

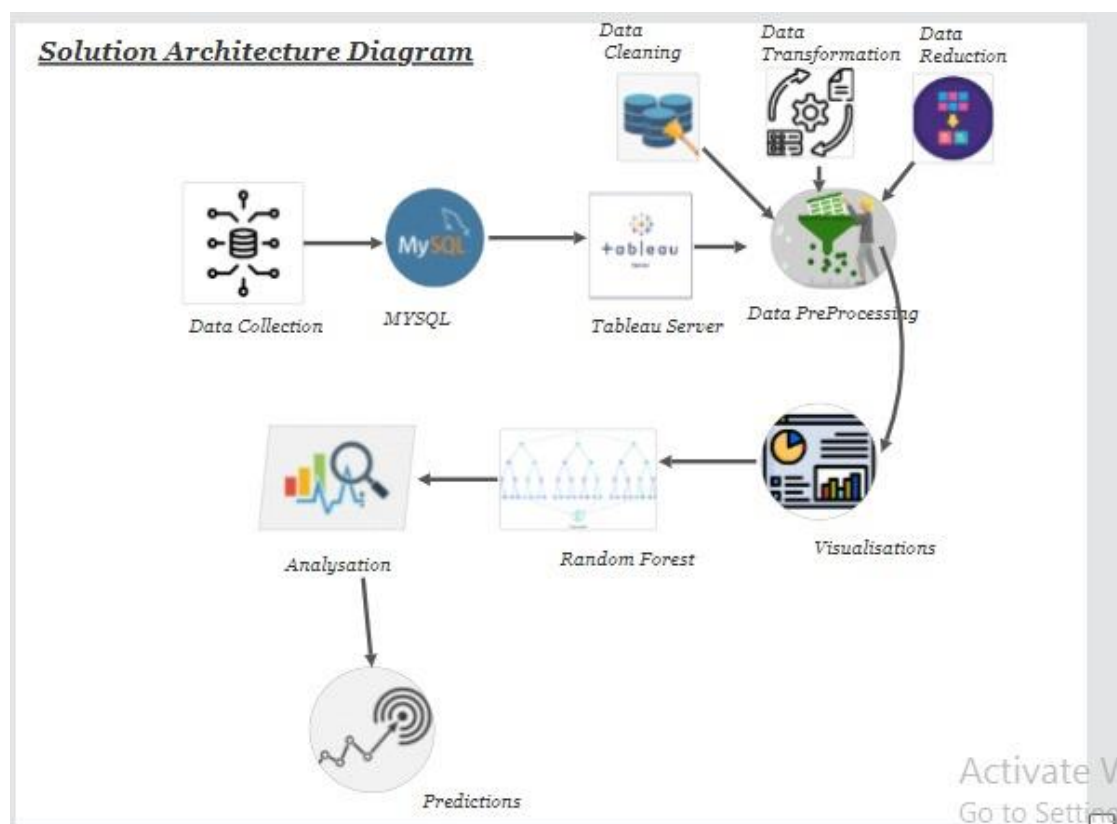
# *Heart Disease Analysis*

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

## **Technical Architecture:**



## **Project Flow:**

To accomplish 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
- **Dashboard**
  - 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 / Problem Understanding**

### **Activity 1: Specify the Business Problem**

Refer Project Description

### **Activity 2: Business requirements**

The health care industry produces a huge amount of data. This data is not always made use to the full extent and is often underutilized. Using this huge amount of data, a disease can be detected, predicted or even cured. The business requirements for analysing the Heart Disease in world include identifying patterns and comparing factors of heart diseases, creating interactive dashboards and reports, identifying areas for improvement, making data-driven decisions, comparing to the current situation and creating forecasting models for future performance. The ultimate goal is to gain insights and improve performance through data visualization techniques.

### **Activity 3: Literature Survey (Student Will Write)**

A literature survey for the Heart disease analysis would involve researching and reviewing previous studies, articles, and reports on the topic. This could include information on the methods and techniques used for analysing heart disease, as well as the results and conclusions of these studies. A comprehensive literature survey should include peer-reviewed journals, scientific databases (e.g., PubMed, Scopus), conference proceedings, and authoritative sources in the field of cardiovascular medicine. The survey should encompass a range of studies, including clinical trials, observational studies, systematic reviews, and meta-analyses, to provide a comprehensive overview of the current knowledge landscape in the field of heart disease.

## Activity 4: Social or Business Impact.

**Social Impact:** Analysing heart disease has profound social impacts, ranging from individual-level health outcomes to community empowerment and public health initiatives. By promoting awareness, prevention, equitable healthcare access, and research advancements, heart disease analysis plays a crucial role in improving the well-being of individuals and society as a whole.

**Business Model/Impact:** Analysing heart disease has substantial business impacts across various sectors, including healthcare, medical technology, pharmaceuticals, digital health, insurance, research, workplace wellness, and consumer products. It creates market opportunities, drives innovation, and influences policy and advocacy efforts in the fight against heart disease.

## Milestone 2: Data Collection & Extraction from Database

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, evaluate outcomes and generate insights from the data.

### Activity 1: Downloading the dataset

Please use the link to download the dataset: [Link](#)

#### *Activity 1.1: Understand the data*

Data contains all the meta information regarding the columns described in the CSV files

#### **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 hemographic. 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.

## Activity 2: Storing Data in DB & Perform SQL Operations

The screenshot displays the MySQL Workbench application. The top window shows the 'Welcome to MySQL Workbench' screen with a sidebar on the left containing icons for connections, schemas, and queries. The main area contains a welcome message and links to documentation, blog, and forums. Below this, the 'MySQL Connections' section shows a connection to 'Local instance MySQL80' with the user 'root' at 'localhost:3306'.

The bottom window shows the 'Local instance MySQL80' connection. The 'Navigator' pane on the left shows the 'heart\_disease' schema with a table 'heart\_new2'. The 'SQL File' editor shows a query:

```
1 use heart_disease;
2 SELECT * FROM heart_new2;
3 SELECT HeartDisease FROM heart_new2;
4 SELECT * FROM heart_new2 order by AgeCategory;
5
```

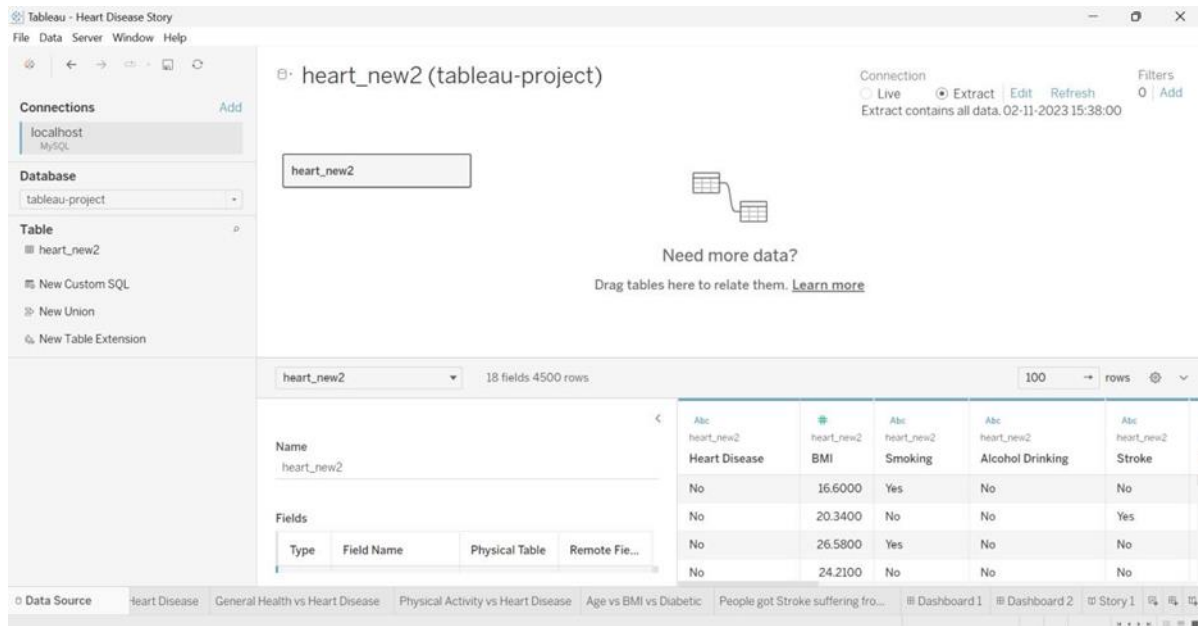
The 'Result Grid' shows the results of the query. The columns are: HeartDisease, BME, Smoking, AlcoholDrinking, Stroke, PhysicalHealth, MentalHealth, DiffWalking, Sex, AgeCategory, Race, Diabetic, Physical. The results are as follows:

