High-Level Design (HLD)

Health Analytics on Heart Disease Data

Business Intelligence

Last Revised Date: 25/08/2021

Document Version Control

Version	Date	Author	Comments
1.	21-Aug-21	Manish Tiwari Samruddhi Deshmukh Subhabrata Nath	Adding Problem statement
2.	25-Aug-21	Samruddhi Deshmukh	Adding Flowchart Updating Index

Reviews:

Version	Date	Reviewer	Comments

Approval Status:

Version	Review Date	Reviewed By	Approved By	Comments

Index

Abstract

Heart disease is the most major health issue that is suffering by many people all over the globe, some of the causes of heart diseases due to hypertension, diabetes, overweight, and an unhealthy lifestyle. This project of Healthcare Analysis on Heart Disease Data is aimed to explore the Heart Disease dataset .The objective is analyze the various features and their relationship with each other and find out their contribution towards getting a heart disease.

Various features such as **Age, Sex, Chest pain type, Blood pressure, Cholestrol,Fasting Blood sugar, Rest ECG, Thalach, Exercise enduced Angina,Major vessels,oldpeak,slope,thal** are present in the dataset.The goal of the project is to find all types of relationships between the features and come out with significant contributors to a heart disease.

1. Introduction

1.1 Scope of the HLD:

The HLD system aims to present the functionality of the overall system. It aims to describe the process workflow along with description of its various components.It describes about the –

- Problem statement under consideration.
- Solution strategy towards the problem.
- Architecture of the process.
- Implementation of the solution.
- Tools used.
- Analysis of the KPI's and coming out with suitable conclusions about the analysis.

The final analysis (dashboard) will be shared through Tableau public.

2. General Description

2.1 Product Perspective and Problem Statement

Healthcare domain is one of the leading domain. Using Data analysis tools and techniques to build a solution for a heathcare domain problem can be counted as a significant contribution to this domain.

The main objective of the project includes :

- Analyzing raw dataset.
- Perform data cleaning and pre processing operations on the raw data.
- Building visualizations on the cleaned features using a BI tool.
- Final conclusions.
- Deployment of the project on Tableau public.

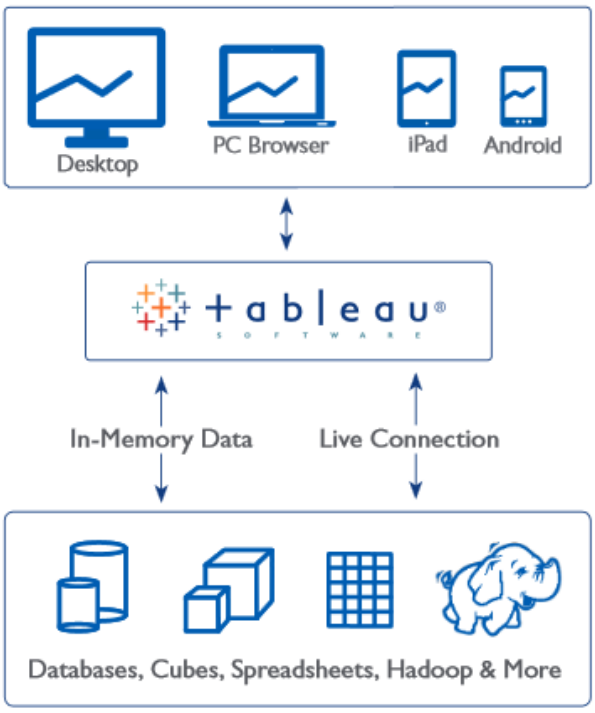
2.2 Tools Used

- Data Pre-processing - Python library pandas ,Excel.
- Data Visualization – Tableau
- Dashboarding/ Deployment – Tableau public.

2.2.1 Tableau Desktop and Tableau Public –

The **Tableau Desktop** is data visualization software that lets you see and understand data in minutes. The professional version of this can transform, process and store huge volumes of data which is responsible for all the data-driven decision making of an organization. Tableau Desktop can connect variety of datasources whether it's big data, a SQL database, a spreadsheet, or cloud apps like Google Analytics and Salesforce. Tableau Desktop has no limit to how many rows of data it can store, process or share. For Tableau Desktop, reports published can be stored on your Local Drive as well as on the Tableau Server .

