

# High Level Design (HLD)

## Crop Production Analysis In India



**BY- Dhanshri Manusmare**

- **Document Control**

Date	Version	Description	Author
31-12-22	1.0	Introduction, Problem Statement	Dhanshri
28-01-23	1.1	Dataset Information, Architecture Description	Dhanshri
25-04-23	1.2	Final Revision	Dhanshri

- **Contents**

Topics	Page Number
Document Control	2
Abstract	3
1. Introduction 1.1. Why this high level design document? 1.2. Scope	4
2. General Description 2.1. Product perspective and Problem statement 2.2. Tools Used	6
3. Design Details 3.1. Functional Architecture 3.2. Optimization	7
4. KPI 4.1. Key Performance Indicators	10
5. Deployment	11

# Abstract

The Agriculture business domain, as a vital part of the overall supply chain, is expected to highly evolve in the upcoming years via the developments, which are taking place on the side of the Future Internet. This paper presents a novel business-to-business collaboration platform from the agriculture sector perspective, which aims to facilitate the collaboration of numerous stakeholders belonging to associated business domains, in an effective and flexible manner. In the world of rising new technology and innovation, the Agriculture industry is advancing with the role of Data Science and Analytics. Data analysis can help them to understand their business in a quite different manner and helps to improve the quality of the service by identifying the weak areas of the business. This study demonstrates how different analysis help to make better business decisions and help analyze customer trends and satisfaction, which can lead to new and better products and services. Different analyses were performed such as Exploratory Data Analysis and Descriptive Analysis on a variety of use cases to get the key insights from this data based on which business decisions will be taken. This dataset provides a huge amount of information on which state is the best crop in India by production and find important insights highlighting key indicators and metrics that influence the customer choice.

## 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
- 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 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.

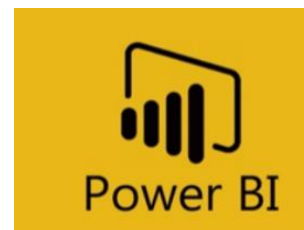
## 2. General Description

### 2.1 Product Perspective & Problem Statement

Health is real wealth in the pandemic time we all realized the brute effects of covid-19 on all irrespective of any status. You are required to analyse this health and medical data for better future preparation. A dataset is formed by taking into consideration some of the information of 303 individuals.

### 2.2 Tools

Business Tools such and libraries such as Power Bi , Numpy , Seaborn, Matplotlib, Plotly, Pandas and Python Programming language are used to build the whole framework



Pandas

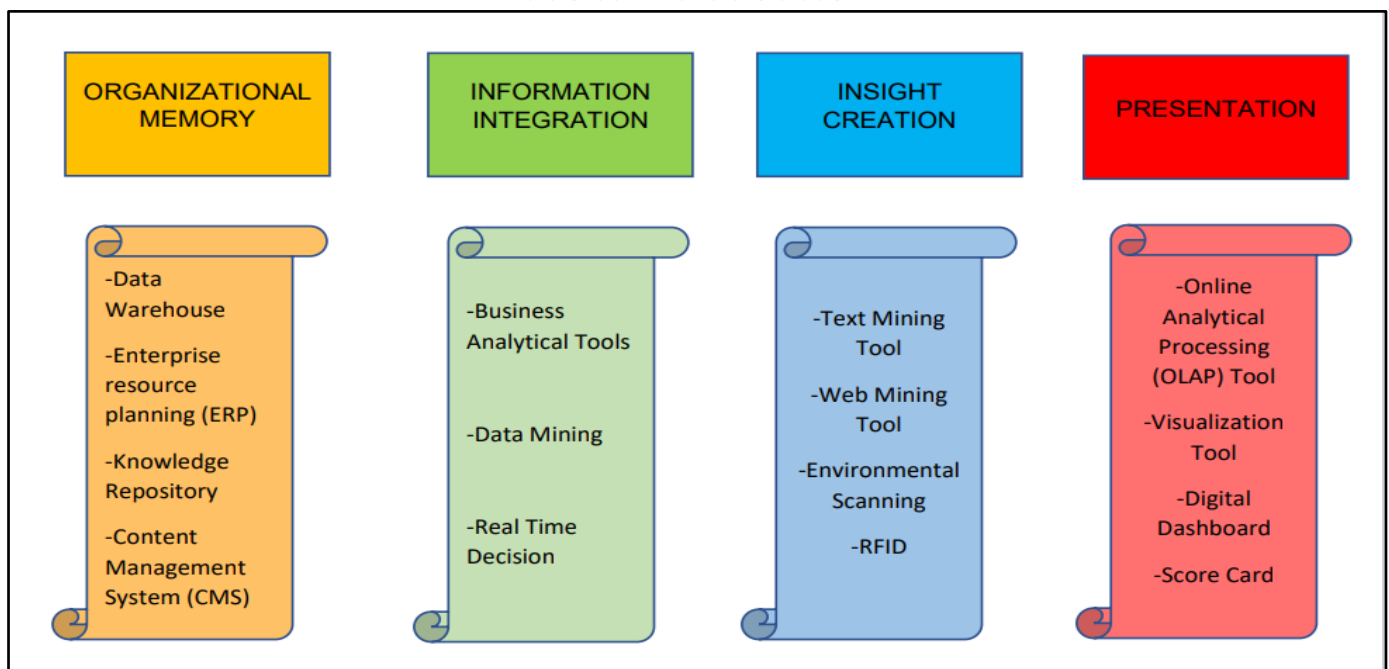


## 3. Design Details

### 3.1 Functional Architecture



### • How BI Works



## 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. KPI

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 Crop Production area and production information based on various parameters –

- Name of state in India.
- Name of district corresponding to state.
- Season and crops in india.
- Area - It refers to the area starting from 0.04 to 8580100
- Production - It simply refers to the Production by state ,season in India.

## 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 analysing 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 Power BI 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 prioritizes choice in flexibility to fit, rather than dictate, your enterprise architecture. Power BI Desktop and Power BI Service leverage your existing technology investments and integrate them 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 Power BI to match your requirements.

