

Data Visualisation Assignment

Dharmarlou Bowen,
student number: 20123076,
total word count: 1400

May 14, 2022

Contents

1	Abstract	4
2	Introduction	5
2.1	Summary	5
2.2	Highlight	5
2.3	Aim	5
2.4	Achievements	5
2.5	Organised	5
3	Motivation and Objectives	6
3.1	Questions about Dataset	6
4	Experimental Results	8
4.1	Categorical Breakdown	8
4.2	Correlation of Ratings	10
4.3	Pricing Strategy	13
4.4	Size Strategy	15
5	Summary	18
6	References	19

List of Figures

1	Most Popular Categories	9
2	Average App Rating	11
3	Average App Rating per Genre	12
4	Free vs Paid	14
5	Light vs Bulky	16
6	Top 2 Bulkiest Apps	17

1 Abstract

The Google App Store has over a variety of 3.45 million apps a user is able to choose from with it being Apple's main App store competitor. With such a huge pool to pick from it must be a challenge for an app to stand out and become popular. Is it due to some niche genre in the mobile market as there is great app saturation? or Do popular apps share certain features or demographics that make them stand out? These are a few of the questions used to initiate my analysis.

Different questions can arise as you may be looking from the perspective of a consumer or an app developer. If you're developing an app being able to identify why apps fail or why others were successful can effect your whole development process will allow for amendments to the development process. A consumer on the other hand would be more concerned about the reviews of an app.

2 Introduction

2.1 Summary

This report is based on the Google App Store's Apps. The Play Store apps data has enormous potential to drive app-making businesses to success. Actionable insights can be drawn for developers to work on and capture the Android market.

2.2 Highlight

This subsection highlight intriguing observations that are discovered in the report:

1. Android Market Share
2. Correlations of Ratings
3. Pricing Strategy

2.3 Aim

The core aim of this report is acquire insight into the traits and attributes of successful apps. We have seen many apps rise and fall over the last few years so developing analytical insight into factors such as the category, price, size and rating.

2.4 Achievements

In this report the reader can find insight into the categories of apps that are currently dominating the android mark as well as how their ratings intertwine with all their attributes which result in their success.

When it comes to taking into account the size of the app the is a moderate factor due to what figure 5 and 6 demonstrates as two of the largest categories mirror the the holistic overview in fig.5 despite the high variation of app sizes a majority still perform well from a rating perspective as there condensed between the 4 to 5 star region.

2.5 Organised

Structure of this report is:

1. Title
2. Subtitle (if applicable)
3. Description
4. Graph (graphs are on own page)

3 Motivation and Objectives

Smartphone make up a significant part of everyday life so it makes sense that the applications we use on them play a even greater role. By developing a better comprehension for what makes these apps popular an successful it may provide invaluable insight.

The data used is from Kaggle, which is free and accessible to anyone, the dataset's name is 'Google Play Store Apps'. The dataset has 10804 rows and 13 columns. Columns from dataset:

1. App
2. Category
3. Rating
4. Reviews
5. Size
6. Installs
7. Type
8. Price
9. Content Rating
10. Genres
11. Last Updated
12. Current Ver
13. Android Ver

3.1 Questions about Dataset

This is a section of questions that we are looking to be answered with expected results:

1. Q: What categories present the most significance in the mobile market?
E: Most likely Games, Communication(Social Media) and Finance
2. Q: Is there any apps that present a high variance in rating?
E: A majority of app are probable to present a shared trend between a particular rating
3. Q: Are there any pricing trends of Apps across varying categories?
E: Apps are generally free with some services offering some sort of subscription service so this may technically vary.

4. Q: How much does an App's size impact rating?
E: As people phone have limited storage consumption of more space can be more detrimental to a apps success
- Q = Question E = Expectation

