

# High Level Design (HLD)

## Amazon Sales Data Analysis.

Prepared By,  
Mahesh Toradmal.

## Contents.

Sr.	Topic	Page No.
1	Introduction	<u>3</u>
2	General Description	<u>4</u>
3	Design Details	<u>5</u>
4	KPIs	<u>7</u>
5	Deployment	<u>7</u>

## Abstract.

Sales management has gained importance to meet increasing competition and the need for improved methods of distribution to reduce cost and to increase profits. Sales management today is the most important function in a commercial and business enterprise. This design document shows the tools and strategies required to get insights form the given dataset.

# 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 analysis. This document is also intended to help detect contradictions prior to analysis, 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 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

Sales management has gained importance to meet increasing competition and the need for improved methods of distribution to reduce cost and to increase profits. Sales management today is the most important function in a commercial and business enterprise. In this project Amazon Sales Dataset is analysed to draw out some key insights from the data set. Below mentioned is the overview of the path.

ETL: Extract-Transform-Load some Amazon dataset and finding the Sales-trends, month wise, year wise, yearly month wise.

We need to find key metrics and factors and show the meaningful relationships between attributes.

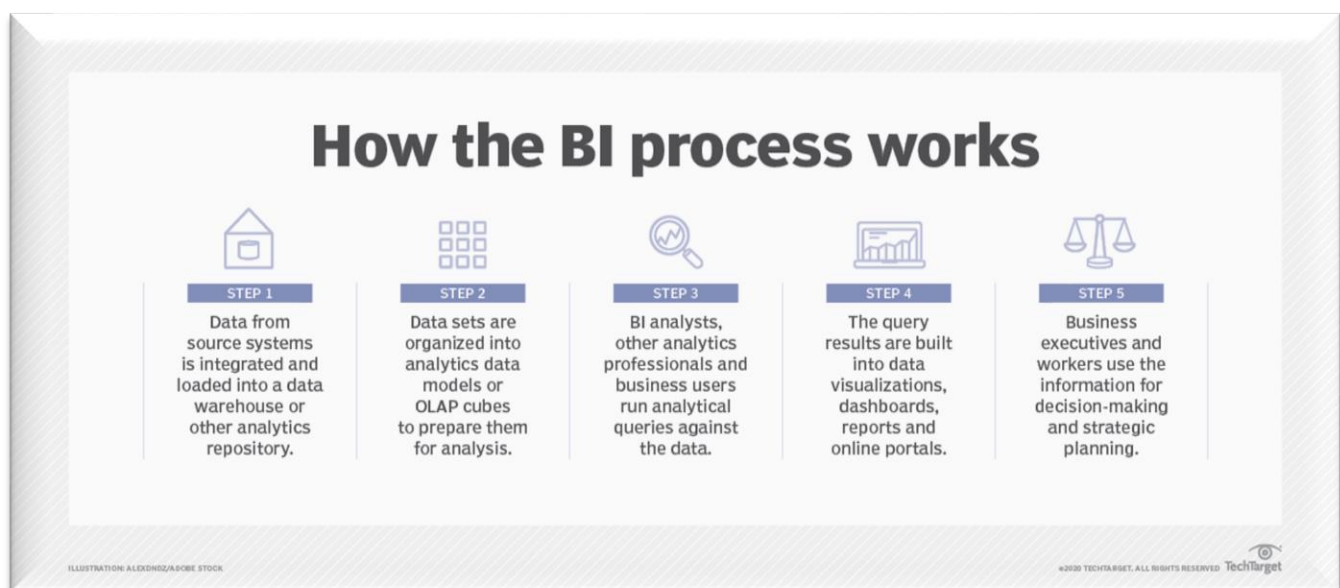
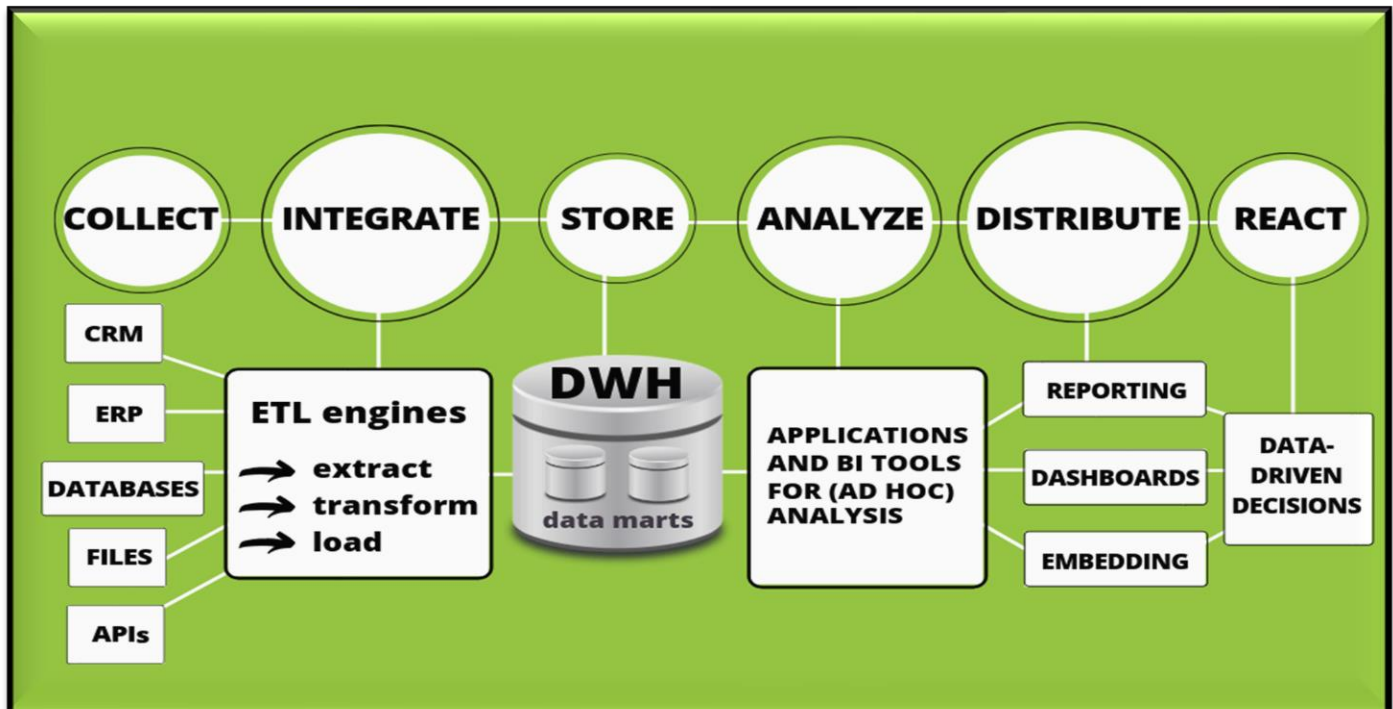
### 2.2 Tools Used

Business Intelligence tools and libraries works such as NumPy, Pandas, Excel, Power BI are used 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.

#### Optimize and materialize 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. [AUTHOR NAME] 9 9 High Level Design (HLD)
- 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>Date Time>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 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 Key Performance Indicators:

- Month wise Sales
- Year wise Sales
- Yearly Month wise Sales.
- Region wise Sales.
- Sales Quantity according to Customer Types. Etc.

## 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 Tableau and 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.