The Advent of Smartphone Applications Insights from a Dataset of 10,000+ Apps

Name: K Lakshmi Nirmala
International Institute of Information Technology, Hyderabad

Overview of the Dataset

Our dataset includes 12,736 mobile applications, each represented by a row with detailed attributes. This extensive dataset provides valuable insights into various factors that influence app success. It covers key metrics such as ratings, reviews, downloads, and more.

Dataset Columns after filtering

"The dataset contains 10 essential columns, each contributing to our understanding of app performance:
For each varibale i included the importance in the jupyter notebook detially
Application: Name of the application.
Ratings: Average user rating on a scale of 5.
Reviews: Number of user reviews.
File Size: Disk size of the app.
Downloads: Total number of downloads.
Category: Indicates if the app is free or paid.
Purchase Price: Cost of the app if it’s a paid download.
Content Rating: Age rating for the app.
Genres: Genre or category of the app.
Android Supported: Android version compatibility

Objective of Analysis

"Our goal is to identify and analyze the key parameters driving the success of mobile applications. By examining factors such as ratings, downloads, reviews, and other attributes, we aim to determine which elements contribute most significantly to an app's success."

Data Preparation:

"To prepare the data for analysis, we performed several preprocessing steps:

Handled Missing Values: Used machine learning techniques to fill in missing ratings and manually updated null values in other columns. and for ratings there are lot of missing values so handled by created a ml model which is giving 0.2 mse which is i think better to fill missing rating values based on the remaining parameters

Data Transformation: Converted file size to MB, normalized Android version data, and cleaned purchase price data.

F**eature Engineering:** Created dummy variables for categorical data and removed unnecessary columns."

Key Parameters and Their Weights

"We assigned weights to different parameters based on their perceived impact on app success. The key parameters and their weights are:

Ratings: 25%

Content Rating: 20%

Downloads: 30%

File Size: 10%

Purchase Price: 5%

Review:5%

Genres: 5%"

