

Team:S.Meghasyam,R.Govardhan Reddy

Team id:593524

Visualizing And Predicting Heart Diseases with An Interactive Dashboard

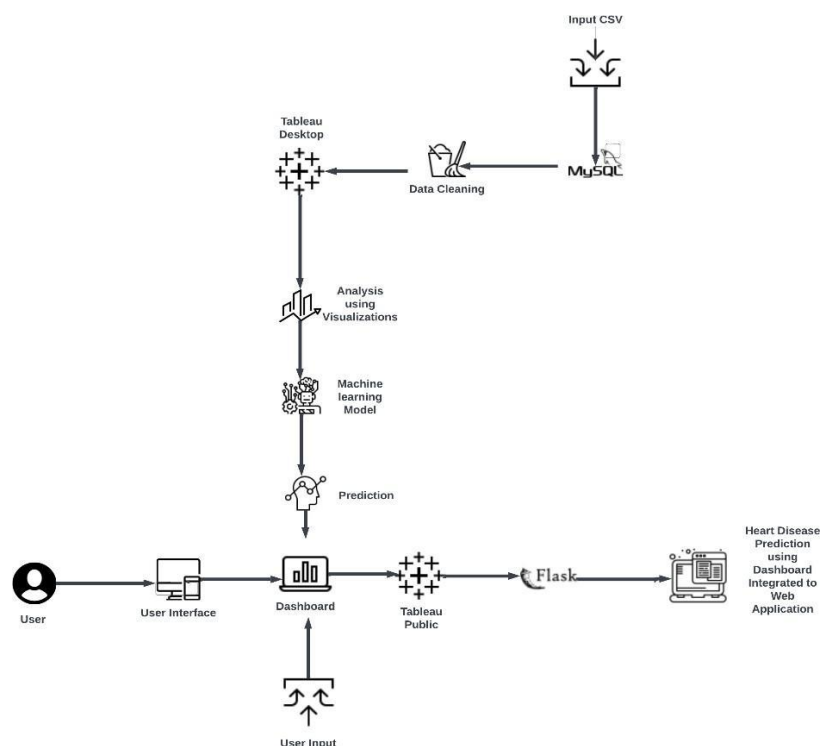
Introduction:

Our project, "Visualizing and Predicting Heart Diseases with an Interactive Dashboard," seeks to address the critical global health issue of cardiovascular diseases. These diseases, which encompass conditions like heart attacks and strokes, are the leading cause of mortality worldwide, imposing a substantial burden on healthcare systems and society at large. The key challenge is the timely identification and accurate prediction of heart diseases to enable effective intervention and prevention.

To tackle this problem, our project employs a multifaceted approach that combines advanced technology with medical expertise. At its core is an interactive dashboard, meticulously designed to integrate cutting-edge data visualization, sophisticated machine learning algorithms, and real-time health monitoring. This innovative tool not only empowers healthcare professionals with advanced diagnostic capabilities but also enables individuals to actively engage in their own cardiovascular health management.

Within the following pages, this document provides a comprehensive examination of our project. We delve into the technology underpinning our initiative, elucidate the sources of data used, describe the development of predictive models, and explore the design and functionality of the interactive dashboard. By the conclusion of this report, readers will have gained profound insights into how our groundbreaking solution contributes to the early identification and prevention of heart diseases, potentially saving numerous lives and offering a significant contribution to mitigating the global burden of cardiovascular diseases.

Technical Architecture:



Project Objectives:

By the end of this project you will:

1. Develop an interactive dashboard for data visualization.
2. Create accurate predictive models for heart diseases.
3. Implement real-time health monitoring features.
4. Utilize advanced data visualization techniques.
5. Enhance accessibility and provide educational resources.

Project Flow:

To accomplish this, we have to complete all the activities listed below,

- Define Problem / Problem Understanding
 - Specify the business problem
 - Business requirements
 - Literature Survey
 - Social or Business Impact
- Data Collection & Extraction from Database
 - Collect the dataset,
 - Storing Data in DB
 - Perform SQL Operations
 - Connect DB with Tableau
- Data Preparation
 - Prepare the Data for Visualization
- Data Visualizations
 - No of Unique Visualizations
- Dash board
 - Responsive and Design of Dashboard
- Story
 - No of Scenes of Story
- Performance Testing
 - Amount of Data Rendered to DB
 - Utilization of Data Filters
 - No of Calculation Fields
 - No of Visualizations/ Graphs
- Web Integration
 - Dashboard and Story embed with UI With Flask
- Project Demonstration & Documentation
 - Record explanation Video for project end to end solution
 - Project Documentation-Step by step project development procedure

Milestone 1: Define Problem

Heart disease (heart disease) is a group of diseases related to cardiovascular diseases, manifested by a violation of the normal functioning of the heart. May be caused by damage to the epicardium, pericardium, myocardium, endocardium, valvular apparatus of the heart, heart vessels.

According to the National Heart, Lung and Blood Institute in Framingham (USA), the most important factors in the development of cardiovascular disease in humans are obesity, sedentary lifestyle and smoking.

In this project we are trying to analyse the Heart Disease related data and be able to extract some insights from the data using Business Intelligence tools. To Extract the Insights from the data and put the data in the form of visualizations, Dashboards and Story we employed Tableau tool.

Milestone 2: Data Collection

Data collection for this project involves gathering medical records, patient demographics, and real-time health data. We acquire anonymized patient records from healthcare institutions, integrate wearable device data, and utilize surveys to collect lifestyle information. This multifaceted approach ensures a comprehensive dataset for accurate heart disease prediction and visualization.

Data set: [link](#)

Column Description of the Dataset:

1. Heart Disease - target trait.
2. BMI – A value that allows you to assess the degree of correspondence between a person's mass and his height, and thereby indirectly judge whether the mass is insufficient, normal or excessive. It is important in determining the indications for the need for treatment.
3. Smoking: It is a major risk factor for cardiovascular disease. When smoke from a cigarette is inhaled, the reaction of the cardiovascular system immediately follows: within one minute, the heart rate begins to rise, increasing by 30% within ten minutes of smoking. The bad habit also increases blood pressure, fibrinogen and platelet levels, making blood clots more likely.
4. Alcohol Drinking - alcohol causes not only temporary disturbances in the functioning of the heart, but also permanent ones. Heart pain after alcohol is not the only health problem associated with alcohol consumption.
5. Stroke - Ischemic stroke occurs 4 times more often than hemorrhagic. One of the leading causes of this suffering is heart disease, which impairs its functioning, as a result of which the blood flow in the arteries is disturbed and the blood supply to the brain is reduced. Another cause of stroke in heart disease is thromboembolism, when clots form in the cavities of the heart (most often with heart failure) - blood clots.
6. Physical Health - how many days in a month did you feel poor physical health.
7. Mental Health - how many days in a month did you feel poor mental health.
8. Diff Walking - difficulty climbing stairs.
9. Sex - gender of a person.
10. Age Category - age category of the subjects.
11. Race- Race is a complex social construct that categorizes people into distinct groups based on certain physical and genetic characteristics
12. Diabetic – Person suffering from Diabetes
13. Physical Activity - adults who reported doing physical activity or exercise during the past 30 days other than their regular job
14. Gen Health - well-being.

