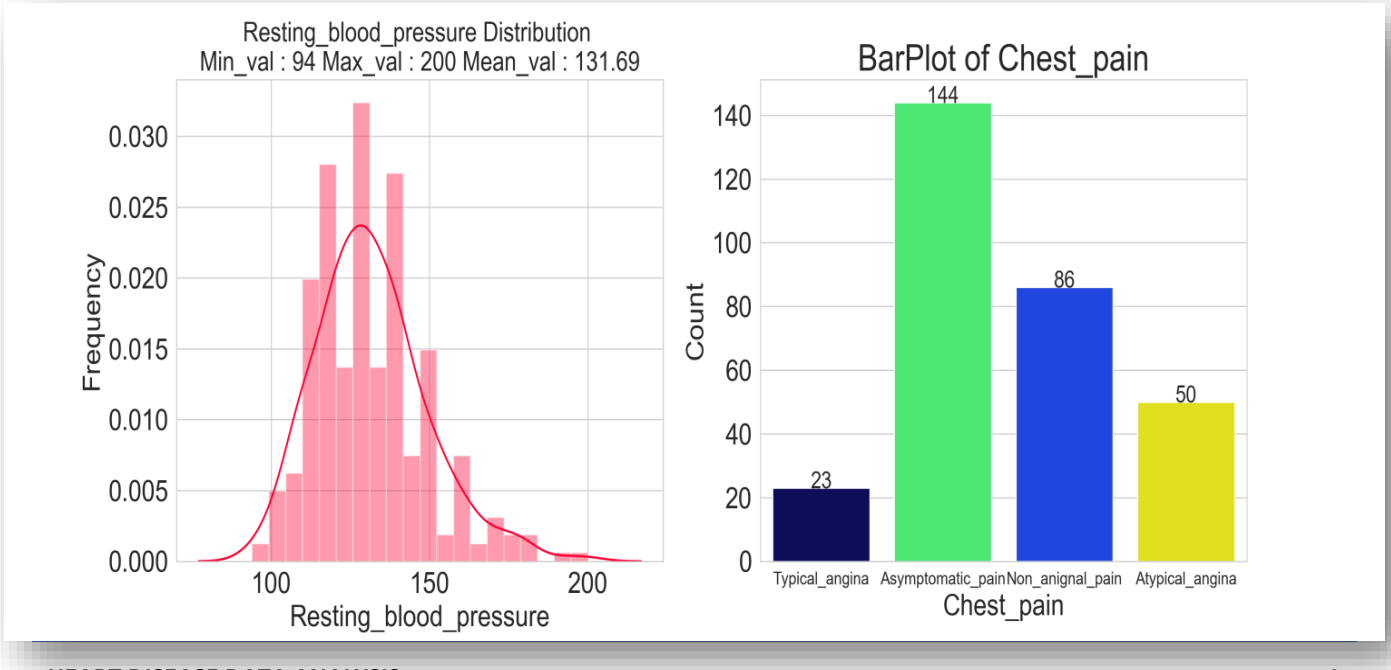
1. How the population is distributed over age and gender?

Younger as well as older people are present in the dataset. Males are more in number than Females.



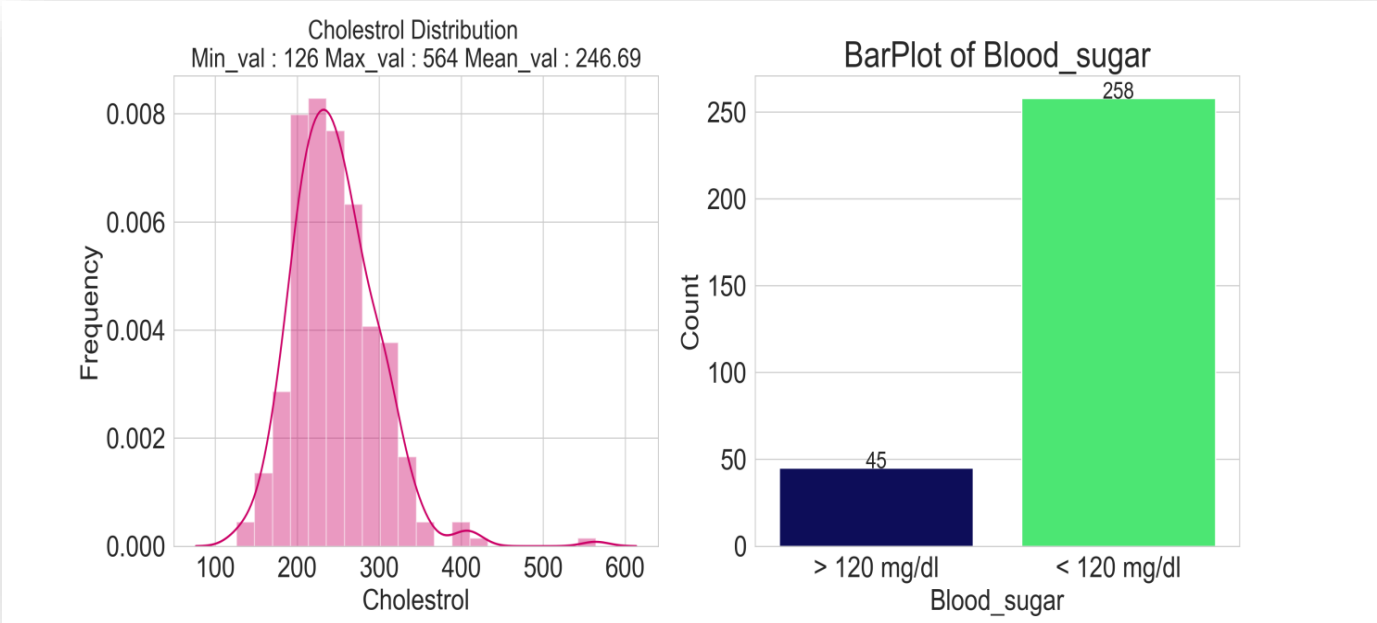
2. How are Resting blood pressure & Chest pain distributed?

