Low Level Design

Amazon Sales Data Analysis

Written By	Saurav Rana
Document Version	0.2
Last Revised Date	02/09/2022

DOCUMENT CONTROL

Change Record:

VERSION	DATE	AUTHOR	COMMENTS
0.1	29- Aug - 2022	Saurav Rana	Introduction and architecture defined
0.2	02 -Sep - 2022	Saurav Rana	Architecture & Architecture description appended and updated.

Reviews:

VERSION	DATE	REVIEWER	COMMENTS

Approval Status:

VERSION	REVIEW DATE	REVIEWED BY		APPROVED BY	COMMENTS

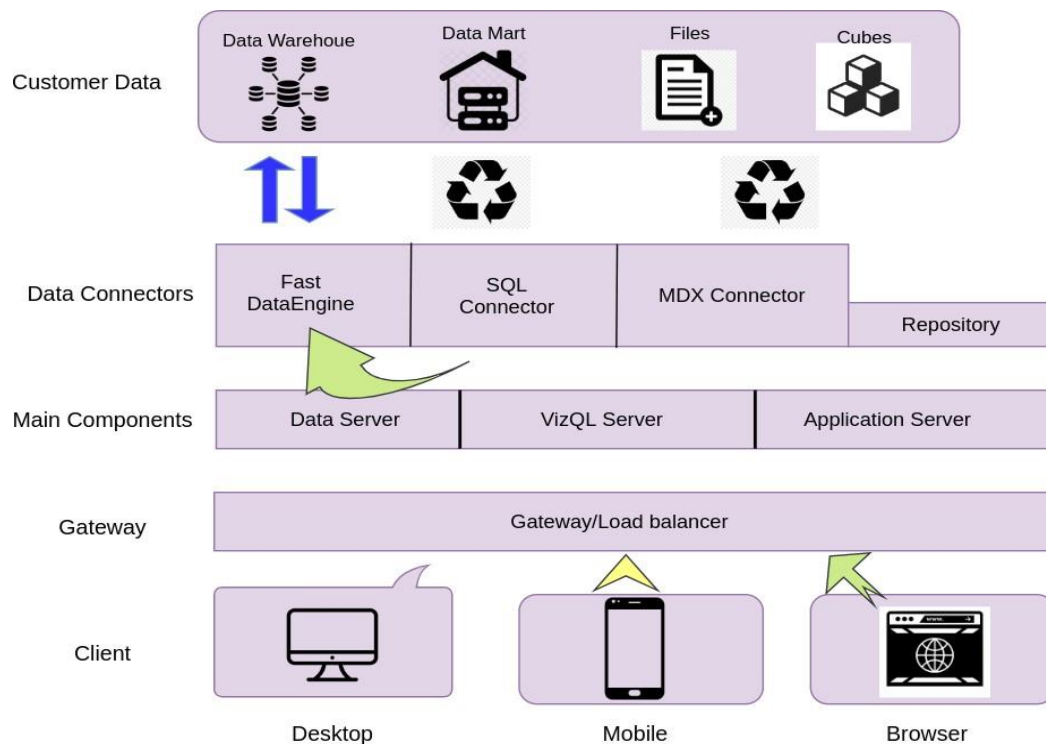## Contents

1. Introduction

What is Low-Level design document?

➢ The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the House Price Prediction dashboard. LDD 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.

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.
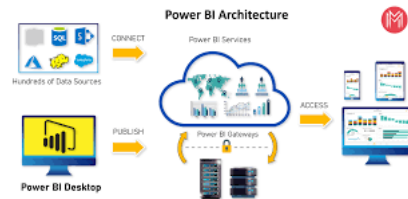
2. Architecture

Power BI Server Architecture

Power BI has a highly scalable, n-tier client-server architecture that serves mobile clients, web clients and desktop-installed software. Power BI Server architecture supports fast and flexible deployments.

The following diagram shows Power BI Server's architecture:

Power BI Server is internally managed by the multiple server

1. Gateway/Load Balancer

It acts as an Entry gate to the Power BI Server and also balances the load to the Server if multiple Processes are configured.

2) Application Server:-

Application Server processes (wgserver.exe) handle browsing and permissions for the Power BI Server web and mobile interfaces. When a user opens a view in a client device, that user starts a session on Power BI Server. This means that an Application Server thread starts and checks the permissions for that user and that view.

