

GOOGLE PLAY APP ANALYSIS PROJECT

PREPARED BY
MADU CHIBUIKE JUSTIN



madujustinc@gmail.com
+2347062579819

May 2022

Professional Background

Passionate, motivated, teamwork-oriented, and enthusiastic data analyst who strives to pose, answer questions and solve business problems with quantitative-driven insights. Through the development of personal projects, valuable training, and work experience.

I have learned the importance of having an iterative, hypothesis-oriented approach to analysis. With significant experience in data visualization, transformation, creating reports, creating dashboards, and sharing insights.

I possess a strong educational background in chemical engineering, which has equipped me with solid mathematical and statistical skills. This enables me to approach data analysis with rigor and a meticulous mindset.

My diverse skillset includes proficiency in various tools and technologies such as Excel, Power BI, Power Automate, SQL, Tableau, Azure, SSIS, and Python, which I use to extract, transform, and analyze data from different sources.

Overall, I am a versatile and resourceful data/BI analyst, committed to delivering high-quality work within tight deadlines. I am passionate about solving real-world problems through data analysis and continually learning new technologies and techniques to improve my skills.

Portfolio Outline



Professional Background	1
Portfolio Outline	2
Google Play App Project Description	3
The Problem	4
Data Design	5
Findings	6
Data Analysis	9
Conclusion	10
Appendix	11



Google Play Store App Performance Project Description

- **Situation:** I was given a real world dataset of apps that are scraped from the Google Play App Store website to work with and understand where opportunities to drive growth and retention may lie and also track their performance in the play store.
- **Task:** The purpose of this report was to understand where opportunities to devise strategies to drive app growth and retention may lie, and track their performances in the android app market. Data was scraped from the Google Play App Store website in a spreadsheet. Data collected are app categories, their ratings, reviews, number of installs, type, date updated, etc.
- **Action:** I uploaded the dataset in Google spreadsheet, performed data cleaning, created pivot tables and charts in excel and finally created more visualizations and dashboard using Tableau Public.
- **Result:** From the report, apps in the Games category has the highest number of reviews and therefore, has high performance while apps in the Events category has the lowest number of reviews. The root cause for underperforming apps in the Google play store is as a result of lack/poor updates of apps by their developer in the app store.

The Problem

- What does the business want to achieve? The business wants to drive growth and retention of apps and track their performance in the Google Play Store
- What is the cause of the problem? Lack/poor updates of apps by their developers in the app store
- What data should be collected? App categories, ratings, reviews, number of installs, type, date updated, and any other relevant data that may help in understanding app growth and retention.
- What will be the impact if the problem is not solved? Apps will continue to underperform, leading to low growth and retention rates, resulting in decreased revenue and potentially losing market to competitors.

Data Design

The Google Play apps data collected was unclean. I removed the duplicates, and the blank cells, ensured my headers are well written and ensured that there is consistency in the data, Below were the steps I used to clean the data;

- **Duplicate Rows:** This is common in datasets and can affect the analysis. I removed duplicates from my spreadsheet (Google Sheets) with the 'Remove Duplicate Function' by selecting the entire datasheet by pressing the ctrl + A on my keyboard. I Clicked on Data > Remove Duplicates. All duplicated rows were removed.
- **Duplicate Column:** Some apps appeared more than once in the App column. So, I deleted the duplicated apps by repeating the steps above but instead of selecting the whole datasets, I only selected the App column.
- **Blank Cells:** Blank cells can create gaps in the data and can make the data inconsistent and dirty. I used the Filter Function to remove blank cells. I selected the entire sheet/dataset and clicked on Data > Create filter. I clicked on the filter icon at the top of one of the columns which look like an inverted pyramid. In the dropdown menu that appeared, I clicked on the 'Clear' which deselects the filters. Then, I clicked on 'Blanks' after I have clicked on 'Clear' in the previous step and then clicked OK. All the blank rows were filtered out together and all the rows with data were hidden. Then I selected all the empty rows and deleted them. After deleting the blank rows, I now removed the filter by clicking on Data > Turn off Filter. The dataset was now free of blank rows.
- **Headers:** I ensured the headers were clear and concise by using underscores between words.
- **Consistent Data:** In the dataset, I noticed that "nan" was used to replace some values of "0". So, I used the Edit > Find and Replace function to replace all values of "nan" with "0", making my dataset consistent.

