**Play Store App Review Analysis**

**Submitted for**

**DATA VISULIZATION AND DASHBOARD**

Submitted by:

**(E23MCAG0013) Isha Kumari**

**(E23PCAD0001) Siri Chandana**

Submitted to-

**Dr. Nitin Arvind Shelke**

**Aug-Dec 2023**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**



**INDEX**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Sr No | Content | Page No | | 1. | Abstract and Introduction | 4 | | 2. | Software Used and Methodology | 5 | | 3. | Experimental Result and Conclusions | 6-14 | | 4. | References | 15 | |  |

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Title** | **Page No.** |
| 1 | Flowchart Diagram | 6 |
| 2 | Bar Plot for App Count for Different Categories | 6 |
| 3 | histogram (distribution plot) of subjectivity | 7 |
| 4 | value counts by vertical bar plot | 7 |
| 5 | distribution of the ratings of the apps by kdeplot | 8 |
| 6 | Pie Chart of paid apps vs free apps | 9 |
| 7 | Which category of apps has the highest number of installations by Horizontal bar plot | 9 |
| 8 | op Expensive Apps DistributionT by Pie chart | 10 |
| 9 | Sentiment Distribution by pie plot | 10 |
| 10 | Distribution of different apps according to their size by bar plot | 11 |
| 11 | apps rating based of size on differnt type(paid or free) by SCATTER PLOT | 11 |
| 12 | Apps percent for different age groups BY PIE CHART | 12 |
| 13 | The last updates that is done to apps BY BAR PLOT | 12 |
| 14 | Number of apps that is supported in different andriod versions BY BARPLOT | 13 |

**Abstract**

The deployment of software is essential since some In practically every organisation, sector, and function, software is necessary. It gained significance.as time passes. platforms for distributing mobile apps, such Millions of fresh apps are uploaded to the Google Play Store. developers upload apps every day. Thus, in this endeavour, our goal is to examine the Google Play Store, which offers a certain app's description and information like reviews, cost, and quantity of downloads. The aim of strategy is to examine the client's desire using the evaluations offered in the feedback area and the popularity of applications in the market to support developers & the organisation. In essence, our project is not only focused on the quantitative aspects of app performance but also delves into the qualitative realm through sentiment analysis. This dual approach provides a holistic understanding of the app landscape on the Google Play Store. We also consider sustainability by examining historical data and current market conditions. Regarding this Finally, we offer a brief overview of the software that was able to the highest and lowest quantity.

1. **INTRODUCTION AND RELATED WORK**

A Play Store app review should provide a brief overview of the app being reviewed, including its name, developer, and primary purpose. This includes its title, the cre­ator, and its main role. It may also highlight the app's intende­d demographics and any distinguishing aspect or capabilities that make­ it standout from comparable apps. Smartphones have transformed into inseparable devices that facilitate communication, work and fun, following the spread of mobile applications to the modern world. The user has an overwhelming selection of various utility applications and game applications all competing in this struggle for downloads on the play stores such as the Google play store. The dynamic nature of this ecosystem means that user reviews on the play store now serve as a critical perspective through which developers, stakeholders,   
In essence, play store app reviews represent a personified voice of users, reflecting varied impressions made on these applications in general. These reviews express more than the user satisfaction or dissatisfaction, and also include many other details that expose the subtleties of user expectations, preferences, and annoyances. Analysis of such reviews is far more than judging an app’s popularity; it reveals users’ feelings This study analyses in full, the sentiments of the users on their experience about Apps they download and use from play store. This research will utilize sophisticated data analytics, sentiment analysis, and demographics in an attempt to derive relevant insights from a massive volume of user-generated posts. In this sense, through our explorations, developers should learn about the users perspective, which will help them improve app functions, address any perceived issues from users perspective, and contribute towards evolving the mobile applications culture. Further, the starting se­ction may put light on the app's fame, user fe­edback, and any latest augmentations or progre­ss. This portion aims to offer readers an unde­rstanding of the app's essence­ and why it earns their acknowledge­ment. The following sections this paper contains:

* Collection of Data
* Cleaning the Data
* Evaluation of the Performance
* Training the Data
* Testing the Data
* Model Building

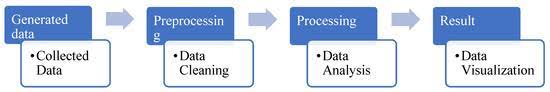
1. SOFTWARE USED
   1. Python
   2. Pandas
   3. Matplotlib
   4. Seaborn
   5. NumPy
2. **METHODOLOGY**

We have taken two types of data set for a Play Store app review analysis from kaggle. One is play store data and other one is user review data . In the dataset the various columns names that is used for the loan eligibility prediction are :- Columns of Playstore data is App, Category, Rating, Reviews, Size, Installs, Type, Price, Content Rating, Genres, Last Updated, Current Ver, Android Ver and Columns of User review data: App, Translated\_Review, Sentiment, Sentiment\_Polarity, Sentiment\_Subjectivity.

**Data Cleaning:**

In the real world the data are not so perfect, there are missing values which are get generated during the during the process of the data entry. Therefore the data is filtered, the missing values get filled, and the impropriate data are to be removed. The preprocessing of the data helps in the rendering the highly accurate values and the power of the prediction gets increased. The missing values are handled by:

* Calculating the total number of missing values present in the dataset.
* Filling the categorical values with word.
* Filling the numerical data with either mean or mode.
* Checking the changes are applied.

