**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| 1)**Sourav Karmakar**  Email: [karmakarsourav2024@gmail.com](mailto:karmakarsourav2024@gmail.com)   * Data Cleaning, Sorting, Manipulation. * Framework establishment and Error Debugging * Analysis of Best Average Rating of categories and Genres * Last Update impact, Rating and Average Reviews Analysis * Overall people’s sentiment analysis according to subjectivity, reviews, genres * Bar plot, reg plot, Histogram, Line plot, PI-plot * Data melting * Project Summary   2)**Meenu Benny**  Email: <meenubenny555@gmail.com>   * Data Visualization * Bar plot * Finding Free and Paid Category App * Data Melting * Sentiment Analysis according to Rating * Approach towards Plan * Technical Documentation     3)**T. Ompriya Subudhi**  Email: <ompriyasubudhi2000@gmail.com>   * Seaborn * Matplotlib * Most Reviews, Rating analysis * Variation of Reviews in particular month, year analysis * Bar plot * pi-plot * Data Visualization   4) **Thasniya Mansoora**  Email: <thasniyamansoora91@gmail.com>   * Analyze factors affecting Installation * Finding Top Content Rating * Co-relation establishment * Matplotlib * Seaborn, bar plot * PPT Presentation |
| **Please paste the GitHub Repo link.** |
| Github Link:- <https://github.com/souravkarma/playstore-app-review-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)** |
| Data has been provided by AlmaBetter .  From these analysis clients can understand the needs of their customers in different aspects of product development and it will help them to solve the problems with the market superiors.  The main objective of this project is to deal with the data provided by the company and to analyze the data in every possible aspect. It will help company to understand consumer behaviors, their buying pattern and give more clarity to developers regarding future important decisions.  **Our Approaches**    **At first** We are following four steps for cleaning our data that is:-  1)Treatment of null values  2)Treatment of Duplicate Rows  3) Conversion of necessary column values from string to required datatype 4)Replacing the missing values with mean, median or mode.  After cleaning the data we simply analyze our first data frame playstore data csv and solve the problems of top average rated category,genres,most reviewed, installed apps etc.  For sentiment analysis we merge two dataframe playstore data and user review data. It helps us to find sentiment subjectivity, Top review sentiments, Genre’s sentiment, co-relation between various columns and many more things.  **Conclusion**  After doing this project we got many interesting conclusions. First, we have done the removal of null value from rows and columns and the same goes with the removal of duplicates from the datasets. Then we did the formatting for each of the required columns in each dataset.  App with the category Family and the genre tools are in large numbers. But interestingly Category EDUCATION and Genres 'Board;Pretend Play' and 'Comics;Creativity' both get the best average rating in playstore.  App Rating is directly proportional with the recent update. From this we can be sure that with the latest update the reviewers are giving better rating.  We also found that Ratings are directly proportional with Average Reviews till 4.5. But after that we observe some deviation and eventually reviews decreases after 4.5 ratings. 4.5 Rating apps have maximum 1.1 million reviews.  we found that most of the opinion on sentiment subjectivity lies high in the range 0.4 to 0.7. It means people give review after some experience.  We also found people's sentiment are more towards paid app less towards free app  GAME category has maximum no of translated reviews but, most positive percentage acquired by HEALTH\_AND\_FITNESS category and GAME has most negative reviews among top 5 categories. So, it concludes that most reviews do not suggests most positive sentiment percentage.  Action Genres got most strong positive sentiment among all genres. It suggests that people prefer Action genres app far more than any other genres.  google playstore has more free apps than paid. and Family category has most free and paid apps in playstore. MEDICAL category has highest number of paid apps considering quantity.  Social,entertainment,event,education,comic\* etc these popular category apps are always free.  We found that in terms of both Top Total and Average Apps Reviews outcomes Facebook,Instagram,whatsapp messanger,Clash of Clans these apps are common. so we can conclude that these 4 apps are most popular and trending among masses.  Subway Surfers app has highest rating along with highest reviews followed by Instagram,clash of clans,Facebook,whatsapp messanger,clean master.  We got highest reviewed apps are those which are last updated in january,2018 and got lowest review in February and march update. It concludes that Most of the companies are majorly updating their apps in January month, so in that months they get most reviews.  we also found that 4.4 rated apps are most reviewed apps in google playstore and from 2.5 to 3.1 there are almost equal numbers of positive, negative, neutral reviews. It states that neither top rating nor less rating makes an app good or bad.  We found that GAME,COMMUNICATION,SOCIAL these category apps got most installation(above 1 crore).  Most heavy app got 10cr installations, among top 10 heavy apps 4 apps got above 1 crore installations so, it concludes that customers still preferred heavy apps.  Most Installed, Rated, reviewed positive sentiment apps are always in range between 4 and 4.5 rating.  Most of the apps which are used by the users have a content rating of ‘Everyone’.  From install,rating,reviews co-relation graph we also found install and reviews are positively co-related means when Installation is more reviews are more. |