Patients with Asymptomatic chest pain are more in number than patients with Non-Anginal chest pain than patients with Atypical chest pain.



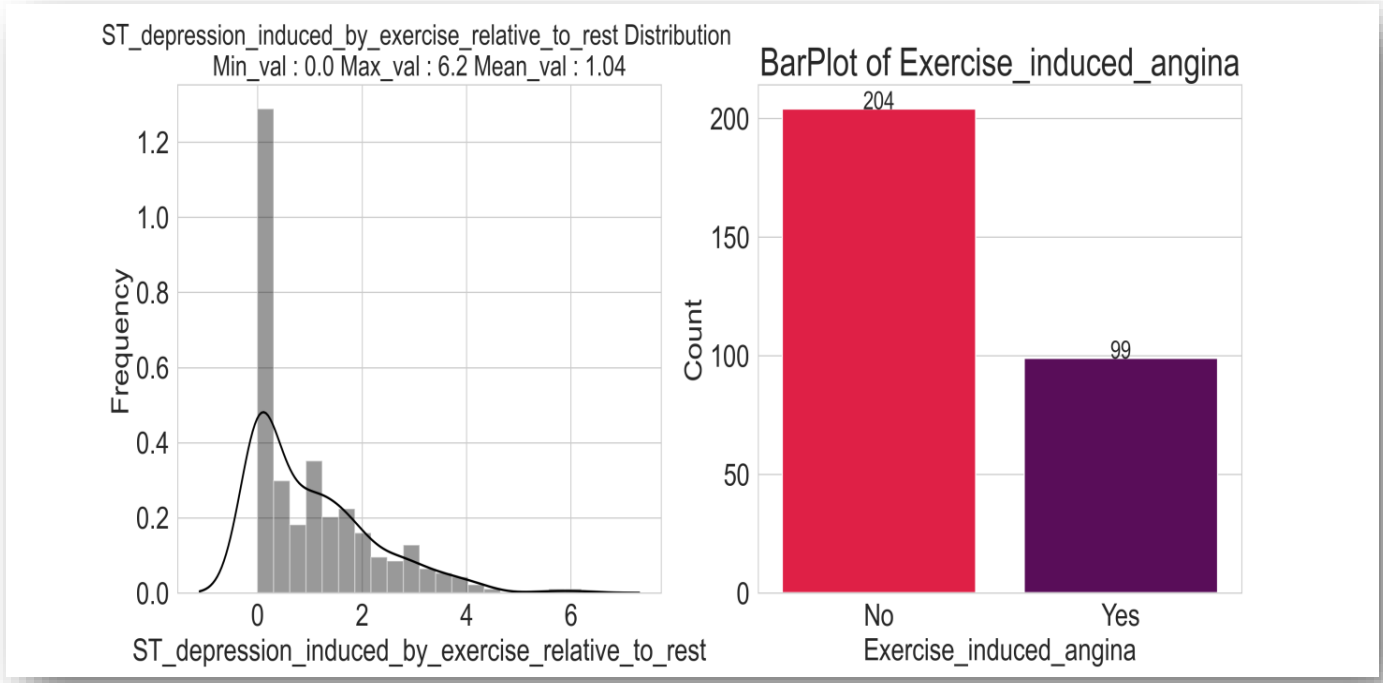
3. What is the level of Cholesterol and Diabetic conditions (calculated in terms of Fasting Blood Sugar) in all the patients in our sample data?

Most patients have Fasting Blood Sugar less than 120 mg/dl that is most patients are non-diabetic.



4. Are patients getting chest pain while exercising (indicated by Exercise induced angina) when they are asked to do exercise on Treadmill test

