Analysis of Heart Disease Factors

Nadia Fantello – February 2024

Overview and Problem Statement

- Heart disease is the leading cause of death for men, women, and people of most racial and ethnic groups in the US1.
- ***Objective***:

  Analyze certain factors associated with heart disease to:

  - Prevent and/or decrease the number of positive diagnosis.
  - Decrease the cost of healthcare services, medicines, and lost of productivity due to death.

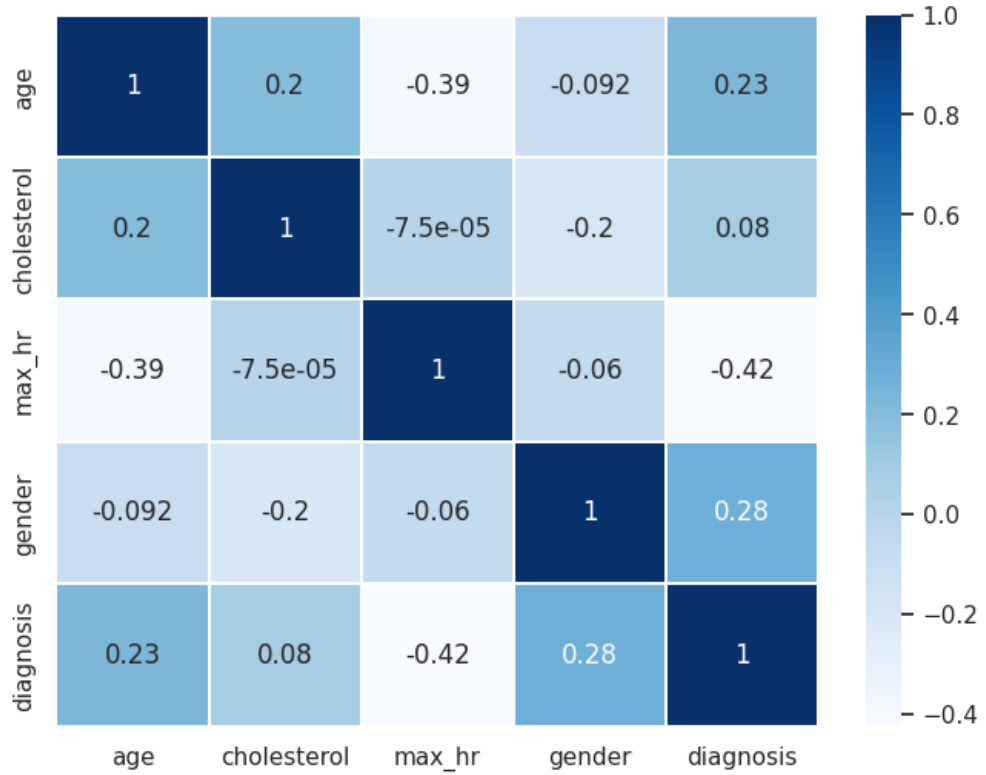
Questions and Dataset

- Factors analyzed: *age*, *cholesterol levels (mg/dl)*, and *max. heart rate*.

1. Is there any correlation between these factors? Which one presents higher correlation with heart disease diagnosis?
2. Do patients with heart disease have higher cholesterol levels than patients without heart disease?
3. Do patients with heart disease are older than patients without heart disease?
4. Do patients with heart disease have higher max. heart rate than patients without heart disease?

