High Level Design (HLD)

Heart Disease Diagnostic Analysis

Revision Number: 1.0

Last date of revision: 13/04/2024

Prakhyat Srivastava

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| **13th April 2024** | 1.0 | First Version of Complete HLD | Prakhyat Srivastava |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contents**

[Document Version Control 2](#_bookmark0)

[Abstract 3](#_bookmark1)

1. [Introduction 5](#_bookmark2)
   1. [Why this High-Level Design Document? 5](#_bookmark3)
   2. [Scope 5](#_bookmark4)
2. [General Description 6](#_bookmark5)
   1. [Product Perspective & Problem Statement 6](#_bookmark6)
   2. [Tools used 6](#_bookmark8)
3. [Design Details 7](#_bookmark9)
   1. [Functional Architecture 7](#_bookmark10)
   2. [Optimization 8](#_bookmark11)
4. [KPIs 9](#_bookmark12)
   1. [KPIs (Key Performance Indicators) 9](#_bookmark13)
5. [Deployment 9](#_bookmark14)

# Abstract

In the panorama of public health, heart disorders present an escalating challenge, particularly within the Indian demographic, where they have ascended as the preeminent cause of mortality. The study encapsulated within this High-Level Document (HLD) pivots on an exhaustive diagnostic analysis of heart disease, employing a robust dataset encompassing both categorical and numerical variables integral to heart disease diagnosis. The meticulous cleaning and preprocessing of this data, inclusive of handling null values and outliers, followed by an incisive Exploratory Data Analysis (EDA), lay the groundwork for a data-driven prognostic model.

This model is poised to significantly bolster the prediction and, consequently, the prevention of heart ailments, thereby heralding a potential decline in the heart disease-related mortality rates that have seen a disquieting upsurge of 34% from 1990 to 2016 in India. Through this endeavor, the document endeavors to underpin the framework of a system adept at pre-empting the onset of heart conditions, catalyzing the advancement of cardiovascular health and therapeutics in the Indian healthcare sector.

# Introduction

## Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

* + - Present all of the design aspects and define them in detail
    - Describe the user interface being implemented
    - Describe the hardware and software interfaces
    - Describe the performance requirements
    - Include design features and the architecture of the project
    - List and describe the non-functional attributes like:
      * Security
      * Reliability
      * Maintainability
      * Portability
      * Reusability
      * Application compatibility
      * Resource utilization
      * Serviceability

## Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

# General Description

## Product Perspective & Problem Statement

This project is orchestrated to navigate through the intricacies of cardiovascular health, focusing on the pivotal task of predicting the likelihood of heart disease in individuals. It harnesses a data set meticulously curated with 303 patient profiles, each annotated with a suite of parameters pertinent to heart health. The crux of the problem statement lies in leveraging these multifaceted data points—ranging from physiological metrics to lifestyle indicators—to ascertain the propensity for heart disease in each case. The project is not just an academic exercise but a stepping stone towards a broader vision of preemptive healthcare, aiming to distill actionable insights from the data that could potentially steer individuals away from the path leading to heart ailments.

## Tools used

Business Intelligence tools and libraries works such as NumPy, Pandas, Seaborn, Tableau, , MS-Excel, Jupyter Notebook and Python Programming Language are used to build the whole framework.





# Design Details

## Functional Architecture

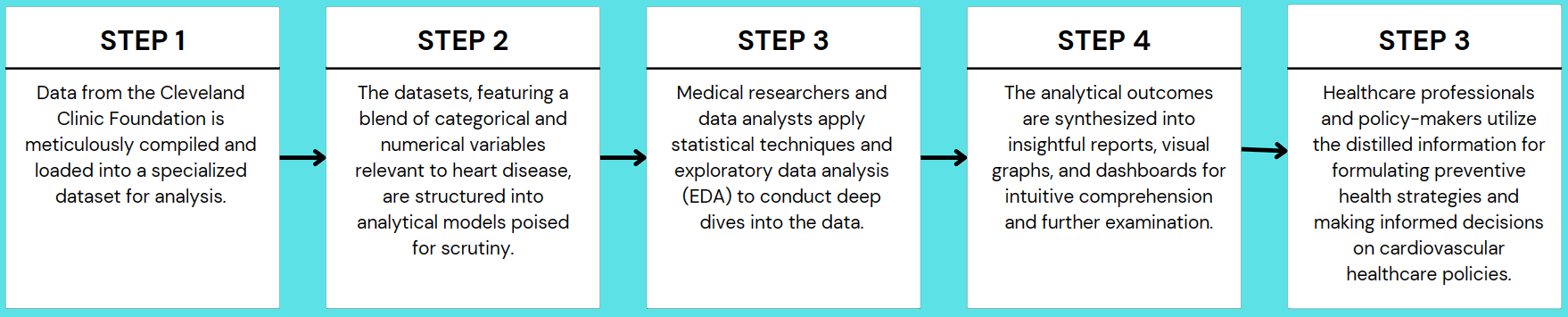
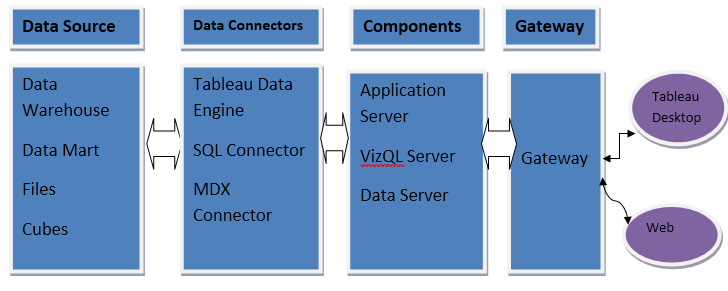
****