HeartDisease	BME	Smoking	AlcoholDrinking	Stroke	PhysicalHealth	MentalHealth	DiffWalking	Sex	AgeCategory	Race	Diabetic	Physical
No	25.01	No	No	No	4	4	No	Male	18-24	White	No	Yes
No	25.83	No	Yes	No	0	15	No	Male	18-24	White	No	Yes
No	22.67	No	No	No	0	7	No	Female	18-24	Black	No	Yes
No	22.83	No	No	No	0	0	No	Female	18-24	Black	No	Yes
No	22.3	No	No	No	0	0	No	Female	18-24	Asian	No	No
No	21.79	No	No	No	7	7	No	Female	18-24	White	No	Yes
No	25.5	No	No	No	0	0	No	Male	18-24	White	No	Yes
No	24.39	No	No	No	2	0	No	Male	18-24	White	No	Yes
No	22.46	Yes	Yes	No	0	30	No	Female	18-24	White	No	Yes
No	22.4	No	No	No	0	2	No	Female	18-24	Hispanic	No	Yes
No	21.52	No	No	No	0	0	No	Male	18-24	Black	No	No
No	19.2	No	No	No	2	0	No	Male	18-24	Asian	No	Yes
No	19.97	No	No	No	0	0	No	Female	18-24	Asian	No	Yes

The 'Output' pane shows the execution of the query, with the following results:

#	Time	Action	Message	Duration / Fetch
1	11:59:38	use heart_disease	0 row(s) affected	0.000 sec
2	11:59:42	SELECT * FROM heart_new2	4500 row(s) returned	0.000 sec / 0.016 sec
3	11:59:47	SELECT HeartDisease FROM heart_new2	4500 row(s) returned	0.000 sec / 0.000 sec
4	11:59:51	SELECT * FROM heart_new2 order by AgeCategory	4500 row(s) returned	0.016 sec / 0.015 sec

## Activity 3: Connect DB with Tableau



## Milestone 3: Data Preparation

### Activity 1: Prepare the Data for Visualization

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency. Since the data is already cleaned we can move to visualization.

## Milestone 4: Data Visualization

Data visualization is the process of creating graphical representations of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

## Activity 1: No of Unique Visualizations

The number of unique visualizations that can be created with a given dataset. Some common types of visualizations that can be used to analyse the performance and efficiency of banks include bar charts, line charts, heat maps, scatter plots, pie charts Maps etc. These visualizations can be used to compare performance, track changes over time, show distribution, and relationships between variables, breakdown of revenue and customer demographics, workload, resource allocation and location of banks.

### *Activity 1.1 - Gender vs HD*

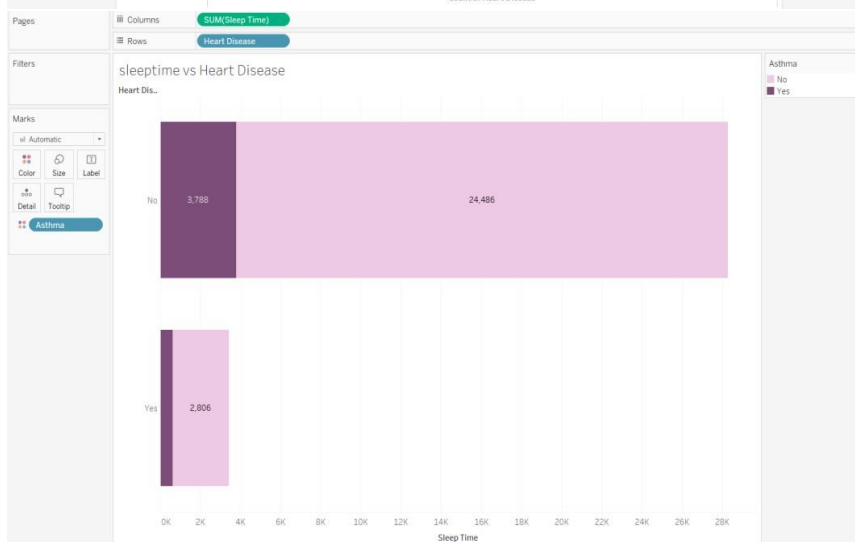
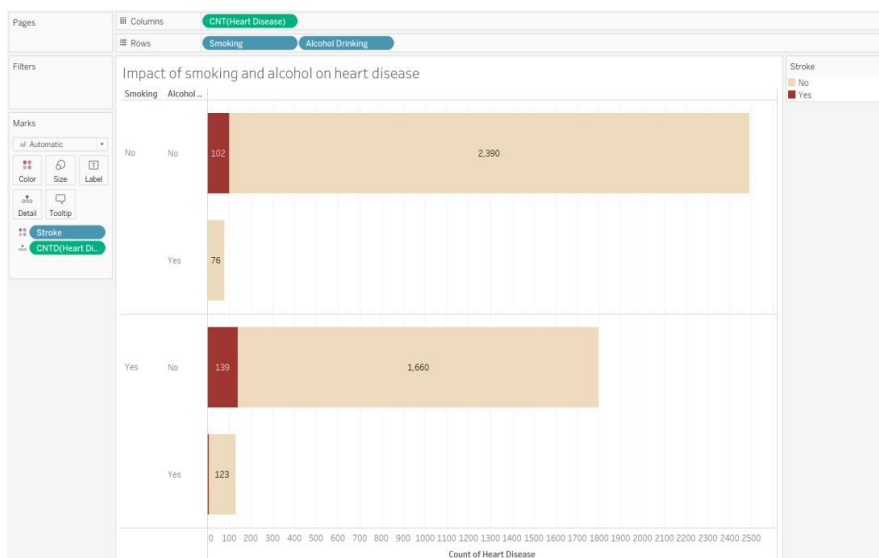


