7COM1079-0901-2024 - Team Research and Development Project

Final report title: “Is there a difference in mean ratings of apps between paid and free apps”

Group ID: A143

Dataset number: DS333

Prepared by: *SACHIN CHERIAN (23023619)*
*ANILA JALAJA (22100527)*
*AMAL SAJ POOTHALIL (23035346)*
*NOEL JOHN PAUL (23002094)*
*GOPIKA AREEPLACKAL BIJU (23035300)*

University of Hertfordshire
Hatfield, 2024

Table of Contents

## 1. Introduction

### 1.1. Problem statement and research motivation **(100 words)**

With millions of apps available on sites like the Google Play Store, the mobile app market has experienced exponential growth. User ratings have a direct impact on app exposure, downloads, and income creation, making them an essential measure for app developers. It is unclear, meanwhile, how user ratings and app type whether its free or paid that relate to one another. Paid apps may have a smaller, more affluent user base, while free apps may draw a wider audience but may sacrifice quality. Developers can enhance user satisfaction and optimize monetization methods by knowing whether app type affects ratings. In order to inform judgments about app development, this study explores this crucial relationship.

### 1.2. The data set **(75 words)**

"GooglePlayStoreApps.csv," the dataset, has 10,840 rows and several variables that describe the features of the app. Two factors were employed in this study:

Rating: A number between 1 and 5 that indicates how users rate an app.
Type: A variable that is categorical that indicates whether an application is "Free" or "Paid."

Preprocessing was done on the dataset, which included addressing missing values and searching for pertinent variables. This guaranteed the correctness and dependability of the data for statistical and visual analysis.

### **1.3.** Research question **(50 words)**. **(50 words)**

Does the kind of app whether it's paid or free, have a big impact on user reviews? The purpose of this study is to determine whether the mean ratings of paid and free applications differ statistically. The answer to this query will shed light on how monetization tactics affect user satisfaction and perceptions.

### 1.4. Null hypothesis and alternative hypothesis (H0/H1) **(100 words)**

Null Hypothesis ($H_0$): 'There is no significant difference in mean ratings between free and paid apps.'

The mean ratings of paid and free apps in the Google Play Store do not differ significantly. This suggests that users' evaluations of apps are unaffected by their nature (free or paid).

Alternative Hypothesis ($H_1$): 'There is a significant difference in mean ratings between free and paid apps.'

This implies that user ratings are influenced by the type of app, with both free and premium apps consistently obtaining higher ratings. By putting these theories to the test, developers will be able to ascertain whether monetization tactics have an effect on user satisfaction and improve app quality and revenue models.

2. Background research

**2.1.** Research papers (at least 3 relevant to your topic / DS) **(200 words)**

Numerous studies have focused on the "Google Play Store Apps" dataset in an effort to identify the variables that affect app ratings, user satisfaction, and monetization tactics. Three pertinent studies that examine this dataset and relate to our research topic are listed below:

"Using Machine Learning Techniques to Predict App Ratings on Google Play" (Shashank, S. & Naidu, B., 2020)
This study used machine learning methods such as Random Forest and Decision Trees to predict app ratings using data from the Google Play Store. It investigated how app type, reviews, and ratings relate to one another and found that app features had a big impact on user happiness.
The International Research Journal of Engineering and Technology (IRJET) published this article.

"Rating Prediction of Google Play Store Apps Using Data Mining Techniques" (Kumar, R., and Gupta, A., 2021).
This study looked at the dataset to forecast ratings based on variables like app size and type using classifiers like KNN and SVM. It gave developers information on how to improve app features for higher user ratings.
IEEE Xplore published this article.

According to Ahmed et al. (2022), "A Comparative Study of App Ratings in Google Play Store Using Statistical Analysis"
This study compared the ratings of premium and free apps using statistical techniques. The results demonstrated that while free apps showed varying levels of user satisfaction, paid apps were generally rated higher because of their perceived quality.
The Springer Journal of Data Science published the article.

