Google App Store Data Analytics

Team:

- Prashul Kumar
- Dikshak Khanna
- Jagadeesh

Summary:

As the mobile app market continues to grow, it becomes imperative for developers and businesses to understand the dynamics of user engagement, retention, and overall app performance to stay competitive. The abundance of apps in various categories and varying user reviews and ratings make it challenging for consumers and developers to identify the best-performing and most popular apps.

To tackle this challenge, we propose developing a comprehensive, user-friendly web application that utilizes a dataset of Google app details to provide insightful analytics on user retention, engagement, and app performance. The application could also allow users to submit their reviews or ratings for the apps available on the Google Play Store. Moreover, this software will be designed to help developers understand the factors contributing to an app's popularity and success and provide consumers with a platform to easily identify the top apps in different categories based on user reviews and ratings.

This can also be utilized to get an idea of the social impact of Apps by reviewing the user reviews and getting a user sentiment accordingly.

Technology Stack (could change later on some parts depending on how we progress):

1. Backend: Django
2. Frontend: React
3. Database: PostgreSQL

Objectives:

The primary objective of this project is to develop a web application that leverages data analytics to provide valuable insights on user engagement, retention, and app performance, ultimately helping developers optimize their apps and consumers make informed decisions.

Features and Functionalities:

1. User Engagement and Retention Analytics:

   Analyze user engagement metrics such as daily active users, session duration, and bounce rate. Provide insights on user retention rates and churn rates. Offer recommendations on how to improve user engagement and retention.

2. App Performance Analytics:

   Evaluate app performance based on user reviews and ratings. Identify trends and patterns in user feedback. Provide actionable insights for app improvement and

optimization.

3. Top App Identification:

   Allow users to submit ratings and reviews for the apps they downloaded from the store, adding to the database of existing reviews. Secondly, they can filter to view the top-performing apps based on various categories. Lastly, it also lets users sort apps based on user reviews, ratings, and other relevant metrics. Highlight the best-performing apps in each category.

Usefulness:

Our database will be helpful as it will help with answering the following -

Usefulness: Firstly, the database and web application will provide valuable insights into user engagement app performance and enable users to add their ratings and reviews to the app as and when they use it, thus making it a one-stop solution for app analytics. The new users can then utilize these ratings and analyze the results to see which app is the best in its respective genre and use the same for downloading as per their needs. Lastly, since we are also tracking the sentiments (reviews), it may help in the future with detailed market research for app developers on what aspects they can improve when upgrading the version of the app or developing a new one to compete with the existing ones.

While there may definitely be an existing database for this type of study, our website and database, in conjunction, will provide a comprehensive data analysis using detailed information about apps user reviews, thereby providing a holistic view of the app. Secondly, we also enable user-generated data, as in the ability for users to submit their own reviews and ratings, adding a dynamic component to this database rather than just a static view, making it more collaborative.

Our target users are spread across multiple groups -

App Developers - Developers aiming to enhance their app's performance or utilize the reviews to build a better version or a new app.

Consumers - People aiming to download high-quality apps across different genres and wanting to submit reviews for the newly launched apps in the Play Store

App Market researchers - People interested in analyzing the Android app market, understanding user sentiment, and tracking trends across categories.

Dataset:

The Kaggle dataset contains 2 files named "apps.csv" and "user_reviews.csv". The data was scraped from the Google Play website. There are many popular datasets for the Apple app

store, but there are few for the Google Play store, mainly due to increased difficulty in scraping the latter compared to the former.

The main purpose of collecting this data is to comprehensively analyze the Android app market by comparing over ten thousand apps in Google Play across different categories.

The "apps.csv" file contains all the details of the applications present on Google Play. There are 13 features that describe a given application. They are App, Category, Rating, Reviews, Size, Installs, Types, Price, Content Rating, Genres, Last Updated, Current version, and Android version.

The "user_reviews.csv" file contains 100 reviews for each application, with the most helpful first. The text in each review has been pre-processed and attributed with 3 new features. i.e., Sentiment, Sentiment Polarity, and Sentiment Subjectivity. The remaining 2 features are application and review.

The dataset was collected from Kaggle.
source: https://www.kaggle.com/datasets/yassershrief/goggle-play-data?select=apps.csv

Communication and Sharing:

We communicate within the team whenever required on key topics and regularly discuss with each other on the Gmeet.

GIT Repo: https://github.com/jagadeesh-chitturi/ADT_project