



LONG-TERMINTERNSHIP



**Dr. LANKAPALLI BULLAYYA COLLEGE
VISAKHAPATNAM**

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, 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 analysis the heart disease related data and 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.

*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.

* 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.

* Data Preparation

Data Preparation is a crucial step in the data analysis process. It involves transforming raw data into a clean, structured format that is suitable for analysis.

* Amount of Data Rendered To DB

- 1) 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.
- 2) 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.

* Web Integration

Publishing help 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.

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

*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 & even cured. The business requirements for analyzing the heart disease in world include identifying patterns and comparing factors of heart disease. The ultimate goal is to gain insights and improve performance through data visualization techniques.

Activity 3: Literature Survey

A literature survey for heart disease would involve researching & reviewing previous studies, articles on topic. This could include information on the methods and techniques used for analyzing heart disease. A comprehensive literature survey should include journals, scientific, conference proceedings. The survey should encompass a range of clinical trials, studies, meta-analyses, to provide a comprehensive overview of the current knowledge landscape in the field of heart disease.

Activity 4: Social & Business Impact

Social Impact - Analyzing heart disease has profound social impacts, ranging from individual-level health outcomes to community empowerment and public health initiatives.

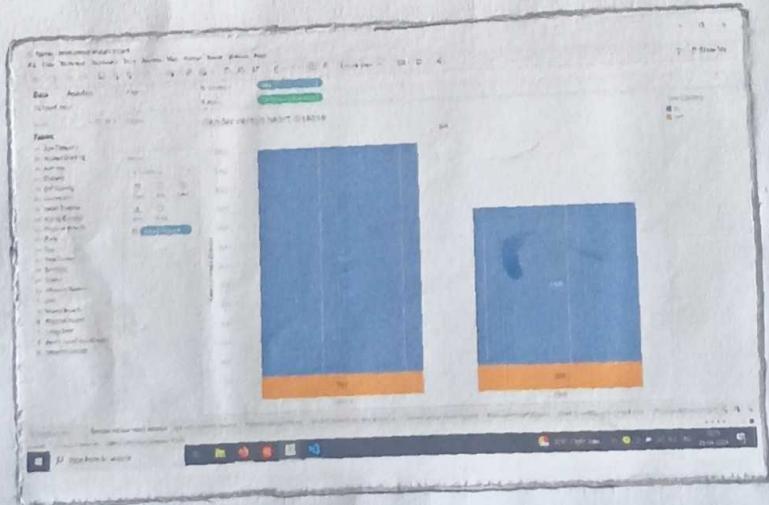
Business Impact - Analyzing heart disease has substantial business impacts across various sectors, including healthcare, medical technology, digital health, insurance, research, workplace wellness, and consumer products.

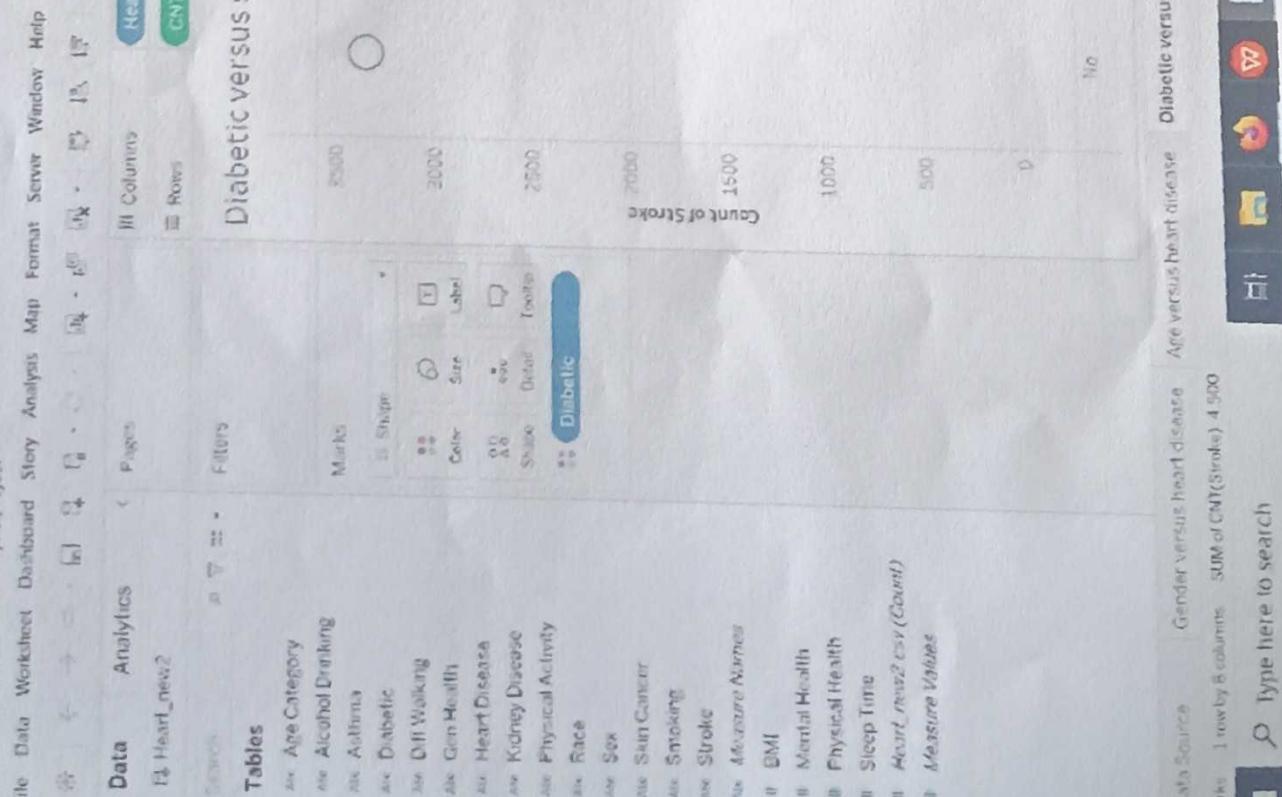
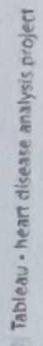
* Data visualization

