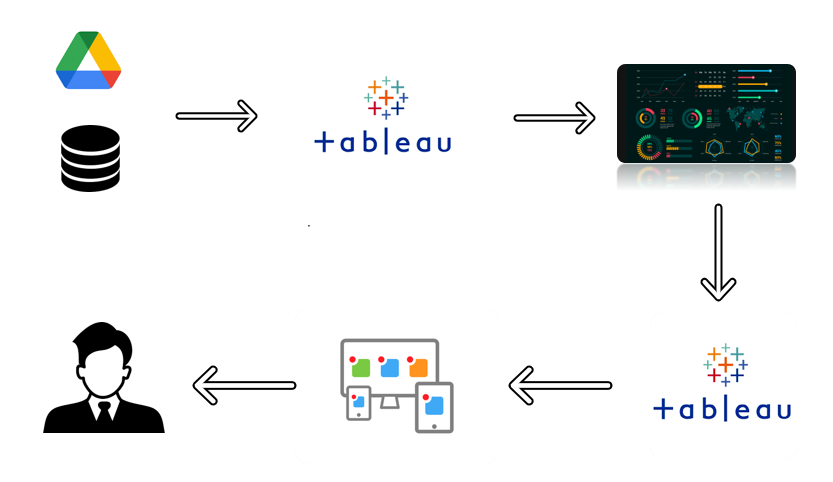
**Heart Disease Analysis**

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 analyze 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 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
* 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 analyzing 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**

Heart rate arrhythmias have become one of the most popular heart issues in human beings at recent years. It has opened a way to many researchers to classify the heart beats issues in the medical field for detecting the arrhythmia signals at early stages which will help in reducing the death rates. The research focuses on classifying the arrhythmias in the ECG signal for the 5 different categories of signals in the MIT-BIH dataset. A lot of research has been done in the classification for arrhythmias signal using various data mining techniques. Logistic Regression and CNN algorithms were implemented to classify the heart beat signals and aims at providing better accuracy in predicting the signals. Both the implemented models have been compared in terms of accuracy, precision, recall and computational time. The evaluation results of the deep leaning model - CNN turned out to be high in accuracy of 99.09% when compared with the machine learning model. This research helps out in classifying the heartbeat rate arrhythmias at early stages for enhanced treatment.

**Activity 4: Social or Business Impact.**

Social Impact- Analyzing 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: Analyzing 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.csv](file:///C:\Users\VIJAY%20RAMARAJ\Downloads\LINK.csv)

**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 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. Difficulty in 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. General Health - well-being.
15. Sleep Time - number of hours of sleep.

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

Explanation video link:

<https://drive.google.com/file/d/1roOS1Ijn-5lKSRkaxdqqdae7Syo5-kIR/view?usp=sharing>

**Activity 3: Connect DB with Tableau**

Explanation video link:

<https://drive.google.com/file/d/1RHOtTX4H26hUTjgcS48m62qzqUJLEvPE/view?usp=sharing>