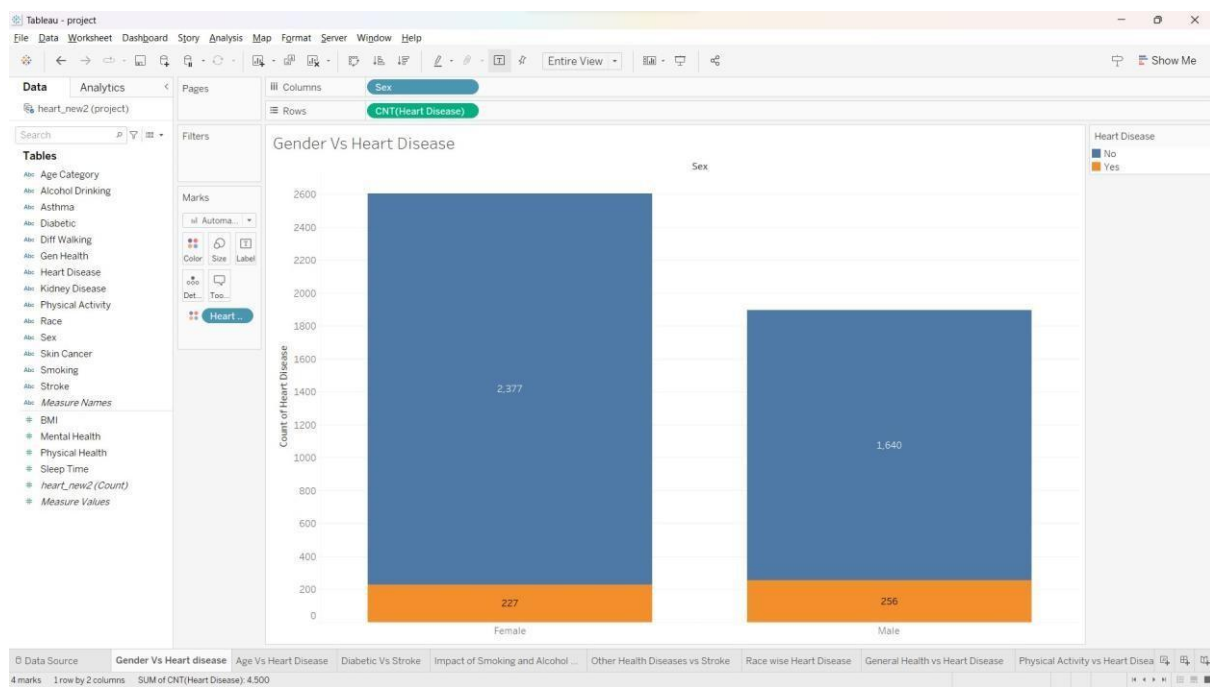
15. Sleep Time - number of hours of sleep.
16. Asthma- Asthma is a chronic respiratory condition due to breathing Issue
17. Kidney Disease – Disease related to Kidney
18. Skin Cancer – People suffering from Skin Cancer

Milestone 3: Data Preparation

The data preparation process for visualization involves several key steps. First, we clean the raw data by addressing missing values, outliers, and inconsistencies. Next, we perform feature engineering to extract relevant variables and create new features for predictive models. Then, we normalize and scale the data to ensure consistency and remove any biases. Subsequently, we apply dimensionality reduction techniques, such as PCA (Principal Component Analysis), to simplify complex datasets. Finally, the prepared data is ready for visualization, where we employ advanced techniques like data aggregation, encoding, and transformation to create informative and user-friendly visual representations in our interactive dashboard. This meticulous data preparation ensures accurate and meaningful insights for heart disease analysis.

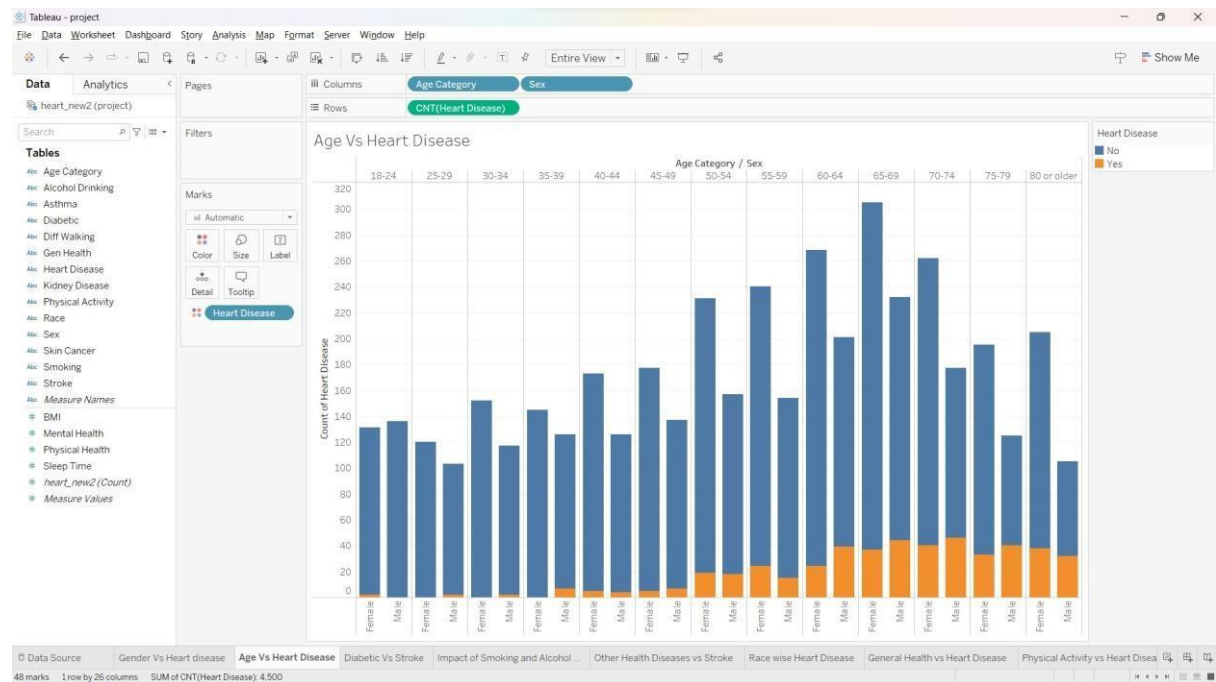
Milestone 4: Data Visualization

Gender vs Heart Disease



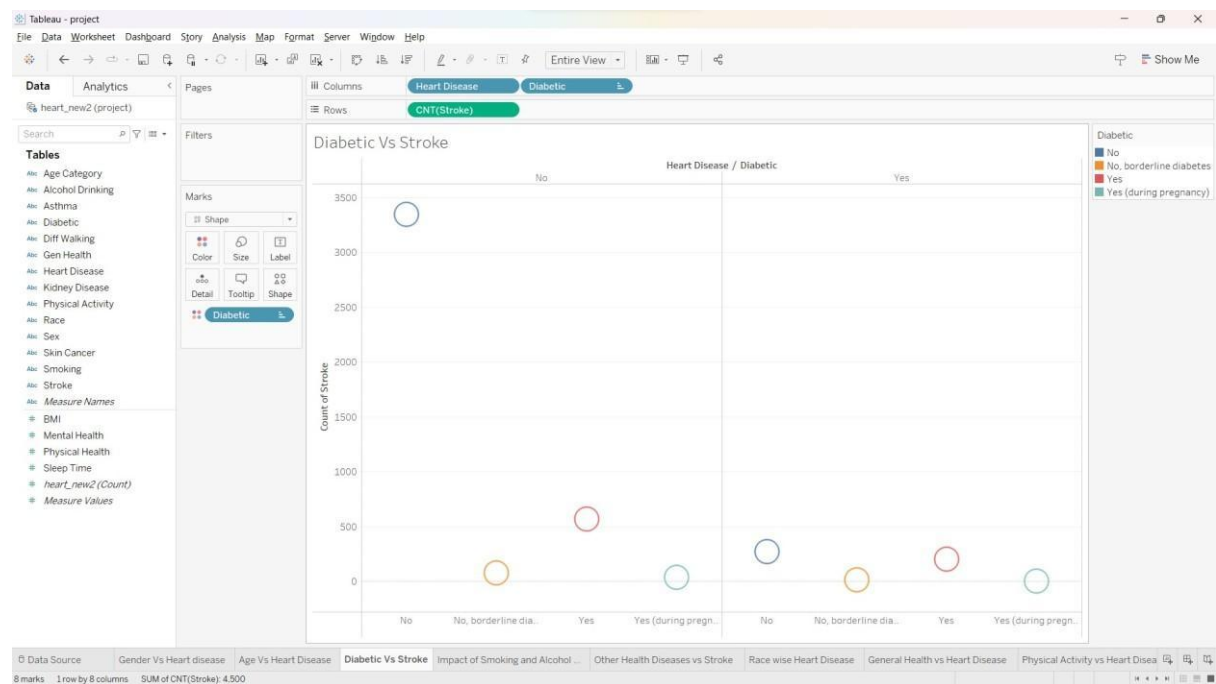
<https://drive.google.com/file/d/1RHJmAZ1iHLLYWBcBRRMX3mtdzc6D02HN/view>