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.

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 maps, charts etc.





Diabetic Versus stroke :-

In this visualization we have to drop Heart Disease, Diabetic to the columns. And drop stroke to the rows. Add Diabetic to the Marks Give a color to it.

In a show me chart, click on circle view. we can see the visualization in a circular shape. click on it.

we observe that

Diabetic : Yes

Heart Disease : NO

Count of stroke : 36

Physical Activity Versus Heart Disease:-

For this visualization, we have to drop Physical Activity in columns.

Heart Disease to the rows. Give a click on it. Select Measure and then select Count, we can see CNT(Heart Disease)

Add filters to Heart Disease.

Add a calculation to the Rows

ARG01(arg0) ARG01(arg0)

Observe the "Marks" column and select the Second ARG01(arg0). Adjust the size, and give a color to it.

In Rows Column select the second ARG01(arg0). Click on it and select Dual Axis.

We observe that Donut chart.

Click on Heart Disease to the label.

Physical Activity : Yes

Count of Heart Disease : 279

arg0) : 0

The screenshot shows a Tableau dashboard titled "Age versus BMI versus Diabetic". The dashboard has a light blue background and features several data visualizations:

- Top Panel:** A chart showing the average BMI for different age groups. The y-axis ranges from 23.906 to 33.710. The x-axis categories are 45-49, 50-54, 55-59, 60-64, 65-69, and 70-74. The values are: 45-49 (23.906), 50-54 (28.000), 55-59 (33.710), 60-64 (30.000), 65-69 (30.000), and 70-74 (30.000).
- Second Panel:** A chart titled "Diabetic: Yes (during pregnancy)" showing the average BMI for pregnant women. The y-axis ranges from 25.29 to 25.54. The x-axis categories are 25-29 and 30-34. The values are: 25-29 (25.29) and 30-34 (25.54).
- Third Panel:** A chart titled "Yes (during pregnancy)" showing the average BMI for pregnant women. The y-axis ranges from 26.24 to 26.59. The x-axis categories are 26-29 and 30-34. The values are: 26-29 (26.24) and 30-34 (26.59).
- Fourth Panel:** A chart titled "Yes (during pregnancy)" showing the average BMI for pregnant women. The y-axis ranges from 26.24 to 26.59. The x-axis categories are 26-29 and 30-34. The values are: 26-29 (26.24) and 30-34 (26.59).
- Bottom Panel:** A chart titled "Yes (during pregnancy)" showing the average BMI for pregnant women. The y-axis ranges from 26.24 to 26.59. The x-axis categories are 26-29 and 30-34. The values are: 26-29 (26.24) and 30-34 (26.59).

The dashboard includes various filters and data source information at the bottom.

Age Versus BMI Versus Diabetic :-

First of all we have to drop Age category in columns. BMI to the Rows. And we have to right click to the measure and select AVG.

Add Diabetic to the filters.

We can see the pop-up box.

