Final Project Report

Heart Disease Analysis Using Tableau

Team ID: LTVIP2026TMIDS38963

Team Leader: Gandla Praneetha

Team Member 1: Gandla Praneetha

Team Member 2: K. Divya

Team Member 3: Sirigiri Sumedha

Team Member 4: Gorukanti Susmitha

INTRODUCTION

Project Overview

The Heart Disease Analysis project focuses on analyzing healthcare data to identify patterns and risk factors related to heart disease. The system uses data analytics and visualization techniques to transform raw medical data into meaningful insights. Interactive dashboards and stories are created to help users understand trends and correlations. The project aims to support early detection and awareness of heart health risks.

Purpose

The main purpose of this project is to provide a user-friendly platform for analyzing heart disease data. It helps users visualize complex medical information in a simple manner. The project also aims to assist in identifying high-risk groups and understanding lifestyle impacts on heart health. This supports better decision-making and preventive healthcare.

This Project Aims To

This project aims to analyze heart disease data using modern analytical tools. It focuses on improving data interpretation through visual dashboards and storytelling. The system aims to detect risk patterns and highlight important health indicators. It also aims to promote awareness and encourage healthy lifestyle choices among users.

1. Ideation Phase

1.1 Problem Statement

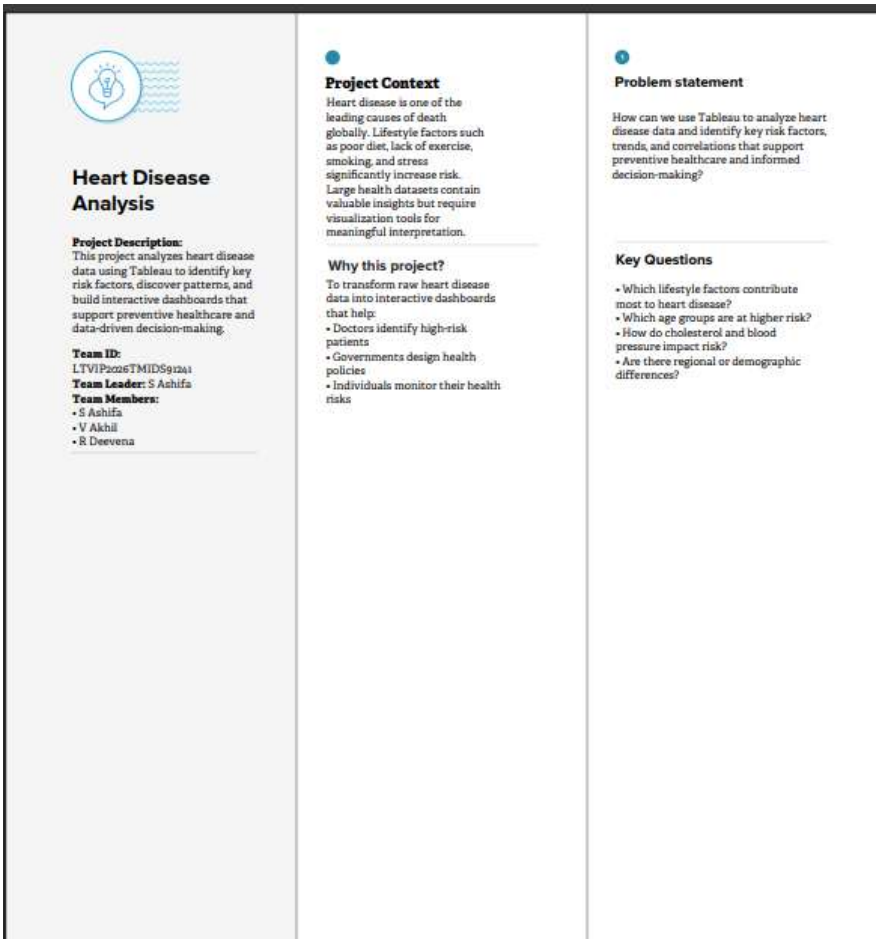
Problem Statement	I am	I'm trying to	But	Because	Which makes me feel
PS-1	a healthcare analyst	predict whether a patient has heart disease	I don't have a clear way to identify high-risk patients early	patient data is complex and risk factors are not obvious	uncertain about making accurate clinical decisions
PS-2	a doctor	understand key factors contributing to heart disease	it is difficult to analyze large patient datasets manually	traditional analysis methods are time-consuming	frustrated and concerned about missing critical insights
PS-3	a patient	know my risk level for heart disease	I don't understand my medical reports properly	medical data is technical and hard to interpret	anxious about my health condition

1.2 Brainstorm

Brainstorm & Idea Prioritization

Many people are unaware of their risk of heart disease until serious symptoms appear, and there is a lack of simple tools that help users understand their heart health using data. Lifestyle factors such as BMI, sleep patterns, alcohol consumption, and smoking habits are often ignored or not properly monitored. Medical reports are usually difficult for common people to interpret without clear visual explanations. Because of this, people fail to take early preventive actions. Using data analytics and visualization for early prediction can help individuals understand risks and prevent life-threatening heart conditions.

Step1 : Team Gathering,project context and problem statement



Heart Disease Analysis

Project Description:
This project analyzes heart disease data using Tableau to identify key risk factors, discover patterns, and build interactive dashboards that support preventive healthcare and data-driven decision-making.

Team ID:
LTVIP2026TMIDS91241
Team Leader: S Ashifa
Team Members:
- S Ashifa
- V Akhil
- R Deevena

Project Context

Heart disease is one of the leading causes of death globally. Lifestyle factors such as poor diet, lack of exercise, smoking, and stress significantly increase risk. Large health datasets contain valuable insights but require visualization tools for meaningful interpretation.

Why this project?

To transform raw heart disease data into interactive dashboards that help:
- Doctors identify high-risk patients
- Governments design health policies
- Individuals monitor their health risks

