Analysing Amazon Sales data

Low-Level Design (LLD)

Abdul Atif

Soumyadip Tikader

Version: 1.0

Date:28-01-2023

Project Title	Analyzing Amazon Sales data
Technologies	Business Intelligence
Domain	E-commerce
Project Difficulties Level	Advanced

Document Version Control

Date Issued	Version	Description	Author
27-01-2023	1.0	First Version on complete LLD	Abdul Atif
28-01-2023	1.1	Revised Version	Soumyadip Tikader

Abstract

In the world of rising new technology and innovation, the Amazon sales is advancing with the role of Data Science and Analytics. Data analysis can help them to understand their business in a quite different manner and helps to improve the quality of the service by identifying the weak areas of the business. This study demonstrates how different analysis help to make better business decisions and help analyse customer trends and satisfaction, which can lead to new and better products and services. Different analyses were performed such as Exploratory Data Analysis and Descriptive Analysis on a variety of use cases to get the key insights from this data based on which business decisions will be taken.

This dataset provides a huge amount of information on the sales of all over the world. Based on the Information the ultimate goal would be to predict the best sales product for common people and find important insights highlighting key indicators and metrics that influence customer choice.

Table of Contents

1. 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 Data Analysis project. LLD 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. 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.

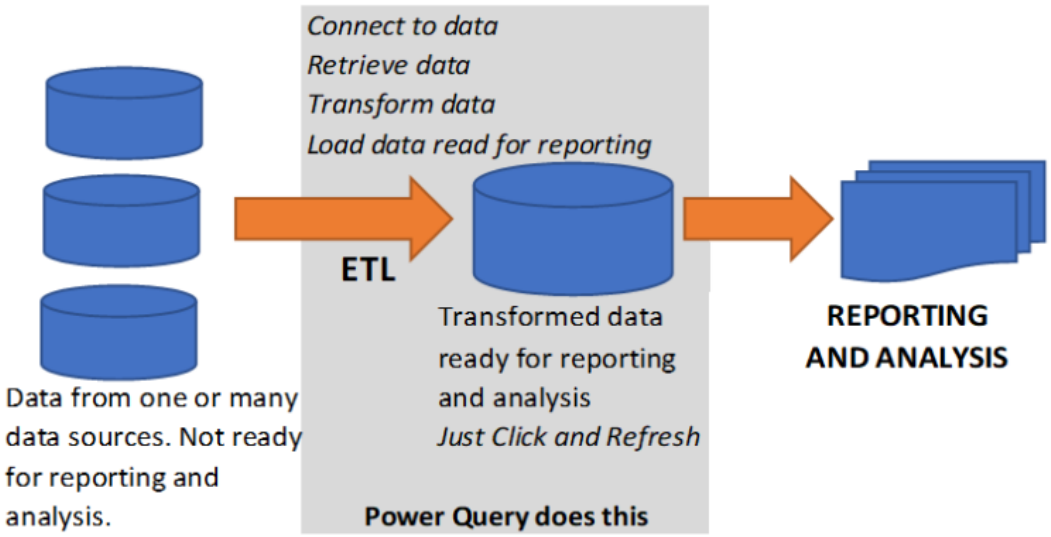
The main objective of the project is to analyse the various aspects with different use cases which covers many aspects of Amazon sales al over the world. It helps in not only understanding the meaningful relationships between attributes but also allows us to do our own research and come up with our findings.

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, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