### *Activity 1.2 – Age vs HD*

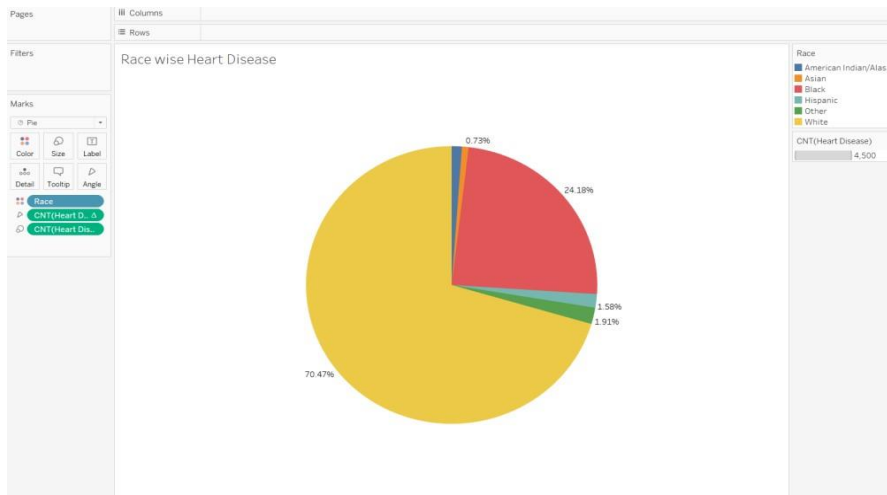
### *Activity 1.3 – Diabetic vs Stroke*

### *Activity 1.4 – Impact of Smoking and Alcohol on HD*



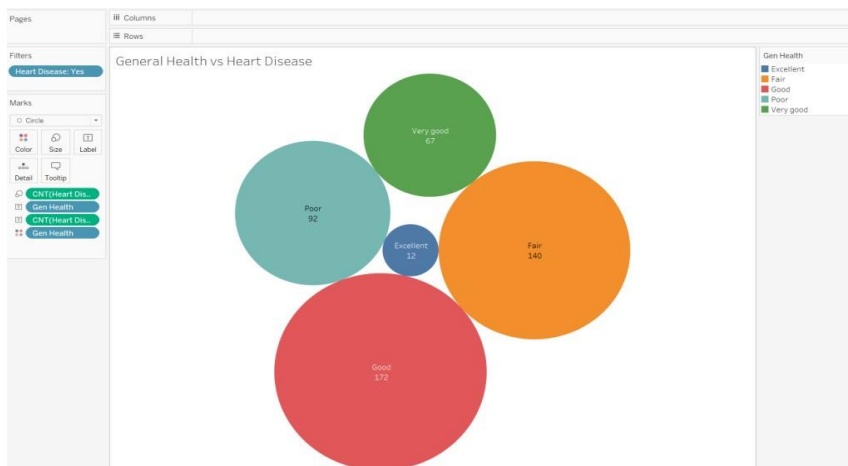


## Activity 1.5 – Other Health disease vs Stroke



## Activity 1.6 – Race wise Health disease

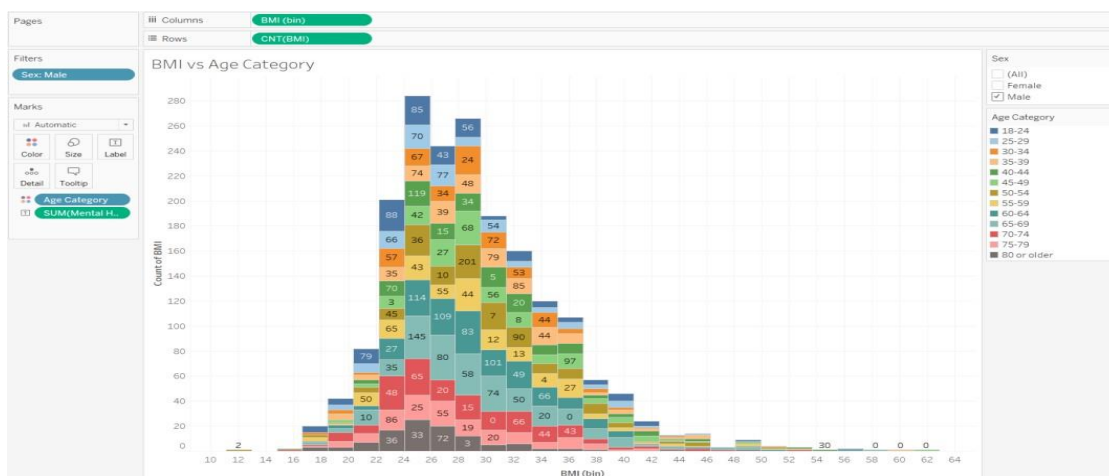
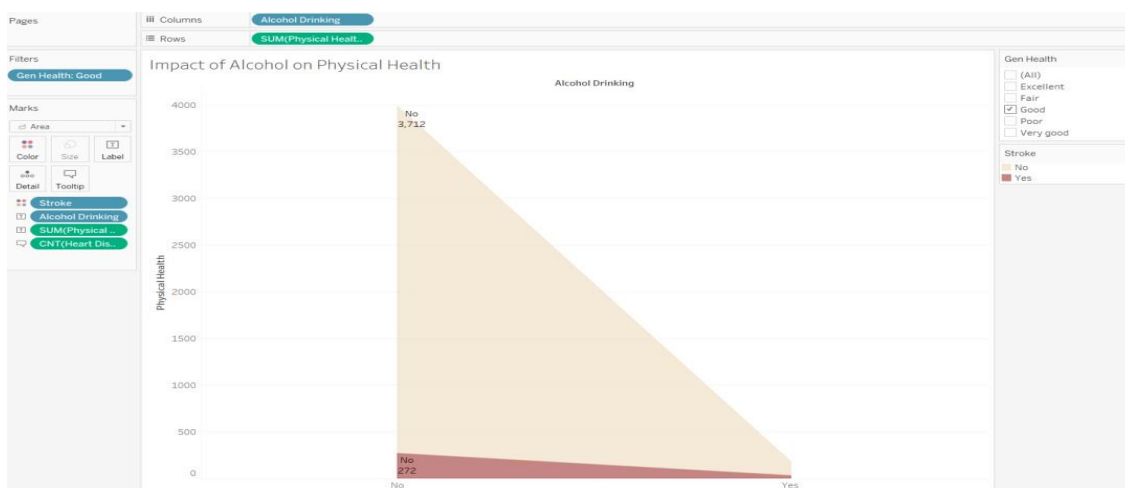
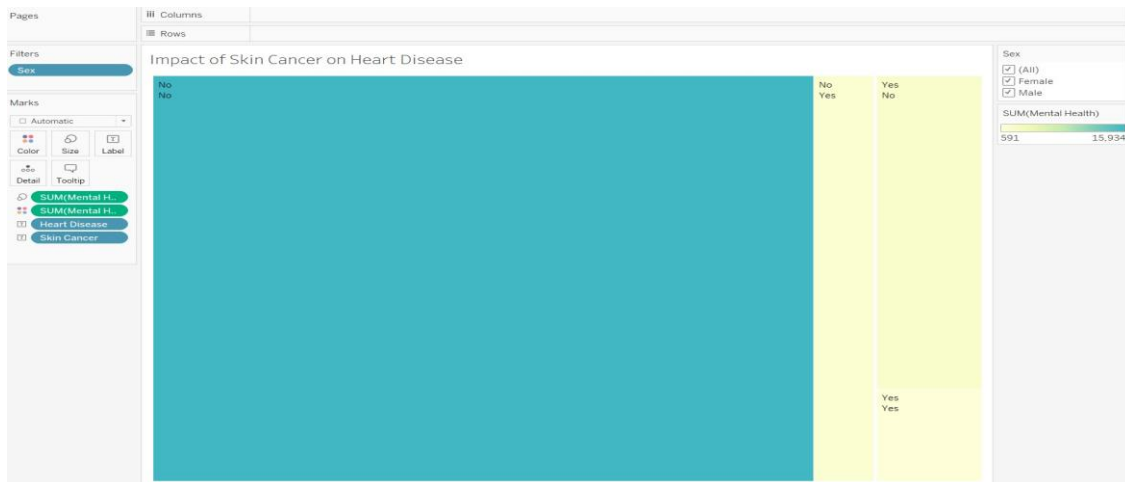
## Activity 1.7 – General Health vs Health disease