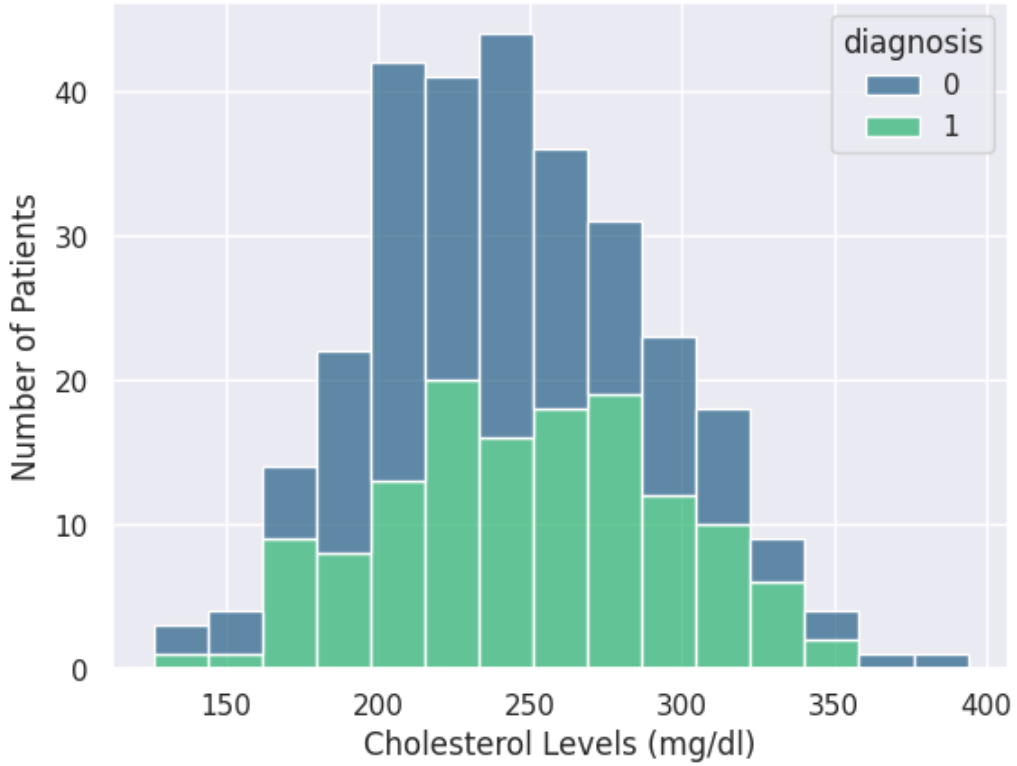
Results – *Correlation of Factors*

- Max. heart rate presents the highest correlation with diagnosis.
- Max. heart rate and cholesterol present the lowest correlation.



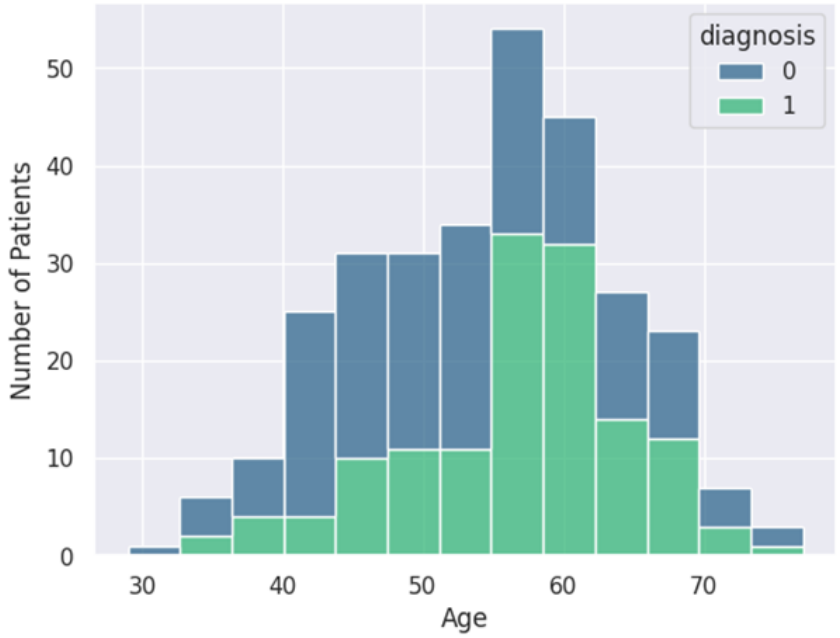
Results – *Cholesterol*

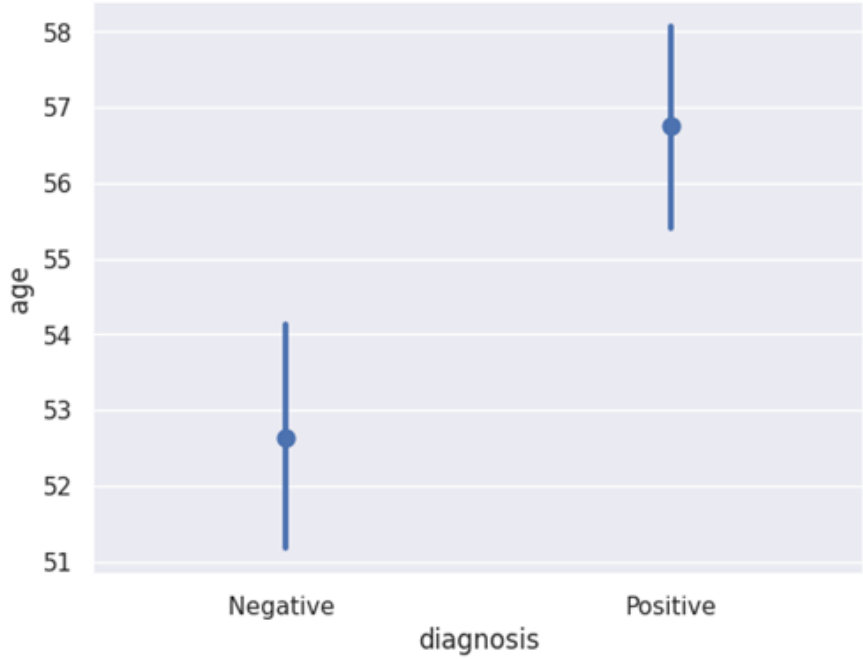
- After running a t-test, we found that we failed to reject the null hypothesis. Therefore, we can conclude that *there is not a significant difference in the means between positive and negative heart disease diagnosis,* for the cholesterol group.



