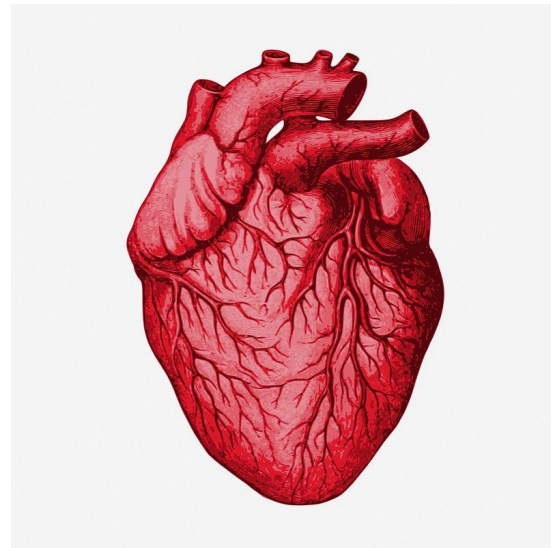# Heart Disease Factors

Learning about the factors that relate to heart disease

by Em Greene, Gabriel Rodriguez, and Noor Zaki

Dataset

- 270 patients were screened on 13 factors related to heart disease.
- We looked at :
  - Number of Vessels Fluro, Chest Pain Type, Max HR, Exercise Angina, ST Depression, Sex, Age, EKG Results, BP, Cholesterol
- We did not evaluate:
  - Fbs over 120, Slope of ST, or Thallium, as these were not clearly defined in the Kaggle resources

Research Questions

1. **Are there correlations between age and heart disease factors such as Cholesterol, max HR, and blood pressure?**
   - There are very tiny correlations between age and heart disease risk factors such as cholesterol levels, maximum heart rate and blood pressure.
   - Cholesterol levels tend to increase with age a small amount.
   - Maximal heart rate also tends to decrease with age which means older individuals may have a harder time achieving and maintaining elevated heart rate.
   - Blood pressure also tends to increase with age.
   - Finally, it is acceptable to conclude that cholesterol, max HR and blood pressure do have a very small correlation in regards to heart disease.

Research Questions

2. What are the top 3 factors related to heart disease?

- Number of vessels fluro
- Chest Pain Type
- Max HR
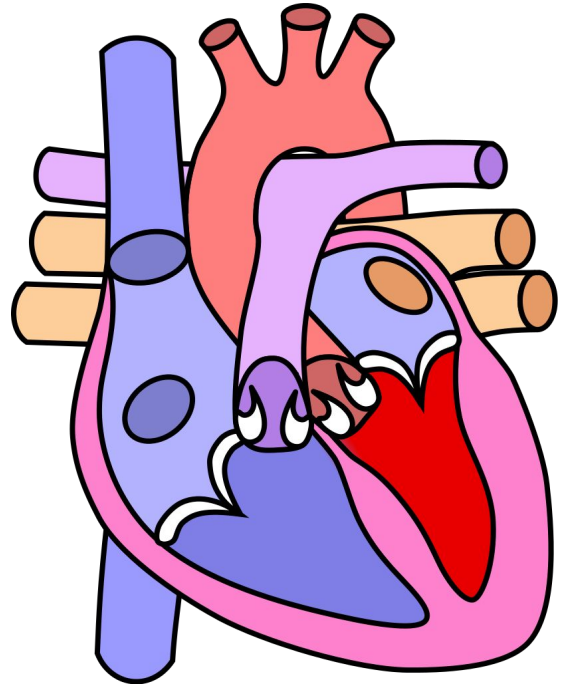
Pie Chart Analysis

- **Sex Distribution Chart:**
- In the Sex Distribution Chart the finding was that there was more men individuals in this study then female individuals.
- 68% of the sex in this study was males (1)
- 32% of the sex in this study was females (0)
- **Chest Pain Type Distribution:**
- In the Chest Pain Type Distribution Chart the finding was that about half of the population of the study experienced some sort of chest pain.
- There's 4 types of chest pain:
- **Value 1:** typical angina (normal pain)
- **Value 2:** atypical angina (abnormal)
- **Value 3:** non-anginal pain (not severe)
- **Value 4:** asymptomatic

