

Low Level Design

Amazon Sales Analysis

Written By	Ankita Gupta, Sanjeev Kumar
Document Version	0.3
Last Revised Date	30/09/2023

DOCUMENT CONTROL

Change Record:

VERSION	DATE	AUTHOR	COMMENTS
0.1	25- July - 2023	Ankita Gupta & Sanjeev Kumar	Introduction and architecture defined
0.2	20 - August - 2023	Ankita Gupta & Sanjeev Kumar	Architecture & Architecture description appended and updated.

Reviews:

VERSION	DATE	REVIEWER	COMMENTS
0.3	30- September - 2021	Ankita Gupta & Sanjeev Kumar	Final Revision

Approval Status:

VERSION	REVIEW DATE	REVIEWED BY		APPROVED BY	COMMENTS

Contents

1.	Introduction	04
1.1	What is Low-Level Design Document?	04
1.2	Scope	04
2.	Architecture	05
3.	Architecture Description	08
3.1	Data Description	08
3.2	Web Scrapping	08
3.3	Data Transformation	08
3.4	Data insertion into database	08
3.5	Connection with SQL server	08
3.5	Export Data from database	12
3.6	Deployment	12
4.	Uni test cases.....	15
	t	

1. Introduction

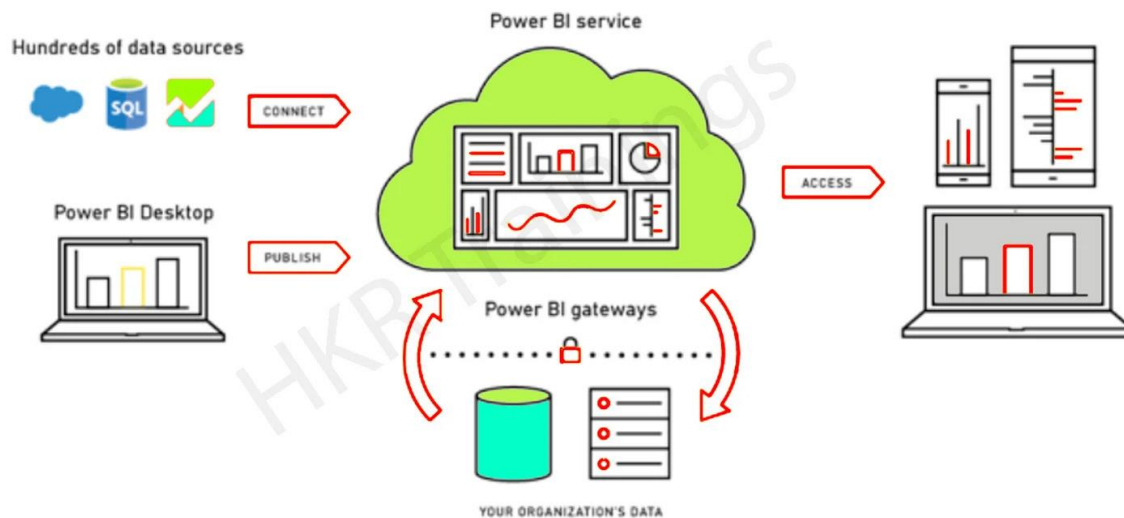
1.1 What is Low-Level design document?

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.

1.2 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 features a highly scalable n-tier client-server architecture that caters to mobile clients, web clients, and desktop-installed software. Power BI Server architecture is designed to facilitate swift and adaptable deployments

The following diagram shows Power BI Server's architecture:

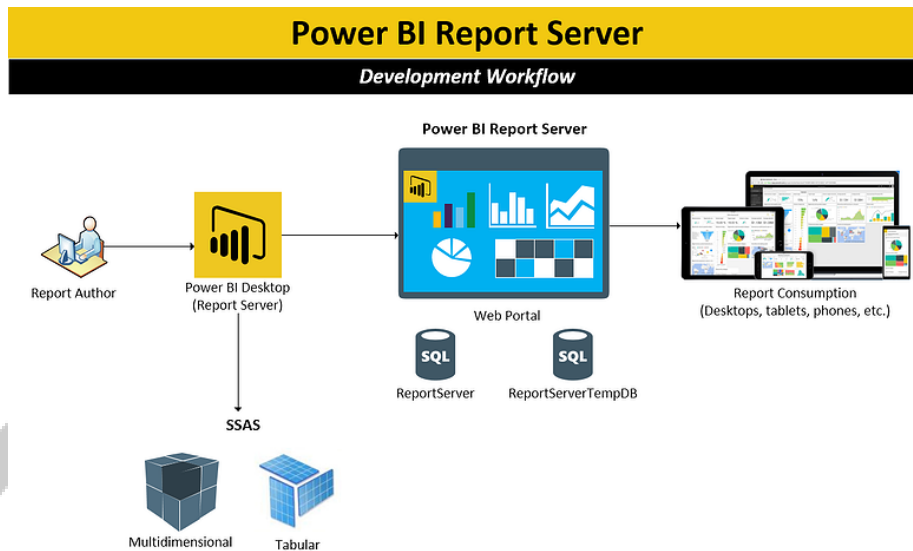


Tableau Server is internally managed by the multiple server

1. Gateway/Load Balancer

It serves as an entry point to the Power BI Server, ensuring efficient access.

Additionally, it balances the load to the server when multiple processes are configured

2) Application Server:-

Application Server processes (e.g., wsgserver.exe) manage browsing and permissions for the Power BI Server web and mobile interfaces. When a user initiates a session on Power BI Server by accessing a view from a client device, an Application Server thread is triggered. The thread is responsible for verifying the permissions of the user and the requested view

3) Repository:-

The Power BI Server Repository is a PostgreSQL database designed to store essential server-related data. This data encompasses details concerning Power BI Server users, groups, group memberships, permissions, projects, data sources, as well as extract metadata and refresh information.

4) VIZQL Server:-

When a view is accessed, the client initiates a request to the VizQL process (vizqlserver.exe). The VizQL process subsequently sends queries directly to the data source and retrieves a result set. This result set is then transformed into images and displayed to the user. Each VizQL Server maintains its own cache, which can be utilized by multiple users for improved performance and efficiency.

5) Data Engine:-

The Data Engine serves as a repository for data extracts and handles query responses

6) Backgrounder:-

The Backgrounder executes server tasks, including scheduled extract refreshes and tasks initiated from external tools like tabcmd. It also manages various background tasks within the server environment.

7) Data Server:-

The Data Server is responsible for overseeing connections to Power BI Server data sources. It also preserves metadata originating from Power BI Desktop, which includes calculations, definitions, and groups.

3. Architecture Description

3.1. Data Description

The dataset comprises Amazon sales data spanning the years 2010 through 2017, providing comprehensive view of sales trends and patterns over this eight-year period.

1. Region: The geographic region where the sales data is recorded.
2. Country: The specific country within the region where the sales occurred.
3. Item Type: The category or type of product sold, such as "Baby Food" or "Cereal."
4. Sales Channel: The method through which the sales were made, either "Offline" or "Online."
5. Order Priority: The priority level assigned to the order, typically denoted by letters like "High" or "Low"
6. Order Date: The date when the order was placed.
7. Order ID: A unique identifier for each order.

8. Ship Date: The date when the products were shipped to the customer.
9. Units Sold: The quantity of units (products) sold in the order.
10. Unit Price: The price of a single unit of the product.
11. Unit Cost: The cost incurred for producing or purchasing a single unit of the product.
12. Total Revenue: The total income generated from the sale of units (Unit Price multiplied by Units Sold).
13. Total Cost: The total cost incurred to produce or purchase the sold units (Unit Cost multiplied by Units Sold).
14. Total Profit: The overall profit earned from the sale (Total Revenue minus Total Cost).

3.2. 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 BeautifulSoup, Scrapy, Selenium, etc.

3.3. 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.

3.4. 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

3.5 Make the SQL connection and set up the data source

Step 1: Configuring Power BI

To connect to PowerBI, start by launching the application on your workstation. Then, navigate to the

"Get Data" section and select "SQL Server" from the list of available data source options. This action will

This action will open a dialog box where you'll need to input the connection details for your SQL Server.

Be sure to provide the essential information about the server hosting your database. If you're looking to connect to a specific contained database, you can also specify its name during this process.

To establish a connection in PowerBI with a port other than the default, you'll need to specify port and server in the following format:

Example query: my_server:8051

There are two primary authentication methods available for signing in to the server. You can either utilize Windows authentication or opt for the username and password approach. The latter is particularly necessary when dealing with a password-protected server within a non-Kerberos environment.

3.6 Data Transformation

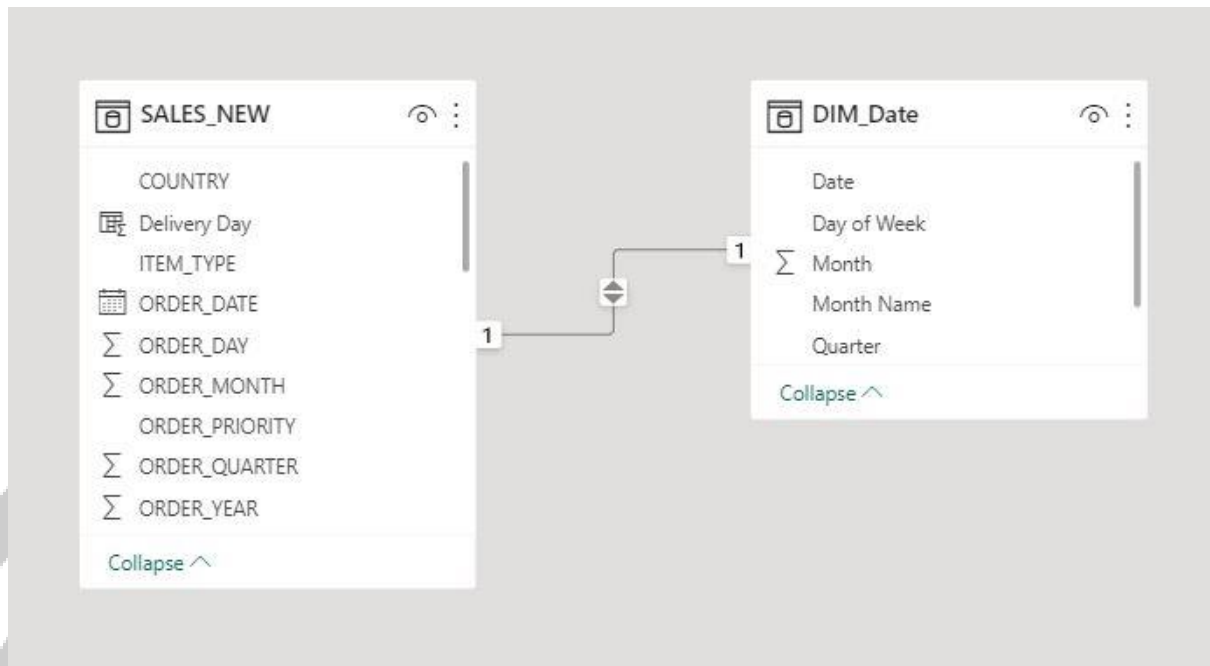
In the Transformation Process, we will convert our original datasets with other necessary attributes format and change the features according to the problem statement on Power BI ETL tool Power Query as the data is in excel format.

A new Dim_date table created for easy time intelligence analysis the features are:

1. Day of Week
2. Month name
3. Year
4. Year quarter
5. Month
6. Date

3.7 Data Modelling

After the data is transformed the data is modeled for visualizing and analysis.

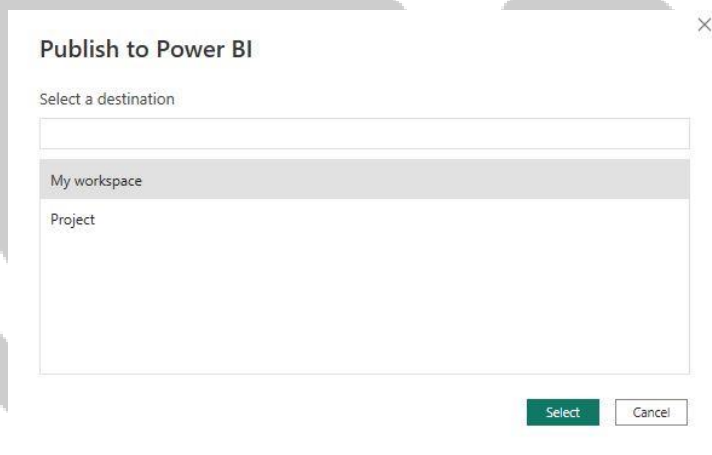


3.8 Deployment

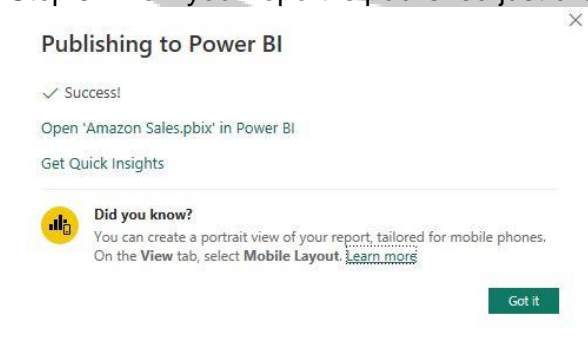
After completing Dashboard Follow the Steps to deploy the report

Step 1:- Click on **Publish** on the ribbon.

Step 2:- A box will pop just click on **my workplace** and **select**



Step 3:- Now your report is published just click on open to check your report.



Step 4:- Report is published now explore your report



3.9 Unit Test Cases

Test Case	Description
Slicer of Year, Country & Quarter	Slicer shows a drop down
Top 20 product slicer	Slicer show top 20 product and select works
Charts	All charts showing 0 error
Tooltips	Tooltips on various page working properly
Page buttons	Page buttons working properly