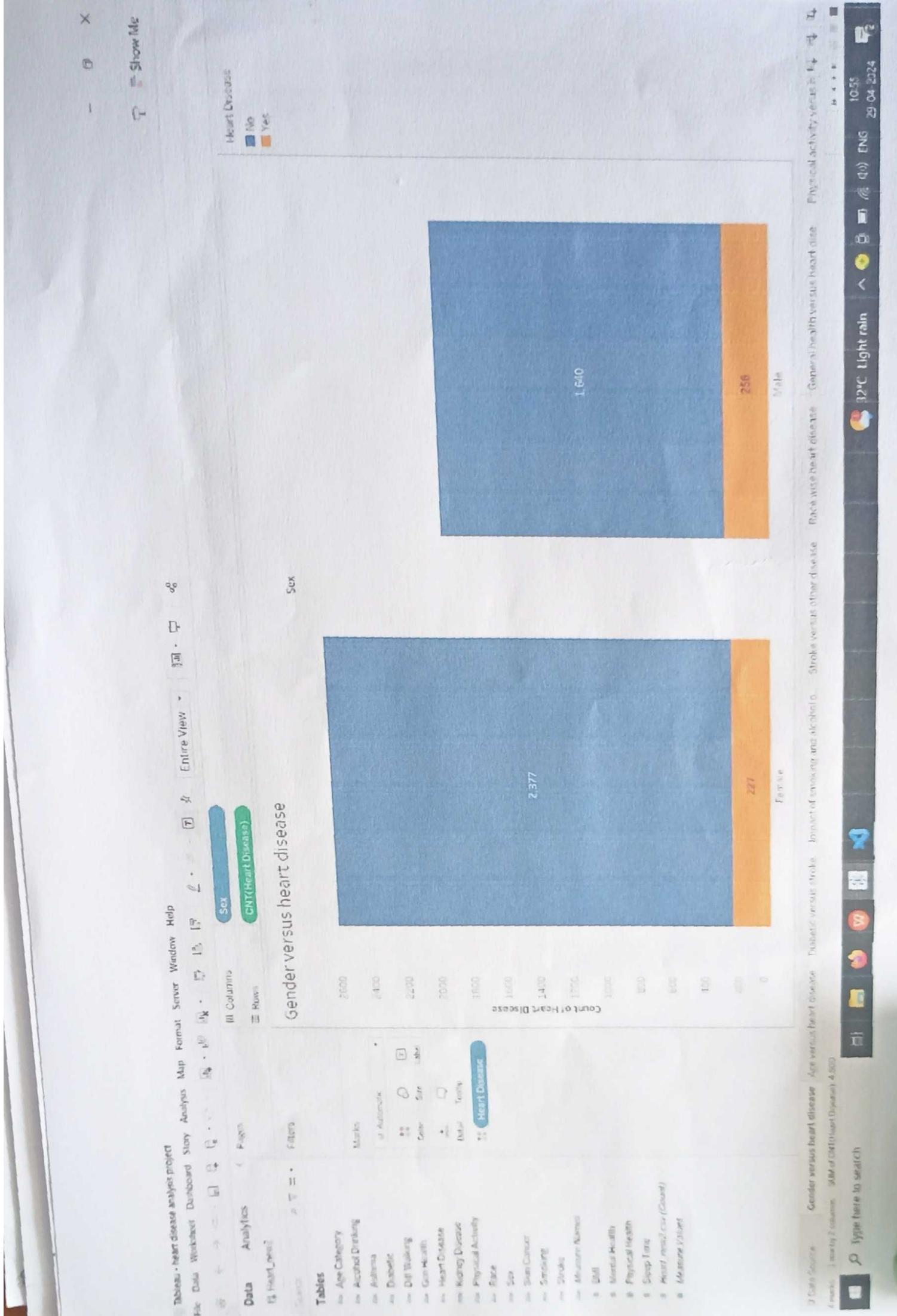
Click on Yes button and then select Tree Map.

Add Diabetic to the Label and we can see that.

The persons who have the age in between age 18-24 :- The Avg BMI :- 30.670

Age :- 25-29 = The Avg BMI :- 32.260

Age - 30-34 = The Avg BMI :- 33.394



Gender Versus Heart Disease :-

first we have to drag 'sex' in column. And Heart Disease in rows.

We see that bar chart.

We observe that male and female categories.

Count of Heart Disease : 227

Sex : female

Heart Disease : Yes

Count of Heart Disease : 256

Sex : Male

Heart Disease : Yes

In female and male categories
we observe that male having heart
disease higher than female.

Age Versus Heart Disease :-

for this visualization, drop age category to the columns. sex category to the columns.

Drop heart disease to the rows, and give a right click on it and select measure to then select count.

In a Marks Column, drop heart disease and add color to it.

We observe that in visualization we can see that sex category, the female and male who have heart disease in a particular age.

Impact of Smoking and Alcohol on Heart Disease

for this visualization, drop heart disease to the columns.

Drop smoking and alcohol drinking to the rows.

In a Marks column, drop stroke, add color to it, and heart disease in detail.