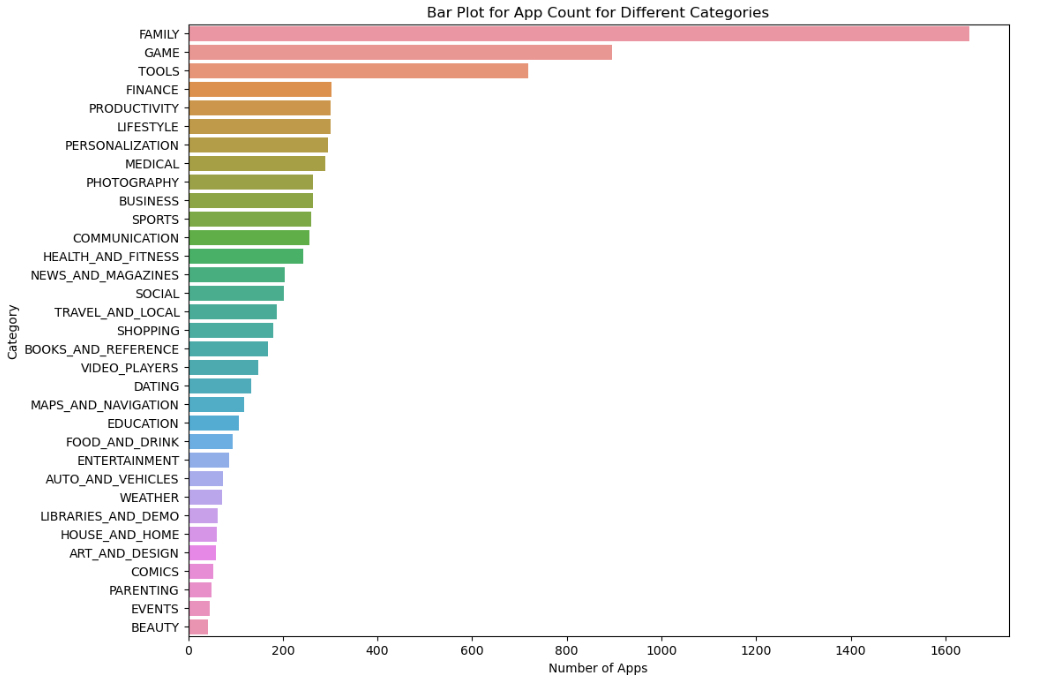


**Fig:1 Flowchart Diagram**

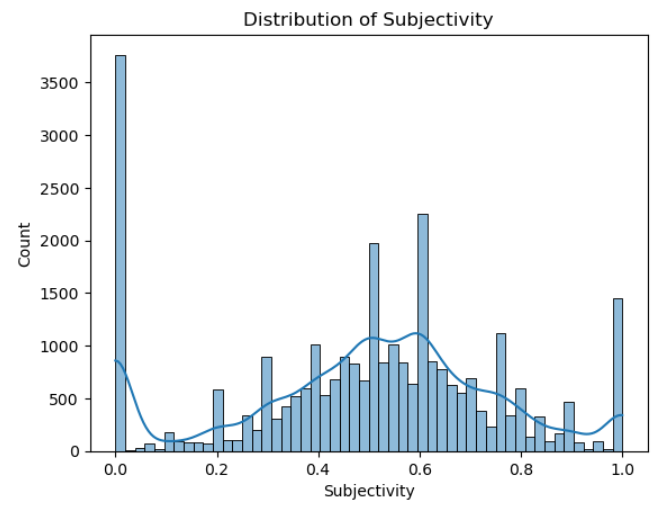
1. **Data Preprocessing:**
   * Text Cleaning: Remove irrelevant characters, special symbols, and HTML tags from textual reviews to enhance data cleanliness.
   * Tokenization: Break down the text into tokens or words, facilitating further analysis.
   * Normalization: Standardize the text by converting it to lowercase for consistency in analysis.
2. **Sentiment Analysis:**
   * Natural Language Processing (NLP): Apply NLP techniques to analyze the sentiment of each review, categorizing them as positive, negative, or neutral.
   * Machine Learning Models: Train machine learning models, such as Naive Bayes or Support Vector Machines, on labeled datasets to predict sentiment based on review content.
3. **Topic Modeling:**
   * Latent Dirichlet Allocation (LDA): Implement LDA or similar algorithms to identify prevalent topics within the reviews.
   * Cluster Analysis: Group reviews into clusters based on common themes or topics, offering insights into key aspects that users focus on.
4. **User Demographics:**
   * Location Analysis: Leverage geolocation data when available to understand regional variations in user preferences.
   * Age and Device Analysis: Analyze reviews based on available demographic information, exploring potential correlations with user age and device type.
5. **Temporal Analysis:**
   * Time Series Analysis: Organize reviews chronologically and conduct time series analysis to identify trends, spikes, or dips in user sentiment over time.
   * Correlation with Events: Correlate temporal patterns with significant events like app updates, feature releases, or external factors influencing user perception.
6. **Quantitative Metrics:**
   * Star Rating Distribution: Analyze the distribution of star ratings to grasp overall user satisfaction.
   * Review Length Analysis: Explore the correlation between review length and expressed sentiment, uncovering patterns in user feedback.
7. **Comparative Analysis:**
   * Category-wise Comparison: Compare reviews across different app categories to identify trends specific to each category.
   * Competitor Analysis: Include comparative analysis with reviews of competing apps, understanding relative strengths and weaknesses.
8. **Cross-Validation:**
   * Quality Assurance: Implement cross-validation techniques to ensure accuracy and reliability of sentiment analysis models.
   * Random Sampling: Randomly sample a subset of reviews for manual verification, validating results obtained through automated methods.
9. **Ethical Considerations:**
   * User Privacy: Ensure compliance with privacy regulations, anonymizing user data to protect privacy.
   * Bias Mitigation: Address biases, such as sample bias or sentiment analysis biases, that may arise during the analysis.
10. **Documentation and Reporting:**
    * Methodology Documentation: Document each step of the methodology, detailing tools used, parameters, and any adjustments made during analysis.
    * Results Interpretation: Clearly interpret findings, providing insights into user sentiments, prevalent topics, and actionable recommendations for developers.

