Low Level Design (LLD)

Analyzing Amazon Sales Data

Written By / Author	Deepak Kumar
Document Version	LLD-V1.0
Last Revised Date	29/11/2022

Document Version Control

Date Issued	Version	Description	Author
29/11/2022	LLD-V1.0	First Version of Complete LLD	Deepak Kumar

Reviews:

VERSION	DATE	REVIEWER	COMMENTS
V1.0	29/11/2022	Deepak Kumar	Document, Version Control and Architectural description added.

Approval Status:

VERSION	REVIEW DATE	REVIEWED BY	APPROVED BY	COMMENTS

Abstract

Amazon Sales data refers to sales, high performing sellers and several other data points. There are millions of Amazon sellers around the world. Nearly half of them are self-employed and live off their ecommerce/retail businesses (47%), and 22% earn income from their Amazon businesses alone. Amazon sales data Analysis focuses on the process of analysing consumer behaviour, sales, and several other attributes in order to make improved, data-driven decisions. It is key to successfully sustaining their businesses and earning profits and for this purpose, they analyse different metrics like sales, Sales Quantity, Discount rate, Sales over years etc. By analysing different metrics, you will be able to increase and improve your performance in terms of sales, Items to be sold and discount rates etc. Analysis of the sales data the main factor that contributes to sellers improving their business and increasing their revenue. They can better understand the market trends and customers' buying behaviours and help them cater to what the customers really want.

In the world of rising new technology and innovation, E-commerce industry is advancing with the role of Data Science and Analytics. Data analysis can help them to understand their business in aquiet different manner and helps to improve the quality of the service by identifying the weak areas of the business. This study demonstrates the how different analysis help to make better business decisions and help analyze customer trends and satisfaction, which can lead to new and better products and services. Different analysis performed to get the key insights from this data based on which business decisions will be taken.

Introduction

1.1 Why this Low-Level design document?

The purpose of this LLD or a Low-Level Design (LLD) document is to give the internal logical design of the actual program code for Amazon Sales Data Analysis project. LLD describes the class diagrams withthe methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document. This document is intended for both the stakeholders and the developers of this project and will be proposed to the higher management for its approval.

The main objective of the project is to analyse the various aspects with different use cases which covers many aspects of amazon sales. It helps in not only understanding the meaningful relationshipsbetween attributes but it also allows us to do our own research and come-up with our findings.

1.2 Scope

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

This study demonstrates the how different analysis help out to make better business decisionsand help analyse customer trends and satisfaction, which can lead to new and better products and services.

1.3 Project Introduction

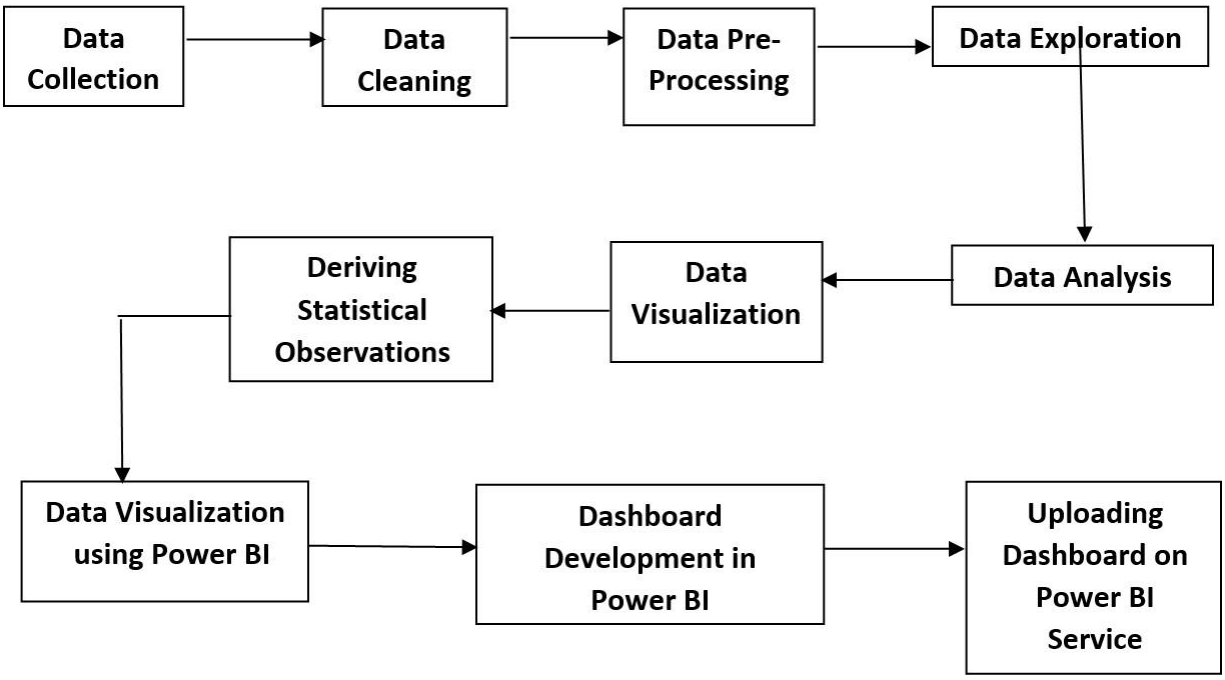
This is a project about Amazon Sales Data Analysis. The report critically evaluates how service-based organizations -Amazon use Management information systems to attain competitive advantage through efficient management andacquisition of information. The purpose of this project is to analyze AmazonSales 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 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
Find key metrics and factors and show the meaningful relationships between attributes.

Do your own research and come up with your findings

3. Architecture



3.1 Architecture Description

3.1.1 Data Collection

As we have seen earlier, in our Dataset preview, we have around hundreds of records with 18 different features. Features are distributed as 8 Continuous features, 9 Categorical features and 1 Geographical feature in our Sales Dataset. The "100 Sales Records.csv" dataset was given in the form of Comma Separated Value (.csv) format and we load this dataset in Power BI.

3.1.2 Data Cleaning

At this stage, based on the given dataset and business problems we have defined the several Use Cases to perform the analysis on and this will definitely help out get the key insights from this data based on which business decisions will be taken. Furthermore, It helps in not only understanding the meaningful relationships between attributes, but it

also allows us to do our own research and come-up with our findings. Fortunately, provided dataset has no missing values, hence we proceededfurther.

3.1.3 Import the Dataset

In this process, we transformed the data into highly understandable format by applying Python's Pandas and Numpy libraries. We eliminated the unnecessary rows and columns. We transformed categorical data into numerical dummy variables in order to make it more suitable for statistical numerical analysis which renders meaning to the data up to higher extent. We emphasized upon removing redundancy in data and made it suitable for performing exploratory data analysis.

3.1.4 Exploratory Data Analysis

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

3.1.5 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 selfexplanatory report because your model will be used by many stakeholderswho are not from a technical background.

- High-Level Design Document (HLD)
- Low-Level Design Document (LLD)
- Architecture
- Wireframe
- Detailed Project Report

3.1.6 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 tostore the data in a database. The Data Model's main focus is on what data isneeded and how we have to organize data rather than what operations we have to perform.

4. Deployment

we extracted data from the datasets into the Power BI Desktop. We prepared various charts and plots based on meaningful data. We established several mathematical relations between the numerical attributes present in the data. We prepared various visuals for different sections of the dataand filtered the charts according to various parameters to make it more user interactive and user friendly.

The resulting visual representation of data makes it easier to identify and share insights about the information represented in the data.