## Activity 1.8 – Physical Activity vs Heart disease

## Activity 1.9 – Age and BMI vs Diabetic

## Activity 1.10 – people got stroke suffering from HD and Diabetic



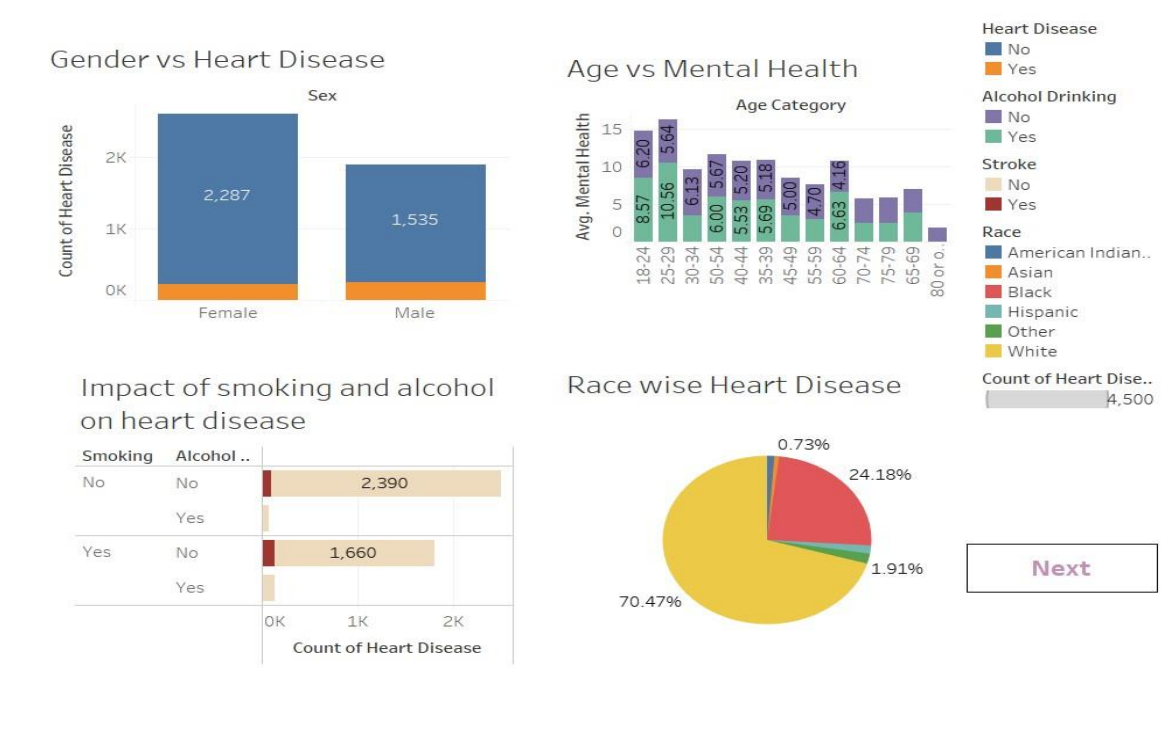
## Milestone 5: Dashboard

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.

### Activity 1- Responsive and Design of Dashboard

Once you have created views on different sheets in Tableau, you can pull them into a dashboard.

**Dashboard is about the smoking affects by the heart disease**



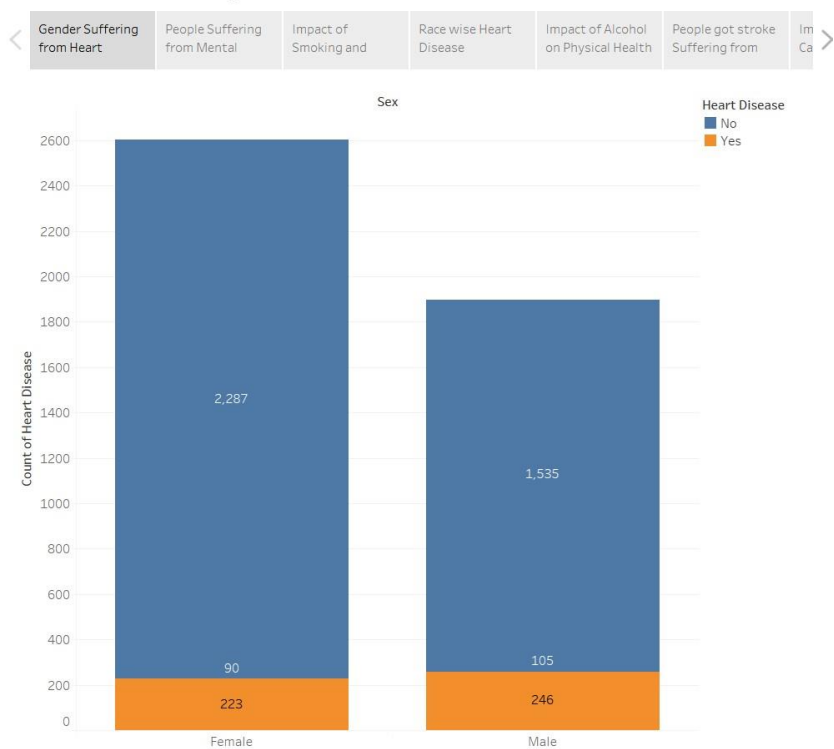
## Milestone 6: Story

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.

**Activity 1: No. of Scenes of Story**

The number of scenes in a storyboard for a data visualization analysis of the Heart disease will depend on the complexity of the analysis and the specific insights that are trying to be conveyed. A storyboard is a visual representation of the data analysis process and it breaks down the analysis into a series of steps or scenes.

