Mobile Apps Playstore Data Analysis

Introduction:

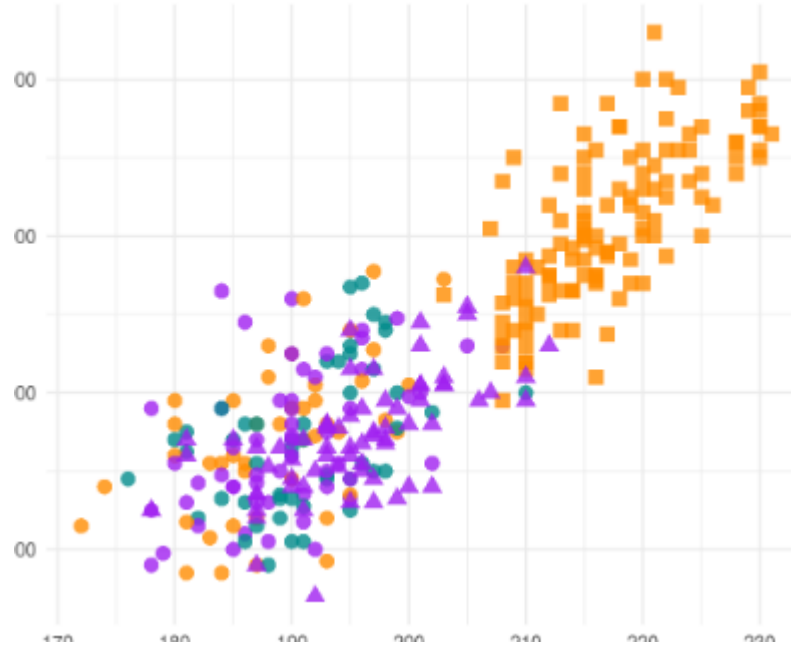
This data analysis project focuses on gaining insights into the future business approaches and development of mobile apps. By analyzing historical data from the Google Play store, we aim to understand app availability, usability, and users' expectations, providing valuable information for decision-making in the industry.

Objective:

The objective of this data analysis project is to extract valuable insights regarding mobile apps' future business approaches and development. By summarizing historical data from the Google Play store, spanning June 2012 to February 2019, we aim to understand the trends related to app availability, usability, and users' expectations. The primary focus is on generating a final dataset that is calculated and presented in a clear and comprehensible manner. This project aims to provide actionable information for decision-making in the mobile app industry and facilitate informed strategies for business growth.

Dataset Overview:

The dataset encompasses various data fields, including app names, categories, Android versions, ratings, reviews, installs, app size, last updated date, and current version. The data covers a timeframe from June 2012 to February 2019, providing a comprehensive view of mobile app information during that period.



Problem & Background

The dataset analysis reveals several issues, including the presence of blank fields and NaN (Not a Number) values. These problems pose challenges to accurate analysis as they can skew results and hinder the extraction of meaningful insights. Addressing these issues becomes crucial to ensure the reliability and validity of the findings. By handling blank fields and NaN values appropriately, we can eliminate data discrepancies and enhance the accuracy of the analysis, ultimately leading to more reliable conclusions and actionable recommendations for mobile app business approaches and development.

Methodology & Project Scope:

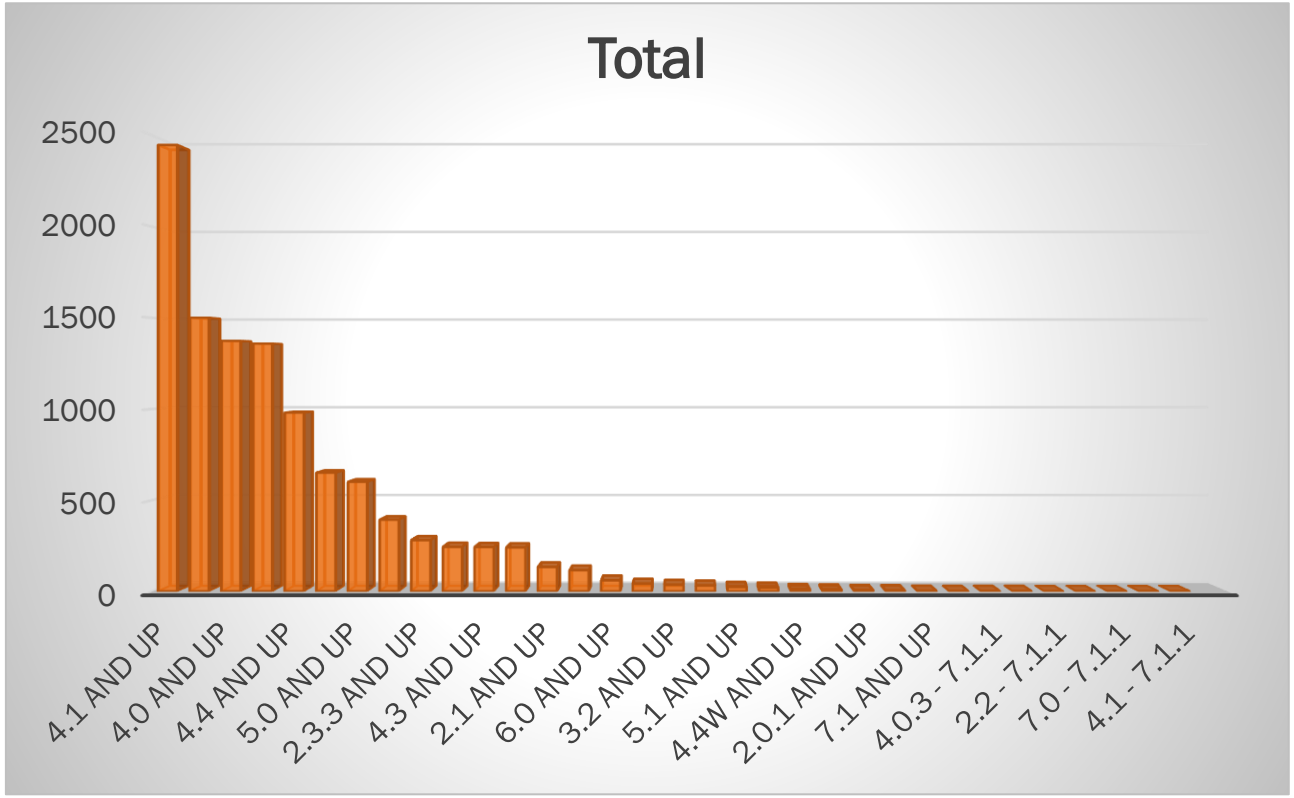
- Discuss the steps involved in the analysis process, including handling blank columns and addressing NaN values using the VLOOKUP formula
- Highlight the use of pivot tables to identify patterns and derive recommendations from the dataset

Recommended Analysis:

Q1: - Which is the most used Android version?

The most used Android version is 4.1 and up (2448).

To determine the most used Android version, we utilized the pivot table function by selecting the Android version as the key variable and calculating the count of Android instances. By employing the pivot table function, we were able to summarize and analyze the distribution of Android versions across the dataset. This analysis allows us to identify the Android version that appears most frequently, indicating the most commonly used version among the mobile apps in the dataset



Q2: - Which all categories of mob apps are existing from this dataset?