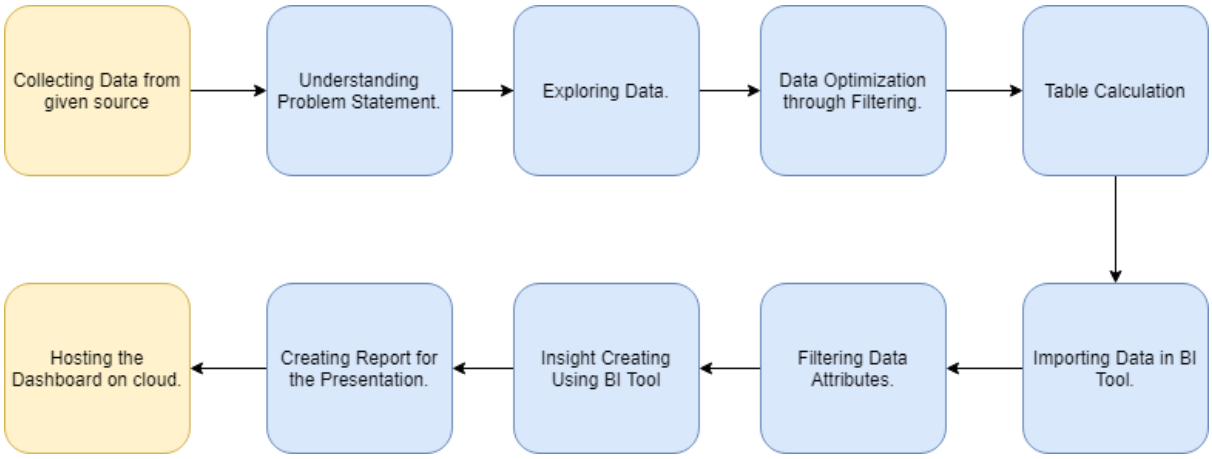
The **Tableau Public** is essentially a free version of Tableau visualization software. It allows you to use most of the software functions. You can create visualizations and connect to CSV, Text and Excel documents. Tableau Public is a free platform to publicly share and explore data visualizations online. Anyone can create visualizations using either Tableau Desktop Professional Edition or the free Public Edition.



3. Design Details:

3.1 Application Architecture.

- The following flow we will cover for the entire application.



3.2 Optimization

3.2.1Data Strategy and Performance.

- Removing duplicate records from the dataset.
- Handling null values.
- Performing feature encoding on the dataset.
- Building interactive filters in tableau to view data as required.
- Creating groups, hierarchy and calculated fields for easy analysis.

3.3 Architecture Description.

- **Understanding Problem Statement:**
  Find /Analyse relations between features and come out with the significant contributors to a Heart Disease. End goal includes building Dashboard for user interaction and a Detailed project report for publishing our findings.
- **Collecting Data:**
  Dataset is already available on the project dashboard. We will use the given dataset for building the project dashboard.
- **Exploring Data:**
  Performing pre-processing with raw dataset and using the cleaned dataset for building graphs. We will use the plotting library and python data manipulation library like pandas for these tasks.
- **Data Optimization through Filtering:**
  We will optimize the given dataset before starting the analysis. We have to drop duplicate columns and reduce the dataset dimensions as needed. We will also be encoding our data into numerical to categorical or vice versa as needed.
- **Table Calculation:**
  We will calculate some new fields that will represent more granular results than raw data. We will find the relationship between attributes and perform the calculation as required.
- **Importing Data In BI Tool:**
  We will use Tableau for creating the Visualization and Dashboard. Tableau is a drag and drop application for creating Dashboard and it is very much efficient in handling a large amount of data. We can plot and show different types of relationships between attributes.
- **Filtering Data Attributes:**
  We will use different types of filters on our data set for plotting purposes and getting the insight information. Using filters can speed up information finding and also it can produce more variety and granular reports.
- **Insight Creation Using BI Tools:**
  This is the most crucial part of the complete process. Here we will find meaningful information by going deep into our dataset. We will plot different types of plots for every

attribute on the Heart Diseases dataset to show the relationship between a patient's health condition depends on having heart disease or not. This is our key finding on the dataset.

