**Capstone Project Submission**

**Instructions:**

I) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Project name:** Play Store App data analysis and review |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| **Problem** **statement**: given is, The Play Store apps data has enormous potential to drive app-making businesses to success. Actionable insights can be drawn for developers to work on and capture the Android market.Each app (row) has values for category, rating, size, and more. Another dataset contains customer reviews of the android apps.Explore and analyze the data to discover key factors responsible for app engagement and success. |
| **Approach**: The basic requirement is that we have to validate the data, find out if there are any null value is there or not, also we have to check any missing value is there or not, this kind of data affect the dataset and make them dull. So here, what we do we just check that where this kind of data is found, and wherever we found, we just make the delete whole row as its will not much affect the data positively.  Then next thing we have to deal that find the null values,  Identify the null values and delete the null values.  We find correlation of app and category and it’s unique and nunique values and count, which will further used to show that on graph to make it understand clear.  Then we find top categories of play store and viewed it on a chart.  Then we find distribution of apps across a various app category and show it on distribution graph.  Then we find a number of apps per categories and shows it on pie chart.  Then we do find application categories with their average rating and view it on a bar graph.  Then we found number of counts of current version and android version and actual count on basis android version to get what we have in bulk quantity.  And then we merge two table to find number of install type wise according to type.  Apart from this we also saw out of total, what amount of content is classified for which kind of age group so that we can decide the range of customer for the app categories.  And also find what kind of apps are dominate in market among paid and free.  **Conclusion**: So, at the end we want to spread the app, make them free, then make it rare, and after that just make them paid. |
| **Team nam**e: Alma Striker  **Team Member Name:**   * + - Adesh Dhole: [adesh.dhole19@gmail.com](mailto:adesh.dhole19@gmail.com)     - Chetan Mangale: [sajanpatki@gmail.com](mailto:sajanpatki@gmail.com)     - Mohit Sonwane: [mohitsonwane@gmail.com](mailto:mohitsonwane@gmail.com)     - Amol Waghmare: [irsamolwaghmare@gmail.com](mailto:irsamolwaghmare@gmail.com)   **Contributor:**  **Adesh Dhole:**   1. Data setup 2. Data Cleaning 3. Gather Data characteristics 4. Visualize the characteristics of column on graph   **Chetan Mangale:**   1. understand the data 2. Create comparison on two columns 3. Make a data frame and visualize it’s on graph 4. Merge the two table and give a characteristics of merged table     **Amol Waghmare:**  **1**  Take a input from all and start analyze on the basis of characteristics   1. Analyze the charts add some extra info to it. 2. Add some extra question to analyze 3. Perform one operation to take instincts   **Mohit Sonwane:**  **1**  collect all the codes, collect the colab notebook and edit it well  2 read all the flow, understand the work, create technical documents  3. Create ppt and finish all the documentation part. |
| **GitHub Repo link.** - <https://github.com/Adesh1999/EDA_Capstone_Project_1> |
| **Drive link: -** [**https://drive.google.com/drive/folders/1MJezLK8M7oSrfzHQtHfUU4M\_KkUK835a?usp=sharing**](https://drive.google.com/drive/folders/1MJezLK8M7oSrfzHQtHfUU4M_KkUK835a?usp=sharing) |