Assignment 2

Heart Disease Analysis

**21st May 2018**

*Prepared by Rahul Gupta (s3635232) and Neha Voora (s3691382)*

Table of Contents

[Introduction 3](#_Toc425663362)

[Methods 3](#_Toc425663363)

[Results 3](#_Toc425663364)

[Discussion 4](#_Toc425663365)

[Conclusion 4](#_Toc425663366)

[References 5](#_Toc425663367)

# Executive Summary

Here we explored the statlog heart data to get insights about heart disease. Features are observed individually, in groups via visualization and machine learning algorithm.

Need to write few lines of analysis

# Introduction

This report explores the (Statlog (Heart) Data Set, 2004) collected from UCI Machine learning repository.

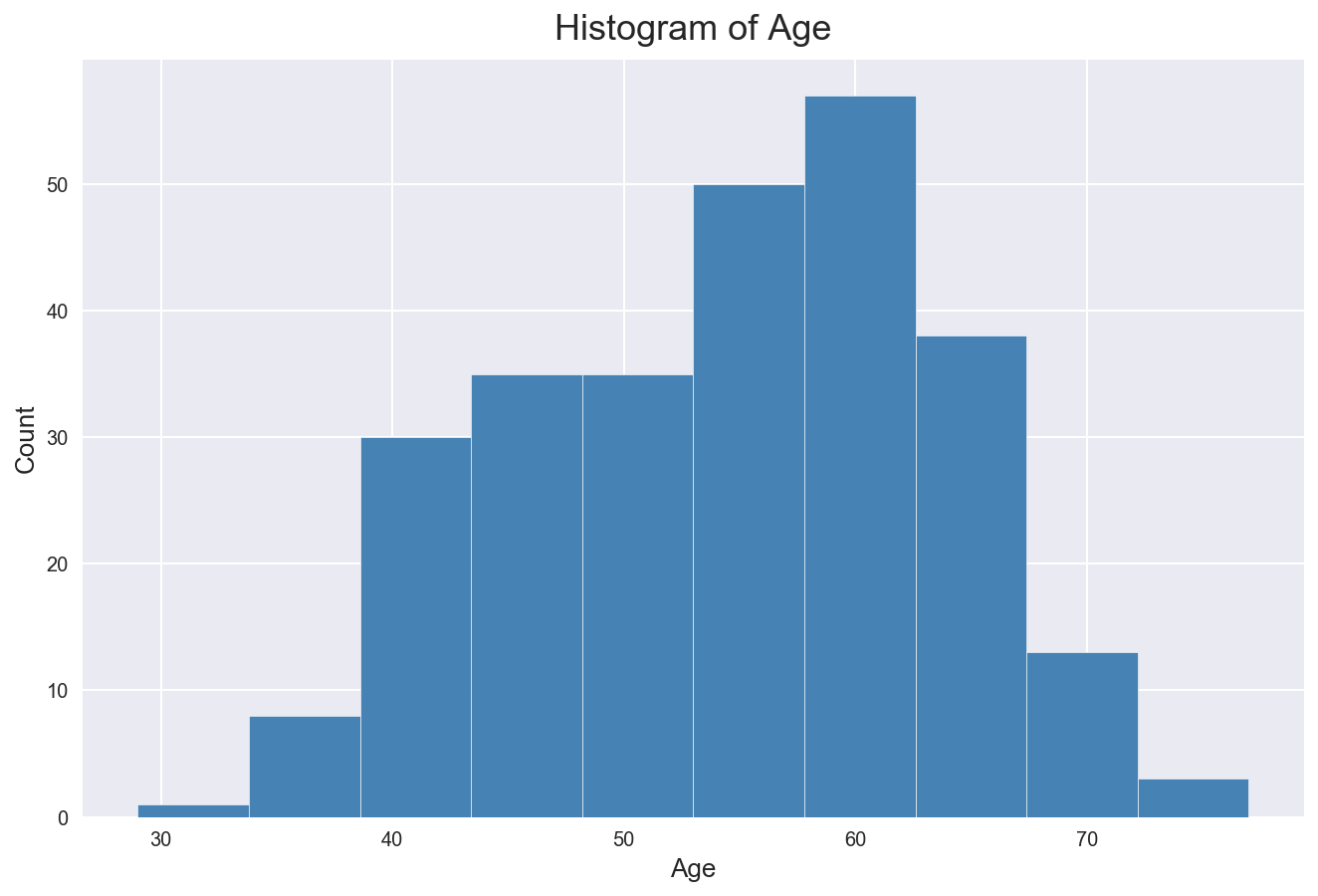
# Data Exploration

Here we’ll explore the features using python pandas and matplotlib library.

Data has no missing values.

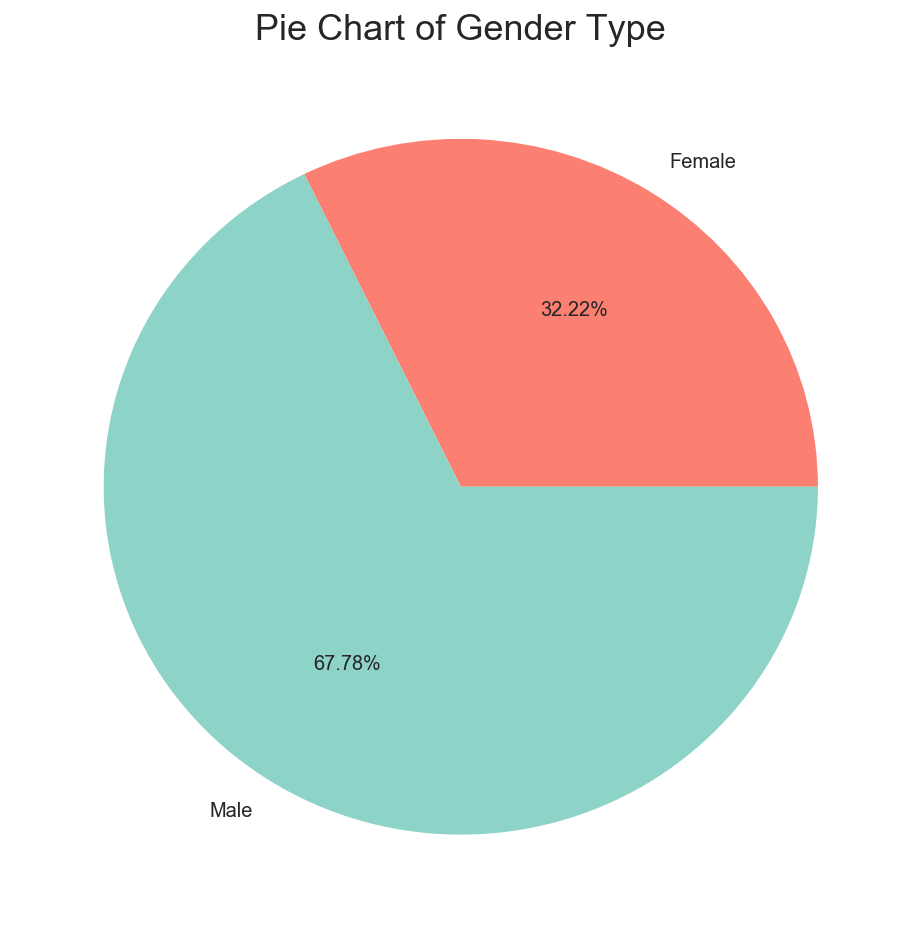
### Feature Visualization – Age

Age is a continuous variable and following visualization shows the distribution of age. Most of the data is between 50 and 70 years.



