

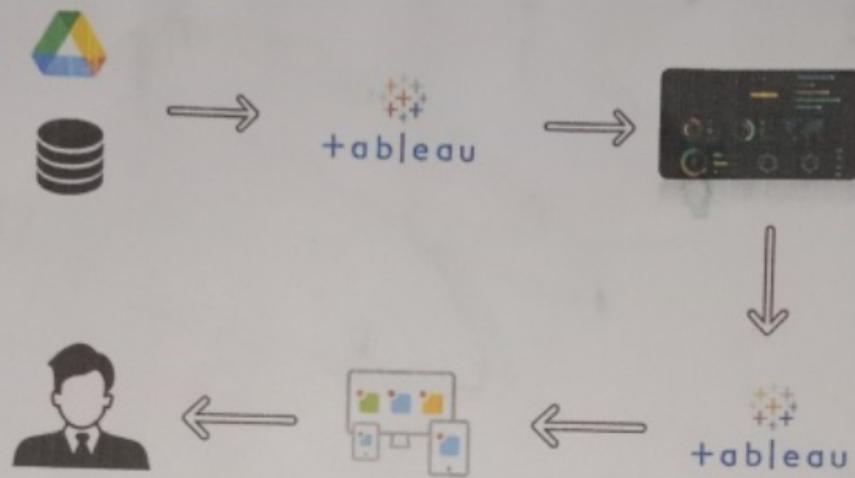
Heart Disease Analysis

Heart disease describes a range of conditions that affect the heart. Heart disease includes blood vessel disease such as coronary artery disease, irregular heart beats (arrhythmias), disease of the heart muscle, heart valve disease.

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

In this Heart Disease Analysis project we are trying to analyze the Heart Disease related data and be able to extract some insights from that data using some tools. To extract the insights from the data and put the data in the form of visualizations, dashboards and story we employed a Tableau tool.

Technical Architecture :-



Project Flow :-

To accomplish this, the listed below activities has to be done.

- Define Problem / Problem understanding.
 - Specify the business problem
 - Literature Survey
 - Business Requirements
 - Social or Business Impact

- Data Collection and Extraction from DataBase
 - Collect the data set
 - Storing data in DB
 - Perform SQL operations
 - Connect DB with Tableau
- Data Preparation
 - Prepare the data for visualizations.
- Data Visualization
 - No. of Unique Visualizations.
- Dash Board
 - Responsive and Design of Dashboard.
- Story
 - No. of stories of a 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
- Perfect Demonstration and Documentation.
 - Record explanation video project end to end solution.
 - Project Documentation- Step by step project development procedure.

Mile Stone 1: Define Problem / Problem understanding

Activity 1: Specify the business problem

Refer for 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 comprising factors of Heart Disease, 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

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 of 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-analysis to provide a comprehensive overview of the current knowledge landscape in the field of heart disease.

Activity 4: Social & 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 health care access, and research advancements, heart disease analysis play a crucial role in improving the well-being of individuals and society as a whole.

Business Impact - Analyzing heart disease has substantial business impacts across various sectors, including health care, 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.

Mile Stone 2: Data Collection and 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

The dataset will be downloaded using the link provided.

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 directly 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 begins to race, 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 only the health problem associated with alcohol consumption.
5. Stroke - Ischemic stroke occurs not only temporary disturbances in the functioning of the heart, but also permanent ones. 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 like 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 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 and perform SQL Operations
Here, we save and store the data in the database and have to perform SQL operations.

Activity 3:

Connect DB with Tableau

Here, we connect database with Tableau and when we open the tableau will be showing a dataset ready.
And that data will be prepared for visualization.

