# Amazon Sales Data Analysis

## LOW LEVEL DESIGN(LLD)

Shubham Ghosh
&
Vipin Bhagat

VERSION:
1.0 DATE
30/8/2023

## Document Version Control

| Date Issued | Version | Description | Author |
|---|---|---|---|
| 12-09-2023 | 1.0 | First Version on complete LLD | Shubham Ghosh & Vipin Bhagat |
| | | | |

## Project Introduction

The project focuses on the critical evaluation of how service-based organizations like Amazon leverage Management Information Systems (MIS) to gain a competitive edge through effective information management and acquisition. The primary objective is to perform a thorough analysis of Amazon Sales Data to extract valuable insights. The provided Sales dataset contains crucial information such as sales amounts, list prices, cost prices, and more. By harnessing MIS and data analysis techniques, this project aims to uncover actionable information that can drive decision-making and strategic initiatives within Amazon's sales operations.

# Table of Contents

## Abstract

Amazon's Sales data encompasses a wide range of information, including sales figures and insights into high-performing sellers. With millions of sellers worldwide, the analysis of this data is crucial for making informed, data-driven decisions. By evaluating metrics such as Total Sales, Sales Quantity, Total Profit, Sales, and comparing them to historical data like Last Year Sales, Amazon aims to enhance its performance. This analytical approach not only helps sustain businesses and boost profits but also provides valuable insights into consumer behavior, market trends, and customer preferences, ultimately guiding Amazon in meeting customer demands effectively.

In the rapidly evolving landscape of technology and innovation, the E-commerce industry is leveraging Data Analytics to drive its progress. Data analysis offers a unique perspective, enabling businesses to enhance service quality by pinpointing areas of weakness. This study illustrates how various analyses contribute to informed decision-making, customer trend analysis, and improved satisfaction, ultimately leading to the development of innovative products and services. Leveraging a comprehensive dataset encompassing Profit, Revenue, Cost, Unit Sold, and more across different regions and countries, the primary objective is to unveil sales trends on a monthly, yearly, and quarterly basis. This analysis aims to extract critical insights, emphasizing key indicators and metrics that shape customer preferences and choices.

# 1. Introduction

## 1.1 Why this Low-Level Design Document?

The Low-Level Design (LLD) document for the Amazon Sales Data Analysis project serves as a blueprint for the actual program code. It provides an internal logical design, including class diagrams with methods and class relationships, as well as program specifications. The LLD is a valuable resource for both stakeholders and developers, facilitating the coding process based on the document's guidance. Ultimately, it will be presented to higher management for approval.

The primary goal of the project is to conduct a comprehensive analysis of various aspects of Amazon sales, encompassing diverse use cases. This analysis not only uncovers meaningful relationships between attributes but also empowers researchers to conduct their investigations and generate valuable findings. The LLD document plays a pivotal role in translating these objectives into actionable code and software development.

## 1.2 Scope

Low-Level Design (LLD) is a systematic approach to designing at the component level, involving a step-by-step refinement process. It encompasses the design of data structures, software architecture, source code, and even performance algorithms. Typically, the initial data organization is outlined during requirement analysis and is further refined during the data design phase.

