Low Level Design (LLD)

Analysing Google Apps Store

Revision Number: 1.0

Last Date of revision: 25/04/2025

Swaraj Ranjan Behera

Document Version Control

Date Issued	Version	Description	Author
26-04- 2025	1.0	First Version of Complete HLD	Swaraj Ranjan Behera

Table of Contents

Abstract – Google Play Store Data Analysis Project

The Google Play Store Data Analysis project focuses on exploring and analyzing various mobile applications available on the Play Store. This analysis helps uncover trends related to app downloads, ratings, reviews, pricing strategies, categories, and user preferences. By leveraging data from real-world apps, this project provides meaningful insights that can assist developers, marketers, and business analysts in making informed decisions.

The dataset is processed using data cleaning and Exploratory Data Analysis (EDA) techniques in Python. Additionally, insights are visualized using tools like Power BI, enabling interactive dashboards and reports. This high-level analysis supports stakeholders in identifying top-performing apps, optimal pricing strategies, and content trends across the Google Play ecosystem.

1. Introduction

1.1 What is a Low-Level Design Document?

The Low-Level Design (LLD) document provides the internal logic, data flow, and component-level design of the **Google Play Store Data Analysis** project. It defines how the data is processed, modeled, and visualized to deliver business insights using BI tools like Power BI.

The document details modules, data structures, steps of transformation, and dashboard deployment strategies. It serves as a blueprint for developers and analysts to understand and implement the dashboard effectively.

1.2 Scope

This LLD outlines the entire analysis pipeline for the Google Play Store dataset, which includes app-related data such as downloads, ratings, reviews, category types, and pricing.

The scope includes:

- Data collection from a cleaned CSV file
- ETL processing using Power BI
- Data modeling for relationships and time intelligence
- Dashboard development
- Publishing and sharing insights

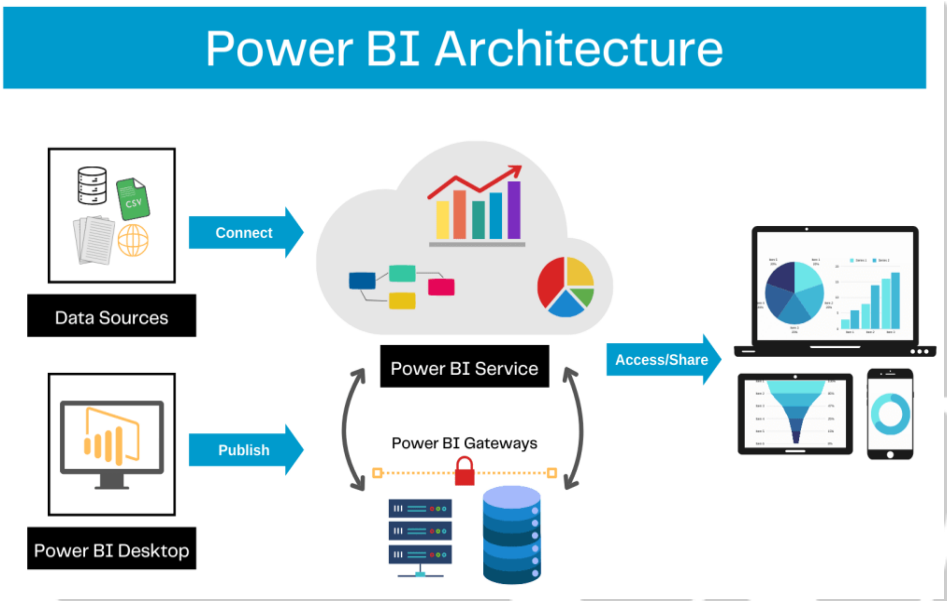
The objective is to provide visual analytics to help understand:

- Top-rated and most downloaded apps
- Category-wise app performance
- Free vs. Paid app trends
- Correlation between reviews, ratings, and installs

2. Architecture

The architecture for the Google Play Store Data Analysis project follows the **ETL + BI Dashboard** flow using **Power BI**. The system consists of the following key components:

- **Power BI Desktop** – Used for data connection, transformation, modelling, and report creation.
- **Power Query Editor** – Handles data cleansing and transformations.
- **Power BI Service** – For publishing and sharing the dashboard online.
- **Power BI Gateway (if needed)** – For connecting to local files securely in real-time.
- **Power BI Mobile** – For viewing dashboards on smartphones or tablets.
- The overall flow is:
  Data Source (CSV) → Power Query (ETL) → Data Modeling → Dashboard Visualization → Publish to Power BI Service



2.2 Tools used

Business intelligence tool Power BI is used to build the whole framework.

3. Architecture Description

3.1 Data Description

The dataset used for this project is a cleaned version of the **Google Play Store** dataset. It contains app-level information such as:

- **App** – Name of the app
- **Category** – App category (e.g., GAME, BUSINESS)
- **Rating** – Average rating
- **Reviews** – Total number of user reviews
- **Size** – Size of the app
- **Installs** – Total number of installs
- **Type** – Whether the app is Free or Paid
- **Price** – Price of the app
- **Content Rating** – Age group suitability
- **Genres** – Sub-category of the app
- **Last Updated** – Last date of update
- **Android Version** – Minimum Android version required

These columns are analyzed to extract insights on app popularity, user feedback, and trends across categories.

3.2 Data Loading

Steps:

1. Open **Power BI Desktop**
2. Click **Get Data** → **Text/CSV**
3. Select the cleaned Play Store dataset
4. Load the data and click **Transform Data** to clean further

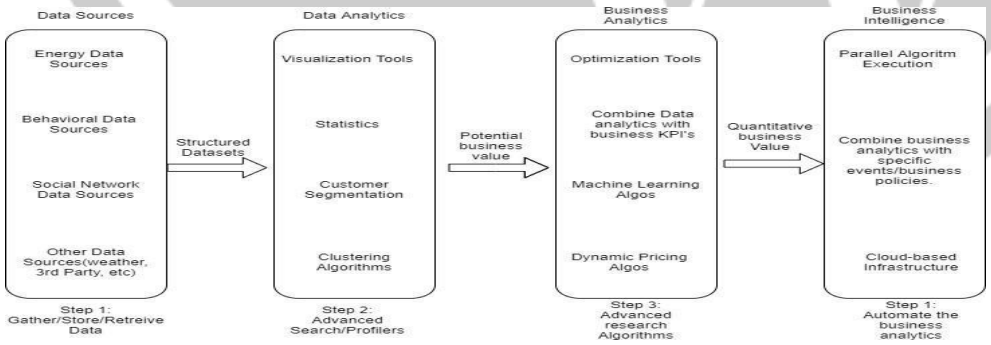
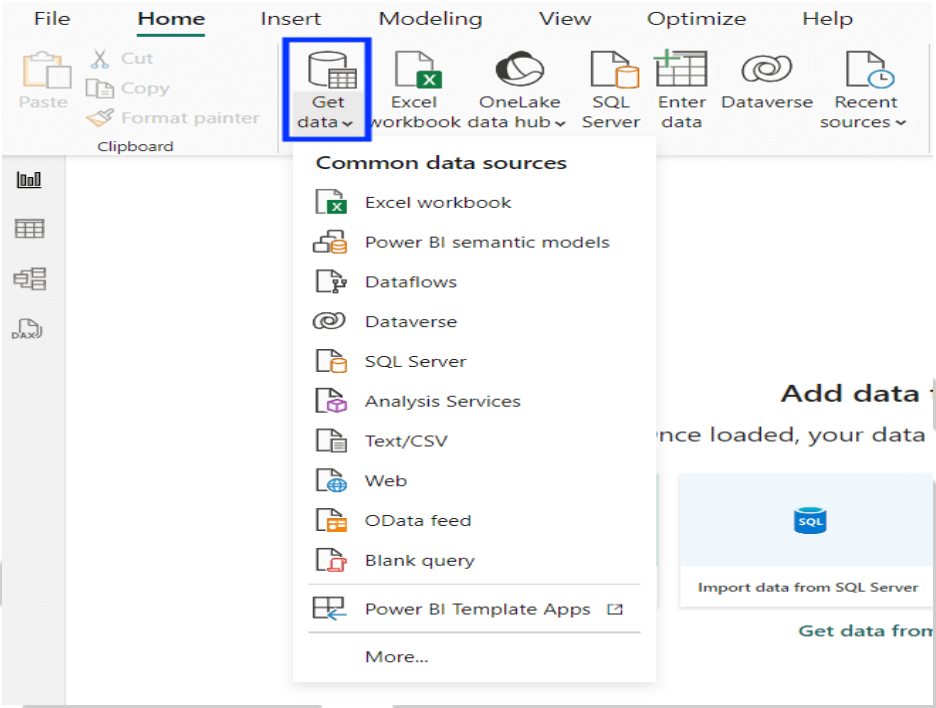


Figure 1: Functional Architecture of Business Intelligence



3.3 Data Transformation

In Power BI's **Power Query Editor**, the following transformation steps were applied:

- Removed duplicates and null values
- Converted columns to appropriate data types
- Extracted year from **Last Updated** for time-based analysis
- Created a new column **Install Range** by binning install counts
- Filtered out outliers from **Rating** (>5) and **Reviews** (>10M)

🔧 Optional: Added a **Date Dimension Table** to support time intelligence if needed.

3.4 Data Modeling

The model is kept simple as the dataset is flat (single table). If Date Dimension is added, a **relationship** is created:

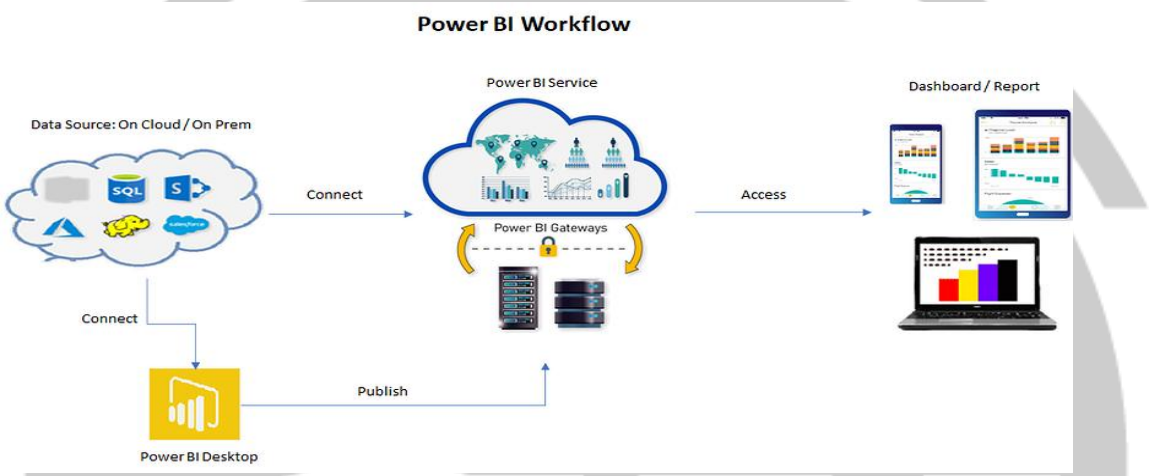
- `Last Updated` ⟶ `Date[Date]`

This enables monthly/yearly trend analysis.

3.5 Deployment

Once the dashboard is finalized in Power BI Desktop, follow these steps to deploy:

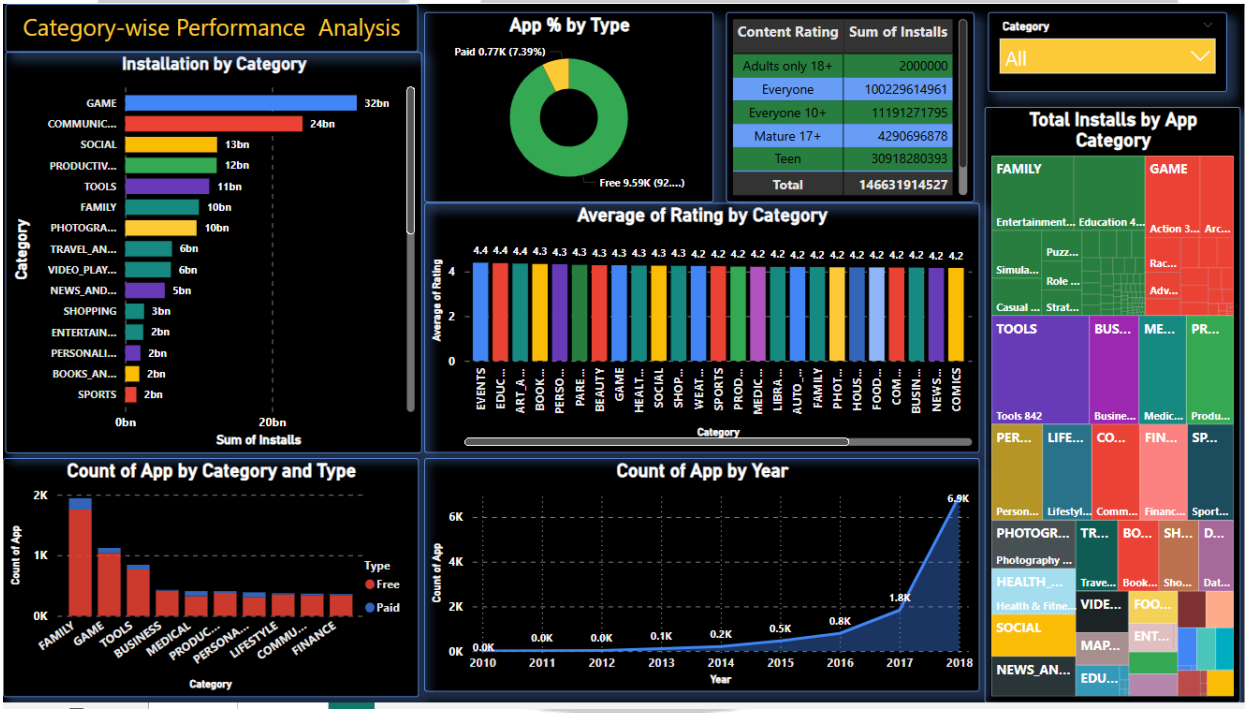
1. Click on the **“Publish”** button in the Home ribbon
2. Sign in to your **Power BI account**
3. Select a workspace (e.g., **My Workspace** or a shared one)
4. Wait for upload and click **“Open”** to view the report online
5. Set up **Scheduled Refresh** if needed to keep data updated



4. Unit Test Cases

Test Case	Description	Status
Slicers (Category, Type, Year)	Drop-down filters working correctly	✓ Working
Charts	All charts load without errors	✓ Working
Tooltips	Show correct values on hover	✓ Working
Page Navigation Buttons	Navigate between pages smoothly	✓ Working
Data Refresh	Dataset updates correctly on refresh	✓ Working


Google Play Store Analysis
Type
Type: All
Free
Paid
9.637K
Total Apps
4.20
Avg Rating
119
Genres
147bn
Total Installs
4bn
Total Reviews
33
Categories
Total Installs by App
Maps - Navigate & Explore 3bn (10.71%)
Subway Surfers 5bn (17.86%)
Instagram 3bn (10.71%)
Google Photos 4bn (14.29%)
Google News 3bn (10.71%)
Google Drive 3bn (10.71%)
Google Chrome: Fast & Secure 3bn (10.71%)
Hangouts 4bn (14.29%)
Top 5 Genres by Rating
Tools 3.4K
Entertainment 2.4K
Education 2.3K
Business 1.8K
Medical 1.7K
Genres
Sum of Rating
Top 5 Apps by Reviews
0.20bn
0.16bn
0.14bn
0.14bn
0.13bn
Sum of Reviews
Instagram
Facebook
Subway...
WhatsA...
Clash of...
App
Top 5 Android version by Installation
80bn
28bn
11bn
10bn
7bn
Sum of Installs
Varies with device
4.1 and up
4.0 and up
4.4 and up
4.0.3 and up
Android Ver
Top 5 Category by Installs
GAME 32bn
COMMUNIC... 24bn
SOCIAL 13bn
PRODUCTIVI... 12bn
TOOLS 11bn
Category
Sum of Installs
Total Reviews by App
Sum of Reviews
App


Category-wise Performance Analysis
Installation by Category
GAME 32bn
COMMUNIC... 24bn
SOCIAL 13bn
PRODUCTIV... 12bn
TOOLS 11bn
FAMILY 10bn
PHOTOGRA... 10bn
TRAVEL_AN... 6bn
VIDEO_PLAY... 6bn
NEWS_AND... 5bn
SHOPPING 3bn
ENTERTAIN... 2bn
PERSONALI... 2bn
BOOKS_AN... 2bn
SPORTS 2bn
Category
Sum of Installs
App % by Type
Paid 0.77K (7.39%)
Free 9.59K (92....)
Content Rating | Sum of Installs
Adults only 18+ | 2000000
Everyone | 100229614961
Everyone 10+ | 11191271795
Mature 17+ | 4290696878
Teen | 30918280393
Total | 146631914527
Category
All
Average of Rating by Category
Average of Rating
Category
Total Installs by App Category
FAMILY
GAME
TOOLS
BUS...
ME...
PR...
PER...
LIFE...
CO...
FIN...
SP...
PHOTOGR...
TR...
BO...
SH...
HEALTH_...
VIDE...
FOO...
SOCIAL
MAP...
ENT...
NEWS_AN...
EDU...
Count of App by Category and Type
Count of App
Type
Free
Paid
Category
Count of App by Year
Count of App
Year