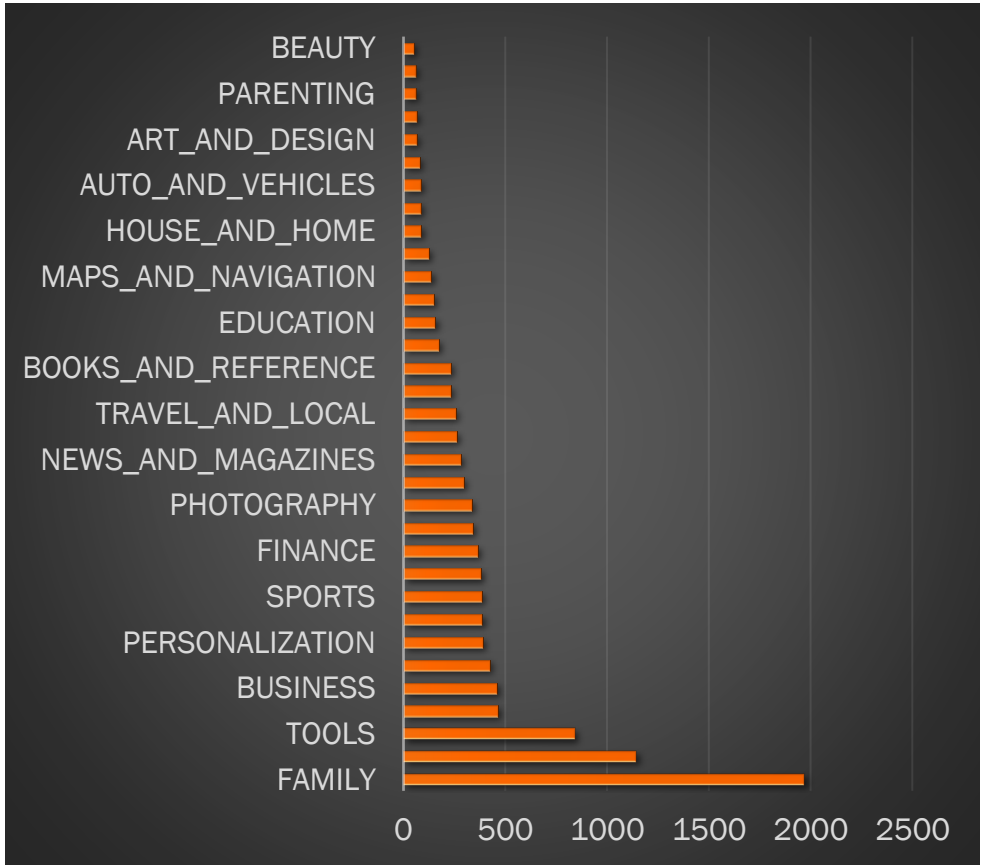
FAMILY	GAME	TOOLS	MEDICAL
BUSINESS	PRODUCTIVITY	PERSONALIZATION	COMMUNICATION
SPORTS	LIFESTYLE	FINANCE	HEALTH_AND_FITNESS
PHOTOGRAPHY	SOCIAL	NEWS_AND_MAGAZINES	SHOPPING
TRAVEL_AND_LOCAL	DATING	BOOKS_AND_REFERENCE	VIDEO_PLAYERS
EDUCATION	ENTERTAINMENT	MAPS_AND_NAVIGATION	FOOD_AND_DRINK
HOUSE_AND_HOME	LIBRARIES_AND_DEMO	AUTO_AND_VEHICLES	WEATHER
ART_AND_DESIGN	EVENTS	COMICS	PARENTING
BEAUTY			

To determine the categories of mobile apps existing in this dataset, we utilized the pivot table function. By setting the category as the key variable and calculating the count of apps within each category, we were able to summarize the data and identify all the unique categories present. This analysis provides insights into the different categories of mobile apps represented in the dataset, along with the number of apps falling under each category. Understanding the breadth of categories helps in gaining a comprehensive view of the app landscape and can inform decisions related to app development, market targeting, and user preferences.

Q3: - Which is the most installed category of mob app from this dataset?

To determine the most installed category of mobile app from the dataset, we utilized the pivot table function. By selecting the category as the key variable and calculating the sum of installs for each category, we were able to analyze the installation counts across different app categories. Based on the pivot table analysis, we observed that the "Family" category had the highest number of installations, indicating its popularity among users. This insight helps in understanding the app preferences and user behavior, enabling developers and businesses to strategize and target the family-oriented app market effectively.

Family app (1969) most installed category of mob app from this dataset



Q4: - How many total categories are there?