In the context of business analysis, this study showcases the value of various analyses in facilitating informed decision-making. By analyzing customer trends and satisfaction, businesses can gain insights that drive the development of new and improved products and services. LLD plays a crucial role in translating these analytical insights into concrete software and system designs to support these business objectives.
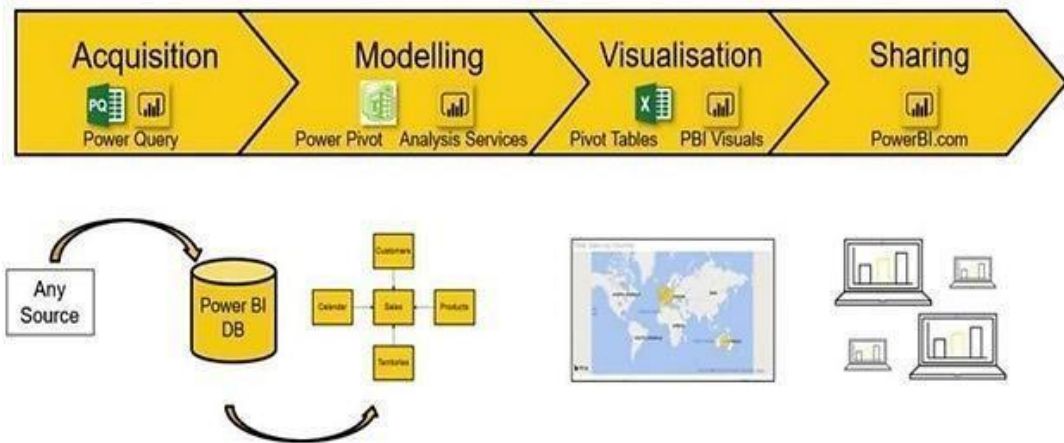
## 2.Architecture
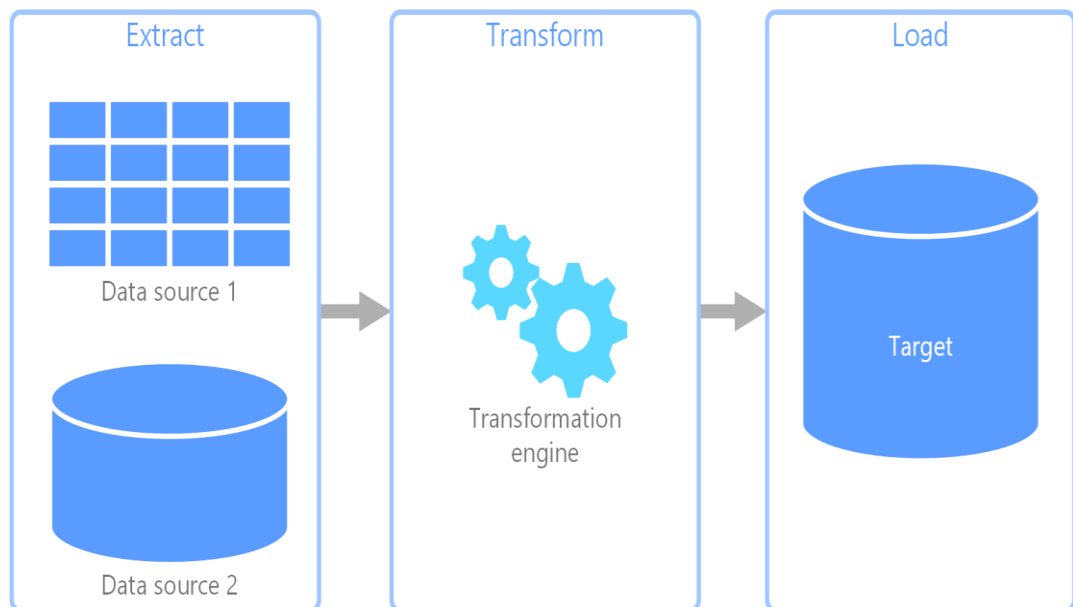
ETL (Extract, Transform, Load) in Power BI is a crucial process for data preparation, involving the extraction of data, its transformation into a suitable format, and loading it into the analysis tool. This process aims to clean and structure data, making it ready for meaningful analysis and visualization.

The results of ETL play a pivotal role in shaping business decisions, which can have significant consequences. When ETL is not executed properly, it can harm a business in various ways, including the loss of clients, inaccurate decision-making, and other operational issues. Conversely, when done effectively, ETL can enhance the efficiency and accuracy of subsequent activities, ultimately leading to better-informed decisions and improved business outcomes. The quality of ETL processes directly impacts the reliability and trustworthiness of data-driven decision-making in organizations..

Below are the following steps to follow for ETL:
1. Data Sourcing
2. Data Cleaning
3. Data Modelling
4. Data Visualization

# 3. Architecture Description

## 3.1 Data Sourcing

- The dataset is in CSV format.

