Low-Level Design (LLD)

Low-Level Design (LLD)

ANALYZING AMAZON SALES DATA

Version – 1.0

25-07-2022

By : Kshitij Yejare

Document Control

Date	Version	Description	Author
25-07-2022	1.0	Introduction, Problem Statement	Kshitij
01-09-2022	1.1	Sales Insights update	Kshitij

Contents

Topic No.	Topic Name	Page No.
1	**Introduction** 1.1 What is Low-Level Design Document? 1.2 Scope 1.3 Project Introduction	5
2	**Problem Statement**	5
3	**Dataset Information**	6
4	**Architecture** 4.1 Architecture Description	7-8

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 programs 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 project about Sales Management. 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 service-based organizations -Amazon use Management information systems to attain competitive advantage through efficient management and acquisition of information. The purpose of this project is to analyze Amazon Sales Data to obtain meaningful information. To do that, a Sales dataset is provided, which includes sales amount, list price, cost price, etc.

2. Problem Statement

Sales management has gained importance to meet increasing competition and the need for improved distribution methods to reduce cost and increase profits. Today, sales management 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. Find key metrics and factors and show the meaningful relationships between attributes.

3. Dataset Information

Invoice Date: Day on which Invoice 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 a difference of 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

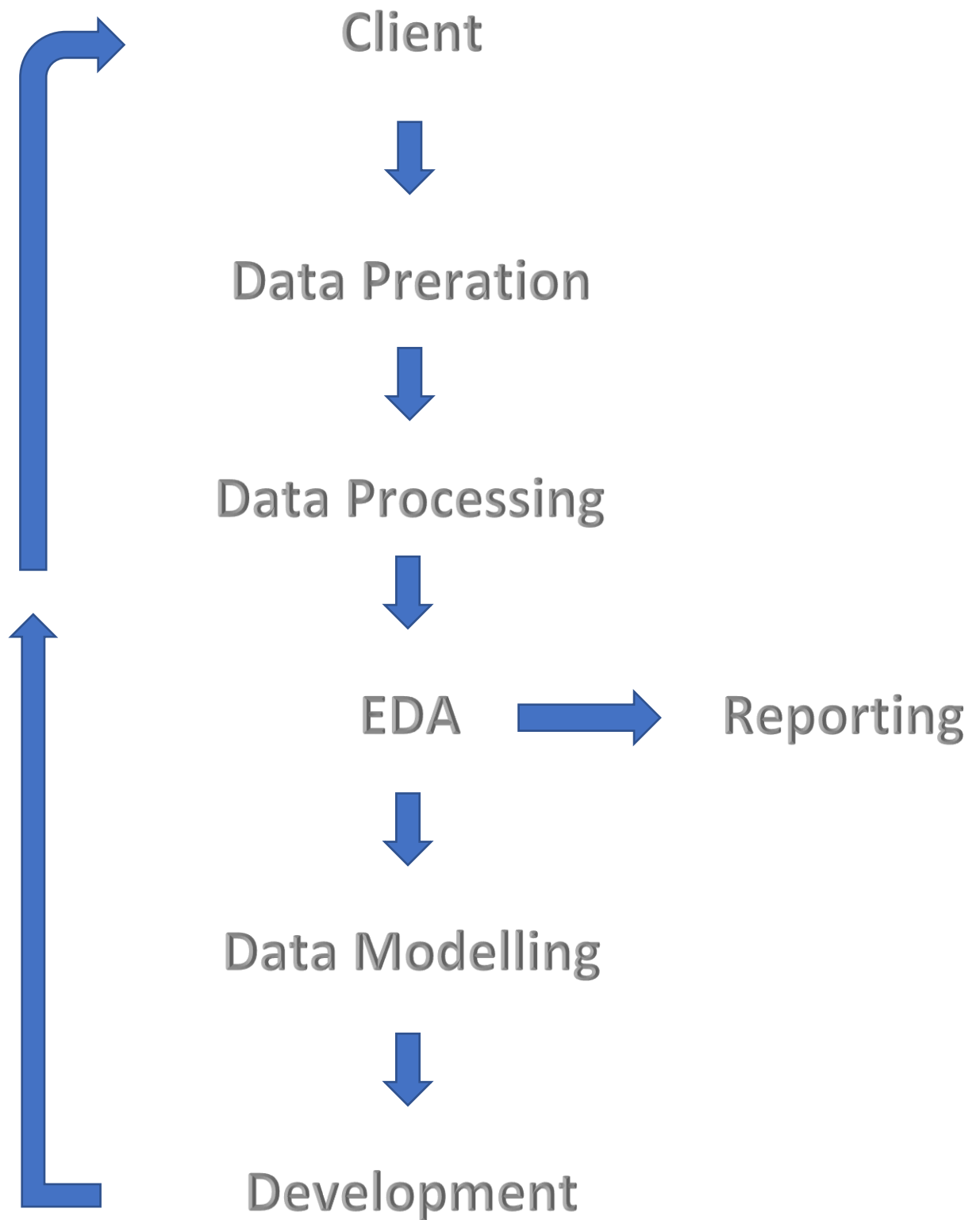**CustKey:** It is a Unique Number on the Invoice that is used to reference customers' account.

Revenue = Sales Amount

Profit = Sales Margin Amount

Total Sales = Sales Amount Based On List Price

Total Quantity = Sales Quantity.

4. Architecture

Client

↓

Data Preration

↓

Data Processing

↓

EDA → Reporting

↓

Data Modelling

↓

Development

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 of a data analytics field. Because being a Data Analyst you should be good in the easy and self-explanatory report 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 the 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