

# Capstone Project Submission

## Play Store Apps Review Analysis

### Instructions:

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

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

Name	Email	Contribution
Raman Kumar	<a href="mailto:ramank445522@gmail.com">ramank445522@gmail.com</a>	Data Visualization & Conclusion
Rajat Chaudhary	<a href="mailto:rajat.25.chaudhary@gmail.com">rajat.25.chaudhary@gmail.com</a>	Data Visualization
Anukriti Shakyawar	<a href="mailto:shakyawaranukriti@gmail.com">shakyawaranukriti@gmail.com</a>	Data Filtering
Deepmala Srivastava	<a href="mailto:srvdeepmala@gmail.com">srvdeepmala@gmail.com</a>	Data Cleaning: Identifying and removing duplicate entries. Removing visual impurities like "+", "\$" sign.

Github Link:- [https://github.com/deepmalasrivastava/Playstore\\_analysis](https://github.com/deepmalasrivastava/Playstore_analysis)

**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 this capstone project we have compared thousands of applications across various categories. We have analyzed the data to discover key factors responsible for app engagement and success helping the developers to work and capture the android market.

We have been provided with 2 Dataset files – 'Play store .csv' and 'User Reviews.csv'. One containing 13 databases namely 'App', 'Category', 'Ratings', 'Reviews', 'Types', 'Size', 'Installs', 'Genres', 'Price', 'Content Rating', 'Last Updates', 'Current Version' and 'Android Version' and another file containing databases namely 'App', 'Translated Review', 'Sentiment', 'Sentiment\_Polarity' and 'Sentiment\_Subjectivity'.

We started by performing some fundamental research on our dataset. When we did this, we found the basic information regarding our data set such as columns, data types and we also found out that we have missing values, data duplication and a few other issues as well. Then we **filtered** it one by one.

After this we performed EDA . We plotted following graphs:

- Number Distribution of Application Type
- Apps per Category
- Number of installs type-wise according to categories
- Top genres in playstore
- Number of installs for each category
- Number of apps per category differentiated by type
- Distribution of Size of app
- Impact of size on the number of installs
- Distribution of App Rating
- Install per Rating
- Distribution of Apps in terms of their Ratings, Size and Type
- Reviews Sentiment
- Distribution of type of reviews as per categories
- Distribution of Subjectivity
- Sentiment subjectivity and Sentiment Polarity
- Content Rating on the basis of Age
- Sentiment polarity relation with type of app
- Content Rating relation with Sentiment Polarity
- Categories Relation with Sentiment Subjectivity

#### **Asking questions from Dataset**

1)What is the top 5 apps on the basis of installs?

Google Street View, Instagram, Google Play Movies & TV, Google Photos, WhatsApp Messenger

2) What is the top 5 reviewed apps?

Facebook, WhatsApp Messenger, Instagram, Messenger Text and Video Chat for Free, Clash of Clans

3) What is the top 5 expensive apps?

I'm RichT '(),/ أناغني/ 我很有錢 , I am Rich Plus, I am Rich!

4) What is the top 3 most installed apps in Game category?

Subway Surfers, Temple Run 2 ,Score! Hero

5) Which 5 apps from the ' category are having the lowest rating?

FE Mechanical Engineering Prep, Truck Driving Test Class 3 BC, Speech Therapy: F, AC REMOTE  
UNIVERSAL PRO, BG TV App

#### **We concluded from analysis:**

- Most of the trending apps are from the categories like family, tools, and games.
- Most preferred Apps by users in point of size or weight are light size apps.
- Users also installed apps on the basis of their ratings.
- These ratings are defined on 2 points- Polarity and Subjectivity.
- This analysis will help developers while preparing for their next apps.