

# DATA VISUALIZATION – CSE3020

## REVIEW 1

TOPIC: ANALYSIS OF GOOGLE PLAY STORE APPLICATIONS

---

TEAM MEMBERS:

20BCE0362 – MUDIT JAIN

20BCE0641 – KARTIKEYA RAWAT

20BCE0684 – SIDHARTH MISHRA

20BCE0244 – SATVIK NADKARNI

20BCE0461 – ANISH DESAI

Faculty: Dr. Jyotismita Chaki

Slot: L39+L40

# ABSTRACT

Now a days, it is becoming difficult for app developers to predict the number of installs of their app or to analyze what is the reach of their apps and what is the current market demand.

The goal of this project is to provide insights that will enable developers better understand what user wants and thus promote the applications. The dataset is chosen from Kaggle. It is of 10k Play Store apps for analyzing the Android market. This dataset provides information on several applications as well as user feedback.



# INTRODUCTION

The purpose of our research is to anticipate the amount of app installs based on app information and reviews. We hope that this project will assist app creators in predicting their number of installs, as well as investors looking for the next big thing. Companies may conduct beta focus groups, or app developers may receive feedback from beta testers and receive a set number of reviews. Developers and company managers might benefit from knowing the number of installs because it allows them to forecast earnings. The outcome of this study may demonstrate the relevance of app reviews in the market, as they may be one of the determining factors for the number of installs.



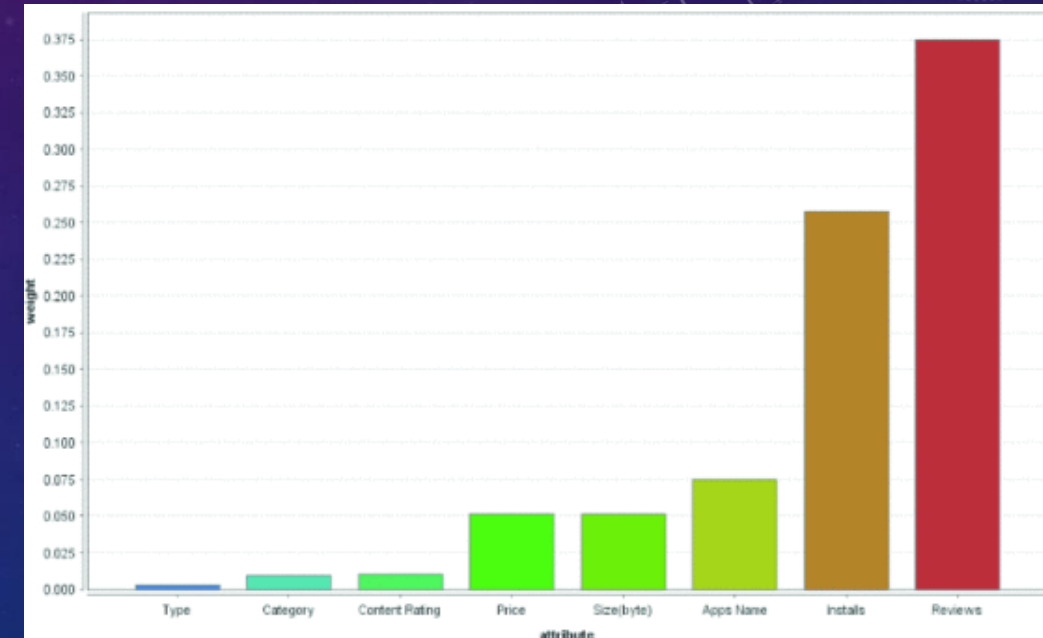
# LITERARY SURVEY

## Paper 1: Apps Rating Classification on Play Store Using Gradient Boost Algorithm

In this paper they have used bar graph data visualization to predict which criteria is most influential in ranking the applications.

### Citation:

O. Lengkong and R. Maringka, "Apps Rating Classification on Play Store Using Gradient Boost Algorithm," 2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), 2020, pp. 1-5, doi: 10.1109/ICORIS50180.2020.9320756.