- Microsoft Excel is used for data loading.

- The dataset is publicly available on GitHub.

- Title: Amazon Sales Data Analysis.csv

- Source: GitHub

## 3.2 Data Overview

- Dataset size: 12.4 KB

- File format: .csv

- Number of rows/records: 100

- Number of attributes: 14

## 3.3 Data Description

The dataset includes the following attributes:

- **Region**

  - Description: Name of the Region

  - Datatype: string

- **Country**

  - Description: Name of different Countries

  - Datatype: string

- **Item Type**

  - Description: Different Product types sold on Amazon

  - Datatype: string

- **Sales Channel**

  - Description: Mode of shopping (Online or Offline)

  - Datatype: string

- **Order Priority**

  - Description: Priority of Sales (Ranges between low to high)

  - Datatype: string

- **Order Date**

  - Description: Date of the Order

  - Datatype: Date

- **Order ID**

  - Description: Order Id of the Varieties of Products

  - Datatype: string

- **Ship Date**

  - Description: Date when the product is dispatched

  - Datatype: date

- **Unit Sold**

  - Description: Number of Units sold per product

  - Datatype: Int

- **Unit Price**

  - Description: Unit Price of the Product

  - Datatype: Int

- **Unit Cost**

  - Description: Unit cost of the Product

- Datatype: string

- **Total Revenue**

  - Description: Amount Incurred after selling different Products

  - Datatype: Decimal

- **Total Cost**

  - Description: Total Cost Incurred by the company for making Products

  - Datatype: Decimal

- **Total Profit**

  - Description: Profit Earned by the Company after subtracting all the expenses from the revenue

  - Datatype: Decimal

## 3.4 Data Loading in Power BI Query Editor

Power Query in Power BI Query Editor is a powerful tool for data loading and preparation, and it exhibits the following characteristics when importing and reshaping data:

1. **Multiple Rows and Columns**: Power Query can handle data tables with multiple rows and columns. Whether your dataset is small or large, Power Query can efficiently process and manipulate it.

2. **Row-Level Data**: Each row in the imported data represents a sample of data or a specific record. Power Query allows you to work with individual rows, making it easy to filter, transform, or aggregate data at the row level.

3. **Column-Based Variables**: Each column in the dataset represents a different variable or attribute that describes the samples (rows). You can perform various operations on these columns, such as renaming, reordering, removing, or creating new calculated columns to enhance your data analysis.

4. **Mixed Data Types:** Power Query is flexible when it comes to data types. It can handle columns with different data types such as numbers, strings, dates, Booleans, and more. You can also change data types, format dates, and handle null values to ensure that your data is consistent and ready for analysis.

5. **Data Transformation**: Power Query provides a wide range of data transformation capabilities. You can perform tasks like filtering rows, sorting data, merging tables, pivoting and unpivoting data, aggregating data, and applying custom transformations using the M or Power Query Formula Language. This flexibility allows you to prepare your data for meaningful visualizations and insights in Power BI.

In summary, Power Query in Power BI's Query Editor is a versatile tool that enables users to load and prepare data from various sources, handling datasets with multiple rows and columns, diverse data types, and offering extensive data transformation options to shape the data for analysis and visualization.

