

# Capstone Project Submission

## Instructions:

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

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

Rajat Chaudhary ([rajat.25.chaudhary@gmail.com](mailto:rajat.25.chaudhary@gmail.com))

Contributions-

- 1- Data Cleaning of User Review Dataset
  - a. Removing all duplicate data from the dataset.
  - b. Removing all null values from the dataset.
- 2- Data Visualization of Play Store Dataset
  - a. Number of apps per category
  - b. Number of apps per category differentiated by type.
  - c. Number of installs according to type-size and categories.
  - d. Number of installs for each category.
  - e. Impact of size on the number of installs.
  - f. Distribution of App Rating.
  - g. Rating on the basis of size.

Anukriti Shakyawar ([shakyawaranukriti@gmail.com](mailto:shakyawaranukriti@gmail.com))

Contributions-

- 1- Data Cleaning of Play Store Dataset.
  - a. Removing duplicate values from the dataset.
  - b. Removing null values.
  - c. Removing impurities from each column of dataset.
- 2- Data Visualization of User Review Dataset.
  - a. Number of review type on the basis of categories.
  - b. Sentiment Subjectivity and Sentiment Polarity.
  - c. Sentiment Polarity relation to the type of app.
  - d. Content Rating relation with Sentiment Polarity.
  - e. Categories Relation with Sentiment Subjectivity.
  - f. Word Cloud of Reviews.

Raman Kumar ([ramank445522@gmail.com](mailto:ramank445522@gmail.com))

Contributions-

- 1- Distribution of Application Type
- 2- Distribution of Size of App
- 3- Installs per Rating
- 4- Reviews Sentiment
- 5- Distribution of Subjectivity
- 6- Content Rating based on Age

Deepmala Srivastava ([srvdeepmala@gmail.com](mailto:srvdeepmala@gmail.com))

Contributions-

- 1- Prepared Presentation

**Please paste the GitHub Repo link.**

Github Link:- <https://github.com/rajat250996/Play-Store-App-Review-Analysis-EDA>

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

In our project Play Store App Review Analysis, we have 2 datasets- Play Store Dataset and User Review Dataset. The problem statement for our analysis is to deliver insights to understand customer demands better and thus help developers to popularize the product.

To solve our problem statement, we performed Exploratory Data Analysis in which first step was to clean all the data by removing duplicate data and null values. After moving on to our Data Visualization we focused on the install pattern of apps for different categories to know about user's behaviour and their preferences while installing apps.

We first identified which category is more preferred after that performed many tests on rating, size and type. While we analysed the preference of users while downloading apps, we also performed analysis on the user behaviour when they give reviews to the apps because it also plays a major role while installing any app. These tests show the sentiment polarity and sentiment subjectivity of reviews, these two factors define whether is app is properly used before the review or not and whether the review is positive or not. If we can improve this factor then it will act as a major part for developers when they develop their new apps.

In our analysis we found that there are some major factors which decides the user's pattern while download apps. These are Type of the App, category from which it belongs, size of the app and lastly rating of the app. In rating factor, we can see whether the rating is given by after using the app or not or whether it is a negative review or not.

From this analysis we can help developers to find which type of apps will be more successful in Play Store and on which factors they should be careful while they prepare for their next android app.