Age vs Heart Disease



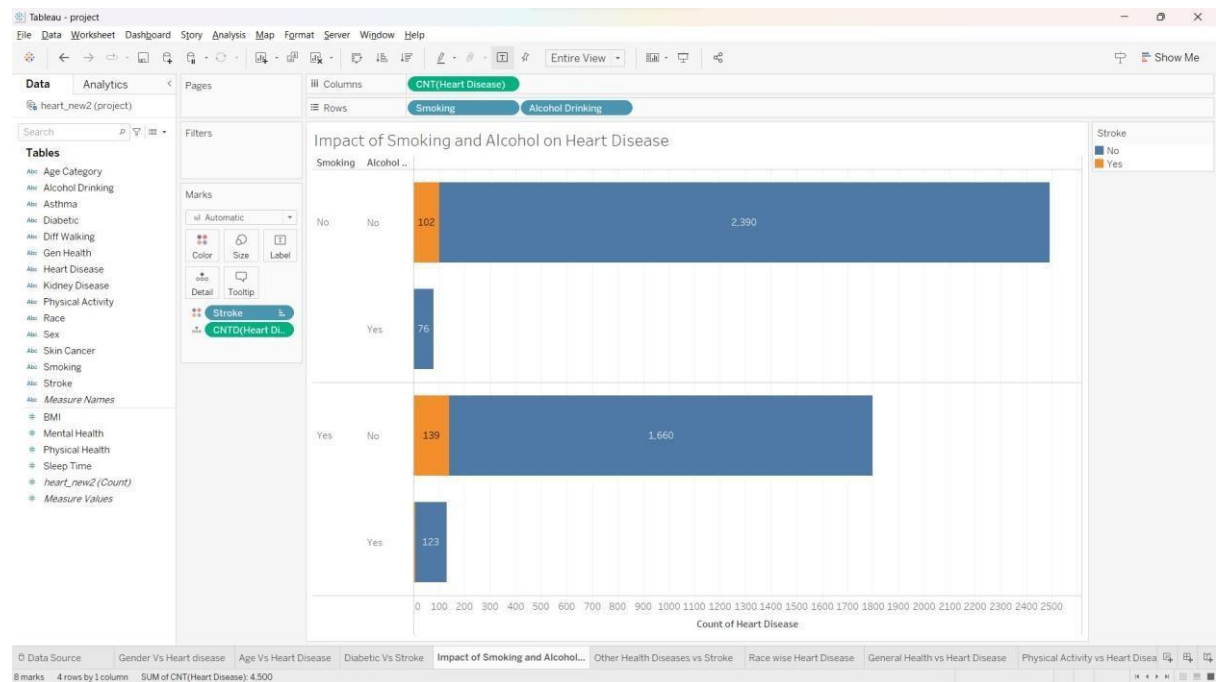
<https://drive.google.com/file/d/10aw8cOcdjjDCj7hRrzmZYJAnhM6YfIKl/view>

Diabetic vs Stroke



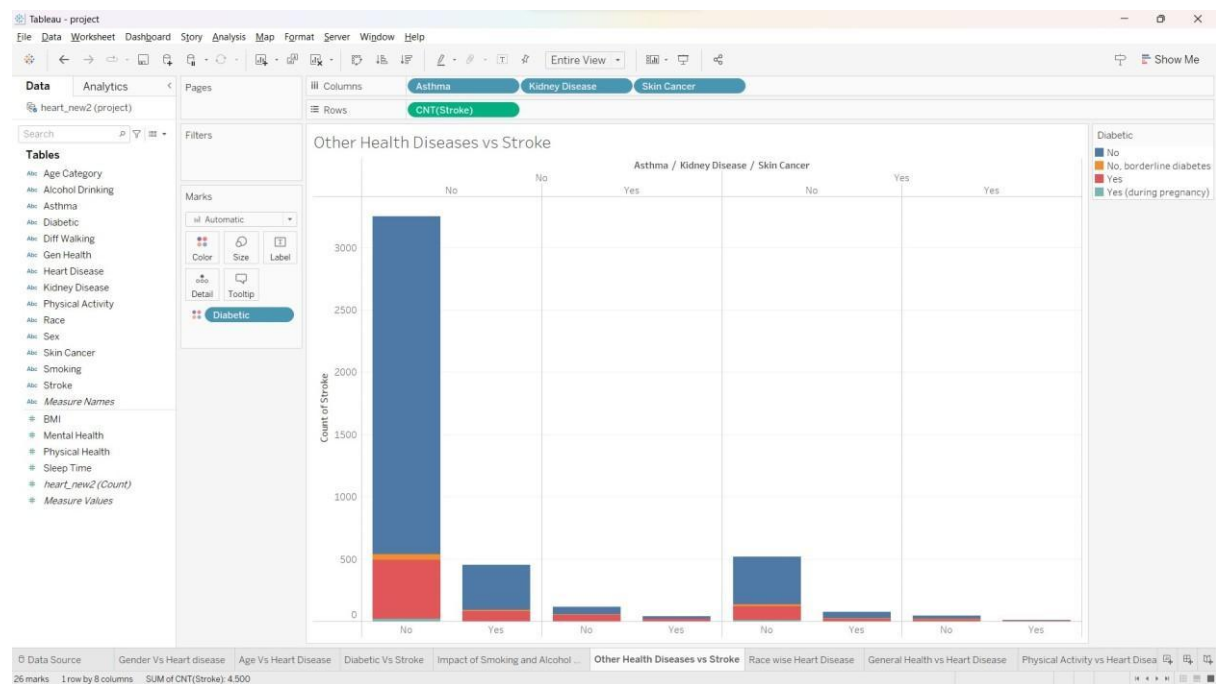
<https://drive.google.com/file/d/1XFgAYyIOmuKdAKsOsZKrSvA2MURFU9zb/view>