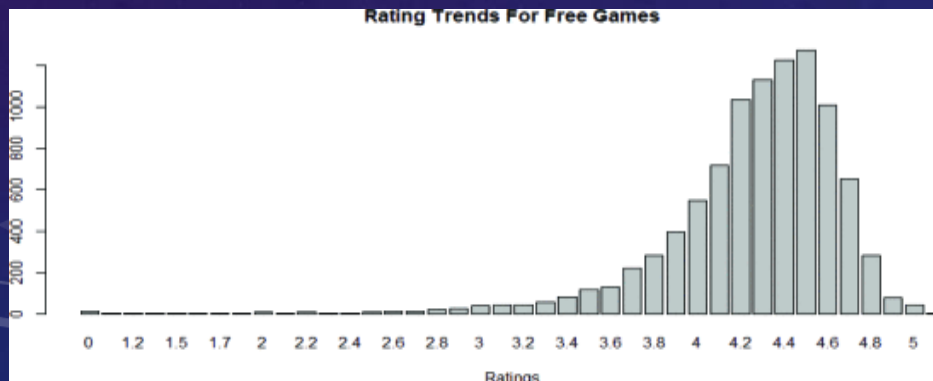
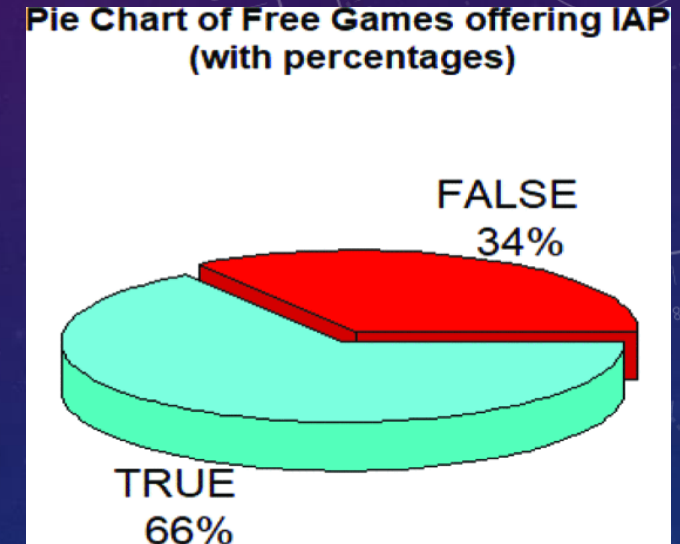


# PAPER 2: DATA SCRAPING FROM GOOGLE PLAY STORE AND VISUALIZATION OF ITS CONTENT FOR ANALYTICS

## A. Free Games InApplicationPurchases (IAP)

In Google play store many games contain such items, products, credits etc. which get by performing actions which are not relevant to the game.

The pie3D chart shows that 66% of free games offer IAP and 34% free games did not offer IAP as shown.



## B. Free Games Rating Value

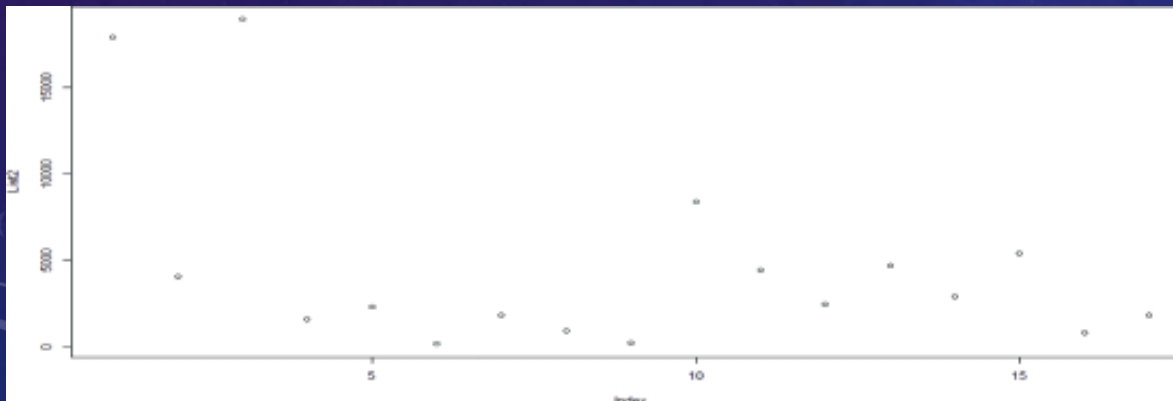
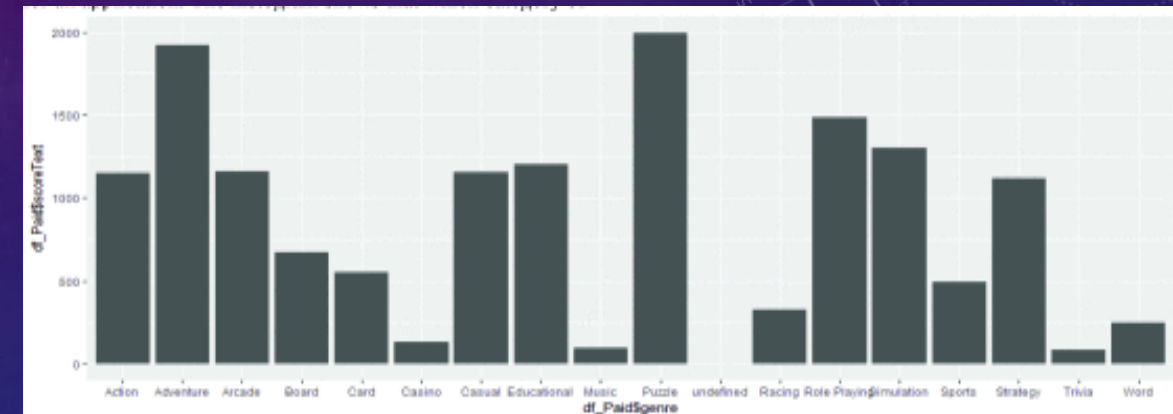
In Google play store users give ratings to the application according to their experience about that application.