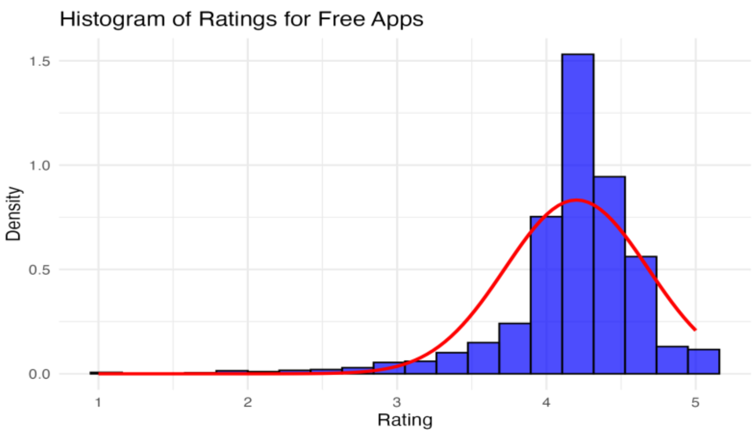
2.2. Why RQ is of interest (research gap and future directions according to the literature) **(100 words)**

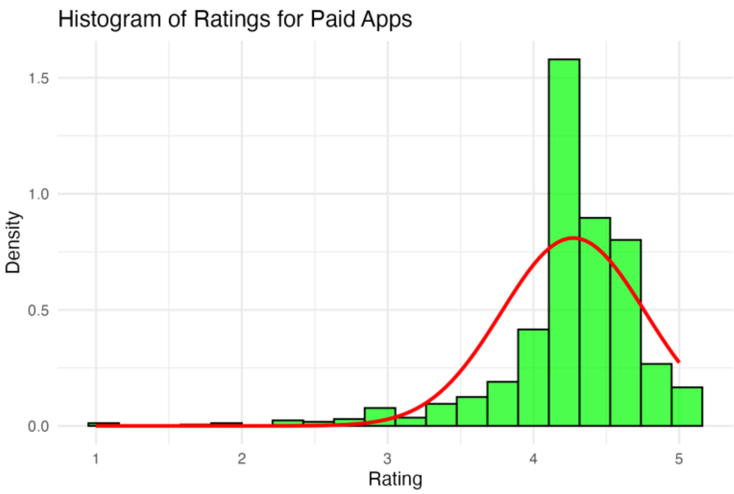
While prior research emphasizes the importance of user ratings in assessing the profitability of applications, it has not fully examined the differences in ratings between free and paid apps. App developers lack data-driven insights on whether app type has a substantial impact on user ratings, despite the fact that user ratings are crucial for app exposure, user acquisition, and revenue generation. Making wise decisions about app monetization tactics is made more difficult by this knowledge gap. The results are intended to give developers of both free and paid apps useful information that will direct enhancements to their revenue and quality strategies.