4 Experimental Results

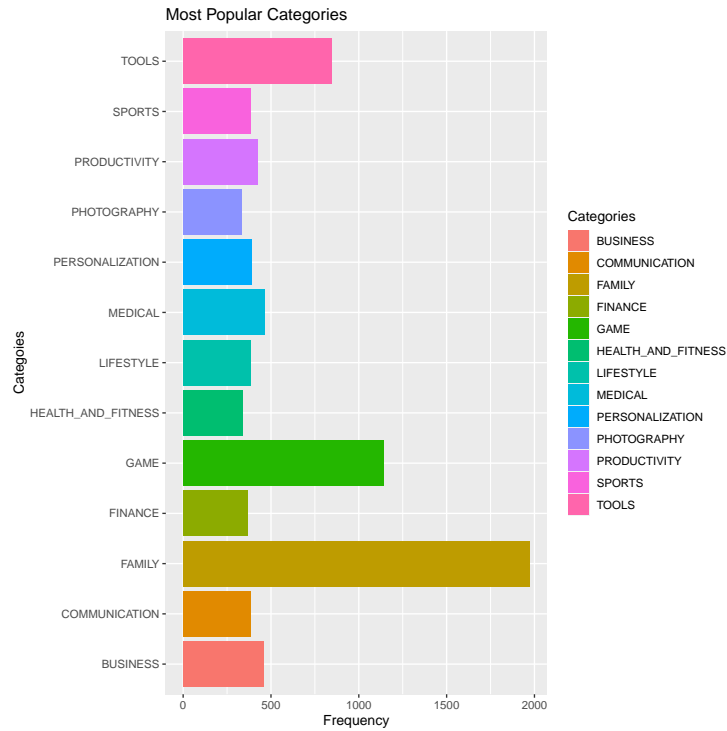
This section is research based. The questions posed in the previous sections are presented and developed upon here with each page either displaying a graph or the topical insight from the analysis.

4.1 Categorical Breakdown

This first sub-section of Experimental result performs a categorical breakdown of the categories found on the Google app store. An expectation was noted for the first question, that the most predominant categories would be Games, Communication(Social Media) and Finance.

A histogram was used to depict top 13 most popular categories based on their occurrence in the dataset. 'Family' and 'Game' apps proved to be present with the greatest market share, with Tools, Business and Medical apps are also catching up.

Figure 1: Most Popular Categories



4.2 Correlation of Ratings

This section attempts to answer the section question posed "Is there any apps that present a high variance in rating?" Two graphs were used to demonstrate the averages. Fig.2 displays a holistic overview of app rating resulting in a average rating of 4.17. In addition to this Fig.3 presents a faceting of Fig.2 provide insight into if there is any particular Genre that performs better or worse on average. From Fig.3 there is not a great amount of variance as all the facets share the trend of most of their ratings ranging between 4 to 5.

Figure 2: Average App Rating

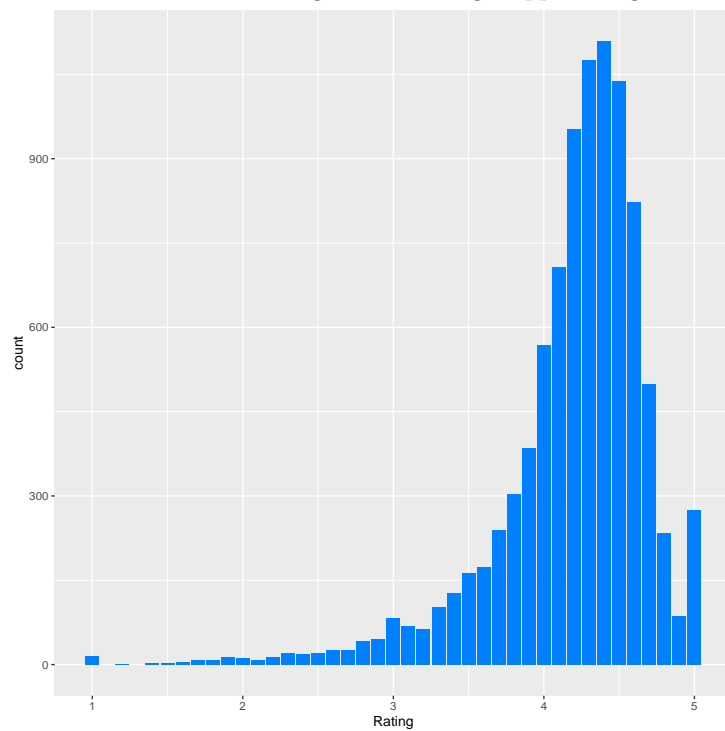
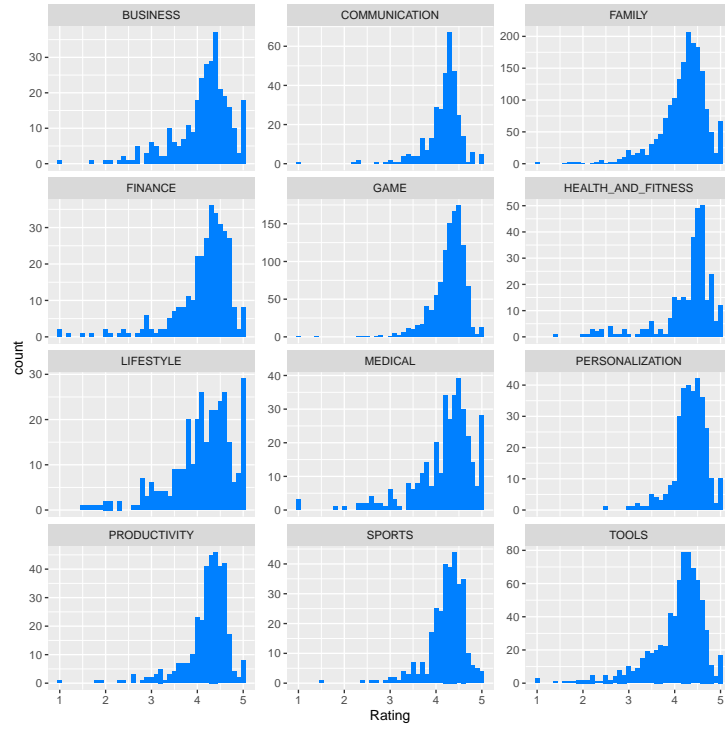


