AMAZON SALES DATA ANALYSIS

HIGH LEVEL DESIGN

A. NOORJAHAN

Abstract

Organizations under the E-commerce industry seek to attain core competence by creating and sustaining a unique process to collect personal information about customers and their purchasing trends. The report critically evaluates how a service-based organization Amazon uses Management information systems as a vibrant tool in attaining competitive advantage through efficient management and acquisition of information. As in today's market without proper sales management, it's very hard to predict how the business is running and how it will be in future. Many companies with proper sales management have shown better growth as they already know which item they have to focus on, which product needs some improvement etc. Sales Management helps in maintaining its customer base for a longer time by providing them attractive offers, as they already have the information's like who are their top customers, whom they have to focus on etc. Sales Management also helps in minimizing the losses. Also, Competition is increasing day by day as many new companies are coming with better management systems and giving tough competition due to that it is now very important to have a proper Sales Management to run any business and to compete with these companies.

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.

Do ETL: Extract-Transform-Load some Amazon dataset and find for me

Sales-trend -> month wise, year wise, yearly_month wise

2.2 Tools used

Business Intelligence tools and libraries works such as NumPy, Pandas, MATPLOTLIB, Power BI are used to build the whole framework.
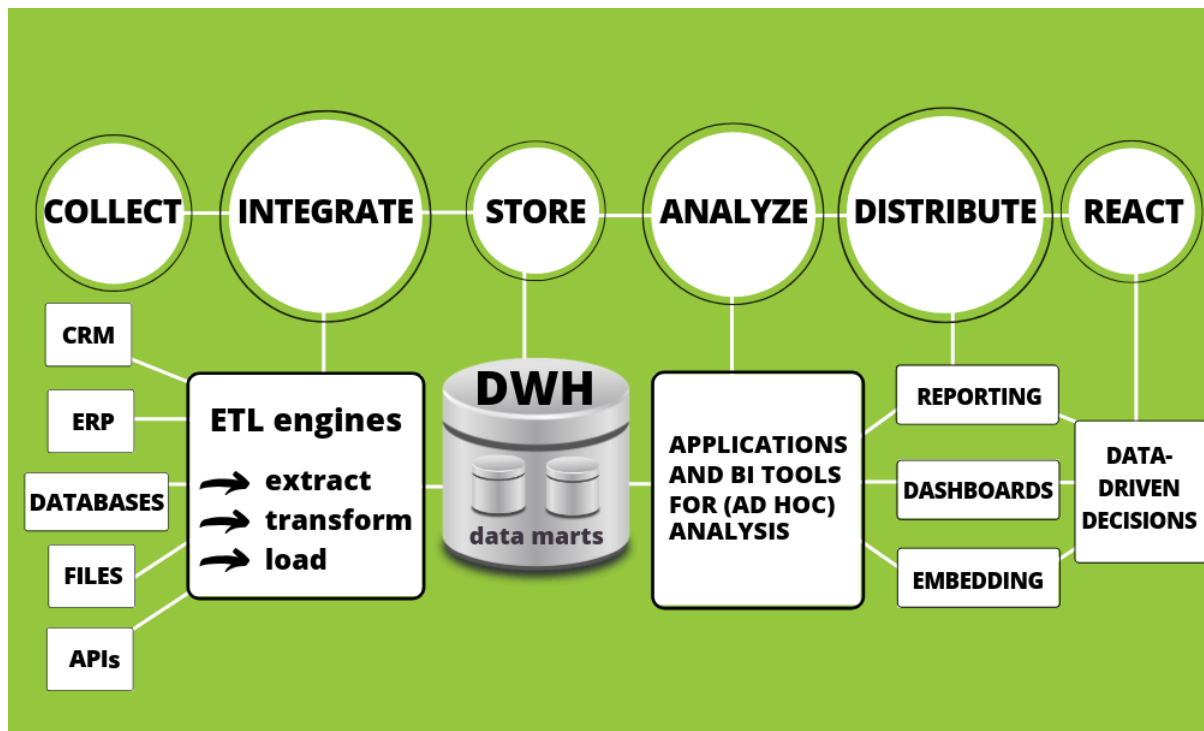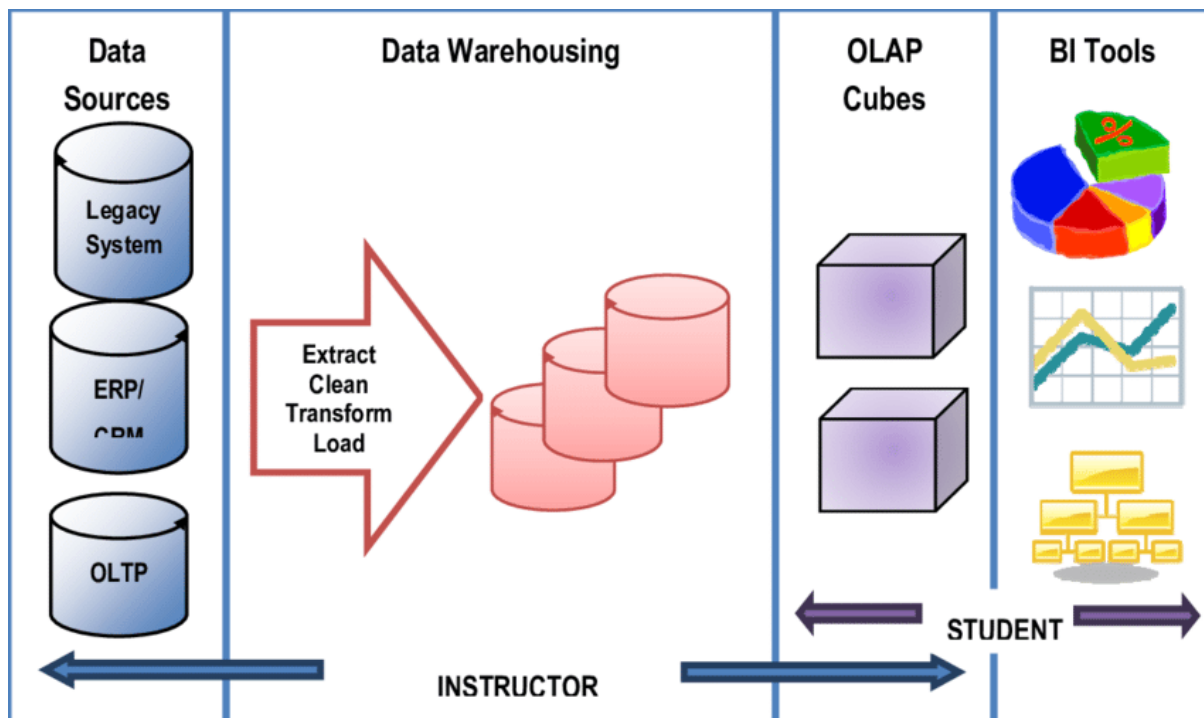
3 Design Details