3) Repository:-

Power BI Server Repository is a PostgreSQL database that stores server data. This data includes information about Power BI Server users, groups and group assignments, permissions, projects, data sources, and extract metadata and refresh information.

4) VIZQL Server:-

Once a view is opened, the client sends a request to the VizQL process (vizqlserver.exe). The VizQL process then sends queries directly to the data source, returning a result set that is rendered as images and presented to the user. Each VizQL Server has its own cache that can be shared across multiple users

5) Data Engine:-

It Stores data extracts and answers queries.

6) Backgrounder:-

The backgrounder Executes server tasks which includes refreshes scheduled extracts, tasks initiated from tabcmd and manages other background tasks.

7) Data Server:-

Data Server Manages connections to Power Bi Server data sources

3. Architecture Description

Data Description

The Dataset contains the sales trend of Amazon :

1. It shows the profit of products according to countries.

2. It shows the sales according to different aspects.

3. Sales management is most useful system for growth of company.

Web Scrapping

Web scraping is a technique to automatically extract content and data from websites using bots. It is also known as web data extraction or web harvesting. Web scrapping is made simple now days, many tools are used for web scrapping. Some of python libraries used for web scrapping are

Beautiful Soup, Scrapy, Selenium, etc.

Data Transformation

In the Transformation Process, we will convert our original datasets with other necessary attributes format. And will merge it with the Scrapped dataset.

Data Insertion into Database

a. Database Creation and connection - Create a database with name passed. If the database is already created, open the connection to the database.

b. Table creation in the database.

c. Insertion of files in the table

Step 2: Configuring Data Source

The data source page loads up after configuring the Power BI connector and successfully signing in. This is how the page looks like:

Region	Country	Item Type	Sales Channel	Order Priority	Order ID	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost	Total Profit
Australia and Oceania	Tuvalu	Baby Food	Offline	H	669165933	9925	255.28	159.42	2533654	1582243.5	951410.5
Central America and the Caribbean	Grenada	Cereal	Online	C	963881480	2804	205.7	117.11	576782.8	328376.44	248406.36
Europe	Russia	Office Supplies	Offline	L	341417157	1779	651.21	524.96	1158502.59	933903.84	224598.75
Sub-Saharan Africa	Sao Tome and Principe	Fruits	Online	C	514321792	8102	9.33	6.92	75591.66	56065.84	19525.82
Sub-Saharan Africa	Rwanda	Office Supplies	Offline	L	115456712	5062	651.21	524.96	3296425.02	2657347.52	639077.5
Australia and Oceania	Solomon Islands	Baby Food	Online	C	547995746	2974	255.28	159.42	759202.72	474115.08	285087.64
Sub-Saharan Africa	Angola	Household	Offline	M	135425221	4187	668.27	502.54	2798046.49	2104134.98	693911.51
Sub-Saharan Africa	Burkina Faso	Vegetables	Online	H	871543967	8082	154.06	90.93	1245112.92	734896.26	510216.66
Sub-Saharan Africa	Republic of the Congo	Personal Care	Offline	M	770463311	6070	81.73	56.67	496101.1	343986.9	152114.2
Sub-Saharan Africa	Senegal	Cereal	Online	H	616607081	6593	205.7	117.11	1356180.1	772106.23	584073.87
Asia	Kyrgyzstan	Vegetables	Online	H	814711606	124	154.06	90.93	19103.44	11275.32	7828.12
Sub-Saharan Africa	Cape Verde	Clothes	Offline	H	939825713	4168	109.28	35.84	455479.04	149381.12	306097.92
Asia	Bangladesh	Clothes	Online	L	187310731	8263	109.28	35.84	902980.64	296145.92	606834.72
Central America and the Caribbean	Honduras	Household	Offline	H	522840487	8974	668.27	502.54	5997054.98	4509793.96	1487261.02
Asia	Mongolia	Personal Care	Online	C	832401311	4901	81.73	56.67	400558.73	277739.67	122819.06
Europe	Bulgaria	Clothes	Online	M	972292029	1673	109.28	35.84	182825.44	59960.32	122865.12
Asia	Sri Lanka	Cosmetics	Offline	M	419123971	6952	437.2	263.33	3039414.4	1830670.16	1208744.24
Sub-Saharan Africa	Cameroon	Beverages	Offline	C	519820964	5430	47.45	31.79	257653.5	172619.7	85033.8
Asia	Turkmenistan	Household	Offline	L	441619336	3830	668.27	502.54	2559474.1	1924728.2	634745.9
Australia and Oceania	East Timor	Meat	Online	L	322067916	5908	421.89	364.69	2492526.12	2154588.52	337937.6
Europe	Norway	Baby Food	Online	L	819028031	7450	255.28	159.42	1901836	1187679	714157
Europe	Portugal	Baby Food	Online	H	860673511	1273	255.28	159.42	324971.44	202941.66	122029.78
Central America and the Caribbean	Honduras	Snacks	Online	L	795490682	2225	152.58	97.44	339490.5	216804	122686.5
Australia and Oceania	New Zealand	Fruits	Online	H	142278373	2187	9.33	6.92	20404.71	15134.04	5270.67
Europe	Moldova	Personal Care	Online	L	740147912	5070	81.73	56.67	414371.1	287316.9	127054.2
Europe	France	Cosmetics	Online	H	898523128	1815	437.2	263.33	793518	477943.95	315574.05
Australia and Oceania	Kiribati	Fruits	Online	M	347140347	5398	9.33	6.92	50363.34	37354.16	13009.18
Sub-Saharan Africa	Mali	Fruits	Online	L	686048400	5822	9.33	6.92	54319.26	40288.24	14031.02
Europe	Norway	Beverages	Offline	C	435608613	5124	47.45	31.79	243133.8	162891.96	80241.84

