

# Capstone Project Submission

## Instructions:

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

Team Member's Name, Email and Contribution:

Name	Email	Contribution
<ul style="list-style-type: none"><li><u>Kumar Abhinav</u></li></ul>	<u>Kumarabhinavthakur274@gmail.com</u>	Data Cleaning ,Data Manipulation, Problem Statements, Insights.
<ul style="list-style-type: none"><li><u>Saumya Dash</u></li></ul>	<u>Saumyadash9@gamil.com</u>	Data Visualization, Data wrangling, Technical Documentation, Ppt.

Please paste the GitHub Repo link.

Github Link:-

[https://github.com/kumarabhinavthakur274/Play\\_Store\\_-App\\_-Review\\_Analysis---Capstone-Project](https://github.com/kumarabhinavthakur274/Play_Store_-App_-Review_Analysis---Capstone-Project)

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

Google Play store is an android application which provides facility of an app store to its users, where user can install various software applications of their use. It contains various applications diverse in class and genre for example entertainment, communication, gaming and what not. There are major key insights that can be drawn from the play store data set which can influence decision making. It has the potential to make business plans and expand the business further as far as app making business is concerned. User review dataset can be used to improve the service for hassle free experience for the users.

We are provided with two data sets that is play store data and user review data as CSV files. We imported these two dataset in google colab and converted it into data frame (df and df1) . Each row of the df (play store df) has attributes namely 'App', 'Category', 'Ratings', 'Reviews', 'Types', 'Size', 'Installs', 'Genres', 'Price', 'Content Rating', 'Last Updates', 'Current Version' and 'Android Version' . Similarly for df1(user review df) has attributes 'App', 'Translated Review', 'Sentiment', 'Sentiment\_Polarity' and 'Sentiment\_Subjectivity'.

Then we performed various data cleaning operation to avoid miscalculation and double counting of data during data analysis. Tasks we performed for this purpose are dropping duplicated instances from both the data frames, dropping or substituting null values as per requirement, replacing numerical columns given as 'object' to 'float' values, Conversion of Different Units given for a particular column to a common one, drop or replace wrong data as per requirement.

After all the cleaning operation is done we framed some problem statements which were based on real life problems faced by service providers and users in app making business. Going by each problem statement we performed various groupby and sorting operations to get useful insights from the data then we plotted the insights on various graphs and charts which were easy to comprehend. The key insights were :

- Top 3 categories with most number of apps are Family, Game and Tools.
- Top 3 categories with most installations are communication, social media and video player.
- There is a gap in number of users in few categories and number of apps available in the market e.g. Medical category.
- Most of the apps have average rating more than 4.
- Gaming apps have slight negative sentiments among the users.
- Apps rated for 'Everyone' has abundance in market but 'Teen' and 'Mature' rated apps are widely downloaded.
- There are very few paid apps available.
- Reviews and Installations are directly correlated.

Based on the insights that we got from the exploratory data analysis we can make plenty of data driven decisions which will help growing and expanding businesses. Based on our analysis we gave few recommendations which can be implemented:

- Some of the Categories like Communication, Social Media, Video Player and Gaming are highly in demand among users but there are very few apps to cater so there is a huge gap in the market where business can be built.
- Gaming category is popular among users and service providers as well but we can observe some negative sentiments among the users for this category. So Gaming firms can improve their services for hassle free experience among users.
- Play store has maximum number of apps rated for everyone but crowd is downloading teen and mature 17+ more. So there is scope of building business around it.
- There is a very few app that are paid in the app store so there is scope of building business around monthly or yearly subscription model. Ed tech, Gaming, Entertainment has enormous potential to act in this regard.