Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
Name= Krushna Chaure Email id= chaurekrushna143@gmail.com Contribution> • Doing project individually. • Connecting and Uploading dataset to the google colab notebook. • Understand and Analyze the data. • Perform Data cleaning over dataset. • Perform Data wrangling over dataset. • Doing Data Visualization and finding insights. • Technical writeup. • Power Point presentation. • Project Summary.
Please paste the GitHub Repo link.
Github Link:- https://github.com/KrushnaChaure/play-store-app-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) The Google play store is one of the largest and most popular Android app stores. It has an enormous amount of data that can be used to make an optimal model. In today's scenario we can see that market has increased to over 3.5 million Apps and around 3000+ apps are being added per day as per a Google survey report. Thus, the market, in turn, led to around 5 billion users downloading all over the world. therefore enormous datasets & variety of insights can be concluded for business improvements. There are various key factors that play a major role in the success & engagement from the user's end. Our problem statement is quite inevitable in comparison with the present Google Play store App market. In this EDA project we were provided with two datasets. • Playstore.csv -> contains all the details of the applications of Google Play. There are 13 features that describe a given app. • User_reviews.csv -> contains 100 reviews for each app, most helpful first. The text in each review has been pre-processed and attributed with three new features. · Sentiments (Positive, Negative, Neutral) · Sentiment Polarity

- Sentiment Subjectivity

At first, we break down the datasets by importing necessary library classes, followed by checking unique values, converting the data types to similar objects, removing special characters as the analysis demands & making the entire dataset ready for analyzing & plotting actionable insights.

After examining null & missing values from the dataset we directly went deep into the visualization steps.

Some observation and conclusion on which we worked are as follows:

- Number of categories of apps available on play store where Family and Game have maximum number of apps available in play store and Comics & Beauty has less number of apps in play store.
- Number of installed apps for each category where gaming and communication type of apps installed top in this list.
- How much number of apps available on play store are free & paid where most of the apps which is 92.6% are free and only 7.4% apps are paid.
- How many apps are accessible according to their specific age group where most of the apps in play store are accessible for everyone, their is no any restriction to use this apps and there min number of apps accessible only Teen, Everyone10+, Mature 17+.
- Average rating of different categories of apps where most of the apps rating are between in 4 to 4.5 and education base & event base apps top in this list.
- Correlationship between apps(free & paid) with respective size where size of apps smaller then it will more install and size of apps less it will less install also free apps installed maximum time compare to paid apps.
- Correlationship between multiple variable such as rating, size, installs, price where positive rating between installs and rating & there negative rating between size and rating.
- Percentage of review sentiments where positive sentiment is more which is 64.1%, negative sentiment is 22.1%, neutral sentiment is 13.8%.
- Correlationship between sentiment and rating where that even its positive, negative or neutral the median(50%) remains near 4.4.
- How much customer subjectivity of sentiment where maximum number of customers give review to the apps according to their experience.