**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.

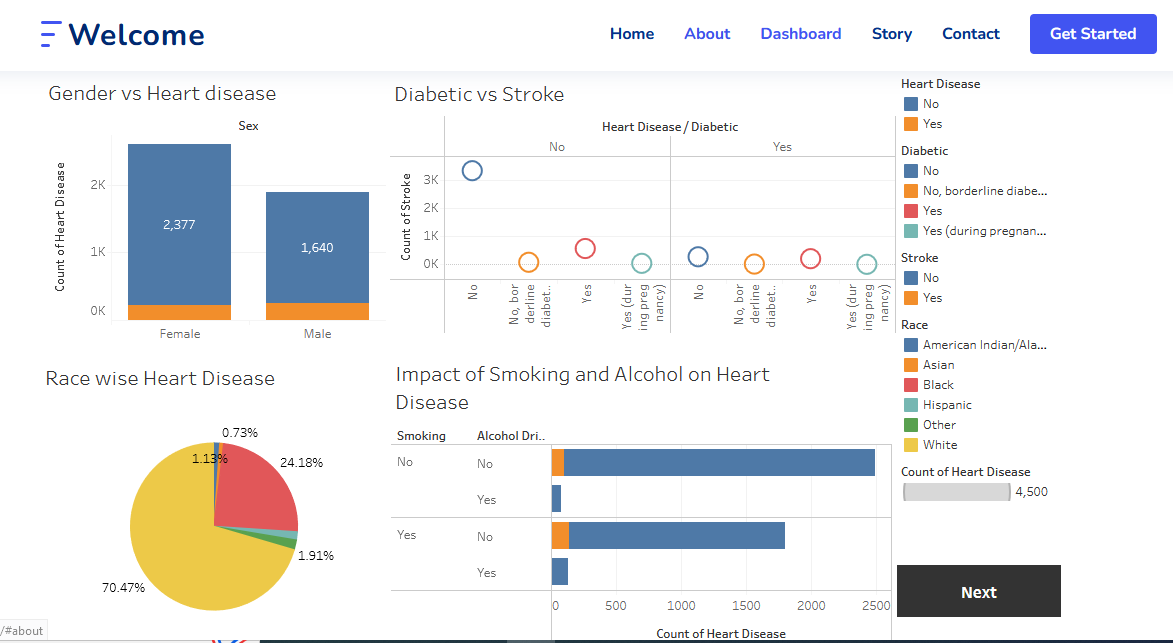
**Milestone 4: Data Visualization**

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 analyze 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.

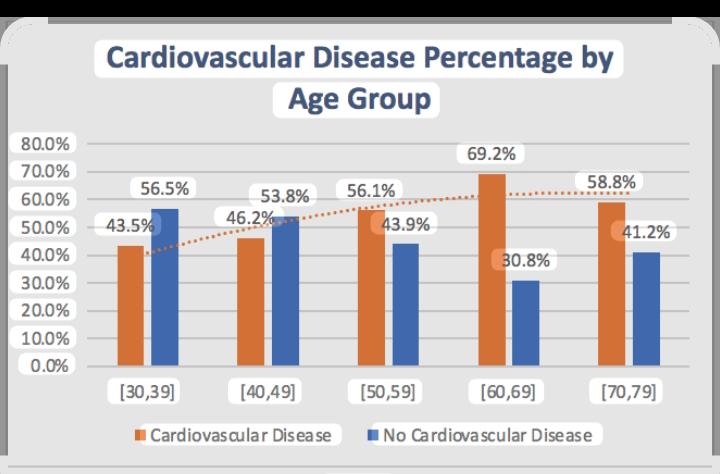
**Activity 1.1: Gender vs heart disease**



**Explanation video link:**

<https://drive.google.com/file/d/1v7JrKBC8wZ0d0J2pklUpHSSrcR7cYywS/view?usp=sharing>