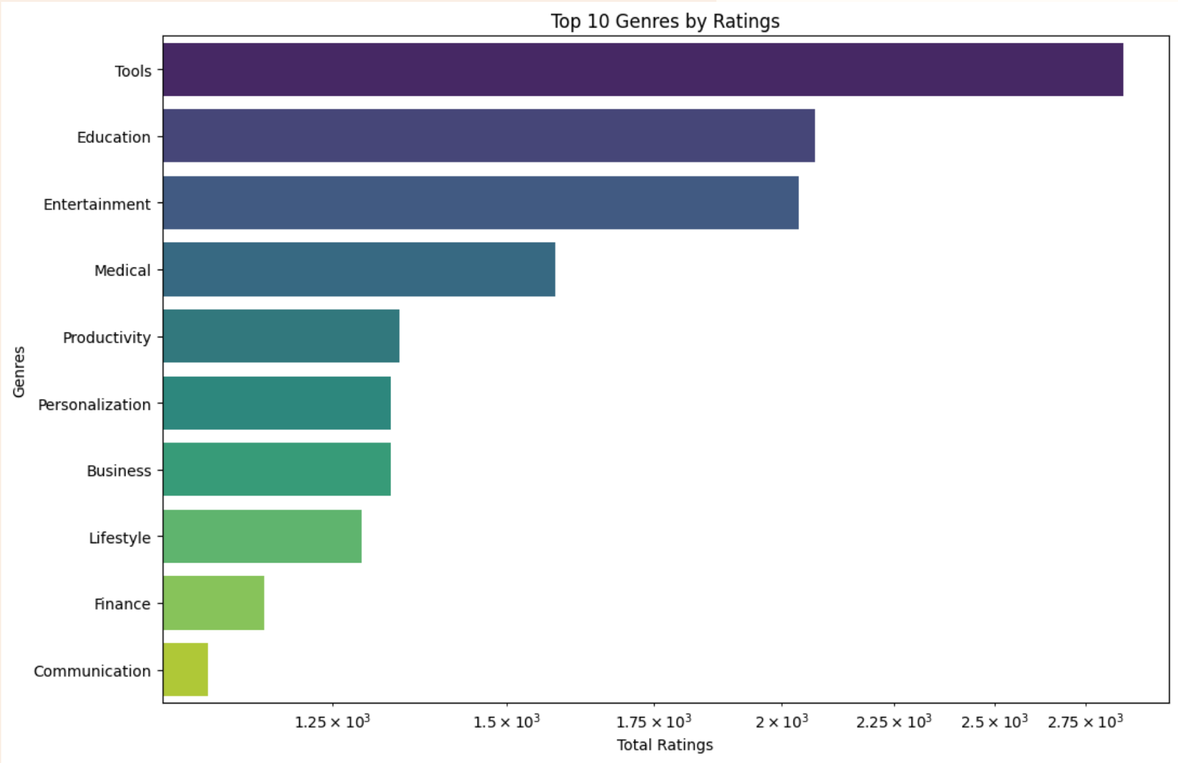
Insights and Visualizations

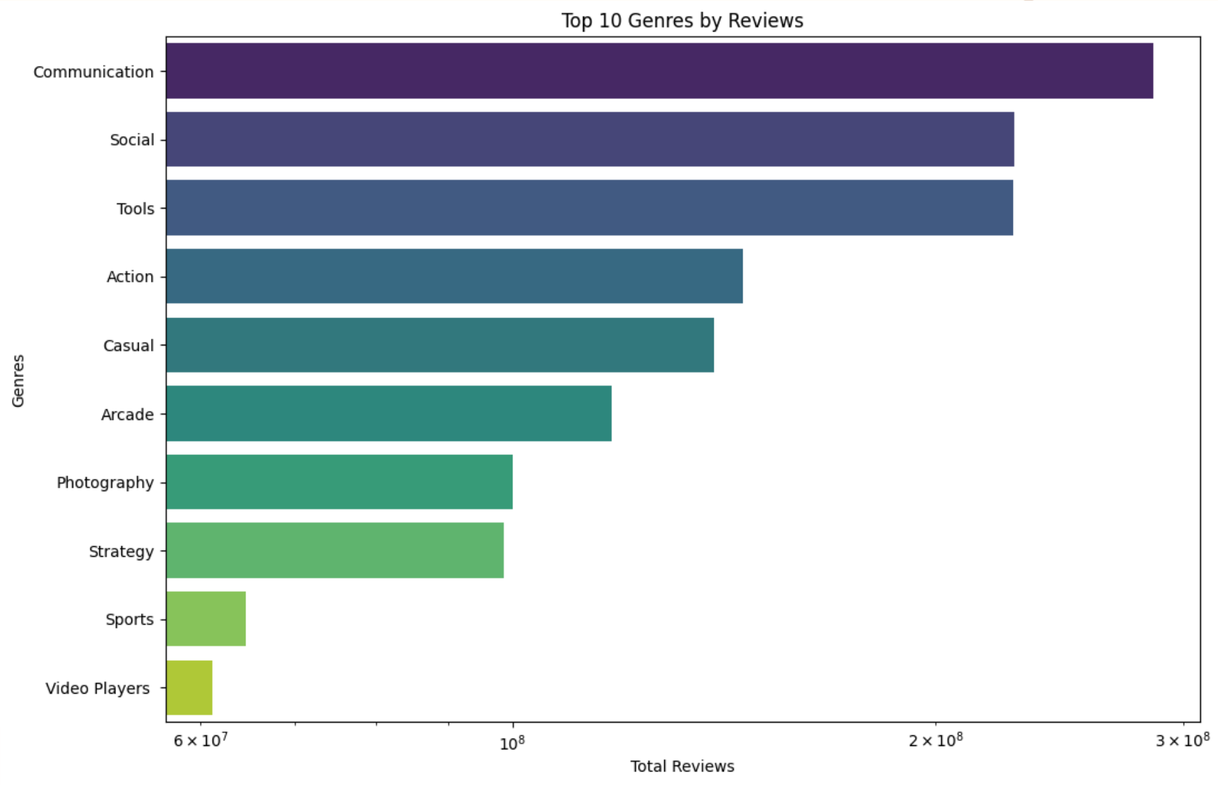
"The analysis reveals several insights:

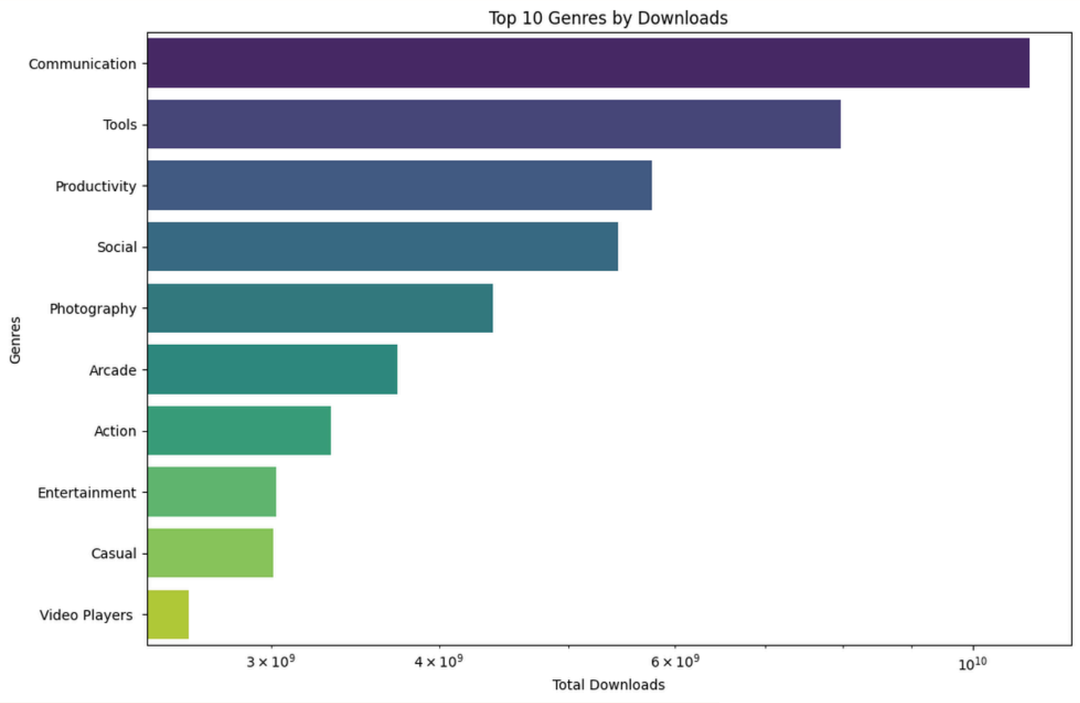
High Ratings and Downloads Correlate with Success: Apps with higher ratings and downloads tend to perform better.

