Google Play store App Data Analysis

Utkarsh Shrivastava,
Omkar Dadasaheb Sargar
Data science trainees (Cohort Amsot),
AlmaBetter

Abstract:

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. We are provided with two datasets, namely: 'Play Store Data' and 'User Reviews'.

'Play Store Data' has details about app category, rating, size, and more, while the 'User Reviews' table contains customer reviews of the android apps. Our goal was to explore and analyze the data to discover key factors responsible for app engagement and success.

1.Problem Statement

This data set contains Play Store App information for a wide variety of apps and its categories, includes information such as how many apps belong to a certain category, app ratings out of 5, how many downloads were made, User reviews for the apps in the dataset, the size of the app, type and Genres etc. The User Reviews dataset contains information about some reviews for some apps including sentiment subjectivity and polarization, which can be useful to provider in future for making important decisions.

Dataset Information:

Play Store Data:

- **Apps**: Contains name of the app
- **Category**: Contains category of the app
- **Rating**: Contains avg rating of the app
- **Reviews**: Contains no. of reviews available for the app
- **Size**: Contains size of the app
- **Installs**: Contains no. of installs to date
- **Type**: Contains free/paid status of the app
- **Price**: Contains price of the app
- **Content Rating**: Contains information about age preference of user base
- **Genre**: Contains genre of the app
- **Last Updated**: Contains last update history for the app
- **Current Ver:** Contains the info about current version of the app
- **Android Ver:** Contains info about the android version required on the device

User Reviews:

- **Apps**: Contains name of the app
- **Translated Review**: Contains the user review
- **Sentiment**: Contains Sentiment information of the review
- **Sentiment Polarity**: Contains sentiment polarity of the review
- **Sentiment Subjectivity**: Contains sentiment subjectivity of the review