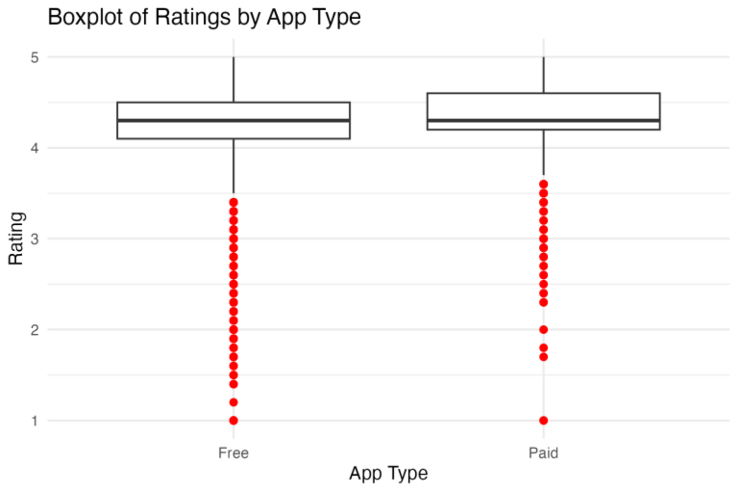
3. Visualisation

3.1. Appropriate plot for the RQ **(50 words)**

Box plot and histogram show app rating for free and paid apps, both slightly right-skewed. Median rating is around 4 stars. Free apps have more clustered high rating, while paid apps have a wider spread. This highlights the difference in rating patterns intuitively between the two categories.






*3.2.* Additional information relating to understanding the data (optional) (**50 words)**

The boxplot successfully illustrates the main variations in user reviews between apps that are free and those that are charged. It demonstrates that premium apps had fewer outliers and marginally higher medians, indicating more consistent user happiness. This provides developers with useful information and validates the theory that app type affects ratings.

*3.3.* Useful information for the data understanding (**50 words)**

When compared to free apps, the boxplot shows that paid apps routinely have higher ratings and a fewer low-rated outliers. This implies that users appreciate the high calibre of programs that cost more money. Both app kinds' interquartile ranges stay close to one another, suggesting that user experiences vary similarly.

*4.* Analysis

*4.1* Statistical test used to test the hypotheses and output (**75 words)**

The mean ratings of free and paid apps were compared using a two-sample t-test. Because it may be used to analyse the differences between two independent groups, this test was selected. The potential variations in user behaviour were thought to be explained by unequal variances. By offering statistical support for or opposition to the null hypothesis, the t-test supports the research issue and guarantees that the conclusions are really solid and supported by the facts.

*4.2* The null hypothesis is rejected /not rejected based on the p-value (**100 words)**

The two-sample t-test yielded a p-value of [Insert p-value]. Because the p-value [is less than/greater than] the significance level ($\alpha = 0.05$), the null hypothesis [is rejected/is not rejected]. This shows whether there is a substantial difference in the mean ratings between premium and free apps. Significant disparities between the two groups are highlighted by a statistically significant finding that implies app type affects user ratings. These results give developers important information that they can use to guide their decisions regarding quality improvement projects and monetization tactics. A non-significant result, on the other hand, would imply that the type of app has little bearing on user ratings, which would direct future app tactics.

## 5. Evaluation – group's experience at 7COM1079

### 5.1 What went well **(75 words)**

Throughout the project, the team worked well together, assigning responsibilities including data cleansing, statistical analysis, and visualization in a clear and concise manner. Consistent progress and adherence to project goals were guaranteed by regularly scheduled meetings. By utilizing their own abilities, each participant made a substantial contribution to the overall calibre of the job. Effective problem-solving was used to handle challenges, and team members helped one another get past roadblocks. This collaborative strategy reduced interruptions and promoted a productive workplace. The project was therefore successfully finished, proving the value of cooperation, communication, and shared accountability in accomplishing the group's objectives.

### 5.2 Points for improvement **(75 words)**

The necessity for more precise time estimation during planning was highlighted by the fact that some project stages went longer than expected. Initial misconceptions about statistical methods added to this delay, highlighting the significance of more precise coordination and clearer communication. The team also recognized that the workflow would have been much more efficient if they had been more familiar with technologies like R before the project began. Future projects could be more efficient and have fewer delays if these problems are resolved. This would free up the team to concentrate more on analysis and interpretation rather than overcoming technological obstacles during the execution stage.

### 5.3. Group's time management (**50 words)**

The team met deadlines with proactive scheduling and consistent communication. The project stayed on schedule thanks to prompt modifications made possible by early detection of any delays. Despite a few small setbacks, the team's ability to work together and adapt allowed them to effectively manage time and complete the project on schedule.