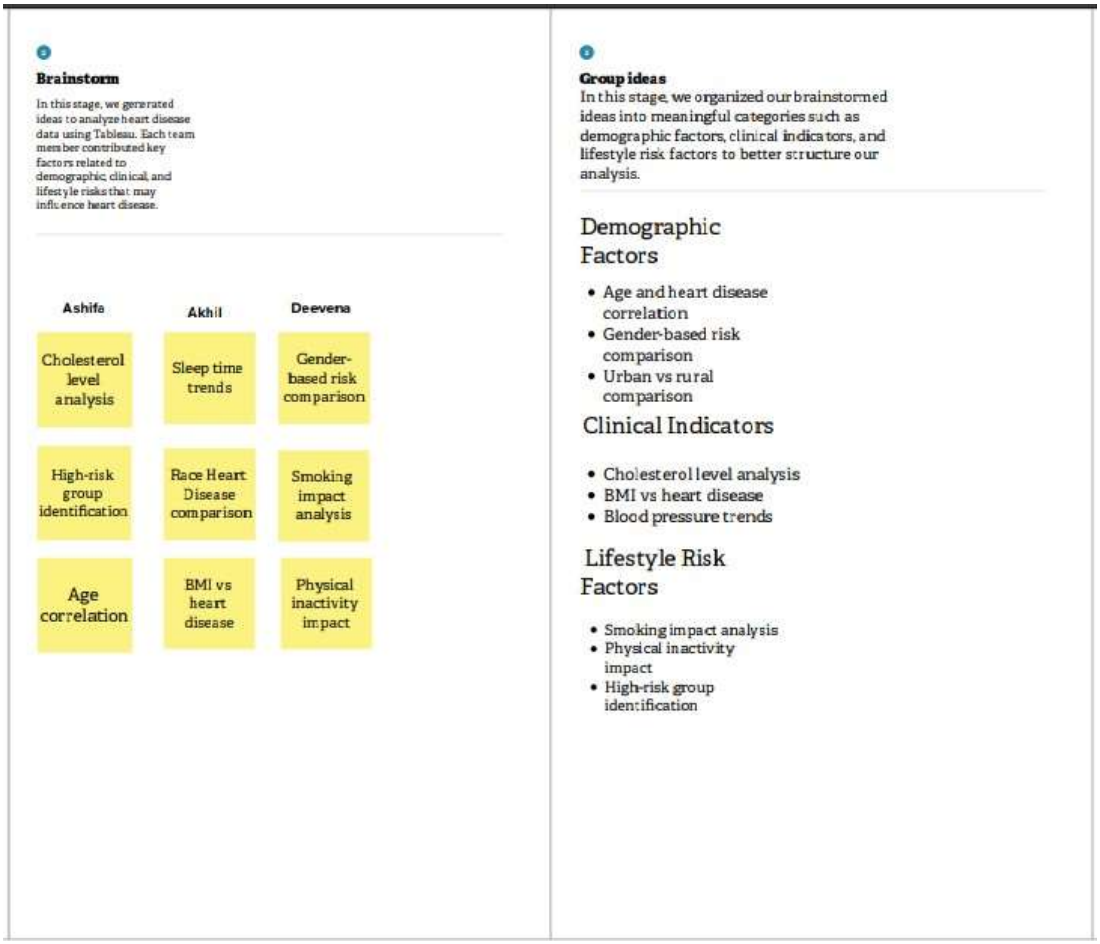
Problem statement

How can we use Tableau to analyze heart disease data and identify key risk factors, trends, and correlations that support preventive healthcare and informed decision-making?

Key Questions

- Which lifestyle factors contribute most to heart disease?
- Which age groups are at higher risk?
- How do cholesterol and blood pressure impact risk?
- Are there regional or demographic differences?

Step 2:Brainstorm,Idea Listing,Grouping and Group Ideas



Brainstorm

In this stage, we generated ideas to analyze heart disease data using Tableau. Each team member contributed key factors related to demographic, clinical, and lifestyle risks that may influence heart disease.

Ashifa	Akhil	Deevena
Cholesterol level analysis	Sleep time trends	Gender-based risk comparison
High-risk group identification	Race Heart Disease comparison	Smoking impact analysis
Age correlation	BMI vs heart disease	Physical inactivity impact

Group ideas

In this stage, we organized our brainstormed ideas into meaningful categories such as demographic factors, clinical indicators, and lifestyle risk factors to better structure our analysis.

Demographic Factors

- Age and heart disease correlation
- Gender-based risk comparison
- Urban vs rural comparison

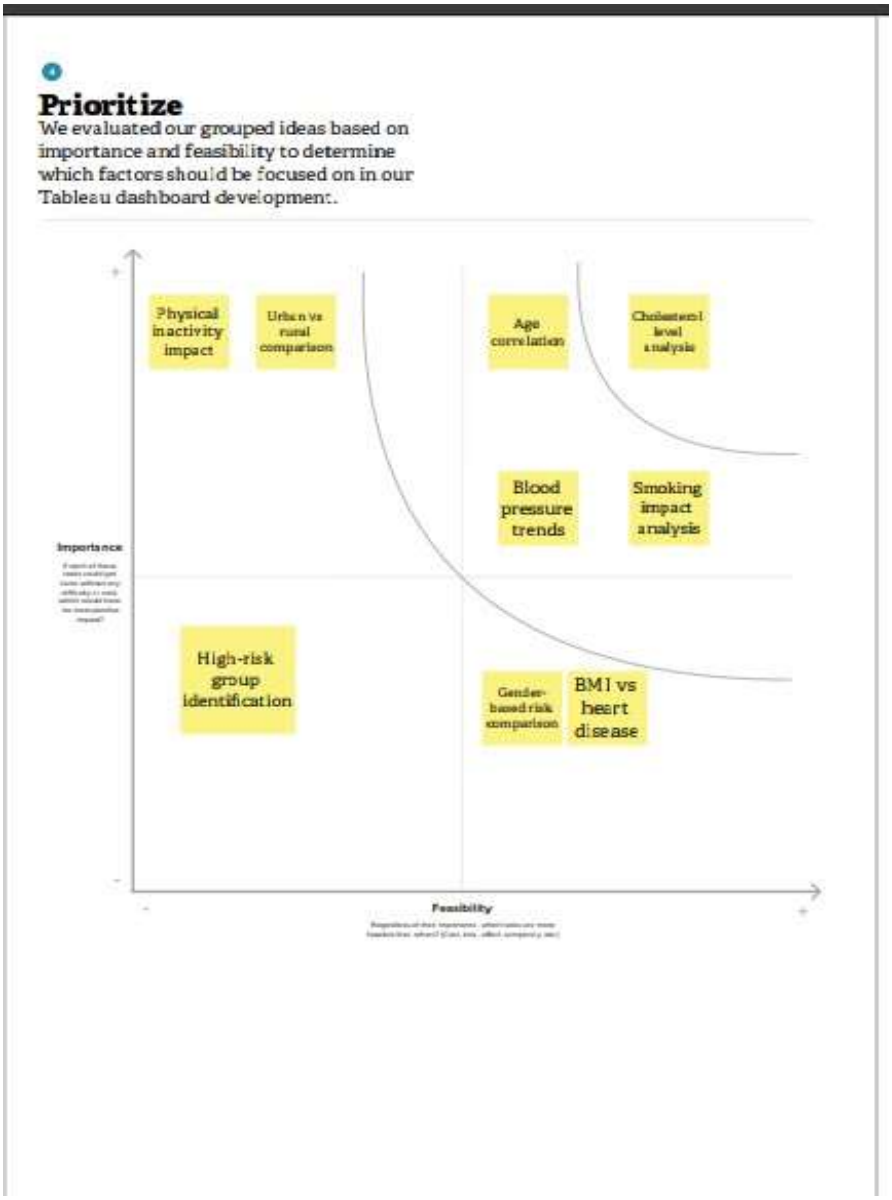
Clinical Indicators

- Cholesterol level analysis
- BMI vs heart disease
- Blood pressure trends

Lifestyle Risk Factors

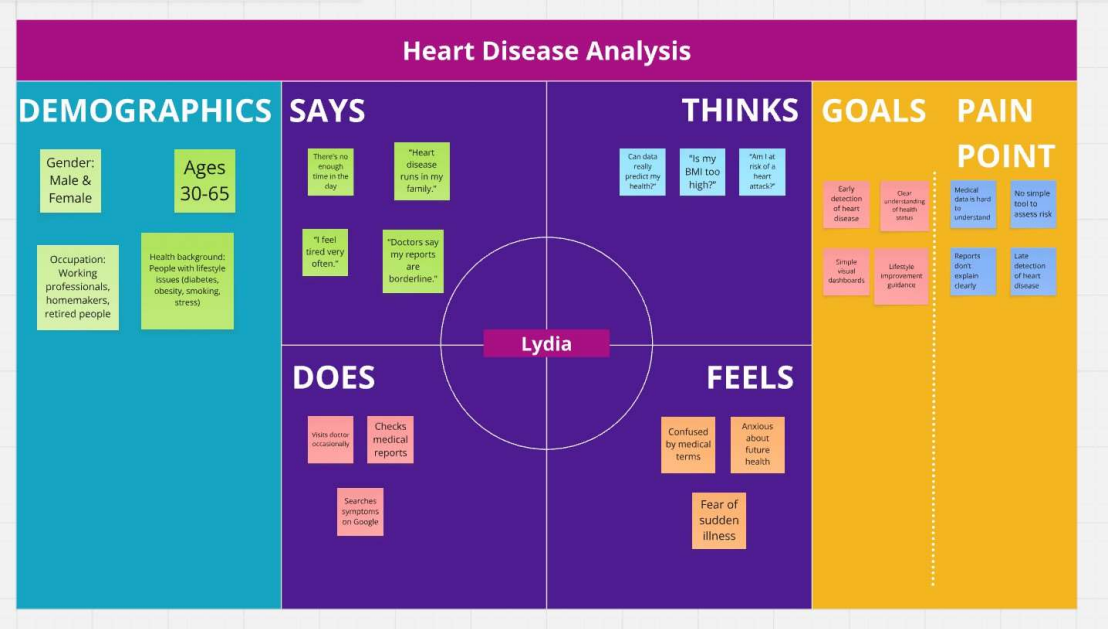
- Smoking impact analysis
- Physical inactivity impact
- High-risk group identification

Step 3:Idea Prioritization



1.3 Empathy Map

The target users are adults with busy lifestyles who are often unaware of their heart health risks. They feel confused by medical reports and find it difficult to understand complex health data. They usually ignore early symptoms and skip regular checkups due to lack of time. This makes them anxious and worried about sudden health problems. Their main goal is to get clear, simple insights that help them detect heart disease early and improve their lifestyle.


