High Level Design (HLD) Amazon Sales Analysis

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| **11th oct 2023** | 1.0 | First Version of Complete HLD | Yash Shinde |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**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)

[Housing prices are an important reflection of the economy, and housing price ranges are of](#_bookmark7) [great interest for both buyers and sellers. In this project, house prices will be predicted given](#_bookmark7) [explanatory variables that cover many aspects of residential houses. 6](#_bookmark7)

* 1. [Tools used 6](#_bookmark8)

1. [Design Details 7](#_bookmark9)
   1. [Functional Architecture 7](#_bookmark10)
   2. [Optimization 8](#_bookmark11)
2. [KPIs 9](#_bookmark12)
   1. [KPIs (Key Performance Indicators) 9](#_bookmark13)
3. [Deployment 9](#_bookmark14)

# 

# Abstract

This Power BI-driven Amazon sales analysis report offers a comprehensive exploration of sales data, presenting crucial insights to guide strategic decision-making and performance enhancement. Providing a detailed overview of sales performance metrics like revenue, units sold, and average order value, the report dissects sales by product categories, highlighting top-selling items and their impact on overall revenue. Geographical analysis pinpoints sales distribution across regions, aiding in identifying market trends and growth prospects. Time-based trends reveal seasonal patterns and growth rates, while customer behavior insights illuminate purchasing habits and their influence on sales. With interactive visualizations including graphs, charts, and maps, this report empowers users to explore and analyze data dynamically, facilitating informed decision-making for Amazon sellers, managers, and stakeholders seeking to optimize strategies and drive business growth.

# 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

## Amazon sellers often struggle to extract meaningful insights from their sales data due to its volume and complexity. Without a clear understanding of sales trends, product performance, or customer behavior, sellers may miss opportunities to enhance their offerings, target the right market segments, or optimize their strategies. Traditional data analysis methods may prove insufficient in handling the vast amounts of data generated by Amazon sales. Therefore, there is a critical need for a robust and user-friendly analytics tool, like the Amazon Sales Analysis in Power BI, that can efficiently process, visualize, and analyze this data to provide actionable insights. This tool addresses the challenge of transforming raw sales data into understandable, actionable intelligence, facilitating data-driven decision-making for Amazon sellers and stakeholders.

## Tools used

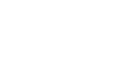
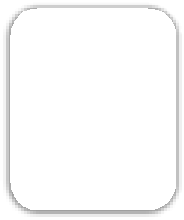
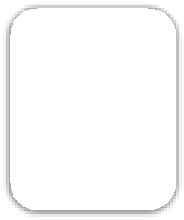
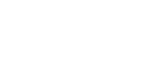
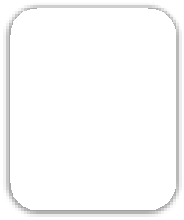
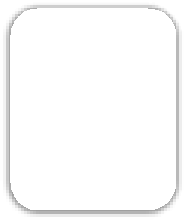
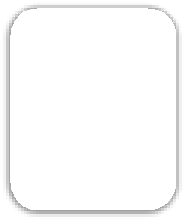
Business Intelligence tools and libraries works such as Numpy, Pandas, Excel, R, Tableau, Power BI are used to build the whole framework.





# Design Details

## Functional Architecture



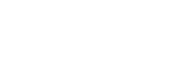
Problem

Evaluation



Data

Gathering



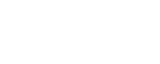
Data

Preprocessing



Data

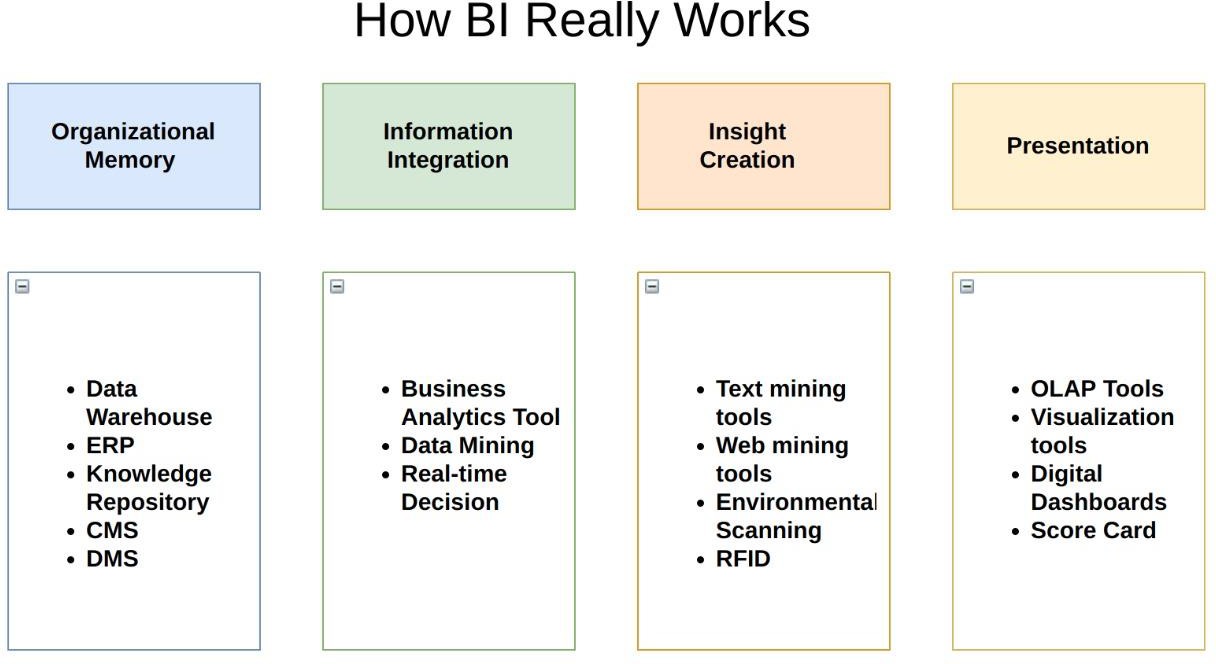
Analysis



Data

Visualization

Figure 1: Functional Architecture of Business Intelligence



**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>Date Time>String

# KPIs & Charts

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the food sales, 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 Amazon Food Sales Data and its relationship with different metrics

1. Total Revenue
2. Total Revenue Last Year
3. Total Profit Margin
4. Total Profit Last Year
5. Customers Last Purchase Date
6. YOY Sales Growth
7. Revenue Contribution %
8. Profit Contribution %

## Charts

Charts displaying an understanding of Amazon Food Sales Data

1. Revenue by country (Pie Chart)
2. Revenue by Region (Pie chart)
3. Top 20 products
4. Top Customers (Matrix Table)
5. Product Overall sales trend (Water fall Chart)

# 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 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.

Amazon Food Sales data is been cleaned using ETL tool Power Query and analyzed on Power BI which will provide a better key insight for the data and tell you a better story of the raw data. The Power BI report is published on workspace where you can play with the data for necessary insights.