

# AMAZON SALES ANALYSIS



VERSION: 1.0  
DATE 14-1-2023

RAVIKUMAR SURAM

## DOCUMENT VERSION CONTROL

### AMAZON SALES ANALYSIS BUSINESS INTELLIGENCE PROJECT

DATE	VERSION	AUTHOR	CHANGE
14-1-2023	1.0	RAVIKUMAR SURAM	FIRST VERSION OF LLD

## Contents:

### Document Version Control

1. Introduction.....	4
1.1 Why this Low-Level Design Document?.....	4
1.2 Scope .....	4
1.3 Project Introduction .....	4
2. Problem Statement.....	4
3. Dataset Information.....	5
4. Architecture.....	6
4.1 Architecture Description.....	6

## 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 Amazon Sales Analysis dashboard. LLD 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

- E-Commerce business modules are growing irrespective of the state and country barriers. At the same time it's difficult to maintain the business profitable quarter-on-quarter and year-on-year due to several reasons.
- Sales analysis is the important part of the company's growth. Selling the products and making profit out of it, is totally depend on cost of manufacturing and improved methods of distribution.
- In this project we will try to understand the Amazon Sales in various countries and scope of improvement.

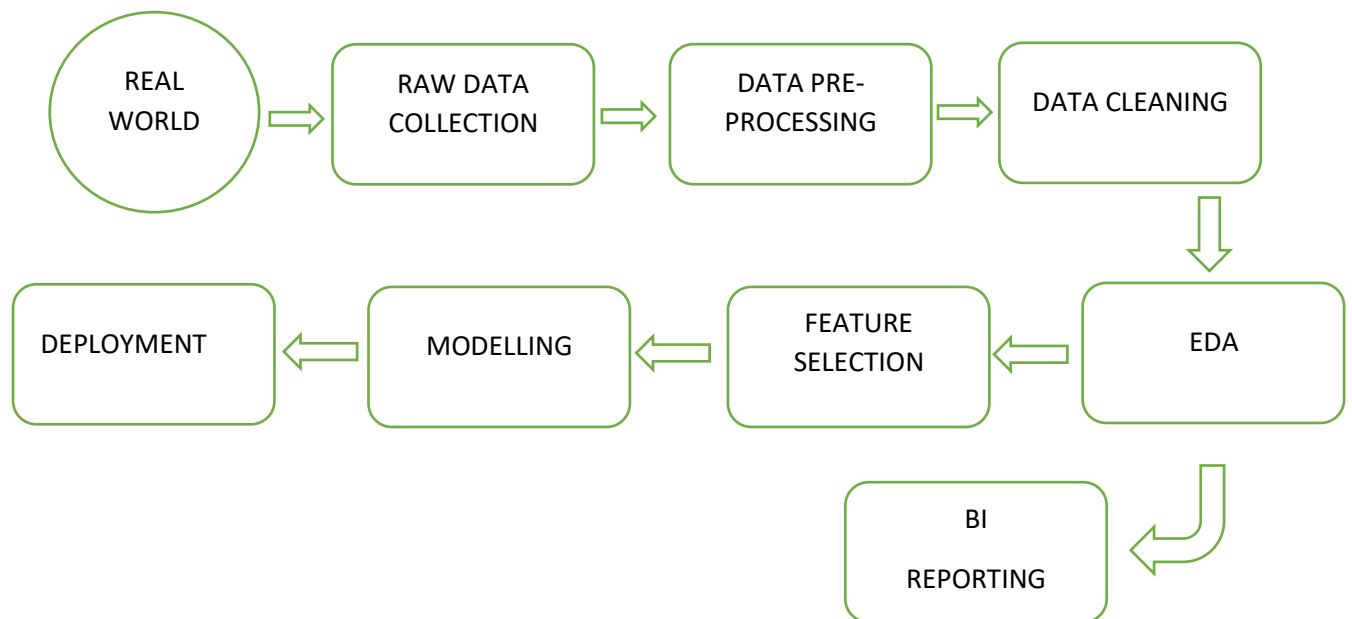
## 2. Problem Statement

- 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.
- To do EDA and to extract the meaningful insights from the data based on the given information.

### 3. Dataset Information

- Region: Categorical data (The regions where amazon is selling the products.)
- Country: Categorical data (Country names)
- Item type: Categorical data (products name)
- Sales channel: Categorical data (offline/online)
- Order priority: Categorical data (H,C,L,M)
- Order date: Date (day when order was placed)
- Order id: Integer (generated order number for each order )
- Ship date: Date (day when product was shipped)
- Units sold: Integer (quantity of the product)
- Unit price: Float (selling price for one unit of product)
- Unit cost: Float (manufacturing cost of the product)
- Total revenue: Float (sales = units sold \* unit price)
- Total cost: Float (total units \*price)
- Total profit: Float (profit = Total revenue – Total cost )

## 4. Architecture



### 4.1 Architecture Description

#### 1. Raw Data Collection

The Dataset is available at

[https://drive.google.com/file/d/1aFnyEfR6hNI76LmNraVi04Pb8iz2YnCn/view?usp=share\\_link](https://drive.google.com/file/d/1aFnyEfR6hNI76LmNraVi04Pb8iz2YnCn/view?usp=share_link)

#### 2. Data Pre-Processing

Before building any model, it is necessary to clean the data to feed the correct data to the model to bring meaningful insights and build a model and predict. Model performance depends on the quality of data is feed to the model to train. This Process includes

- a) Handling Null/Missing Values
- b) Handling Skewed Data
- c) Outliers Detection and fixing

### 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) Fixing the 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 hypothesis and to 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 easy and self-explanatory report because your model will be used by many stakeholders who are not from technical background.

- a) Low Level Design Document (HLD)
- b) Low Level Design Document (LLD)
- c) Architecture
- d) Wireframe
- e) Detailed Project Report
- f) Power Point Presentation

### 6. Modelling

Data Modelling is the process of analysing the data objects and their relationship to the other objects. It is used to analyse the requirements that are required for the business processes. The data models are created for the data to be stored 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

### Power BI Dashboard.

