Problem Statement:

Train a model to predict the app ratings.

Also Visualize the distribution and correlation between the features.

Data Description:

googleplaystore.csv

================

App-Application name

Category-Category the app belongs to ss

Rating-Overall user rating of the app (as when scraped)

Reviews-Number of user reviews for the app (as when scraped)

Size-Size of the app (as when scraped)

Installs-Number of user downloads/installs for the app (as when scraped)

Type-Paid or Free

Price-Price of the app (as when scraped)

Content Rating-Age group the app is targeted at - Children / Mature 21+ / Adult

Genres-An app can belong to multiple genres (apart from its main category). For eg, a musical family game will belong to Music, Game, Family genres.

Last Updated-Date when the app was last updated on Play Store (as when scraped)

Current Ver-Current version of the app available on Play Store (as when scraped)

Android Ver-Min required Android version (as when scraped)

googleplaystore\_user\_reviews.csv

=============================

App-Name of app

Translated\_Review-User review (Preprocessed and translated to English)

Sentiment-Positive/Negative/Neutral (Preprocessed)

Sentiment\_Polarity-Sentiment polarity score

Sentiment\_Subjectivity-Sentiment subjectivity score