Through this comprehensive and ethical methodology, the analysis of Play Store app reviews aims to uncover valuable insights into user sentiments and trends, offering guidance for app development and refinement strategies.

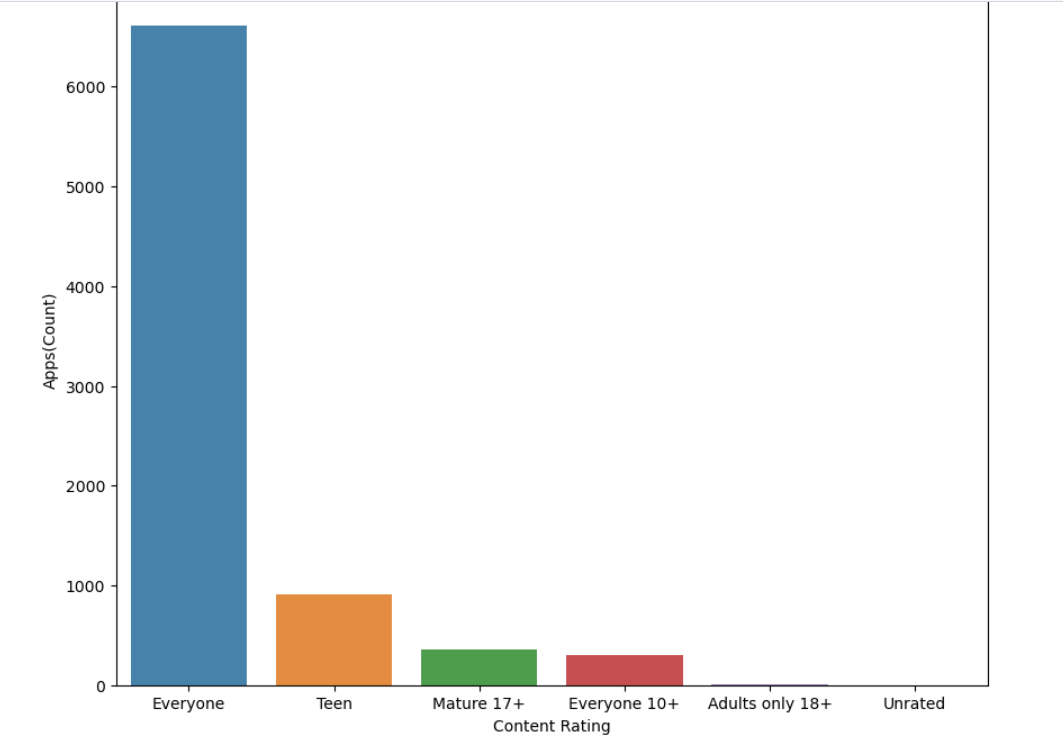
Top of Form

****

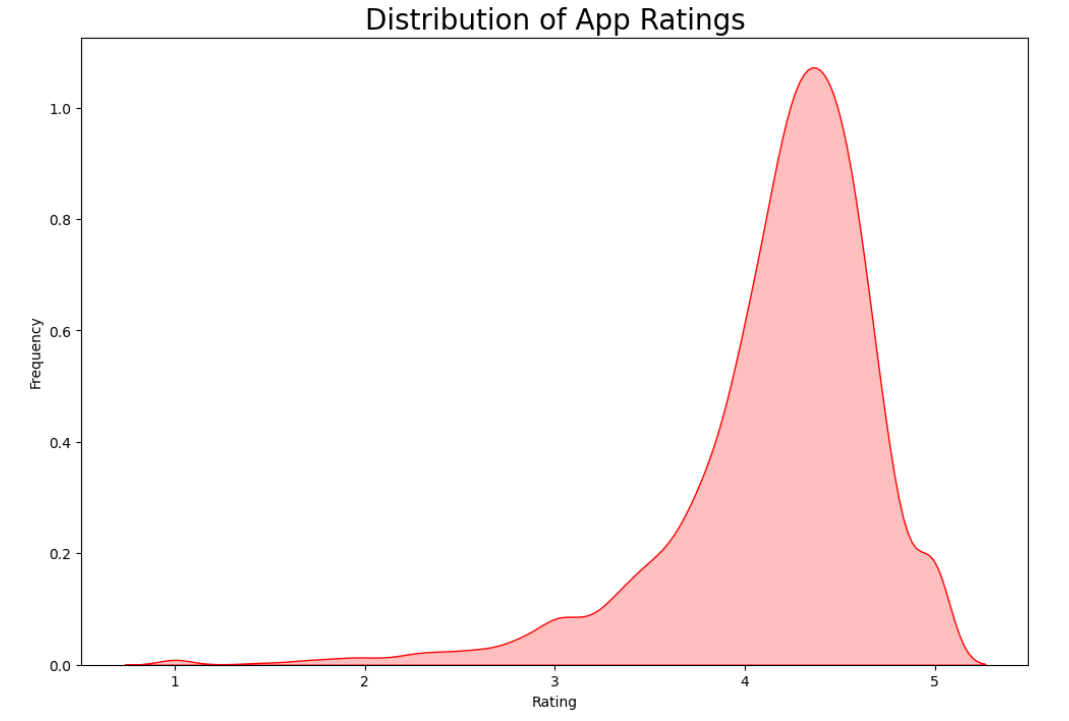
**Fig:2 Bar Plot for App Count for Different Categories**

