# High Level Design (HLD)

# Heart Disease Prediction

Revision Number: 1.0

Last date of revision: 13/08/2021

AMIT KUMAR

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | Version | Description | Author |
| 13th AUGUST 2021 | 1 .0 | First Version of Complete HLD | Amit Kumar |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Contents:

Document Version Control.........................................................................................................2

Abstract......................................................................................................................................4

1 Introduction .............................................................................................................................5

1.1 Why this High-Level Design Document?.........................................................................5

1.2 Scope ...............................................................................................................................6

2 General Description ................................................................................................................6

2.1 Product Perspective & Problem Statement .....................................................................6

Heart disease is one of the biggest causes of morbidity and mortality among the population of the world. The objective of the project is to perform data visualization techniques to understand the insight of the data. This project aims apply various Business Intelligence tools such as Tableau or Power BI to get a visual understanding of the data........................................6

2.2 Tools used........................................................................................................................7

3 Design Details.........................................................................................................................8

3.1 Functional Architecture ....................................................................................................9

3.2 Optimization .....................................................................................................................10

4 KPIs.........................................................................................................................................11

4.1 KPIs (Key Performance Indicators) .................................................................................12

5 Deployment.............................................................................................................................13

**ABSTRACT:**

I am creating a Data Analysis Project on Heart Disease Prediction. The project uses raw data in form of a .csv file and transforms into Data Analysis. This project is an attempt of data analyzing Heart Disease Prediction with the help of data analytics in Microsoft Power Bi.

Heart disease is one of the biggest causes of morbidity and mortality among the population of the world. Prediction of cardiovascular disease is regarded as one of the most important subjects in the section of clinical data analysis. The amount of data in the healthcare industry is huge. Data mining turns the large collection of raw healthcare data into information that can help to make informed decisions and predictions.

Coronary Heart Disease (CHD) is the most common type of heart disease, killing over 370,000 people annually. Every year about 735,000 Americans has a heart attack. Of these, 525,000 are a first heart attack and 210,000 happen in people who have already had a heart attack. This makes heart disease a major concern to be dealt with. But it is difficult to identify heart disease because of several risk factors such as diabetes, high blood pressure, high cholesterol, abnormal pulse rate, and many other factors. Because of these factors, scientists have turned towards modern approaches like Data Mining and Machine Learning for predicting the disease.

# 1. Introduction

# 1.1 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:

o Security

o Reliability

o Maintainability

o Portability

o Reusability

o Application compatibility

o Resource utilization

o Serviceability

# 1.2 Scope6

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.

# 2. General Description

# 2.1 Product Perspective & Problem Statement

Heart disease is one of the biggest causes of morbidity and mortality among the population of the world. Prediction of cardiovascular disease is regarded as one of the most important subjects in the section of clinical data analysis. The amount of data in the healthcare industry is huge. Data mining turns the large collection of raw healthcare data into information that can help to make informed decisions and predictions.

The objective of the project is to perform data visualization techniques to understand the

insight of the data. This project aims apply various Business Intelligence tools such as Tableau

or Power BI to get a visual understanding of the data.

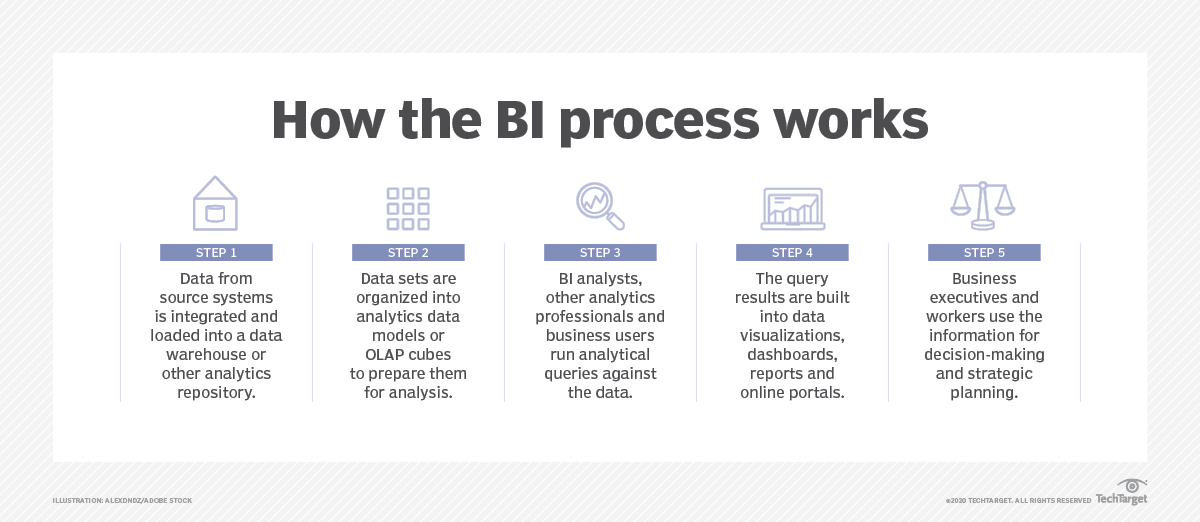
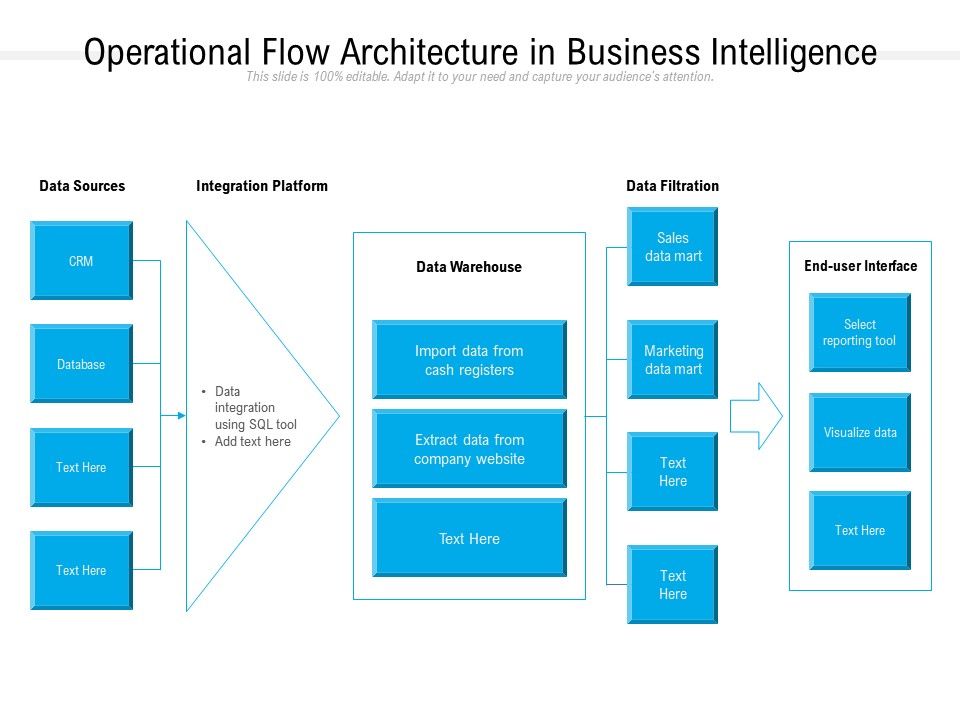
# 2.2 Tools used