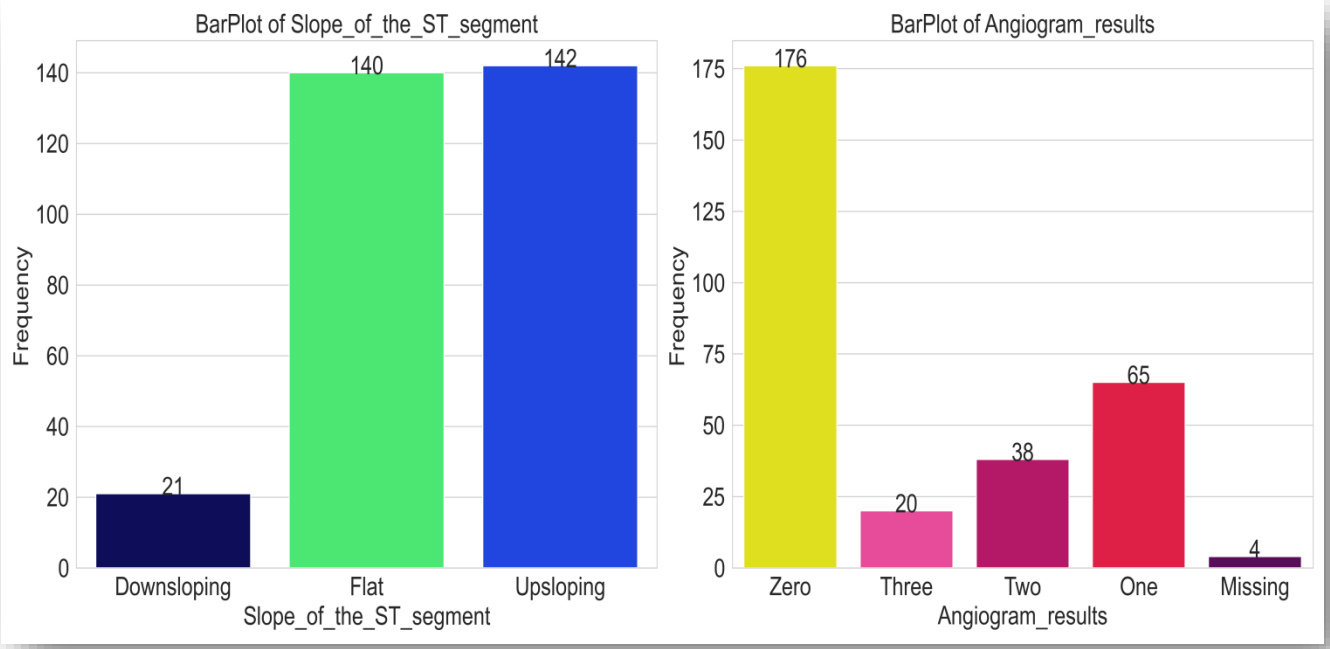
Most patients have normal ECG results whereas a small chunk of patients have ST-T wave abnormality.



5. **How are the major vessels of patients that supply blood, oxygen & nutrients to the heart distributed in the sample dataset (indicated by angiogram results)? What is the distribution of ST depression induced by exercise relative to rest?**

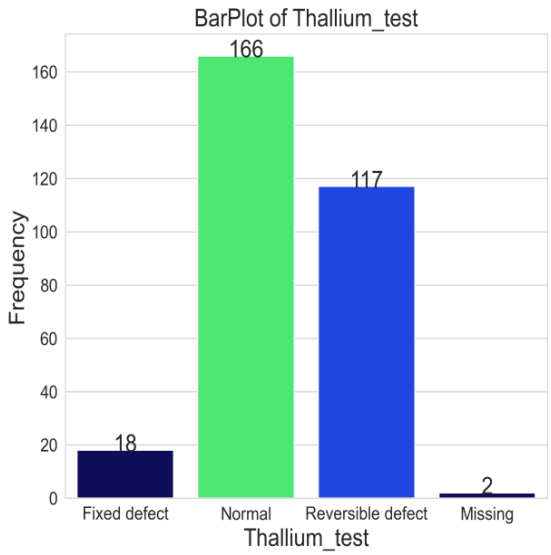
   Most patients are not having an angina/chest pain because of an Exercise (indicated by Treadmill test results).

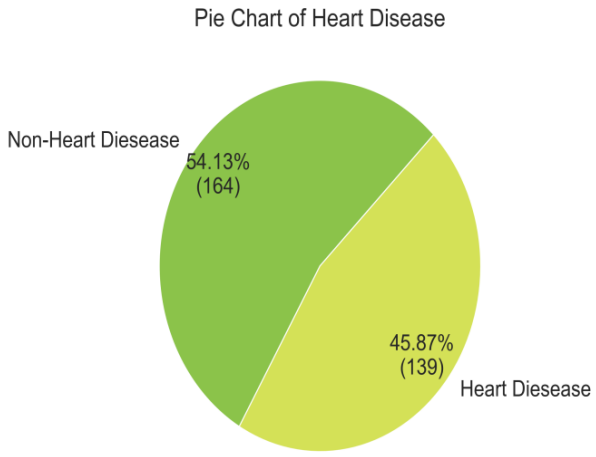
- After exploration of the data we found that Thallium test column has an unknown category as "100000" which we are defining for now missing. Same thing was found for Angiogram results as well.
- About 40% of the patients' blood vessels are blocked by Plaque.



6. **What is the percentage of Heart disease patients and Non-heart disease patients in our sample?**

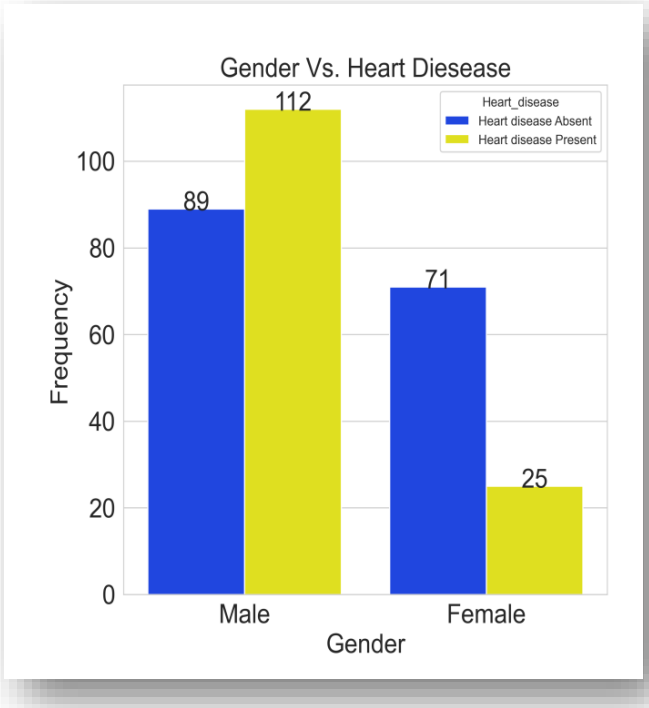
   Approximately 54% & 46% of the patients are Non-heart disease patients and Heart disease patients.