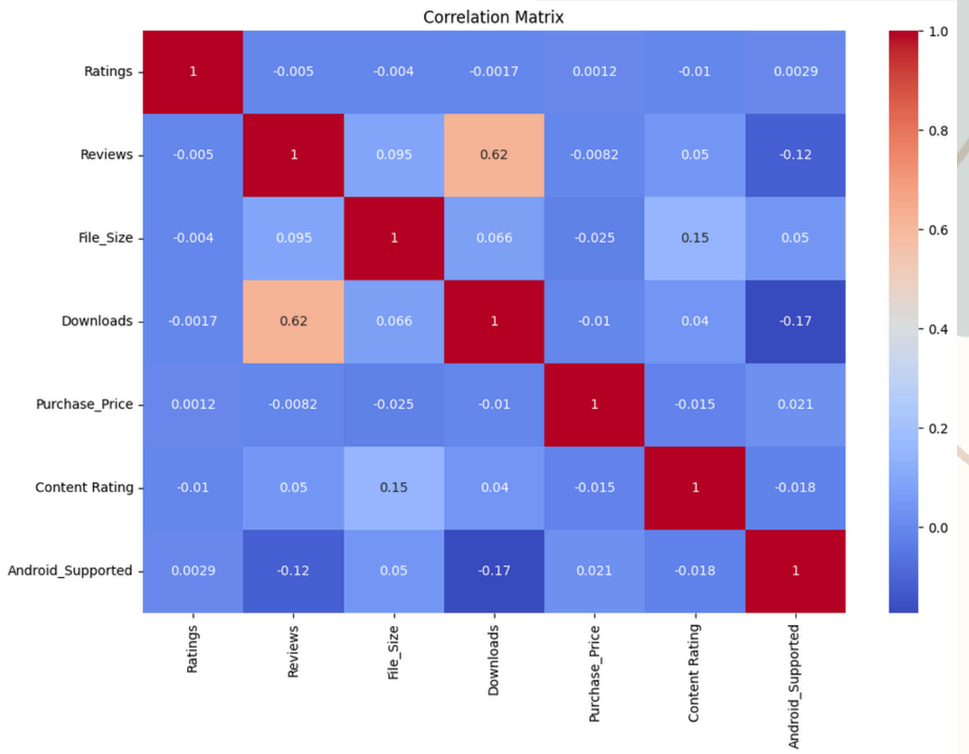
Free Apps Have More Downloads: Free apps generally have a larger user base compared to paid apps.

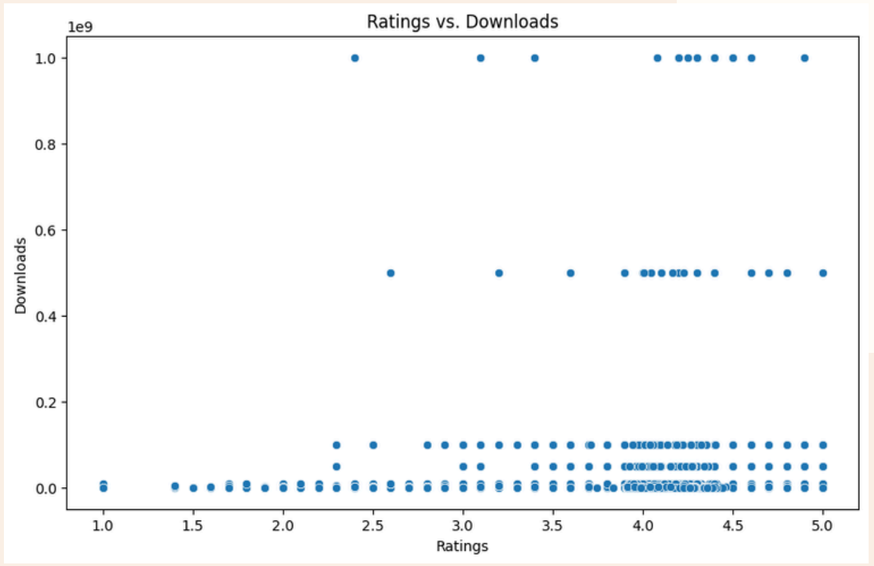
Genre Impact: Certain genres attract more downloads and reviews than others."

