The Google Playstore Apps Dataset

A Dataset derived from information collected by Python and Scrapy running on a cloud vm instance.

This is a huge dataset containing information about google play store apps which was collected from Kaggle(https://www.kaggle.com/gauthamp10/google-playstore-apps). Originally this dataset was uploaded by the author in his github here. Regarding source of data, the author stated that he had collected the data with the help of Python and Scrapy running on a cloud vm instance on June 2021.

Dataset Naming

The name for this dataset is Google-Playstore.

Miscellaneous Details

📢 Origin

The origin of the Google Playstore Apps data is **Kaggle**.

📢 Usage

This dataset may be used for **Assessment** and **Predictions**.

📢 Number of Cases

The dataset contains a total of **600K** cases.

📢 Variables

Each case contains **23** attributes:

- App Name
- App Id
- Category
- Rating
- Rating Count
- Installs
- Minimum Installs
- Maximum Installs
- Free
- Price
- Currency
- Size
- Minimum Android
- Developer Id
- Developer Website
- Developer Email
- Released
- Privacy Policy

- Last Updated
- Content Rating
- Ad Supported
- In app purchases
- Editor Choice

👉 Modifications

Our aim is to use 550 cases out of 600K and only **5** columns out of these 23 for each case:

- Category
- Size
- Installs
- Rating Count
- Rating