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

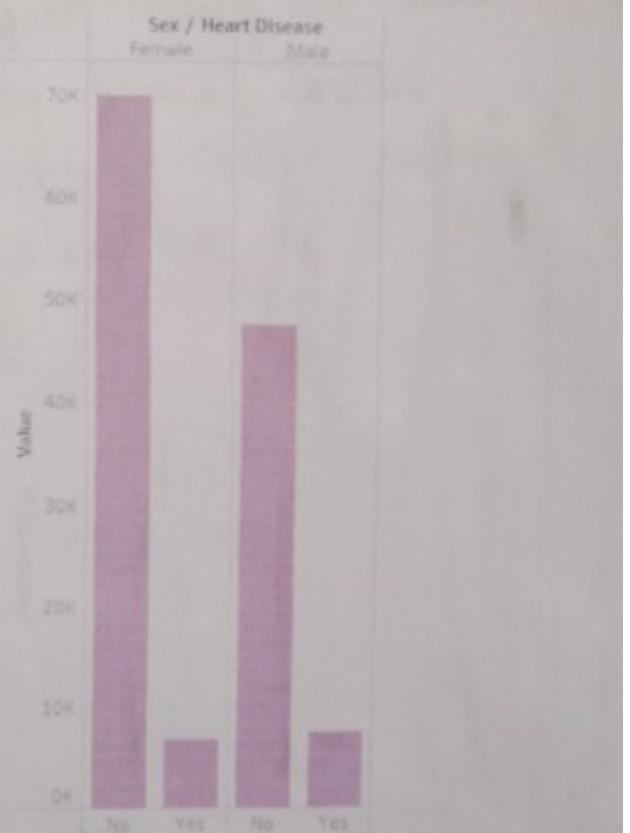
Data visualization is the process of creating graphical representation of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive and easier to interpret. By using visual elements such as charts, graphs, and maps. 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 curated 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

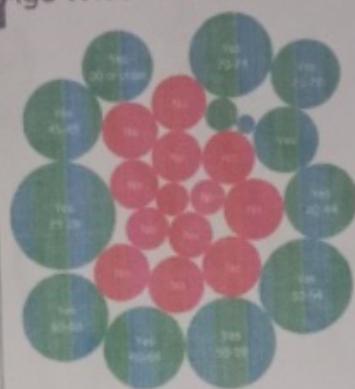
Men typically develop this plaque in the arteries that supply blood to the heart. Women are more likely to develop the buildup in blood vessels, known as microvascular disease. In both sexes is only partly accumulation of cholesterol.



Activity 1.2 : Age Vs Heart Disease

In men, the risk of heart disease or heart attack increase significantly after the age of 45. In women, heart attacks are more likely to occur after the age of 50.

Age Wise Heart Disease



Activity 1.3 : Diabetes Vs Stroke

Diabetes increases the chance of having a stroke, which can damage brain tissue and cause disability or even death. To prevent stroke, people with diabetes should manage blood pressure, cholesterol and weight.

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

Alcohol consumed to excess over several years can produce an alcoholic cardiomyopathy, in which alcohol acts as a toxin to weaken the heart.

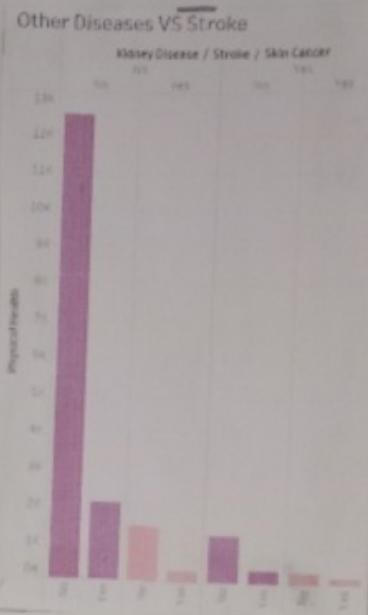
Cigarette smoking also is a strong risk factor for congestive heart failure in general population.

Activity 1.5 Other Health Disease Vs Stroke.

Stroke is a condition in which the brain can't get enough blood flow.

This happens because one or more blood vessels leading to the brain are blocked or have burst.

Other forms of include coronary heart disease.



Activity 1.6 - Race wise Heart Disease

It tells about the Race wise Heart disease.