Heart Disease Analysis
DEMOGRAPHICS
Gender: Male & Female
Ages 30-65
Occupation: Working professionals, homemakers, retired people
Health background: People with lifestyle issues (diabetes, obesity, smoking, stress)
SAYS
"There's no enough time in the day"
"Heart disease runs in my family."
"I feel tired very often."
"Doctors say my reports are borderline."
THINKS
Can data really predict my health?"
"Is my BMI too high?"
"Am I at risk of a heart attack?"
Lydia
DOES
Visits doctor occasionally
Checks medical reports
Searches symptoms on Google
FEELS
Confused by medical terms
Anxious about future health
Fear of sudden illness
GOALS
Early detection of heart disease
Clear understanding of health status
Simple visual dashboards
Lifestyle improvement guidance
PAIN POINT
Medical data is hard to understand
No simple tool to assess risk
Reports don't explain clearly
Late detection of heart disease

2. Requirement Analysis

2.1 customer Journey Map

2.2 Solution Requirement

Functional Requirements

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through form (Name, Email, Password)
		Registration through Gmail
		Registration through LinkedIn
FR-2	User Confirmation	Confirmation via Email
		Confirmation via OTP
FR-3	Login & Authentication	Login using Email and Password
		Logout from system
FR-4	Dashboard Access	View Scenario 1 Dashboard

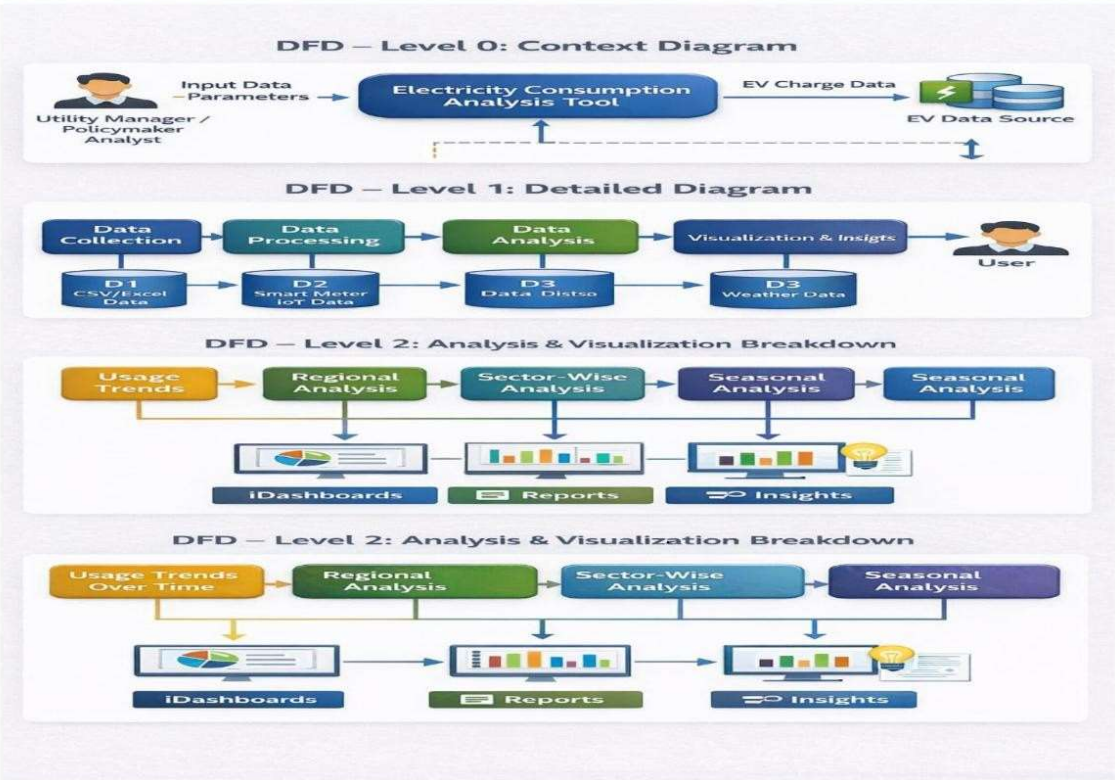
		View Scenario 2 Dashboard
		View Scenario 3 Dashboard
FR-5	Data Input	Enter patient details (age, BP, cholesterol, etc.)
FR-6	Risk Analysis	System calculates heart disease risk
FR-7	Visualization	Display charts, graphs, and KPIs
FR-8	Story View	View story (narrative insights)
FR-9	Report Generation	Download analysis report in PDF
FR-10	Admin Management	Upload datasets
		Remove invalid users

Non-Functional Requirements

Following are the non-functional requirements of the proposed solution.

NFR No.	**Non-Functional Requirement**	**Description**
NFR-1	Usability	The system should be easy to use and understandable by non-technical users.
NFR-2	Security	User data must be protected using authentication and secure access.
NFR-3	Reliability	The system should provide accurate and consistent results without failure.
NFR-4	Performance	Dashboards and reports should load within 2–3 seconds.
NFR-5	Availability	The system should be available 24/7 with minimal downtime.
NFR-6	Scalability	The system should support an increasing number of users and data records.

2.3 Data Flow Diagram



User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance Criteria	Priority	Release
Customer (Web User)	Registration	USN-1	As a user, I can register by entering my name, email, and password to access the Heart Disease Analysis system.	User can successfully create an account and reach the dashboard.	High	Sprint-1
Customer (Web User)	Registration	USN-2	As a user, I receive a confirmation email after registration.	Confirmation email is sent and account is activated.	High	Sprint-1
Customer (Web User)	Registration	USN-3	As a user, I can log in using Google account.	User logs in and accesses dashboards using Google.	Medium	Sprint-2

Customer (Web User)	Login	USN-4	As a user, I can log in using email and password.	User successfully logs into system.	High	Sprint-1
Customer (Web User)	Dashboard Access	USN-5	As a user, I can view heart disease analysis dashboards.	Dashboards load correctly with charts and filters.	High	Sprint-1
Customer (Web User)	Data Input	USN-6	As a user, I can enter health parameters like age, BMI, BP, cholesterol.	System accepts inputs and updates analysis.	High	Sprint-1
Customer (Web User)	Insights	USN-7	As a user, I can see risk level and predictions.	Risk score and insights are displayed clearly.	High	Sprint-1
Customer (Web User)	Reports	USN-8	As a user, I can download reports of my analysis.	PDF report is generated and downloadable.	Medium	Sprint-2
Customer (Web User)	Story View	USN-9	As a user, I can view the story (narrative insights) about heart disease.	Story section loads with explanations.	Medium	Sprint-1