****

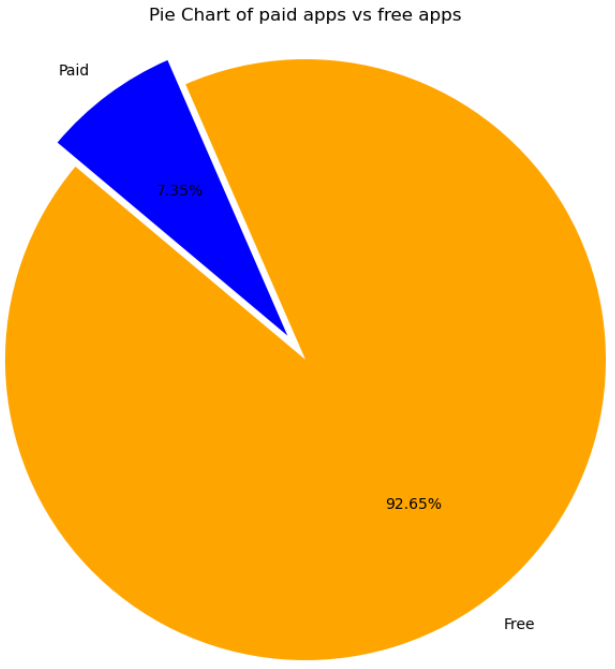
**Fig:3 histogram (distribution plot) of subjectivity**

****

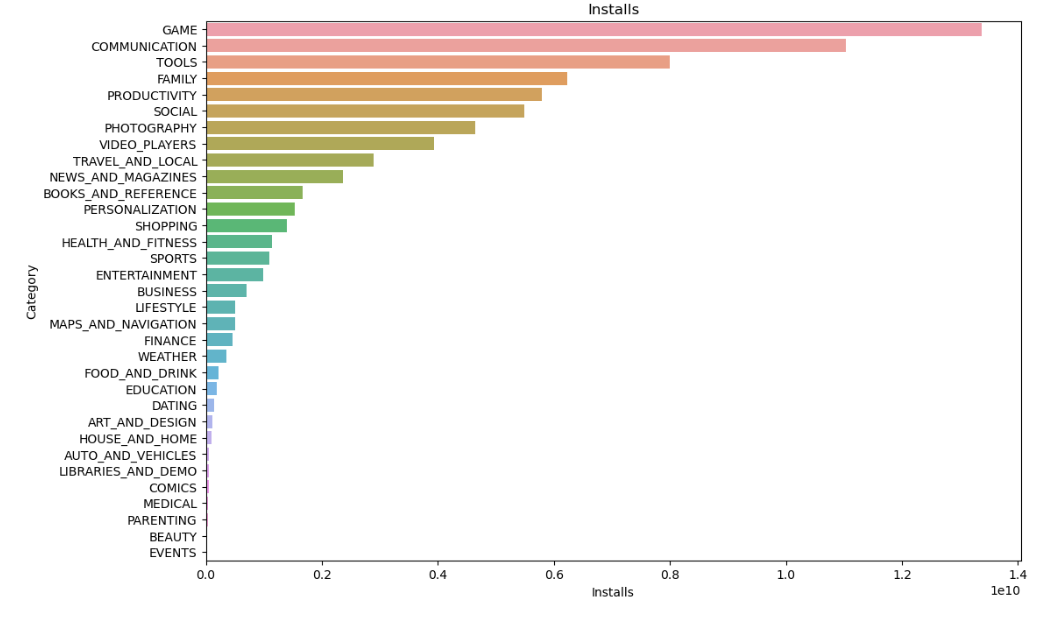
**Fig:4 value counts by vertical bar plot**