Impact of Smoking and Alcohol on Heart Disease



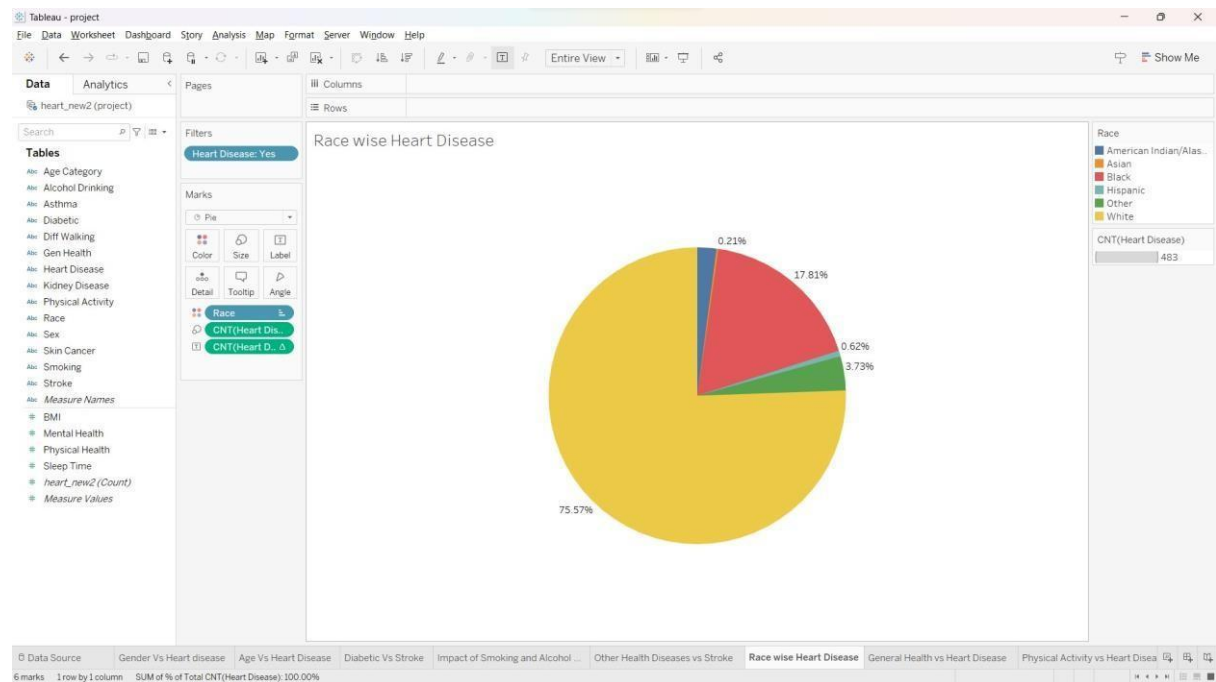
https://drive.google.com/file/d/1tHtuhgsuDOWB9enaP2Y-JGBs_vC3bJSZ/view

Health disease vs Stroke



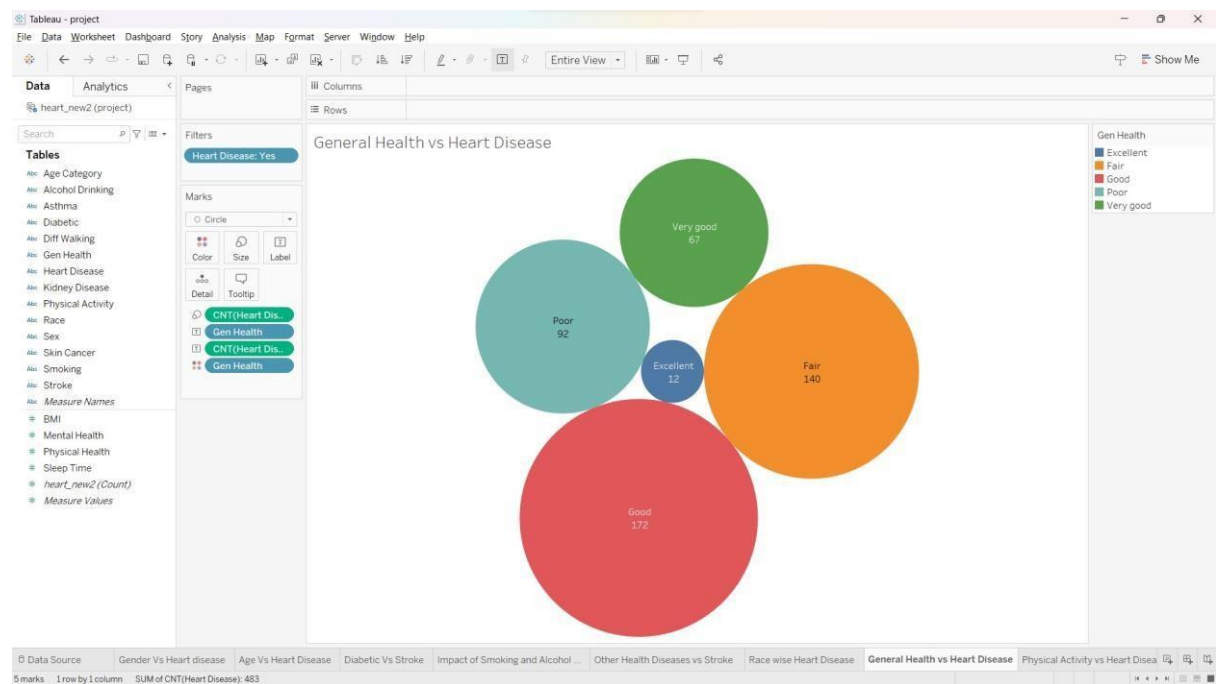
<https://drive.google.com/file/d/14KJmZpP2r8qzgWCLALRRQOGxE-eZtMNH/view>

Race Wise Heart Disease



<https://drive.google.com/file/d/1gHbNq6KdGUQ8RI7Umods1nODKiGyYaVt/view>

General Health vs Heart Disease



https://drive.google.com/file/d/1vRv9fjvQIHrRaN_Nw7Gu2PYs-XlrQLPm/view

Tableau - project

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Search heart_new2 (project)

Tables

- Age Category
- Alcohol Drinking
- Asthma
- Diabetic
- Diff Walking
- Gen Health
- Heart Disease
- Kidney Disease
- Physical Activity
- Race
- Sex
- Skin Cancer
- Smoking
- Stroke
- Measure Names
- BMI
- Mental Health
- Physical Health
- Sleep Time
- heart_new2 (Count)
- Measure Values

Columns

Rows

AGG(avg(0)) AGG(avg(0))

Filters

Heart Dise...

Marks

All

Physical Activity vs Heart Disease.

Physical Activity

- No
- Yes

CNT(Heart Disease)

483

279 204

0 avg(0)

0 avg(0)

0 Data Source nder Vs Heart disease Age Vs Heart Disease Diabetic Vs Stroke Impact of Smoking and Alcohol Other Health Diseases vs Stroke Race wise Heart Disease General Health vs Heart Disease Physical Activity vs Heart Dise...

3 marks 1 row by 1 column SUM of AGG(avg(0))

Tableau - project

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Search heart_new2 (project)

Tables

- Age Category
- Alcohol Drinking
- Asthma
- Diabetic
- Diff Walking
- Gen Health
- Heart Disease
- Kidney Disease
- Physical Activity
- Race
- Sex
- Skin Cancer
- Smoking
- Stroke
- Measure Names

Filters

Diabetic

Marks

Automatic

Color Size Label

Detail Tooltip

AVG(BMI)

AVG(BMI)

Age Category

Age and BMI vs Diabetic

AVG(BMI)