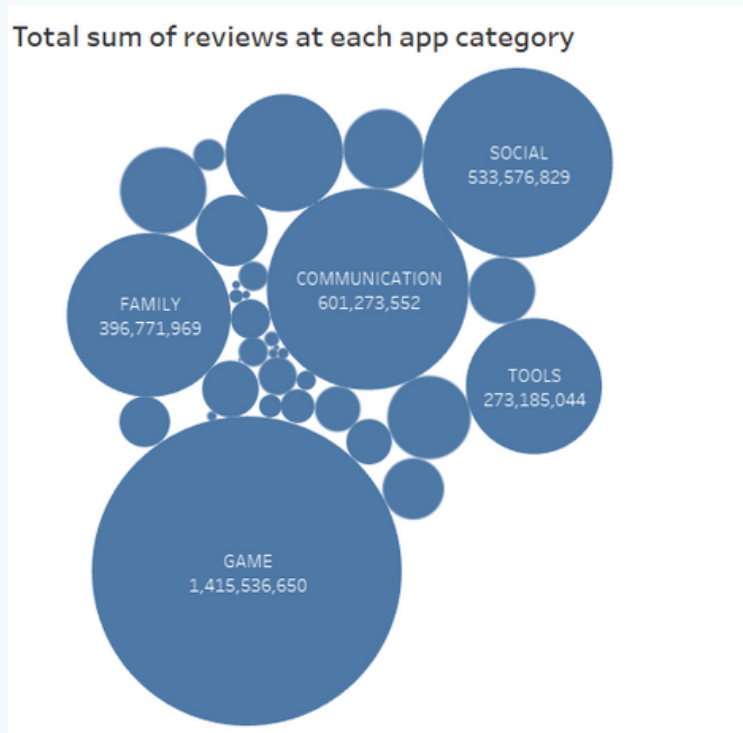
After I cleaned the data, I now analyzed the dataset by creating pivot tables and charts with the Google sheet and designed a dashboard with Tableau Public for more insight into the data.

Findings

- From the dashboard, the Games category has the highest number of reviews with an overall review sum of 1,415,536,650. While the Events category has the least number of reviews (161,018) followed by the Beauty (396,240) and the Parenting categories (958,331).
- Apps with 1,000,000+ installs have the highest installs in the Google play store with a total of 1,488 unique installs.
- Apps in the Family category have the highest count of ratings and apps in the Beauty and Comics have the least count of ratings.
- There are more free apps than paid apps in the app play store.
- Age category/Content rating 'Everyone' has rated the apps more than any others.