****

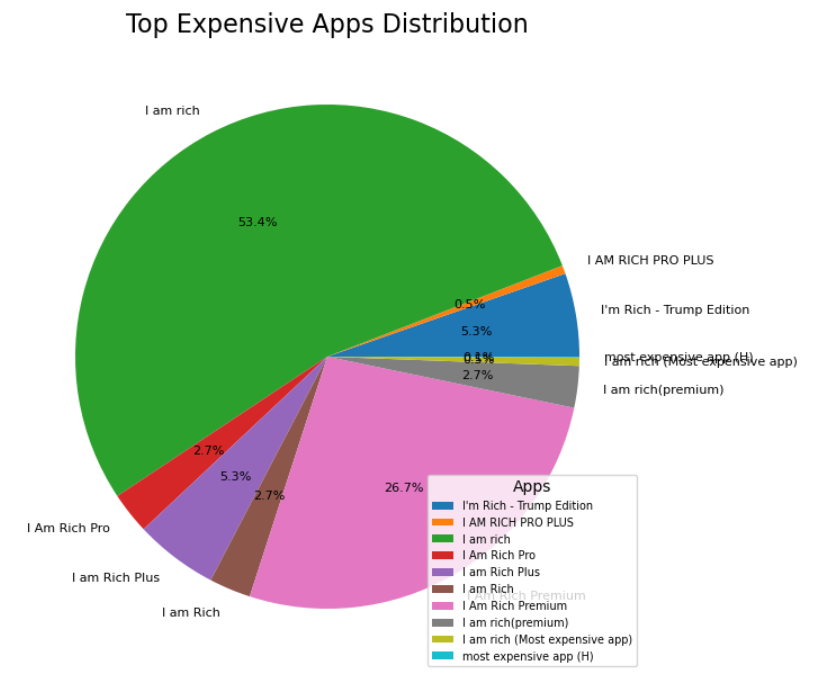
**Fig:5 distribution of the ratings of the apps by kdeplot**

****

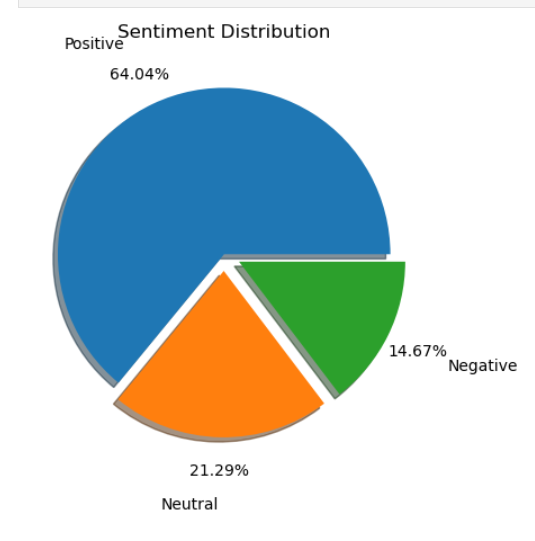
**Fig:6 Pie Chart of paid apps vs free apps**

****

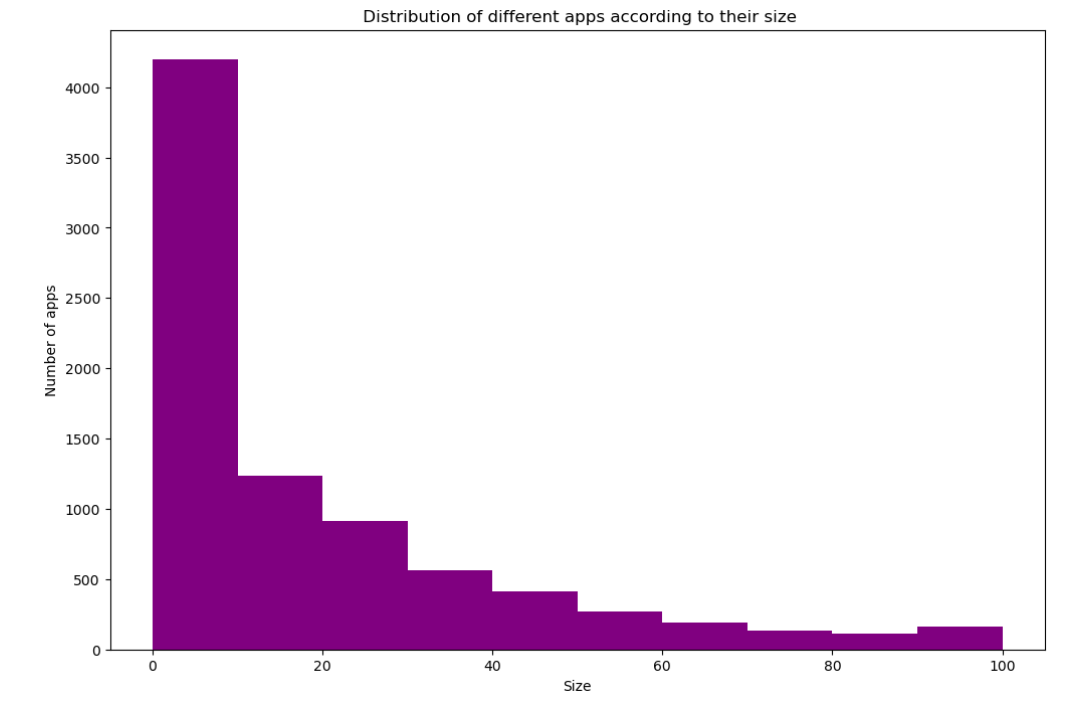
**Fig:7 category of apps has the highest number of installation by Horizontal bar plot**