### Gender Type

Age is a nominal variable. 2/3 of sample collected are from males while remaining 1/3 are females



### Chest Pain Type

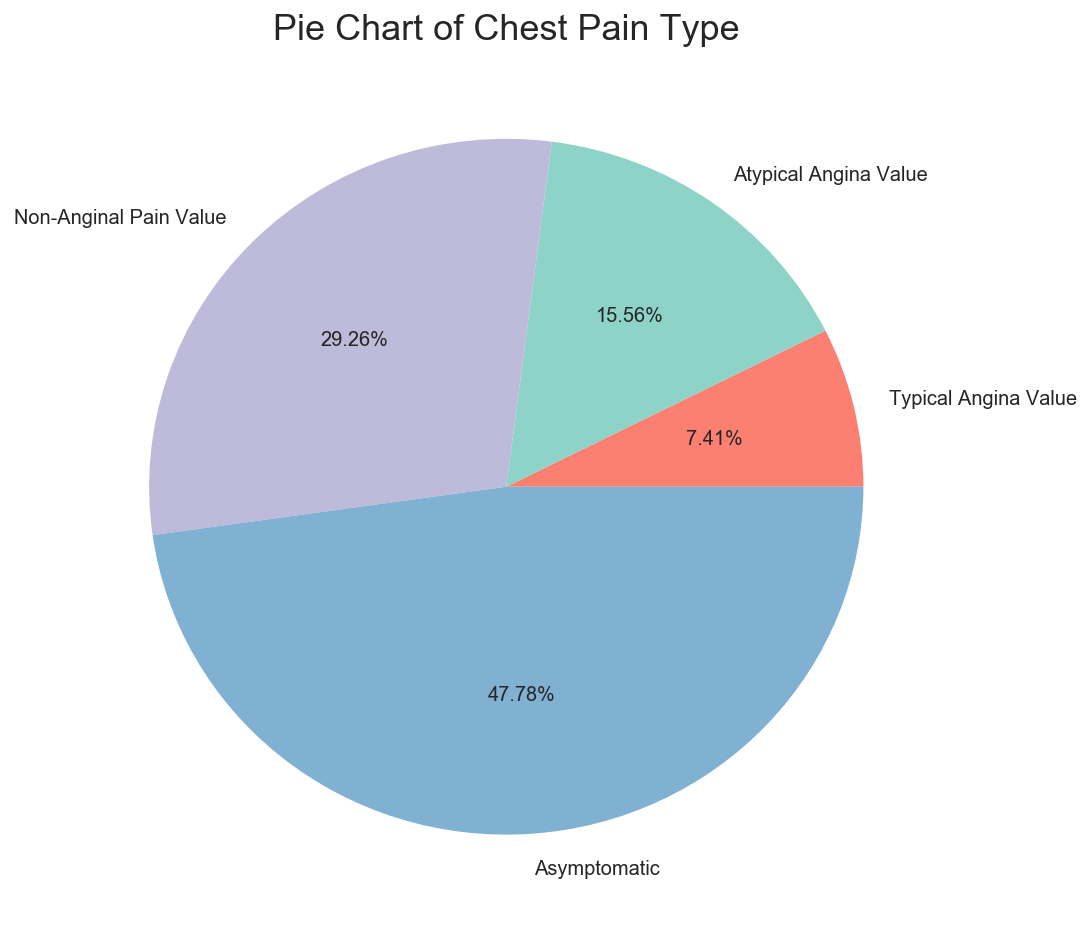
This is categorical variable. We have 4 different types of chest pain:

* Asymptomatic
* Non-Anginal Pain Value
* Atypical Angina Value
* Typical Angina Value

Chest pain is important factor for heart disease. Chest pain can also be due to problems in lungs, ribs, oesophagus, muscles, or nerves. *Asymptomatic* means neither causing nor exhibiting symptoms of any disease.

Non-angina chest pain is of short duration, typically less than 30 minutes or less than 5 seconds. it can be relieved immediately on lying down. Typical angina pain is presence of substernal chest pain or discomfort that was provoked by exertion or emotional stress. (Gore, 2010). Atypical angina pain can last up to days while typical angina pain lasts from 3-15 minutes.