**Activity 1.2: Age vs heart disease**



**Explanation video link**

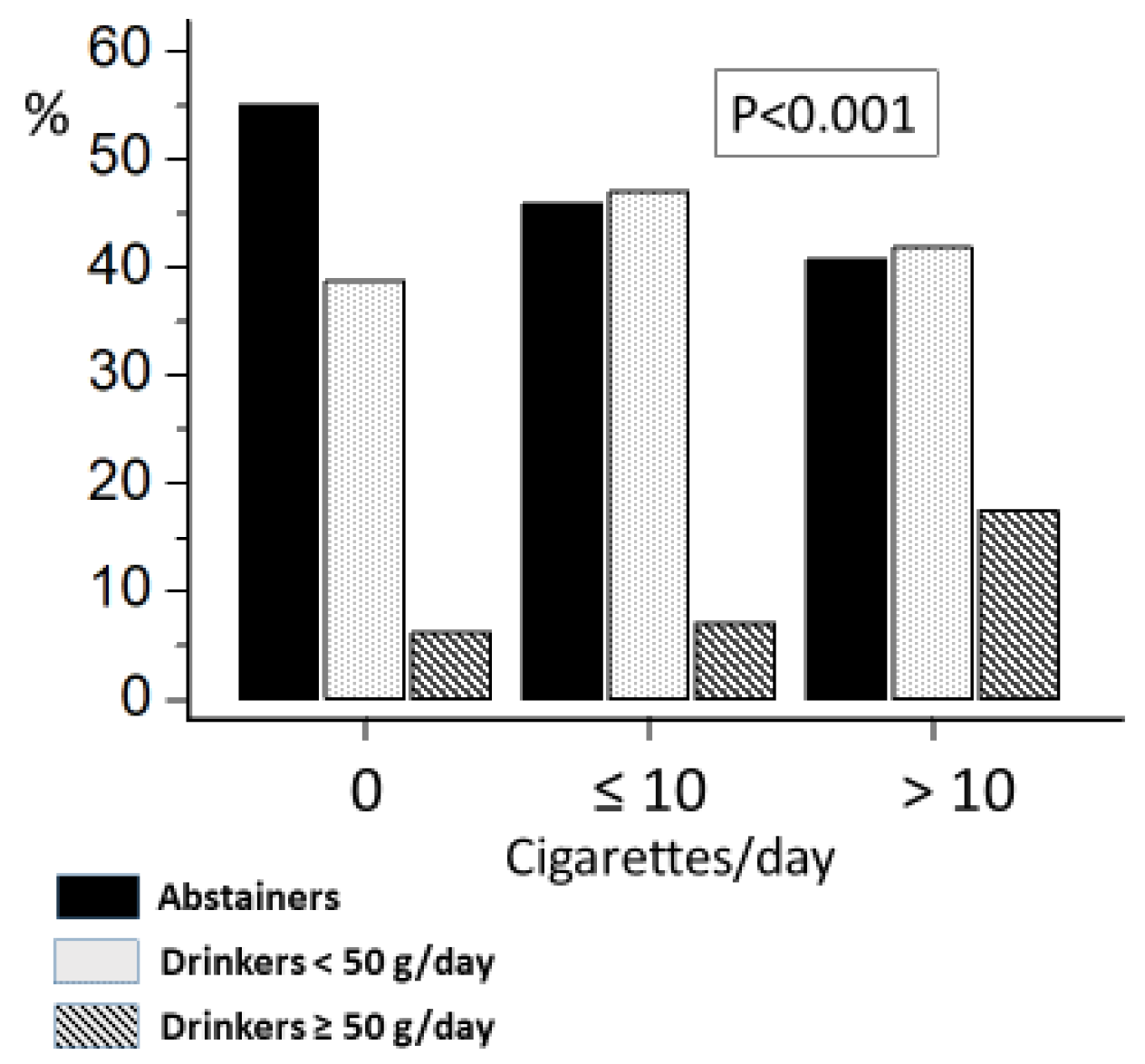
<https://drive.google.com/file/d/1qgwzyYlDgseOFRyHtWH9OKK19DYlgiGW/view?usp=sharing>

**Activity 1.3: Diabetic vs Stroke**

**Explanation video link:**

<https://drive.google.com/file/d/1WgmsSrTT78kTnpfrWHaMVGtpnGo0qmCb/view?usp=sharing>

**Activity 1.4: Impact of Smoking and Alcohol on Heart Disease**

****

**Explanation video link:**

<https://drive.google.com/file/d/1cD1eaKRfKnIobK8LIskW9Q2o4UkS3yew/view?usp=sharing>