| | Region | Country | Item Type | Sales Channel | Order Priority | Order Date | Order ID | Ship Date | Units Sold |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Australia and Oceania | Tuvalu | Baby Food | Offline | H | 28-05-2010 | 669165933 | 27-06-2010 | |
| 2 | Central America and the Caribbean | Grenada | Cereal | Online | C | 22-08-2012 | 963881480 | 15-09-2012 | |
| 3 | Europe | Russia | Office Supplies | Offline | L | 02-05-2014 | 341417157 | 08-05-2014 | |
| 4 | Sub-Saharan Africa | Sao Tome and Principe | Fruits | Online | C | 20-06-2014 | 514321792 | 05-07-2014 | |
| 5 | Sub-Saharan Africa | Rwanda | Office Supplies | Offline | L | 01-02-2013 | 115456712 | 06-02-2013 | |
| 6 | Australia and Oceania | Solomon Islands | Baby Food | Online | C | 04-02-2015 | 547995746 | 21-02-2015 | |
| 7 | Sub-Saharan Africa | Angola | Household | Offline | M | 23-04-2011 | 135425221 | 27-04-2011 | |
| 8 | Sub-Saharan Africa | Burkina Faso | Vegetables | Online | H | 17-07-2012 | 871543967 | 27-07-2012 | |
| 9 | Sub-Saharan Africa | Republic of the Congo | Personal Care | Offline | M | 14-07-2015 | 770463311 | 25-08-2015 | |
| 10 | Sub-Saharan Africa | Senegal | Cereal | Online | H | 18-04-2014 | 616607081 | 30-05-2014 | |
| 11 | Asia | Kyrgyzstan | Vegetables | Online | H | 24-06-2011 | 814711606 | 12-07-2011 | |
| 12 | Sub-Saharan Africa | Cape Verde | Clothes | Offline | H | 02-08-2014 | 939825713 | 19-08-2014 | |
| 13 | Asia | Bangladesh | Clothes | Online | L | 13-01-2017 | 187310731 | 01-03-2017 | |
| 14 | Central America and the Caribbean | Honduras | Household | Offline | H | 08-02-2017 | 522840487 | 13-02-2017 | |
| 15 | Asia | Mongolia | Personal Care | Offline | C | 19-02-2014 | 832401311 | 23-02-2014 | |
| 16 | Europe | Bulgaria | Clothes | Online | M | 23-04-2012 | 972292029 | 03-06-2012 | |
| 17 | Asia | Sri Lanka | Cosmetics | Offline | M | 19-11-2016 | 419123971 | 18-12-2016 | |
| 18 | Sub-Saharan Africa | Cameroon | Beverages | Offline | C | 01-04-2015 | 519820964 | 18-04-2015 | |
| 19 | Asia | Turkmenistan | Household | Offline | L | 30-12-2010 | 441619336 | 20-01-2011 | |
| 20 | Australia and Oceania | East Timor | Meat | Online | L | 31-07-2012 | 322067916 | 11-09-2012 | |
| 21 | Europe | Norway | Baby Food | Online | L | 14-05-2014 | 819028031 | 28-06-2014 | |
| 22 | Europe | Portugal | Baby Food | Online | H | 31-07-2015 | 860673511 | 03-09-2015 | |
| 23 | Central America and the Caribbean | Honduras | Snacks | Online | L | 30-06-2016 | 795490682 | 26-07-2016 | |
| 24 | Australia and Oceania | New Zealand | Fruits | Online | H | 08-09-2014 | 142278373 | 04-10-2014 | |
| 25 | Europe | Moldova | Personal Care | Online | L | 07-05-2016 | 740147912 | 10-05-2016 | |
| 26 | Europe | France | Cosmetics | Online | H | 22-05-2017 | 898523128 | 05-06-2017 | |
| 27 | Australia and Oceania | Kiribati | Fruits | Online | M | 13-10-2014 | 347140347 | 10-11-2014 | |
| 28 | Sub-Saharan Africa | Mali | Fruits | Online | L | 07-05-2010 | 686048400 | 10-05-2010 | |
| 29 | Europe | Norway | Beverages | Offline | C | 18-07-2014 | 435608613 | 30-07-2014 | |
| 30 | Sub-Saharan Africa | The Gambia | Household | Offline | L | 26-05-2012 | 886494815 | 09-06-2012 | |
| 31 | Europe | Switzerland | Cosmetics | Offline | M | 17-09-2012 | 249693334 | 20-10-2012 | |
| 32 | Sub-Saharan Africa | South Sudan | Personal Care | Offline | C | 29-12-2013 | 406502997 | 28-01-2014 | |
| 33 | Australia and Oceania | Australia | Office Supplies | Online | C | 27-10-2015 | 158535134 | 25-11-2015 | |
| 34 | Asia | Myanmar | Household | Offline | H | 16-01-2015 | 177713572 | 01-03-2015 | |
| 35 | Sub-Saharan Africa | Djibouti | Snacks | Online | M | 25-02-2017 | 756274640 | 25-02-2017 | |
| 36 | Central America and the Caribbean | Costa Rica | Personal Care | Offline | L | 08-05-2017 | 456767165 | 21-05-2017 | |
| 37 | Middle East and North Africa | Syria | Fruits | Online | L | 22-11-2011 | 162052476 | 03-12-2011 | |