Sex Distribution


Chest Pain Type Distribution
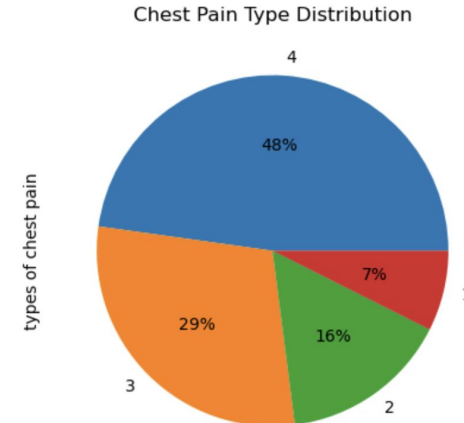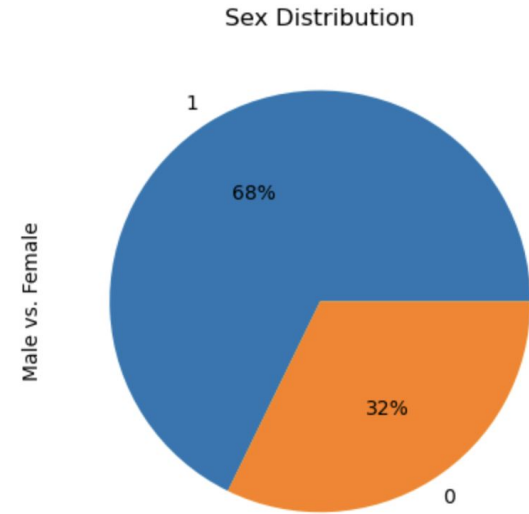
- **Exercise angina Chart:**
- In the Exercise angina chart the finding is that out of all individuals under study, 67% had exercise induced angina present vs. only 33% had exercise induced angina absent.
 - Value 1: yes
  - Value 0: no
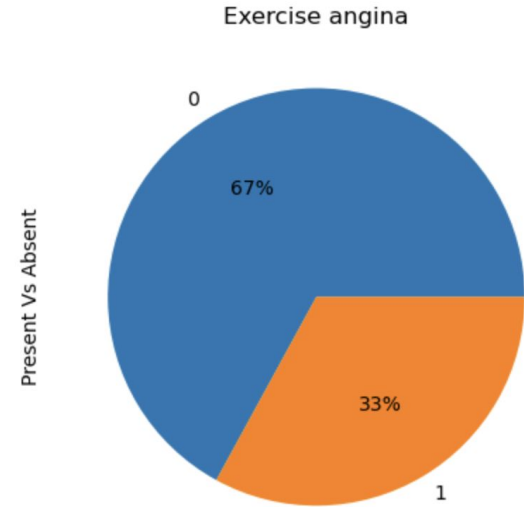
- **Heart Disease Distribution:**
- In this Heart disease distribution chart the finding is that heart disease was present in a little less than half of the 270 patients.
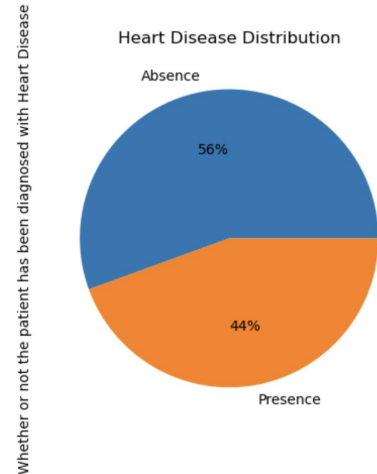
- **EKG Results Distribution Chart:**
- In this EKG results chart the finding is that EKG results were positive for atypical heart rhythm for about half of the population.
- **Value 0:** normal
- **Value 1:** having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV)
- **Value 2:** showing probable or definite left ventricular hypertrophy by Estes' criteria