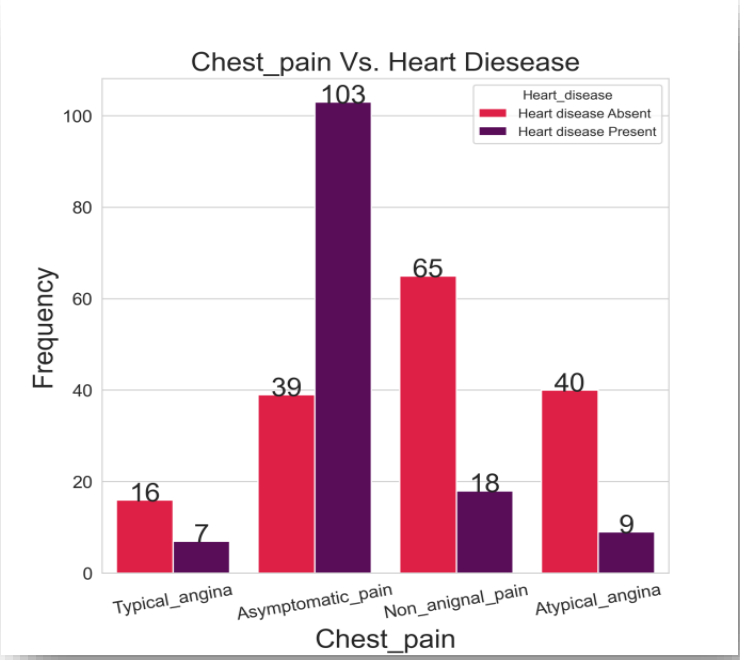
7. Which gender is more likely to get affected by Heart disease? Is gender an important factor in explaining the heart disease?

Both Men & Women are at greater risk of heart disease but the number for males is higher overall (indicated in below image). Gender is very important feature since the Critical value was coming very high using chi-square test.



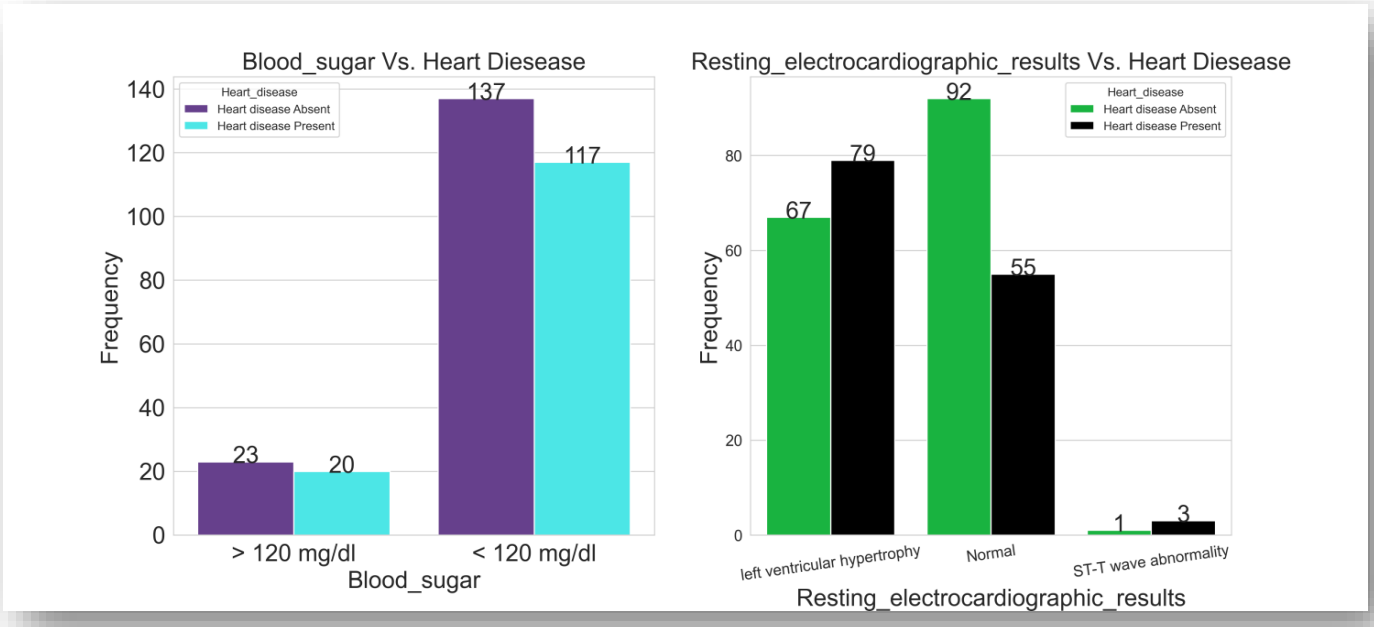
8. What chest pain type is causes heart disease more? Is chest pain an important indicator/factor in explaining the heart disease?