Activity 1.7 - General Health Vs Heart Disease

This visualization tells us about the General Health Vs Heart Disease.

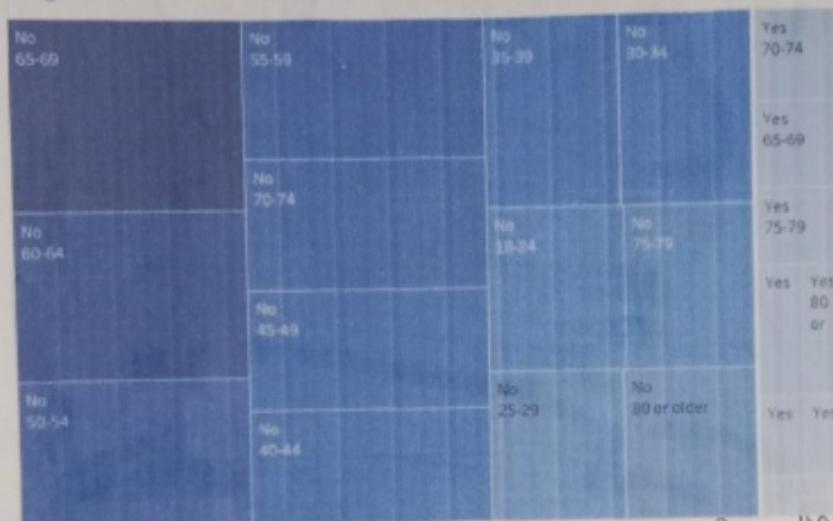
Activity 1.8 - Physical Activity Vs Heart Disease

This visualization shows about the Physical activity Vs Heart Disease.

Activity 1.9 - Age and BMI vs Diabetic

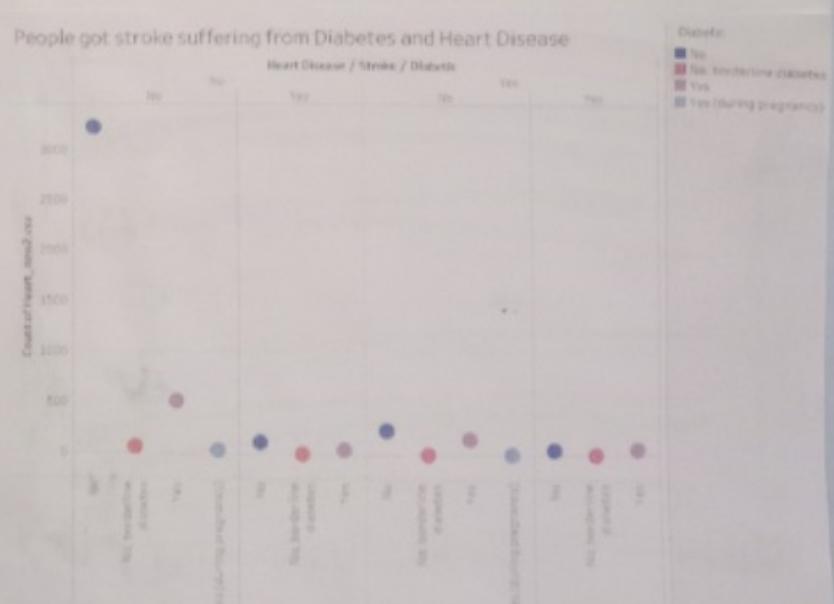
Higher BM was associated with increased risk of AMI and CHD among both men and women.

This visualization tells us about Age and BMI Vs Heart Disease



Activity 1.10 - People got stroke suffering from Heart Disease and Diabetic.

