Low Level Design

GOOGLE PLAY STORE APP

Written By	Preeti Jain
Document Version	0.3
Last Revised Date	**22/07/23**

DOCUMENT CONTROL

Change Record:

VERSION	DATE	AUTHOR	COMMENTS
0.1	15-July - 2023	Preeti Jain	Introduction to Data Set and cleaning & Statical Analysis
0.2	20-July - 2023	Preeti Jain	Exporting Cleaned data for Dashboard building

Reviews:

VERSION	DATE	REVIEWER	COMMENTS
0.2	22-July-2023	Preeti Jain	Power Bi Buttons to be added

Approval Status:

VERSION	REVIEW DATE	REVIEWED BY		APPROVED BY	COMMENTS

Contents

1. Introduction

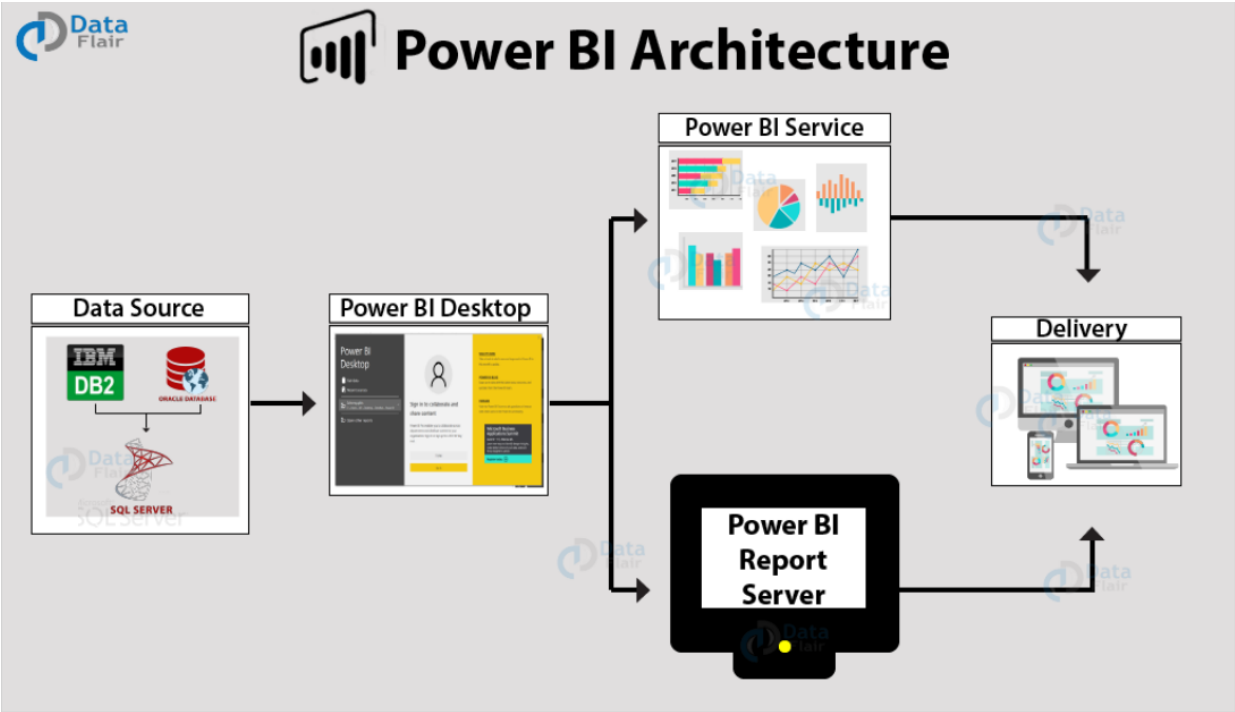
1.1 What is a Low-Level design document?

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.

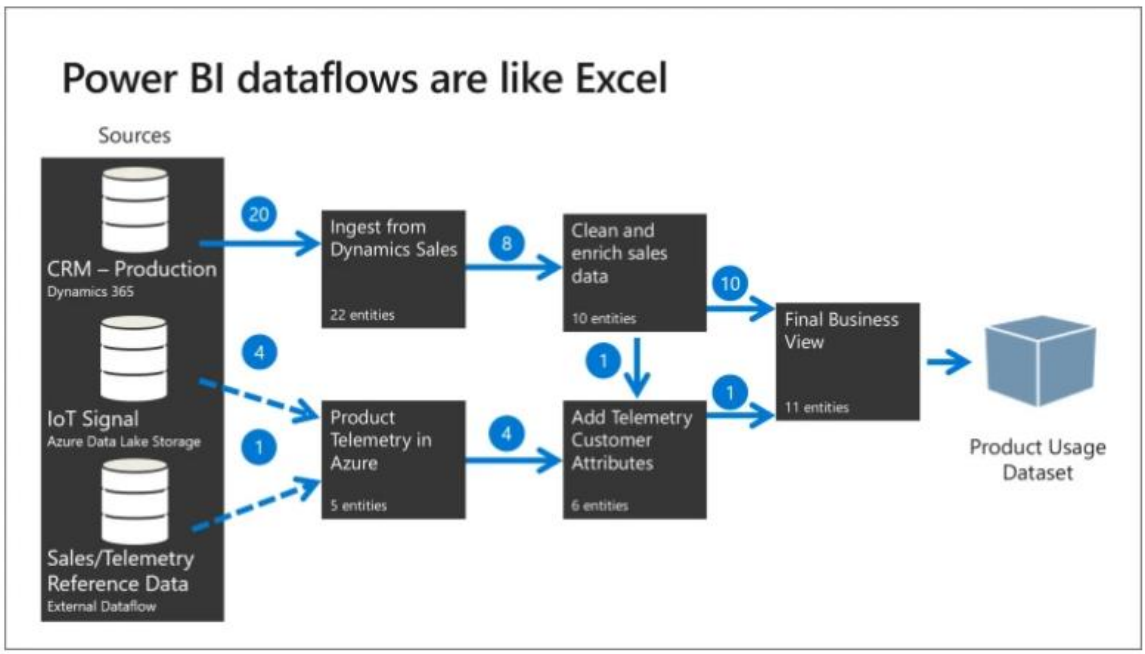
2. Architecture



PowerBi Server Architecture

Power BI is a business suite that includes several technologies that work together. To deliver outstanding business intelligence solutions, Microsoft Power BI technology consists of a group of components such as:

- Power Query (for data mash-up and transformation)
- Power BI Desktop (a companion development tool)
- Power BI Mobile (for Android, iOS, Windows phones)
- Power Pivot (for in-memory tabular data modeling)
- Power View (for viewing data visualizations)
- Power Map (for visualizing 3D geo-spatial data)
- Power Q&A (for natural language Q&A)



1. Data Sources

An important component of Power BI is its vast range of data sources. You can import data from files in your system, cloud-based online data sources or connect directly to live connections. If you import from data on-premise or online services there is a limit of 1 GB. Some commonly used data sources in Power BI are:

- Excel
- Text/CSV
- XML
- JSON
- Oracle Database
- IBM DB2 Database
- MySQL Database
- PostgreSQL Database
- Sybase Database
- Teradata Database
- SAP HANA Database
- SAP Business Warehouse server
- Amazon Redshift
- Impala
- Google BigQuery (Beta)
- Azure SQL Database
- Salesforce Reports
- Google Analytics
- Facebook
- GitHub

2. Power BI Desktop

Power BI Desktop is a client-side tool known as a companion development and authoring tool.

This desktop-based software is loaded with tools and functionalities to *connect to data sources, transform data, data modeling and creating reports.*

3. Power BI Service

Power BI Service is a web-based platform from which you can *share reports made on Power BI Desktop, collaborate with other users, and create dashboards.*

4. Power BI Report Server

The Power BI Report Server is similar to the Power BI Service. The only difference between these two is that Power BI Report Server is an on-premise platform. It is used by organizations who do not want to publish their reports on the cloud and are concerned about the security of their data.

5. Power BI Gateway

This component is used to connect and access on-premise data in secured networks. Power BI Gateways are generally used in organizations where data is kept in security and watch. Gateways help to extract such data through secure channels to Power BI platforms for analysis and reporting.

6. Power BI Mobile

Power BI Mobile is a native Power BI application that runs on iOS, Android, and Windows mobile devices. For viewing reports and dashboards, these applications are used.

7. Power BI Embedded

Power BI Embedded offers APIs which are used to embed visuals into custom applications.

8. Working of Power BI Architecture

Now that we have understood the individual components of Power BI, let us learn how all of these components work in tandem. We will understand the Power BI architecture with the help of this diagram**.**

3. Architecture Description

3.1. Data Description

The Google Play Store app dataset contains a comprehensive collection of information about various mobile applications available on the Google Play Store. The dataset includes structured data points and features that provide insights into app characteristics, user ratings, reviews, and other relevant attributes. Here is a description of some key data fields commonly found in a Google Play Store app dataset:

1. App Name: The name of the mobile application.
2. Category: The category or genre to which the app belongs (e.g., Games, Social, Productivity).
3. Rating: The average user rating for the app.
4. Reviews: The total number of user reviews/ratings received.
5. Installs: The estimated number of app installations or downloads.
6. Size: The size of the app in terms of storage space.
7. Price: The cost or pricing model of the app (e.g., Free, Paid, In-app purchases).
8. Content Rating: The age-based content rating assigned to the app (e.g., Everyone, Teen, Mature).
9. Last Updated: The date when the app was last updated.
10. Developer: The name of the app developer or development company.

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 scraping is made simple nowadays, many tools are used for web scraping. Some of the python libraries used for web scraping are Beautiful Soup, 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 Tableau

Launch Tableau on your workstation and select SQL Server from the connect column on the left. This will open a dialogue box where you need to provide the connection details for SQL Server.

To connect with tableau, you will need to provide information about the server which hosts your database. If you want to connect to a contained database, you can also specify the name of the database.

To connect with a port other than the default port, you need to specify the port and server as follows:

```
<server_name><port_number>
```

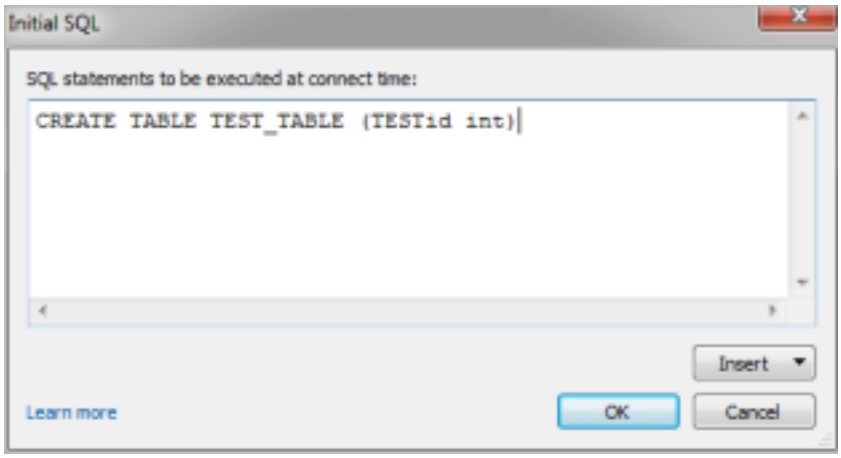
Example query: my_server 8051

There are two ways in which you can sign-in to the server, either by using Windows authentication or by using the username and password. Using the username and password becomes a must if you're working with a password-protected server in a non-Kerberos environment.

Click on Sign in to establish a connection. This will enable a connection without SSL. To establish an SSL enabled connection, click the Require SSL checkbox before you sign in.

SQL Server provides an option to let the user queries access the modified rows even before they have been committed. This option is called Read Uncommitted data. It saves time by preventing complex queries such as extract refreshes from locking the database and causing a delay. If this option is unchecked, Tableau makes use of default isolation levels.

If you want to run a specific SQL command every-time a new connection is established, you can use the Initial SQL option. This will open a dialogue box, where you can specify your desired SQL query.

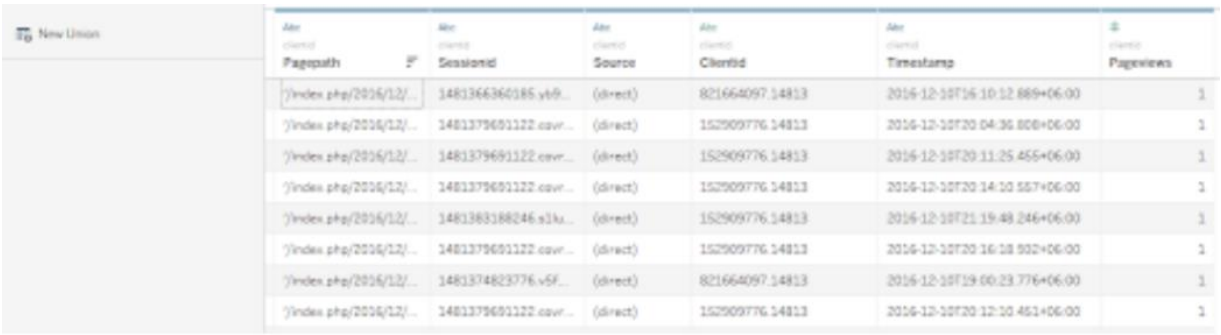


Step 2: Configuring Data Source

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

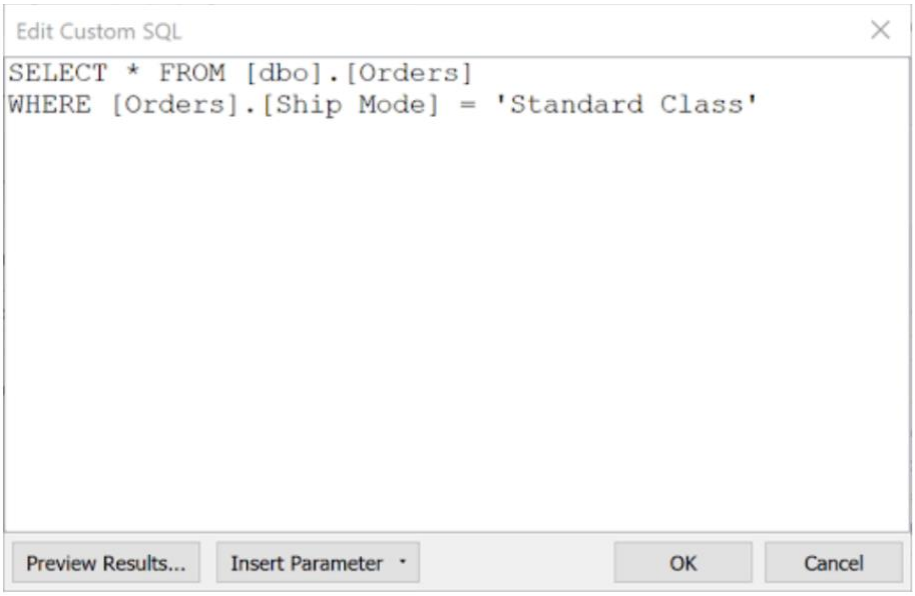
Select the data source name option and give a unique name to the database you are using. It's considered a good practice to have a unique name as it makes it much easier for users to identify the database from which data is being fetched.

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 Tableau. 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.



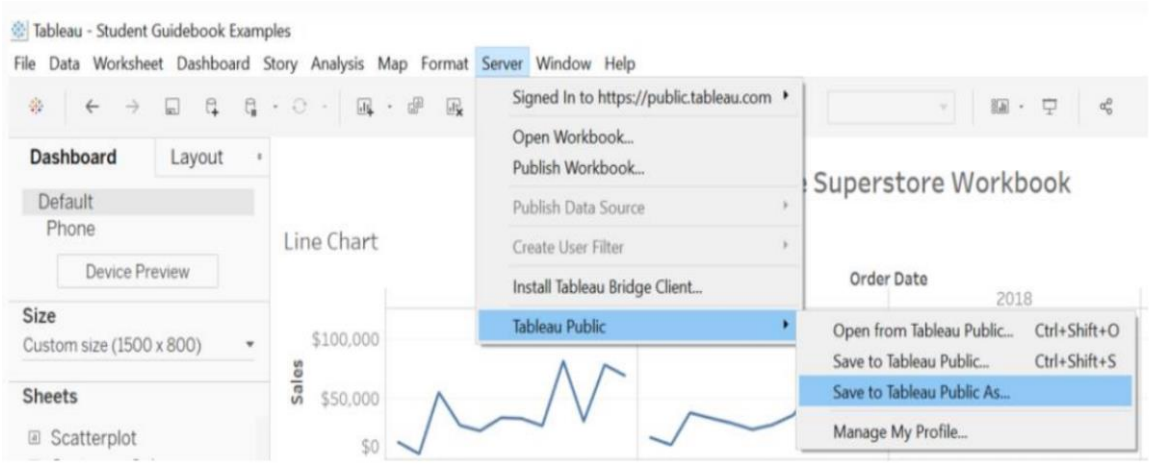
3.5. Export Data from Database

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

3.6 Deployment.

Once you've completed your dashboard, follow these steps:**- Server, Tableau Public, Save to Tableau Public As**

You may be prompted to log into your Tableau Public profile first if this is your first time publishing.



Next, fill out the title you want your viz to have and click “save”.

This message means that your connection to the Sample-Superstore data set is a live connection. Tableau Public cannot host live connections, so you’ll need to convert your connection to an extract (like a frozen screenshot of your data).

Here in the below screenshot, we can see that out workbook has been published to tableau public.

4. Unit Test Cases

Year parameter slicer	When clicked on the slicer, a dropdown should occur which has various years from 2010 to 2018.
Top 10 Categories	A Bar graph representation of Top 10 Categories of Apps
Relation Between Number of Apps and Content Rating	A Bar graph representation of Number of Apps vs Content Rating
Information of Total Application, Average Rating, Average Size	Various Info categories are shown and a visualization is created which shows the Total Application , Avg. Rating and Avg. Size.
Relation between Rating and Application Parameters on various apps	The visual should show a bar graph diagram of relation between Rating & Application.
Free Vs Paid App	A pie chart representation of free apps vs paid apps.