**Activity 1.5 Other Health Diseases vs Stroke**

**Explanation video link:**

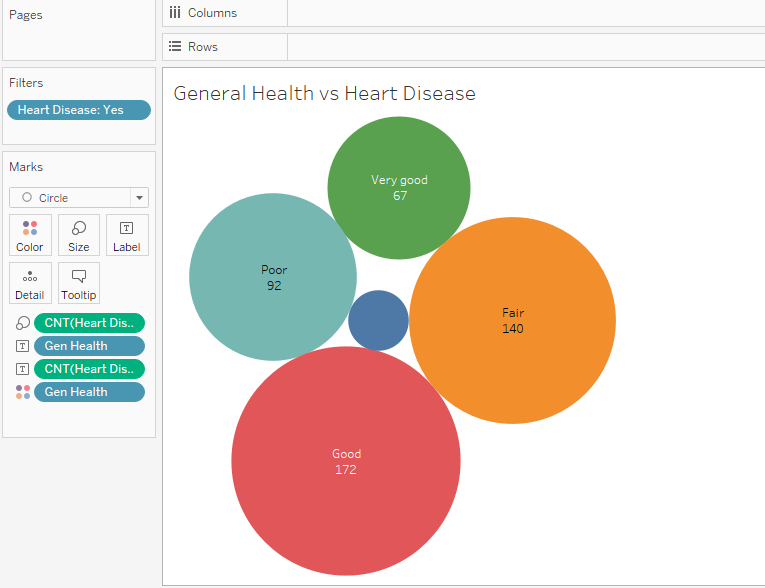
<https://drive.google.com/file/d/1n-jIA4EDBj_9vFGcAWSfJheuVSs7izTn/view?usp=sharing>

**Activity 1.6: Race wise heart disease**

Explanation video link:

<https://drive.google.com/file/d/1ejftMI8ELmCb4N6Y9eE_n1JQEBMVQkV8/view?usp=sharing>

**Activity 1.7: General Health vs heart disease**



Explanation video link:

<https://drive.google.com/file/d/1NT9myDf4vQYuzapfIWHq_GQTs9NiKtTS/view?usp=sharing>

**Activity 1.8: Physical Activity vs heart disease**

Explanation video link:

<https://drive.google.com/file/d/1JW0CJL-yVdJuBGVQMQEeUl9NTOeukufu/view?usp=sharing>

**Activity 1.9: Age and BMI vs Diabetic**

Explanation video link:

<https://drive.google.com/file/d/1n-qiD9IeCBX2_B_yYXGLuE0zscxdpUO-/view?usp=sharing>

**Activity 1.10: People got stroke suffering from heart disease and Diabetic**

Explanation video link:

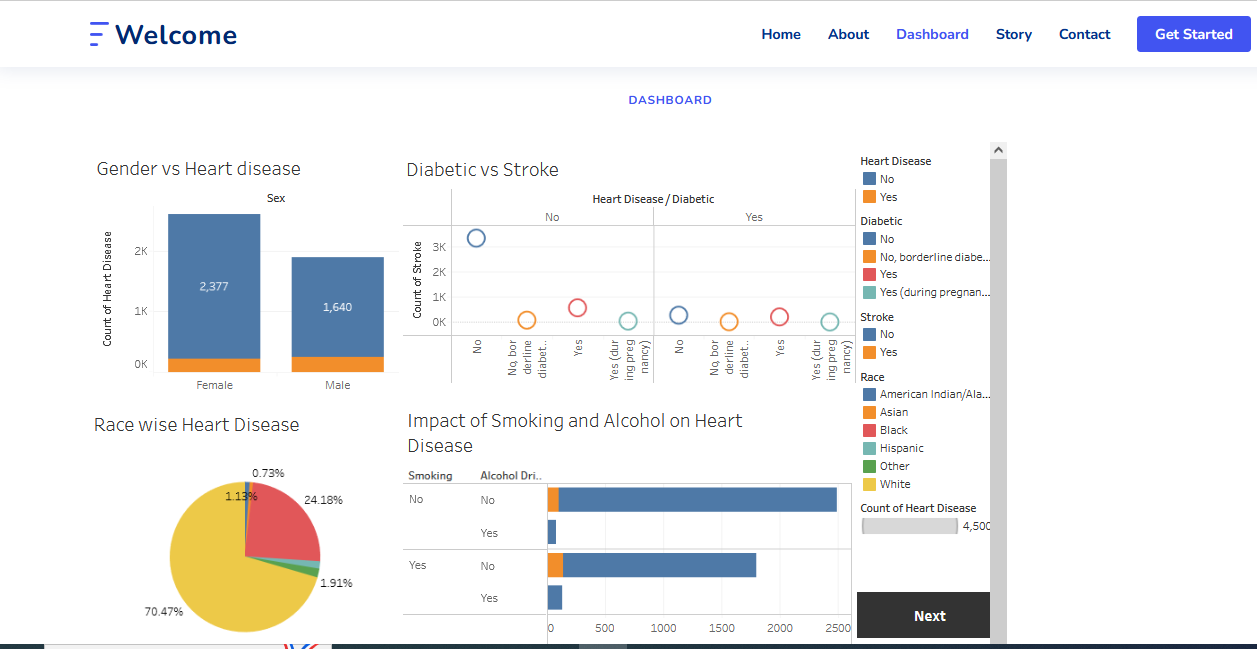
<https://drive.google.com/file/d/17D34bsRWsvMguZIf4CHKfKF-N9ZuwRFL/view?usp=sharing>