## 3.5 Data to Insights through Visualizations and Data Analysis
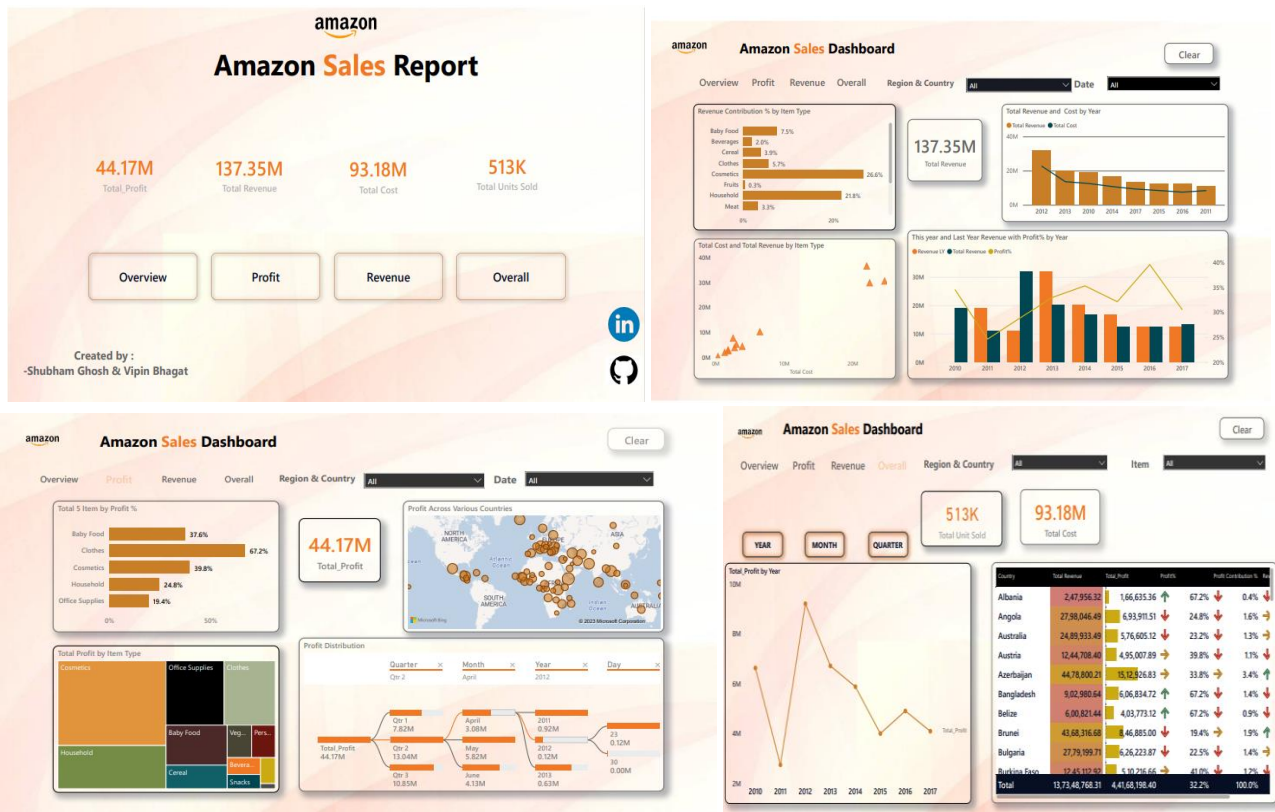


## 4. Deployment in Power BI Service