Capstone Project Submission

Team Member's Name, Email and Contribution:
Name- Saurabh Ravindra Shinkar **Email**- saurabhshinkar26@gmail.com **Contribution**- Everything in the project

Please paste the GitHub Repo link.
GitHub Link: https://github.com/saurabhshinkar/Play-Store-App-Review-Analysis--EDA

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 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. Our problem statement was to explore and analyse the data to discover key factors responsible for app engagement and success. First step was to clean the dataset and make it in proper format. Next to get the solution of the problem statement it was necessary to understand each feature in the both play store and user reviews dataset. After analysing each feature separately, we started to see how different features impact on app at the same time. We analysed each feature through output of relevant data frames and from different with the help of pandas, matplotlib and seaborn.

Conclusions from project-

- Most of the app's belongs to Family category. They are approximately 19% among all categories.
- Ratings are negatively skewed with mean rating of 4.19
- Sizes are positively skewed with median size of 12 MB.
- Prices are positively skewed with mean price of 1, median price of 0 and max price of 400.
- Approximately 92% apps are free.
- Tools, Entertainment and Education are the top 3 genres.
- Approximately 64% sentiments are positive, 22% are negative and 14% are neutral.
- Mean Sentiment Polarity is 0.180904
- Mean Sentiment Subjectivity is 0.493767
- There are 271 apps which have 5-star rating. In which 67 apps belong to the Family category.
- Among Top 20 install apps 30% belongs to the communication category and 15% belongs to the social category.
- Events category have the maximum mean rating of 4.395313
- Genres of Communication, Tools, Productivity and Social have the maximum no. of installs.
- Apps belonging to Finance category have the highest mean price (8.408203) among all categories.
- Family category has the highest no. of paid apps.
- Game And Family Category Have the Highest No. Of Installs of Paid Apps.
- Installs are positively correlated with reviews with correlation value of 0.63