2.4 Technology Stack

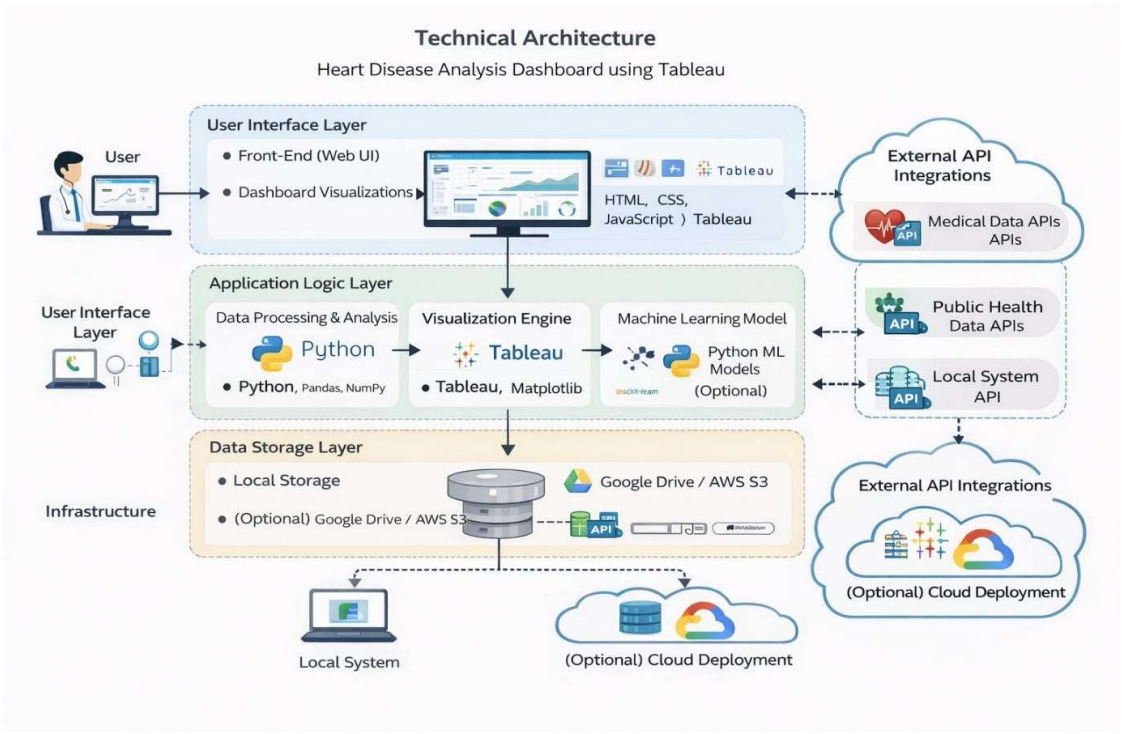


Table 1: System Components

S.No	Component	Description	Technology
1	User Interface	Interface through which users interact with the system to view dashboards and insights	HTML, CSS, JavaScript, Bootstrap
2	Application Logic-1	Handles user authentication, navigation, and routing	Python (Flask)
3	Application Logic-2	Processes data and generates analytical insights	Python (Pandas, NumPy)
4	Application Logic-3	Handles visualization embedding and story flow	Tableau Public Integration
5	Database	Stores heart disease dataset and user details	CSV / SQLite
6	Cloud Database	Stores dashboards and stories online	Tableau Public
7	File Storage	Stores reports and static files	Local File System
8	External API-1	Used for embedding interactive dashboards	Tableau Embed API
9	External API-2	Used for animations and UI effects	AOS (Animate on Scroll)
10	Machine Learning Model	Predicts heart disease risk based on patient features	Scikit-learn (Random Forest / Logistic Regression)
11	Infrastructure (Server / Cloud)	Hosts the application and dashboards	Local System / Tableau Cloud

Table 2: Application Characteristics

S.No	Characteristics	Description	Technology
1	Open-Source Frameworks	Frameworks used to develop the application	Flask, Bootstrap, Pandas
2	Security Implementations	Protects user data and access	Login authentication, session management
3	Scalable Architecture	Supports multiple users and future expansion	3-Tier Architecture (UI, Logic, Data)
4	Availability	Application accessible anytime via web	Tableau Public, Web Browser
5	Performance	Fast loading of dashboards and reports	Optimized queries, caching

3. Project Design

3.1 problem Solution

1. Customer Segments

The primary customers are adults aged 25–60 who are concerned about their heart health. This includes working professionals, elderly people, and individuals with unhealthy lifestyles such as poor diet, lack of exercise, smoking, or family history of heart disease. Secondary users include doctors, healthcare analysts, and researchers who need data-driven insights for medical decision-making.

2. Jobs To Be Done / Problems

Users want to understand their heart disease risk, monitor key health indicators, and take preventive actions before serious conditions occur. They need simple tools to analyze medical data, visualize trends, and get meaningful insights without requiring technical or medical expertise.

3. Customer Constraints

Users face constraints such as lack of medical knowledge, limited time for regular health checkups, difficulty in understanding complex reports, and absence of easy-to-use analytical tools. Many users also hesitate to visit doctors frequently due to cost and accessibility issues.

4. Problem Root Cause

The root cause of the problem is the complexity of medical data and the lack of user-friendly systems that translate raw health data into understandable insights. Traditional healthcare systems focus more on treatment than early detection and preventive analysis.

5. Available Solution Behavior

Currently, users depend on hospital visits, generic medical reports, and online symptom searches. These methods are often manual, time-consuming, and do not provide personalized or predictive insights.

6. Your Solution

The proposed solution is a data-driven Heart Disease Analysis system that uses interactive dashboards and machine learning models to predict risk levels. The system provides clear visualizations, simple KPIs, and narrative stories to help users understand their health status and take early preventive measures.

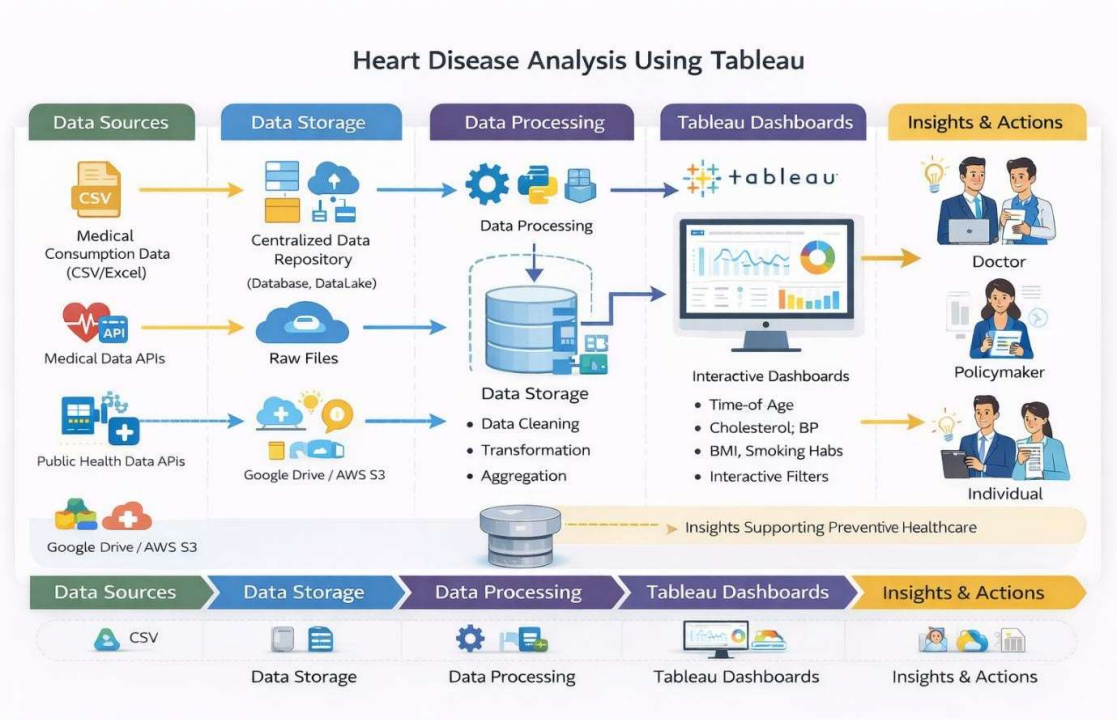
3.2 Proposed Solution

Proposed solution template:

S.No	Parameter	Description
1	Problem Statement (Problem to be solved)	Heart disease is a leading cause of death, but early detection is difficult due to complex medical data and lack of simple analytical tools for common users.
2	Idea / Solution Description	Develop a web-based Heart Disease Analysis system using data analytics and machine learning to predict risk and visualize health insights through interactive dashboards.
3	Novelty / Uniqueness	Uses storytelling dashboards and predictive analytics to present health data in an easy-to-understand visual format instead of traditional reports.
4	Social Impact / Customer Satisfaction	Helps users become aware of heart health, encourages preventive care, and improves decision-making, leading to better health outcomes.