- **Creating Report for the Presentation:**
  After we find all the possible information that can be found from the dataset we will create a detailed report for presenting and publishing our works. We will include all the visual plots and will write key findings from the plots.
- **Hosting the Dashboard on Cloud:**
  After creating a Detailed Project report and creating the dashboard we will host our dashboard on the cloud platform for global use. For this, we will use the Tableau Public Platform. Here we can host our dashboard and it will create a sharable link for access globally.

4. Dashboard

For showing the report user can use our Dashboard. The dashboard will be implemented to display and indicate certain KPIs and relevant indicators.

The dashboard will show the relationship between the Patient's health condition like blood sugar, blood pressure with having heart disease or not. The dashboard will also be capable of showing filter reports as user required and will be able to show the relationship between different attributes.

4.1 Key Performance Indicators (KPIs)

KPIs will put a summary of the findings of the Heart Diseases Dataset and the relationship between different metrics. Inside the dataset, we have information like a patient's age, gender, and his or her different medical condition like blood pressure, blood sugar e.t.c, and all the health conditions are targeting at that particular patients have Heart Diseases or not.

The KPIs are listed below…

- Impact of Age on having heart diseases.
- How gender is varying in heart disease
- How blood sugar and blood pressure are putting an impact.
- Influence of heart diseases on heart rate.

5. Deployment

Tableau Public Server will be used for deployment. It's a free cloud service where we can host our dashboard for making it publically available. There is some advantage of choosing this server it is required less maintenance, can be host directly form local machine, and the deployment process is also very easy.Deployment can be done from Tableau Workspace only.

Benefits of Tableau Public:

- Easy to use and Free
- Fast and more reliable
- No software installation is required for the deployment
- Easy to share of globally accessible link
-