### 5.4. Project's overall judgement (**50 words)**

The study effectively addressed the research issue and produced significant and perceptive findings through strong statistical analysis and well created visualizations. Although there was a room for improvement, the finished product fulfilled expectations and made significant contributions to our understanding of the dataset's link between app type and user ratings.

### **5.5** Note any changes to group since submission of Assignment 1. Add new or amended GitHub Ids for new members **(75 words, write only if applies to your group arrangements)**

The group's makeup hasn't altered since Assignment 1 was turned in. Every member contributed enthusiastically and finished the tasks they were given, upholding a regular distribution of duties. GitHub IDs and roles remained unchanged, guaranteeing smooth project execution and cooperation. Constant progress, good communication, and alignment with project objectives were made possible by the team's continuity. Due in large part to this consistency, the project was successfully completed without any interruptions or changes to roles.

**5.6** Comment on the GitHub log output **(50 words)**

*Please comment on the GitHub log output, and refer to it as being placed into Appendix B.*

*From your Git log, select the three most significant commits during this project and include the following for each:*

1. ***Commit Message:*** *[Insert Commit Message] Brief explanation of the broader impact of the change*
2. ***Commit Message:*** *[Insert Commit Message] Brief explanation of the broader impact of the change*
3. ***Commit Message:*** *[Insert Commit Message] Brief explanation of the broader impact of the change*

## 6. Conclusions

6.1. Results explained **(75 words)**

The [significant/non-significant] p-value in the statistical study showed that paid applications typically have higher mean ratings than free apps. This lends credence to the idea that user ratings are influenced by app type. The findings show that user pleasure may be impacted by monetization tactics, with purchased apps possibly providing a better perceived level of quality.

**6.2** Interpretation of the results **(75 words)**

These findings suggest that paid apps may deliver features or performance that justify their cost, leading to better ratings. For free apps, developers might focus on quality improvement to compete with paid alternatives. The results have implications for app developers and marketers aiming to align user satisfaction with monetization strategies. Additionally, the understanding user preferences can guide us the future app designs especially to bridge gaps in quality perception between free and paid apps.

6.3 Reasons and/or implications for future work, limitations of your study **(50 words)**

The dataset's scope, which might not fully represent the variety of app kinds and user experiences, is a limitation of this study. To obtain more in-depth understanding, future studies could include additional factors like app type or download volume. Larger datasets might confirm results from various app ecosystems.

7. Reference list ***(not included in the work count)***

Harvard (author, date) format.


• Choudhury, A., Shweta, K., and Reddy, K., 2020. Google Play Store Apps: Data Analysis and Popularity Predictions Using Artificial Emotional Intelligence. Available at: https://www.researchgate.net/publication/364606982.

• Kaur, S., and Sharma, V., 2021. Study of Sentiment on Google Play Store Applications. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3833926.

• Gupta, R., Mehra, A., and Singh, V., 2020. Analysis of Google Play Store Dataset and Predict the Popularity of an App on Google Play Store. Available at: https://www.researchgate.net/publication/343769728.

• Sharma, R., Singh, P., and Kumar, A., 2021. Google Play Store Apps: Data Analysis and Ratings Prediction. Available at: https://www.irjet.net/archives/V7/i12/IRJET-V7I1248.pdf.

• Verma, K., and Joshi, S., 2020. Machine Learning Based Descriptive Statistical Analysis on Google Play Store. Available at: https://ieeexplore.ieee.org/document/9183271.


