High Level Design

Analyzing Google App Store

Revision Number: 1.0

Last date of revision: 20/01/2023

Document Version Control

Date Issued	Version	Description	Author
20th April 2024	1.0	First Version of Complete HLD	Utkarsh Shrikhande

Contents

1 Introduction

1.1 Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and the architecture of the project • describe the non-functional attributes like:
    - Security ○ Reliability ○ Maintainability ○ Portability ○ Reusability ○ Application compatibility ○ Resource utilization
    - Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non- technical to mildly-technical terms which should be understandable to the administrators of the system.

2 General Description

2.1 Problem Statement

Technology is the increasing need nowadays and used everywhere. One of the features of Technology is android. Which we all use in our daily life. Android is a mobile operating system based on a modified version of the Linux kernel and other open source software, designed primarily for touchscreen mobile devices such as smartphones and tablets.

Do ETL: Extract-Transform-Load the dataset and find for me some information from this large data. This is form of data mining.

What all information can be achieved by mining this data, would be brainstormed by the interns.

Find key metrics and factors and show the meaningful relationships between attributes.

Do your own research and come up with your findings.

2.2 Tools used

Business Intelligence tools and Python libraries works such as NumPy, Pandas, Excel, Power BI are used to build the whole framework.

3 Design Details

3.1 Functional Architecture

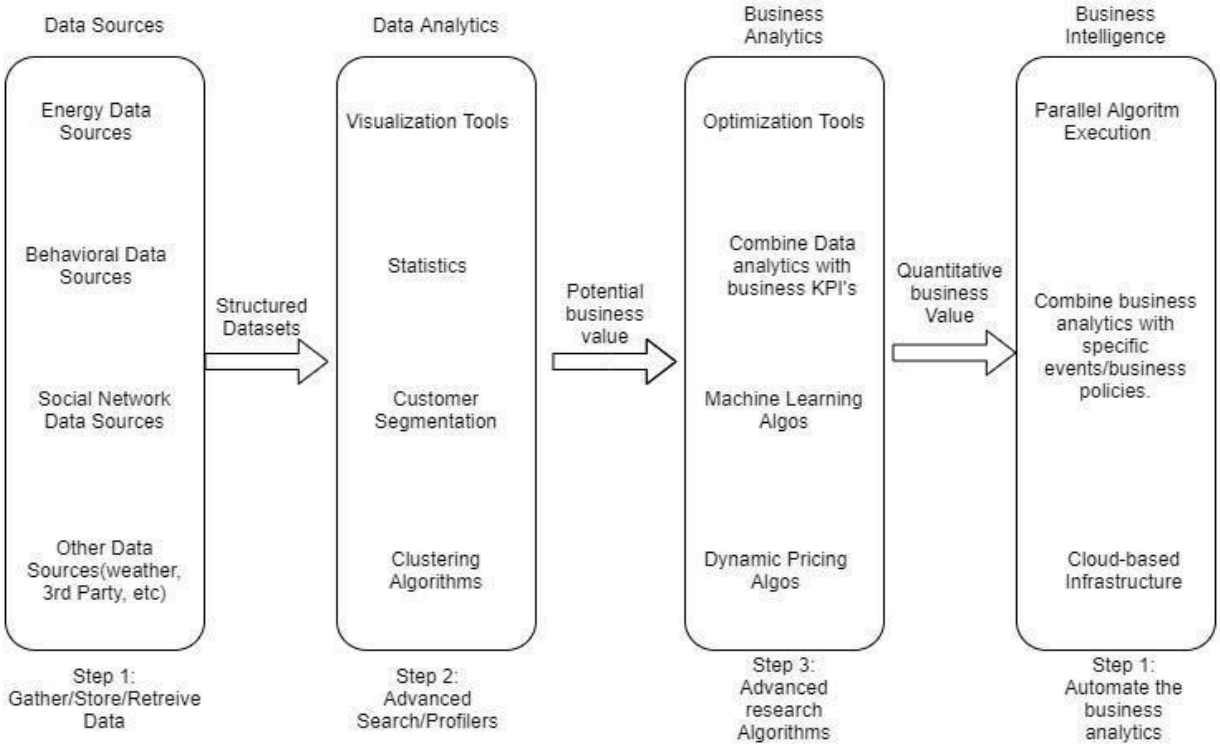
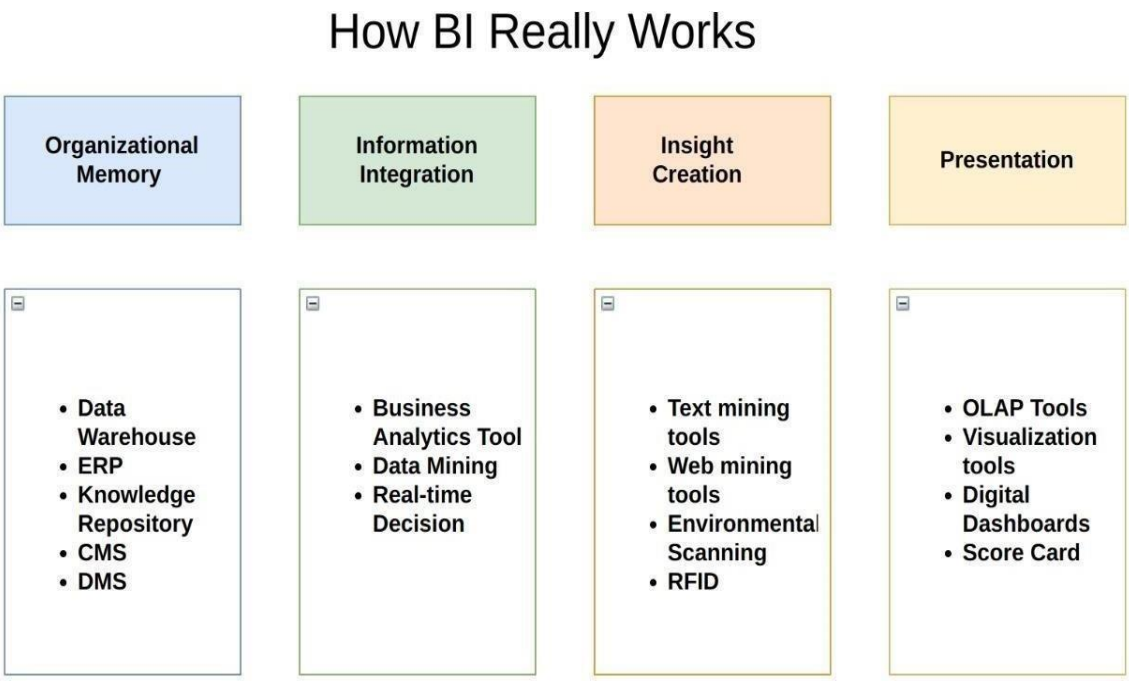