In visualization we observe that,

Alcohol Drinking : NO

smoking : Yes

Stroke : Yes

Count of Heart Disease : 139

Distinct count of Heart Disease : 2

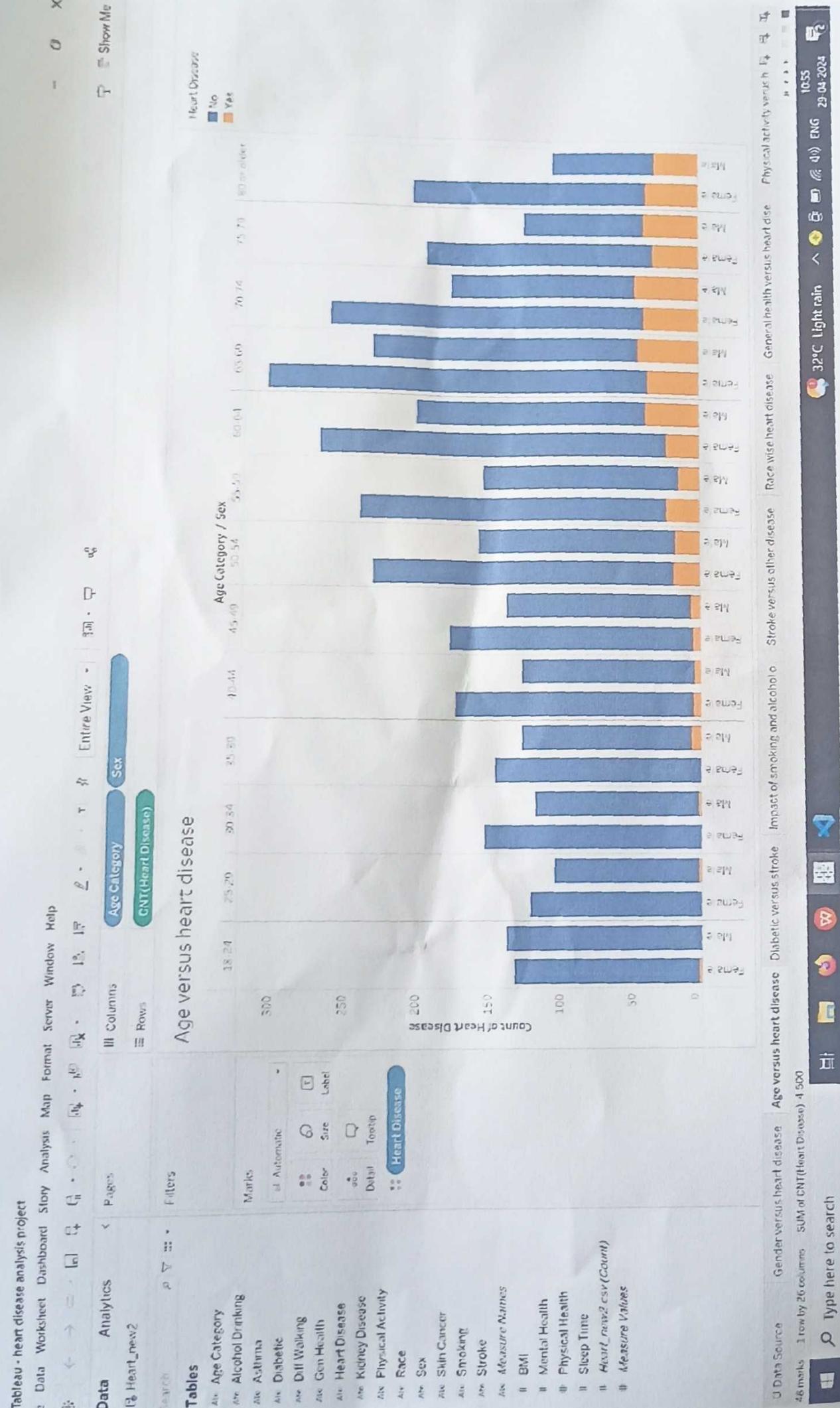
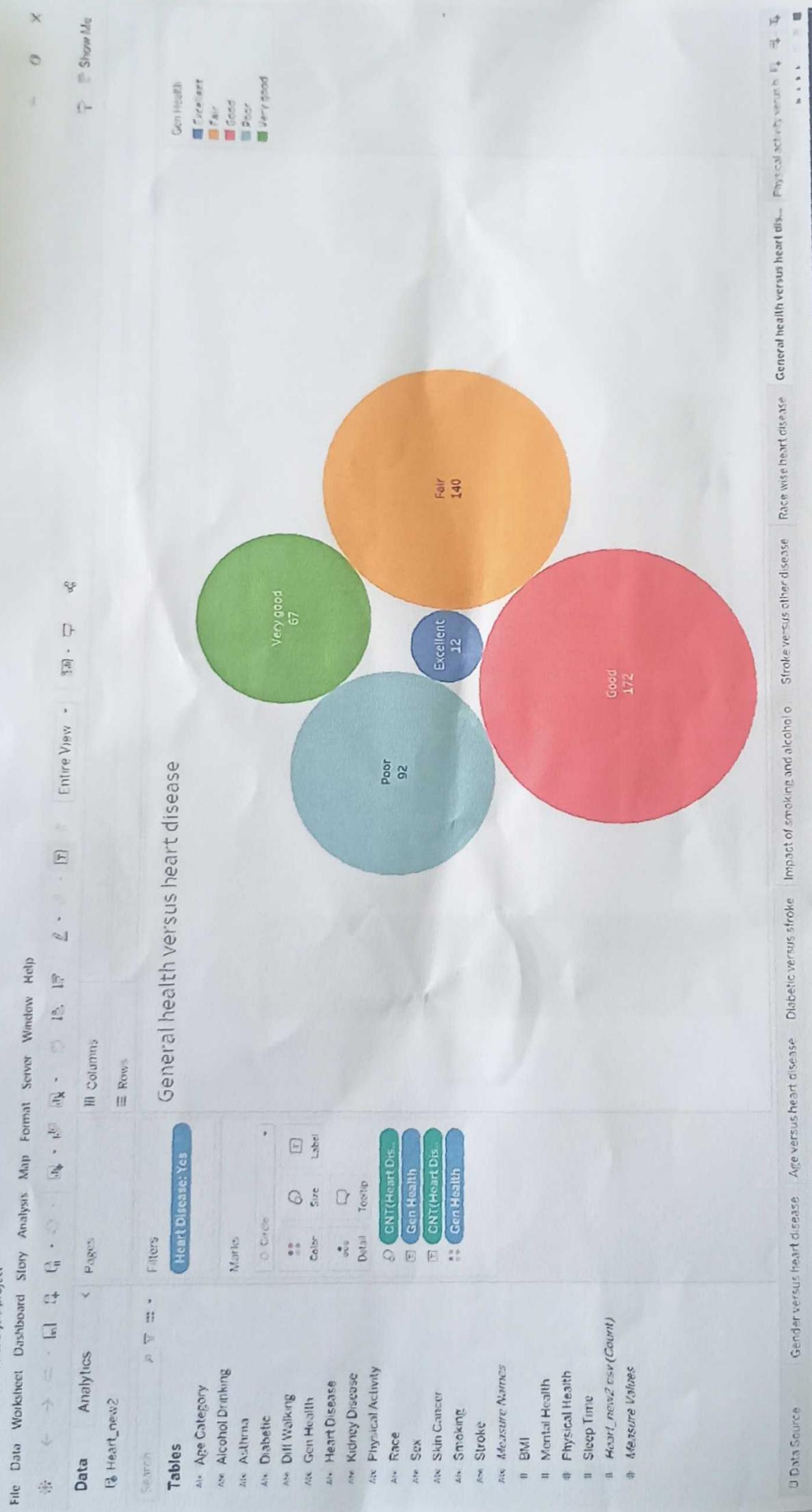