Exercise angina

Heart Disease Distribution

EKG results Distribution
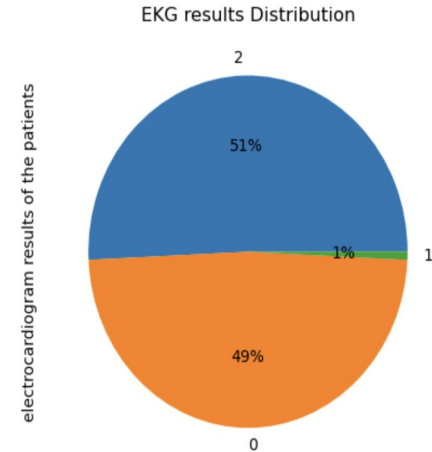
Scatter Plot Analysis

Based on the scatter plot representation and the trend line indicator, there is a general increase in blood pressure as age increases, with most of the data points falling between ages 40 and 70 and blood pressure values ranging from 120 to 140.

Outliers could be a representation genetics, lifestyle factors, or pre-existing health conditions.

Age vs. Blood Pressure
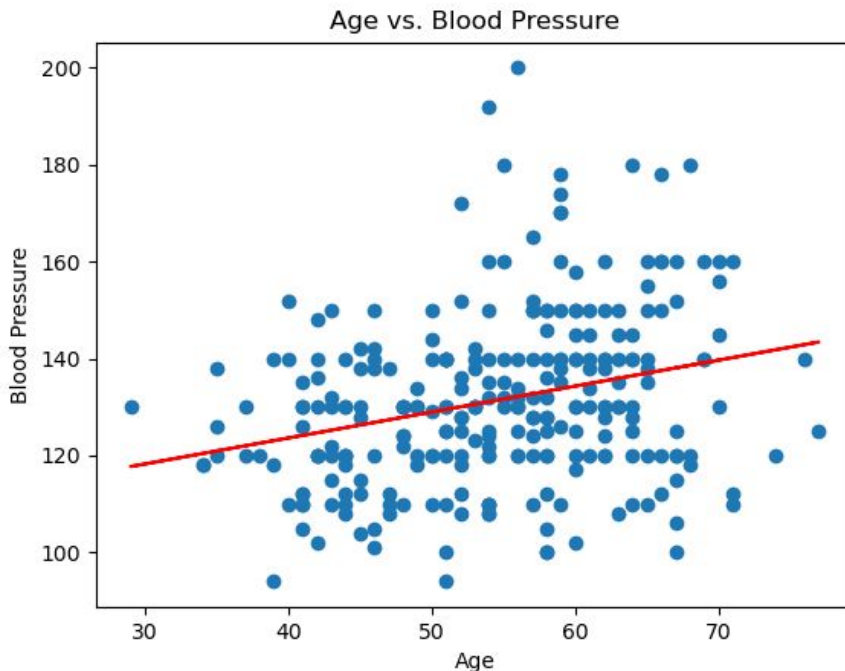
Correlation between age and cholesterol show a trend rising cholesterol following the rise of age.

Factors taken into consideration in regards to outliers are lifestyle, genetics and pre-existing health conditions.