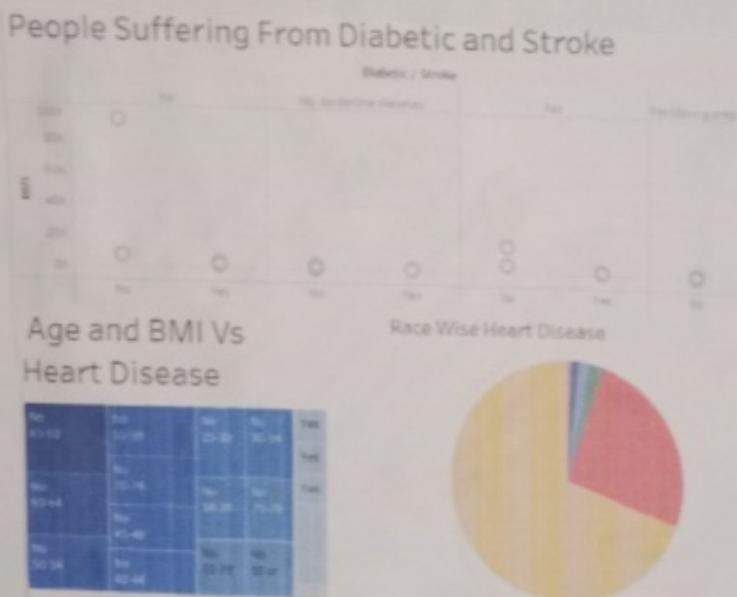
This visualization shows us about on tells us about the people got stroke suffering from Heart Disease and Diabetic.



Milestone 5: Dash Board.

A dashboard is a graphical user interface 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 typically designed for a specific purpose. They can be used to track key performance indicators, monitor performances metrics and display data in the form of charts, graphs and tables.

ACTIVITY 1 - Responsive and Design of Dash Board.

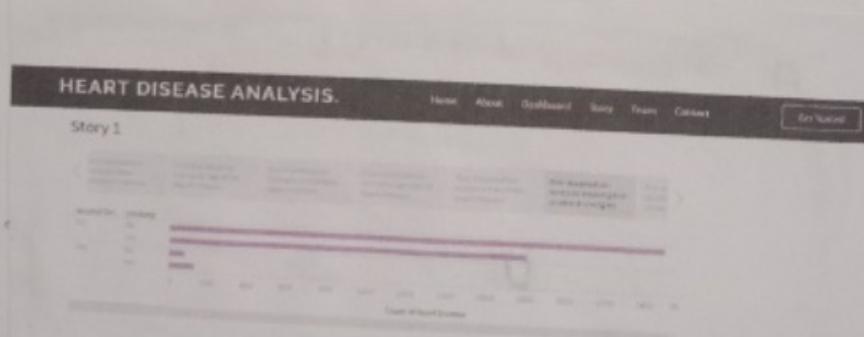


Milestone 6: Story.

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. Data stories can be told using a variety of mediums, such as reports, presentations, 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 or 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.



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 (?) button to get the information related to table such as column, table, rows etc.

SCHEMAS

Filter objects

- ▼ data
 - Tables
 - heart_new2
 - Columns
 - Indexes
 - Foreign Keys
 - Triggers
 - Views
 - Stored Procedures
 - Functions
- ▶ demo
- ▶ sys