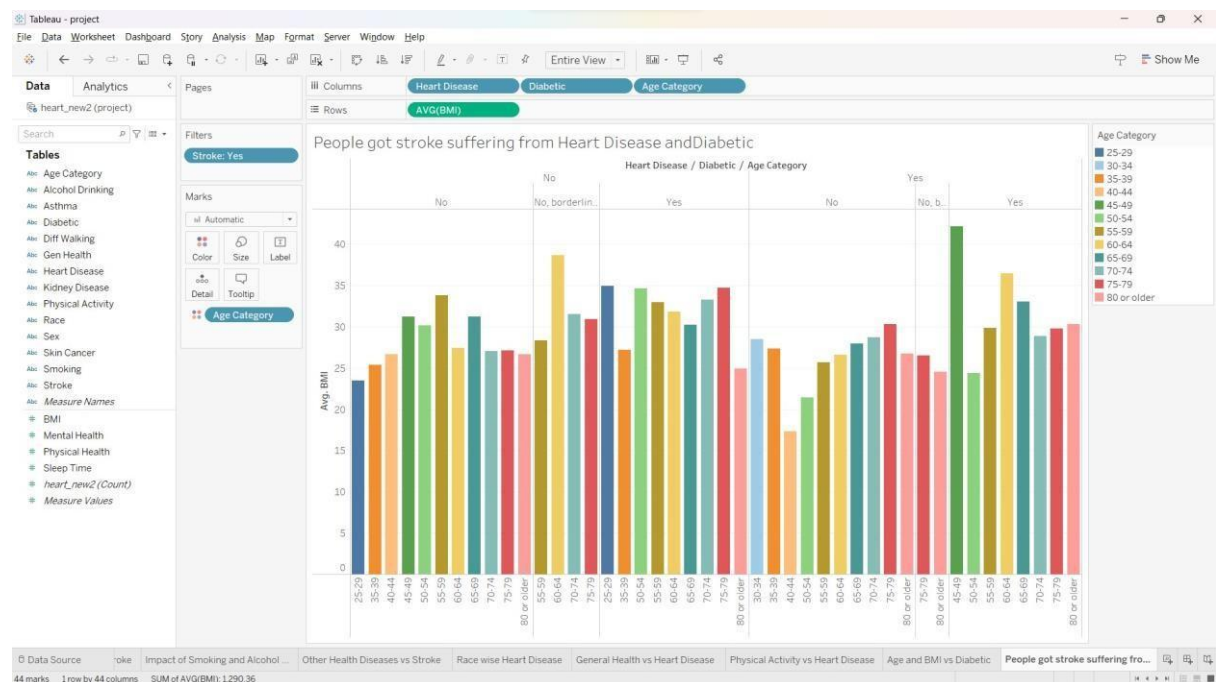
28.92 39.76

Age	BMI
18-24	30-34
18-24	35-39
18-24	40-44
18-24	45-49
18-24	50-54
18-24	55-59
18-24	60-64
18-24	65-69
18-24	70-74
18-24	75-79
18-24	80 or older

© Data Source | e Vs Heart Disease | Diabetic Vs Stroke | Impact of Smoking and Alcohol ... | Other Health Diseases vs Stroke | Race wise Heart Disease | General Health vs Heart Disease | Physical Activity vs Heart Disease | Age and BMI vs Diabetic

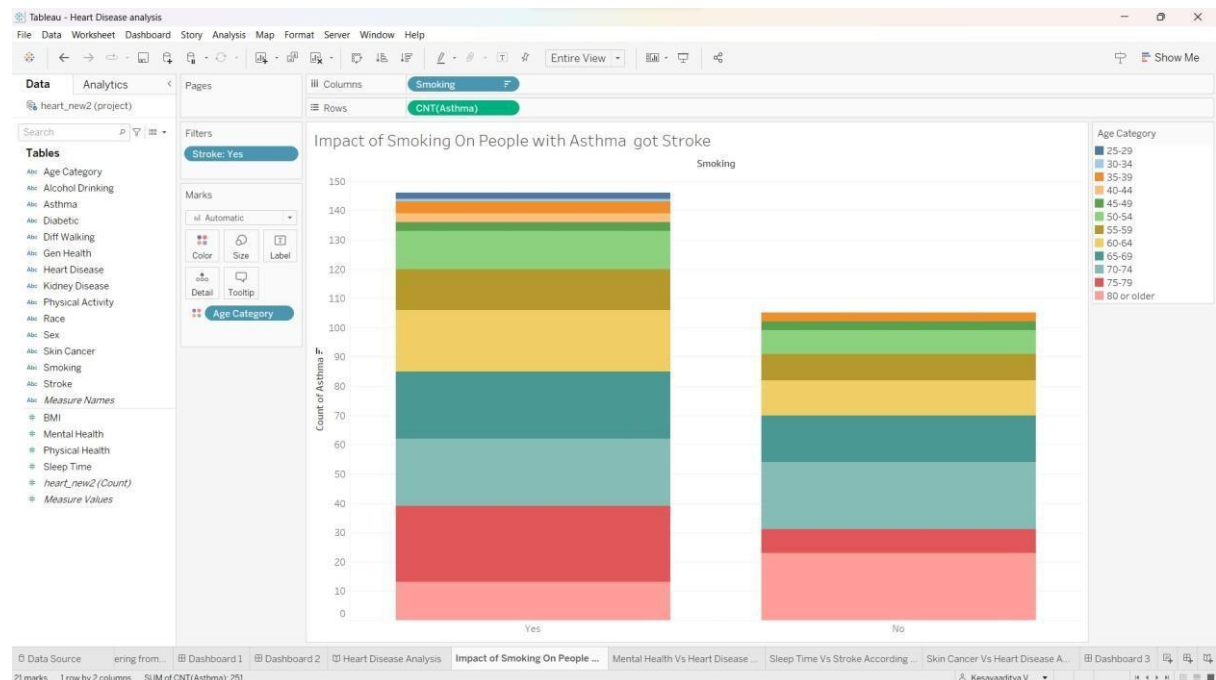
13 marks | 1 row by 1 column | SUM of AVG(BMI): 430.90

People got Storke suffering from Heart Disease and Diabetic



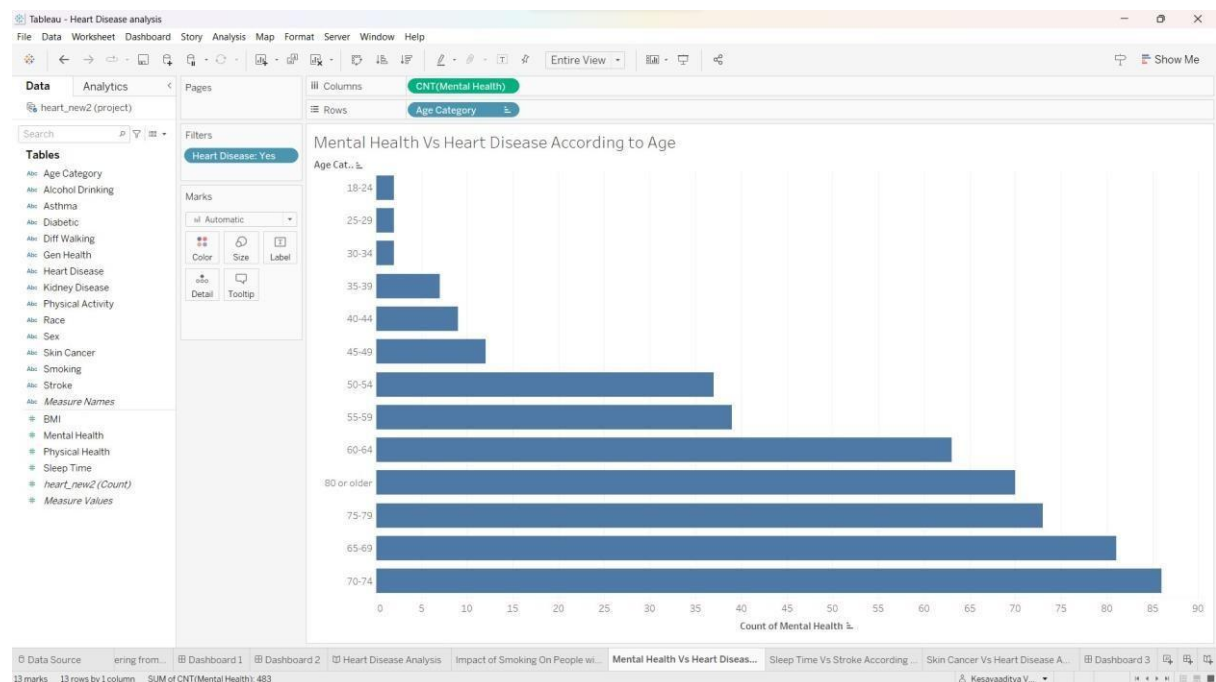
<https://drive.google.com/file/d/193mcmvsoTdawIKLwPLl4aRR71Slax-Ag/view>