Tableau - heart disease analysis project

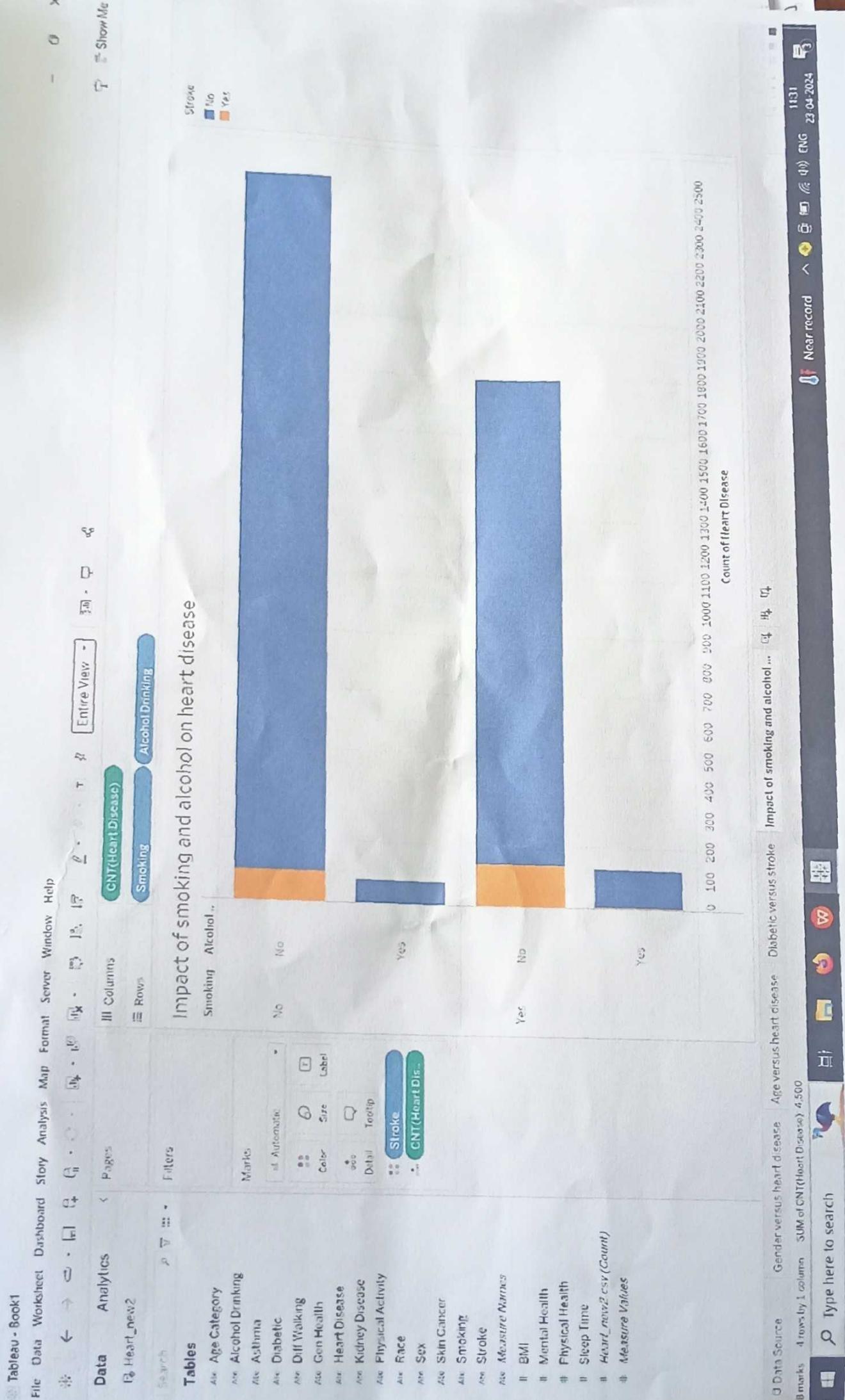


Tables

All	Age Category	No	Yes
All	Alcohol Drinking	2000	2000
All	Asthma	2000	2000
All	Diabetic	2000	2000
All	Difl Walking	2000	2000
All	Gen Health	2000	2000
All	Heart Disease	2000	2000
All	Kidney Disease	2000	2000
All	Physical Activity	2000	2000
All	Race	2000	2000
All	Sex	2000	2000
All	Skin Cancer	2000	2000
All	Smoking	2000	2000
All	Stroke	2000	2000
All	Measure Names	1500	1500
All	BMI	1000	1000
All	Mental Health	1000	1000
All	Physical Health	1000	1000
All	Sleep Time	1000	1000
All	Count_heart.csv (Count)	500	500
All	Measure Values	500	500

Tables







Dashboard :-

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 are typically designed for creating a beautiful dashboard.

