

Heart Disease Diagnostic Analysis

DATA SCIENCE: UNIFIED MENTOR

MADE BY ARINDAM

Introduction

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 analyze this health and medical data for better future preparation.

The objective of this project is to analyze the given dataset and figure out the patterns that might arise from it, in order to better prepare for the future.

If we find out the rate of heart disease in accordance with different demographics such as age, gender and overall population examined etc. we can find out which people are at higher risk of being diagnosed by a heart disease.

Using Python to write a code to aid with the data analysis, we can find some trends

The Scope and Data Source

Scope:

This dataset has details of multiple patients and the readings of their cardiac conditions, including details such as cholesterol level and blood sugar levels of the patients along with whether they have a heart disease or not. By analyzing the data we can find the relation between different the condition and having heart disease. The analysis includes data cleaning, ETL (Extract- Transform & Load) data, EDA (Exploratory Data Analysis), and visualization techniques to interpret and present the findings in a visual manner.

Data Source:

The dataset used in this analysis is the same as provided in the project letter. [It](#) can be accessed through this [link](#).

Methodology

Data Cleaning: Handling missing values, ensuring data consistency, and preparing the dataset for analysis.

Extract- Transform Load: Extracting the important data columns and finding out

Exploratory Data Analysis: Calculating summary statistics, identifying key metrics, and exploring relationships between variables.

Visualization: Using line plots, bar charts, and heatmaps to represent data trends and insights visually.

Key Attributes

Key Columns that we used for the analysis:

1. Age.
2. Sex.
3. Resting Blood Pressure.
4. Fasting Blood Sugar.
5. Cholesterol level.
6. Exercise induced agnia.
7. Target (used for the comparison between those who have heart disease and those who do not.)

Data Cleaning and Preparation

Handling Missing Values:

No missing values were found in the dataset and hence there was no need to handle the missing value. By default, the way to handle the missing value is to set the value to zero (0).

Data Type Adjustments:

All the data is made to be in integer values so the analysis can be made easily, the units have been removed and only numerical data has been kept.

Additional Processing:

Unnecessary data that we deemed irrelevant has been not considered in the analysis such as number of major blood vessels etc.

Exploratory Data Analysis (EDA)

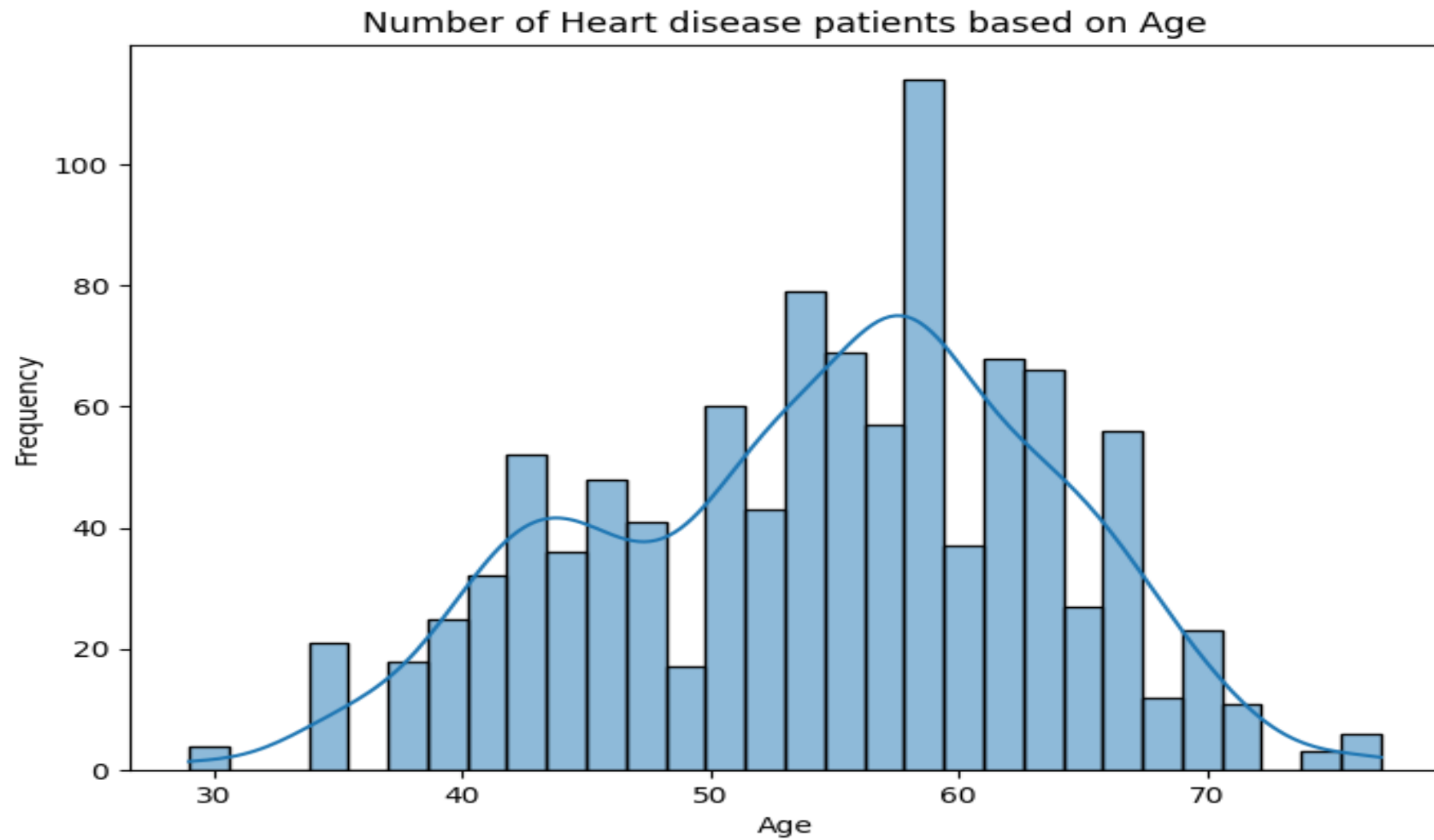
Summary Statistics:

Different details regarding the people's cholesterol levels, blood sugar and whether or not they get exercised induced angina, age and gender when graphed against different columns will let us derive different conclusions.

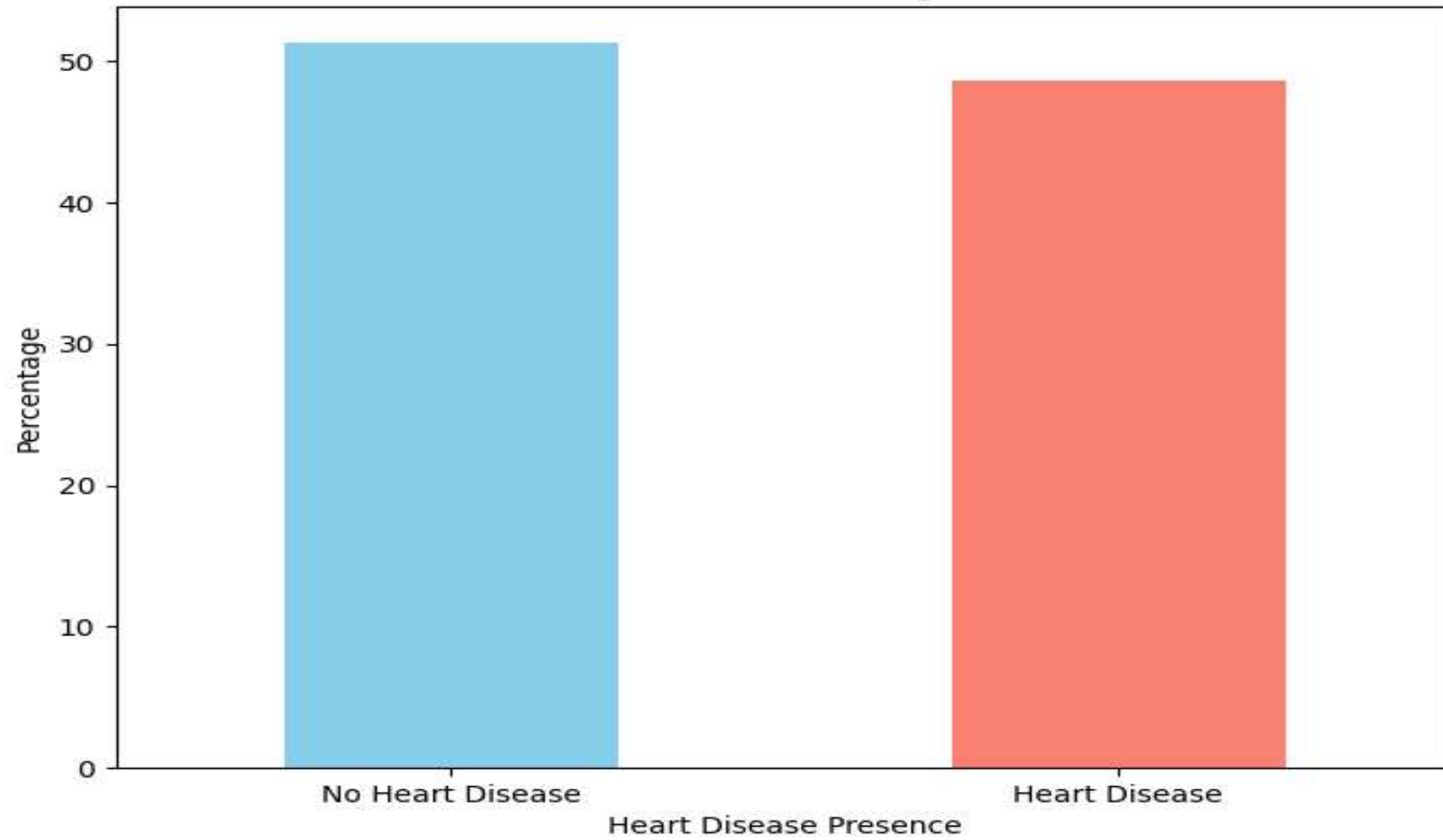
Key Metrics:

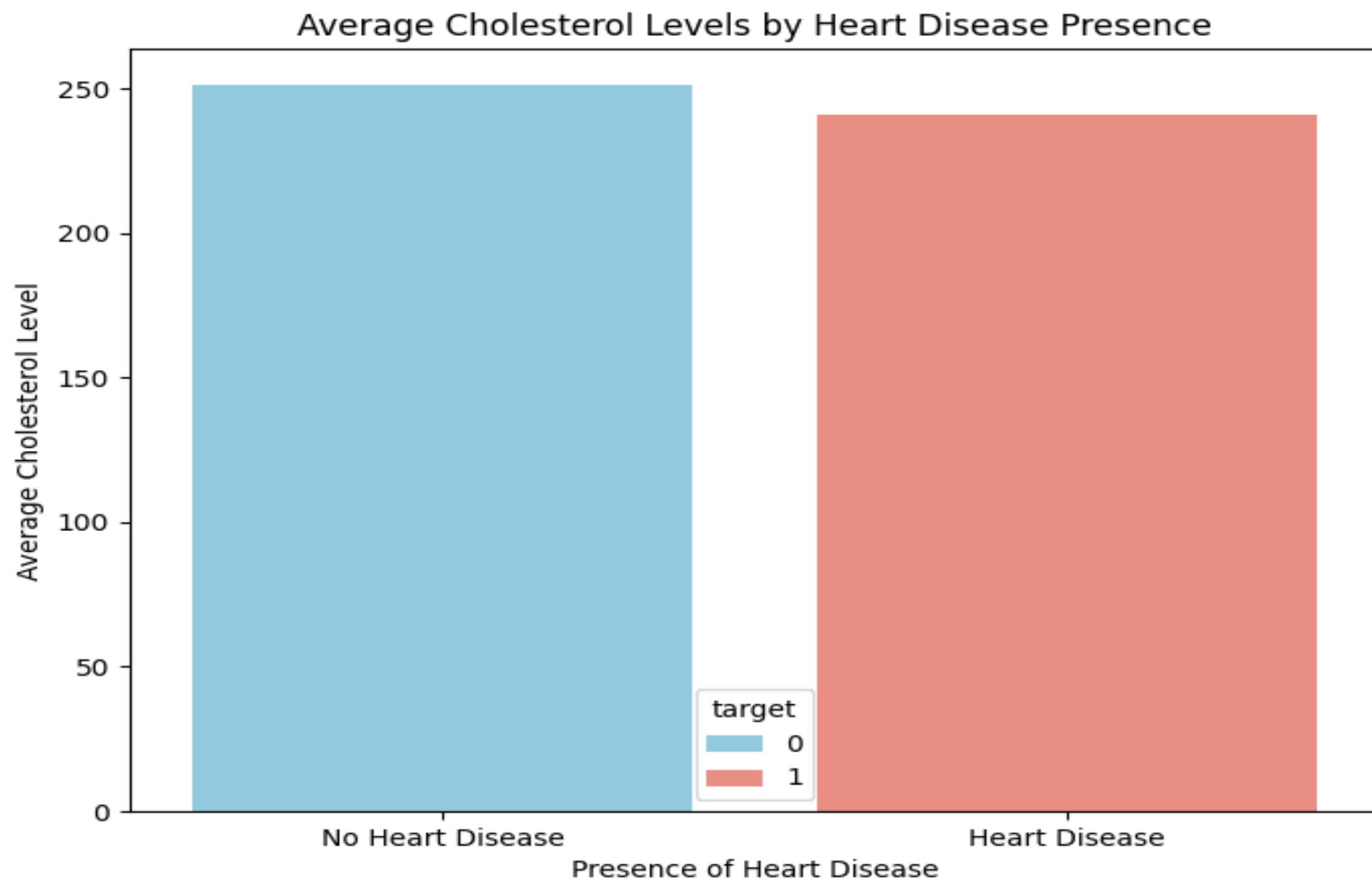
The key metric in analyzing the data was whether or not the patients have a heart disease or not. All other metrics were compared against each other with respect to the number of patients having or not having heart disease.