Impact of Smoking on People with Asthma got Stroke



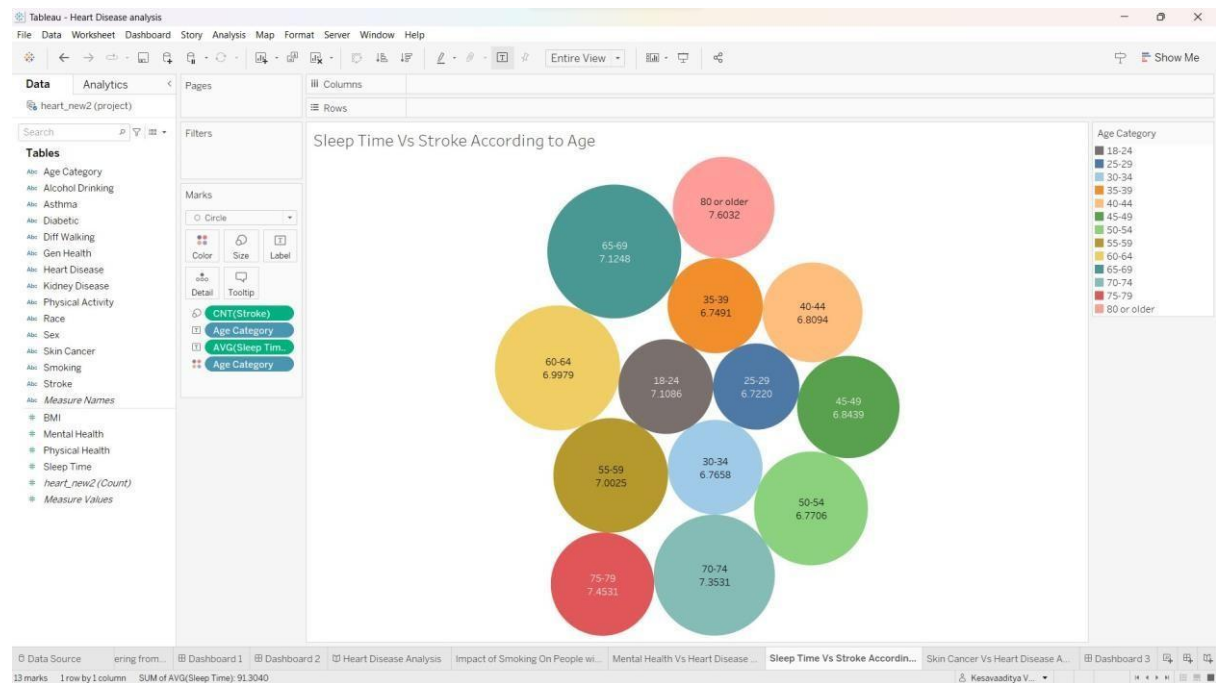
<https://drive.google.com/file/d/1cO1RuOijbwWh1D-hGsmcOCi-jMpwElD0/view>

Mental health vs Heart Disease according to Age



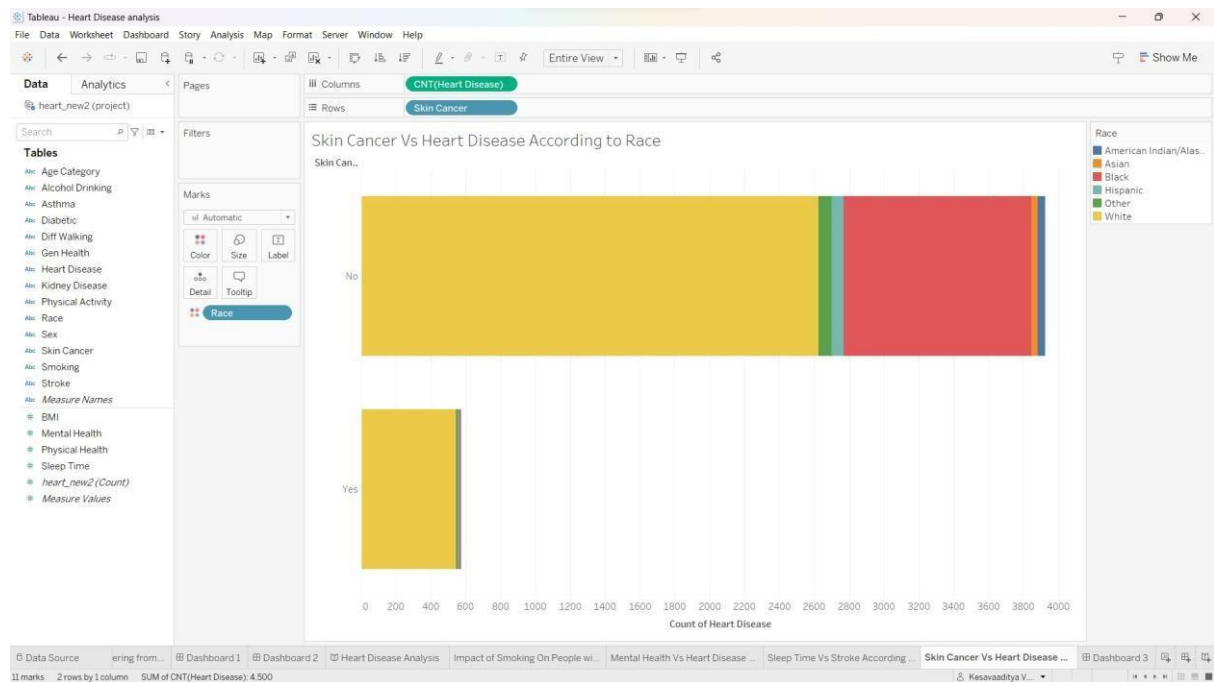
https://drive.google.com/file/d/1tMsp_5ZF7whb4001uYtxSHiOXd5S9XqT/view

Sleep Time vs Stroke according to Age



<https://drive.google.com/file/d/1Se7zbgpbU98nabnt3D4HEverbvzGnr-C/view>

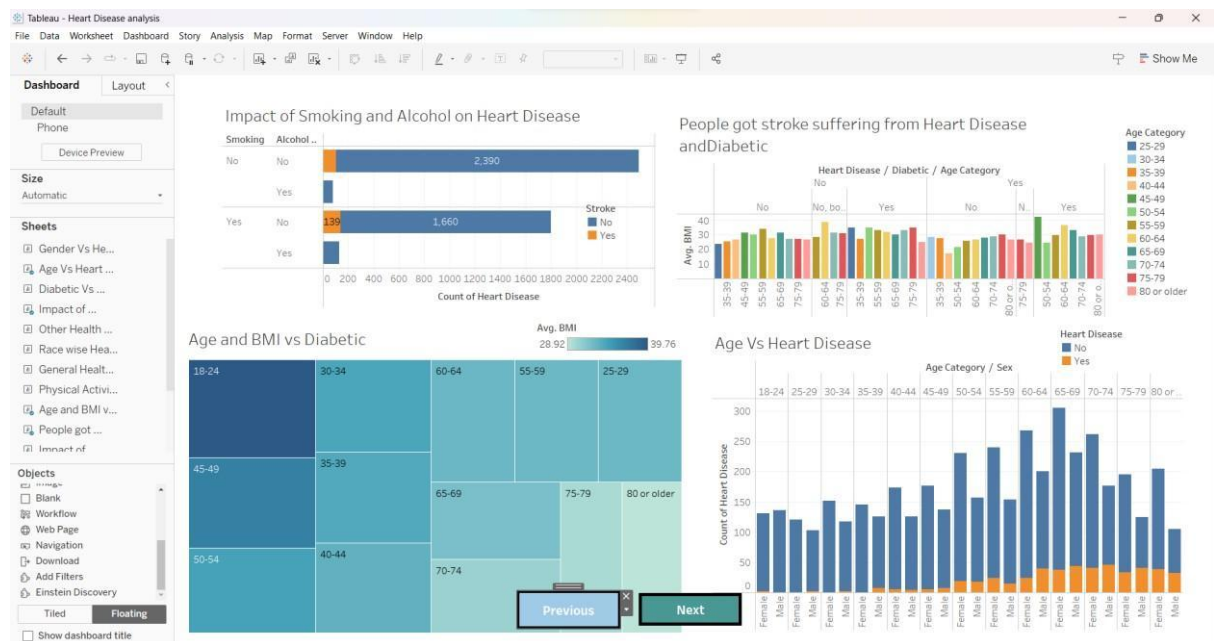
Skin Cancer vs Heart Disease according to Race