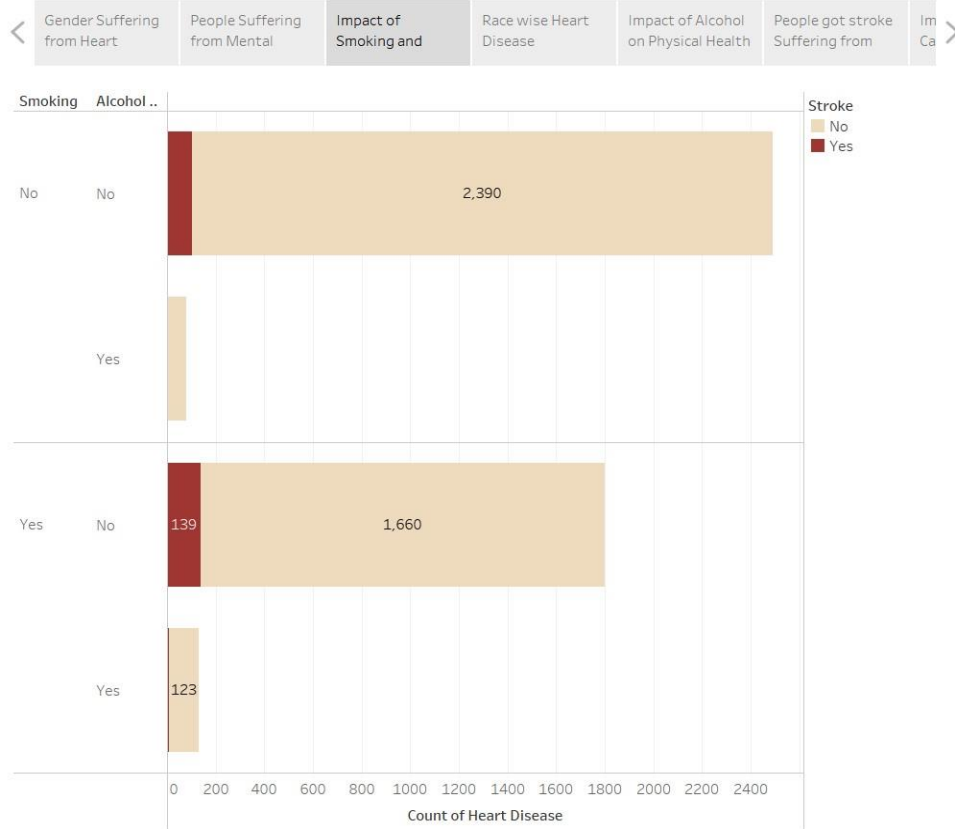
Heart Disease Story



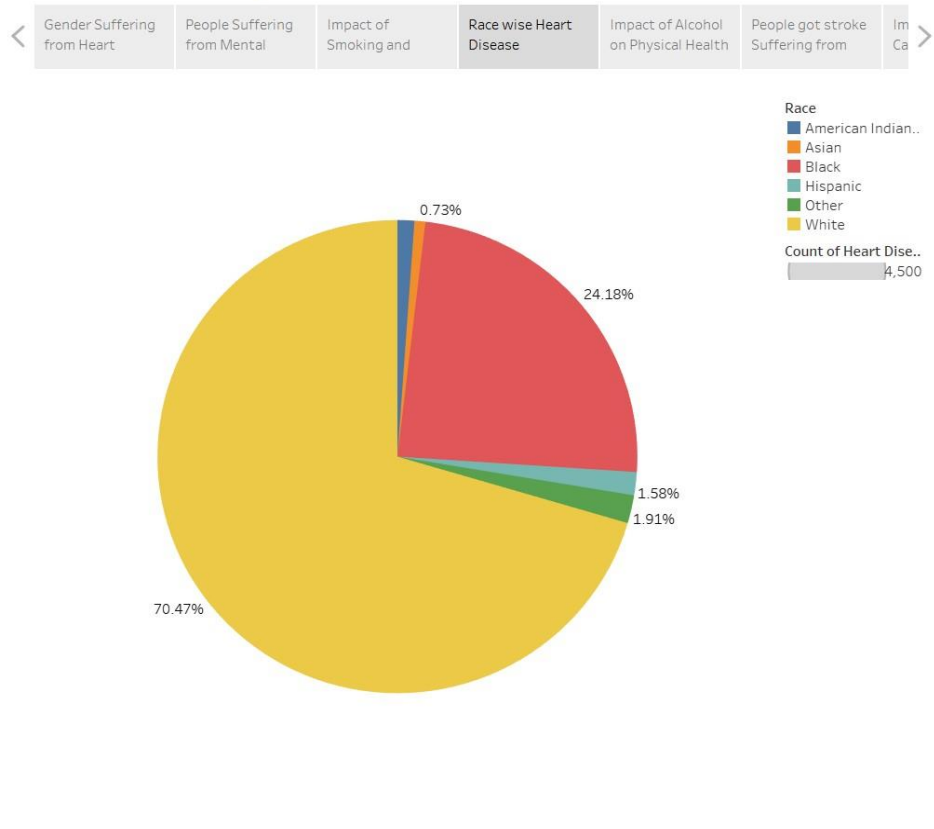
## Heart Disease Story



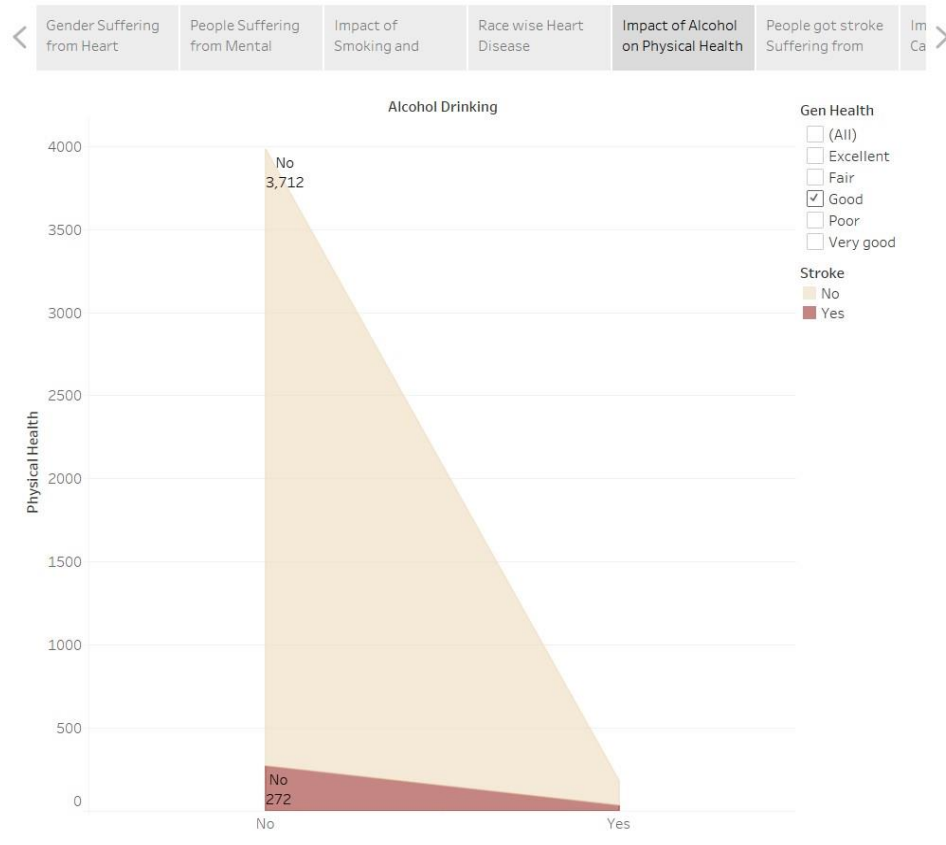
## Heart Disease Story



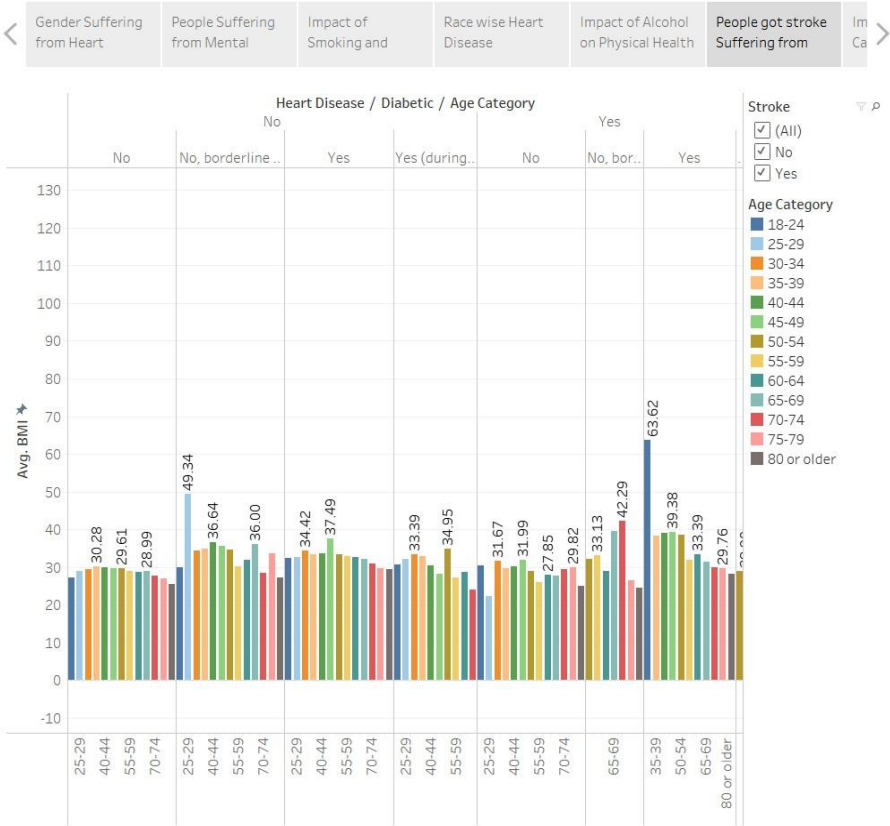
## Heart Disease Story



## Heart Disease Story



Heart Disease Story



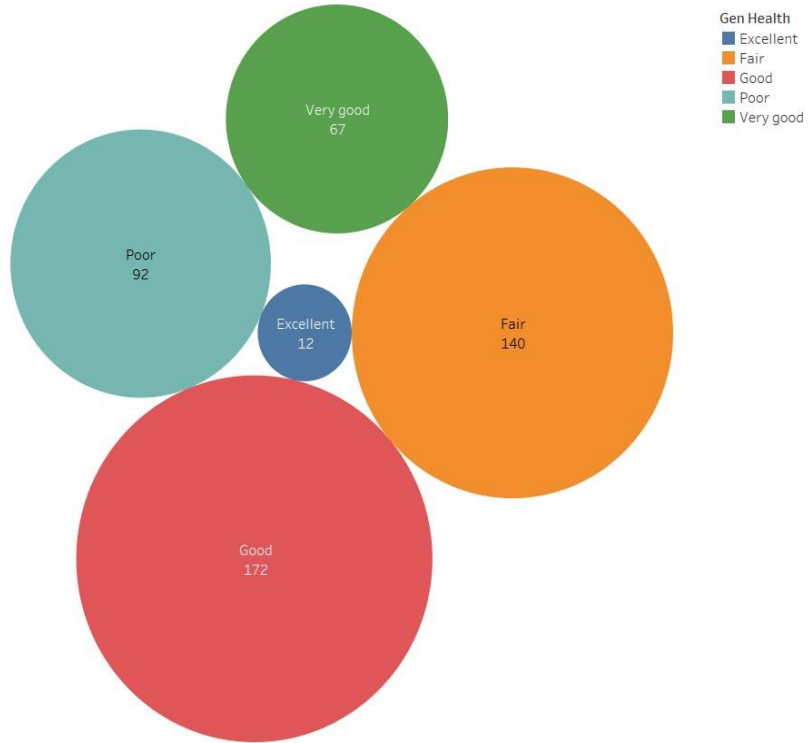
Heart Disease Story





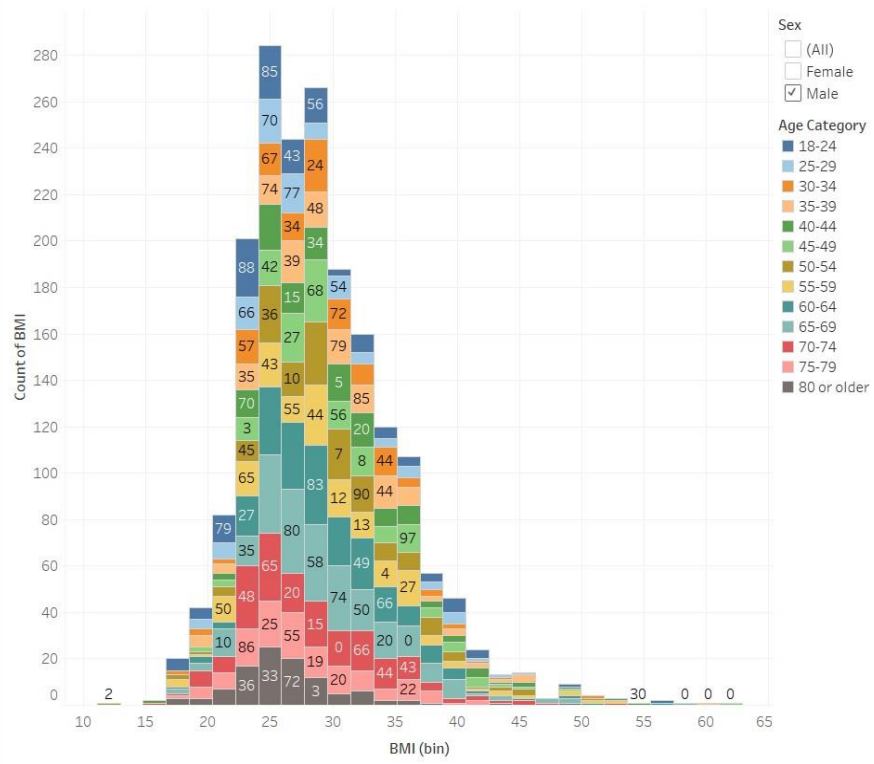
# Heart Disease Story

< Impact of moking and Race wise Heart Disease Impact of Alcohol