Asymptomatic chest pain is the main cause of heart disease among all the types of chest pains. Chest pain is a very important variable which was indicated by high chi-square critical value.



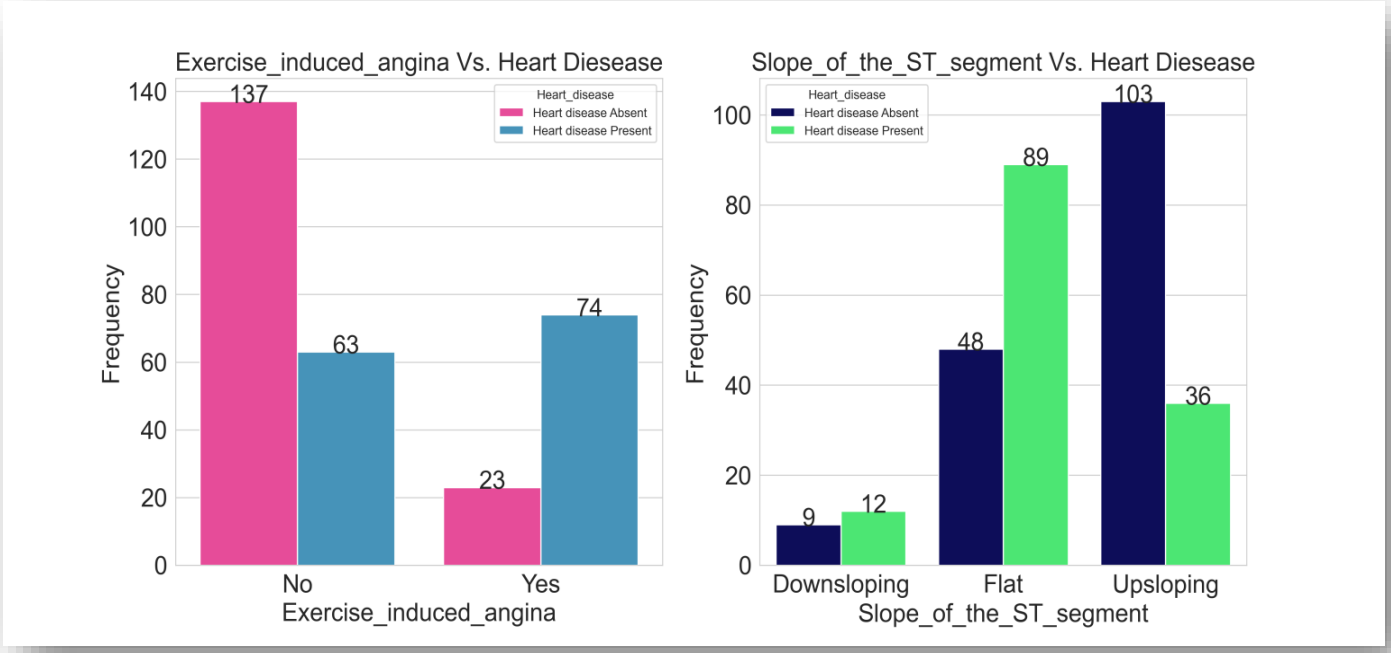
9. Are Blood sugar & Resting Electrocardiographic results important factors?

Neither Blood sugar nor Resting Electrocardiographic results variables are important since by chi-square test of independence these variables have no connection with Heart Disease.



10. What do Treadmill test result (indicated by Exercise induced angina) & Slope of ST segment say about Heart disease for the sample? Are they important factors?

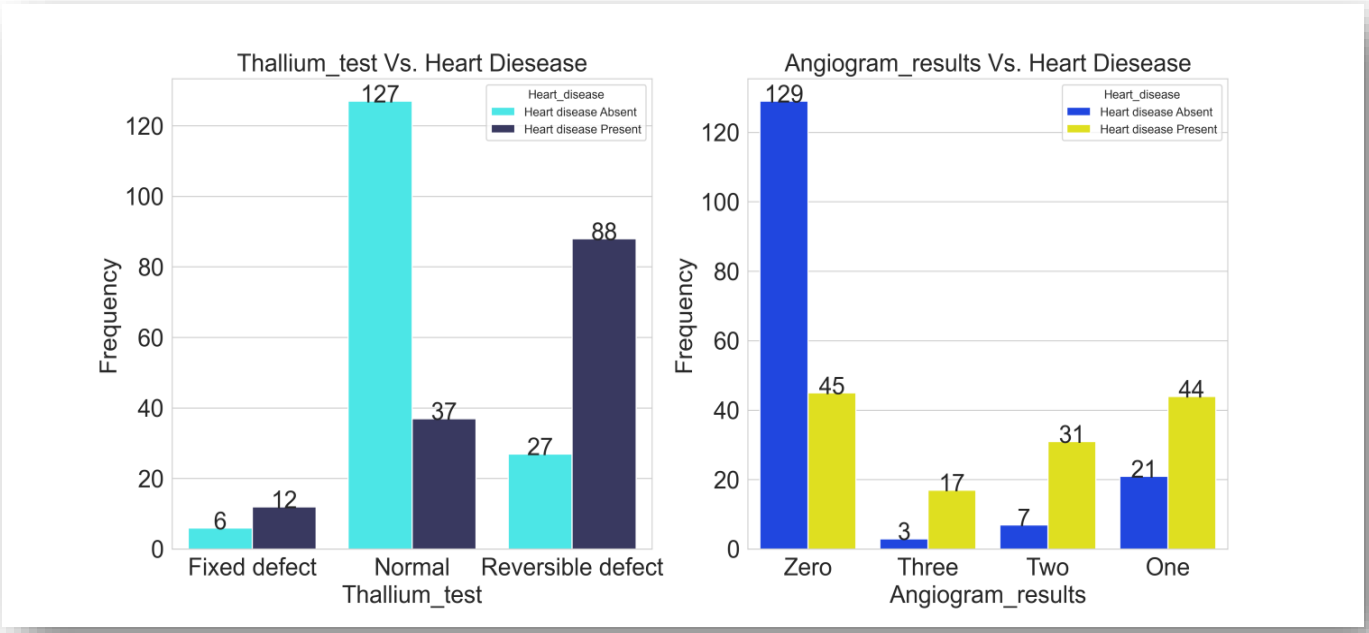
More number of heart disease patients experience angina while doing treadmill test, whereas less number of heart disease patients are experiencing angina at rest. Since the chi-square critical value is coming very high, therefore, we conclude Exercise induced angina to be an important feature. Slowly upsloping ST segment usually indicates heart attack. Horizontal ST Segment depression is considerable abnormal response. Down sloping ST Segment depression represents severe heart attack. We observe more number of patients having abnormal Slope of the ST segment (indicated by Flat slope of the ST segment). Number of patients with down-sloping Slope of ST segment is considerably very less than Flat slope segment patients. This feature is also very important indicated by high chi-square critical value.