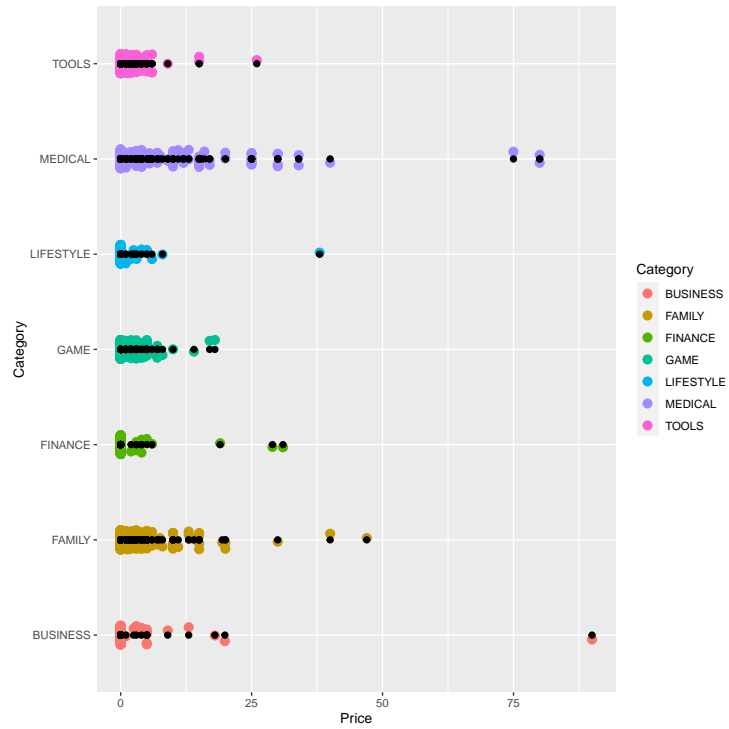
Figure 3: Average App Rating per Genre



4.3 Pricing Strategy

When considering pricing strategy it would be asserted that apps are generally free given the saturation of the mobile market. A scatter graph (Fig.4) in combination with jitter presents 7 of the most popular application categories prices. Despite prices exceeding the 50 pounds the two largest categories (Game and Family) stay below said price. The jitter in the graph was used to provide clarity to the black points that are condensed near the 0 point.

Figure 4: Free vs Paid



4.4 Size Strategy

The last factor taken into account is app size as storage capacity can act as a potential push factor if a app consumes too much space. Fig.5 provides a holistic overview of all the apps sizes and ratings. (Size is in MB) At 37.5MB with ratings of 3.2 and greater are common. Fig.6 Provides an emphasis on Game and Family with them being the largest categories. Most bulky apps (>50MB) belong to the Game and Family category. Despite this, these bulky apps are fairly highly rated indicating that they are bulky for a purpose.