Select the data source name option and give a unique name to the database you are using. It's considered a good practice to have a unique name as it makes it much easier for users to identify the database from which data being fetched.

To select the desired schema, you can use the schema drop-down list from the column on the left. You can also perform a text-based search to find the desired option. Now similarly find and select the desired table and drag it onto the canvas.

This is how you can connect SQL Server with Power BI. Now click on the sheets tab to begin the analysis. Custom SQL features can be used to focus on specific SQL statements, rather than querying the entire database Click on the Custom SQL option from the panel on the left. A new dialogue box will now open up, where you can provide the query you want to execute.
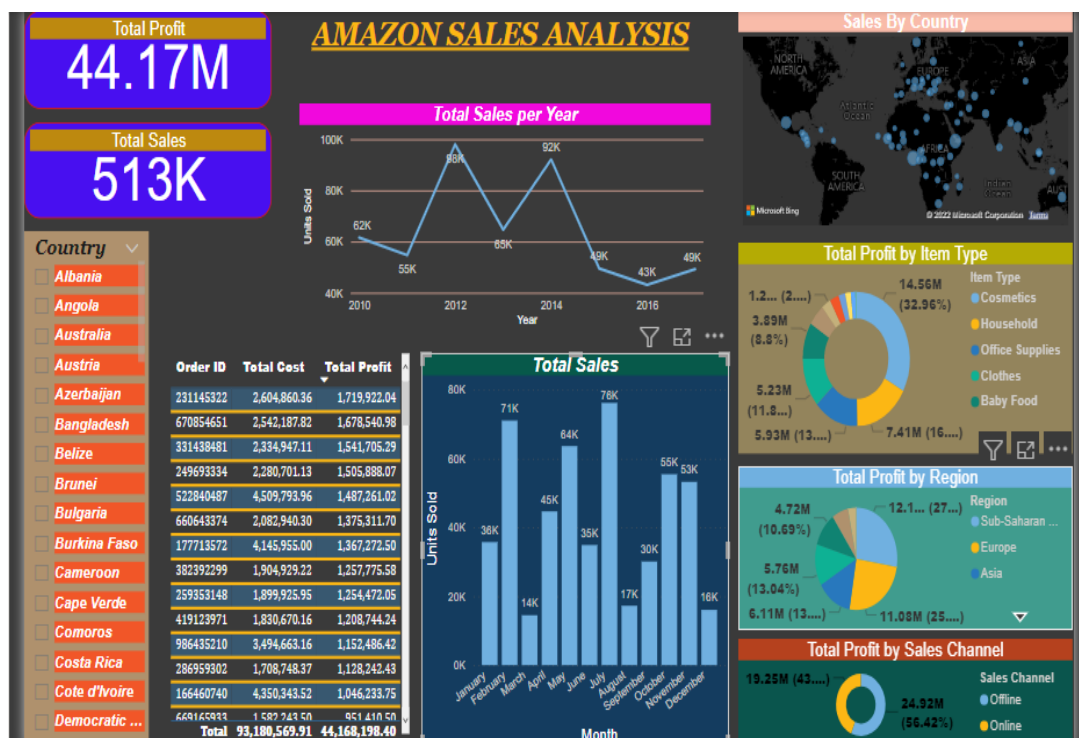
Export Data from Database

Data Export from Database - The data in a stored database is exported as a CSV file to be used for

Data Pre-processing.