Schemas

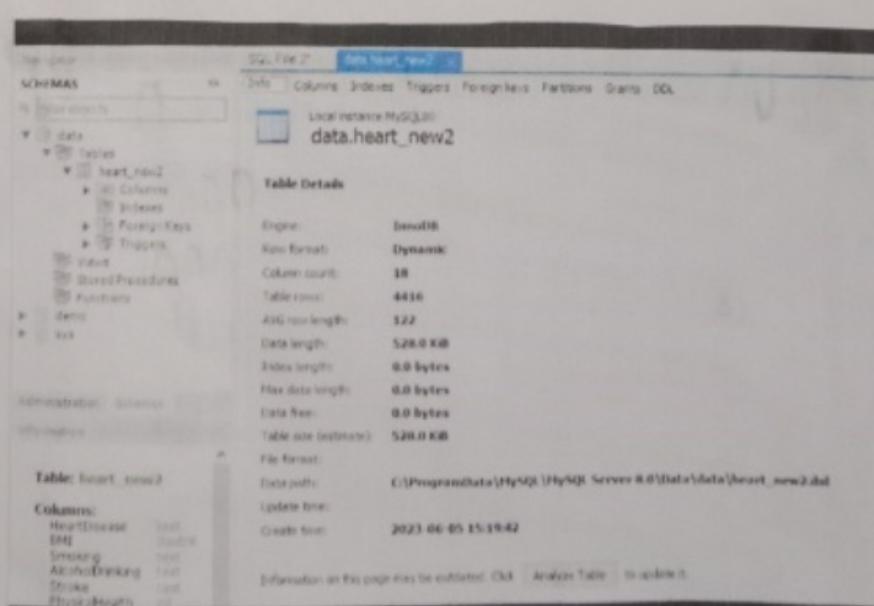
Local instance MySQL 8.0

data.heart_new2

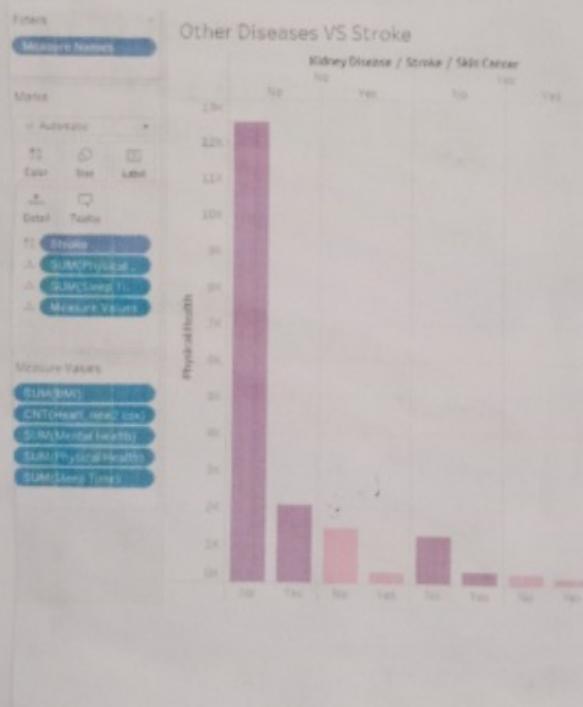
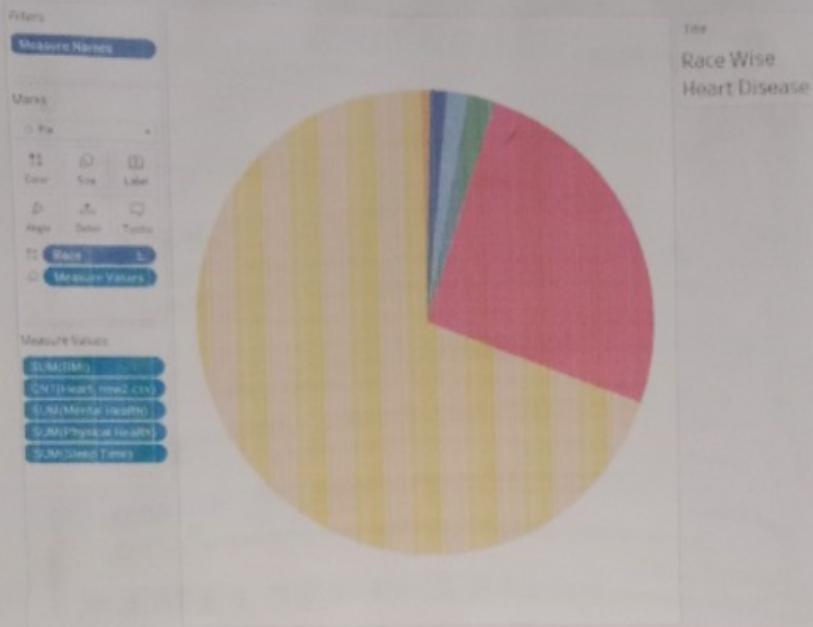
Table Details

Engine	InnoDB
Row format	Dynamic
Column count	18
Table rows	4416
Avg row length	322
Data length	528.0 KB
Index length	0.0 bytes
Max data length	0.0 bytes
Data free	0.0 bytes
Table size (estimated)	528.0 KB
File format	
Data path	C:/ProgramData/MySQL/MySQL Server 8.0/Data/data/heart_new2.fld
Update time	
Create time	2023-06-05 15:19:42

Information on this page may be outdated. Click Analyze Table to update it.