It can be clearly visualized in the histogram that most people give 4.5 ratings to the free game as shown.

# PAPER 2: DATA SCRAPING FROM GOOGLE PLAY STORE AND VISUALIZATION OF ITS CONTENT FOR ANALYTICS

## C. All Categories of Paid Games Ratings

In Google play store ratings matter a lot in decision making for an application. It can be clearly visualized by a bar graph that most people give ratings to the puzzle category as shown.



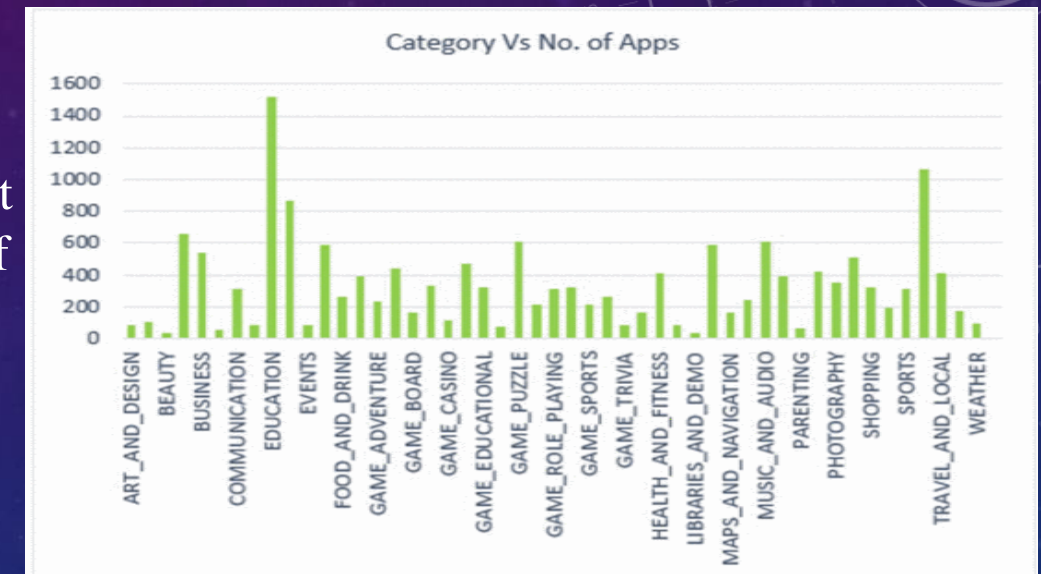
## D. All Categories of Paid Games Installs

If in Google play store find category which installs is most, guide developers to try their skills in that category. As Arcade category have a greater number of installs with respect to other categories as shown.

# PAPER 3: ANDROID APPS SUCCESS PREDICTION BEFORE UPLOADING ON GOOGLE PLAY STORE

## A. Category Vs App

The Category column in the dataset used has 34 different types of categories. Here they have shown the number of different types of apps against the categories. Figure shows the Bar Chart of the number of categories.



## B. Category Vs Average Rating

The average rating of apps are presented with a Bar Chart. It shows the Bar Chart of all category against the Average Rating.



# MOTIVATION

While many public datasets give Apple App Store data, there are very few Google Play Store app datasets available anywhere on the web. Further investigation revealed that the iTunes App Store website has a beautifully indexed appendix-like structure to facilitate web scraping. Google Play Store, on the other hand, employs advanced modern-day techniques (such as dynamic page loading) that make scraping more difficult. Hence, we decided to take this dataset and tried to make it less complex in an organized manner accompanied by data visualization.



# PROPOSED APPROACH

Discussion of Google play store dataset will involve various steps such as:

- Loading the data into data frame
- Cleaning the data
- Extracting statistics from the dataset
- Exploratory analysis and visualizations
- Questions that can be asked from the dataset
- Conclusion

# REFERENCES

- <https://ieeexplore.ieee.org.egateway.vit.ac.in/document/9320756>
- <https://ieeexplore.ieee.org.egateway.vit.ac.in/document/8673523>
- <https://ieeexplore.ieee.org.egateway.vit.ac.in/document/9068071>
- <https://dl.acm.org/doi/abs/10.1145/3152494.3152500>
- <https://ieeexplore.ieee.org/abstract/document/7045413>
- <https://www.tableau.com/learn/articles/data-visualization>
- <https://www.ibm.com/analytics/data-visualization>
- <https://towardsdatascience.com/data-science-a-deep-analysis-on-google-play-store-apps-from-kaggle-8283bbc508b0>