Age vs. Cholesterol
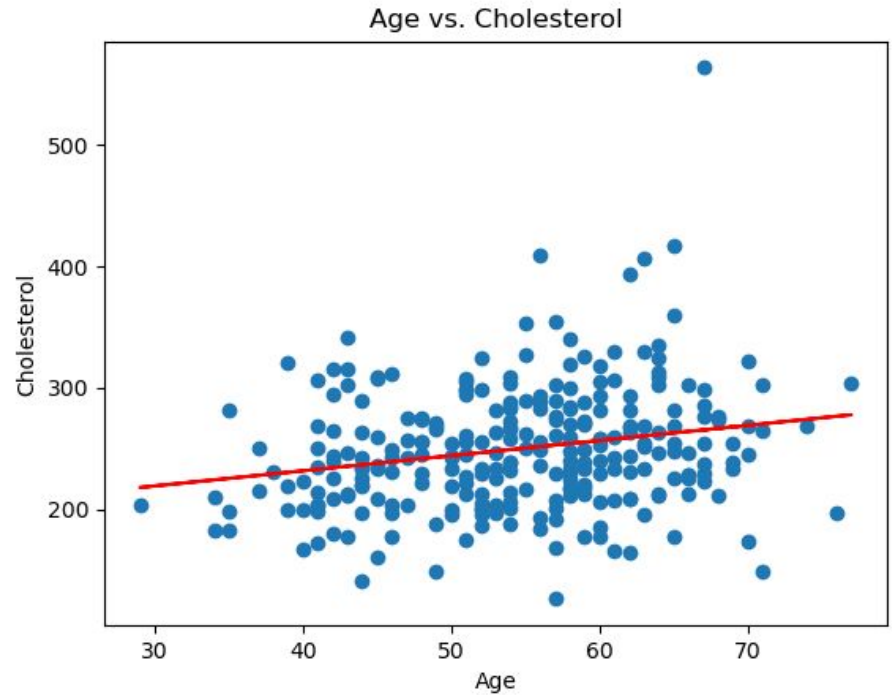
Trend line representation displays indication of less max heart rate the older people become in age.

The severity in max heart rate output (weakness or strength) can be reflection upon an individual's lifestyle and health conditions that aid or hinder someone's heart rate capacity.

Age vs. Max HR
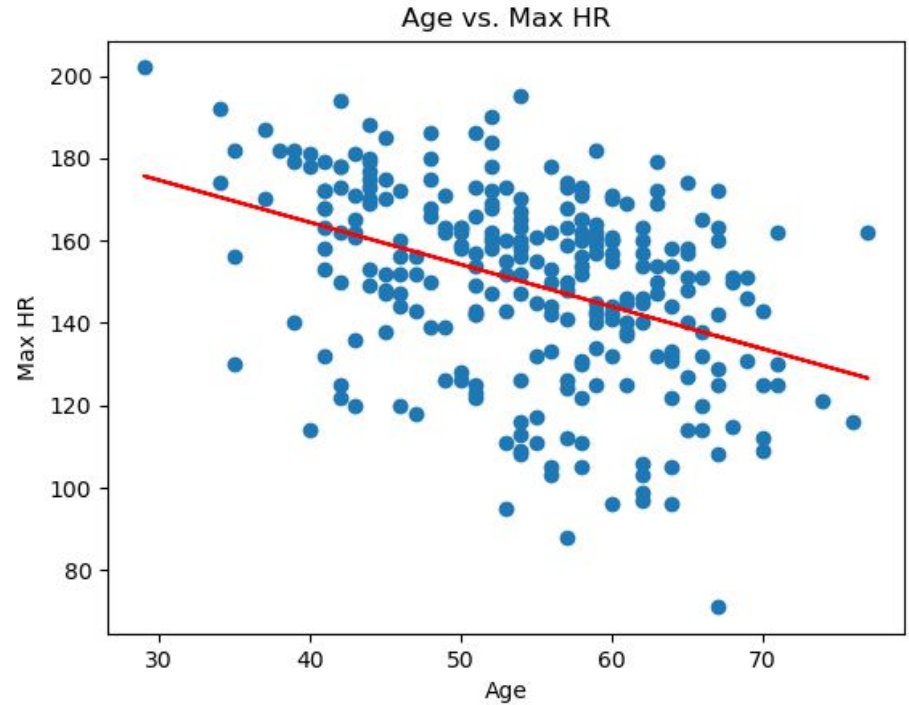
The scatter plot indicates a mix of no change in ST Depression as age increases and also a average range were depression hovers around the ranges of 1.5 to 3.

The mix of outliers tells that not only age but an individual's lifestyle as they get older that we cannot see from the visualisation alone could play a big role in the spike in depression or lack thereof.

But simply, we do see static data points and a spike of data points, many more factors could be involved.