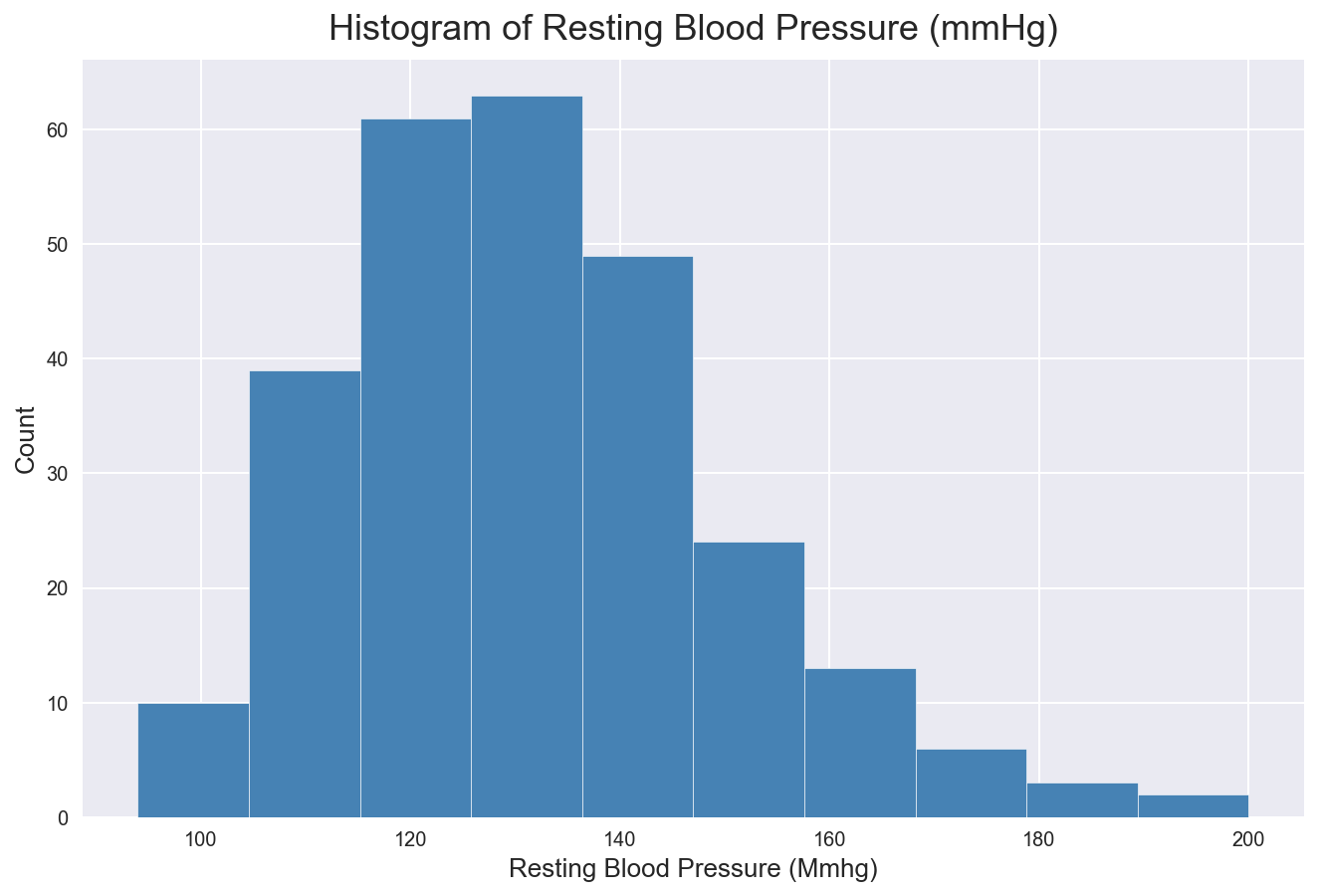
From the visualization, we can see that almost half samples have Asymptomatic chest pain. 30% have non-angina pain, while 15% have atypical angina pain.