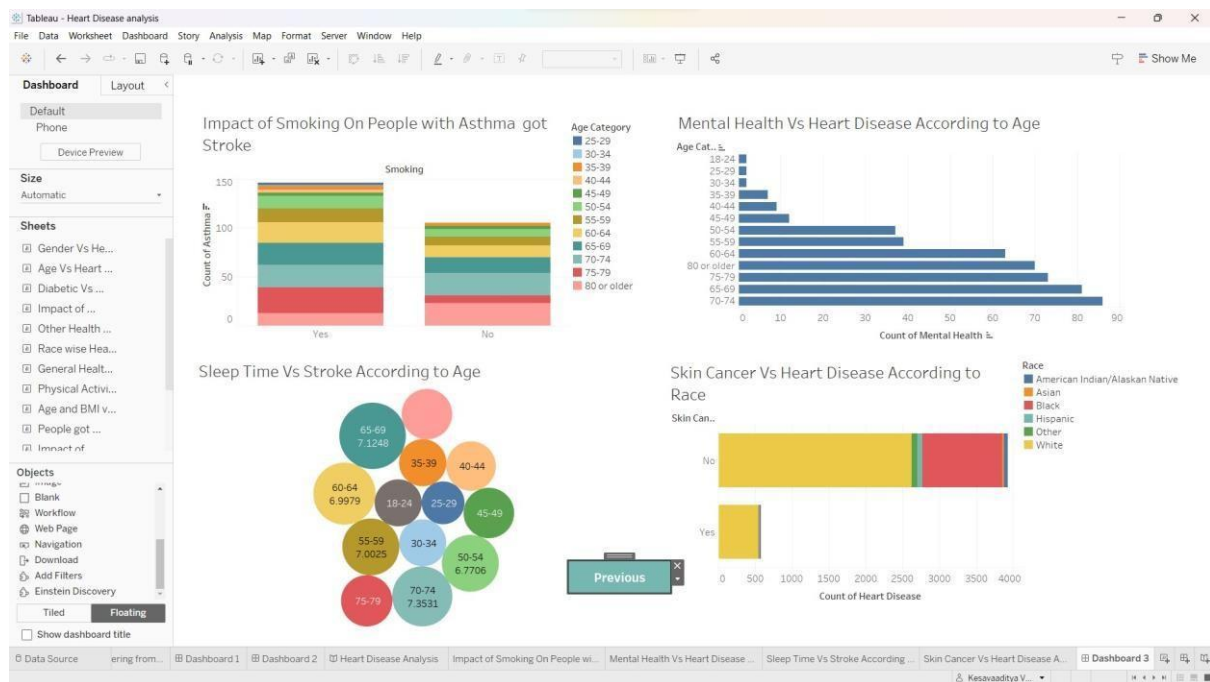
https://drive.google.com/file/d/1B_k-H_FhZpiDINHGKcN84aaw8fEzLQnf/view

Milestone 5: Dashboard

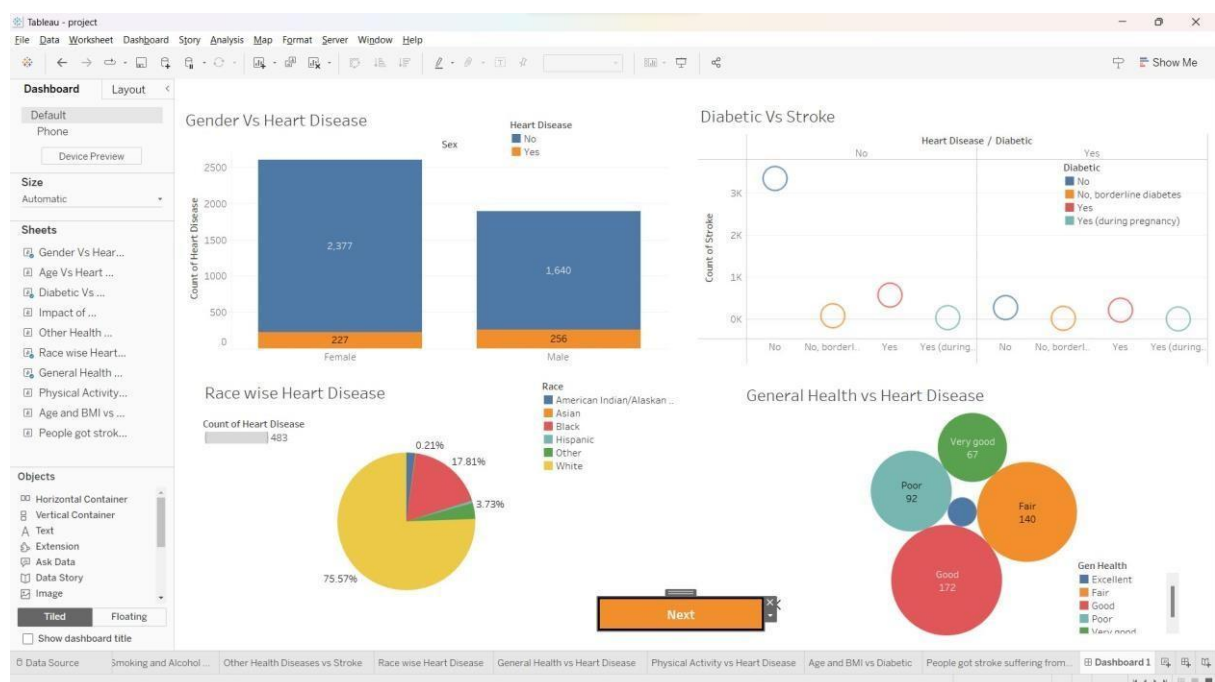
A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.



<https://drive.google.com/file/d/1ECaTgicdzCFArFbWCm9dHnXkPgOqGw3/view>



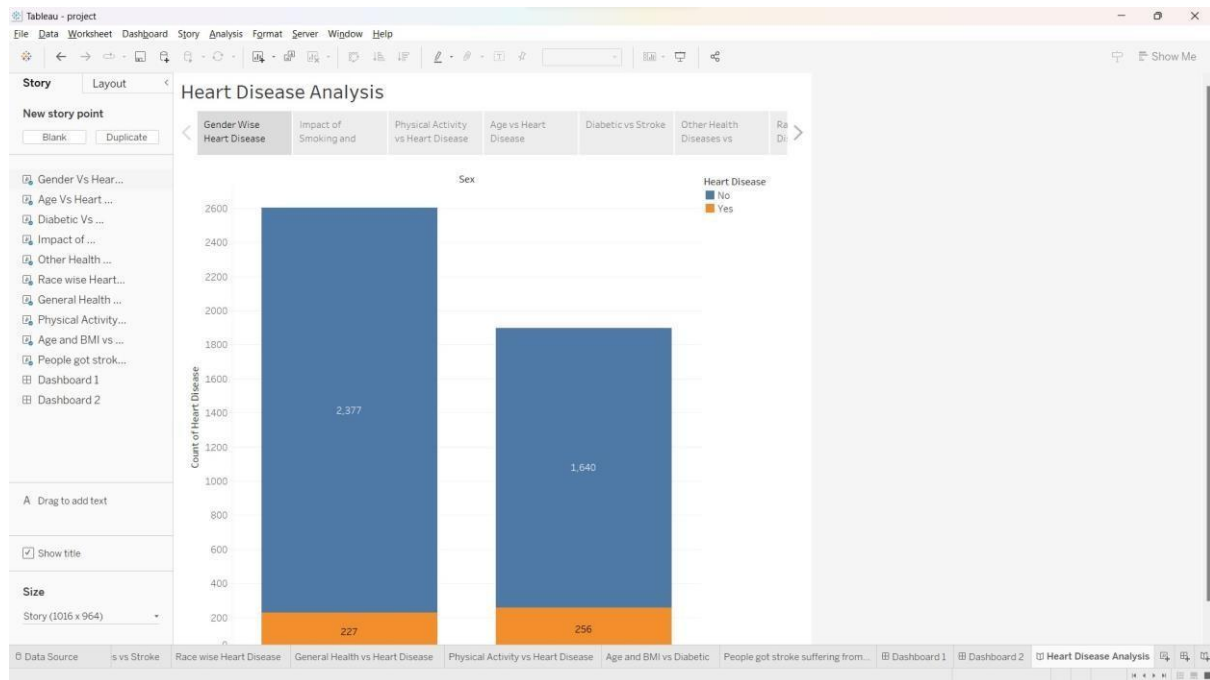
<https://drive.google.com/file/d/1hjz1YP|xvbUC3q8MAj696By|ZOaw|A1S/view>