5	Business Model (Revenue Model)	Can follow a subscription model for users and healthcare institutions, with premium features like personalized reports and recommendations.
6	Scalability of the Solution	The system is web-based and can easily scale to support more users, datasets, and future healthcare features.

3.3 Solution Architecture



4. Project Planning

4.1 Performance Testing

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the system by entering name, email, and password.	2	High	S Ashifa, V Akhil
Sprint-1	Registration	USN-2	As a user, I will receive confirmation email after registration.	1	High	R Deevena
Sprint-1	Registration	USN-3	As a user, I can register using Gmail.	2	Medium	V Akhil
Sprint-1	Login	USN-4	As a user, I can login using email and password.	1	High	S Ashifa
Sprint-1	Dashboard	USN-5	As a user, I can view Scenario 1 dashboard.	3	High	R Deevena
Sprint-2	Dashboard	USN-6	As a user, I can view Scenario 2 dashboard.	3	Medium	S Ashifa
Sprint-2	Dashboard	USN-7	As a user, I can view Scenario 3 dashboard.	3	Medium	V Akhil
Sprint-2	Analysis	USN-8	As a user, I can see heart disease risk insights.	3	High	R Deevena
Sprint-3	Reports	USN-9	As a user, I can download analysis report.	3	Medium	S Ashifa
Sprint-4	Story	USN-10	As a user, I can view story and narrative insights.	3	Low	V Akhil

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the system by entering name, email, and password.	2	High	S Ashifa

Sprint-1	Registration	USN-2	As a user, I will receive confirmation email after registration.	1	High	R Deevena
Sprint-1	Registration	USN-3	As a user, I can register using Gmail.	2	Medium	V Akhil
Sprint-1	Login	USN-4	As a user, I can login using email and password.	1	High	S Ashifa
Sprint-1	Dashboard	USN-5	As a user, I can view Scenario 1 dashboard.	3	High	R Deevena
Sprint-2	Dashboard	USN-6	As a user, I can view Scenario 2 dashboard.	3	Medium	S Ashifa
Sprint-2	Dashboard	USN-7	As a user, I can view Scenario 3 dashboard.	3	Medium	V Akhil
Sprint-2	Analysis	USN-8	As a user, I can see heart disease risk insights.	3	High	R Deevena
Sprint-3	Reports	USN-9	As a user, I can download analysis report.	3	Medium	S Ashifa
Sprint-4	Story	USN-10	As a user, I can view story and narrative insights.	3	Low	V Akhil

Velocity Calculation

Velocity is the number of story points completed by the team in one sprint. It helps measure the team's productivity and is used for future sprint planning.

From the project tracker:

- Sprint 1 completed = 20 story points
- Sprint 2 completed = 20 story points
- Sprint 3 completed = 20 story points
- Sprint 4 completed = 20 story points

Average Velocity (AV) Formula:

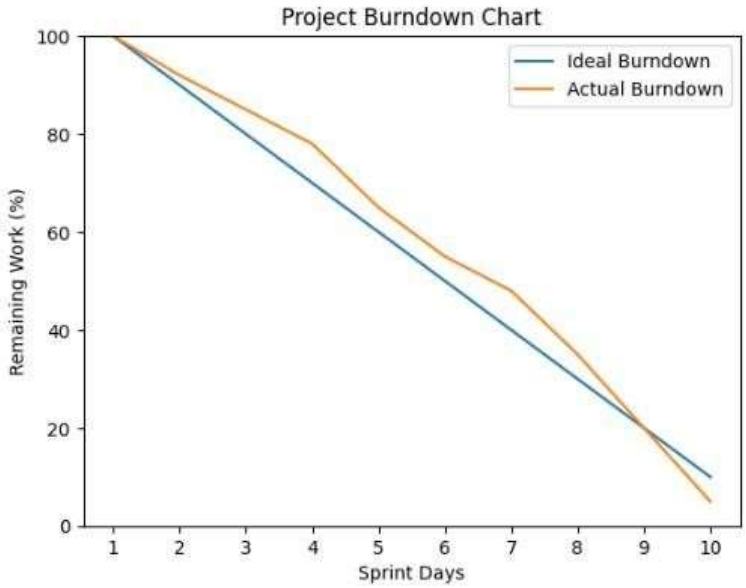
$$AV = (\text{Sprint1} + \text{sprint2} + \text{sprint3} + \text{sprint4})/\text{number of sprints}$$

$$AV = \frac{20 + 20 + 20 + 20}{4}$$

$$AV = 80$$

$$AV = \frac{\quad}{4} = 20$$

Burndown chart:



5. Project development Phase

5.1. Data Connectivity

The system connects to the heart disease dataset using structured data files such as CSV. The data is loaded into the analysis environment and connected to Tableau for visualization. This ensures real-time interaction with the dataset. The connectivity allows seamless data flow between backend processing and frontend dashboards.

Link of dataset:

https://drive.google.com/file/d/15ujsgAL8_vhkogwlegRYfHhoEcD4EQes/view?usp=sharing

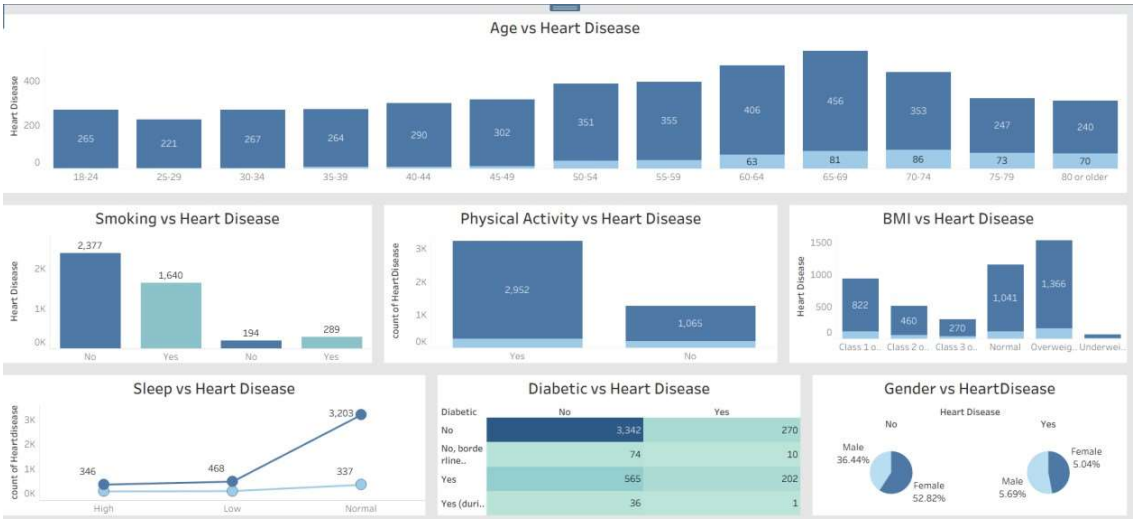
5.2. Data Preparation

The raw dataset is cleaned by handling missing values, removing duplicates, and correcting inconsistent data. Feature engineering is performed by creating meaningful categories such as age groups and BMI levels. Data transformation ensures the dataset is suitable for analysis and visualization. This step improves the accuracy and reliability of insights.