**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.



Explanation video link:

<https://drive.google.com/file/d/1WRs9Bnt0lEWhL5w6LQnFLUJeGQ6sQSEU/view?usp=sharing>

**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.

**Explanation video link:**

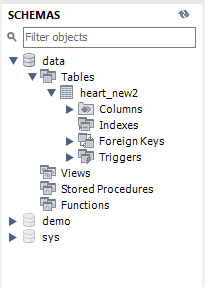
<https://drive.google.com/file/d/17lrCoLKDi6Omikcy6_11MfhCsdd6ySdb/view?usp=sharing>

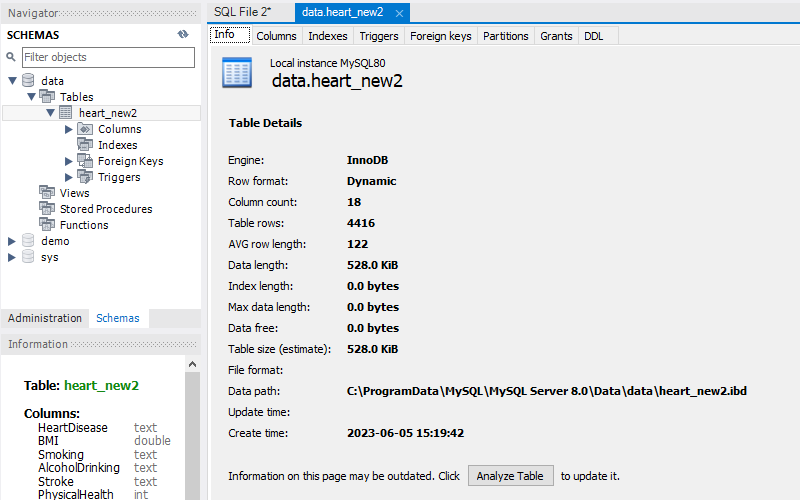


**Milestone 7: Performance Testing**

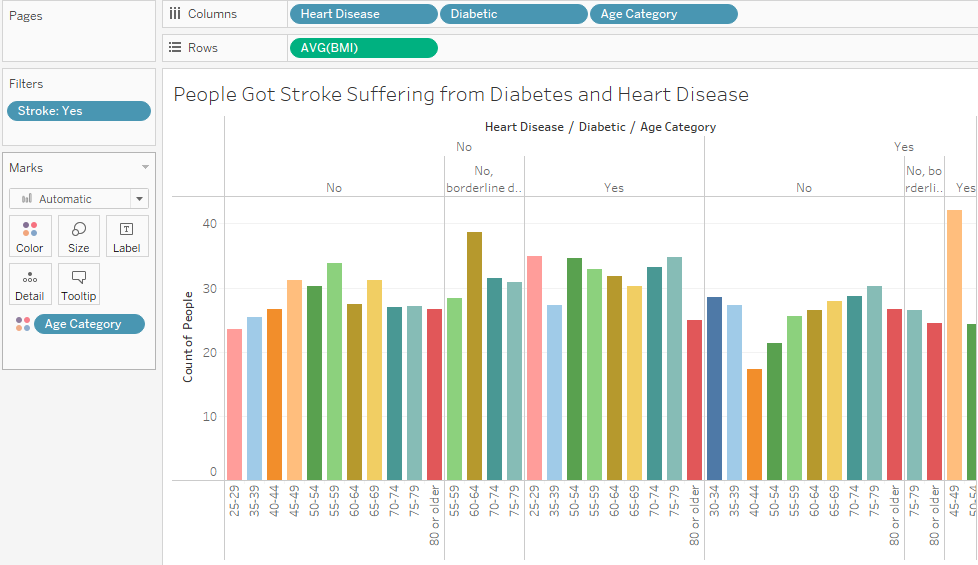
**Activity 1: Amount of Data Rendered to DB**

* The amount of 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 database then click to expand the tables, select the table and click on (i) button to get the information related to table such as column count, table rows etc.