<https://drive.google.com/file/d/1ZAGVtxYmBmxlja|wsvlGW-rZQlBzhe7/view>

Milestone 6: Story

A data story is a way of presenting data and analysis in a narrative format, intending to make the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis logically and systematically, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.



<https://drive.google.com/file/d/123de9p4h0AmB2zZhs4TedvyjNzIzTloK/view>

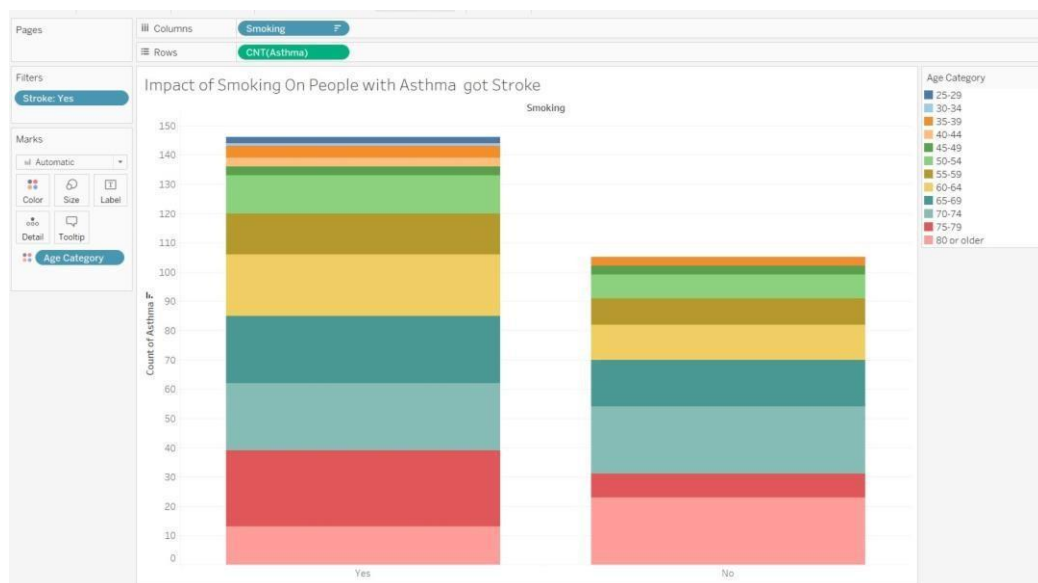
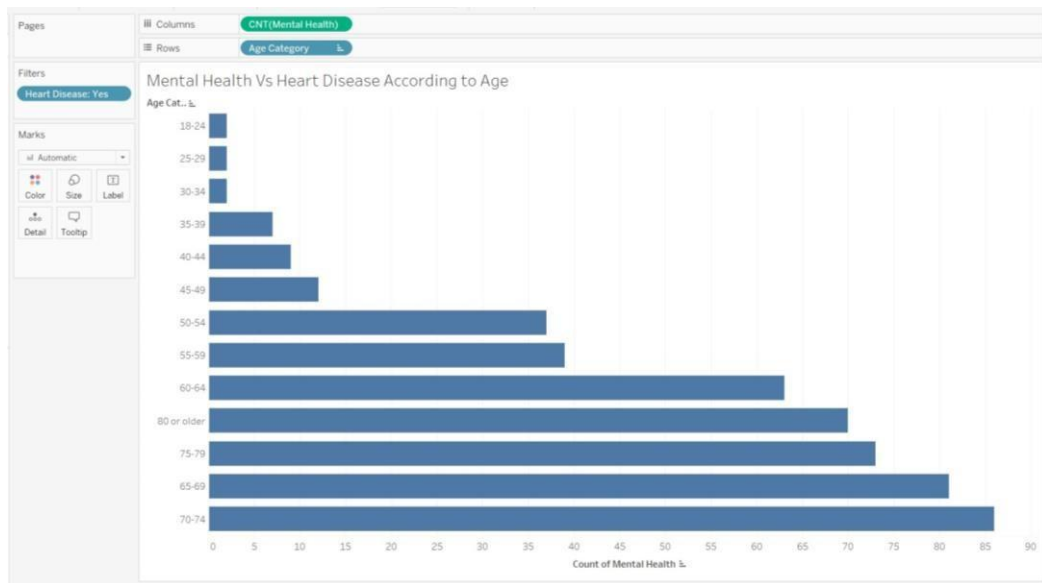
Milestone 7: Performance Testing

Amount of Data rendered to DB

<p>Navigator</p> <p>SCHEMAS</p> <p>Filter objects</p> <ul style="list-style-type: none"> project <ul style="list-style-type: none"> Tables <ul style="list-style-type: none"> heart_new2 Views Stored Procedures Functions superstore sys <p>Administration Schemas</p> <p>Information</p>	<p>day-1 project* project.heart_new2</p> <p>Info Columns Indexes Triggers Foreign keys Partitions Grants DDL</p> <p>Local instance MySQL80 project.heart_new2</p> <p>Table Details</p> <p>Engine: InnoDB</p> <p>Row format: Dynamic</p> <p>Column count: 18</p> <p>Table rows: 4554</p> <p>AVG row length: 348</p> <p>Data length: 1.5 MiB</p> <p>Index length: 0.0 bytes</p> <p>Max data length: 0.0 bytes</p> <p>Data free: 0.0 bytes</p> <p>Table size (estimate): 1.5 MiB</p> <p>File format:</p> <p>Data path: C:\ProgramData\MySQL\MySQL Server 8.0\Data\project\heart_new2.ibd</p> <p>Update time:</p> <p>Create time: 2023-10-30 11:14:44</p> <p>Auto increment:</p> <p>Table collation: utf8mb4_0900_ai_ci</p> <p>Information on this page may be outdated. Click Analyze Table to update it.</p>
--	---

<https://drive.google.com/file/d/1dHCloElAP6ln1F11fC4pf0GIlyiLOUR6/view>

Utilization of Data Filters



No. of Visualizations

1. Gender vs Heart Disease
2. Age vs Heart Disease
3. Diabetic vs Stroke
4. Impact of Smoking and Alcohol on Heart Disease
5. Health disease vs Stroke
6. Race Wise Heart Disease
7. General Health vs Heart Disease
8. Physical Activity vs Heart disease
9. Age and BMI vs Diabetic
10. People got Storke suffering from Heart Disease and Diabetic
11. Impact of Smoking on People with Asthma got Stroke
12. Mental health vs Heart Disease according to Age
13. Sleep Time vs Stroke according to Age
14. Skin Cancer vs Heart Disease according to Race

Milestone 8: Web Integration