Business Intelligence tool used is Microsoft Power BI to build the whole framework.



# 3 Design Details

# 3.1 Functional Architecture



# 3.2 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. 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. Computers process integers and Booleans (t/f) much

faster than strings.

• Use parameters and action filters. 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.

o LODs - Look at the number of unique dimension members in the calculation.

o Table Calculations - the more marks in the view, the longer it will take to

calculate.

• Where possible, use MIN or MAX instead of AVG. 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. 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. Computers can

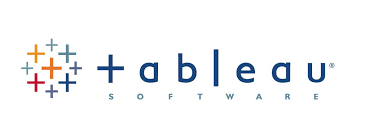
process integers and Booleans (t/f) much faster than strings.

Boolean>Int>Float>Date>DateTime>String

# 4 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.

# 4.1 KPIs (Key Performance Indicators)

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

metrics

1. Occurrence of cholesterol according to age group and gender.

2. Occurrence of Chest pain according to age group and gender.

3. Impact of blood pressure for both gender.

4. Impact of blood sugar according to gender and age.

5. Observing blood disorder called thallium ,according to gender.

6. Resting electrocardiographic results according to age and gender.

# 5 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.

Power BI is a Microsoft business analytics service. It provides interactive visualizations and business intelligence capabilities with an interface that Microsoft says is simple enough for end users to create reports and dashboards. It is part of the Microsoft Power Platform.

Power BI provides [cloud](https://en.wikipedia.org/wiki/Cloud_computing)-based BI (business intelligence) services, known as "Power BI Services", along with a desktop based interface, called "Power BI Desktop". It offers [data warehouse](https://en.wikipedia.org/wiki/Data_warehouse) capabilities including [data preparation](https://en.wikipedia.org/wiki/Data_preparation), [data discovery](https://en.wikipedia.org/wiki/Data_discovery) and interactive dashboards. In March 2016, Microsoft released an additional service called Power BI Embedded on its [Azure](https://en.wikipedia.org/wiki/Microsoft_Azure) cloud platform.[[3]](https://en.wikipedia.org/wiki/Microsoft_Power_BI#cite_note-3) One main differentiator of the product is the ability to load custom visualizations.

types of deployment in power bi:-

* **On-Premises:** Refers to data, applications and infrastructure entirely owned by client at client data center and client has complete control over it.
* **Cloud**: Refers to data, infrastructure and/or services residing in a public cloud environment and completely managed /controlled by third party. Microsoft Azure and web based Power BI service are examples of the cloud offerings.
* **Hybrid**: This denotes to the implementation which  spans both on premises and cloud sources which can be services, infrastructure and data sources

**Power BI Service**

* Data is either from the on premises corporate applications or it might be born in cloud. It can even mix of these two
* Data preparation and report creation occurs in Power BI Desktop or excel
* Completed Power BI reports are then published to Power BI service
* Report consumption, sharing, security, collaboration, data refresh happens in Power BI service
* Dashboards are created in Power BI service and reports can also be edited or created in Power BI service

**Custom Application Integration**

* Data is either from the on premises corporate applications or it might be born in cloud. It can even mix of these two
* Data preparation and report creation occurs in Power BI Desktop or excel
* Completed Power BI reports are then published to Power BI service
* With Power BI API , these reports can be published in custom web application or mobile app  within iFrame
* If user interacts with this report, he/she will be redirected to Power BI service
* Application can be on premise or cloud application

**Public Website**

* Data is either from the on premises corporate applications or it might be born in cloud. It can even mix of these two
* Data preparation and report creation occurs in Power BI Desktop
* Completed Power BI reports are then published to Power BI service
* An embed code is generated by Power BI service for selected report and this code can be embedded in web page of the website within iFrame
* Here no security is maintained as its public website, hence suitable for the data which can be made publicly available.