Graphs visualising the trends
in the value

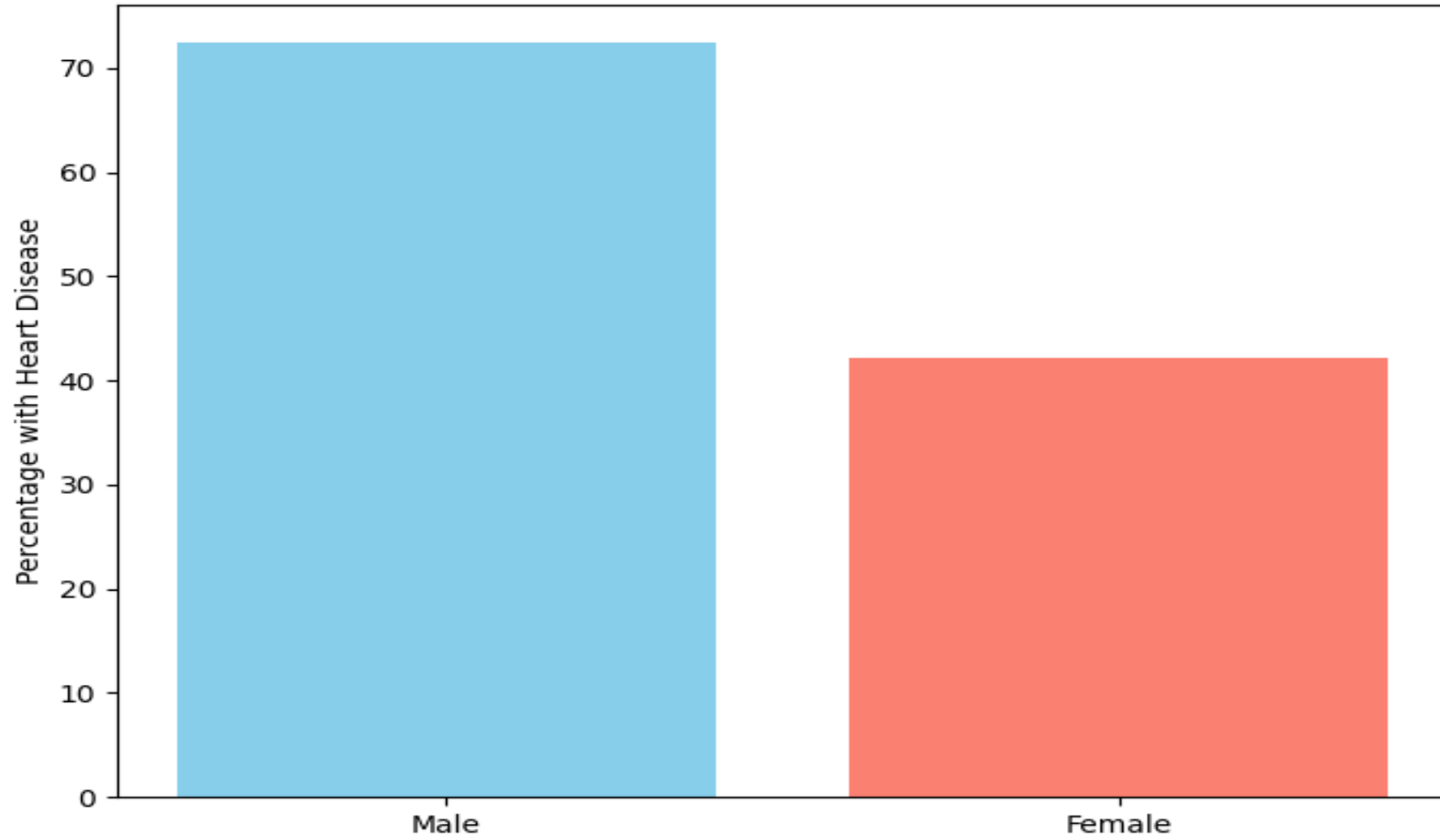


Rate of Heart disease in the given Dataset

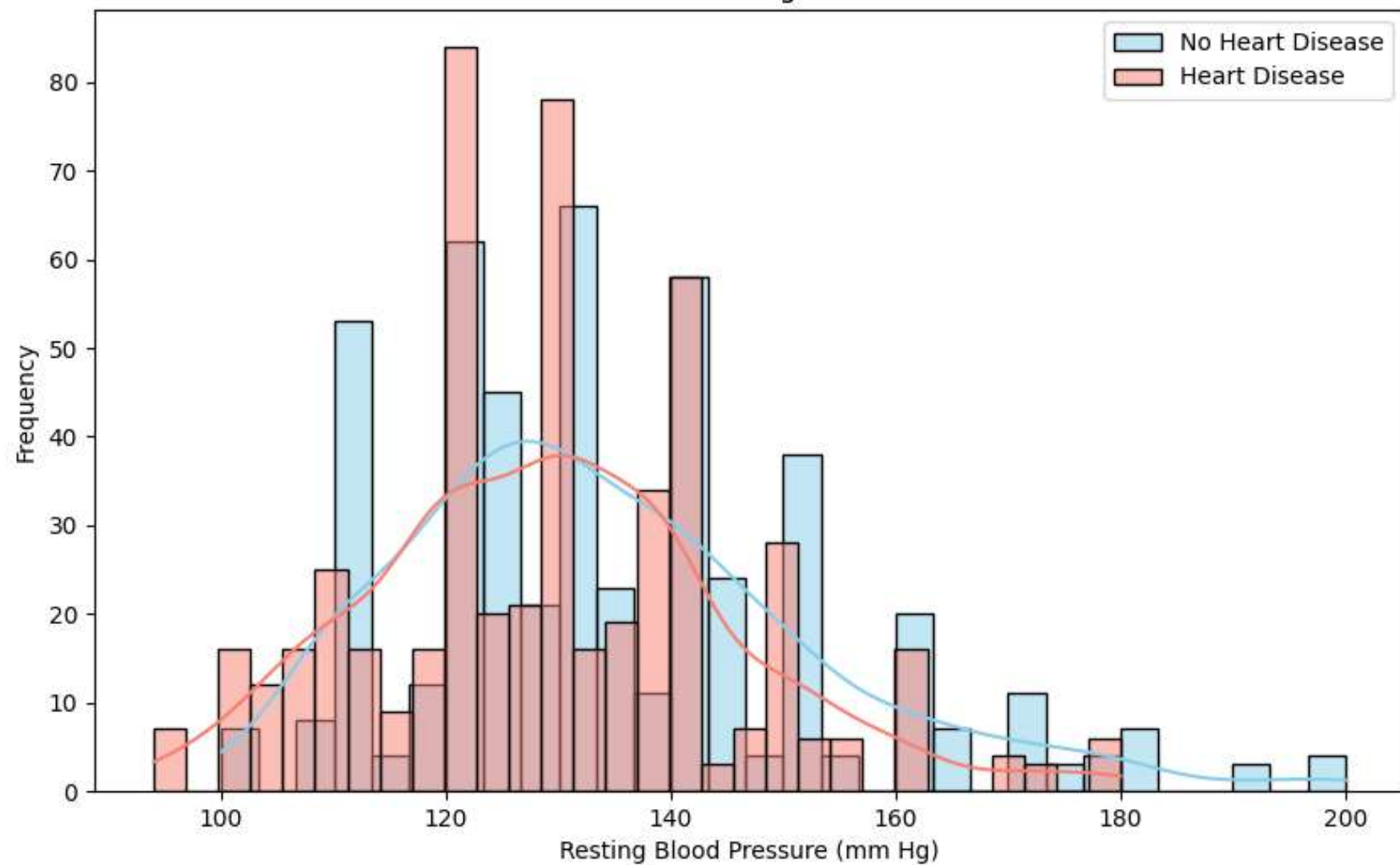




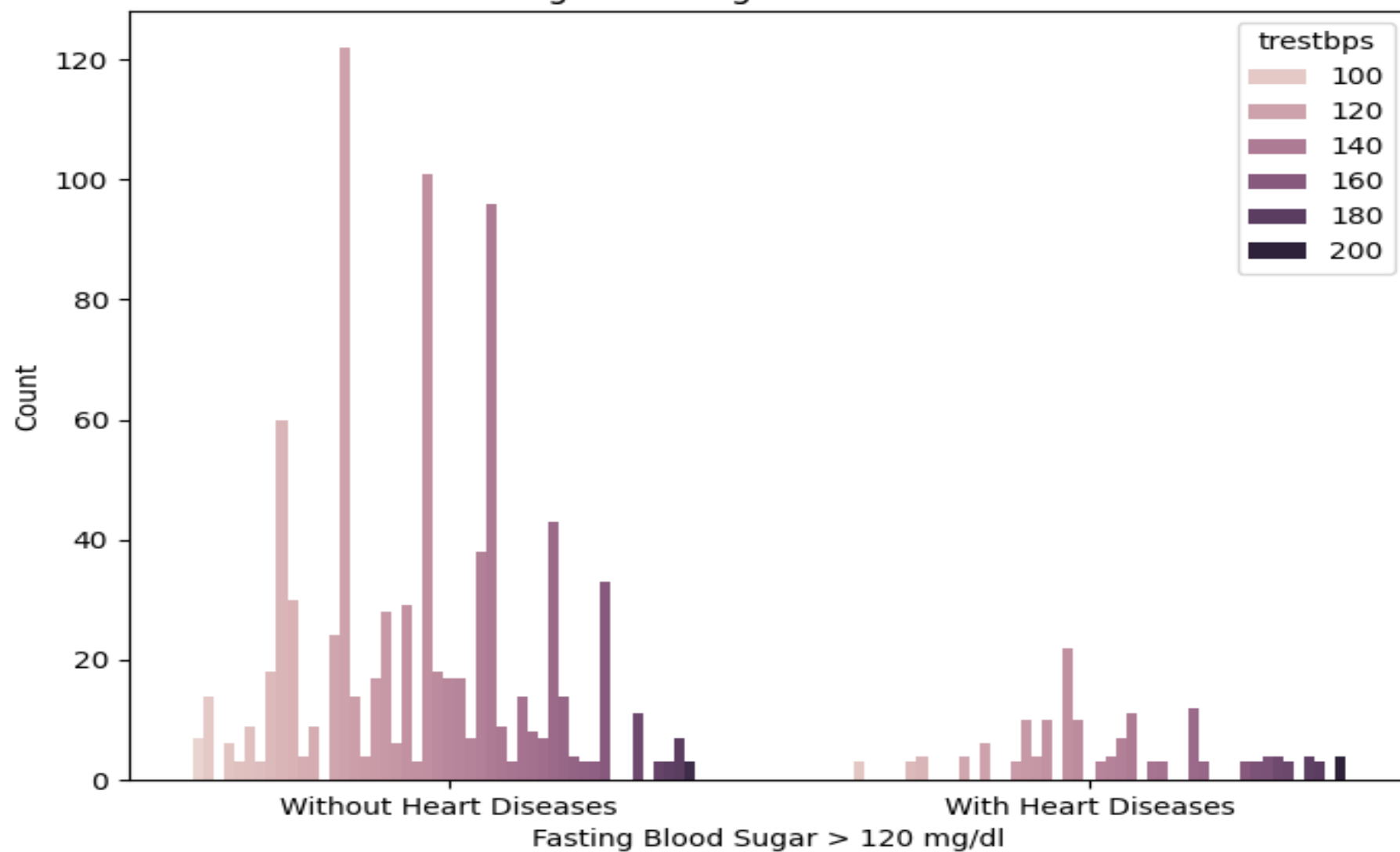
Heart Disease Rate by Sex