Finding 1

Figure 1: Total Reviews by Category



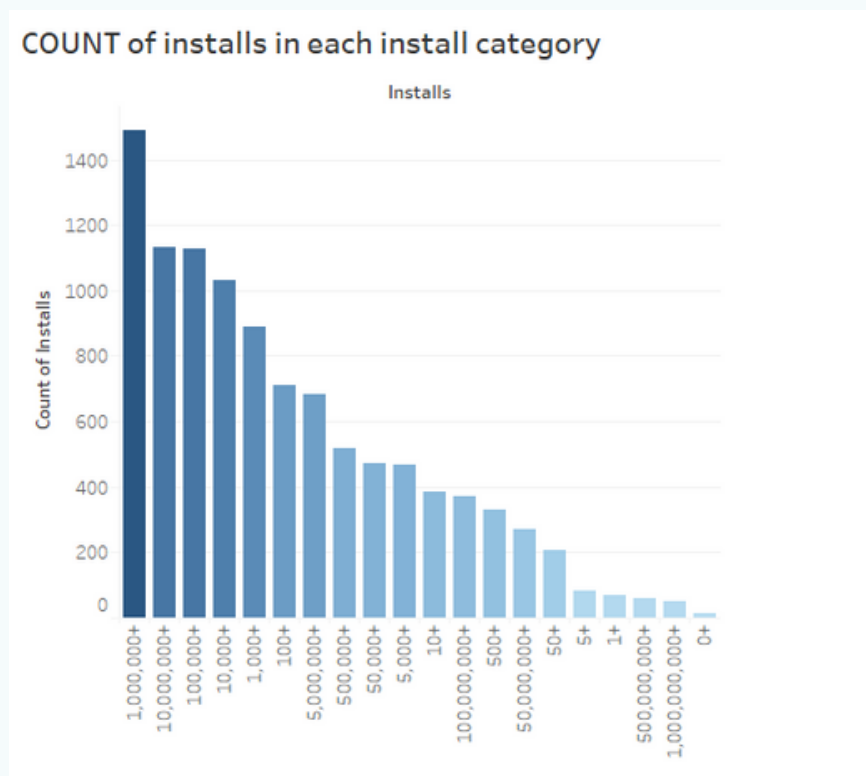
The Games category has the highest app reviews while the Events category has the least reviews followed by the Beauty and the Parenting categories.

Table 1: Top 20 Most Reviewed Apps

App	Category	Rating	Review	Type	Installs	Year_Updated
Facebook	SOCIAL	4.1	78158306	Free	1,000,000,000+	2018
WhatsApp Messenger	SOCIAL	4.1	78128208	Free	1,000,000,000+	2018
Instagram	COMMUNICATION	4.4	69119316	Free	1,000,000,000+	2018
Messenger – Text and Video Chat for Free	COMMUNICATION	4.4	69109672	Free	1,000,000,000+	2018
Clash of Clans	SOCIAL	4.5	66577446	Free	1,000,000,000+	2018
Clean Master- Space Cleaner & Antivirus	SOCIAL	4.5	66577313	Free	1,000,000,000+	2018
Subway Surfers	SOCIAL	4.5	66509917	Free	1,000,000,000+	2018
YouTube	COMMUNICATION	4	56646578	Free	1,000,000,000+	2018
Security Master - Antivirus, VPN, AppLock, Booster	COMMUNICATION	4	56642847	Free	1,000,000,000+	2018
Clash Royale	GAME	4.6	44893888	Free	100,000,000+	2018
Candy Crush Saga	GAME	4.6	44891723	Free	100,000,000+	2018
UC Browser - Fast Download Private & Secure	FAMILY	4.6	44881447	Free	100,000,000+	2018
Snapchat	TOOLS	4.7	42916526	Free	500,000,000+	2018
360 Security - Free Antivirus, Booster, Cleaner	GAME	4.5	27725352	Free	1,000,000,000+	2018
My Talking Tom	GAME	4.5	27724094	Free	1,000,000,000+	2018
8 Ball Pool	GAME	4.5	27723193	Free	1,000,000,000+	2018
DU Battery Saver - Battery Charger & Battery Life	GAME	4.5	27722264	Free	1,000,000,000+	2018
BBM - Free Calls & Messages	GAME	4.5	27711703	Free	1,000,000,000+	2018
Cache Cleaner-DU Speed Booster (booster & cleaner)	VIDEO_PLAYERS	4.3	25655305	Free	1,000,000,000+	2018
Twitter	VIDEO_PLAYERS	4.3	25623548	Free	1,000,000,000+	2018