Figure 1: Functional Architecture of Business Intelligence



3.2 Optimization

Your data strategy drives performance

- Minimize the number of fields
- Minimize the number of records
- Optimize extracts to speed up future queries by materializing calculations, removing columns and the use of accelerated views

Reduce the marks (data points) in your view

- Practice guided analytics. There's no need to fit everything you plan to show in a single view. Compile related views and connect them with action filters to travel from overview to highly-granular views at the speed of thought.
- Remove unneeded dimensions from the detail shelf.
- Explore. Try displaying your data in different types of views.

4 KPIs

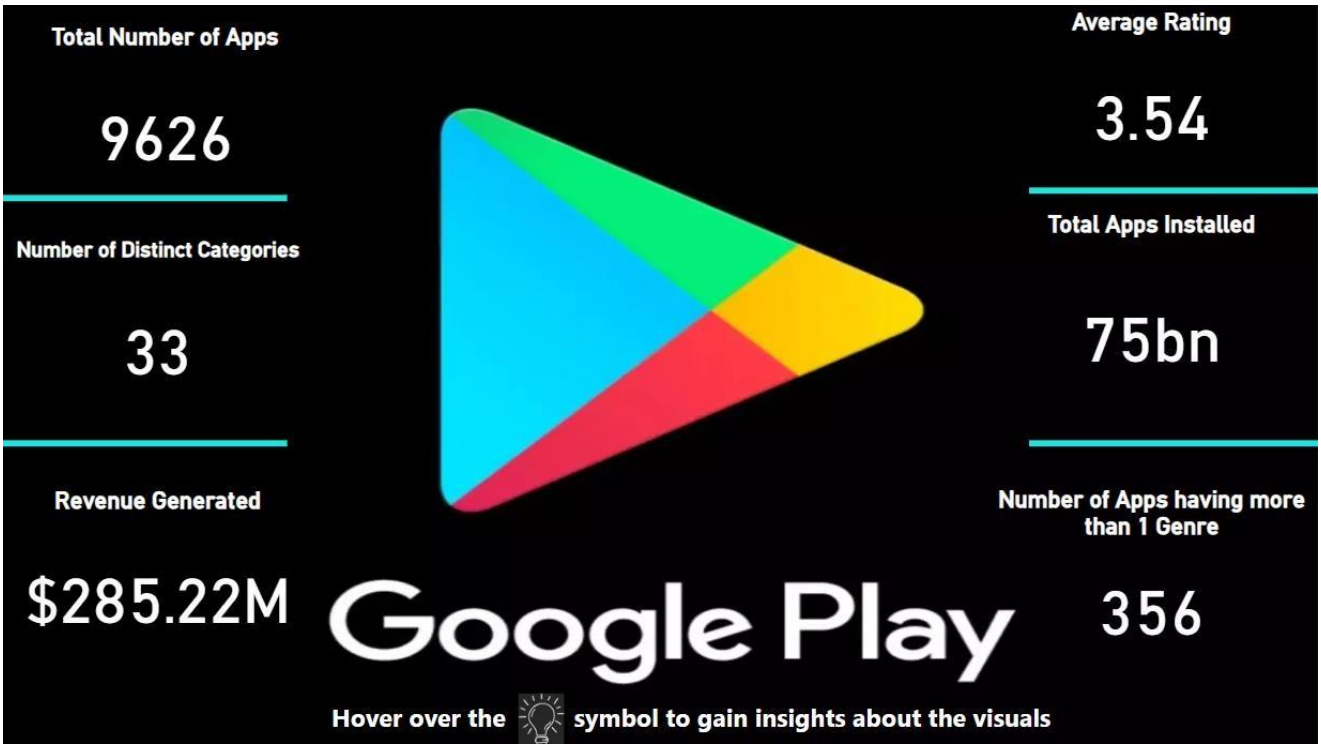
KPI stands for key performance indicator, a quantifiable measure of performance over time for a specific objective. KPIs provide targets for teams to shoot for, milestones to gauge progress, and insights that help people across the organization make better decisions. From finance and HR to marketing and sales, key performance indicators help every area of the business move forward at the strategic level.