Activity 2: Utilization of Data Filter



Activity 3: No. of calculation fields.

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

Activity 4: No. of visualizations / graphs.

1. Gender vs Heart Disease

2. Age vs 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 vs heart Disease

7. General health vs Heart Disease

8. Physical Activity vs Heart Disease.

9. Age and BMI vs Heart Disease

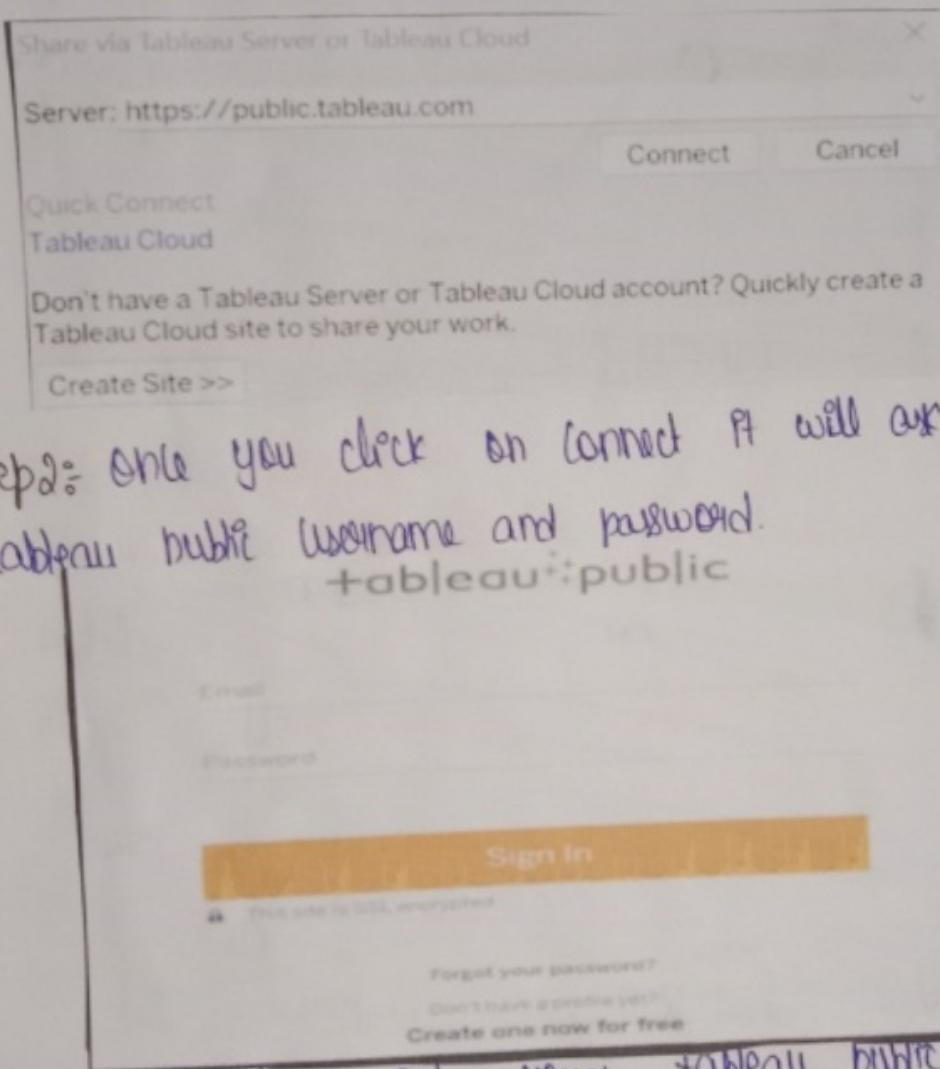
10. People get stroke suffering from Diabetes and Heart Disease.