The screenshot shows a Tableau dashboard titled "Tableau - heart disease analysis project". The dashboard includes the following components:

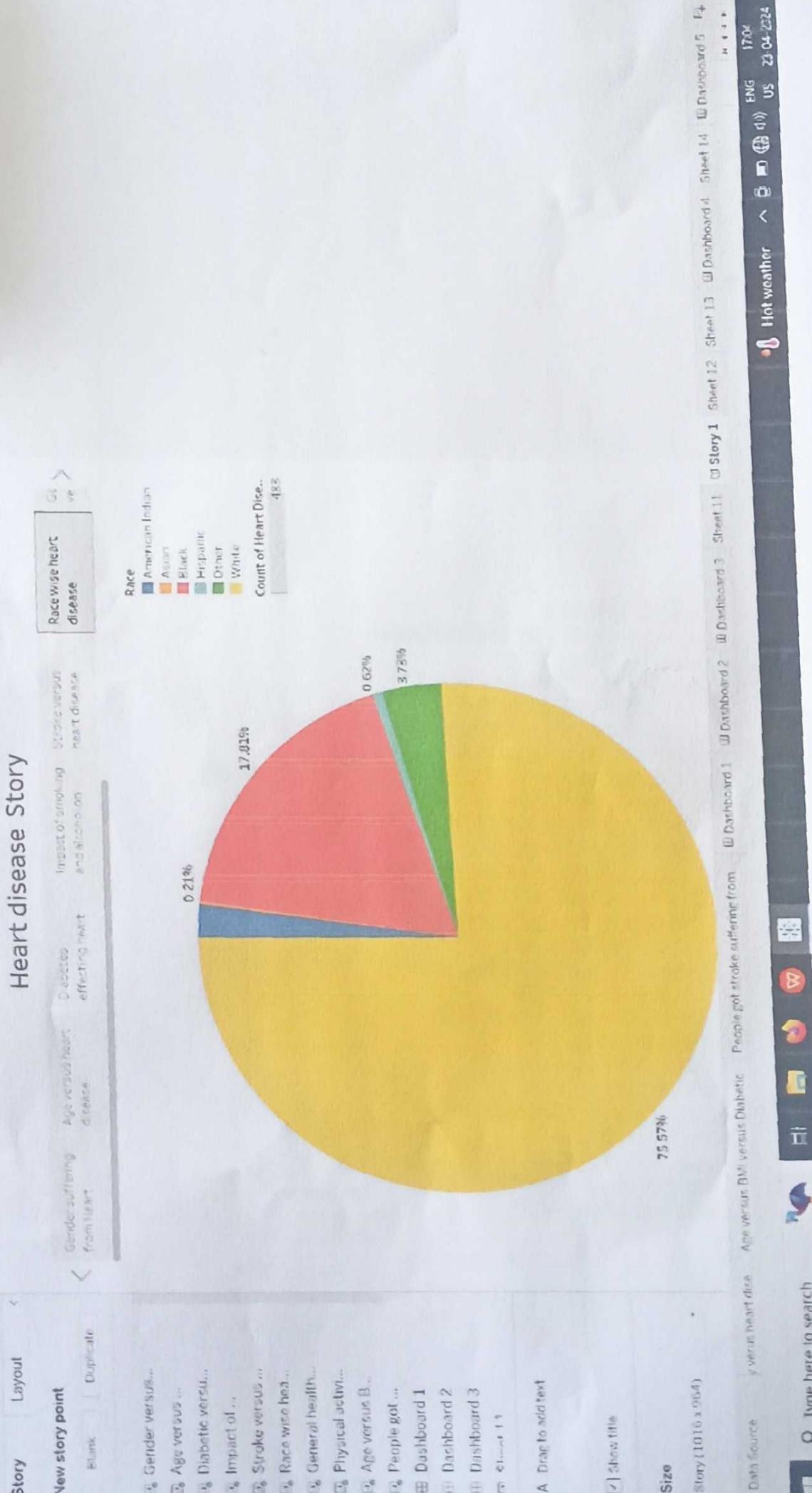
- Sheets:** Default, Phone, Dashboard Layout, Stroke versus other disease, Race wise heart disease, General health versus heart disease, Physical activity versus heart disease, Impact of ..., and Diabetes versus ...
- Widgets:**
 - Stroke versus other disease:** A bar chart comparing the count of stroke cases across different diseases. The y-axis ranges from 0 to 500.
 - Race wise heart disease:** A pie chart showing the distribution of heart disease by race. The categories and percentages are: American Indian (0.21%), Asian (17.01%), Black (75.57%), Hispanic (3.73%), Other (0.21%), and White (0.21%).
 - General health versus heart disease:** A bubble chart showing the relationship between general health (Poor, Fair, Good) and heart disease. The bubbles are labeled with counts: Poor (279), Fair (140), and Good (172).
 - Physical activity versus heart disease:** A donut chart showing the distribution of physical activity levels (Excellent, Fair, Good, Poor, Very poor) among those with heart disease. The counts are: Excellent (0), Fair (279), Good (204), Poor (0), and Very poor (0).
- Dashboard Layout:** Includes sections for "Race", "Gender", "Age", "Diabetic", and "Impact of ...".
- Bottom Navigation:** Includes icons for floating, tiled, dashboard title, show dashboard title, data source, and help.