Finding 2

Figure 2: The Count of Installs

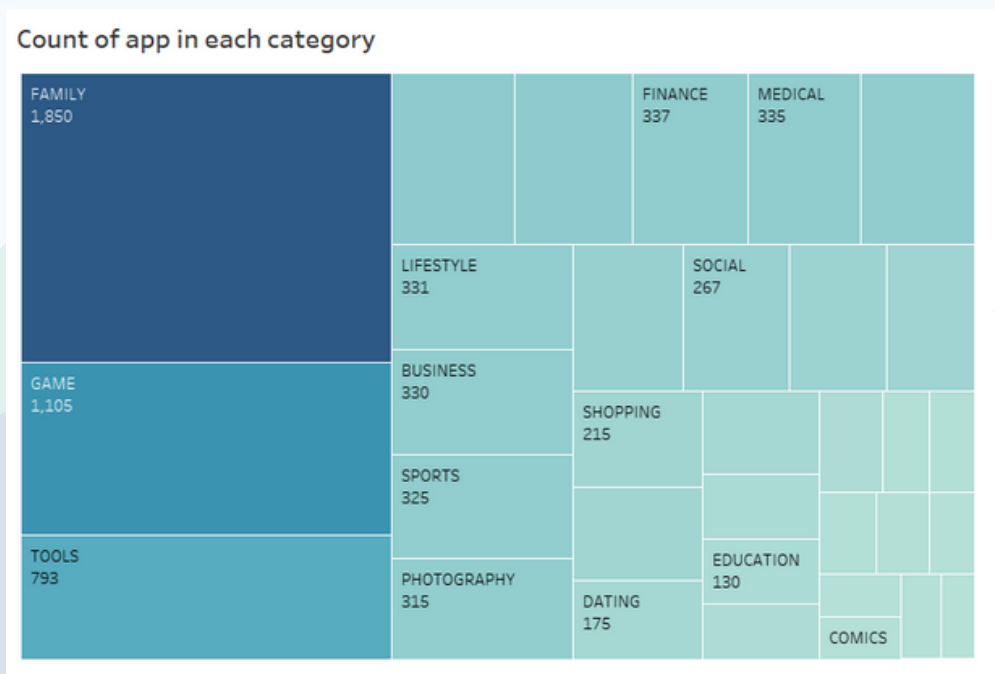


Apps with 1,000,000+ installs has the highest installs in the app play store with a total of 1,488 unique installs

Table 2: The Count of Free and Paid Apps

Type	COUNTA of Type
Free	9,593
Paid	765
Grand Total	10,358

Figure 3: The Count of Apps by Category



Apps in the Family category has the highest number of apps in the play store with 1,850 number of apps, and apps in the Beauty category has the lowest number of apps in the play store, having total number of 50 apps

Analysis

We will use the 5 Whys technique to determine the root cause of poor app performance in the Google Play Store by repeatedly asking the question “Why”.

The Root Cause Analysis (The Whys):

- Why do some app categories have low performance in the Google Play App Store?
Because they have a low number of installs.
- Why do apps have a low number of installs?
Because they have low to no ratings and reviews.
- Why do they have low/no ratings and reviews?
Because there can be bugs, poor UI, or fake apps.
- Why do we have bugs, poor UI, and fake apps in Google Play? Because of the non-update of apps to the latest version by the app creator or developer.
- Why have the apps not been updated regularly?
Because the developers do not have a clear understanding of the user needs and preferences.

Conclusion

Based on the insights gathered from the dashboard and the root cause analysis, it is clear that app developers need to focus on updating their apps regularly to improve their performance in the Google Play Store. To drive growth and retention, app developers should aim to fix bugs, improve the user interface, and ensure that their apps meet the needs and preferences of their users.

In addition, it is recommended that apps with low reviews be optimized, particularly those in the Social, News & Magazines, and Shopping Categories. Monetization of some apps could also be considered to increase revenue and support the development and improvement of the apps.

Finally, by addressing the root cause of the problem and optimizing and monetizing apps where necessary, app developers can improve app performance, increase customer reviews, ratings, and installs, and ultimately drive growth and retention in the Google Play Store.

Appendix



Google Sheets Data Set for Google Play App Performance Project

[\[Link here\]](#)

Tableau Data Visualization For Google Play App Performance Project

[\[Link here\]](#)