



# PROBLEM STATEMENT

Google Play Store Data Analysis



# PROBLEM STATEMENT

The objective of this analysis is to perform an Exploratory Data Analysis (EDA) on a dataset containing information about apps available on the Google Play Store.

The goal is to extract valuable insights that can guide app developers, marketers, and businesses in understanding app performance, user preferences, and industry trends.

Specifically, the analysis focuses on evaluating app characteristics such as ratings, reviews, pricing models, and categories, as well as identifying any correlations or patterns that might influence user behavior and app success.

# BUSINESS PROBLEM OVERVIEW

The mobile app industry, particularly the Google Play Store, is a highly competitive and dynamic marketplace where app developers, businesses, and marketers are constantly striving to create high-quality apps that meet user needs and preferences.

With over 10,000 apps available across a wide range of categories, the challenge is not only to develop an app that users will love but also to ensure that it stands out in a crowded marketplace. Success in the app market is often determined by factors such as app quality, user ratings, number of reviews, category performance, and the pricing model (free vs. paid).

For businesses and developers, understanding the factors that drive user satisfaction and app success is essential to making data-driven decisions that optimize app development, marketing strategies, and monetization approaches.

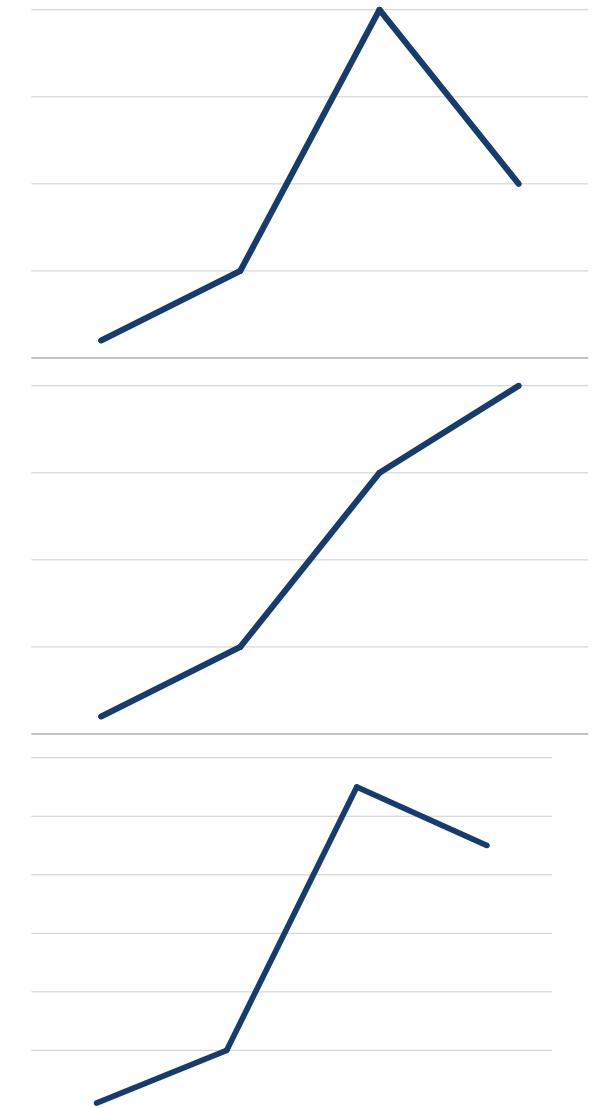
However, without in-depth insights into the performance and trends within the Play Store, businesses may struggle to identify profitable niches, understand user preferences, or allocate resources efficiently.



The Google Play Store is an official digital distribution platform for mobile apps developed for Android devices. It serves as the primary marketplace where users can browse, download, and install applications (apps) for their Android smartphones, tablets, and other compatible devices. The Play Store is operated by Google and offers a wide range of content, including apps, games, music, books, movies, and more.

In the context of the provided Exploratory Data Analysis (EDA) code, the dataset is sourced from the Google Play Store and contains essential information about the apps available on the platform. The Google Play Store dataset is used to analyze various app attributes, such as app categories, ratings, reviews, pricing, and the number of installs.

The primary goal of this EDA is to uncover valuable insights from the Play Store app data, which can help businesses, developers, and marketers optimize their strategies for app development, user engagement, and monetization.



# UNDERSTANDING & DEFINING DATASET

# PROJECT PIPELINE

The project pipeline can be briefly summarized in the following steps:

- Data Understanding: Here, we need to load the data and understand the features present in it. This would help us choose the features that we will need for your final model.
- Exploratory data analytics (EDA): Normally, in this step, we need to perform univariate and bivariate analyses of the data, followed by feature transformations, if necessary. For the current data set, because Gaussian variables are used, we do not need to perform Z-scaling. However, you can check if there is any skewness in the data and try to mitigate it, as it might cause problems during the model-building phase.





# THANK YOU