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

2. Architecture



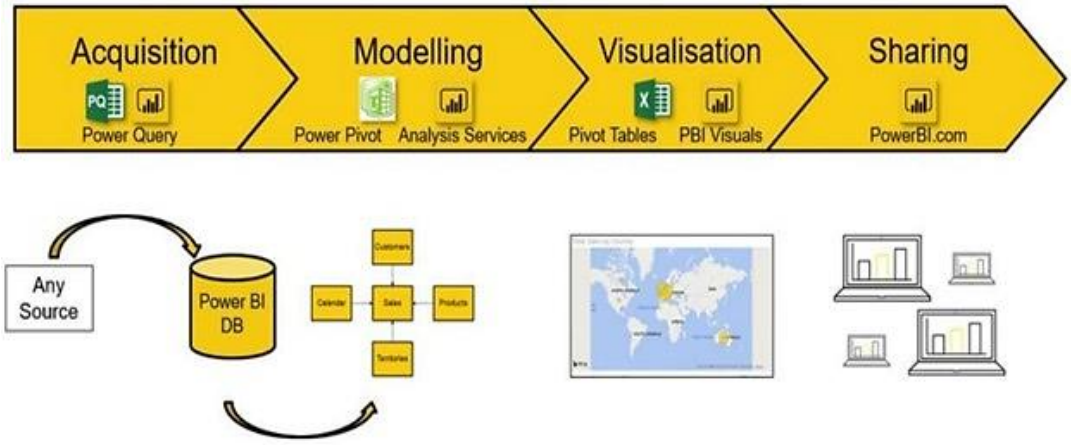
ETL (extract, transform and load) in Power BI use the preparation of data sets for analysis by removing irregularities in the data. It also involves data visualization to draw meaningful patterns and insights.

Based on the results of ETL, companies also make business decisions, which can have repercussions later.

- ✓ If ETL is not done properly then it can damage the business a lot in many ways such as loss of clients which whom we are working, the decision making will go completely wrong, and many more issues.
- ✓ If done well, it may improve the efficacy of everything we do next.

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 (Comma Separated Values) format. MS Excel is used to load the data.

This dataset is publicly available for research purposes.

Title: Amazon Sales Dataset.csv

Source:

3.2 Data Overview

- The dataset is of size 19 KB
- It includes a single file in ".csv" format.
- Number of rows/records: 100
- Number of attributes: 15

3.3 Data Description

The following attributes describes the dataset.

- **Region:**

  Region describe the Continents. Where all the sales has happen.

- **Item type:**

  Item type say what type of our been sold.
  Ex: Baby food, Fruits, Clothes etc.

- **Sales Channel:**

  There are only two Channel where amazon sales happen Offline , Online.

- **Order data and Ship date:**

  The order data and shipping data we can say that how many days did it took to ship the order.

- **Unit sold :**

  This parameter gives us the count of number unit sold based of item type.

- **Unit price and Unit cost:**

  Unit price is price of good sold including all taxes, production cost and profit.
  Unit cost is actual price to prepare the goods or services. (Excluding tax and profit).

- **Total Revenue and Total cost:**

  Total Revenue say how much income we got based of the product type.

Total cost refers to the cost of preparing the number of product.

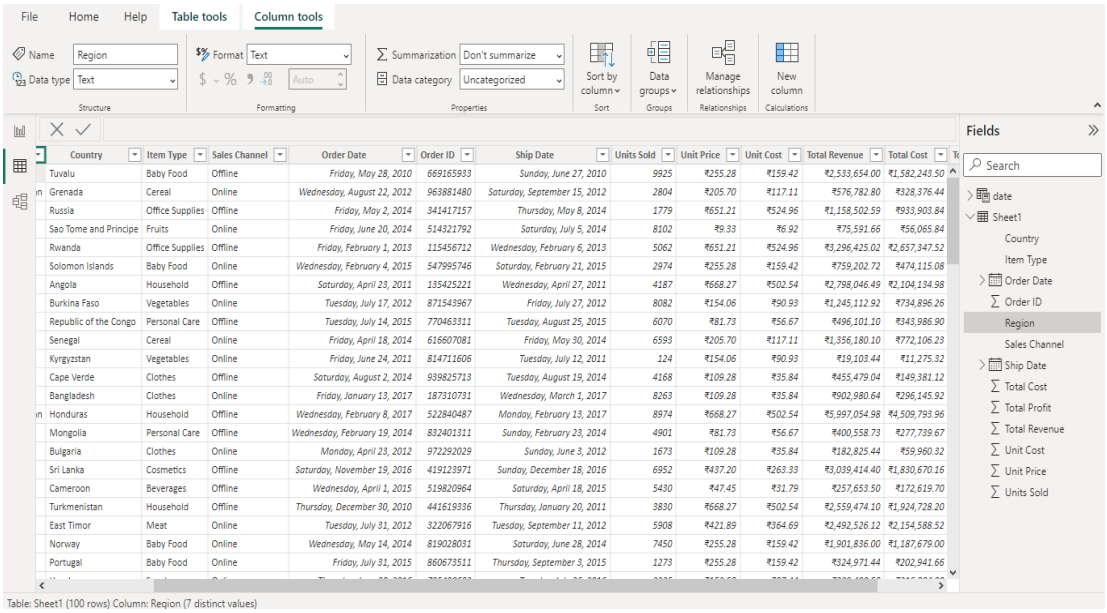
- **Profit:**

By subtracting the Total revenue and total cost the profit column.