### Resting Blood Pressure

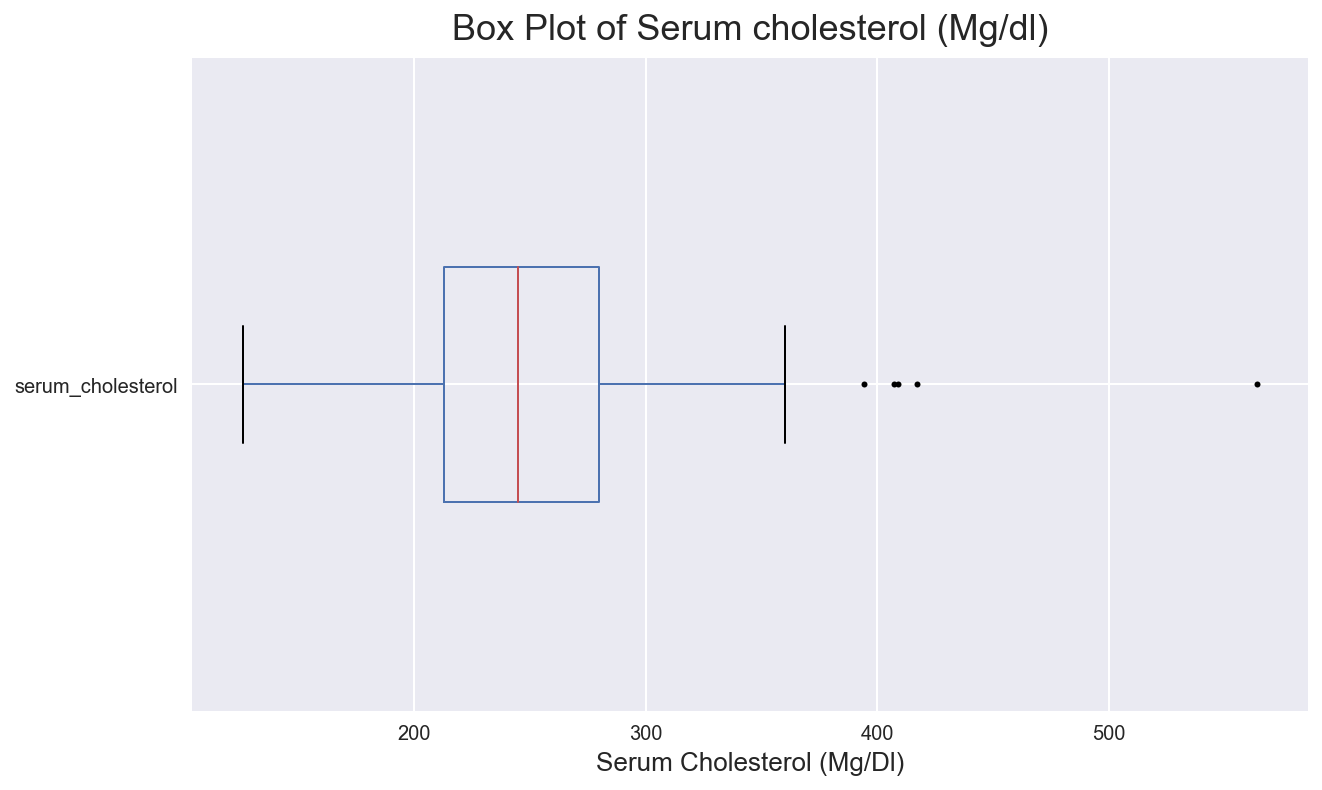
Blood pressure is the pressure of circulating blood in the walls of blood vessels. measured in millimetres of mercury (mmHg), above the surrounding atmospheric pressure. Normal resting blood pressure in an adult is around 120 systolic, and 80 diastolic.

Sample data has positive skewed distribution. Most of the pressure is between 120 and 160.