5.3. Dashboard

Scenario 1 – Demographic & Lifestyle Analysis

This scenario focuses on understanding how demographic factors and lifestyle habits impact heart disease. It analyzes attributes such as age, gender, smoking, alcohol consumption, and physical activity. The dashboard helps identify high-risk groups based on daily habits and population segments. This provides an overview of who is more vulnerable to heart disease.

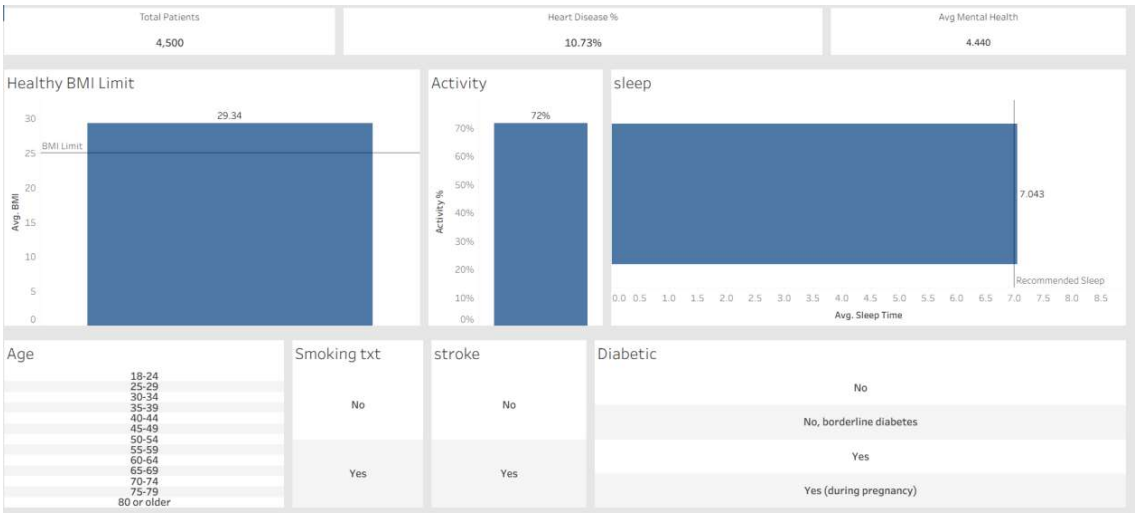


Scenario 2 – Health Indicators Analysis

This scenario analyzes key health parameters such as BMI, sleep duration, blood pressure, and cholesterol levels. It helps understand how these medical indicators influence heart disease risk. The dashboard highlights unhealthy patterns and their correlation with heart problems. This supports early detection through measurable health metrics.

Scenario 3 – Risk & Trend Analysis

This scenario focuses on overall heart disease risk and trends across different user groups. It compares multiple factors to identify common risk combinations. The dashboard shows patterns over time and helps predict future risk levels. This scenario supports decision-making and preventive actions.



5.4. Story

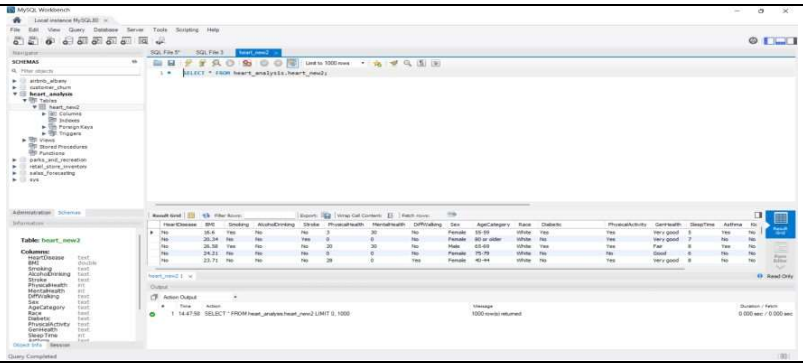
A story is created using Tableau to present insights in a narrative format. The story explains key findings step by step using visuals and text. This helps users interpret complex data easily. It acts as a guided explanation of heart disease trends and risks.

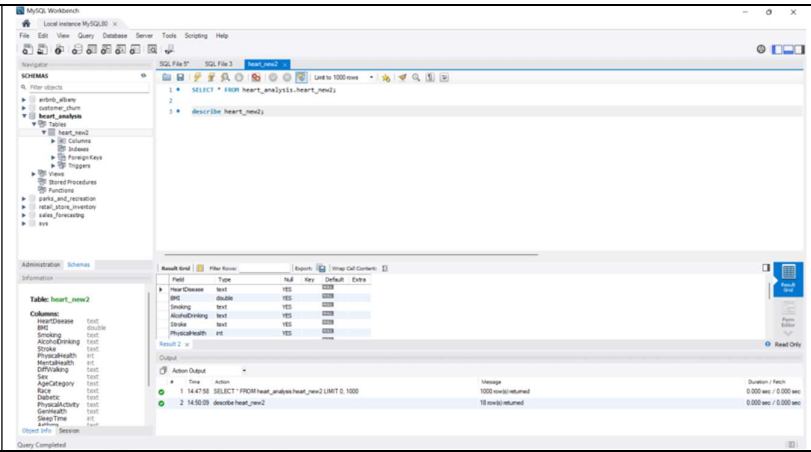
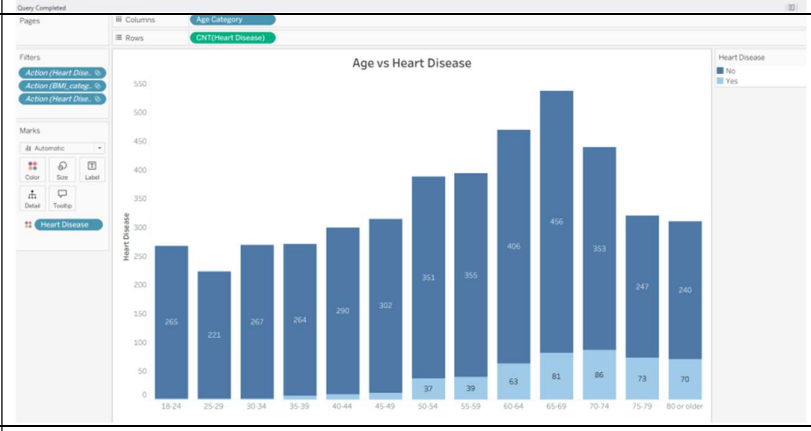
5.5. Creativity (Font and Style)

The project uses consistent fonts, colors, and layouts to improve readability and user experience. A clean and minimal design style is followed to avoid clutter. Visual elements such as icons and animations enhance engagement. This creative approach makes the dashboards more professional and user-friendly.