11. What do Angiogram results & Thallium test say about heart disease?

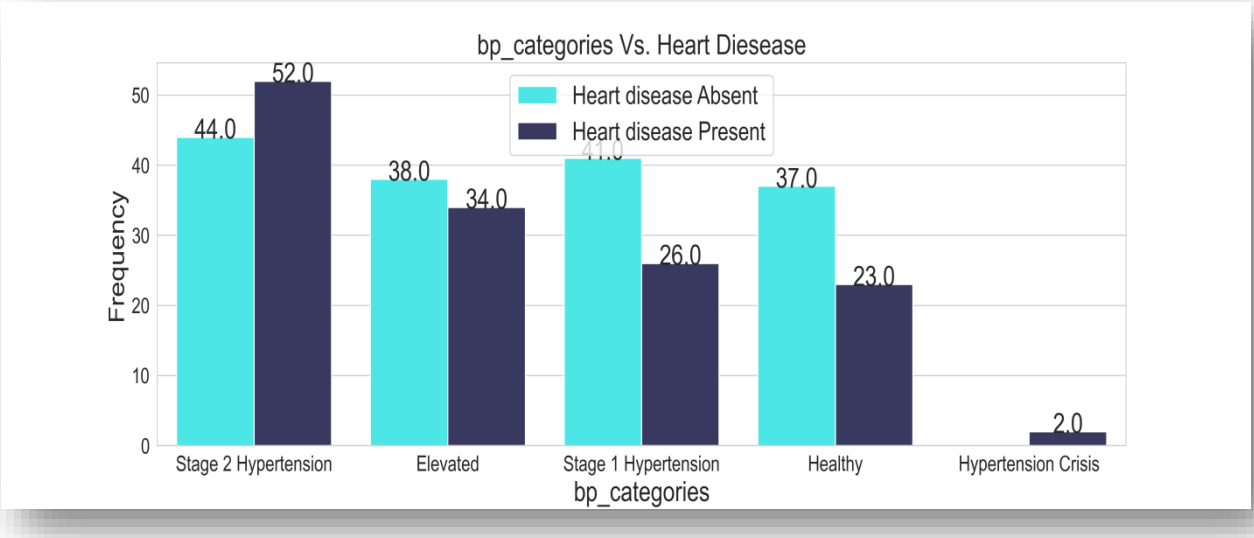
We notice that more number of non-heart disease patients with zero/no defect in their major vessels that supply blood, oxygen & nutrients to the heart than heart disease patients. But as the number of major vessels blocked increases we observe more number of heart diseases patients than with non-heart disease patients. These patients need an immediate attention (indicated by angiogram results). We observe a greater number of heart disease patients with reversible

