**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| 1. Akshay Kumar Saini   E-mail: [ak663348@gmail.com](mailto:ak663348@gmail.com)   * Data Cleaning  1. Price String to Integer 2. Converting reviews string to integer  * Data Visualization  1. Making Bar plot 2. Stacked Bar Plot 3. Klib Distplot 4. Categorical Plot using Klib 5. Correlation heat map 6. Scatter Plot  * Data analyzing  1. Analyze the play store dataset 2. Finding Relation Between 3. Rating and Category of app 4. Rating and size of app 5. Category and No. of installation 6. Analyzing Overall sentiment of user  * Presentation * Summary Documentation  1. Kumkum Singh   E-mail: [kumkum24091999@gmail.com](mailto:kumkum24091999@gmail.com)   * Data Cleaning (Play Store Dataset)  1. Converting Install String into integer Datatype 2. Cleaning size it and string to integer 3. Cleaning the second dataset 4. Replacing the N/A mean value from the average rating 5. Dropping the Null Value  * Data Visualization  1. Making Pie plot 2. Pair plot using Klib. 3. Word Cloud  * Data analyzing  1. Analyzing the data Category Wise 2. Category by no of Install 3. Rating by Category 4. Effect on No of install with increasing app size 5. Analyzing sentiment category wise  * Technical Documentation * Presentation |
| **Please paste the GitHub Repo link.** |
| GitHub Link: <https://github.com/Kumkum24/Playstore-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)** |
| Summary:  The Google Play Store App Overview provides some useful details regarding app trends in the  Google Play. According to the visualizations above, most of the popular applications (in terms of users  installation) are from categories like GAME, COMMUNICATION and TOOL, although the quantity  available applications from these categories are twice as few as in the FAMILY category, but still used  most. The trends of these apps are probably due to their nature of being able to entertain or help  The analysis on the data set of google play store throws the light on the which category has the maximum applications, which category of apps people prefer to install more. Which apps has highest rating.  Many libraries like klib, pandas, ast, matplotlib, numpy, seaborn, datetime and wordcloud were used to analyze the data.  Analysis showed that**: -**   * Family category has maximum numbers of apps, accounts for 28.8%. * Communication category has highest interaction with the users, which accounts for approx. 3.5 million. * Number of free applications accounts for 92.6% whereas number of paid applications accounts for only 7.4%. * Mostly apps are rated by users in range of 4 – 4.5. * people prefer to install apps belonging to gaming most which accounts for 3.5 million. * most of the apps including education, Dating and health and fitness app have good number of positive reviews i.e. more than 50% whereas game app have received 50% positive reviews and 50% negative reviews.   Key factors responsible for app engagement and success:   1. If it Belong to the Game, Communication, Productivity ,Social category ,then it have a high probability to success as compared to other categories 2. Free apps are tends to be more installed as compare to free apps 3. Size of the Apps should be around 40 MB or 70-80MB to increase the probability of Success 4. App Content Rating should be in ‘Everyone’ as ,Everyone Content rating apps are most installed app. |