Figure 5: Light vs Bulky

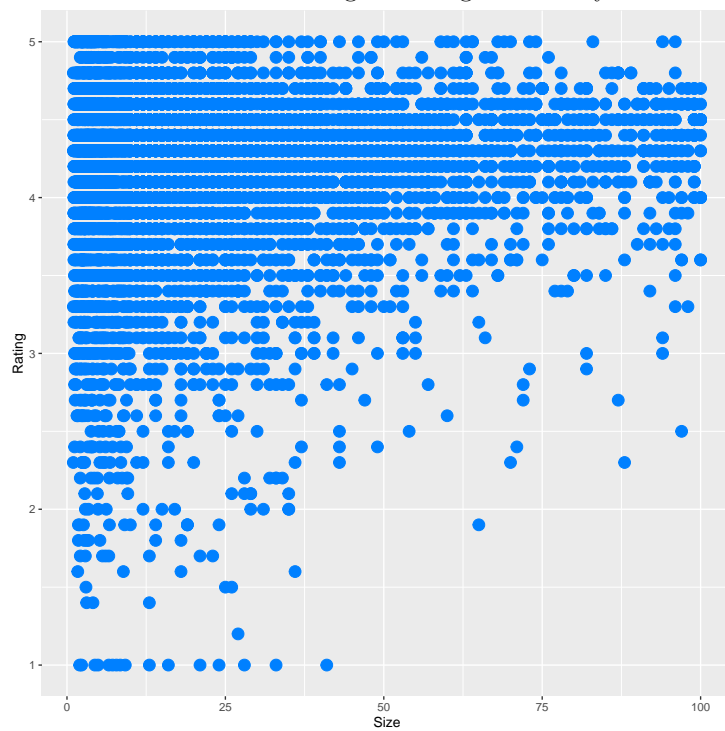
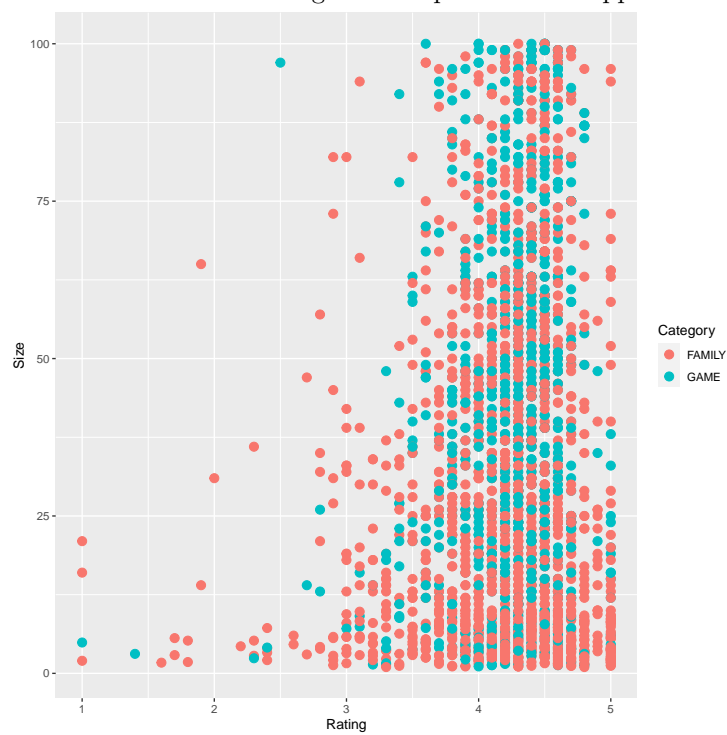


Figure 6: Top 2 Bulkiest Apps



5 Summary

This report has searched for insight into what factors can effect the success of an app based on certain shared trends among popular apps taking into account factors such as App size, rating, market share and price. As each factor was analysed certain expectations were met, with many of the graphs providing a degree of confirmation for our questions providing the conclusion that:

1. Free apps or apps below the price of 25 pounds greatly dominate the app market with the exception of a few outliers. Due to market saturation with free apps the way that these apps make a return is through subscription service or micro transactions once the app has been downloaded.
2. When it comes to taking into account the size of the app there is a moderate factor due to what figure 5 and 6 demonstrates as two of the largest categories mirror the holistic overview in fig.5 despite the high variation of app sizes a majority still perform well from a rating perspective as there is condensed between the 4 to 5 star region.

Findings:

1. Q: What categories present the most significance in the mobile market?
E: Most likely Games, Communication(Social Media) and Finance
F: 'Family' and 'Game' apps proved to be present with the greatest market share, with Tools, Business and Medical apps are also catching up.
2. Q: Is there any apps that present a high variance in rating?
E: A majority of apps are probable to present a shared trend between a particular rating
F: There is not a great amount of variance as all the facets share the trend of most of their ratings ranging between 4 to 5.
3. Q: Are there any pricing trends of Apps across varying categories?
E: Apps are generally free with some services offering some sort of subscription service so this may technically vary.
F: Despite prices exceeding the 50 pounds the two largest categories (Game and Family) stay below said price. The jitter in the graph was used to provide clarity to the black points that are condensed near the 0 point.
4. Q: How much does an App's size impact rating?
E: As people's phones have limited storage consumption of more space can be more detrimental to an app's success
F: Most bulky apps (>50MB) belong to the Game and Family category. Despite this, these bulky apps are fairly highly rated indicating that they are bulky for a purpose.

Q = Question E = Expectation F = Findings

6 References

Here are my results. Referring to section

Appinventi (2022). Top Google Play Store Statistics 2022 You Must Know.
Appinventi Available at:
<https://appinventiv.com/blog/google-play-store-statistics/> [Accessed 09 May. 2022].

Kaggle (2022). Google Play Store Apps. Kaggle Available at:
<https://www.kaggle.com/datasets/lava18/google-play-store-apps?resource=download>