Age vs. ST Depression
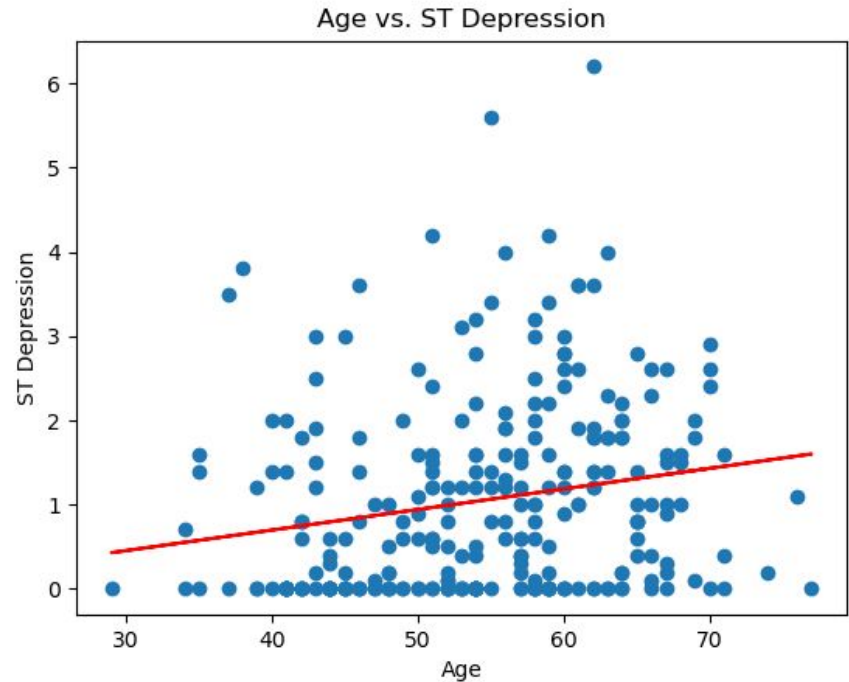
T-test Results

- Number of Vessels Fluro (7.98)
- Chest Pain Type (7.66)
- Max HR (-7.39)
- Exercise Angina (7.30)
- ST Depression (7.17)
- Sex (5.27)
- Age (3.61)
- EKG Results (3.04)
- BP (2.53)
- Cholesterol (1.97)
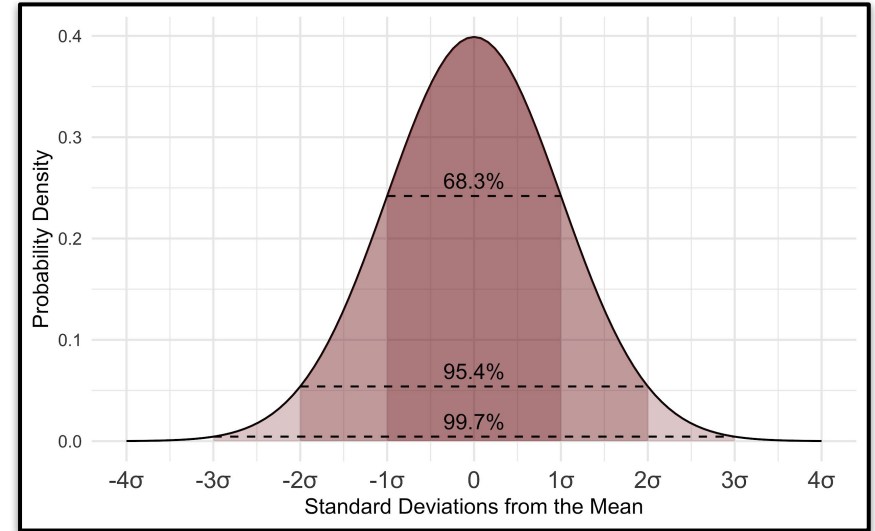
All were significant beyond the p-value of < 0.05, and are listed in order of significance.

Resources

We used Kaggle to get our dataset:

https://www.kaggle.com/datasets/thedevastator/predicting-heart-disease-risk-using-clinical-var

This dataset was provided by Robert Hoyt MD