4.1 KPIs (Key Performance Indicators)

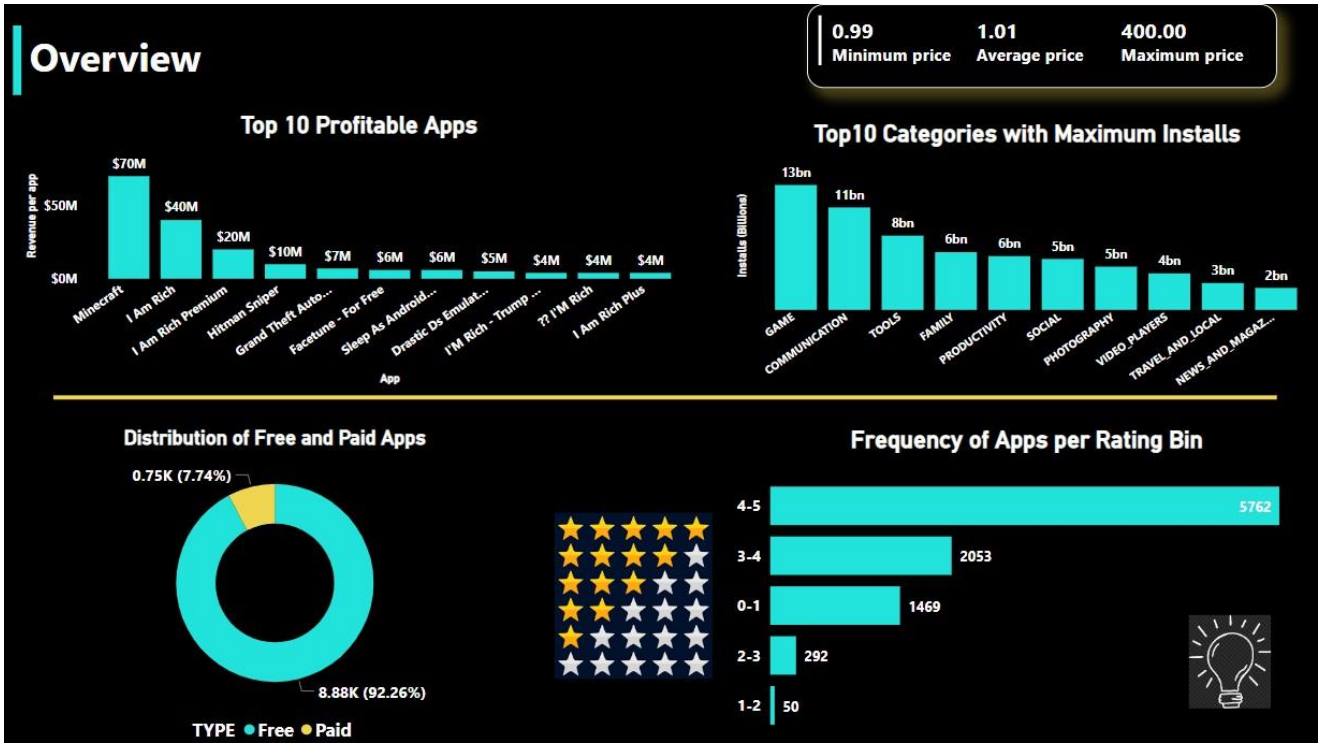
Key indicators displaying a summary of the apps attributes and its relationship with users reviews to understand the sentiment of the users.

