Low-Level Design (LLD)

ANALYZING AMAZON SALES DATA

Revision Number – 1.0

Last Date of Revision - 15/11/2024

Suresh Gupta

Document Control

Date	Version	Description	Author
15/11/2024	1.0	Entire Document of LLD	Suresh Gupta

Contents

1. Introduction

1.1 What is Low-Level Design Document?

The goal of the Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Sales Analysis dashboard. LLDD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 What is Scope?

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code, and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

1.3 Project Introduction

This is a Sales Management project. Organizations in the e-commerce business strive for core competence by developing and keeping a distinct approach for collecting personal information about clients and their shopping habits. The research critically examines how service-based firms, such as Amazon, use management information systems to gain a competitive edge through effective information management and acquisition. The goal of this project is to analyze Amazon Sales Data in order to gain useful information. A Sales dataset is provided to accomplish this, which includes sales amount, list price, cost price, and so forth.

2. Problem Statement

To meet increased competition and the need for enhanced distribution techniques to decrease costs and increase profits, sales management has grown in importance. Today, the most crucial role in a commercial and business firm is sales management. Do ETL: Load some Amazon datasets and extract-transform them for me. Sales trend -> month-by-month, year-by-year, year-by-year. Find essential measurements and factors and demonstrate the meaningful relationships between them.

3. Dataset Information

Invoice Date: The day on which the Invoice was generated.

Discount Amount: Total discount provided on any item.

Sales Amount: Total Sales Price of an Item.

Sales Margin Amount (Profits): Sales Margin Amount is the difference between Sales Cost Amount & Sales Amount.

Sales Cost Amount: Total Cost Price of any Item.

Sales Price: Sales Price of any particular Item.

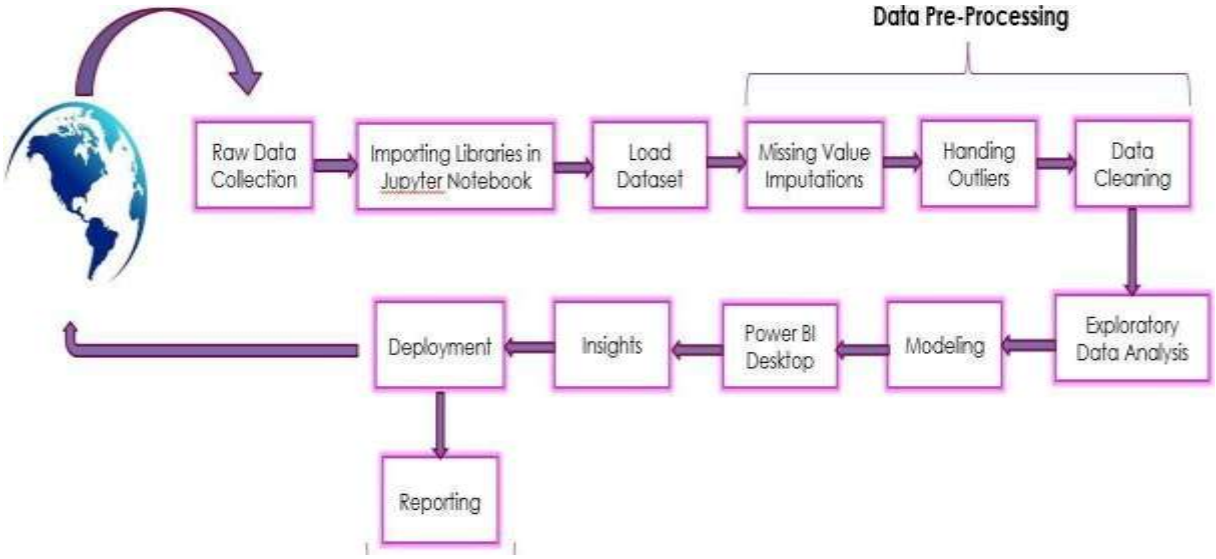
List Price: Basic Price of an Item as published on the price list.

Sales Rep: A person whose job is to sell products or services for a company.

U/M: Unit of Measure

CustKey: It is a Unique Number on the Invoice that is used to reference customers' accounts.

4. Architecture



4.1 Architecture Description

1. Raw Data Collection

The Dataset was taken from iNeuron's Provided Project Description Document. https://docs.google.com/spreadsheets/d/1h3EsPf-fTLrzpP7sGeyuRnGBXrdJRcXY/edit?usp=sharing&ouid=105519103382792804653&rtpof=true&sd=true

2. Data Pre-Processing

Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data fed to the model to train. This process includes-

a) Handling Null/Missing Values

b) Handling Skewed Data

c) Outliers Detection and Removal

3. Data Cleaning

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

a) Remove duplicate or irrelevant observations

b) Filter unwanted outliers

c) Renaming required attributes

4. Exploratory Data Analysis (EDA)

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypotheses and check assumptions with the help of summary statistics and graphical representations.

5. Reporting

Reporting is a most important and underrated skill in the data analytics field.
Because being a Data Analyst you should be good with the easy and self-explanatory reports because your model will be used by many stakeholders who are not from a technical background. a) High-Level Design Document (HLD)

b) Low-Level Design Document (LLD)

c) Architecture

d) Wireframe

e) Detailed Project Report

f) PowerPoint Presentation

6. Modelling

Data Modelling is the process of analyzing the data objects and their relationship to the other objects. It is used to analyze the data requirements that are required for business processes.

The data models are created to store the data in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

7. Deployment

We created a Power BI Dashboard