

Capstone Project Submission

Team Member's Name, Email and Contribution:

1. Mohammad Ammaar

Email: mohammad.ammaar1@gmail.com

- Data visualization.
- Approach towards plain
- Data sorting
- Pi-plot and Histogram plot
- Sample PPT
- Data analysis.
- Frame work of project.
- Debug all Errors
- Technical documentation.
- Bar plot and Heat map.
- PPT presentation.
- Project summary template

2. Mohammad Shahzeb Khan

Email: szeb1997@gmail.com

- Data visualization.
- Data sorting
- Pi-plot and Histogram plot
- Debug all Errors
- Technical documentation.
- Bar plot and Heat map.
- PPT presentation.
- Project summary template

3. Gayatri Gupta

Email: gayatrigupta9761@gmail.com

- Data visualization.
- Data sorting
- Pi-plot and Histogram plot
- Debug all Errors
- Technical documentation.
- Bar plot and Heat map.
- PPT presentation.
- Project summary template

4. Nada Nasser

Email: nadanasser2468@gmail.com

- Data visualization.
 - Approach towards plain
- . • Data sorting
 - Pi-plot and Histogram plot
 - Sample PPT
 - Data analysis.
 - Frame work of project.
 - Debug all Errors
- . • Technical documentation.
 - Bar plot and Heat map.
 - PPT presentation.
 - Project summary template

5. Mansur Shikalgar

Email: mansurshikalgar.963@gmail.com

- Data visualization.
- Data sorting
- Pi-plot and Histogram plot
- Debug all Errors
- . • Technical documentation.
 - Bar plot and Heat map.
 - PPT presentation.
 - Project summary template

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

- 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 insights on which we worked are as follows:

- Top 10 categories which have the most number of application installations.
- Correlations check among 'Install' label and other labels of the datasets.
- Visualization using violin & density plot.
- Visualization using Histograms and Bar plots
- Insights over the content rating with the number of apps.
- Word cloud for better understanding of repetitive words
- Analyzing using a pie chart to understand the significant difference in percentage distributions of sentiments in the Data.
- Category Line plot for positive as well as negative sentiments opinion.
- Average rating of (active) apps on Google Play Store is 4.2.
- 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.

GitHub Repository link.

https://github.com/szebk/PLay_Store_App_Review_Datazoids_Project1_EDA