Results – *Age*

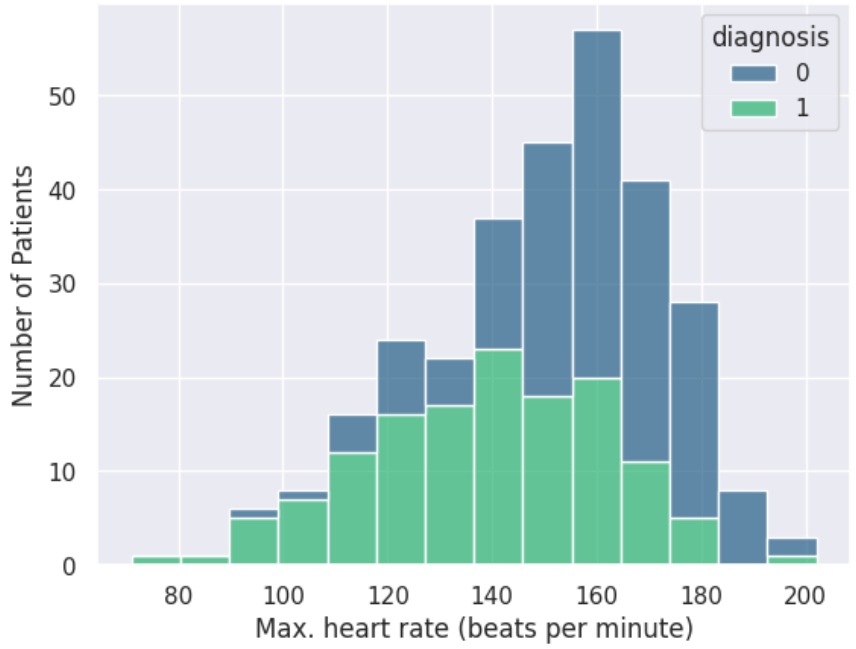
- With 95% confidence, we can say that for the *Age* group, the difference in the average between positive and negative diagnosis is between 2 and 6 years.

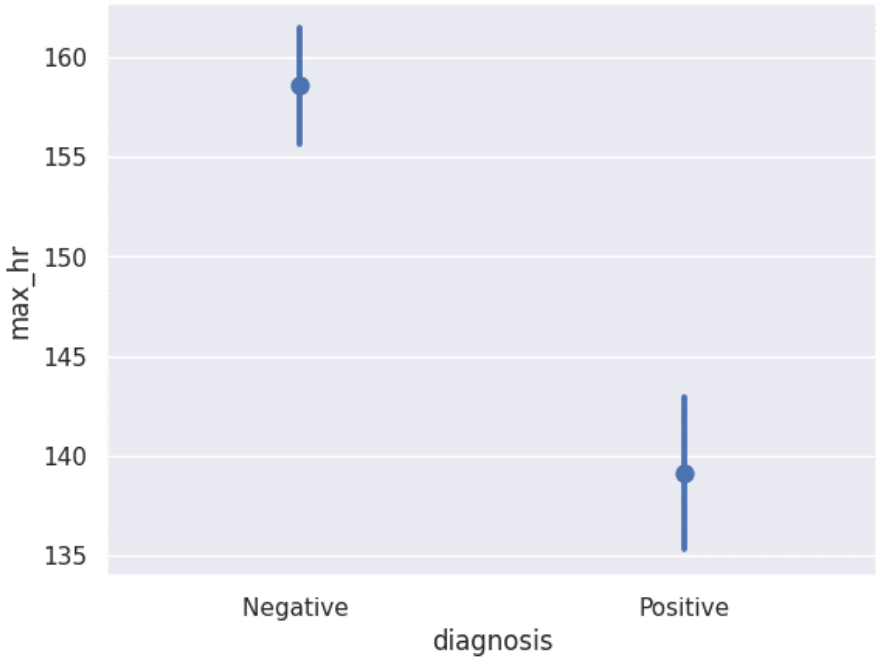




Results – *Max. HR*

- With 95% confidence, we can say that for the *max. hr* group, the difference in the average between positive and negative diagnosis is between 14.66 bpm and 24.28 bpm.





Conclusions and Recommendations

We can conclude that AGE and Max. HR are good metrics for determining or predicting whether someone has heart disease.

Recommendations:

- Increase number of samples
- Add ethnicity/race information to the dataset.