

High Level Design

Analyzing Amazon Sales Data

Revision number: 1.0

Last date of revision:
29/11/2022

Deepak Kumar

Document Version Control

Date Issued	Version	Description	Author
29/11/2022	HLD-V1.0	First Version of Complete HLD	Deepak Kumar

Content

Document Version Control	2
Abstract.....	3
1. Introduction	5
1.1. Why this High-Level Design Document?.....	5
1.2. Scope.....	5
2. General Description	6
2.1 Product Perspective & Problem Statement.....	6
2.2 Tools Used	6
3. Design Details	7
3.1 Functional Architecture	7
3.2 Optimization.....	8
4. KPI's.....	9
4.1. KPI's (Key Performance Indicators).....	9
5. Deployment.....	9

Abstract

Amazon Sales data refers to sales, high performing sellers and several other data points. There are millions of Amazon sellers around the world. Amazon sales data Analysis focuses on the process of analyzing consumer behavior, sales, and several other attributes in order to make improved, data-driven decisions. It is key to successfully sustaining their businesses and earning profits and for this purpose, they analyze different metrics like sales, Sales Quantity, Discount rate, Sales over years etc. By analyzing different metrics, you will be able to increase and improve your performance in terms of sales, Items to be sold and discount rates etc. Analysis of the sales data the main factor that contributes to sellers improving their business and increasing their revenue. They can better understand the market trends and customers' buying behaviors and help them cater to what the customers really want. In the world of rising new technology and innovation, E-commerce industry is advancing with the role of Data Analytics. Data analysis can help them to understand their business in a quiet different manner and helps to improve the quality of the service by identifying the weak areas of the business. This study demonstrates the 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 analysis performed to get the key insights from this data based on which business decisions will be taken.

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. Genral 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.

The objective of the project is to Analyze Amazon Sales data to get a substantial data which will help in bringing changes in a business in the future. It will help to reveals flaws in the business model or in the way that one is going about conducting business. Sellers will be able to clearly see where they're losing money, what the problem is, and reduce their losses accordingly. It facilitates coming up with strategic solutions to problems. This project aims to provide visual understanding of the data using Microsoft Power Bi.

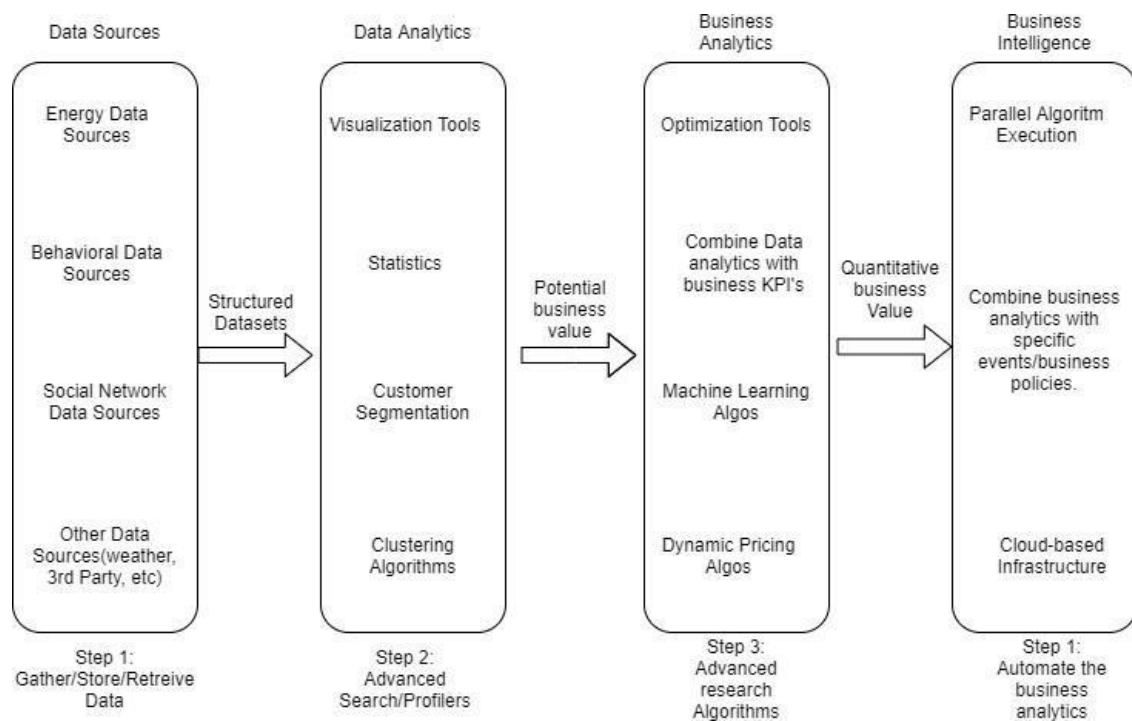
2.2 Tools Used

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



3. Design Details

3.1 Functional Architecture



Functional Architecture of Business Intelligence

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 datasources).

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.
 - LODs - Look at the number of unique dimension members in the calculation.
 - 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 Power BI'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 about the sales of the products in various years, Sales representative according to the number of the sales.

4.1 KPIs (Key Performance Indicator)

Key indicators displaying a summary of Sales Data and its relationships with different metrics.

1. Yearly, Quarterly, Monthly Ups and Downs in Sales & Profits.
2. Items That Generated Highest Sales, Profit etc.
3. Top 5 Items that generated highest Sales and Top 5 Items by Quantity.
4. Bottom 5 Items that generated Lowest Sales and Bottom 5 Items by Quantity.

5. Forecasting.

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. Analyzing Data in Microsoft Power BI empowers you to understand your data through natural language queries that allow you to ask questions about your data without having to write complicated formulas. It will provide answers with stunning visuals such as Charts along with Tooltips, Tables with Conditional Formatting, Maps that can then be placed into the report.