**Activity 2: Utilization of Data Filters**



**Activity 3: No of Calculation Fields**

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

**Activity 4: No of Visualizations/ Graphs**

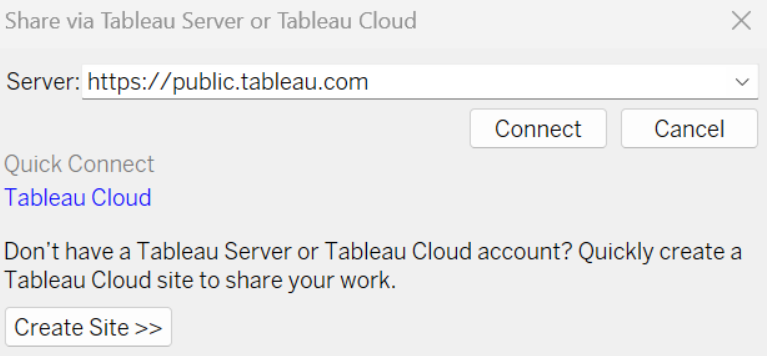
1. Gender wise heart disease
2. Age wise heart disease
3. People Suffering from Diabetic and Stroke
4. Impact of Smoking and alcohol drinking on heart disease
5. Other Diseases vs Stroke
6. Race wise heart disease
7. General Health vs heart disease
8. Physical activity vs heart disease
9. Age and BMI 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

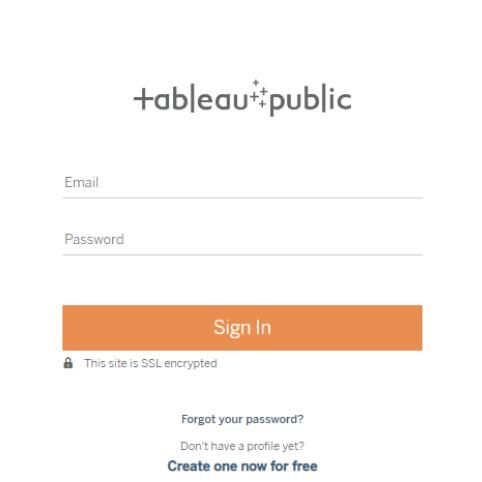
****

Give the server address of your tableau public account and click on connect.