High-Level Design (HLD)

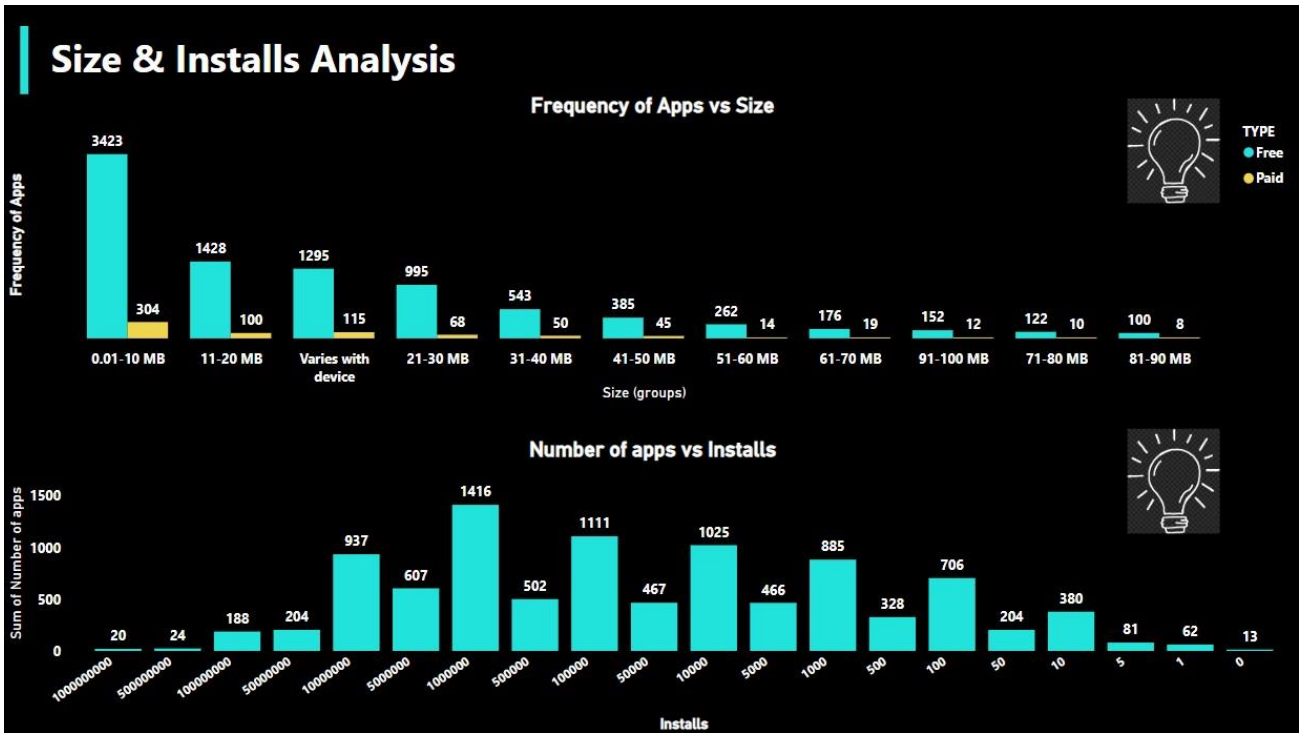
- Overview



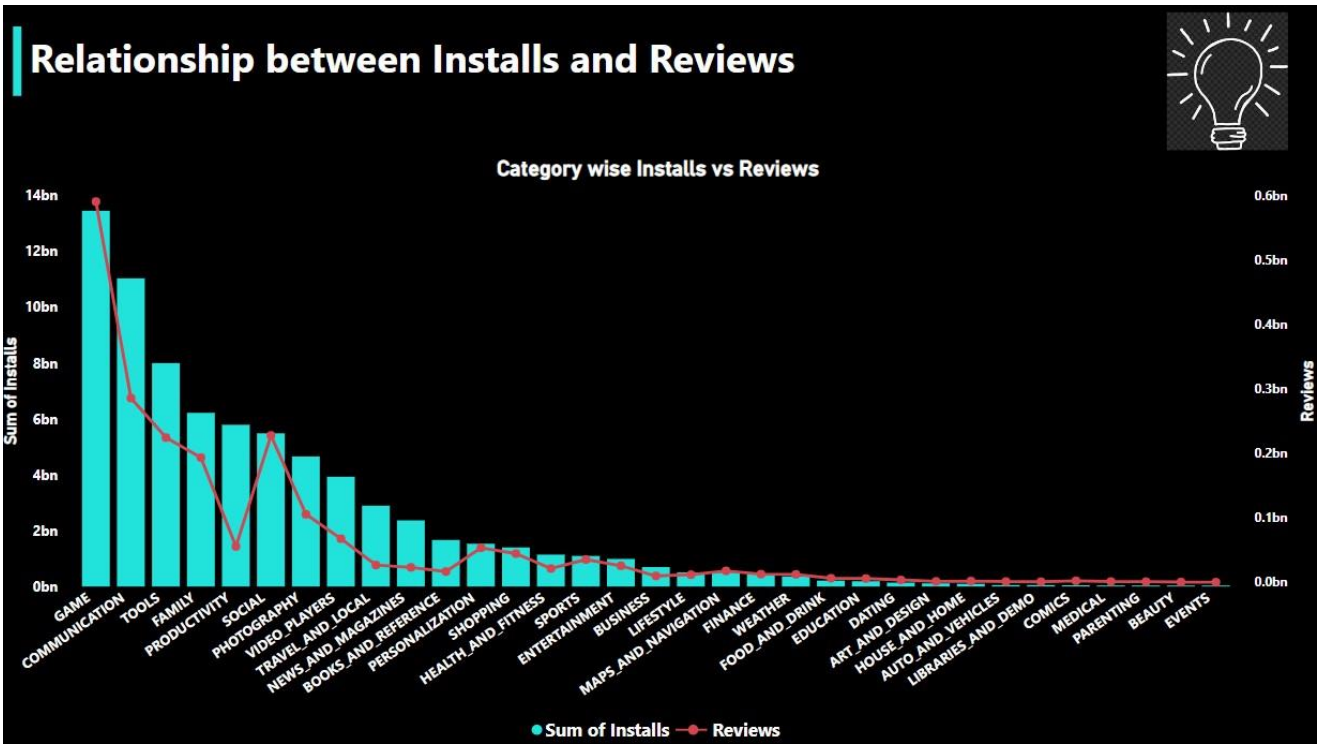
- The bin 4-5 of ratings accounted for 59.8 % of the total ratings received on all the apps. This shows that more than half of the customers were satisfied with Google Play Store Apps and its services.