on Physical Health People got stroke Suffering from Impact of Skin Cancer on Heart **People Suffering from Heart** Count Based >



# Heart Disease Story

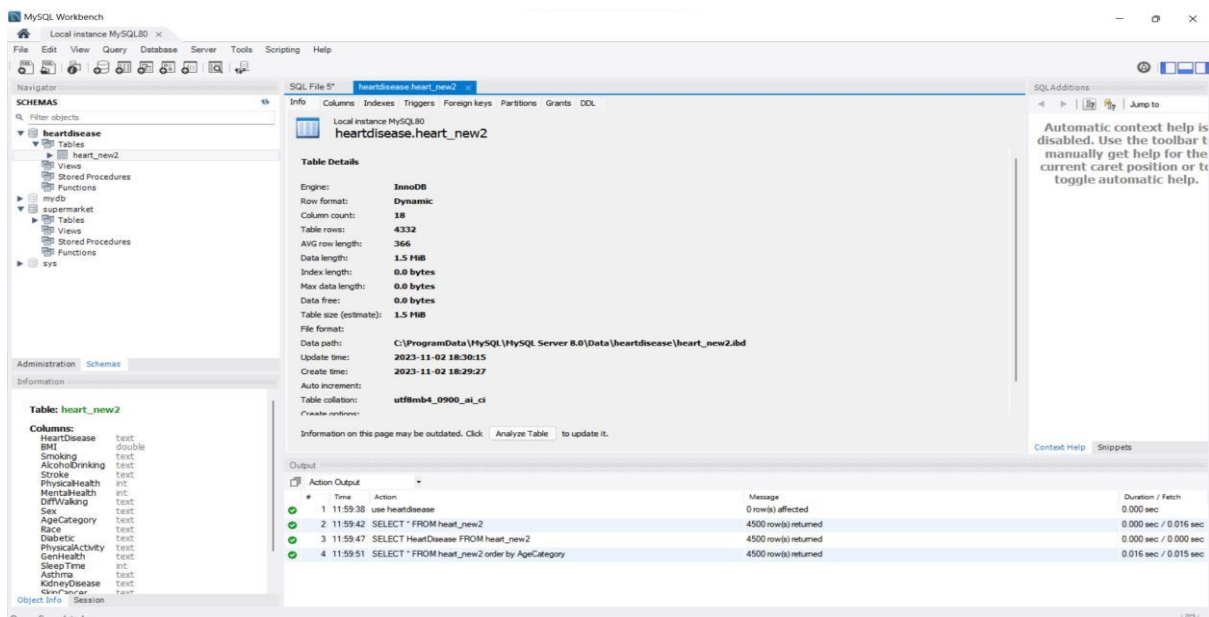
< Race wise Heart Disease Impact of Alcohol on Physical Health People got stroke Suffering from Impact of Skin Cancer on Heart People Suffering from Heart **Count of BMI Based on Age** >



## Milestone 7: Performance Testing

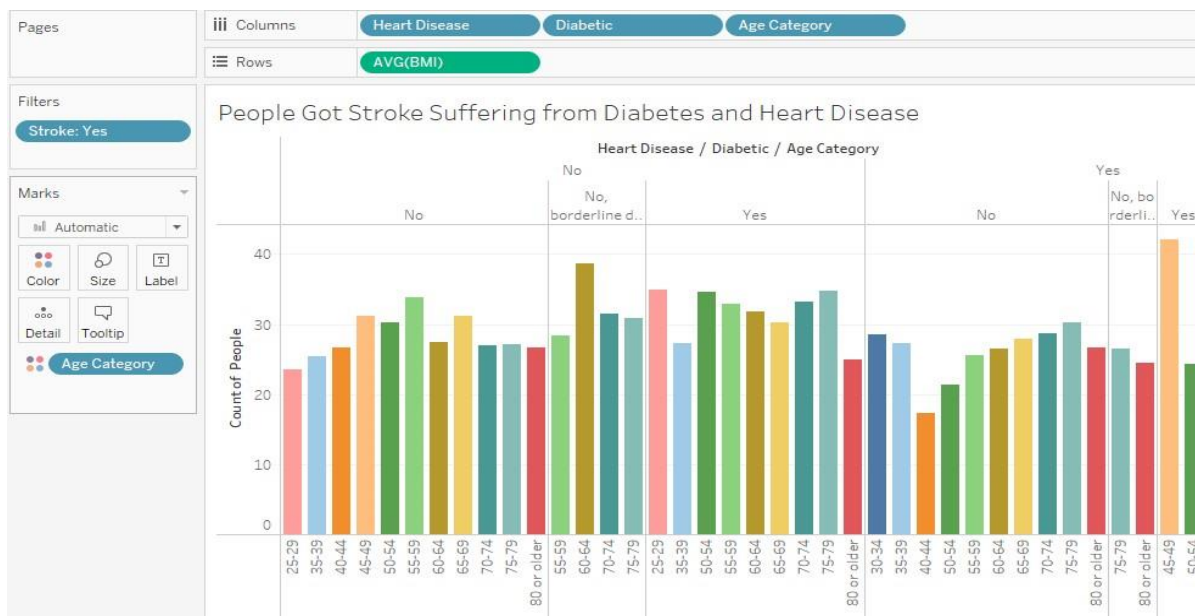
### Activity 1: Amount of Data Rendered to DB