6. Performance Testing

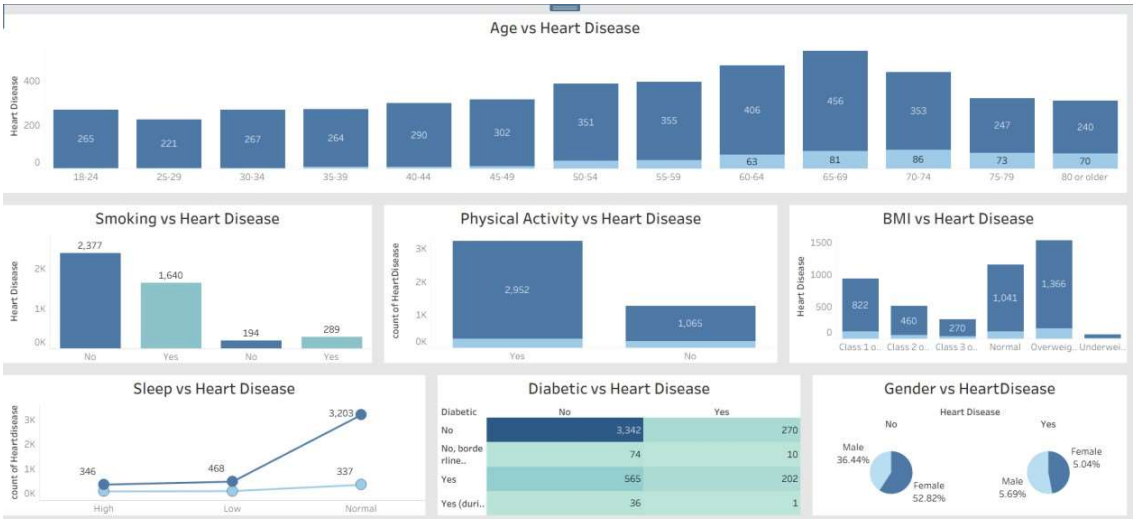
Model Performance Testing:

S.No.	Parameter	Screenshot/Value
1.	Data Rendered	

2.	Data Preprocessing	
3.	Utilization of filters	
4.	Calculation fields used	BMI_Category,Sleep_category,Physical Activity %,Number of records,heart disease %
5.	Dashboard Design	No.of Visualizations/Graphs-24
6.	Story Design	No.of Visualizations/Graphs-1