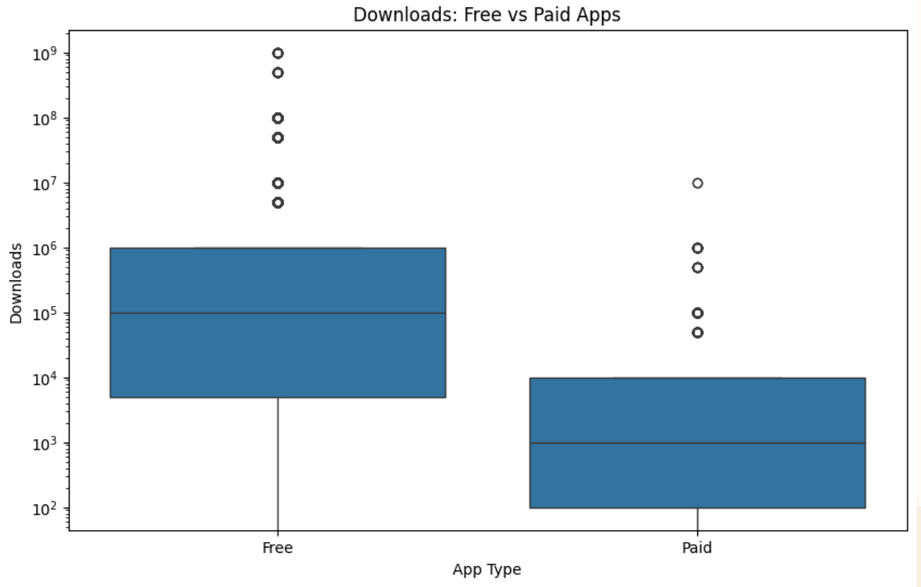





Top 10 Genres by Downloads
Communication
Tools
Productivity
Social
Photography
Arcade
Action
Entertainment
Casual
Video Players
Genres
3×10^9
4×10^9
6×10^9
10^{10}
Total Downloads


Correlation Matrix
Ratings 1 -0.005 -0.004 -0.0017 0.0012 -0.01 0.0029
Reviews -0.005 1 0.095 0.62 -0.0082 0.05 -0.12
File_Size -0.004 0.095 1 0.066 -0.025 0.15 0.05
Downloads -0.0017 0.62 0.066 1 -0.01 0.04 -0.17
Purchase_Price 0.0012 -0.0082 -0.025 -0.01 1 -0.015 0.021
Content Rating -0.01 0.05 0.15 0.04 -0.015 1 -0.018
Android_Supported 0.0029 -0.12 0.05 -0.17 0.021 -0.018 1
Ratings
Reviews
File_Size
Downloads
Purchase_Price
Content Rating
Android_Supported
1.0
0.8
0.6
0.4
0.2
0.0

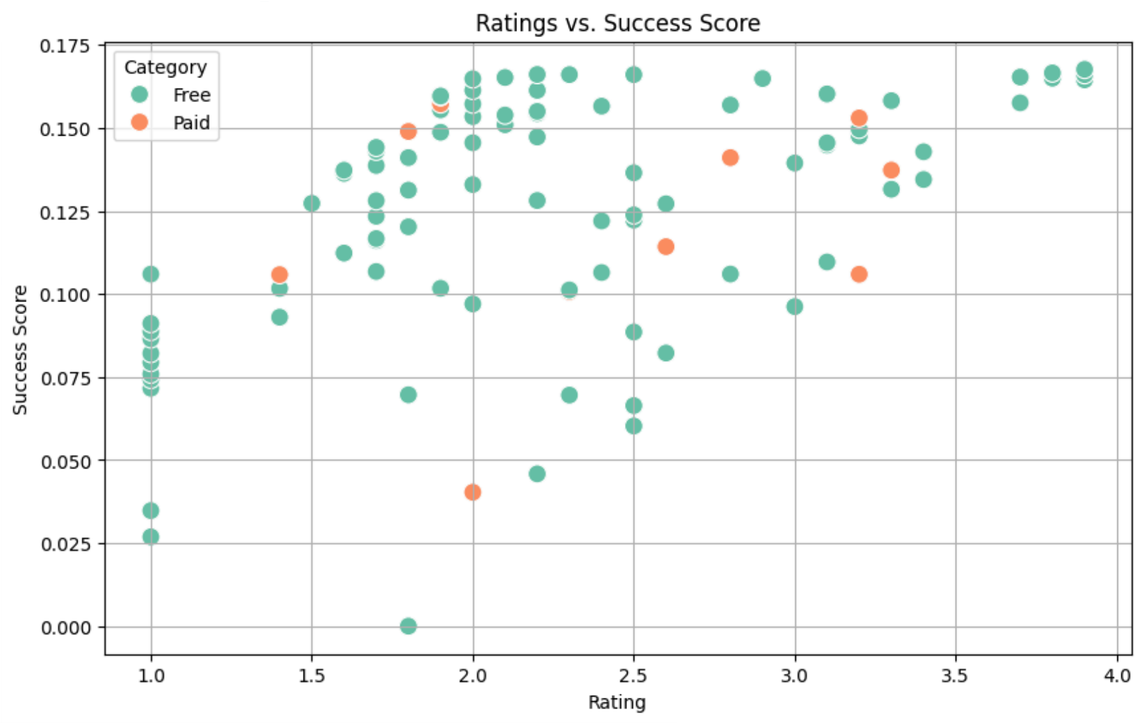

Ratings vs. Downloads
1e9
1.0
0.8
0.6
0.4
0.2
0.0
Downloads
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Ratings