3.4 Data Loading in Power BI Query Editor

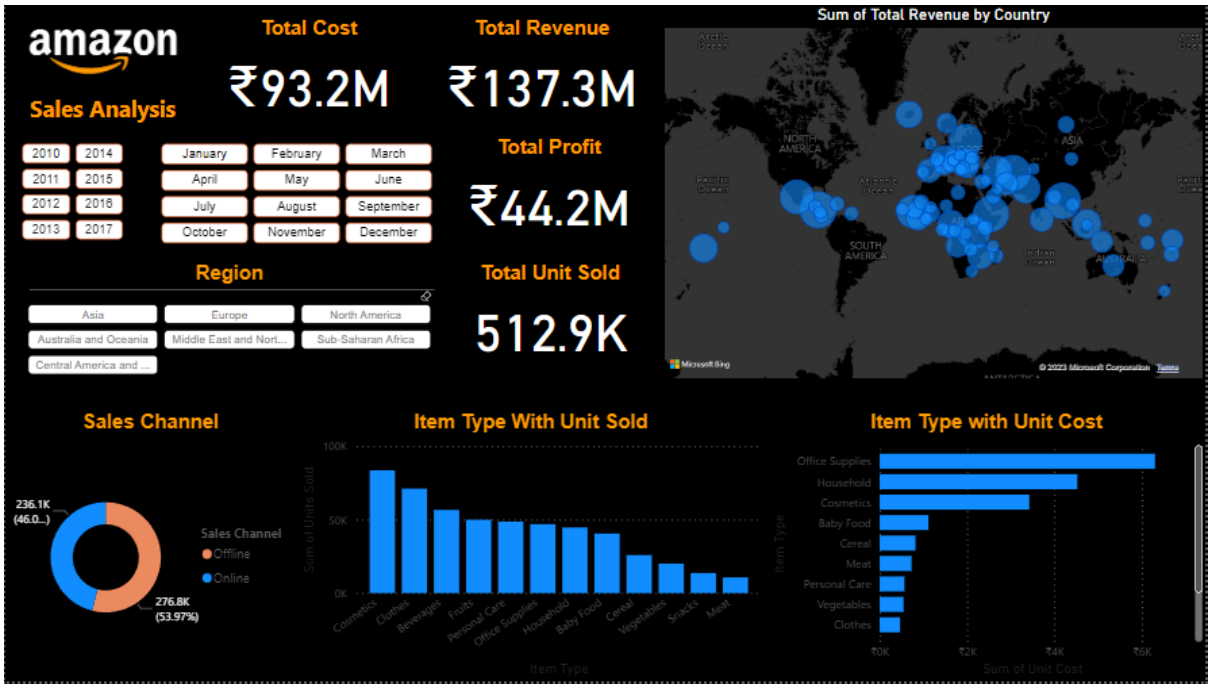
Power Query is the data connectivity and data preparation technology that enables end users to seamlessly import and reshape data from within a wide range of Microsoft products, including Excel, Power BI, Analysis Services, data verse, and more with the following characteristics.

- There can be multiple rows and columns in the data.
- Each row represents a sample of data,
- Each column contains a different variable that describes the samples (rows).
- The data in every column can be a different type of data like numbers, strings, dates, Boolean etc.