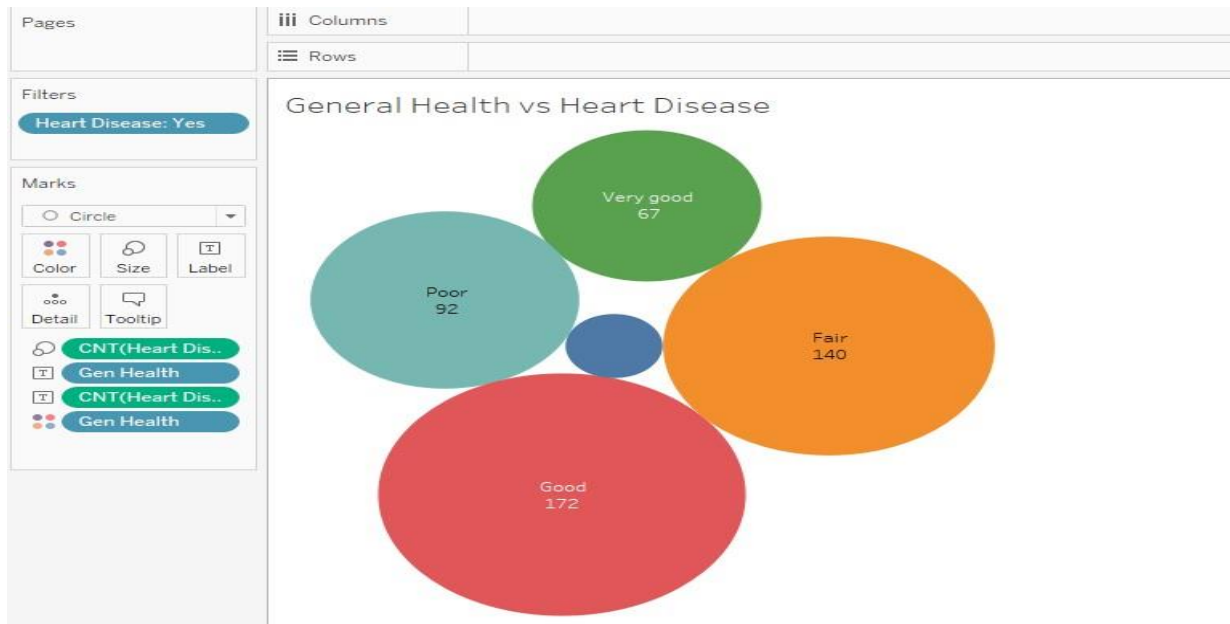
- The amount of the data that is rendered to a database depends on the size of the dataset and the capacity of the database to store and retrieve data.
- Open the MYSQL Workbench, go to the data base then click to expand the tables, select the table and click on the button to get the information related to such as column, count, table rows etc



:

### Utilization of Data Filters:





### Activity 3: No. of Calculations Fields

In this analysis we have not created any new column using calculation filed as data found in the dataset was clean and sufficient for analysis.

### Activity 4: No. of Visualizations/Graphs

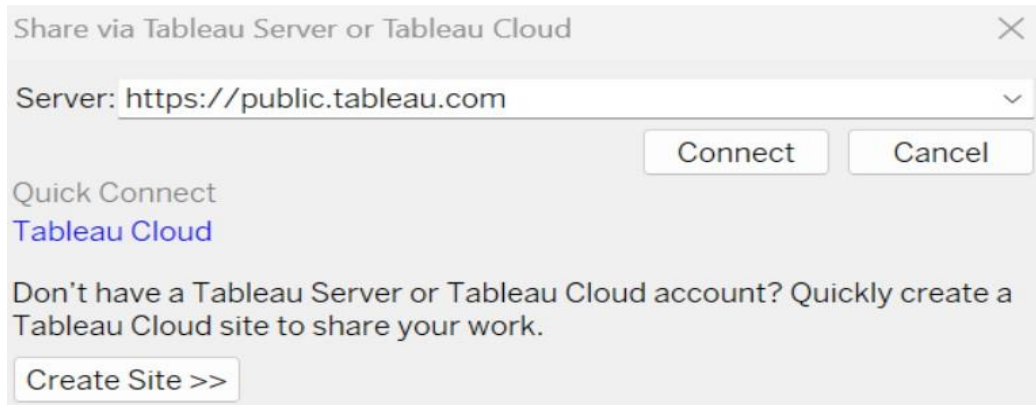
1. Gender wise Heart Disease
2. Age wise Heart Disease
3. People suffering from Diabetic and Stroke
4. Impact of Smoking and alcohol drinking on heart disease
5. Other Disease vs Stroke
6. Race wise heart disease
7. General Health vs heart disease
8. Physical activity vs Heart disease
9. Age and BMI vs Heart Disease
10. People got stroke suffering from Diabetes and Heart disease

### Milestone 8: Web Integration

Publishing helps us to track and monitor key performance metrics and to communicate results and progress. Help a publisher stay informed, make better decisions, and communicate their performance to others.

**Publishing dashboard and reports to tableau public**

**Step 1:** Go to Dashboard/story, click on the share button on the top ribbon  
Give the server address of your tableau account and click on connect.



Share via Tableau Server or Tableau Cloud

Server:

Quick Connect  
**Tableau Cloud**

Don't have a Tableau Server or Tableau Cloud account? Quickly create a Tableau Cloud site to share your work.

**Step2:** once you click on connect it will ask you for the tableau public username and password



tableau public

Email

Password

 This site is SSL encrypted

[Forgot your password?](#)  
[Don't have a profile yet?](#)  
[Create one now for free](#)