- Around 39% of the Apps are between 0.01 MB and 10 MB, which shows that majority of the apps will not consume much space. Also, there are very less apps which may consume huge space. Communication genres has the highest installs among the category.
  There are 20 apps with maximum number of installs i.e. 1000 Million. The Play Store even includes 13 apps which have not been installed by anyone.

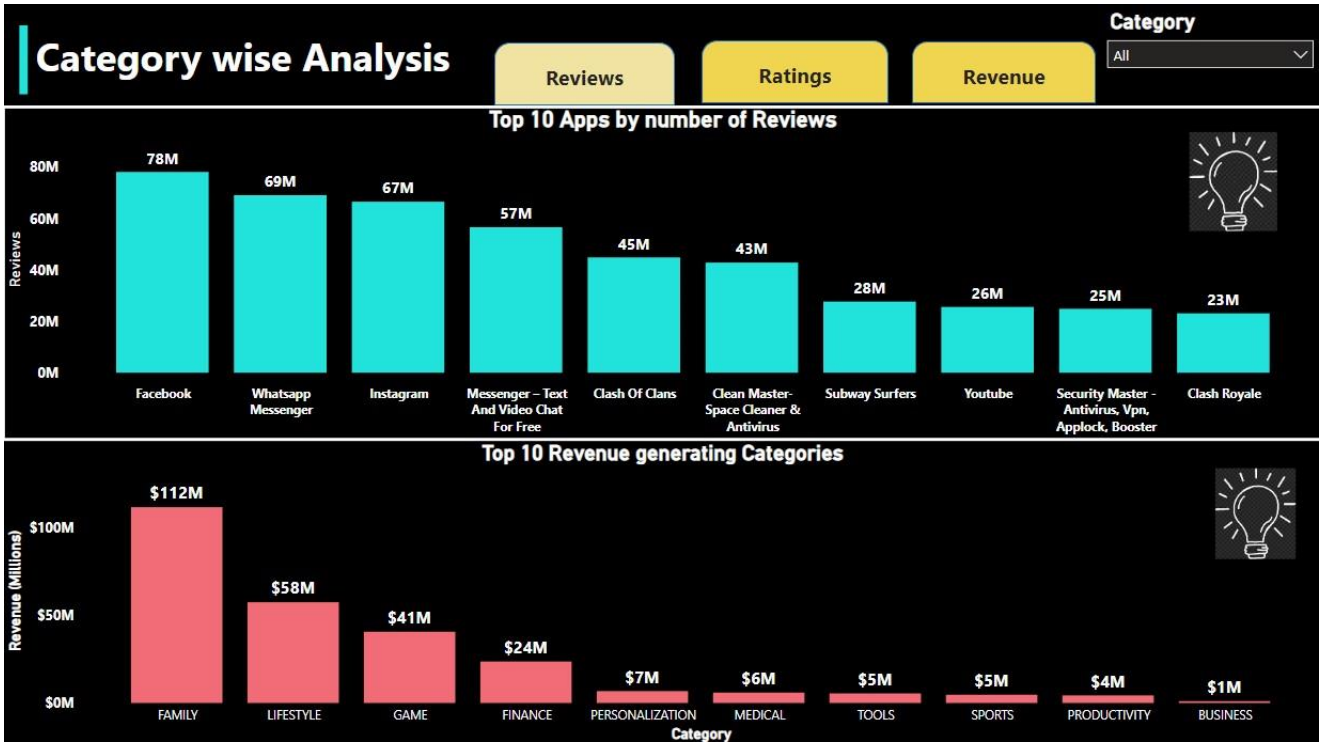


- Category wise Installs and Reviews are positively correlated with each other as shown in the figure. The Category GAME accounted for 17.89 % of total Installs and 28.2 % of total Reviews received.