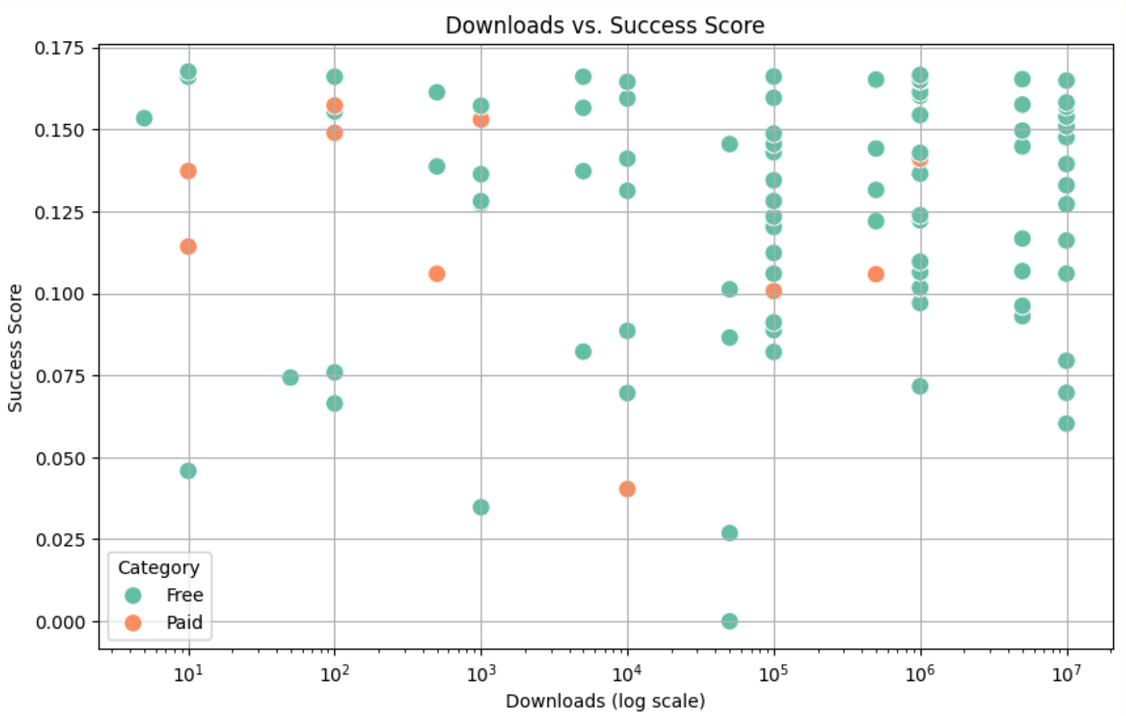

Downloads: Free vs Paid Apps
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
Downloads
Free
Paid
App Type

