Low-Level Design (LLD)

ANALYZING AMAZON SALES DATA

Revision Number – 1.0

Last Date of Revision - 27/01/2024

Yash Singh

Document Control

Date	Version	Description	Author
27/01/2024	1.0	Entire Document of LLD	Yash Singh

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

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.

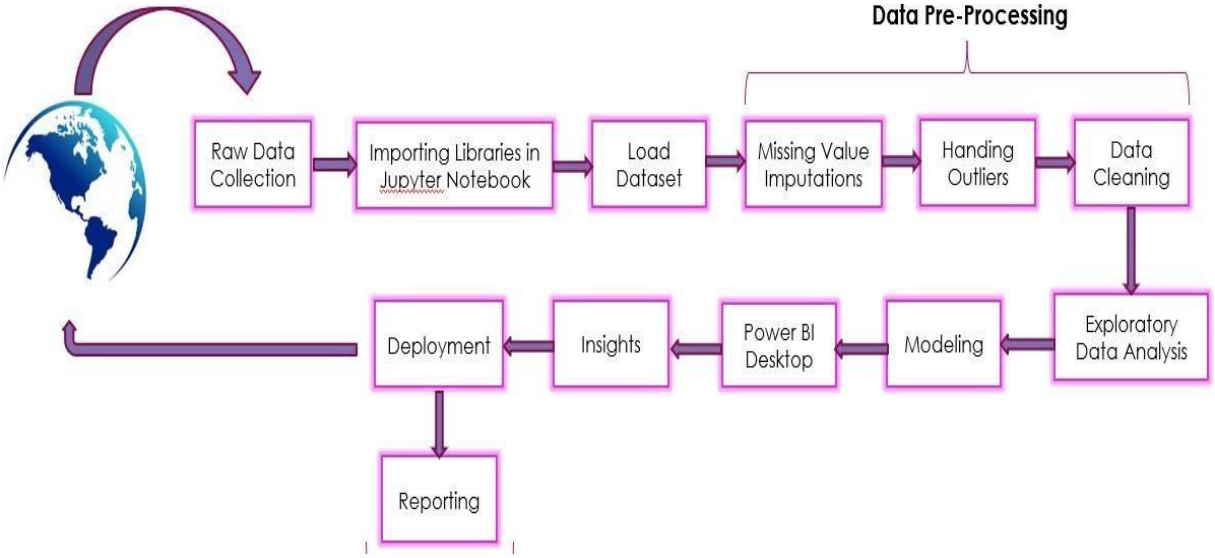
2. Problem Statement

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 : Extract-Transform-Load some Amazon dataset and find for me Sales-trend -> month wise , year wise , yearly_month wise

3. Dataset Information

1. User_ID: Unique identifier for each user.
2. Cust_name: Customer's name.
3. Product_ID: Unique identifier for each product.
4. Gender: Gender of the customer (M for Male, F for Female).
5. Age Group: Age group of the customer.
6. Age: Age of the customer.
7. Marital_Status: Marital status of the customer (1 for married, 0 for unmarried).
8. State: State where the customer is located.
9. Zone: Geographical zone of the state.
10. Occupation: Occupation of the customer.
11. Product_Category: Category of the purchased product.
12. Orders: Number of orders placed.
13. Amount: Amount spent on the purchases.
14. Status: Status information (seems to be empty or not provided).
15. unnamed1: Unnamed column with no clear purpose.

4. Architecture



4.1 Architecture Description

1. Raw Data Collection

https://drive.google.com/drive/folders/1FkmFVL8wlJmQWP1z52TD8PlhOJhitTyI?usp=sharing

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

I created a Power BI Dashboard