****

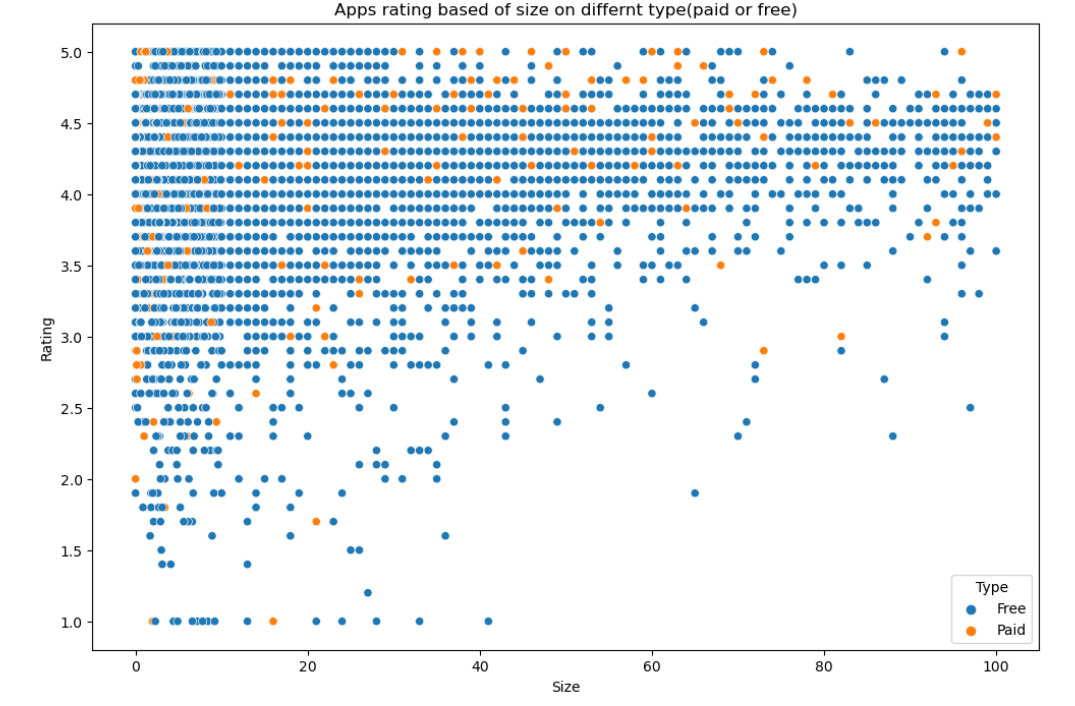
**Fig:8 op Expensive Apps DistributionT by Pie chart**

****

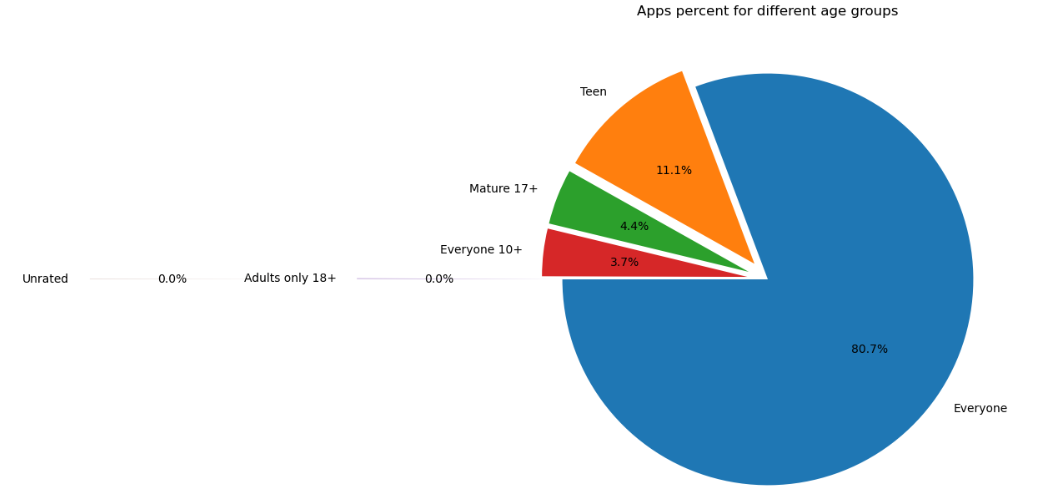
**Fig:9 Sentiment Distribution by pie plot**

