

Capstone Project Submission

Instructions:

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

Team Member's Name, Email and Contribution:

1. Saurabh A. Funde

E-mail – Saurabh.funde@hotmail.com

1. Dataset's Exploration
2. Data Wrangling
 - 2.1 Filling missing values
 - 2.2 Changing columns datatypes
3. Q&A On Merged Dataset
4. Viz. Distribution Plots, Bar plots
5. PPT

2. Kalyani Motkari

E-mail – Kalyani.motkari@gmail.com

1. Data Wrangling
2. Data Visualization
3. Q&A Play Store Dataset
4. PPT

3. Sameer Ambekar

E-mail – sameerambekar222@gmail.com

1. Data Wrangling
2. Viz. Scatterplot, Regplot
3. Q&A
4. PPT

Problem Statement:

Play store is a marketplace (App) for downloading Android applications for smartphones. Smartphones sells increasing YoY across world. It's very lucrative market for App developers as users looking for comfort and their needs increasing like entertainment, games etc. It's important to find out what type of apps people downloading before developing an App for users and list that APP on Play Store for downloading. It's important to explore and analyze the data to discover key factors responsible for app engagement and success like app size, type and other various factors.

EDA Approach:

- Loading required packages for EDA
- Loading datasets
- Dataset's exploration
- Dataset's wrangling
- Dataset's merging
- Questions and answers for data analysis

- Data Visualization's

In Play store app dataset, there are total 13 columns. Play store app dataset contains **9659 Unique** app name in App column. **Rating** is the only column which has maximum **1474 null values** which contributes **13.60%** of the values.

In User reviews dataset, there are total 5 columns and **865 Unique apps** in App column. This user reviews dataset contains sentiment, sentiment polarity of Apps etc. columns.

Conclusion:

- **92.19%** apps are **Free** and 7.81% apps are paid in type.
- **81.10%** apps have **Everyone** content rating.
- **Events** category has a **highest mean rating of 4.39** and Dating category has lowest 4.05 rating.
- **Family, Game and Tools are top three** categories having 1906, 926 and 829 app count.
- Tools, Entertainment, Education, Business and Medical are top Genres.
- **8783 Apps** are having size less than or equal to **50 MB**.
- **7749 Apps** has rating **more than 4.0** including both type of app.
- **Overall sentiment count** of merged dataset in which **Positive sentiment count is 64%, Negative 22% and Neutral 13%**.
- Price, Rating, Size **has no or very less correlation** with **Sentiment Polarity**.
- It's good to develop a **Free type** app and having a content rating for **Everyone**.
- Most of the apps are fall in **0 to 50 MB Size group**. It's good to develop of app this size bracket.

Please paste the GitHub Repo link.

GitHub Link: - https://github.com/Saurabhf/Play-Store-App-Reviews-Analysis/blob/main/Play_Store_App_Review_Analysis.ipynb