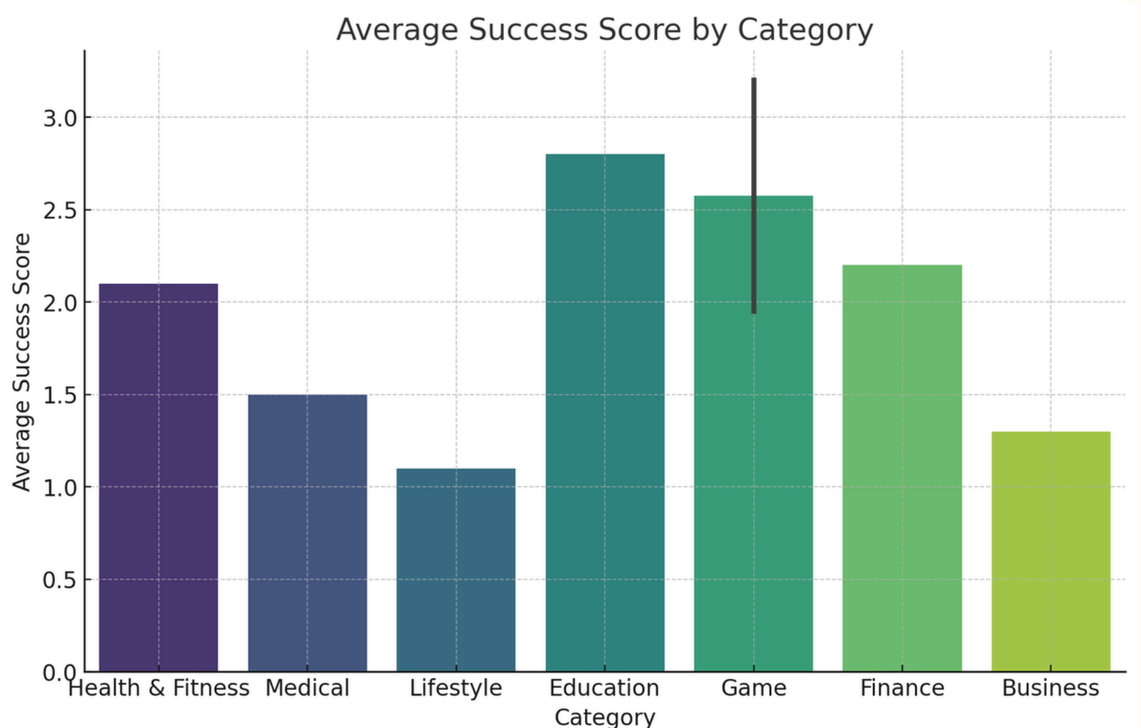
Reasons for Failure for Bottom 100 apps

1.Low Ratings and High Downloads:
Apps like "Miraculous Ladybug & Cat Noir - The Official Game" and "Sweet Baby Girl - Dream House and Play Time" show that high downloads do not necessarily equate to success if user ratings are low. These apps likely suffer from poor user experience or technical issues.
2. Low User Engagement:
Apps like "Chakra Cleansing" and "RadRevision: Anatomy on CT" have moderate downloads but extremely low user engagement, as indicated by minimal reviews and low ratings. This suggests a lack of compelling content or features.
3. File Size vs. Utility:
Apps such as "Audiowalk EB" and "Hojiboy Tojiboyev Life Hacks" have large file sizes relative to their utility, which likely contributes to their poor adoption rates.

