

Capstone Project Submission

Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1) Satyam Jyoti Sankar

E-mail: satyamjyoti99@gmail.com

- Data visualization.
- Approach towards plain.
- Data sorting.
- Bar plot and Heat map.
- PPT presentation.
- Project summery template.

2) Krushnagopal Brahma

E-mail: krbrahma54@gmail.com

- Data analysis.
- Frame work of project.
- Debug all Errors
- Pi-plot and Histogram plot
- Sample PPT.
- Technical documentation.

Problem definition:

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. Each app (row) has values for category, rating, size, and more. Another dataset contains customer reviews of the android apps. Objective of the project to Explore and analyze the data to discover key factors responsible for app engagement and success.

EDA on given Data set:

There are two dataset:

1. **Play Store Data**(App, Category, Rating, Review, Size, Install, Type, current rating ,genres , Last update, Current Var ,Android Var)
2. **User Review Data**(App, Sentiment ,Sentiment Polarity, Sentiment Subjectivity)

Digging into data we understand that

- There are 13 columns of properties with 10841 rows of data.
- Column 'Reviews', 'Size', 'Installs' and 'Price' are in the type of 'object'
- Values of column 'Size' are strings representing size in 'M' as Megabytes, 'k' as kilobytes and also 'Varies with devices'.
- Values of column 'Installs' are strings representing install amount with symbols such as ',' and '+'.
- Values of column 'Price' are strings representing price with symbol '\$'.

Conclusion:

The Google Play Store Apps report provides some useful details regarding the trending of the apps in the play store. As per the graphs visualizations shown above, most of the trending apps (in terms of users' installs) are from the categories like GAME, COMMUNICATION, and TOOL even though the amount of available apps from these categories are twice as much lesser than the category FAMILY but still used most. The trending of these apps are most probably due to their nature of being able to entertain or assist the user. Besides, it also shows a good trend where we can see that developers from these categories are focusing on the quality instead of the quantity of the apps.

Some important point:-

- Average rating of (active) apps on Google Play Store is 4.17.
- If we see individually app wise the communication app like Facebook and what sup get highly reviewed app it shown that people regularly active on that and give there feedback also on that.
- Medical and Family apps are the most expensive and even extend up-to 80\$.
- Users tend to download a given app more if it has been reviewed by a large number of people.
- More than half users rate Family, Sports and Health & Fitness apps positively. Apps for games and social media get mixed reviews, with 50 percent positive and 50 percent negative responses.

Please paste the GitHub Repo link.

Github Link:- <https://github.com/brahma102/Play-Store-App-Review-Analysis.git>

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)