3.1 Functional Architecture

Working of POWER BI

3.2 Optimization

1. 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

2. 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.

3. 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).

4. 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>Date Time>String

4 KPIs

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators

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 Sales Data Analysis and its relationship with different metrics

1. Impact of discount amount on sales quantity, profit.

2. Impact of cost amount on sales margin amount

3. Influence of sales rep on sales quantity.

4. Items That Generated Highest Sales, Profit etc.

5. Top 5 Items and Top 5 customers by profit.

6. Bottom 5 products by profit and Bottom 5 Items by Quantity.

7.Top 5 cost amount for products and comparing that with profit

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 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. Power BI prioritizes choice in flexibility to fit, rather than dictate, your enterprise architecture. Power BI Desktop and Power BI Service 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
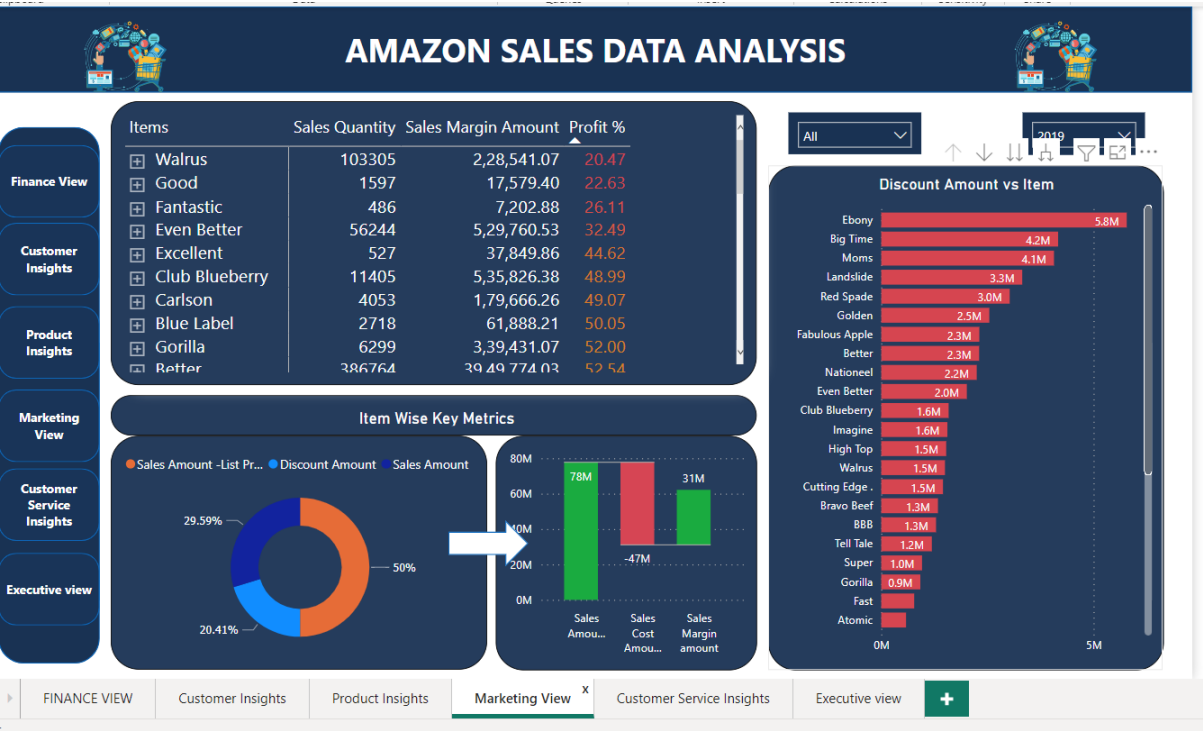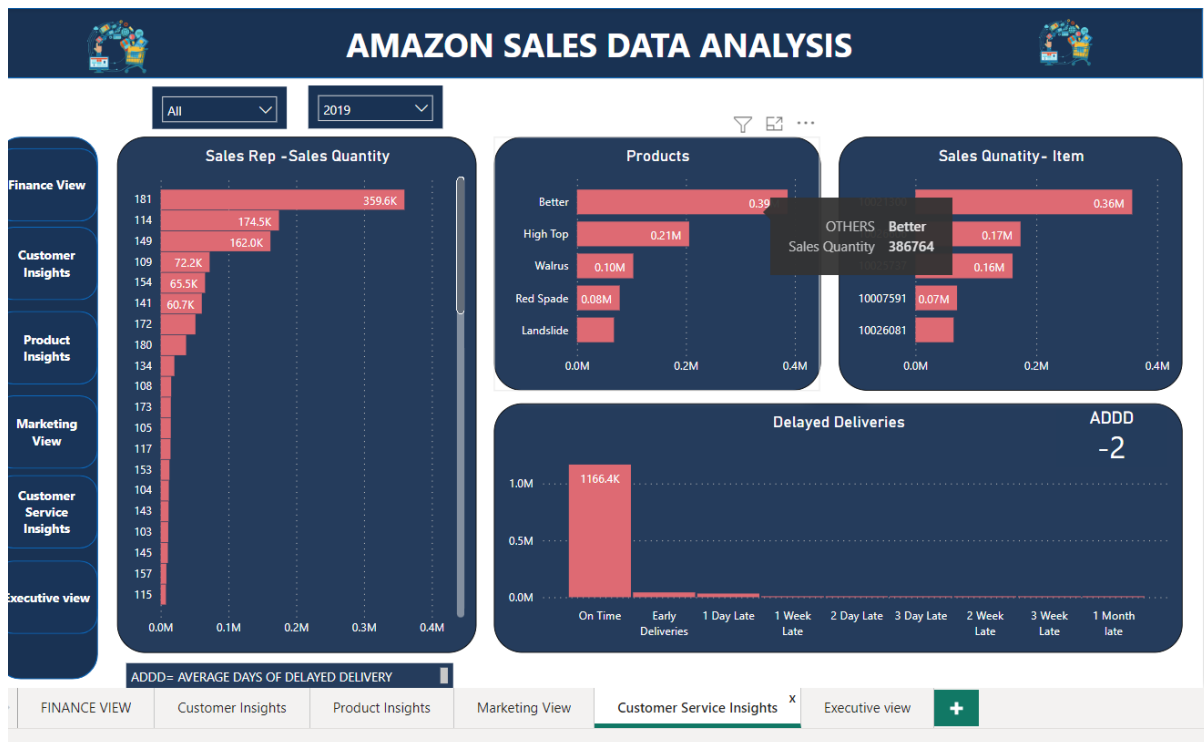
FINANCE VIEW:
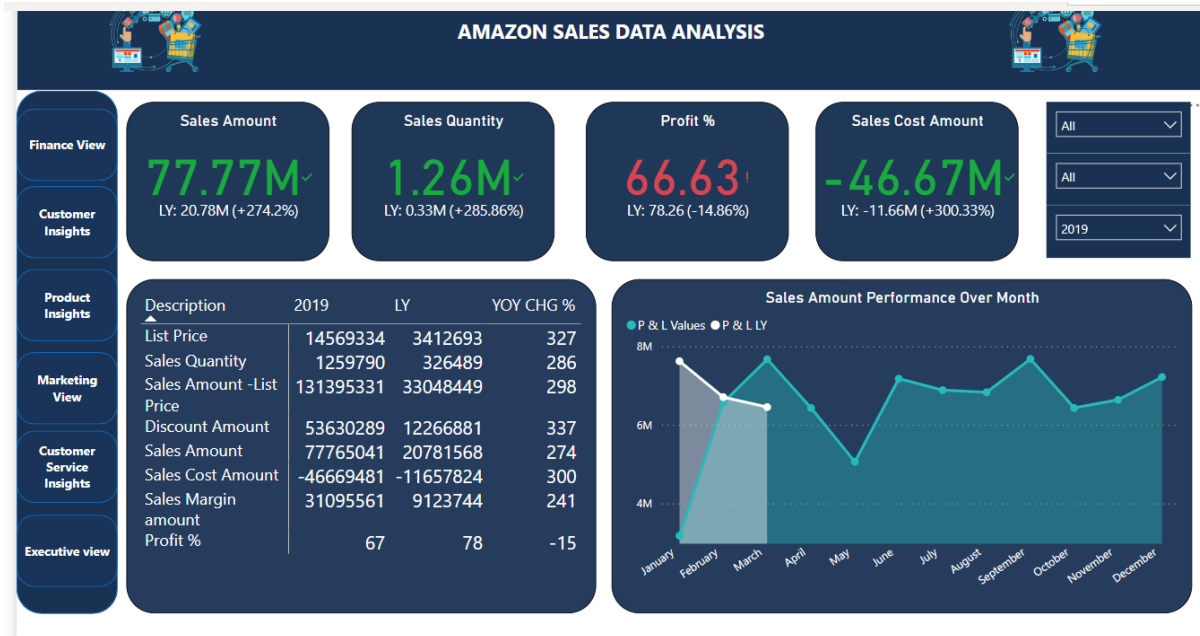
EXECUTIVE VIEW:

MARKETING VIEW:

PRODUCT INSIGHTS:
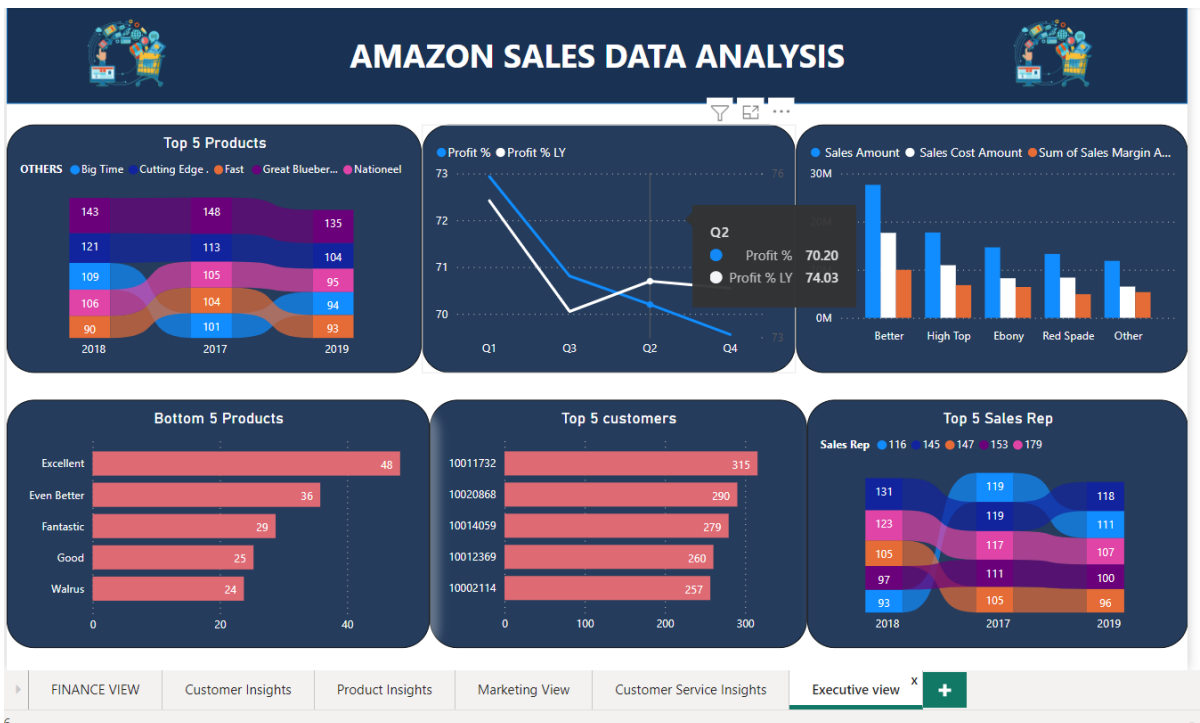
CUSTOMER SERVICE INSIGHTS:

CUSTOMER INSIGHTS: