Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
Pratik D. Jori Email: pratikjori.jori@gmail.com Contribution: 1. Cleaning and exploring the data 2. Finding the null values, dropna and fillna 3. Visualization of the given data :- free vs paid app installation ratio, no of app in the category, top 10 installed apps in any category sentiment ratio of the reviews, correlation heatmap. 4. Ppt making, technical document. Pritesh Tambat Email: pritesh.tambat1@gmail.com Contribution: 1. Cleaning and exploring the data 2. Apply method used to convert the object type into the numerical type 3. Visualization of the data :- factors that affecting the ratings graph, top 10 expensive app, top categories in the playstore, Distribution of apps on there type, size and rating. 4. Ppt making, technical document. Sarvesha Kapse Email: sarvesha6899@gmail.com Contribution: 1. Cleaning and exploring the data 2. Visualization of the given data 3. Technical document 4. Ppt making Someshwar Hydrabade Email: hydrabades@gmail.com Contribution: 1. Cleaning and exploring the data 2. Visualization of the given data 3. Technical document 4. Ppt making
Please paste the GitHub Repo link.
Pratik's Github link:- https://github.com/pratikjori20/Play-Store-App-Review-Analysis Pritesh's Github link:- https://github.com/pritesh2292/play-store-app-review-analysis Sarvesha's Github link:- https://github.com/Sarvesha11/-Cy-of-Play-Store-App-Review-Analysis Someshwar's Github link:- https://github.com/Someshwar15/Cy-of-Play-Store-App-Review-Analysis

Github Link:- https://github.com/Link/to/Repo

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

The data is in form scrap from google play store then dataset is created from that scrap data and we preprocessed the given dataset and this visualize to the user through interface which we have built. There is having the null/missing values in dataset. However, the raw data extracted need to be preprocessed to turn it into some valuable information.

To be able to perform EDA on our dataset, installs, type, size, reviews, price, were converted to integers, size of apps are in kilobytes were converted to megabytes to have consistent units. Further to find out the sentiments of users the review text was processed using sentiment polarity and subjectivity of each of the review were found with polarity values between -1 & 1 and subjectivity values between 0 & 1.

With the help of matplotlib library we visualize using bar plot, pie chart, scatter plot, subplot, heatmap for finding the correlation between variables and pie chart for finding top 10 expensive apps in play store. We visualize that the reviews, size, installs, prize these are the factors which are affecting the rating of the application. We bar plot the highest earning app present in the play store.

We visualize review sentiments ratio which is gives emotions of the writers that is positive, negative and neutral.

The overall objective of this analysis effort is to provide users an interface where users can check the genuineness of application which uploaded on the google play store.