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

Step1: 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.



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

+tableau@public

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

Activity 1: Embed Dashboard and Story with flask.



HEART DISEASE ANALYSIS.

If happens suddenly when a artery becomes wider pressure builds up on the arteries that supply your heart muscle with blood. The plaque narrows an artery over time so the heart muscle will not need to other heart problems. Angina - where pain from lack of blood flow.

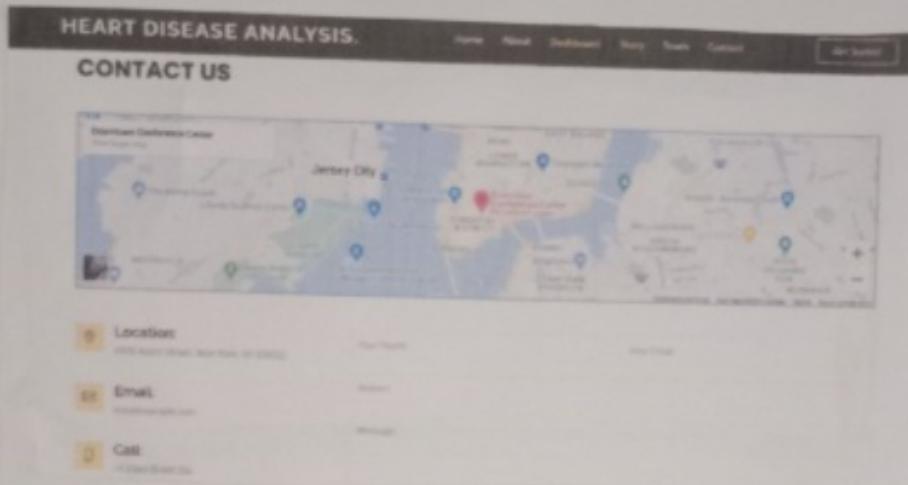
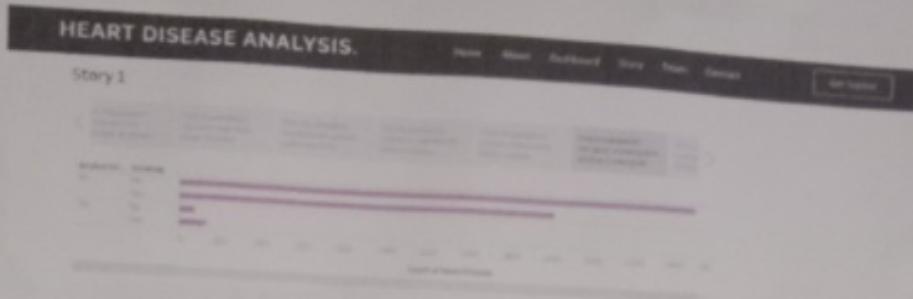
The name "heart disease" refers to several types of heart conditions. The most common type of heart disease is one caused by plaque in coronary artery disease (CAD), which affects the blood flow to the heart.

Heart disease is the top cause of death in the United States. In 2021, heart disease caused 1 in 5 deaths. That's nearly 700,000 people. Coronary artery disease, the most common type of heart disease, caused almost 375,000 of those deaths.

A type of disease that affects the heart or blood vessels. The risk of having heart disease may be increased by smoking, high blood pressure, high cholesterol, unhealthy diet, lack of exercise, and obesity.

Heart disease treatment depends on the cause and type of heart damage. Healthy lifestyle habits — such as eating a low-fat, low-salt diet, getting regular exercise and good sleep, and not smoking — are all important parts of treatment.





Milestone 9: Project Demonstration and Documentation
Below mentioned deliverables to be submitted along with
other Deliveries.

Activity 1:
Record explanation video for the project's end to end solution

Activity 2:
Project Documentation Step by Step project development procedure.