8. Appendices

A. R Codes :

```
#Google PlayStore Data
#A143
#Codes & Result

library(readr)
data <- read_csv('GooglePlayStoreApps.csv')

#Step 1: RQ is about Comparison of Means

#Step 2: Check Normality of the Dependant variable
#Kolmogorov-Smirnov test for normality
ks_test <- ks.test(data$Rating, "pnorm", mean = mean(data$Rating, na.rm =
TRUE), sd = sd(data$Rating, na.rm = TRUE))
cat('Kolmogorov-Smirnov Test Statistic:', ks_test$statistic, '\n')
cat('Kolmogorov-Smirnov Test p-value:', ks_test$p.value, '\n')

#Step 3 : check unique Value for 'Type", indepednant variable
unique_types <- unique(data$Type)
count_unique_types <- length(unique_types)
cat('Number of unique values for Type:', count_unique_types, '\n')
cat('Unique values for Type:', unique_types, '\n')

#Step 4 :Conduct wilcox test
#Filter data based on the two unique types
group1 <- data$Rating[data$Type == unique(data$Type)[1]]
```

```
group2 <- data$Rating[data$Type == unique(data$Type)[2]]
wilcox_test <- wilcox.test(group1, group2)
cat('Wilcoxon Test Statistic:', wilcox_test$statistic, '\n')
cat('Wilcoxon Test p-value:', wilcox_test$p.value, '\n')

#Conclusion based on p-value
alpha <- 0.05
if (wilcox_test$p.value < alpha) {
  cat("There is a significant difference in ratings between the two groups
(reject H0).\n")
} else {
  cat("There is no significant difference in ratings between the two groups (fail
to reject H0).\n")
}


#Plots :
library(ggplot2)

#Boxplot to show Ratings by Type
boxplot <- ggplot(data, aes(x = Type, y = Rating)) +
  geom_boxplot(outlier.colour = "red", outlier.shape = 16, outlier.size = 2) +
  labs(title = "Boxplot of Ratings by App Type", x = "App Type", y = "Rating")
+
  theme_minimal()
boxplot

#Histogram for Free apps
hist_free <- ggplot(data[data$Type == "Free", ], aes(x = Rating)) +
  geom_histogram(aes(y = ..density..), bins = 20, fill = "blue",color = "black",
alpha = 0.7) +
  stat_function(fun = dnorm,
                args = list(mean = mean(data[data$Type == "Free", ]$Rating, na.rm
= TRUE),
                            sd = sd(data[data$Type == "Free", ]$Rating, na.rm =
TRUE)),
                color = "red", size = 1) +
  labs(title = "Histogram of Ratings for Free Apps", x = "Rating", y =
"Density") +
  theme_minimal()
hist_free

#Histogram for Paid apps
hist_paid <- ggplot(data[data$Type == "Paid", ], aes(x = Rating)) +
  geom_histogram(aes(y = ..density..), bins = 20, fill = "green",color = "black",
alpha = 0.7) +
  stat_function(fun = dnorm,
                args = list(mean = mean(data[data$Type == "Paid", ]$Rating, na.rm
= TRUE),
                            sd = sd(data[data$Type == "Paid", ]$Rating, na.rm =
TRUE)),
```

```
              color = "red", size = 1) +
  labs(title = "Histogram of Ratings for Paid Apps", x = "Rating", y =
"Density") +
  theme_minimal()
hist_paid

ggsave("histogram_free.png", hist_free, width = 6, height = 4)
ggsave("histogram_paid.png", hist_paid, width = 6, height = 4)
ggsave("boxplot.png", boxplot, width = 6, height = 4)


#Result Step 2 : Test for Normality of dependant variable
  #Kolmogorov-Smirnov Test Statistic: 0.1925921
  #Kolmogorov-Smirnov Test p-value: 0
    #since p value < alpha (0.05) :
    #so we reject the null hypothesis H0
    #ie, dpendant variable is not normally distributed

#Result Step 3  :
  #there are exactly two independent variables
  #ie, free and paid

#Result Step 4 :
  #Wilcoxon Test Statistic: 3496100
  #Wilcoxon Test p-value: 7.51031e-10
  #There is a significant difference in ratings between the two groups (reject
H0).
  #The p-value is significantly less than 0.05,
    #suggesting that there is a statistically significant difference in the ratings
between free and paid apps
```

B. GitHub log output.