Heart Disease Exploratory Data Analysis

By – Bhawesh Sinha
(Data Scientist)

Introduction

Heart disease remains one of the leading causes of mortality worldwide, imposing a significant burden on public health systems and individuals alike. With its multifaceted nature and diverse risk factors, understanding the epidemiology and patterns of heart disease is crucial for effective prevention, diagnosis, and management strategies. In this exploratory data analysis (EDA) project, we delve into a comprehensive dataset to uncover insights into various aspects of heart disease. By leveraging statistical techniques and data visualization, we aim to elucidate patterns, correlations, and potential predictive factors associated with heart disease. Through this endeavor, we seek to contribute valuable insights that could inform healthcare professionals, policymakers, and individuals in their efforts to combat this pervasive and debilitating condition.

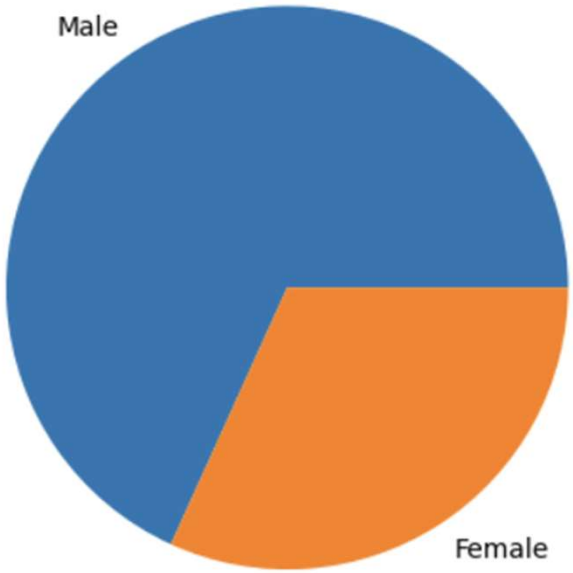
Details about the Data.

1. Age
2. Sex (1= Male, 0=Female)
3. Cp = Chest pain type: Type 0 Angina - Stable Angina Type 1 Angina - Ustable Angina Type 2 Angina - Variant Angina Type 3 Angina - Refractory angina
4. Trestbps = Resting blood pressure
5. Chol = Serum Cholestoral in mg/dl
6. Fbs = fasting blood sugar > 120 mg/dl
7. Restecg = Resting Electrocardiographic Results: Class 0 - Normal Class 1 - Slightly Abnormal Class 2 - Abnomal
8. Thalach = Maximum Heart Rate Achieved
9. Exang = exercise induced angina
10. Oldpeak = ST depression caused by activity in comparison to rest.
11. Slope = Slope of the peak exercise ST Segment.
12. Ca = Calcium Score.
13. Thal = Thalassemia: Value 0 - Normal Value 1 - Fixed Defect Value 2 - Reversable Defect
14. Target = Having Disease or Not: Value 0 - Not Having Heart Disease Value 1 - Having Heart Disease

Main KPIs

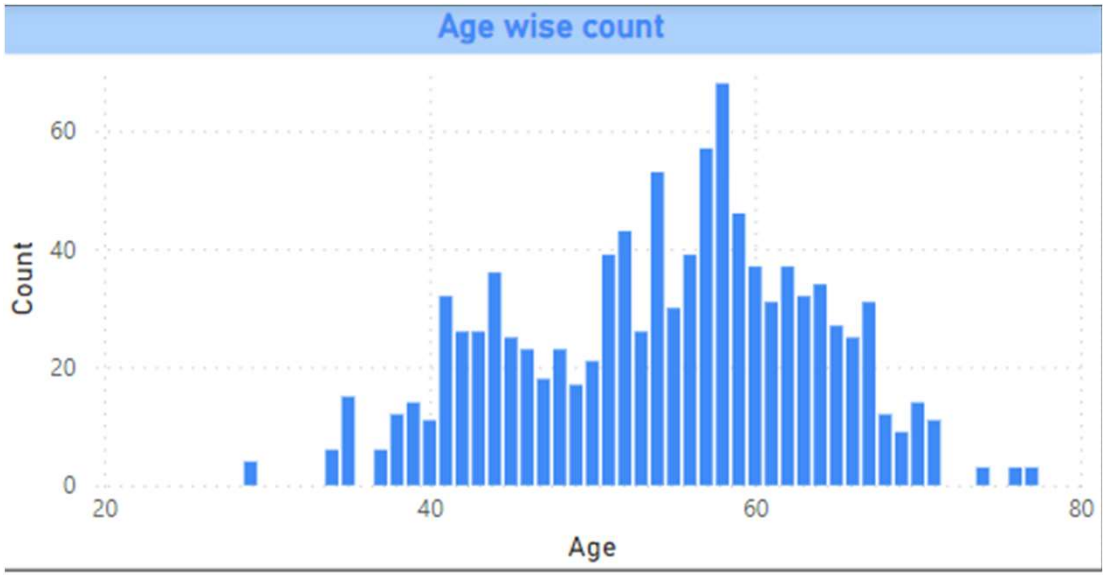
- Age – Average age group having the heart disease.
- Sex – Gender distribution in terms of heart disease.
- Chest pain type – Type of pain they have suffered.
- Target – Are they having the disease or not.