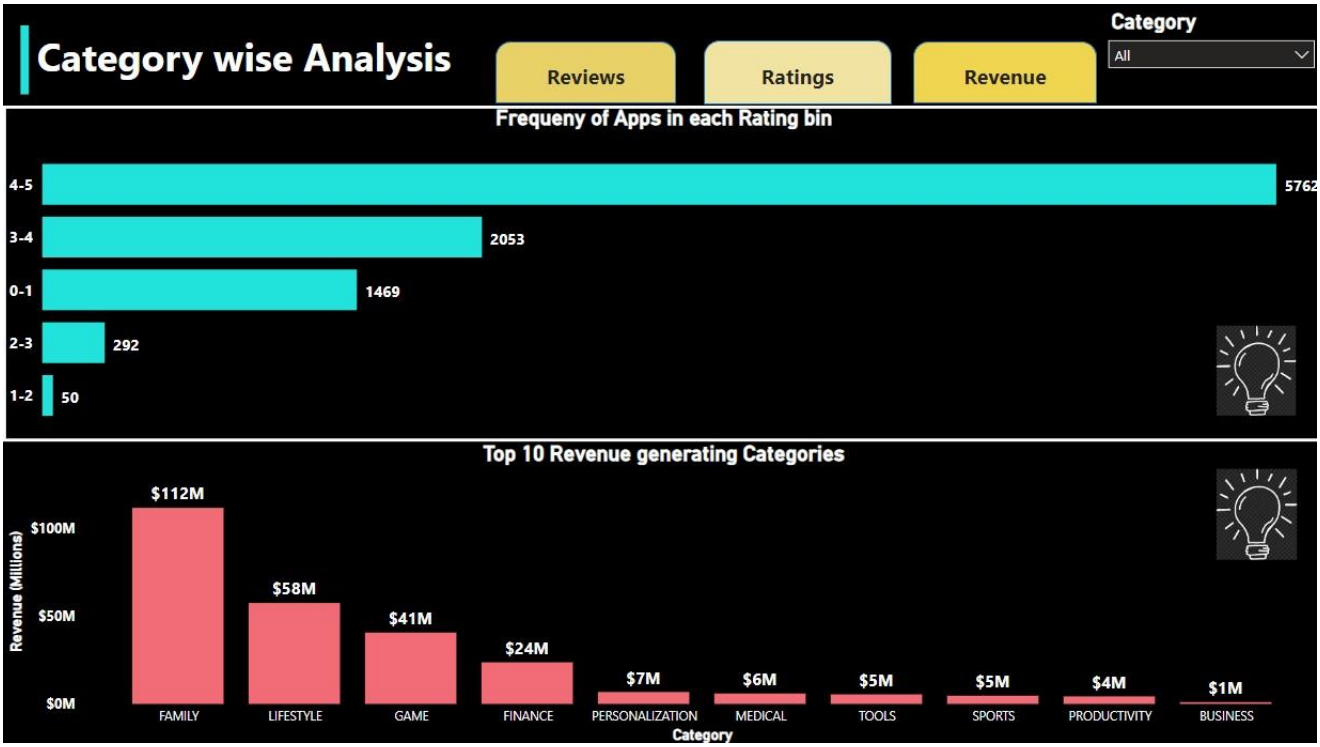


10 High-Level Design (HLD)

- All the social media apps including Facebook, WhatsApp Messenger, Instagram and Messenger contributed the most to the number of reviews. They accounted for 12.9% of the total reviews.

