

## **2018 Google play store apps performance and trends analysis**

### **Introduction**

With the rising of smartphone usage and internet connectivity, there has been a surge of mobile applications favored by technological advancements such as robust operating systems and an increase in the number of application developers. Understanding a better niche for successful apps is necessary for app development organizations to develop business strategies and developers to make better app features. This project will provide essential analytics of app trends and credentials of mobile apps with aim of breaking through the app market.

### **Proposed Project**

During this study, we will explore data about published apps found on the Google Play store in 2018 with the aim of harnessing factors of a successful mobile application based on app metrics such as category, the number of installs, and price as well as qualitative and quantitative descriptors from user reviews. I will use two datasets:

1. Google play store app list contains the app's name, Category, Rating, Reviews score, Size, Installs, Type, Price, Content Rating, Genre, when it was lastly updated, its current version, and Android versions supporting it. This dataset will be used to explicitly find out app performance metrics and analyze the trend-setting metrics.
2. Google Playstore app user reviews dataset includes app's name, English-translated review, sentiment, and sentiment Polarity. I will use the dataset to find out keywords related to app descriptors from the app list dataset. This would help to investigate the reason for app performance based on what users said.

All datasets were sourced from <https://www.kaggle.com/lava18/google-play-store-apps> That is from scraped data from the Google Play store in 2018 by Lavanya (username: lava18) on Kaggle.com.

### **Potential Impact & Risk**

The potential risk of the project is a limited number of metrics to draw high fidelity conclusions and impure data entries. To minimize the latter, we will employ data validation techniques including direct inspection of data from the original data source, Google Playstore.

### **Project timeline**

Step1: Data formatting

Step2: Data validation and cleaning

Step3: Data analysis

Step4: data visualization