

# Capstone Project Submission

## Instructions:

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

## **Team Member's Name, Email and Contribution:**

Pratik D. Jori

Email: [pratikjori.jori@gmail.com](mailto:pratikjori.jori@gmail.com)

Contribution:

1. Cleaning and exploring the data
2. Finding the null values, dropna and fillna
3. Visualization of the given data :- free vs paid app installation ratio, no of app in the category, top 10 installed apps in any category sentiment ratio of the reviews, correlation heatmap.
4. Ppt making, technical document.

Pritesh Tambat

Email: [priteshtambat1@gmail.com](mailto:priteshtambat1@gmail.com)

Contribution:

1. Cleaning and exploring the data
2. Apply method used to convert the object type into the numerical type
3. Visualization of the data :- factors that affecting the ratings graph, top 10 expensive app, top categories in the playstore, Distribution of apps on there type, size and rating.
4. Ppt making, technical document.

Sarvesha Kapse

Email: [sarvesha6899@gmail.com](mailto:sarvesha6899@gmail.com)

Contribution:

1. Cleaning and exploring the data
2. Visualization of the given data
3. Technical document
4. Ppt making

Someshwar Hydrabade

Email: [hydrabades@gmail.com](mailto:hydrabades@gmail.com)

Contribution:

1. Cleaning and exploring the data
2. Visualization of the given data
3. Technical document
4. Ppt making

## **Please paste the GitHub Repo link.**

Pratik's Github link:- <https://github.com/pratikjori20/Play-Store-App-Review-Analysis>  
Pritesh's Github link:- <https://github.com/pritesh2292/play-store-app-review-analysis>  
Sarvesha's Github link:- <https://github.com/Sarvesha11/-Cy-of-Play-Store-App-Review-Analysis>  
Someshwar's Github link:- <https://github.com/Someshwar15/Cy-of-Play-Store-App-Review-Analysis>

Github Link:- <https://github.com/Link/to/Repo>

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

The given dataset contains Play Store Data and User Review Data. Various app and its review, rating, size data are provided, with 13 columns of different variables and its 10841 observations in the row. Each observation shows the app name and its rating, review etc. we clean the data which is in raw form drop the null values from it.

The problem statement was to identify the sentiment analysis on the app that generated most positive and negative sentiments of app in the play store market basis of the given data.

We have started to initialize the given data and explore the null and missing values and correct them.

There is some column has different data type which we fixed using the apply method to the column using pandas operation. Which we require for the plotting various plots.

Second we involved analyzing numerical type features, with the help of different visualization methods such as bar plot, pie chart, scatterplot, subplot, heatmap etc. Finding the correlation between the variables and also find the percentage for the free and paid apps.

User review data has the sentiment of the user which shows the emotion of the user of app it can be positive, negative or neutral it help to improve the app rating and give the developer better understanding to improve on the application. It also shows that the free app are mostly downloaded than the paid app.