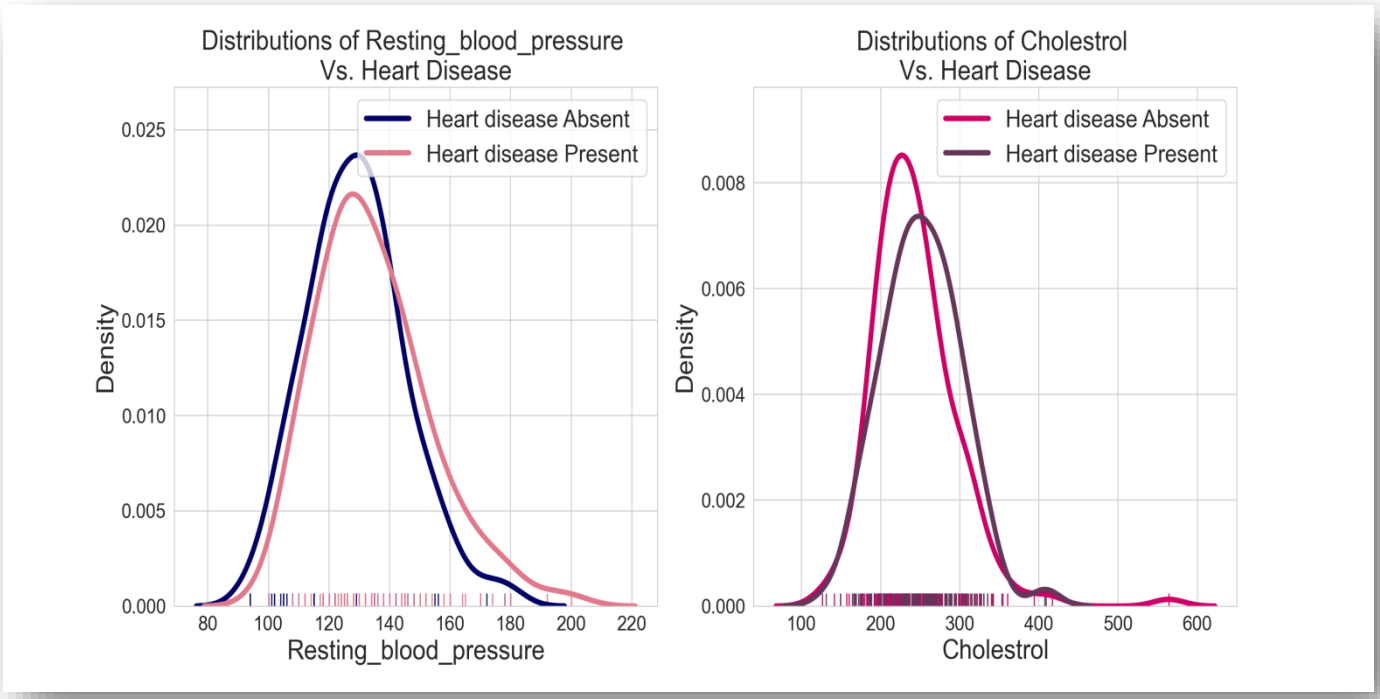
defect (indicated by Thallium test). And both features are important indicated by a very high chi-square critical value.



12. How are distributions of Resting blood pressure & cholesterol level distributed for Heart & Non-heart disease patients? Is there any statistical difference between average cholesterol level between heart & non-heart disease patients?

From both the figures given below we conclude that both heart disease & non-heart disease patients are suffering from high blood pressure (Hypertension). Heart disease patients are more at risk of stage2 hypertension than non-heart disease patients. We also notice that more healthy patients do not have heart disease than patients with heart disease.

