Project Overview

This project explores factors influencing app ratings on the Apple Store. Using SQL, we analyzed app categories, pricing models, language support, description length, and market competition to derive insights that can help developers improve their app ratings and visibility.

Dataset Source

- The dataset was obtained from **Kaggle** and consists of **two tables**:
  - **AppleStore**: Contains app metadata, ratings, pricing, and category information.
  - **AppleStore_Description_Combined**: Merges multiple description tables to analyze the impact of description length.

SQL Analysis & Key Insights

1. Checking Data Quality

- Counted unique apps in both tables to ensure consistency.
- Identified missing values in key fields (app names, ratings, descriptions).
- Ensured data integrity before proceeding with analysis.

2. App Ratings Overview

- **Minimum Rating**: 0.0
- **Maximum Rating**: 5.0
- **Average Rating**: ~3.5

❖ **Implication**: Apps should aim for a rating above **3.5** to stand out.

```sql
-- Get an overview of Apps' ratings
SELECT min(user_rating) AS MinRating,
       max(user_rating) AS MaxRating,
       avg(user_rating) AS AvgRating
 FROM AppleStore
```

3. Do Paid Apps Have Higher Ratings?

- **Finding**: Paid apps have higher average ratings than free apps.
- **Recommendation**: Developers should consider a freemium or premium model if their app delivers high value.

```sql
-- Check whether paid apps has higher rating than free apps

SELECT CASE
 WHEN price > 0 THEN 'Paid'
 ELSE 'Free'
 END AS app_type,
 avg(user_rating) AS avg_Rating
 FROM AppleStore
 GROUP BY app_type

```

4. Do More Supported Languages Improve Ratings?

- **Finding**: Apps supporting 10–30 languages receive the highest rating
- **Recommendation**: Focus on strategic language localization instead of maximizing translations.

```sql
 -- Check if apps with more supported languauge have higher rating
SELECT CASE
     WHEN  lang_num < 10 THEN '<10 languages'
      WHEN lang_num BETWEEN 10 AND 30 THEN '10-30 languages'
       ELSE '>30 languages'
       END AS Language_bucket,
       avg(user_rating) AS avg_Rating
 FROM AppleStore
 GROUP BY Language_bucket
 ORDER BY avg_Rating DESC

```

5. Which App Categories Have the Lowest Ratings?

- **Finding**: Finance and Book apps have the lowest ratings, indicating potential market gaps.
- **Recommendation**: Developers in these genres should improve user experience and address pain points.

```sql
--check genre with low ratings

SELECT prime_genre,
   avg(user_rating) AS avg_Rating
   FROM AppleStore
   GROUP BY prime_genre
   ORDER BY avg_Rating ASC
   LIMIT 10

```

6. Does Description Length Affect Ratings?

- **Finding**: **Longer app descriptions** correlate with higher ratings.
- **Recommendation**: Apps should have detailed, well-structured descriptions to set clear user expectations.

```sql
-- check if there is a correletion betwen length of the app description and user rating
SELECT CASE
 WHEN length(b.app_desc) <500 THEN 'Short'
 WHEN length(b.app_desc) BETWEEN 500 AND 1000 THEN 'Meduim'
 ELSE 'Long'
 END AS Description_length_Bucket,
 avg(a.user_rating) AS avg_Rating

FROM AppleStore AS A
JOIN
appleStore_description_Combined AS B
ON
A.id = B.id

GROUP BY Description_length_Bucket
ORDER BY avg_Rating DESC

```

7. Which Apps Are Top-Rated in Each Genre?

- **Finding**: Extracted highest-rated apps per genre using SQL ranking functions.
- **Recommendation**: New apps should analyze top performers in their category to understand what drives high ratings.

```sql
--Check the top rated apps for each genre
SELECT
prime_genre,
track_name,
user_rating

FROM (
   SELECT
        prime_genre,
        track_name,
        user_rating,
        Rank() OVER (PARTITION BY prime_genre ORDER BY user_rating DESC, rating_count_tot DESC)
        AS Rank

        FROM AppleStore
  ) AS A
WHERE
A.rank = 1

```

Final Key Insights

This project presents a data-driven analysis of app ratings, highlighting key factors that influence user satisfaction. Using SQL, we explored trends in app performance based on various attributes such as pricing, language support, category, description length, and competition levels.

1. Paid Apps Have Better Ratings

Our analysis indicates that paid apps generally receive higher ratings compared to free apps. This trend suggests that users who pay for an app may engage with it more actively and perceive greater value, leading to better reviews. If an app maintains high quality, introducing a paid model could be a viable strategy.

2. Apps Supporting 10–30 Languages Perform Better

Interestingly, apps that support between 10 and 30 languages tend to achieve the highest average ratings. This suggests that strategic language selection is more important than sheer quantity. Developers should focus on including languages that best align with their target audience.

3. Finance and Book Apps Tend to Have Lower Ratings

Our data shows that finance and book apps generally receive lower ratings. This indicates potential gaps in user satisfaction, presenting an opportunity for improvement. Creating high-quality apps in these categories that address user needs more effectively could lead to better ratings and market penetration.

4. Longer App Descriptions Correlate with Higher Ratings

Apps with detailed and well-crafted descriptions tend to receive better ratings. Users appreciate clear and comprehensive information about an app’s features before downloading. Providing a well-structured description can help set expectations and improve user satisfaction.

5. Aiming for an Average Rating Above 3.5

The average app rating across all categories is 3.5. To stand out in the competitive market, new apps should aim for ratings above this threshold. A strong rating increases visibility and credibility, making the app more appealing to users.

6. High Competition in the Games and Entertainment Category

The games and entertainment sector is highly saturated, with a vast number of apps competing for user attention. While this category has strong demand, entering the market can be challenging. Success in this space requires unique offerings and innovative strategies to differentiate from competitors.

Conclusion

This SQL-based data analysis provides actionable insights into what makes an app successful in terms of ratings. By considering factors such as pricing strategy, language support, app description quality, and category-specific challenges, developers can make informed decisions to enhance user satisfaction and improve their app’s market performance.

This project showcases data-driven decision-making in app development, demonstrating proficiency in SQL and analytical reasoning.