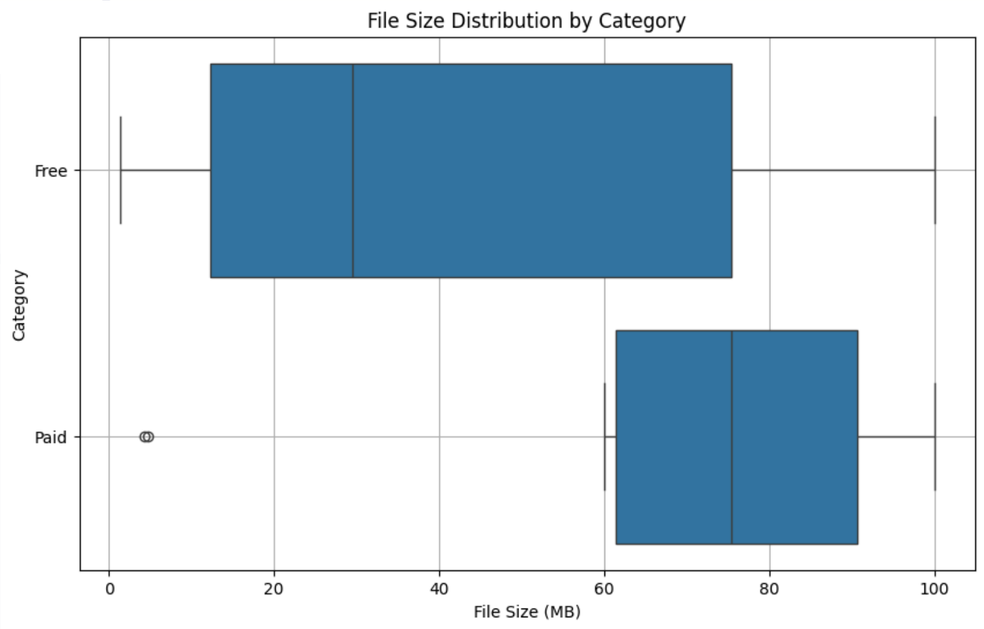
Reasons for Failure for Bottom 100 apps

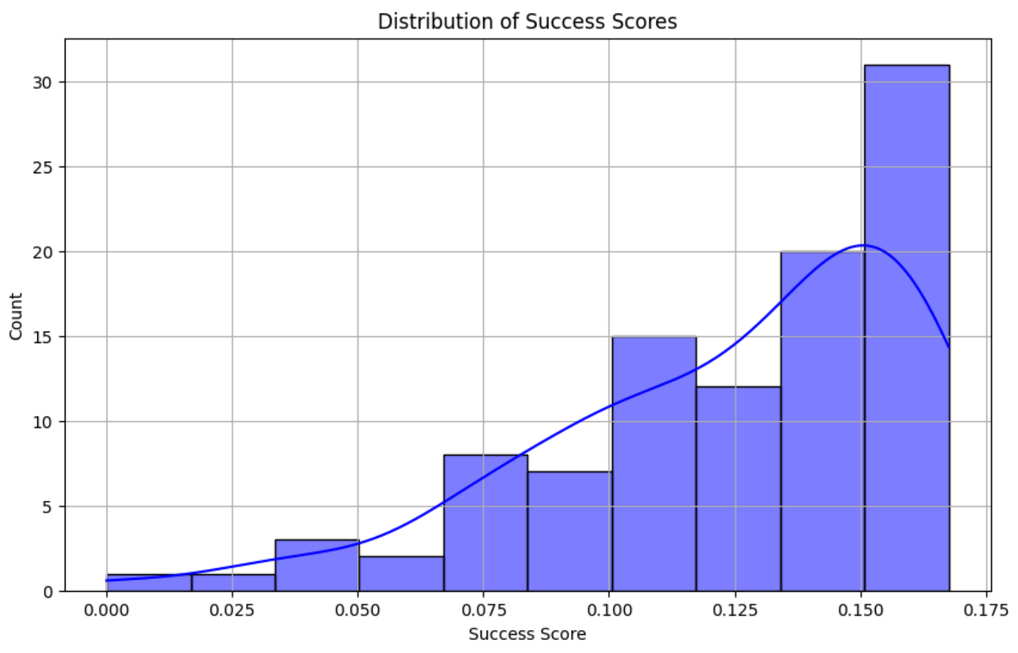




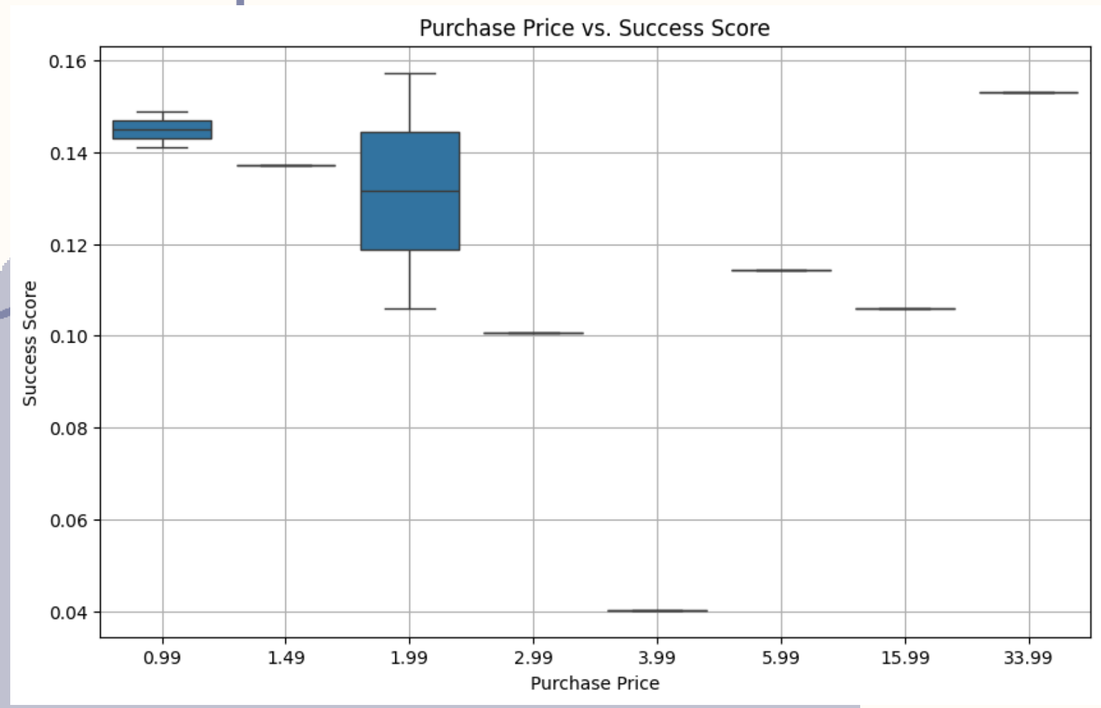


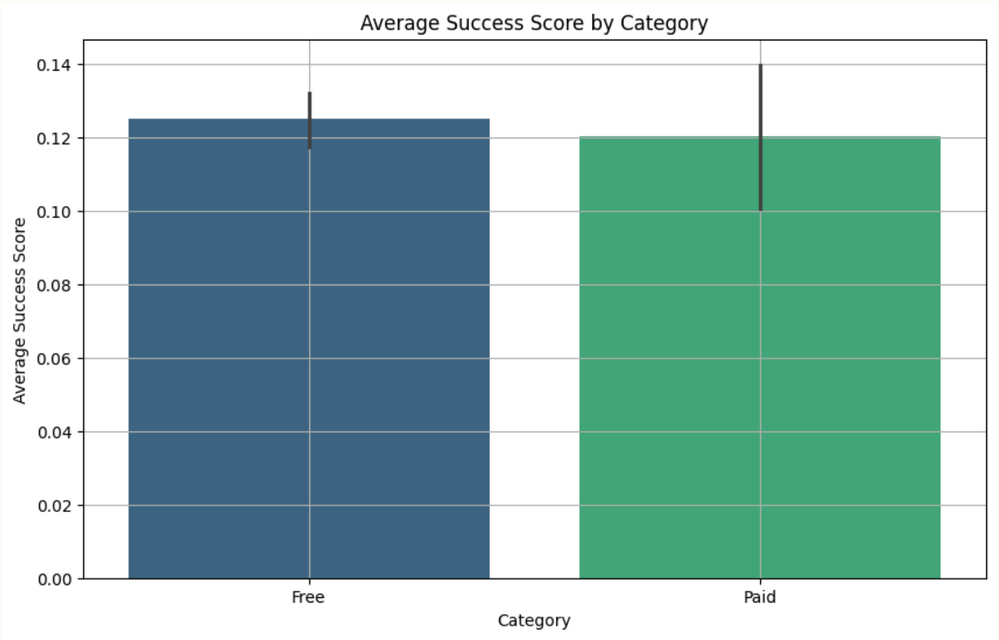
Reasons for Failure for Bottom 100 apps

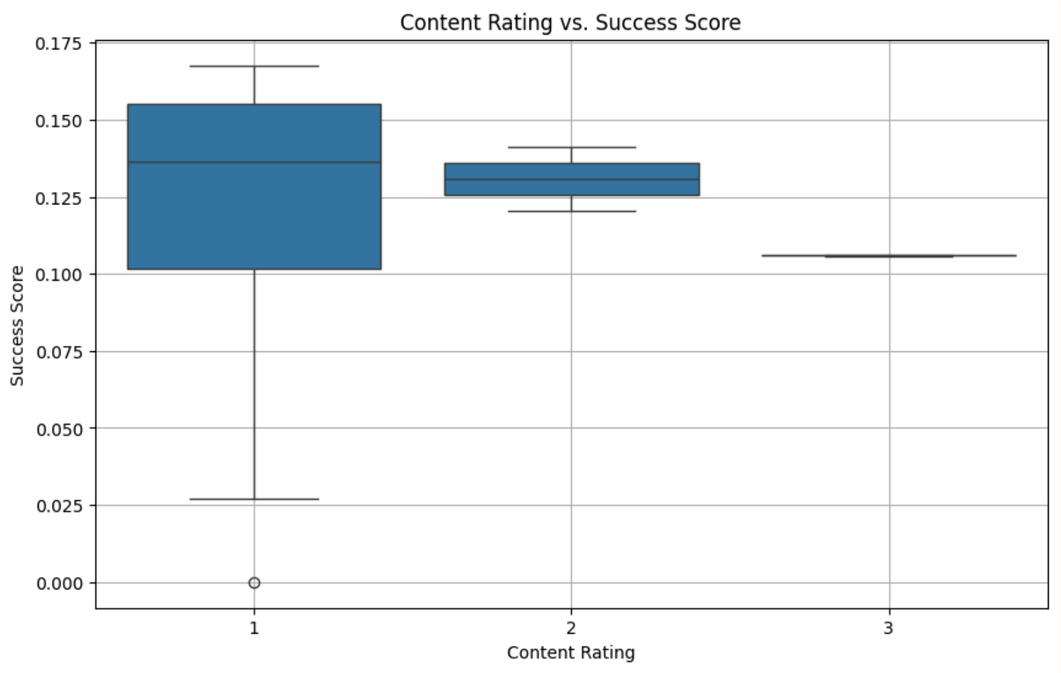




Reasons for Failure for Bottom 100 apps







Conclusions for successful app:

1. High User Ratings

Top Apps: Successful apps generally have high user ratings, typically above 4.0. High ratings often correlate with better user satisfaction, leading to higher success scores.

Recommendation: Focus on delivering a quality user experience and actively address user feedback to maintain high ratings.

2. Positive Correlation with Downloads

Top Apps: Apps with millions of downloads tend to have higher success scores, indicating that popularity and user base size are critical for success.

Recommendation: Invest in marketing strategies to increase app visibility and downloads.

3. Low or Reasonable Pricing

Top Apps: Successful apps often have free or low pricing. Paid apps with high success scores typically provide substantial value to justify the cost.

Recommendation: Consider offering free versions with in-app purchases or ensure that paid apps provide significant value to justify their price.

4. Optimized File Size

Top Apps: Successful apps often maintain an optimized file size that balances functionality with storage efficiency. Large file sizes can deter users with limited device storage.

Recommendation: Optimize app performance and file size to accommodate a broader user base, especially in regions with limited device storage.

5. Targeted Content Rating

