

Capstone Project Submission

Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1. Sanjeev Hegde:
 - Mail: sanjeevhegde454@gmail.com
 - Contributions: Data cleaning, essential modifications to the data, visual representation of analyzed data, PowerPoint Presentation, Project Summary and Project Report

Please paste the GitHub Repo link.

1. Github Link for Sanjeev Repository: -
<https://github.com/Sanjeevhegde009/EDA-Capstone-Project-Playstore-App-Data-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)

Play store is the official app store for devices running on android platform developed by Google LLC. The platform lets android users download and install secure apps developed by various developers or companies. Several invaluable insights can be drawn from analysis of data available on the Play Store. In this project, We are provided with two data sets, one containing publicly available information such as average ratings, number of reviews etc. while other data set contains processed data of user reviews based on the sentiments in the review.

As an initial step to EDA on the given data sets, data cleaning and essential modifications have been performed on the data sets wherein, misaligned values have been arranged appropriately, NaN values in numeric data have been replaced by their median values or Mode values or have been dropped accordingly. Further, Categorical values in the data set have been replaced by their mode values, quantified values of the data were assigned with appropriate Units of measurement, non-essential characters such as commas and "+" etc. were cleaned. Also, Columns have been converted to the appropriate data types while unimportant columns with already processed data were removed from the data set.

Several meaningful insights were drawn from the Exploratory Data Analysis of the two data sets. Among 9701 unique apps, it can be noted that Family category of apps are mostly present in the data set followed by Games category. Average Ratings for apps is about 4.2 stars while average size of apps is about 16 MB. Also, Games category apps were most installed by the users while all category of apps had above 4 star rating. Users have tendency to install apps with more reviews in comparison to apps with lesser reviews. Further, apps with size less than 25MB were most installed. Data sets mostly contain apps developed and rated mostly for everyone. Paid apps reviews are more of positive sentiments

compared to free apps. Also, paid app reviews are slightly subjective in nature compared to free apps wherein paid apps reviews also contain extremely subjective as well as objective reviews as compared to free app reviews.

From the above analysis, several insights have been drawn and are of great help to the developers in design and development of apps, their pricing models and thereby impacting the revenue generated from the apps.