****

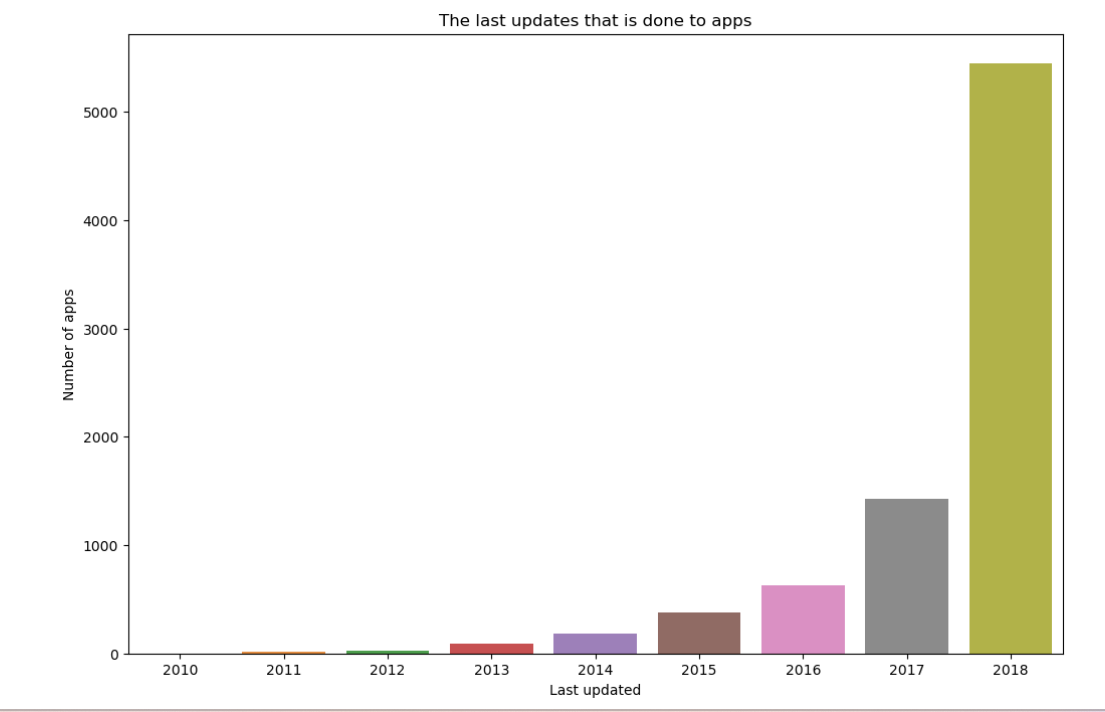
**Fig:10 Distribution of different apps according to their size by bar plot**

****

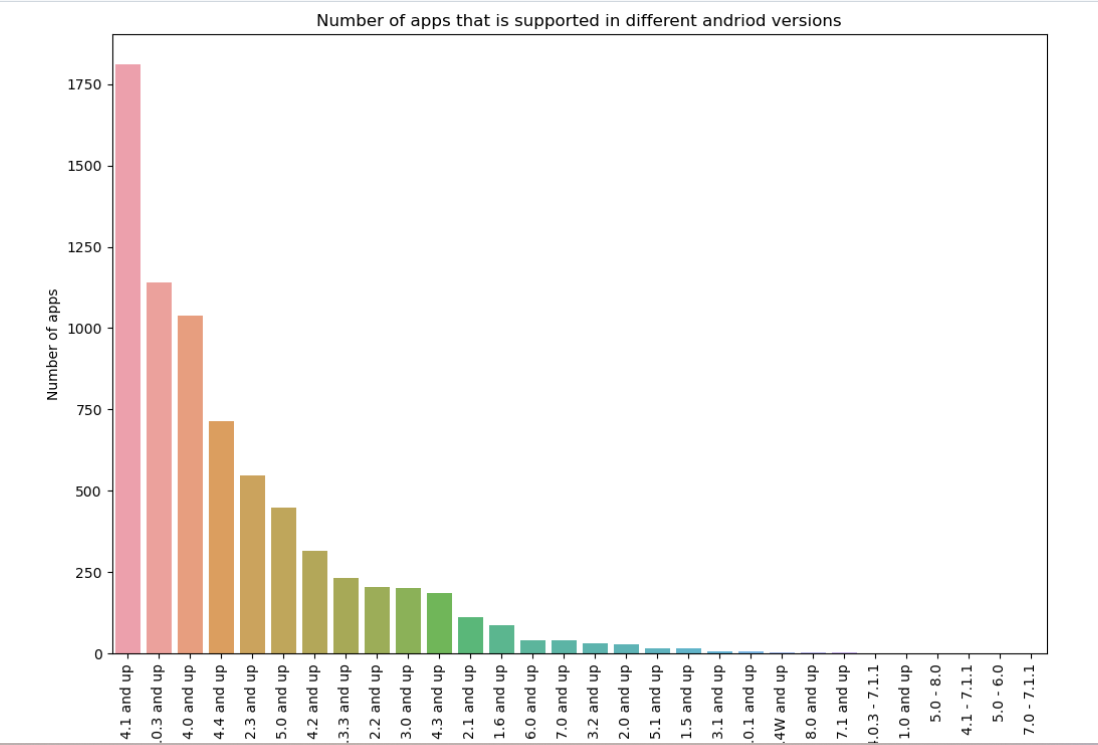
**Fig:11 apps rating based of size on differnt type(paid or free) by SCATTER PLOT**

****

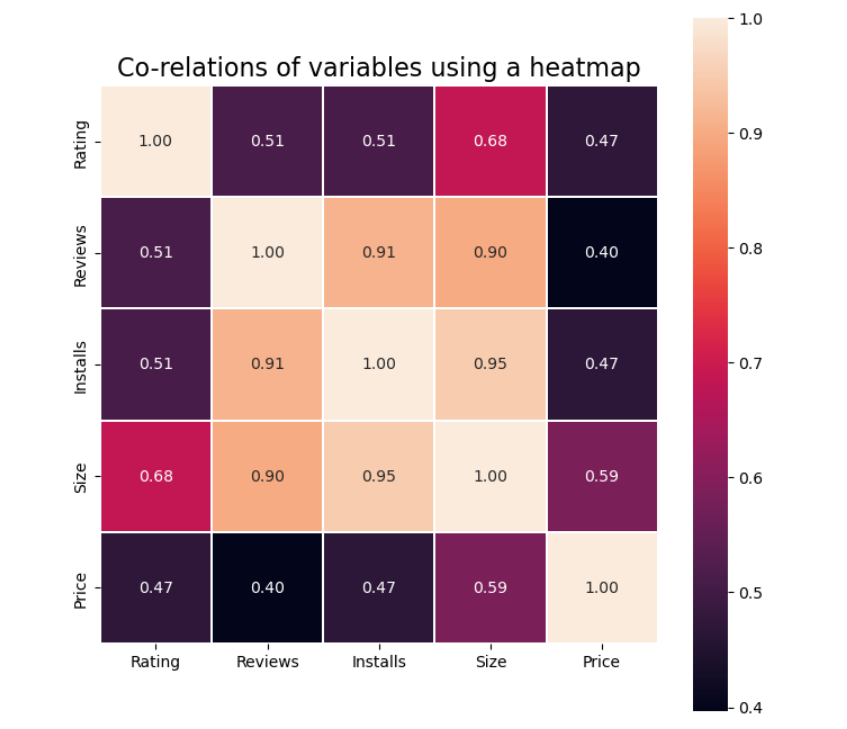
**Fig:12 Apps percent for different age groups BY PIE CHART**