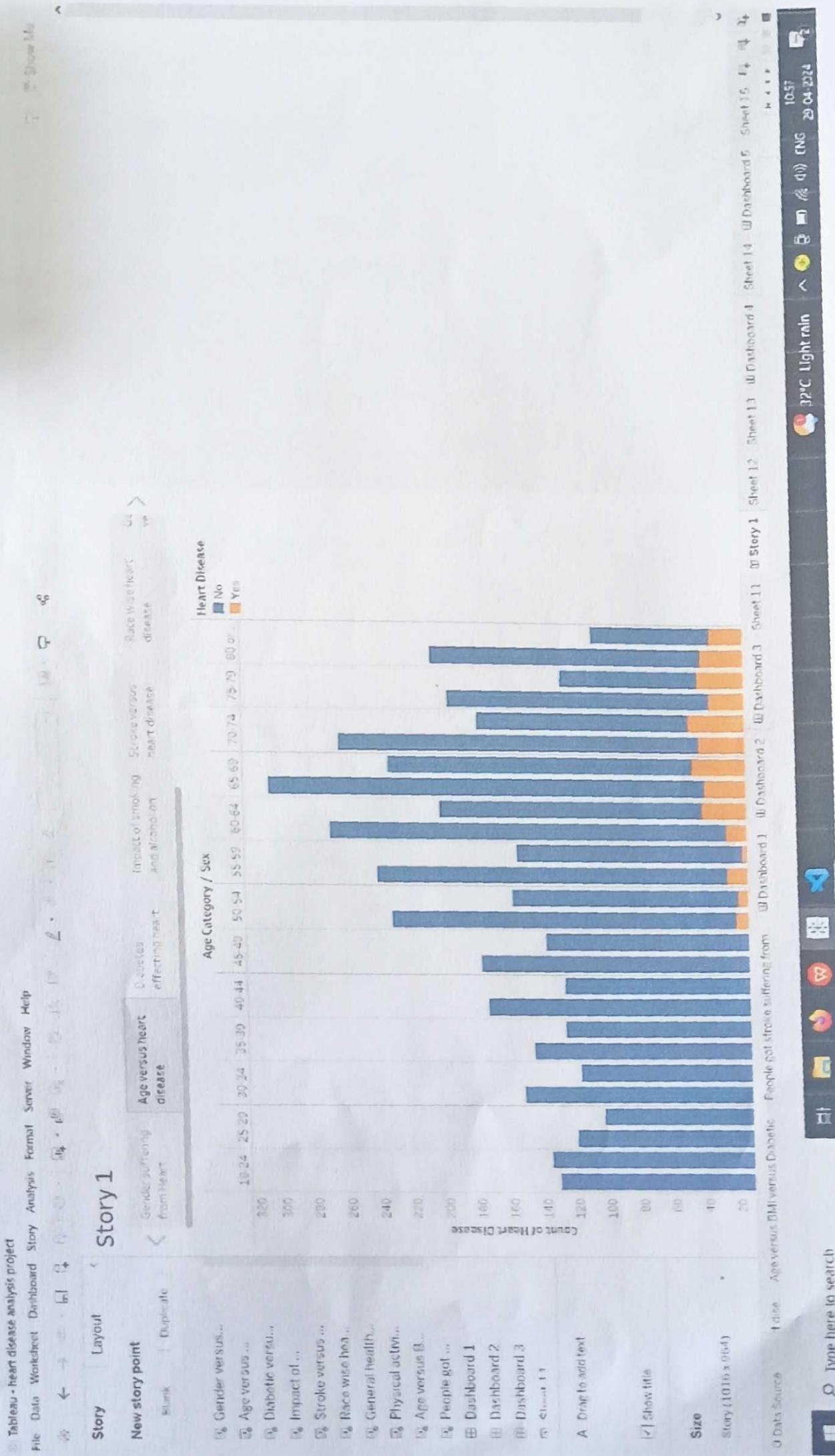
*Story

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. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

Activity :- No of scenes of story
 the NO of scenes in a storyboard for a data visualization Analysis of the heart disease will depend on the specific insights that are trying to conveyed.