Mock up dashboard



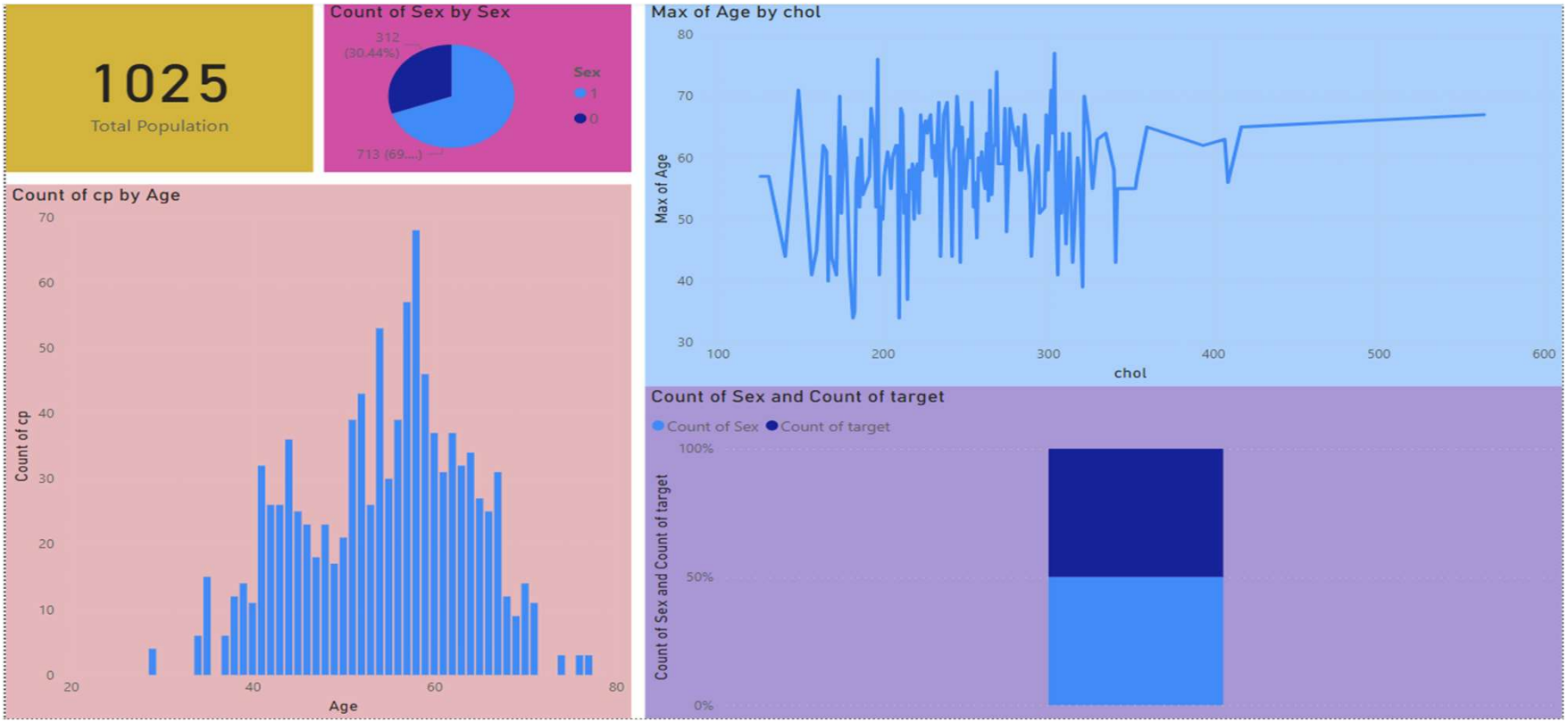
- 70% of the entire survey are Male.
- 30% of the entire survey are Female.

Age Wise Count among the population.



- Maximum people are suffering from the heart disease are falls between age of 50 to 60 years.
- Disease also got captured among the youngster of age less than 30.

My Design


1025
Total Population
Count of Sex by Sex
312
(30.44%)
713 (69....)
Sex
1
0
Max of Age by chol
Max of Age
chol
Count of cp by Age
Count of cp
Age
Count of Sex and Count of target
Count of Sex
Count of target
Count of Sex and Count of target
100%
50%
0%

Thank you