**Capstone Project Submission**

Play Store App Review Analysis

Mobile apps are everywhere. They are easy to create and can be lucrative, because of these two factors, more and more apps are being developed. In this notebook, we will do a comprehensive analysis of the Android app market by comparing over ten thousand apps in Google play store across different categories

The Google play store app dataset consists of enormous data that can be used to create effective insights. 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. To , explore and analyze datasets are provided, one with basic information and the other with user reviews for the respective app and discover all the key factors responsible for app engagement and success.

Upon doing several pieces of research it can be seen that every day around 3000+ apps is being added to the play store library. Therefore enormous datasets & variety of insights can be concluded for business improvements.

In this EDA project we were provided with two datasets

1. Playstore.csv -> contains all the details of the applications of Google Play. There are 13 features that describe a given app.

They are :

* App :- Name of the App
* Category :- Category under which the App falls.
* Rating :- Application's rating on playstore
* Reviews :- Number of reviews of the App.
* Size :- Size of the App.
* Install :- Number of Installs of the App
* Type :- Whether the App is free/paid
* Price :- Price of the app (0 if it is Free)
* Content Rating :- Appropriate Target Audience of the App.
* Genres:- Genre under which the App falls.
* Last Updated :- Date when the App was last updated
* Current Ver :- Current Version of the Application
* Android Ver :- Minimum Android Version required to run the App

1. 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
* Then by basic inspection of dataset.
* Then data cleaning and clean that attributes which have not to be useful and Replace all the Null values with the Average of their Columns or with not null values.
* Then followed by Data manipulation and handling duplicate data 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.
* As the data become ready for analyze we went into the visualization steps.

Some insights on which we worked are as follows:

1. Getting the average Rating of the Apps.
2. Checking the count of application in each category
3. Getting the number of installs in each Category
4. Checking Corelation between the variables
5. Getting App Size Distribution
6. Getting the number of installs with respect to the size of the application
7. Lets see the major Type of app distribution in playstore.
8. Lets see the apps dependence on size and type and it effect on rating too.
9. Review sentiments in all the app dataset.
10. Let's see a more depth understanding of sentiment Polarity and Subjectivity.
11. Does sentiments Polarity is proportional to sentiments subjectivity .

**Contributor Roles**

1. **Sridhar bhagaban Padhy (www.sridharpadhy45@gmail.com)**

* Upload dataset to Google colab and explain dataset to team members.
* Analyze null values and filter them.
* Data cleaning.
* Correction of data types
* Data wrangling and manupulation
* Data Visualizations ( 6- 11 Insights )
* PowerPoint presentation
* Project summary

1. **Pranit Ransingh ( www.pranit.ransingh24@gmail.com)**

* Upload dataset to Google colab and explain dataset to team members.
* Analyze null values and filter them.
* Data cleaning.
* Data handling
* Correction of data types
* Data Visualizations (1-6 Insights)
* Technical documentation
* Project summary

|  |
| --- |
| **Please paste the Github Repo link:** |
|  | | |