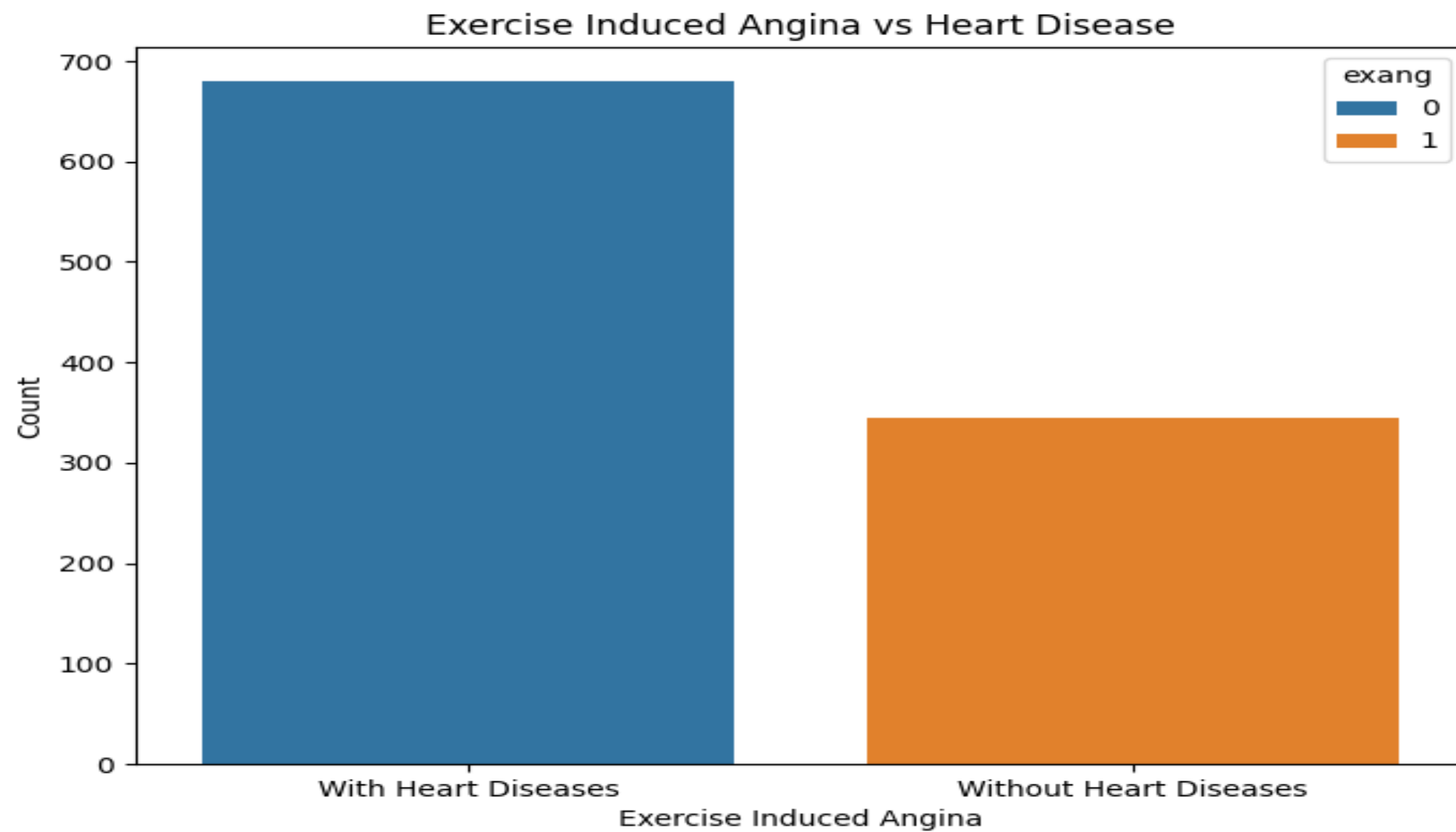


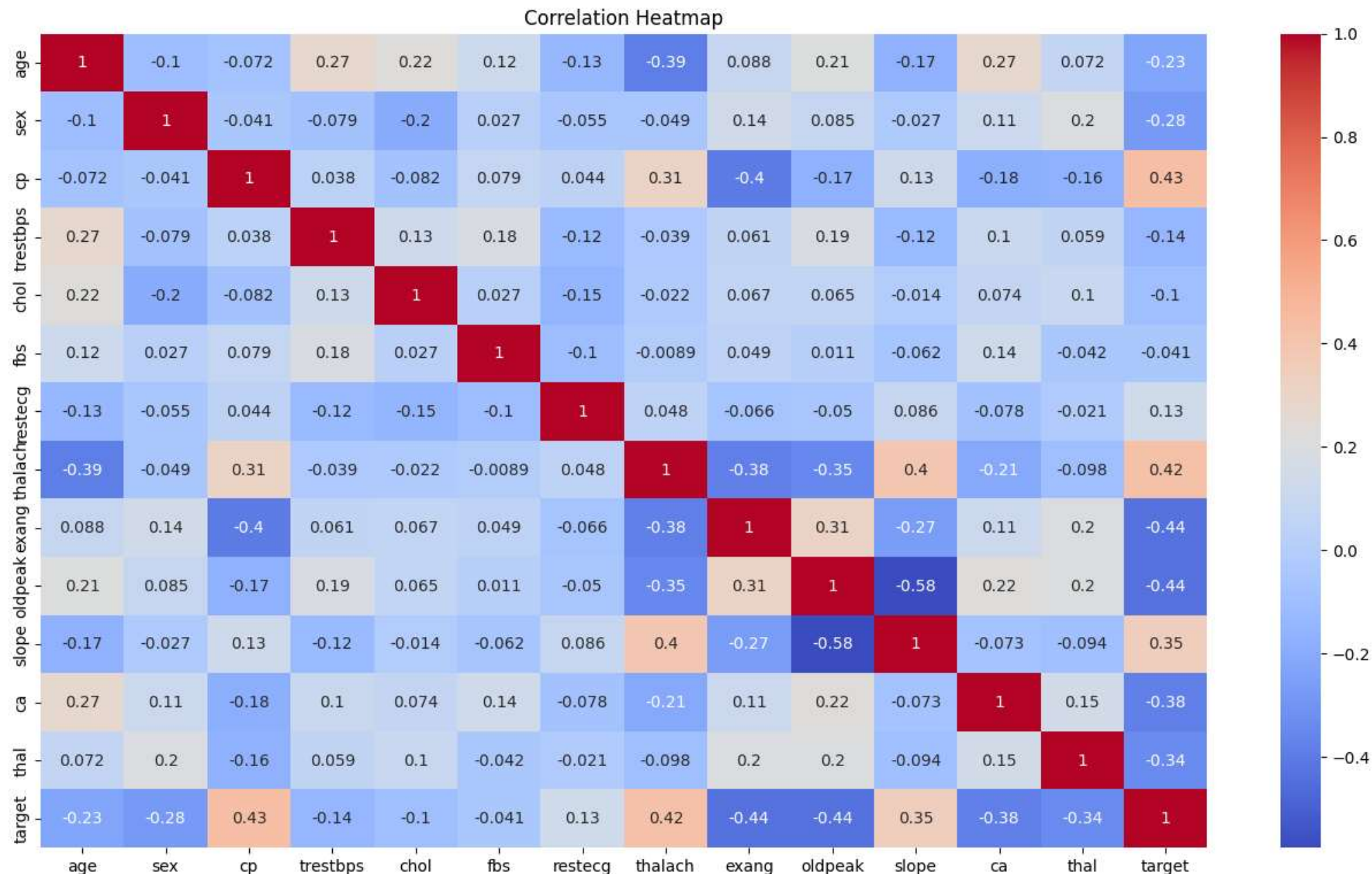
Distribution of Resting Blood Pressure



Fasting Blood Sugar vs Heart Disease







Conclusion

Looking at the visuals obtained and then relating it to the dataset, we can conclude the following details:

1. People in their late 50's have higher chance of getting a heart disease.
2. Heart diseases are fairly common, as nearly half of the patients seem to have heart diseases.
3. Heart diseases isn't directly caused by having higher cholesterol as the people with heart diseases don't necessarily have higher cholesterol than those without heart diseases.
4. Male population is at higher risk of getting heart diseases than the female population.
5. People with heart diseases have higher resting blood pressure.
6. People with heart diseases have lower fasting blood sugar levels.
7. People with heart diseases are prone to getting Exercise induced agnia.

Thank You

MADE BY: ARINDAM