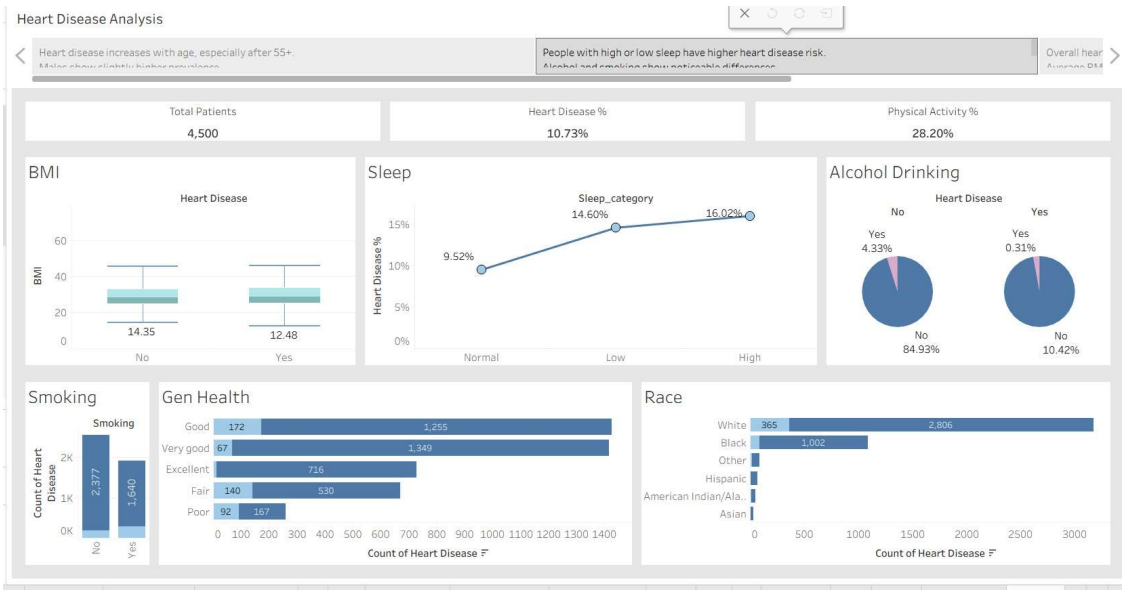
7. Output

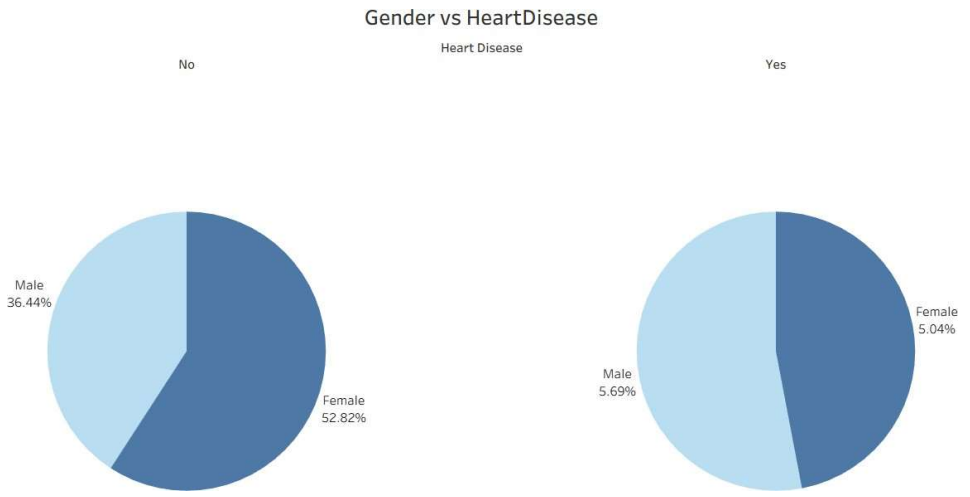
Output with screenshots

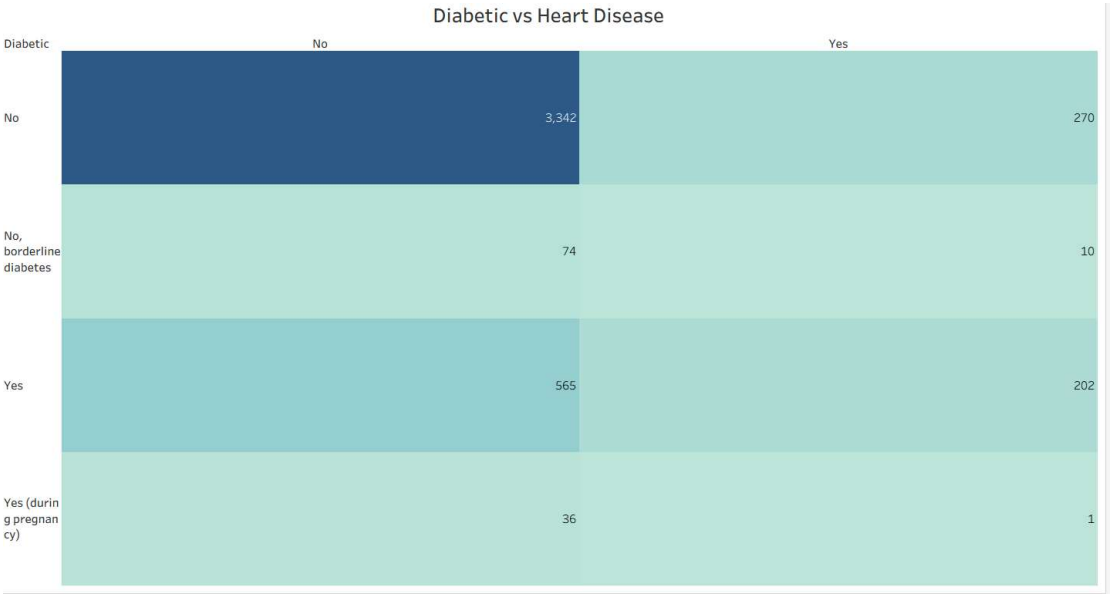


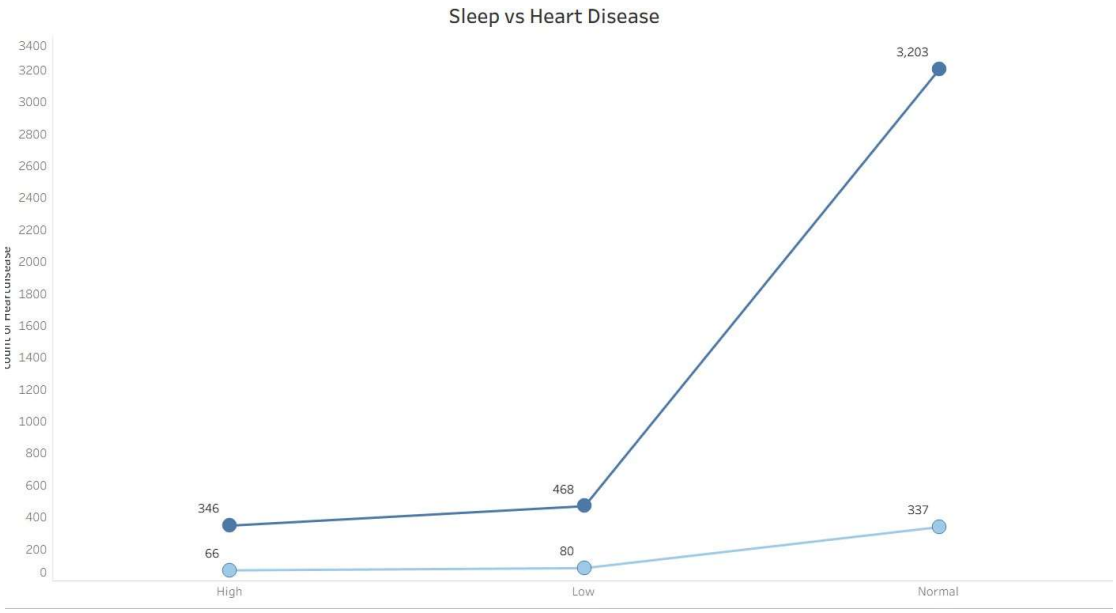

Total Patients
4,500
Heart Disease %
10.73%
Physical Activity %
28.20%
BMI
Heart Disease
75.82
63.62
14.35
12.48
No
Yes
Sleep
Sleep_category
14.60%
16.02%
9.52%
Heart Disease %
Normal
Low
High
Alcohol Drinking
Heart Disease
No
Yes
Yes 4.33%
No 84.93%
Yes 0.31%
No 10.42%
Smoking
Count of Heart Disease
2,377
1,640
Gen Health
Good 172 1,255
Very good 67 1,349
Excellent 716
Fair 140 530
Poor 92 167
Count of Heart Disease
Race
White 365 2,806
Black 86 1,002
Other
Hispanic
American Indian/Ala..
Asian
Count of Heart Disease

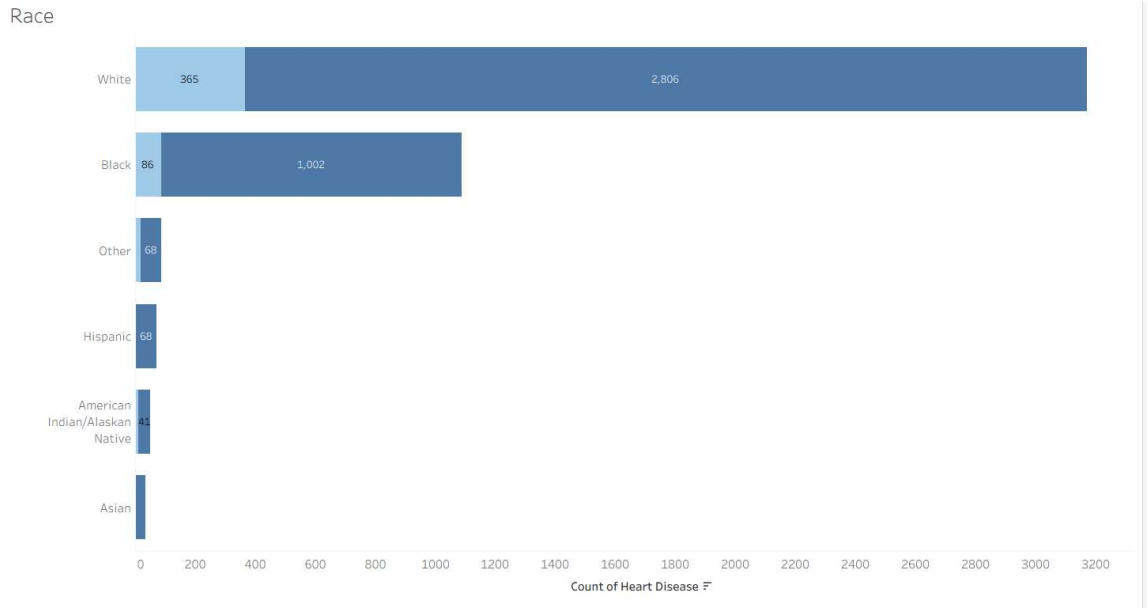

Total Patients
4,500
Heart Disease %
10.73%
Avg Mental Health
4.440
Healthy BMI Limit
BMI Limit
29.34
Avg. BMI
Activity
72%
Activity %
sleep
7.043
Recommended Sleep
Avg. Sleep Time
Age
18-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80 or older
Smoking txt
No
Yes
stroke
No
Yes
Diabetic
No
No, borderline diabetes
Yes
Yes (during pregnancy)


Heart Disease Analysis
Heart disease increases with age, especially after 55+.
People with high or low sleep have higher heart disease risk.
Total Patients
4,500
Heart Disease %
10.73%
Physical Activity %
28.20%
BMI
Heart Disease
75.82
63.62
14.35
12.48
Sleep
Sleep_category
14.60%
16.02%
9.52%
Alcohol Drinking
Heart Disease
Yes 4.33%
No 84.93%
Yes 0.31%
No 10.42%
Smoking
2,377
1,640
Gen Health
Good 172 1,255
Very good 67 1,349
Excellent 716
Fair 140 530
Poor 92 167
Race
White 365 2,806
Black 1,002
Other
Hispanic
American Indian/Ala..
Asian


Gender vs HeartDisease
Heart Disease
No
Yes
Male
36.44%
Female
52.82%
Female
5.04%
Male
5.69%


Diabetic vs Heart Disease
Diabetic
No
Yes
No
3,342
270
No, borderline diabetes
74
10
Yes
565
202
Yes (during pregnancy)
36
1





8. Advantages & Disadvantages

8.1 Advantages

- Provides clear visual insights instead of complex medical reports.
- Easy to use and understand for non-technical users.
- Supports early risk detection and preventive care.
- Interactive dashboards improve user engagement.

8.2 Disadvantages

- Accuracy depends on the quality of the dataset used.
- Does not replace professional medical diagnosis.
- Limited to the available data and predefined parameters.

9. Conclusion

The Heart Disease Analysis project successfully demonstrates how data analytics and visualization can improve healthcare awareness. The system transforms raw data into meaningful insights through dashboards and stories. It helps users understand heart disease risk factors in a simple and effective way. Overall, the project proves the importance of data-driven approaches in healthcare.

10. Future Scope

In the future, this system can be enhanced by integrating real-time data from hospitals and wearable devices. Advanced machine learning models can be implemented for more accurate predictions. A mobile application can be developed for wider accessibility. Personalized health recommendations and alert systems can also be added to make the solution more impactful.

Appendix

Source code(dataset link)

https://drive.google.com/file/d/15ujsgAL8_vhkogwlegRYfHhoEcD4EQes/view?usp=drive_link

Github & project Demon Link

Team Leader github link: Link of

github:

https://github.com/G-Praneetha/Heart-Disease-Analysis

Link of video:

https://drive.google.com/folderview?id=1J7xjte7EcFHvFNtyb2N4ELXrmyoTuj-l