Top Apps: Apps with content ratings like "Everyone" or "Teen" tend to have higher success scores. These ratings broaden the potential user base.

Recommendation: Ensure that the app's content rating aligns with the intended audience to maximize reach.

Conclusions for successful app:

6. Effective Use of Categories
Top Apps: Certain categories like "Productivity," "Health & Fitness," and "Social" have a higher concentration of successful apps. Choosing the right category can enhance visibility.
Recommendation: Select an appropriate category that matches the app's functionality and target market.
7. High Engagement (**Reviews and Active Users**)
Top Apps: Apps with higher numbers of reviews and active users typically score better. This engagement reflects a strong user community and ongoing relevance.
Recommendation: Foster user engagement through regular updates, community building, and responding to reviews.
8. Strong Correlation with Age
Top Apps: Apps that maintain relevance over time by staying updated with new features and trends tend to score higher.
Recommendation: Regularly update the app to keep it relevant and maintain user interest.
9. High-quality Visual and Functional Design
Top Apps: Visual appeal and intuitive design often contribute to higher ratings and, consequently, higher success scores.
Recommendation: Invest in professional UI/UX design to make the app both aesthetically pleasing and easy to use.
10. Social Proof and Branding
Top Apps: Strong branding and positive word-of-mouth can significantly enhance an app's success. Well-known brands or apps with significant social proof tend to have higher success.
Recommendation: Build a strong brand presence and encourage positive user testimonials and reviews.

Thank you!

For giving this Opportunity