Project Description

Business Context

The Play Store apps data has enormous potential to drive app-making businesses to success. Actionable insights can be drawn for developers to work on and capture the Android market. Each app (row) has values for category, rating, size, and more. Another dataset contains customer reviews of the android apps. Explore and analyse the data to discover key factors responsible for app engagement and success.

Main Libraries to be used:
- Pandas for data manipulation, aggregation
- Matplotlib and Seaborn for visualisation and behaviour with respect to the target variable. Use at least 5 different visualisations.
- NumPy for computationally efficient operations

Play Store Data Description

Application Name	The name of the mobile application.
Category	The primary category to which the app belongs (e.g., Games, Productivity, Education).
Rating	The overall user rating of the app, represented as a numeric value (e.g., 4.5).
Reviews	The total number of user reviews for the app, as recorded when the data was scraped.
Size	The size of the app file (e.g., 15MB, 50MB), as recorded when the data was scraped.
Installs	The total number of times the app has been downloaded/installed by users, as recorded when the data was scraped.
Type	The cost of the app if it is paid, represented as a numeric value (e.g., $2.99). If the app is free, the price will be 0.
Price	The cost of the app if it is paid, represented as a numeric value (e.g., $2.99). If the app is free, the price will be 0.
Content Rating	The age group the app is targeted at (e.g., Everyone, Teen, Mature 21+, Adult).
Genres	Additional categories or genres the app belongs to, apart from its main category (e.g., Musical, Family, Game).

User Reviews

App	The name of the mobile application being reviewed.
Translated_Review	The user review of the app, which has been preprocessed (e.g., cleaned, tokenized) and translated into English.
Sentiment	The sentiment of the user review, categorized as Positive, Negative, or Neutral, based on text analysis.
Sentiment_Polarity	A numeric score representing the sentiment polarity of the review, ranging from -1 (most negative) to 1 (most positive).
Sentiment_Subjectivity	A numeric score indicating the subjectivity of the review, ranging from 0 (objective) to 1 (subjective), showing how subjective or opinion-based the review is.