**Explanation Video: -**

<https://drive.google.com/file/d/1DIXlLAPYgJARUQlAubENQjB53E95YxsH/view?usp=sharing>

**Step 2:** Once you click on connect it will ask you for the tableau public username and password



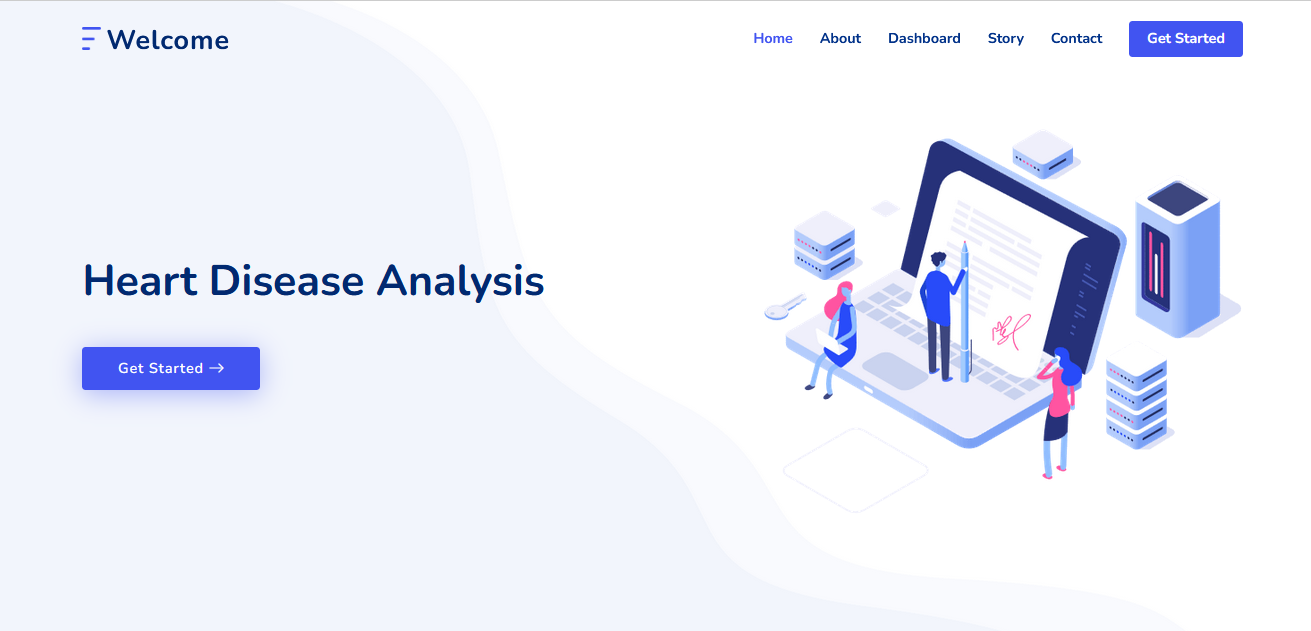
Once you login into your tableau public using the credentials, the particular visualization will be published into the tableau public