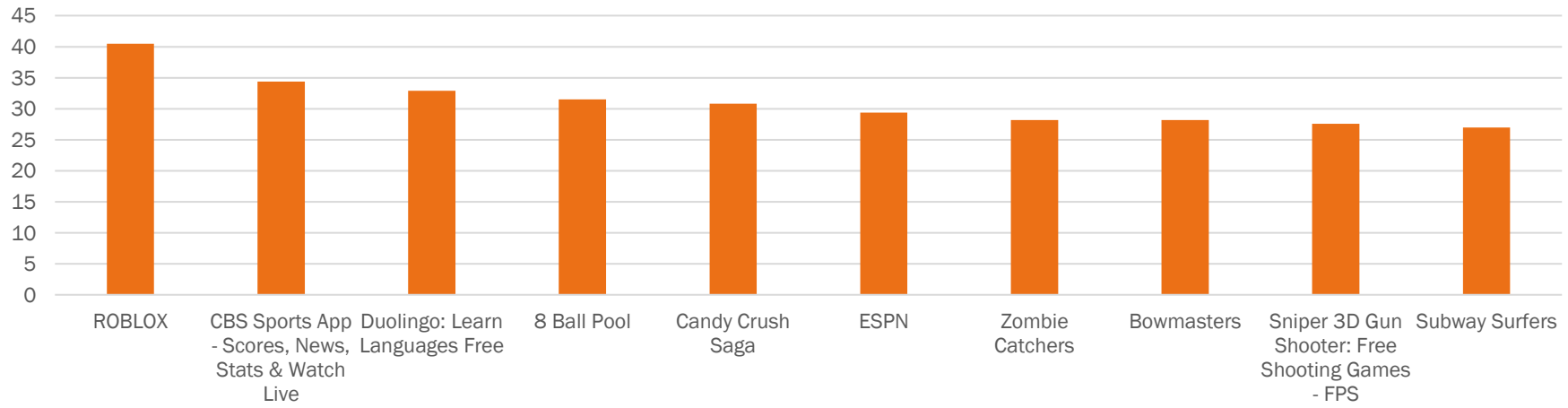
After analyzing the dataset, we have identified a total of 33 categories of mobile apps .

Q5: - Which apps have got 5-star ratings

After analyzing the data, it has been found that a total of 274 apps received a 5-star rating. These apps were highly regarded by users, indicating their exceptional quality and positive user experiences. The significant number of apps receiving a 5-star rating suggests a strong presence of high-quality applications in the dataset, which can be considered as a positive indicator for app developers and users seeking top-rated apps.

To identify the apps that have received a 5-star rating, we can use the filter option in the dataset. By applying the filter and selecting the rating column, we can easily isolate and view only the entries with a 5-star rating. This filtering process allows us to quickly find the apps that have achieved the highest rating and obtain the total count of such apps. By leveraging the filtering capabilities of the dataset, we can efficiently answer the question and provide insights into the number and characteristics of apps with a 5-star rating.

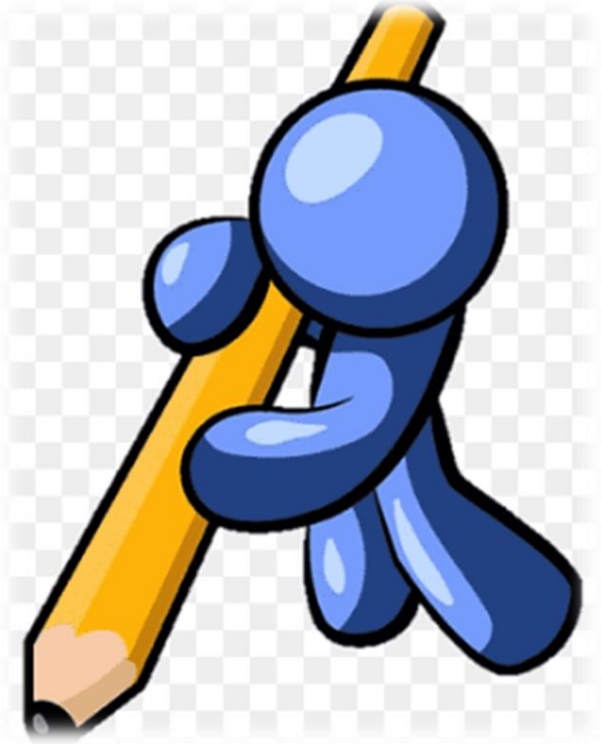
Q6: - Which are the top 10 mob apps based on ratings?



To determine the top 10 mobile apps based on ratings, we can utilize the value filter option in the pivot table. By selecting the rating column as the value field in the pivot table and applying the value filter, we can choose to display only the top 10 values. This filter will sort the apps based on their ratings and show the highest-rated apps at the top. By clicking on the filter and selecting the top 10 option, we can instantly see the names of the top 10 mobile apps based on their ratings

Conclusion:

The project aims to create a clear and accurate dataset by handling blank and NaN values. Pivot tables are utilized to extract insights for informed mobile app business approaches. The goal is to provide valuable information for decision-making in the industry.



Project Owner: Shubham koundal