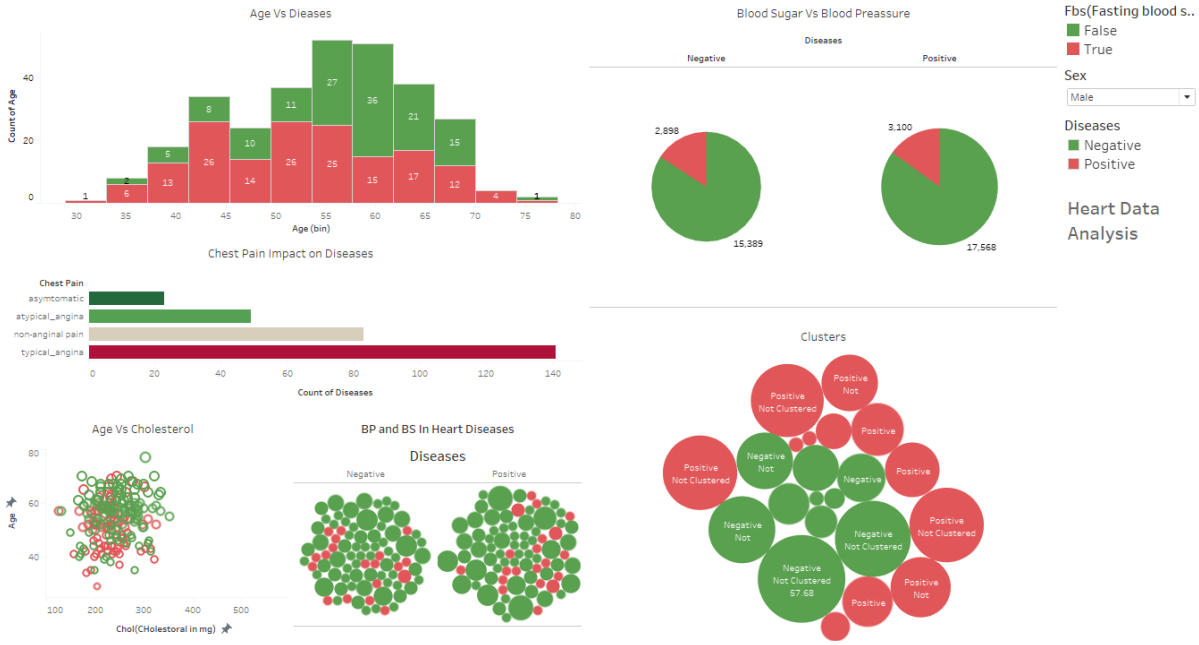
5.1 Steps of Deployment

We can deploy our dashboard directly from our workspace, we must have registered on Tableau Public before that. And after saving the file it will automatically deploy on Tableau Public Server.

Steps 1 – Login

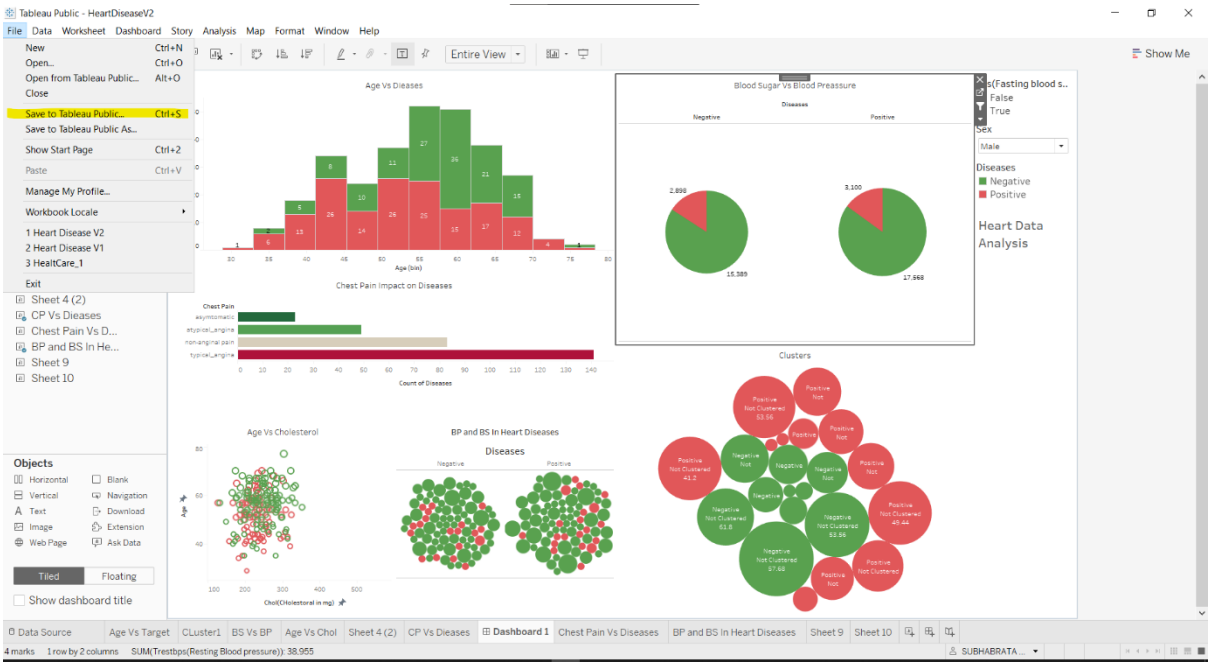
The user has to enter his registered email address and password for the login.

Step 2 – Building the Dashboard.



Users need to Build the Dashboard before Deploying on tableau public. And have to save the dashboard on the tableau workspace. We have to select and have to present on the desire dashboard or worksheet that we want to deploy.

Step 3 – Save the Dashboard from Tableau Workspace.



Step 4 – Save the Dashboard

After click on the save the Tableau dashboard will save and deploy on the tableau public server.