High Level Design (HLD)

Amazon Sales Data Analysis

Revision Number: 1.0
Last date of revision: 06/05/2023

Harshita
Tanwar

Document Version Control

Date Issued	Version	Description	Author
06/05/2023	1.0	First Version of Complete HLD	Harshita Tanwar

Contents

Abstract

A sales report, also known as a sales analysis report, is a document that summarizes a business's sales activities. This report typically includes information on sales volume, leads, new accounts, revenue and costs for a given period. It may also analyse this information along each step of the sales funnel and indicate your sales team's performance (or any gaps therein). These reports might help your company modify its sales approach and other growth initiatives. They can provide insights into sales methodology successes, predictions of future sales data, analyses of performance compared to previous periods, and greater understanding of customer motivations. This work discuss the Detailed analysis of the Sales of Amazon for the year of 2017 to Year 2019 by using the Power bi tool which is capable of showcasing key insights of the sales from the Given data.

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
- 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

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. We need to extract all the Amazon sales datasets, transform them using data cleaning and data preprocessing and then finally loading it for analysis. We need to visualize sales trend month-wise, year-wise and yearly-month wise. Moreover, we need to find key metrics and factors and show meaningful relationships between attributes.

2.2 Tools used

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

3 Design Details

3.1 Functional Architecture

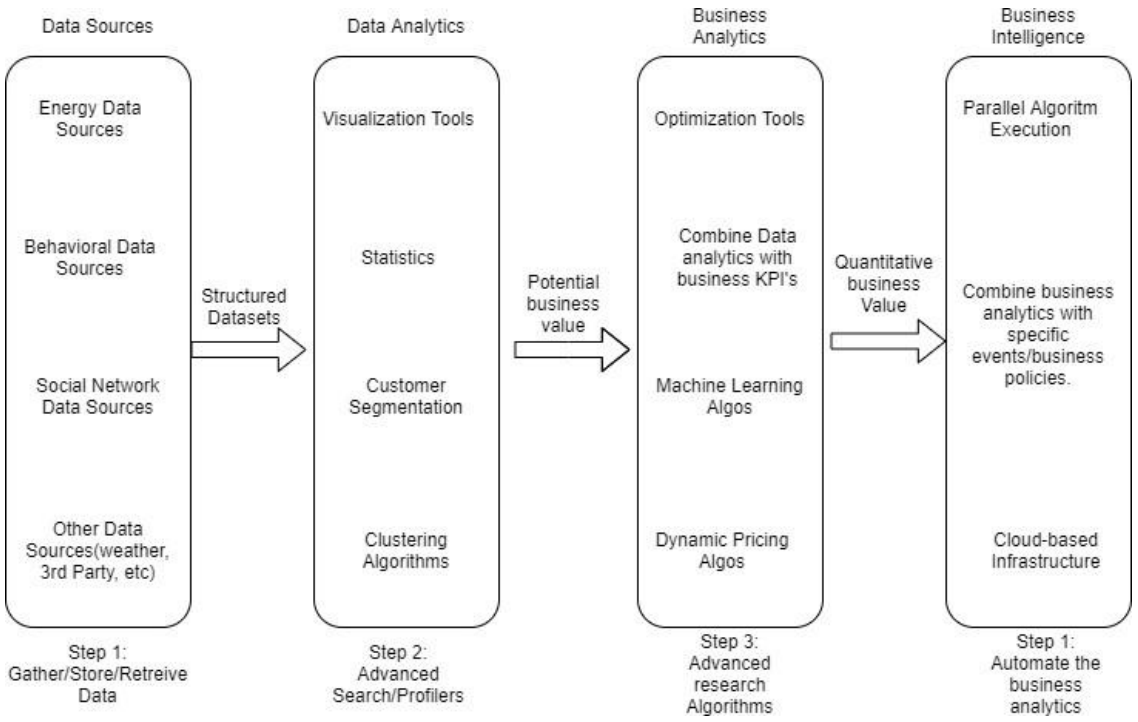
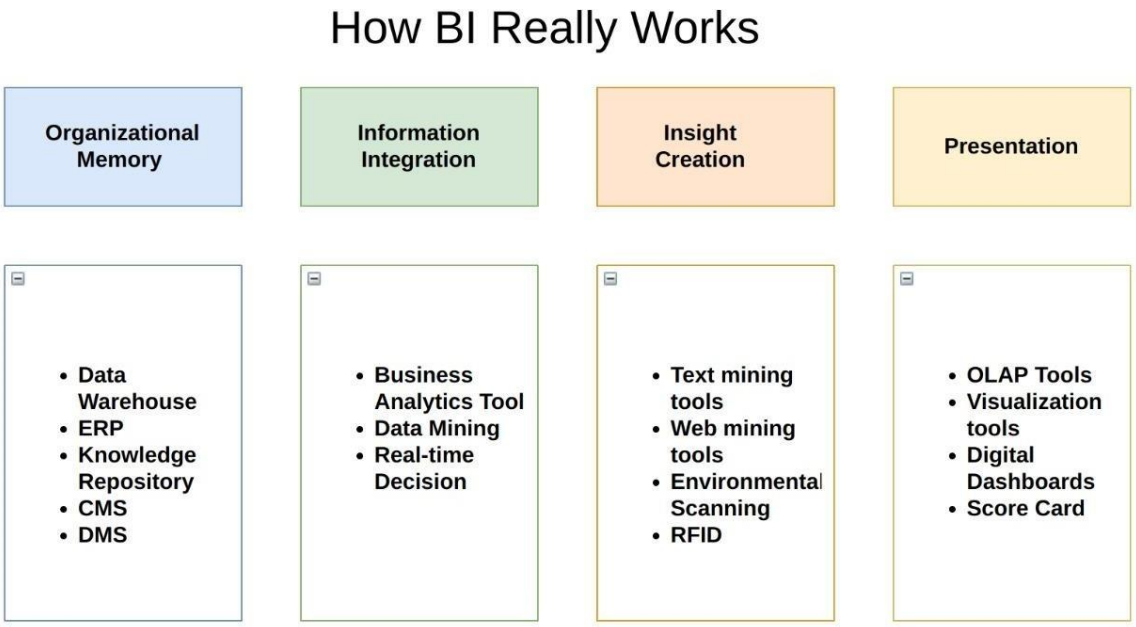


