**High Level Design (HLD)**

**Amazon Sales Data**

**Revision Number – 1**

**Last Date of Revision –29/01/2023**

**Shreyas Patil**

H HIGH

**Document Version Control**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| 29/01/2023 | 1.0 | Abstract,  Introduction,  General Description  Design Detail, KPI,  Deployment  Final Revision | Shreyas Patil |

HIGH LEVEL DESIGN (HLD

**Contents**

Document Version Control…………………………………………………………………………2

Abstract……………………………………………………………………………………………….3

* Introduction……………………………………………………………………………………….4

1.1 Why this High-Level Design Document?..............................................................4

1.2 Scope……………………………………………………………………………………4

2.2 Tools Used……………………………………………………………………………...5

* Design Detail……………………………………………………………………………………...6

3.1 Functional Architecture………………………………………………………………...6

3.2 Optimization………………………………………………………………………….….7

* KPI……………………………………………………………………………………………….…8

4.1 KPIs (Key Performance Indicators) ……………………………………………..…...9

* Deployment……………………………………………………………………………………......9

**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.

DESIGN (HLD)

**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 before 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

**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**

The goal of this project is to analyse to predict sales trends over the month, year and year month. Sales trend is basically to analyse whether the sales is going in profit or loss. 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

**2.2 Tools used**

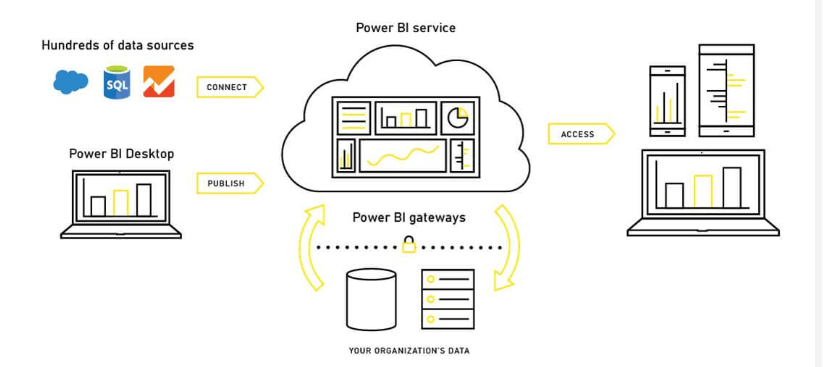
Business Intelligence tools and libraries works such as NumPy, Pandas, Matplotlib, MS-Excel, Power BI, Google Colab and Python Programming Language are used to build the whole framework.

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

* Google Collab is used as IDE.
* Python is the Programming Language used.
* EDA is done using Numpy& Pandas.
* Visualizations were done using Matplotlib.
* Power BI is used for dashboard creation.

**3 Design Details**

**3.1 Functional Architecture**

****

**How does Power BI Work?**

Microsoft Power BI technology is made up of a variety of parts, including the following, to create superior business intelligence solutions:

* Power Query (for data mash-up and transformation)
* Power BI Desktop (a companion development tool)
* Power BI Mobile (for Android, iOS, and Windows phones)
* Power Pivot (for in-memory tabular data modeling)
* Power View (for viewing data visualizations)
* Power Map (for visualizing 3D geo-spatial data)
* Power Q&A (for natural language Q&A)

Power BI user obtains information from various data sources, including files, Azure sources, web services, DirectQuery, and gateway sources. Afterward, they use a client development tool like Power BI Desktop to deal with the data. The imported data is cleaned and modified per the user's requirements.

The data is prepared for usage in report visualizations once it has been transformed and formatted. An assortment of visualizations, including graphs, charts, tables, filters, and slicers, make up a report.

The reports you create on Power BI desktop are published on two platforms: Power BI Service and Power BI Report Server.

Unlike Power BI Report Server, which is an on-premises platform protected by a firewall, Power BI Service is a cloud-based public platform.

Using the visuals from your published reports as a starting point, you can build dashboards on these platforms. Finally, leverage distribution methods like a web browser, Power BI on iPad, tablets, laptops, phones, etc., to collaborate and share your dashboards and reports with other users inside or outside your organization.

**Assumptions**

It is a task that is trivially performed by investor analyst to analysis on the previous historic 17 years dataset and find the equilibrium investment.

Our analysis assumes that all the data provided was true without any corruption and the features mentioned in the raw dataset are the only driving factors of an equilibrium investment.

HIGH LEVEL DESIGN (HLD)

**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 including 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 data 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).

**Performance**

Investment analytics determines the historic all the previous data and it should be as accurate as possible. So that it will not mislead to the future investor. Also, model retraining is very important to improve the performance.

**Security**

Since the investment care analysis consists of years data, the information should be secured.

**Reusability**

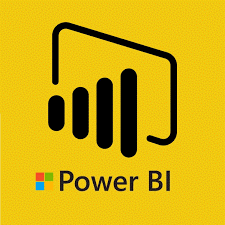
The code written and the components used should have the ability to be reused with no problems.

**Resource utilization**

When any task is performed, it will likely use all the processing power available until that function is finished.

**4 KPI**

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



As and when the system starts to capture the historical/periodic data for a Year, 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 Investment Analysis and its relationship with different metrics

* Month wise Sales trends
* Year Wise sales trends
* Year Month wise sales trends

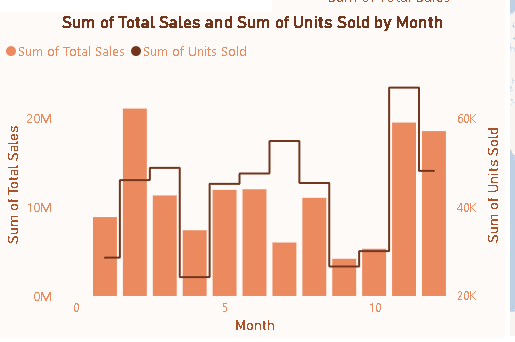
**5 Deployment**

Power BI prioritizes choice in flexibility to fit, rather than dictate, your enterprise architecture. Power BI 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.

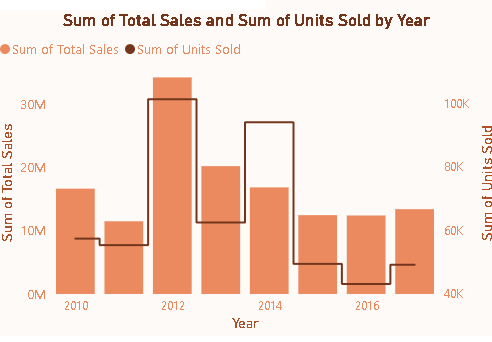
The Dashboard is published on Power BI and an auto-refresh mode has been set so that the dashboard keeps on updating as the real time data loads into the log file.

The Dashboard showcases the multiple insights that has been drawn from the log files as follows:

* **month wise-**This dashboard shows the sales trend over month
* **Year wise:** This dashboard shows the sales trend over years.
* **Year month wise :-**This dashboard shows the sales trend over year months.
* **Month Wise Sales Trends :-**



* **Year Wise Sales Trend**



* **Month Year Vs Sales Trends -**

