**Analysis & Findings**

### **1. Dataset Overview**

- The dataset contains records about patients, with columns indicating various health metrics and whether or not they have heart disease.

- Number of Rows 303

- Number of Columns 14

### **2. Data Inspection**

- No missing values were found in the dataset.

- There were some duplicate records, which were identified and removed to ensure data integrity.

### **3. Statistical Summary**

- The dataset was summarized using descriptive statistics, providing insights into the distribution, mean, and standard deviation of each column.

### **4. Correlation Analysis**

- A correlation matrix was generated to examine the relationships between variables. Strong correlations were observed between some variables, which might be useful for predictive modeling.

### 

### **5. Heart Disease Distribution**

- The target variable indicates the presence of heart disease.

- People with Heart Disease 165

- People without Heart Disease 138

- A bar chart was plotted to visualize this distribution.

### **6. Gender Distribution**

- The dataset includes a binary gender variable (0 for Female, 1 for Male).

- The majority of the records are male.

- Male 207

- Female 96

- Gender distribution was visualized using a bar chart.

### **7. Heart Disease by Gender**

- An analysis of heart disease prevalence by gender showed that males had a higher incidence of heart disease compared to females.

- A count plot was created to show the distribution of heart disease by gender, highlighting the difference between males and females.

### 

### **Recommendations**

### 

### **1. Targeted Health Interventions**

- Given the higher prevalence of heart disease among males, targeted health interventions and awareness programs should be directed towards men, especially those in high-risk age groups.

### 

### **2. Further Research on Correlated Variables**

- Variables with strong correlations to heart disease should be further studied to understand their impact. This could involve more detailed statistical analysis or even clinical studies.

### **3. Data Quality Improvement**

- Regularly check for and address duplicate entries in health records to maintain data quality and reliability in analysis.

### **4. Customized Healthcare Plans**

- Develop customized healthcare plans based on gender-specific risk factors to improve patient outcomes and prevent heart disease more effectively.

### 

### **5. Predictive Modeling**

- Use the findings from the correlation analysis to build predictive models that can help in early detection of heart disease, potentially improving patient prognosis through early intervention.

These findings and recommendations will help in understanding the current state of heart disease among patients in the dataset and guide strategic decisions for improving healthcare outcomes.