**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| Dhawal Subhash Khandait(dhawalskhandait6701@gmail.com)  1. Data Cleaning:  a. App review dataset  b. Play Store app dataset(some part)  2. Objectives:  a. Which type of apps are users willing to pay for  b. Which type apps has more / best reviews ?  c. How do reviews affect the user's decision to download apps?  d. Are app updates important?  e. Are sentiment influences the final rating of the app?  f. How is the app rating distributed?  g. How many apps are there in each category?  h. Which category apps are installed most?  Sakshant Rajendra Gongal (sakshantgongal@gmail.com)  1. Data Cleaning:  a. Play store app dataset  2. Data transformation  3. Objectives:  a. Which age groups do different categories target?  b. Do paid apps get better ratings?  c. Does rating change with increasing price?  d. Does the size of an app influence the number of downloads?  e. Sentiment analysis for free and paid apps.  f. What percentage of apps are paid?  g. Correlation matrix |
| Drive Link-  https://drive.google.com/drive/folders/1SC2SPMxc-0zbJ4Xk2wqJ4vtEYGcS-TBv?usp=share\_link |
| **Please paste the GitHub Repo link.** |
| Github Link:-  https://github.com/DhawalKhandait/Play-Store-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)** |
| Google play store is immersed with a few thousands of new applications regularly with a progressively huge number of designers working freely or on the other hand in a group to make them successful, with the enormous challenge from everywhere throughout the globe. In this EDA project, we were provided with two datasets i.e play store data & user reviews.  As the first step, we perform data wrangling over the raw data. Further we divided the complete project into three main parts i.e User analysis, Sentiment analysis & bivariate analysis.  After transforming data, we started to extract data insights from given data. In user wise analysis, we focused on distribution of app rating, most installed apps according to category, targeted age group, percentage paid apps, is paid apps performance over free apps etc. in short, we were trying to extract each and every details from data which will help to tell what kind of apps user really want and which will help developers to design apps according to needs of user.  In sentiment analysis, our main goal was to analyze how sentiment influences performance off app on Google Play Store. In this analysis, we observe how sentiment influences final rating of the app, sentiment analysis for free and paid apps etc. based on these observations we safely can say that user sentiment has really huge effect on performance of an app on google play store.  In Bivariate analysis, we have compared different features of the data. Like we have observed, does rating change with increasing prize? Does size influences no. of downloads? Do high rated apps attract more user? Apps update’s importance and is review affect user decision or not.  In the entire analysis we have trying to observe each feature which influences success of an app on Google play store. After Analyzing the dataset, we have got answers to some of the serious & interesting questions which any of the android users would love to know. We have also learned that the following things might affect the rating. Rating is depending variable, whether it’s free or paid, number of the reviews, sentiment of the review comments. |