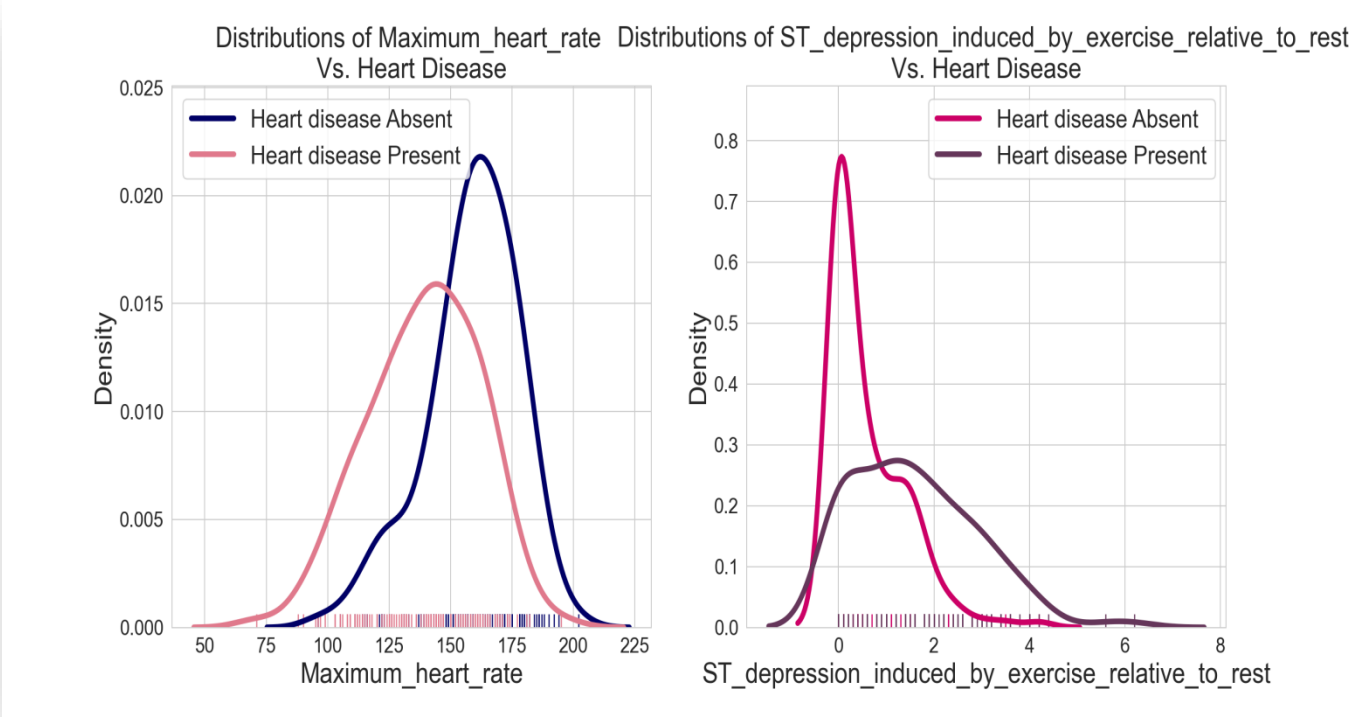




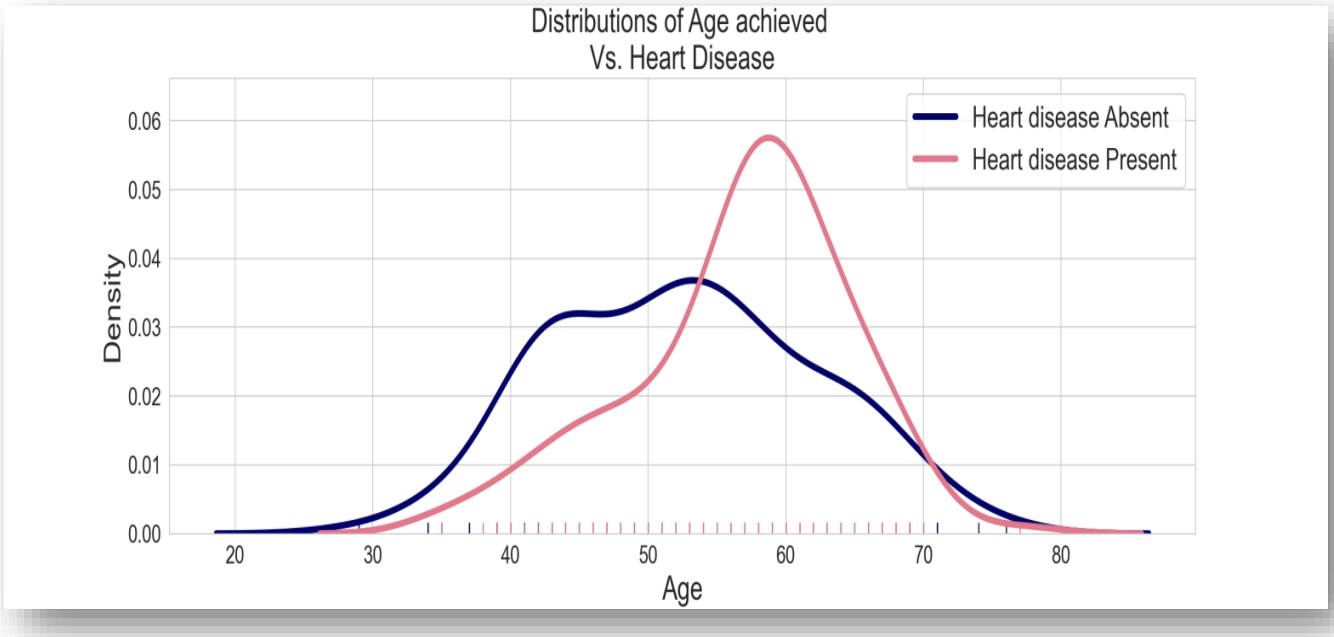
But this variable is not significant indicated by chi-square test when we created the blood pressure categories. There is also statistical difference between Average cholesterol level between heart disease & non-heart disease patients. The ideal total cholesterol level should be less than 200 mg/dl but we observe more number of heart disease patients beyond 250 mg/dl level.

13. How do the distributions of maximum heart rate & ST depression relative to rest look like for heart disease & non-heart disease patients?

There is a problem called Arrhythmia which is an irregular heart rhythm. It occurs with a normal heart rate or with slow heart rate that may include coronary artery disease, heart attack, blood imbalances & more. There are many types of arrhythmias which can occur with a normal heart rate or with heart rates that are slow or rapid. Abnormal heart rate reflects abnormal tone & is associated with higher risk of heart disease. In the sample dataset, we notice that more number of heart disease patients with lower maximum heart beats (beats per minute) & more number of non-heart disease patients with higher maximum heart rate.



14. What about Age? Generally, it is observed that older people (with age > 50 years) suffer from Heart disease. Is it true?



We can clearly see that within 50 to 70 years old age more number of heart disease patients than with age less than 50 years old (indicated in above plot). We have broken down the Age variable into 2 categories/groups/bins, i.e., age less than 50 years & age greater than 50 years.

We found Age group variable to be important which was indicated by chi-square test of independence. We also calculated the percentage of patients having heart disease given age is above 50 years old & percentage of patients not having heart disease which were coming to be roughly 46% & 53& respectively. The percentage of patients having heart disease & patients not having heart disease were found to be roughly 31% & 69% given the age is less than 50 years old.

These observations conclude that as a person becomes older the risk of damaged & narrowing arteries also increases. It also weakens or thickens heart muscles that contributes to Ischemic heart disease & thus lead to heart attack. We also see that patients with Non-heart disease are higher than heart disease patients but with not much greater difference. The gap in percentage between patients having heart disease & non-heart disease is much bigger. We, therefore, conclude that age is an important feature.

Findings/Conclusion

- Gender, Age, Cholesterol, Resting blood pressure, Maximum heart rate, Chest pain & some test results like Thallium test, Angiogram results, Exercise induced angina, slope of ST segment are the important factors in determining the Heart disease.
- As people becomes older they need to maintain their blood pressure, cholesterol level, Heart rates and they should visit to a doctor as well to get check their health check. To avoid any heart disease they can do some of the following things like avoid smoking, do exercise, avoid high fat consumption diet and adopt low fat diet, eat raw green vegetables, maintain their stress level. In short, people should change their lifestyle & adopt healthy habits.