High Level Design (HLD)

Amazon Sales Analysis

Revision Number: 1.0
Last date of revision: 20/01/2022

Freeva Dsouza

Document Version Control

Date Issued	Version	Description	Author
20th Jan 2022	1.0	First Version of Complete HLD	Freeva Dsouza

Contents

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.

In this Sales Analysis project, I have used Python Pandas and Matplotlib to analyze data to answer key questions foe what drives business performance. The data contains hundreds of thousands of food orders broken down by Sales Price, Order number, Delivery Date, Discount Amount, etc.

The detailed analysis gives an idea of overall sales trend, product which is sold the most and what products sold in terms of profit or loss.

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

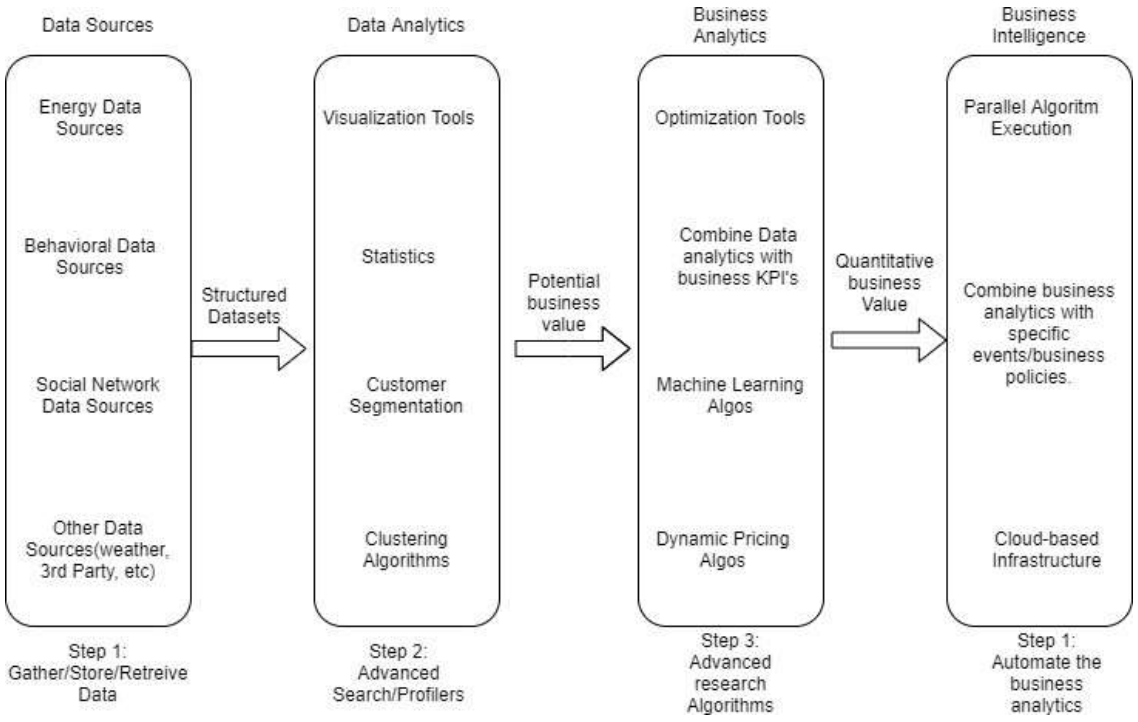
The objective of the project is to creative rich interactive visualizations through data interpretation and analysis and identified the sales trends yearly, monthly. The stakeholders would like to know most profitable products, expensive products ordered, revenue, etc.

2.2 Tools used

Python – Data Cleaning and Data Analysis

3 Design Details

3.1 Functional Architecture



3.2 Optimization

Your data strategy drives performance

- Minimize the number of fields
- Imputing missing values
- Dropping unwanted columns and null values.

4 KPIs

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

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 Product analysis and its relationship with different metrics

1. Overall sales Trend
2. MoM growth
3. YoY Growth
4. Products in terms of Profit and Loss
5. Product most sold
6. Total Revenue

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 Tableau at scale, as well as organizing, orchestrating, and unifying disparate sources of data for business users and experts alike to author and consume content.

The data is collected and prepared by the company. The data has been cleaned and preprocessed in order to make suitable for analysis.

The cleaned data is used to perform EDA (Exploratory Data Analysis) to find as many insights as possible and establish trends and correlations between various attributes and features of data.