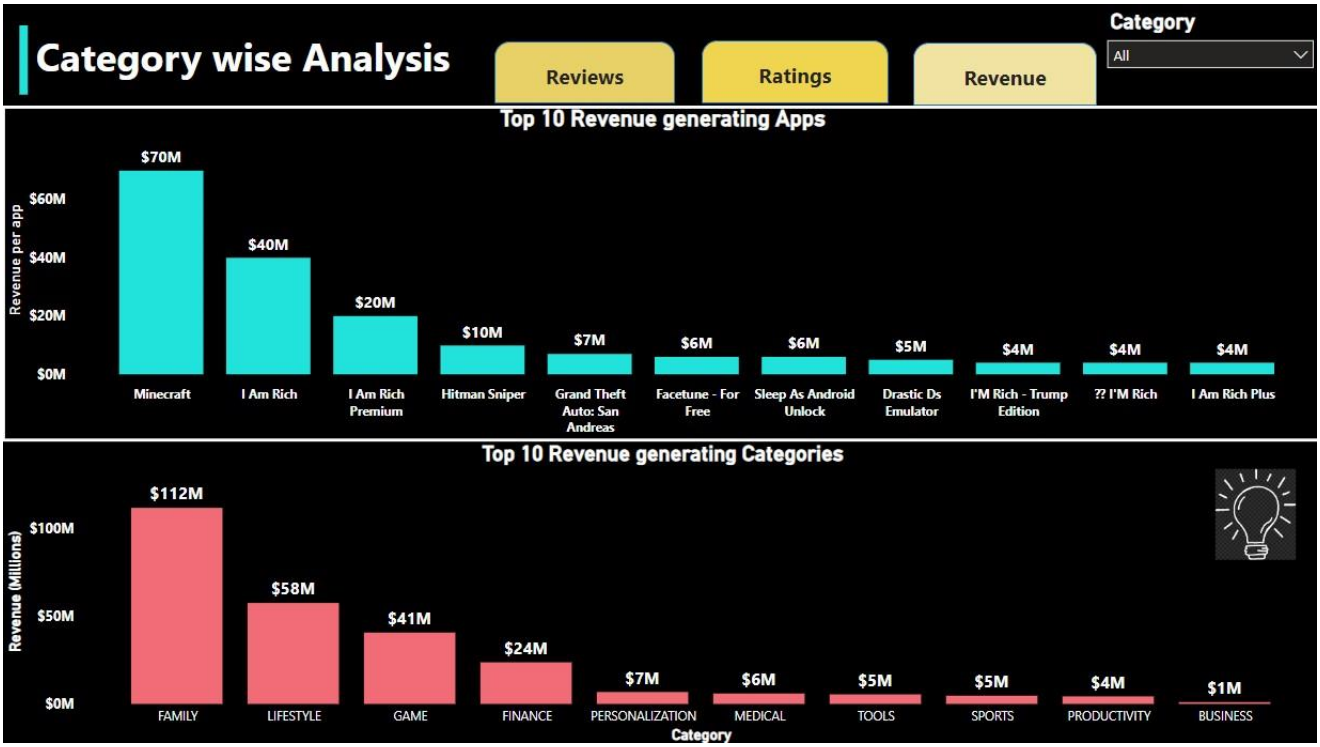


- 59.86% of the ratings received were between 4 and 5. The ratings between 4 and 5 was 11,424 % higher than the ratings received between 1 and 2.