### Serum cholesterol

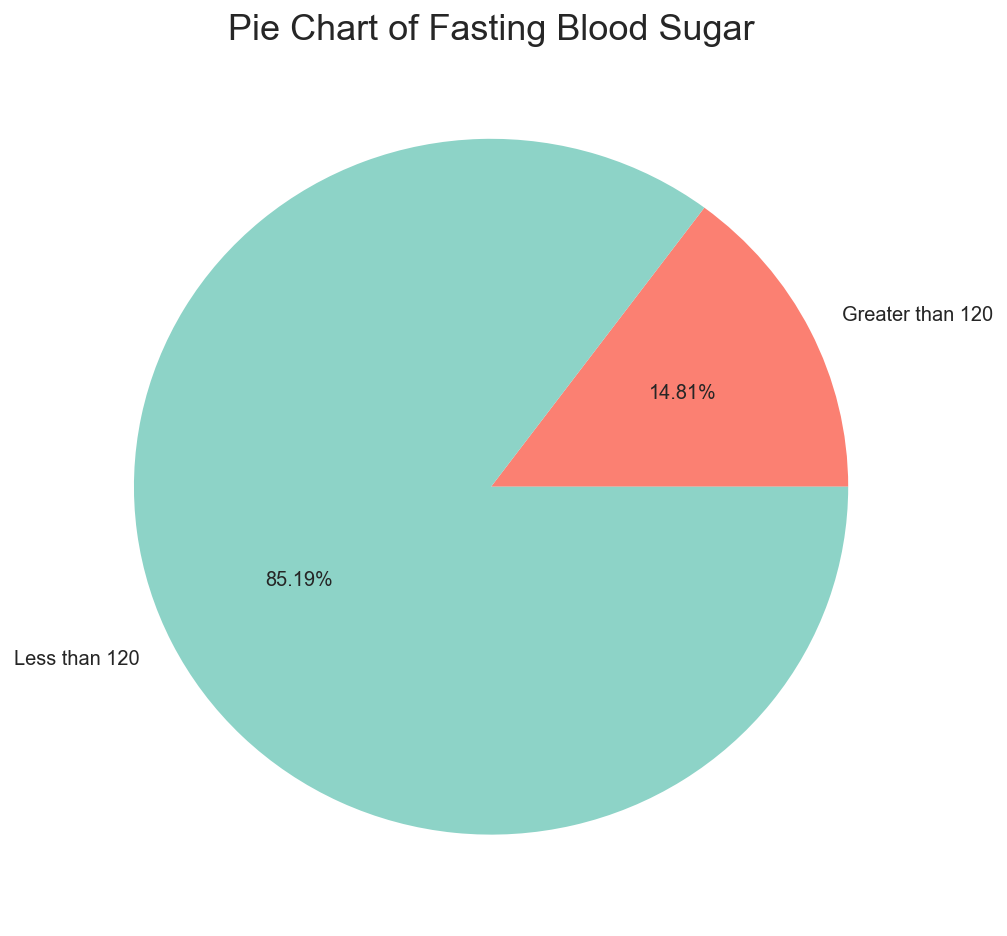
Cholesterol is a type of fat, also called as lipid. It travels through your bloodstream in molecules that can build up in arteries and restrict or block blood flow. *This is often associated with heart disease.* healthy serum cholesterol should be less than 200 mg/dl. Here the sample, most of the observation has high level of cholesterol, median is around 240 mg/dl. We have few outliers with more that 400 units.



### Fasting Blood Sugar

Blood sugar is the amount of glucose in blood. This is one of the major factor for Diabetes. Here the continuous variable is used as nominal variable. Fasting blood sugar level less than 100 mg/dL is normal. Up to 125 is considered as pre-diabetic.

Data is divided into greater than or less than 120. This will help up to identify if heart disease chances increase with blood sugar levels.



# Discussion

# Conclusion

# References

Gore, J. M. (2010, June 1). *Typical Angina vs. Atypical Chest Pain*. Retrieved from Journal Watch: https://www.jwatch.org/jc201007070000002/2010/07/07/typical-angina-vs-atypical-chest-pain

*Statlog (Heart) Data Set*. (2004). Retrieved from UCI Machine Learning Repository: http://archive.ics.uci.edu/ml/datasets/statlog+(heart)