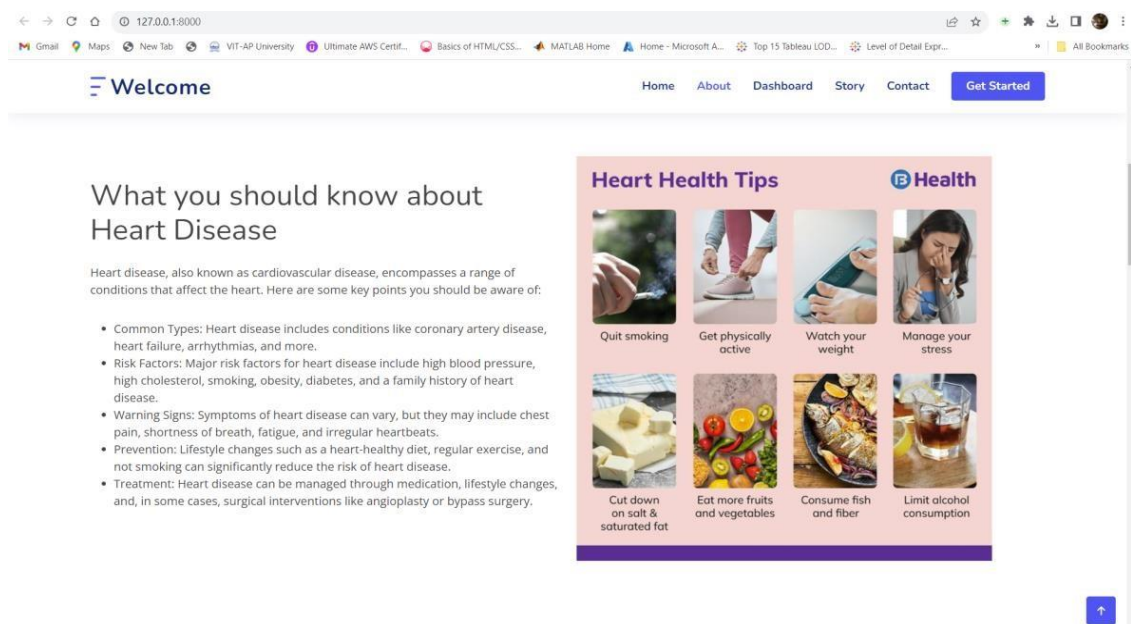
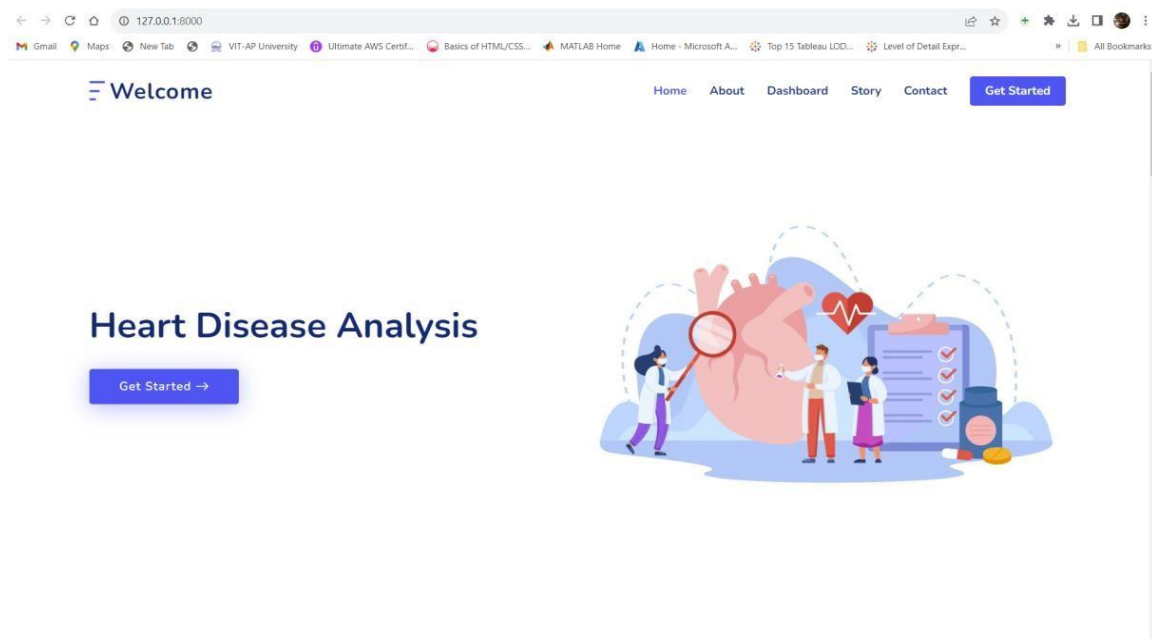
Once you login into your tableau public using the credentials, the particular visualization will be published into the tableau public

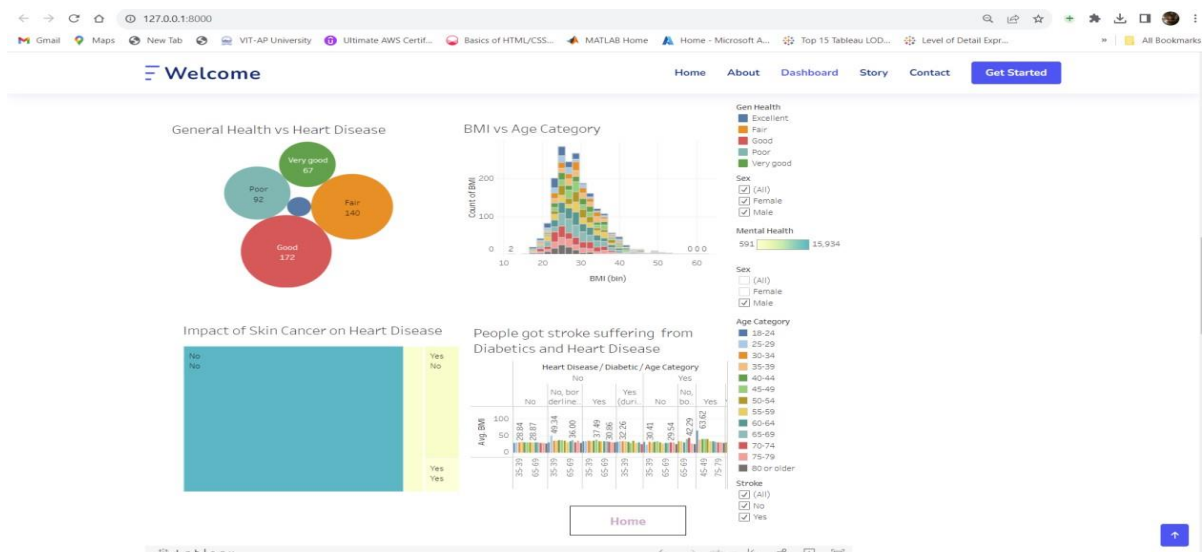
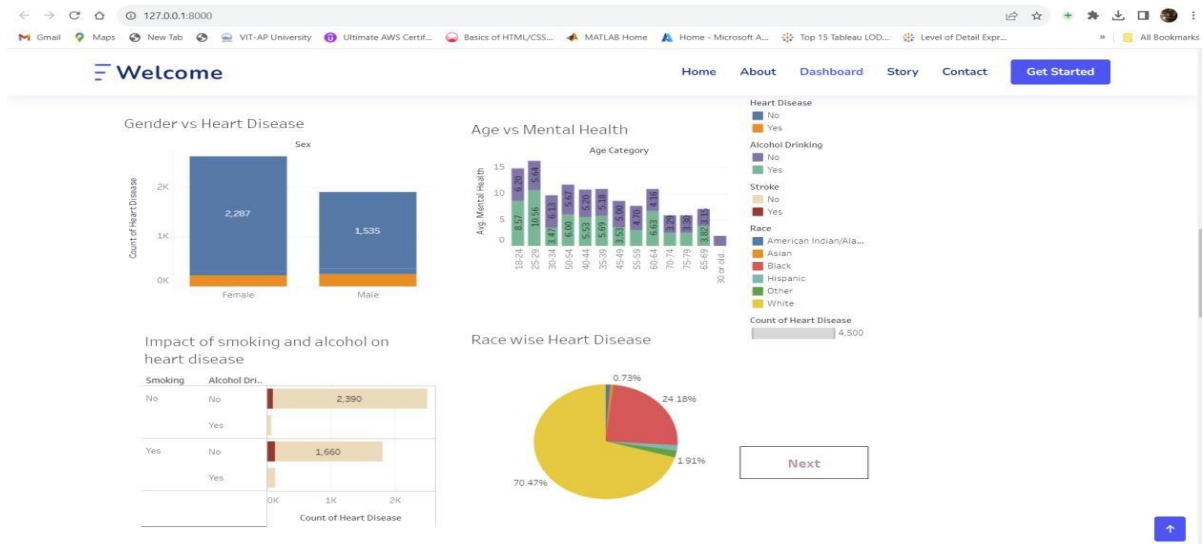
These are the dashboard and story links which we have Published

[https://public.tableau.com/app/profile/vishnupriya.pragada/viz/Dashboard\\_16989481972700/Dashboard1?publish=yes](https://public.tableau.com/app/profile/vishnupriya.pragada/viz/Dashboard_16989481972700/Dashboard1?publish=yes)

[https://public.tableau.com/app/profile/vishnupriya.pragada/viz/Story\\_16989482739400/Story1?publish=yes](https://public.tableau.com/app/profile/vishnupriya.pragada/viz/Story_16989482739400/Story1?publish=yes)

## Activity 1: Embed Dashboard and Story with Flask





Welcome Home About Dashboard Story Contact Get Started

## CONTACT

### Contact Us

**Address**  
Andharpradesh,  
INDIA, 522124

**Call Us**  
+1 5589 55488 55  
+1 6678 254445 41

**Email Us**  
info@example.com  
contact@example.com

**Open Hours**  
Monday - Friday  
9:00AM - 05:00PM

Your Name

Your Email

Subject

Message

Send Message