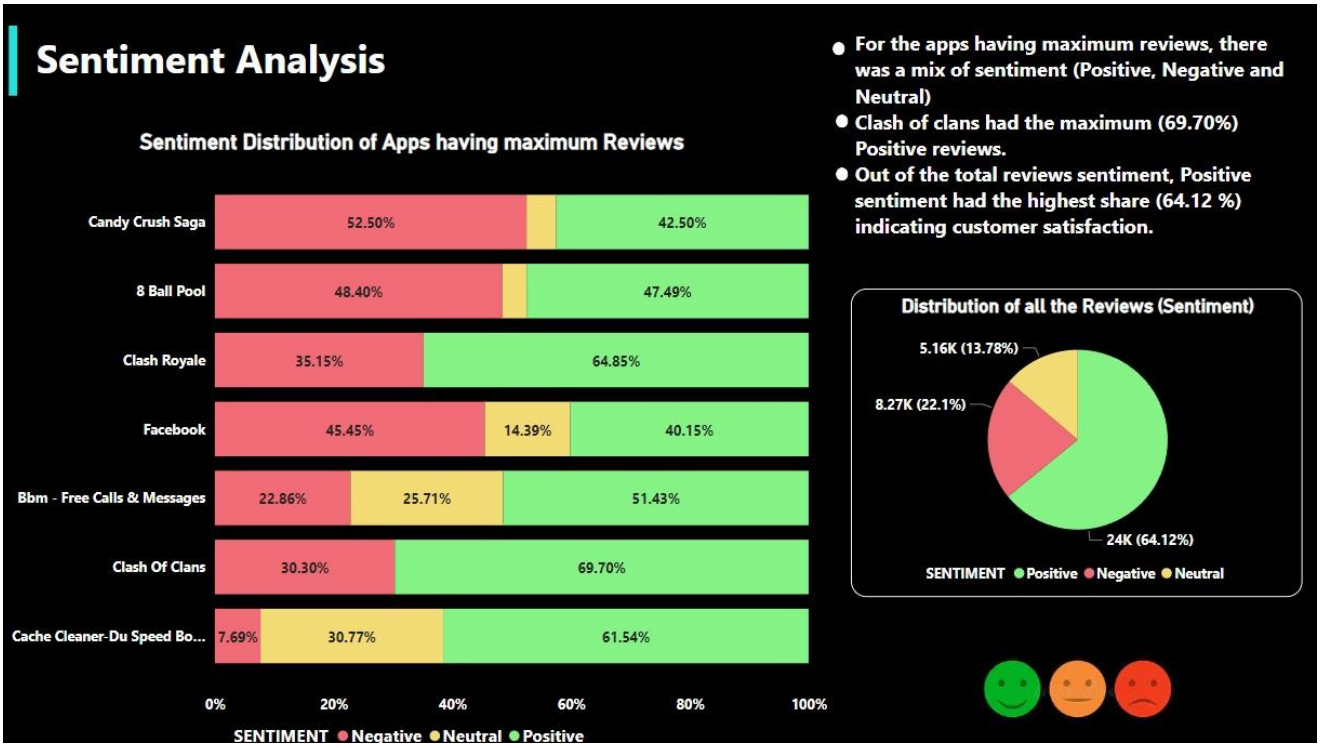


11 High-Level Design (HLD)

- The Category FAMILY accounted for 39.29 % of the total revenue generated which is the highest among all the Categories.



- Sentiments:

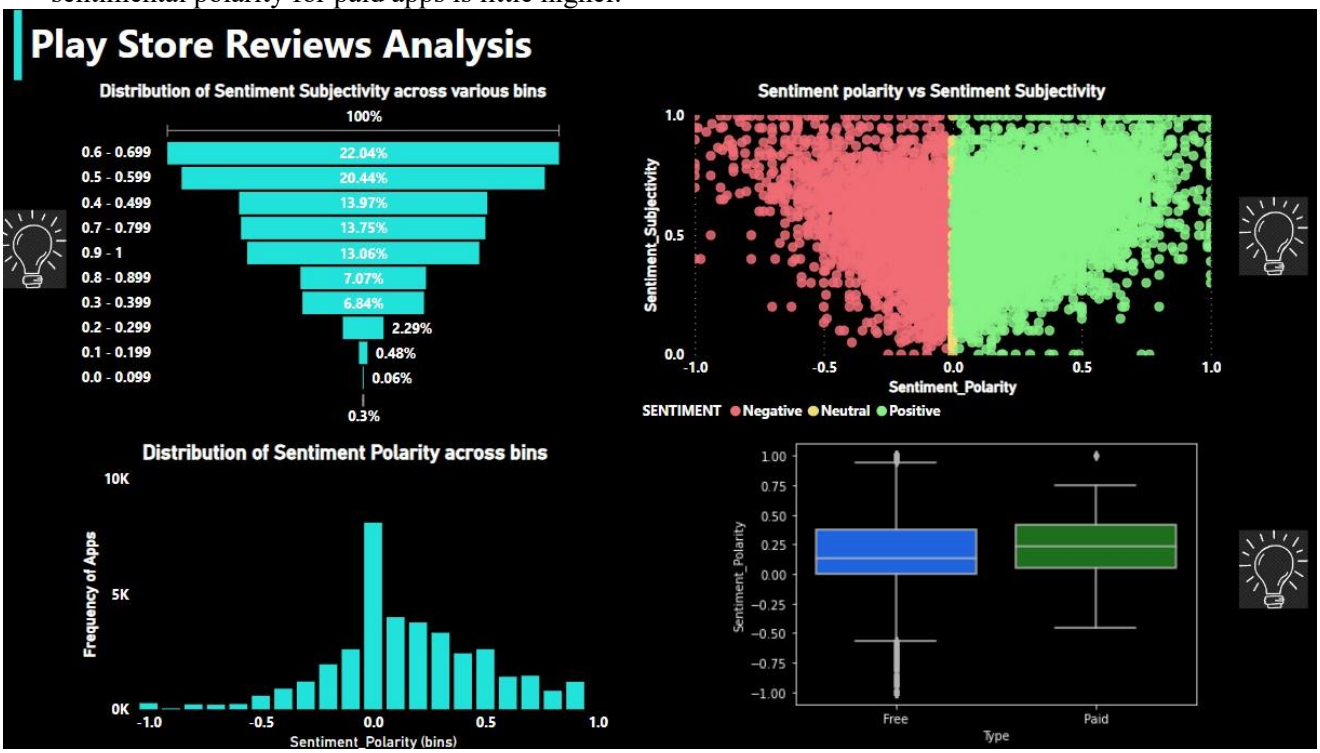


12 High-Level Design (HLD)

- The Apps with Sentiment Subjectivity between 0 and 0.3 account for only 2.83 % of the total. It means there are few users who have given reviews based on the personal opinions. Most of the users have given reviews based on factual information.

  Most of the reviews either good or bad are based majorly on their experiences of the users after using the app.

  We can observe through the outliers that there are a lot of free apps having extremely negative sentiment polarity. This may indicate that free apps are not suited to users as much as paid apps. The median sentimental polarity for paid apps is little higher.



5. Deployment

Prioritizing data and analytics couldn't come at a better time. Your company, no matter what size, is already collecting data and most likely analyzing just a portion of it to solve business problems, gain competitive advantages, and drive enterprise transformation. With the explosive growth of enterprise data, database technologies, and the high demand for analytical skills, today's most effective IT organizations have shifted their focus to enabling self-service by deploying various BI tool.