Figure 1: Functional Architecture of Business Intelligence

**How Tableau Works**



## Optimization

### Your data strategy drives performance

* + - Minimize the number of fields
    - Minimize the number of records
    - Optimize extracts to speed up future queries by materializing calculations, removing columns and the use of accelerated views

### Reduce the marks (data points) in your view

* + - Practice guided analytics. There’s no need to fit everything you plan to show in a single view. Compile related views and connect them with action filters to travel from overview to highly-granular views at the speed of thought.
    - Remove unneeded dimensions from the detail shelf.
    - Explore. Try displaying your data in different types of views.

### Limit your filters by number and type

* + - Reduce the number of filters in use. Excessive filters on a view will create a more complex query, which takes longer to return results. Double-check your filters and remove any that aren’t necessary.
    - Use an include filter. Exclude filters load the entire domain of a dimension, while include filters do not. An include filter runs much faster than an exclude filter, especially for dimensions with many members.
    - [Use a continuous date filter](http://onlinehelp.tableau.com/current/pro/online/mac/en-us/help.htm#filtering_add_dragfields_dates.html). Continuous date filters (relative and range-of-date filters) can take advantage of the indexing properties in your database and are faster than discrete date filters.
    - [Use Boolean or numeric filters](http://www.tableau.com/learn/tutorials/on-demand/logical-calculations). Computers process integers and Booleans (t/f) much faster than strings.
    - Use [parameters](http://onlinehelp.tableau.com/current/pro/online/en-us/help.htm#parameters.html) and [action filters](http://onlinehelp.tableau.com/current/pro/online/en-us/help.htm#actions.html). These reduce the query load (and work across data sources).

### Optimize and materialize your calculations

* + - Perform calculations in the database
    - Reduce the number of nested calculations.
    - Reduce the granularity of LOD or table calculations in the view. The more granular the calculation, the longer it takes.
      * LODs - Look at the number of unique dimension members in the calculation.
      * Table Calculations - the more marks in the view, the longer it will take to calculate.
    - [Where possible, use MIN or MAX instead of AVG](http://onlinehelp.tableau.com/current/pro/online/windows/en-us/help.htm#calculations_aggregation.html). AVG requires more processing than MIN or MAX. Often rows will be duplicated and display the same result with MIN, MAX, or AVG.
    - [Make groups with calculations](http://kb.tableau.com/articles/knowledgebase/creating-groups-using-calculated-fields). Like include filters, calculated groups load only named members of the domain, whereas Tableau’s group function loads the entire domain.
    - [Use Booleans or numeric calculations instead of string calculations](http://onlinehelp.tableau.com/current/pro/online/mac/en-us/help.htm#functions_functions_string.html). Computers can process integers and Booleans (t/f) much faster than strings. Boolean>Int>Float>Date>DateTime>String

# KPIs

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the disease.



As and when, the system starts to capture the historical/periodic data for a user, the dashboards will be included to display charts over time with progress on various indicators or factors

## KPIs (Key Performance Indicators)

Key indicators displaying a summary of the Housing Price and its relationship with different metrics

1. Proportion of individuals diagnosed with heart disease
2. Distribution of heart disease across various age groups and genders
3. Comparison of heart disease prevalence between male and female patients
4. Types of chest pain experienced by heart disease patients
5. Analysis of blood pressure, cholesterol levels, and max heart rate by age in heart disease patients
6. ST depression levels in relation to age and presence of heart disease

# Deployment

Prioritizing data and analytics couldn’t come at a better time. Your company, no matter what size, is already collecting data and most likely analyzing just a portion of it to solve business problems, gain competitive advantages, and drive enterprise transformation. With the explosive growth of enterprise data, database technologies, and the high demand for analytical skills, today’s most effective IT organizations have shifted their focus to enabling self-service by deploying and operating Tableau at scale, as well as organizing, orchestrating, and unifying disparate sources of data for business users and experts alike to author and consume content.

Tableau prioritizes choice in flexibility to fit, rather than dictate, your enterprise architecture. Tableau Server and Tableau Online leverage your existing technology investments and integrate into your IT infrastructure to provide a self-service, modern analytics platform for your users. With on-premises, cloud, and hosted options, there is a version of Tableau to match your requirements.