Country	Item Type	Sales Channel	Order Date	Order ID	Ship Date	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost
Tuvalu	Baby Food	Offline	Friday, May 28, 2010	669165933	Sunday, June 27, 2010	9925	₹255.28	₹159.42	₹2,533,654.00	₹1,582,243.50
Grenada	Cereal	Online	Wednesday, August 22, 2012	963881480	Saturday, September 15, 2012	2804	₹205.70	₹117.11	₹576,782.80	₹328,376.44
Russia	Office Supplies	Offline	Friday, May 2, 2014	341417157	Thursday, May 8, 2014	1779	₹651.21	₹524.96	₹1,158,502.59	₹933,903.84
Sao Tome and Principe	Fruits	Online	Friday, June 20, 2014	514321792	Saturday, July 5, 2014	8102	₹9.33	₹6.92	₹75,591.66	₹56,065.84
Rwanda	Office Supplies	Offline	Friday, February 1, 2013	115456712	Wednesday, February 6, 2013	5062	₹651.21	₹524.96	₹3,296,425.02	₹2,657,347.52
Solomon Islands	Baby Food	Online	Wednesday, February 4, 2015	547995746	Saturday, February 21, 2015	2974	₹255.28	₹159.42	₹759,202.72	₹474,115.08
Angola	Household	Offline	Saturday, April 23, 2011	135425221	Wednesday, April 27, 2011	4187	₹668.27	₹502.54	₹2,798,046.49	₹2,104,134.98
Burkina Faso	Vegetables	Online	Tuesday, July 17, 2012	871543967	Friday, July 27, 2012	8082	₹154.06	₹90.93	₹1,245,112.92	₹734,896.26
Republic of the Congo	Personal Care	Offline	Tuesday, July 14, 2015	770463311	Tuesday, August 25, 2015	6070	₹81.73	₹56.67	₹496,101.10	₹343,986.90
Senegal	Cereal	Online	Friday, April 18, 2014	616607081	Friday, May 30, 2014	6593	₹205.70	₹117.11	₹1,356,180.10	₹772,106.23
Kyrgyzstan	Vegetables	Online	Friday, June 24, 2011	814711606	Tuesday, July 12, 2011	124	₹154.06	₹90.93	₹19,103.44	₹11,275.32
Cape Verde	Clothes	Offline	Saturday, August 2, 2014	939825713	Tuesday, August 19, 2014	4168	₹109.28	₹35.84	₹455,479.04	₹149,381.12
Bangladesh	Clothes	Online	Friday, January 13, 2017	187310731	Wednesday, March 1, 2017	8263	₹109.28	₹35.84	₹902,980.64	₹296,145.92
Honduras	Household	Offline	Wednesday, February 8, 2017	522840487	Monday, February 13, 2017	8974	₹668.27	₹502.54	₹5,997,054.98	₹4,509,793.96
Mongolia	Personal Care	Offline	Wednesday, February 19, 2014	832401311	Sunday, February 23, 2014	4901	₹81.73	₹56.67	₹400,558.73	₹277,739.67
Bulgaria	Clothes	Online	Monday, April 23, 2012	972292029	Sunday, June 3, 2012	1673	₹109.28	₹35.84	₹182,825.44	₹59,960.32
Sri Lanka	Cosmetics	Offline	Saturday, November 19, 2016	419123971	Sunday, December 18, 2016	6952	₹437.20	₹263.33	₹3,039,414.40	₹1,830,670.16
Cameroon	Beverages	Offline	Wednesday, April 1, 2015	519820964	Saturday, April 18, 2015	5430	₹47.45	₹31.79	₹257,653.50	₹172,619.70
Turkmenistan	Household	Offline	Thursday, December 30, 2010	441619336	Thursday, January 20, 2011	3830	₹668.27	₹502.54	₹2,559,474.10	₹1,924,728.20
East Timor	Meat	Online	Tuesday, July 31, 2012	322067916	Tuesday, September 11, 2012	5908	₹421.89	₹364.69	₹2,492,526.12	₹2,154,588.52
Norway	Baby Food	Online	Wednesday, May 14, 2014	819028031	Saturday, June 28, 2014	7450	₹255.28	₹159.42	₹1,901,836.00	₹1,187,679.00
Portugal	Baby Food	Online	Friday, July 31, 2015	860673511	Thursday, September 3, 2015	1273	₹255.28	₹159.42	₹324,971.44	₹202,941.66

3.5 Data to Insights through Visualizations and Data Analysis



4. Deployment in Power BI Service