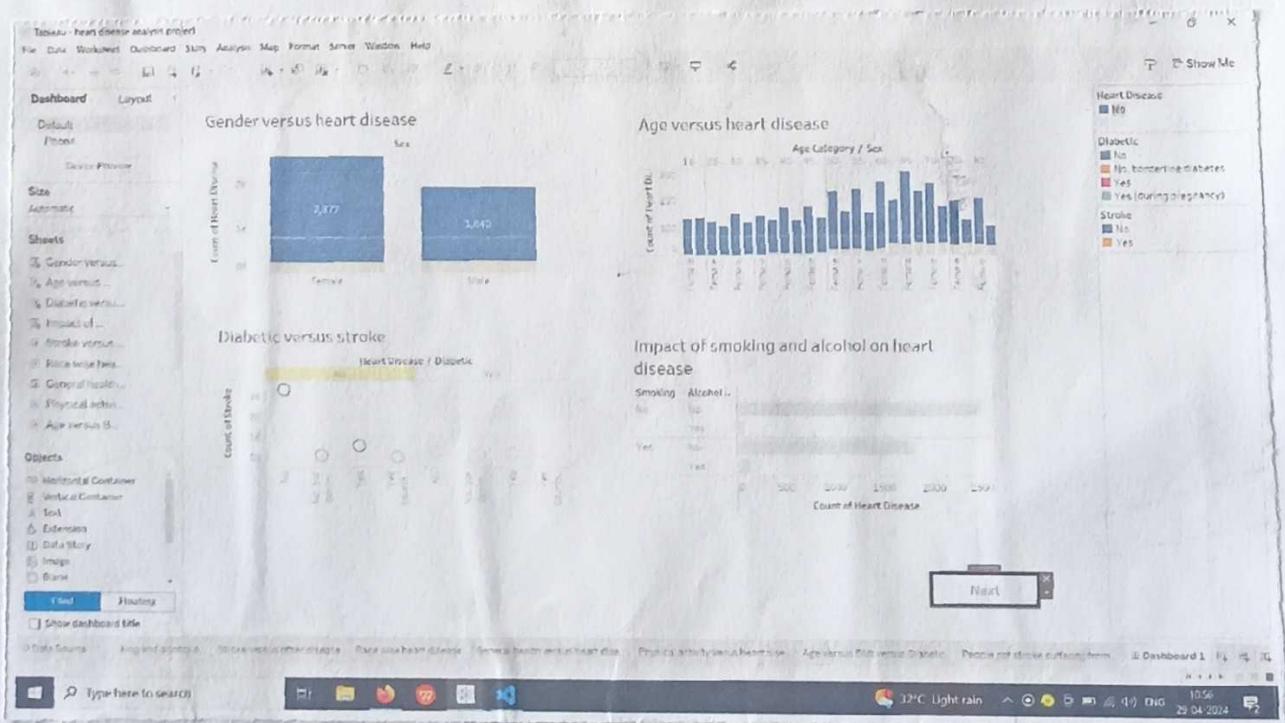
* Performance :-

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 MySQL, go to database and click to expand tables, select table and click on (i) button to get information related to table such as column count, table rows etc.

Activity - 2: Utilization of Data filters



Activity 3: NO of calculation

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

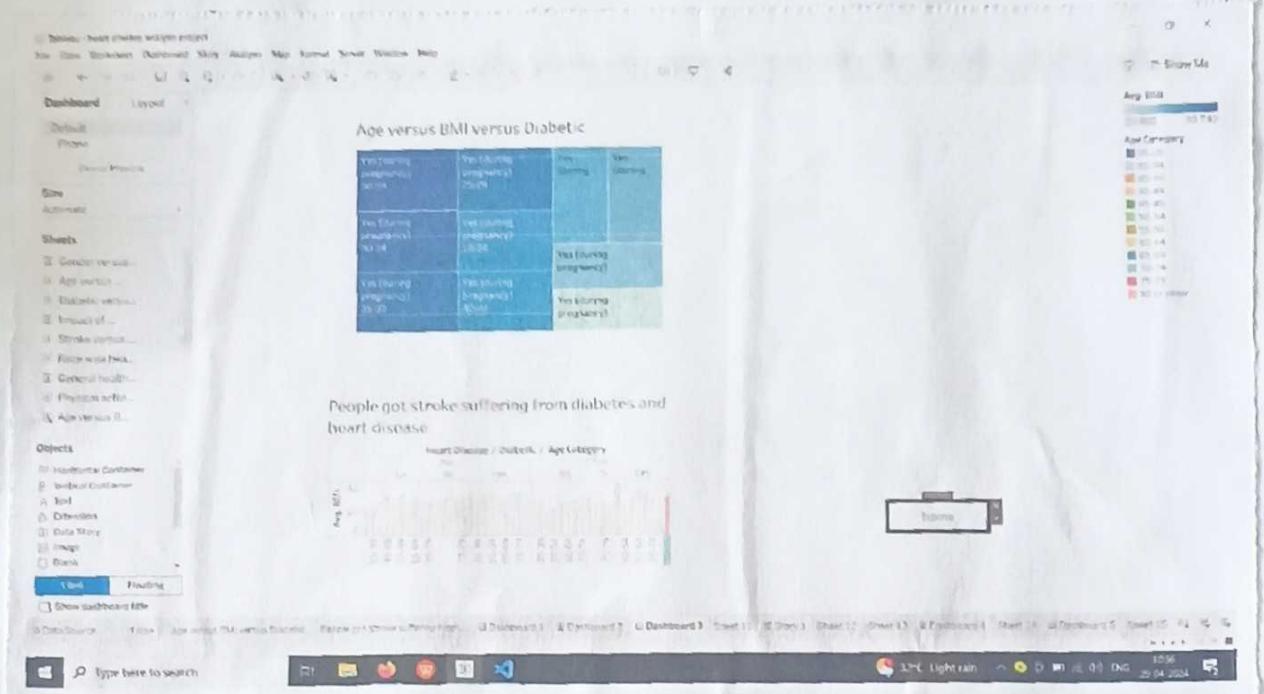
Activity 4: NO of visualizations /Graphs

1. Gender 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

* 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

etc.



Publishing dashboard and reports tableau public
 step 1: Go to dashboard/story, click on share button
 on the top ribbon.

share via Tableau server on Tableau cloud

server : <https://public.tableau.com>

Quick connect
 Tableau cloud

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

[create site >>](#)

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

