

# Google Play Store

Rubin Thomas, Maria Perez de la Llave and Meredith Dobrzynski

rt359, mp668, mmd272

## Question

What factors do customers weigh most heavily when deciding whether or not to download an app on the Google Play Store?

## Importance

Developing successful applications is becoming increasingly difficult as competition grows. Currently, there are around three million apps available on the Google Play Store, creating a saturated market. In order to stand out, developers need better insight into customers' demand. More specifically, developers need to know which factors play the biggest role in determining which apps customers download or purchase. Our objective is to help ensure a higher number of installations.

## Data

Google Play Store Apps (<https://www.kaggle.com/lava18/google-play-store-apps>)

The dataset we have chosen will allow us to explore the gaps in the market and the patterns which make apps successful in the Google Play store because:

1. Data Size: 10,841 entries
2. Data Coverage: Genre, Ratings(Out of 5), Content Rating(Audience), Size, Reviews(Count) to name a few of the 13 columns

The large data size will allow us to adequately analyze the question at hand, meaning our model will have a large enough training, validation and test set. On the other hand, it will not be too large to work with, as the population size of 3 million would be.

Ample data coverage will allow as us to set the number of "Installs" as our dependent variable and examine how the rest of independent variables above impact that value. Each independent variable can potentially be modeled with a different method depending if the variable is continuous, discrete, etc.