Figure 1: Functional Architecture of Business Intelligence



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 data sources).

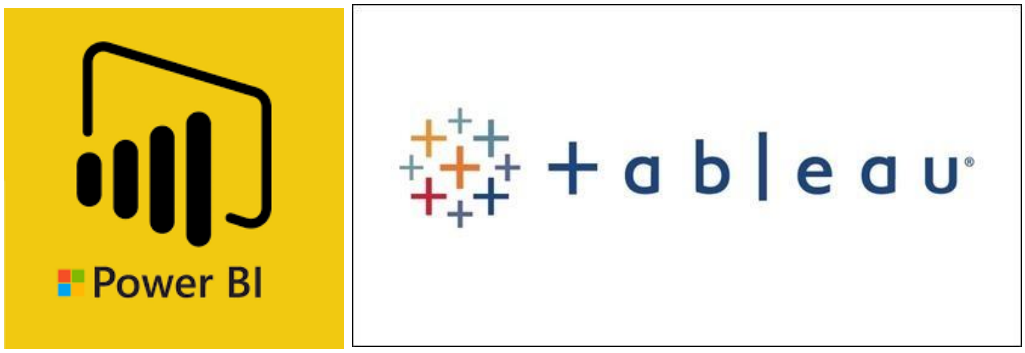
Optimize and materialize your calculations

- 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.
    - 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 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 KPIs

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the sales and product units.



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 Housing Price and its relationship with different metrics

1. Profit by customer sales.
2. Top 5 profits by product.
3. Bottom 5 profits by product.
4. Total sales in different years.
5. Total product units sold in different years.
6. Influence of sales on different sales.
7. Influence of sales in different markets.

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. 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 PowerBi 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 is a Business Intelligence and Data Visualization tool for converting data from various data sources into interactive dashboards and analysis reports. Power BI offers cloud-based services for interactive visualizations with a simple interface for end users to create their own reports and dashboards.

Different Power BI versions like Desktop, Service-based (SaaS), and mobile Power BI apps are used for different platforms. It provides multiple software connectors and services for business intelligence.

In this Power BI training, you will learn all the important concepts of Power BI and develop a foundational understanding of how to use Power BI tool.

Why use Power BI?

Depending on your organizational roles and responsibilities, Tableau Server should be installed by a systems administrator and the designated Tableau Server Administrator in coordination with the appropriate IT roles. For Tableau Online, you will integrate with your existing technology and configure the site settings. The Data & Analytics Survey, completed by business teams, identifies and prioritizes data use cases, audience size, and users. You will use the information collected in both surveys to plan your deployment strategy, including sizing, installation, and configuration of your Tableau Server or integration and configuration of Tableau Online. In addition to installing Tableau Server or configuring Tableau Online, administrators will also need to plan for the client software installation of Tableau Prep Builder, Tableau Desktop, Tableau Mobile, and Tableau Bridge for Tableau Online where applicable.