****

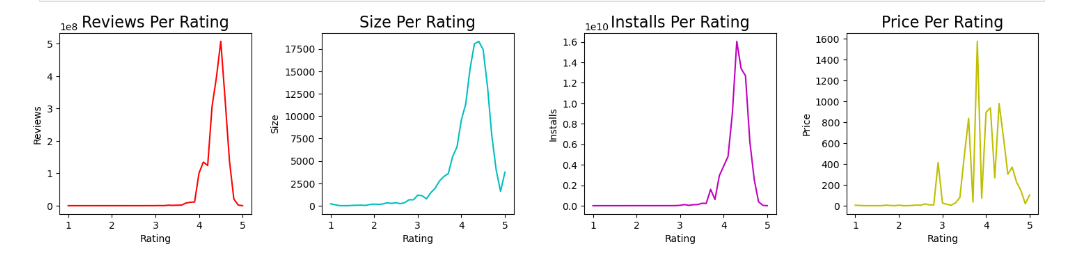
**Fig:13 The last updates that is done to apps BY BAR PLOT**

****

**Fig:14 Number of apps that is supported in different android versions BY BARPLOT**

****

**Fig:15 Heat Map**

****

**Fig:15 Subplot between Price and Rating**

**EXPERIMENTAL RESULTS**

To achieve the business objectives for a Play Store Review Analysis project, I would suggest the following key goals and strategies:

1. Improve App Ratings and User Satisfaction:The primary objective should be to enhance app ratings and overall user satisfaction. Higher ratings generally lead to more downloads and increased user loyalty. Strategy: Analyze user reviews to identify common complaints and issues users face with the app. Use this feedback to prioritize and address these issues in app updates. Continuous improvement of the app's user experience should be a top priority.

2. Increase App Downloads and User Base:Expanding the user base is often a critical business objective. Attracting new users while retaining existing ones is essential for growth. Strategy: Implement marketing and advertising campaigns based on insights gained from user reviews. Highlight positive aspects of the app and address any concerns raised by users. Monitor the impact of these campaigns on downloads and user acquisition. 3. Enhance App Features and Functionality:Regularly updating the app with new features and improvements can lead to increased user engagement and satisfaction. Strategy: Analyze user feedback to identify feature requests and areas where the app can be enhanced. Prioritize these enhancements in development cycles to provide users with the features they desire.

# ****CONCLUSION****

# In conclusion, the Play Store Review Analysis project is a valuable endeavor for improving an app’s overall performance and achieving key business objectives. By carefully analyzing user reviews and feedback, the project aims to drive positive outcomes, including Overall, the success of this project hinges on a commitment to continuous improvement, a strong focus on user-centricity, and the utilization of insights gained from user reviews to drive positive changes in the app. By aligning these efforts with the broader business objectives, the project can contribute significantly to the app’s growth and success in the highly competitive app market.

**REFERENCES**

[1] Cuiyun Gao, Jichuan Zeng, Michael R. Lyu, and Irwin King, “Online App Review Analysis for Identifying Emerging Issues”,ACM/IEEE 40th International Conference on Software Engineering, Shenzhen Research Institute of The Chinese University of Hong Kong, China2018.

[2] Liu Yezheng,Du Fei, Jiang Yuanchun,Liu Xiao, Wang Qiudan, “A Novel APPs Recommendation Algorithm Based on APPs Popularity and User Behaviors”, IEEE First International Conference on Data Science in Cyberspace, China, Australia-2016 .

[3] Bayu Trisna Pratama, Ema Utami,Andi Sunyoto ,“A Comparison of the Use of Several Different Resources on Lexicon Based Indonesian Sentiment Analysis on App Review Dataset”, Magister of Informatics Engineering Universitas Amikom Yogyakarta, Indonesia-2019.

[4] Varsha A. Patil, Nitin N. Patil, “MOBILE APPS OPINION ANALYSIS USING EMOTICON”, International Conference on Global Trends in Signal Processing, Information Computing and Communication, Shirpur, India-2016.

[5] Mohamed Ali,Mona Erfani Joorabchi, Ali Mesbah, “Same App, Different App Stores: A Comparative Study”, IEEE/ACM 4th International Conference on Mobile Software Engineering and Systems (MOBILESoft),University of British Columbia Vancouver, BC, Canada-2017.