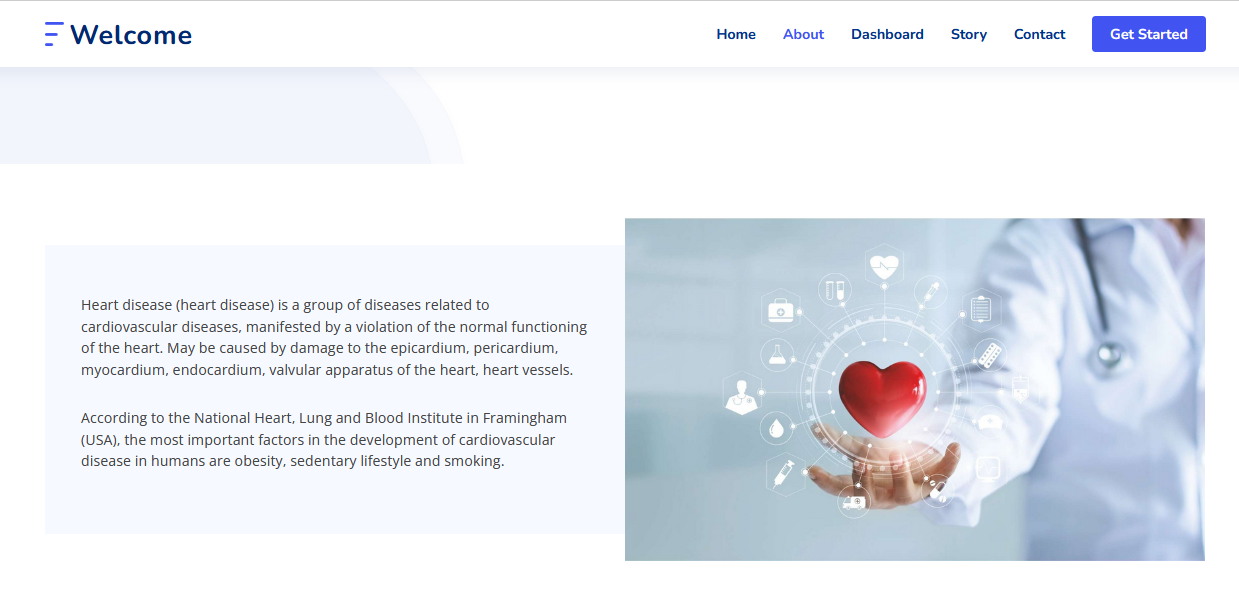
**Note: While publishing the visualization to the public, the respective sheet will get published when you click on the share option.**

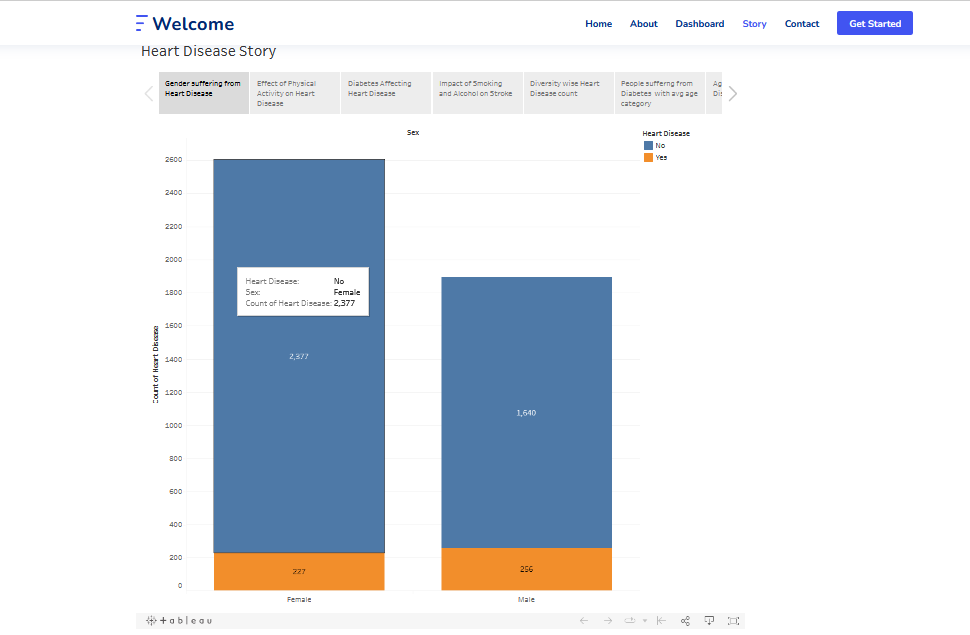
**Activity 1: Embed Dashboard & Story with Flask**

**Explanation video link:**

<https://drive.google.com/file/d/18TXWaVBldKV4r1TQ4mLit9aFQzpaJm_L/view?usp=sharing>







**BY**

**T. URMIKA DEVI (**TEAM LEADER)

**R. GOWRI PUSHPA**

**B. VASU**

**B. RAMBABU**