Heart Disease Analysis Project

Data Preprocessing Steps

1. Data Collection

Collected heart disease dataset from a reliable medical dataset source. The dataset includes patient health parameters such as age, gender, blood pressure, cholesterol level, chest pain type, and other medical indicators.

2. Data Cleaning

Removed incorrect, inconsistent, and incomplete records from the dataset to improve data quality and ensure reliable analysis.

3. Handling Missing Values

Checked the dataset for missing or null values. Missing values were replaced using statistical methods such as mean or median, or removed if the missing data was too large.

4. Removing Duplicate Records

Duplicate patient entries were identified and removed to prevent repeated data from affecting analysis and prediction accuracy.

5. Data Transformation

Converted categorical data into numerical format so that it can be used in machine learning and analysis. For example, gender and chest pain categories were converted into numeric values.

6. Data Normalization / Scaling

Scaled numerical features such as cholesterol and blood pressure to a standard range. This ensures that all features contribute equally to the model and improves performance.

7. Feature Selection

Selected the most relevant health parameters that influence heart disease prediction and removed unnecessary or irrelevant columns.

8. Outlier Detection

Identified abnormal or extreme values in the dataset and handled them by removal or adjustment to maintain realistic data distribution.

9. Data Integration

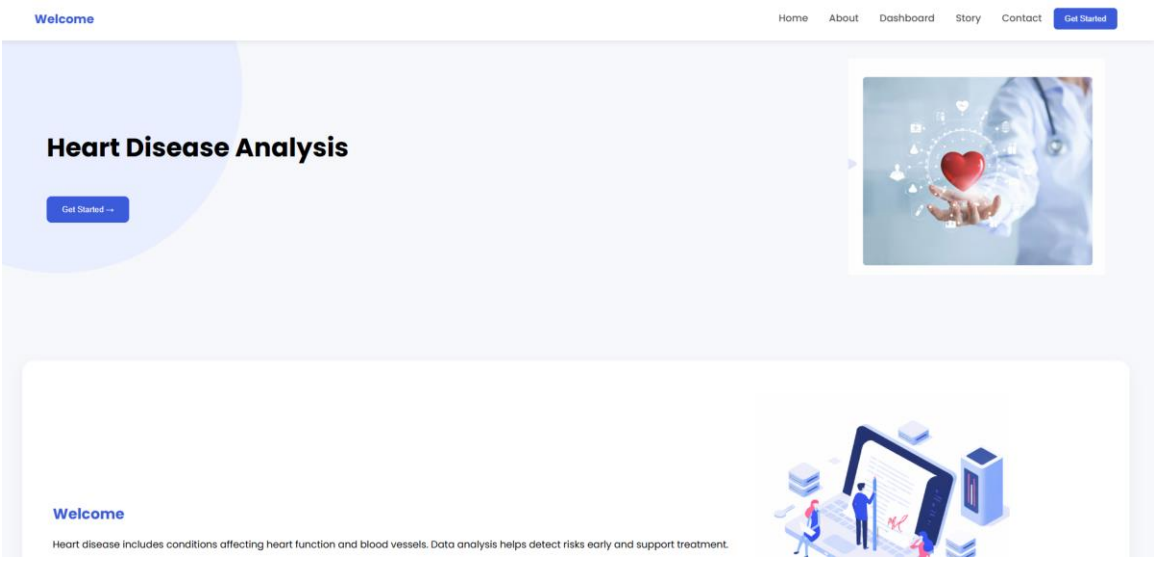
Organized and structured the dataset into a consistent format suitable for visualization, analysis, and prediction.

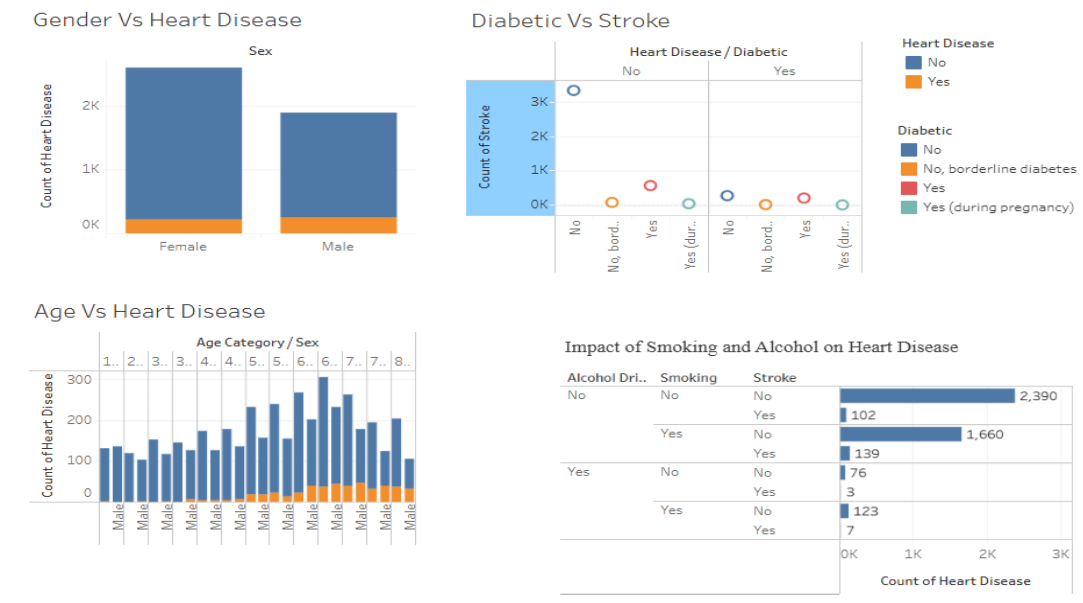
10. Data Splitting

Divided the dataset into training and testing sets. The training set is used to train the model, and the testing set is used to evaluate performance.

11. Data Validation

Performed final verification to ensure the dataset is clean, consistent, and ready for analysis and model development.

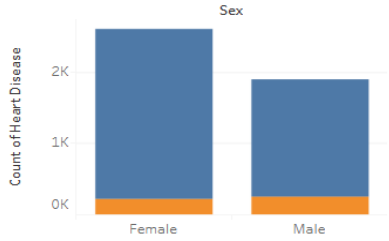




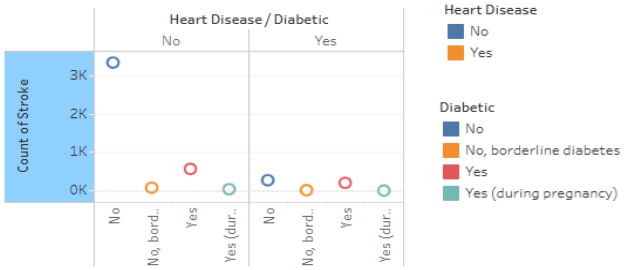
Story 1

Heart Analysis

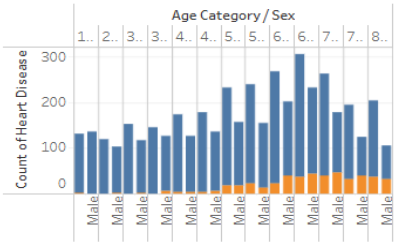
Gender Vs Heart Disease



Diabetic Vs Stroke



Age Vs Heart Disease



Impact of Smoking and Alcohol on Heart Disease