Deployment.

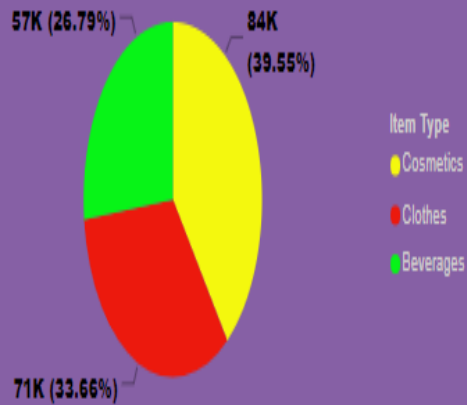Once you've completed your dashboard, follow these steps:-**Server, Power Bi Public, Save To Power BI Public As**

You may be prompted to log into your Power BI Public profile first if this is your first time publishing.
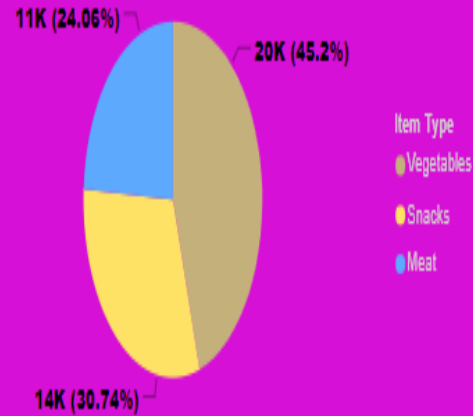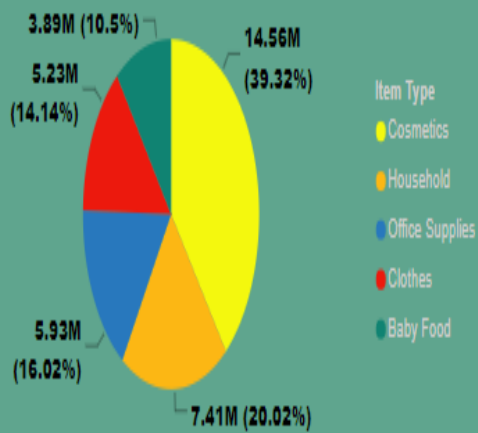
Product Performance Analysis
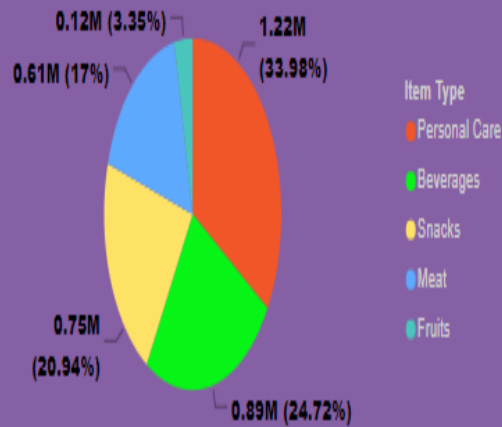
Top 3 Selling Products

57K (26.79%)
84K (39.55%)
71K (33.66%)

Item Type
- Cosmetics
- Clothes
- Beverages

Least 3 Selling Products

11K (24.06%)
20K (45.2%)
14K (30.74%)

Item Type
- Vegetables
- Snacks
- Meat

Top 5 Profitable Products

3.89M (10.5%)
14.56M (39.32%)
5.23M (14.14%)
5.93M (16.02%)
7.41M (20.02%)

Item Type
- Cosmetics
- Household
- Office Supplies
- Clothes
- Baby Food

least 5 Profitable Products

0.12M (3.35%)
1.22M (33.98%)
0.61M (17%)
0.75M (20.94%)
0.89M (24.72%)

Item Type
- Personal Care
- Beverages
- Snacks
- Meat
- Fruits

4. Unit Test Cases

Test case	Results
Click on donut chart charts	It shows the profit by different aspects like region, order type etc.
Click on line chart	Shows total sales per year
Bar chart	Shows total sales per month