Kickstarter Cluster Model

KO-JEN WANG

Overview

Our endeavor in this data science initiative is to meticulously dissect and comprehend the intricate array of attributes that categorize Kickstarter's diverse project spectrum. We aim to extract actionable insights that not only illuminate the patterns of success and failure but also serve as a compass for navigating the crowdfunding domain.

Clustering Model

By analyzing successful and failed Kickstarter projects, I refined our clustering approach and improved the silhouette score from 0.06 to 0.24. I used K-means to identify five distinct clusters and applied PCA and t-SNE for visualization, integrating our findings back into the original dataset for deeper insight. Below is an overview of cluster characteristics and insights.

Cluster 0 — The Diverse Pioneers (473 projects): Cluster 0 exhibits a rich blend of geographic diversity, with significant contributions from the US, GB, and CA. This cluster is characterized by a higher median name and blurb length, suggesting a tendency for more descriptive and possibly complex projects. Despite being the smallest cluster, its broad category range, particularly in tech, and the even distribution of project creation and launch times, indicate a group of early adopters who may not follow conventional Kickstarter trends. Projects here show varied outcomes, presenting a balanced mix of success and failure, which could be pivotal for understanding the nuances of project impact factors.

Cluster 1 — The British Innovators (3288 projects): Dominance in Great Britain sets this cluster apart, with a significant number of projects falling under the web, software, and hardware categories. Despite its large size, Cluster 1 has a lower median goal but a high median pledged amount, indicating a strong success rate with backers. Projects in this cluster have a relatively shorter period from creation to launch, which may suggest better pre-launch preparation or more streamlined project planning. The success rate here is notable, and examining the key factors such as category choice, goal setting, and campaign duration could yield actionable insights for new creators.

Cluster 2 — The American Mainstream (7260 projects): This US-exclusive cluster is the largest, suggesting Kickstarter's strong foothold in the American market. Categories such as hardware, web, and software dominate, reflecting the backers' interests. A pronounced spike in projects created during the afternoon and launched on Tuesdays suggests that creators are tapping into optimal times for engagement. Furthermore, the shorter launch-to-deadline duration indicates a trend towards more concise campaigns, which may resonate better with backers. The success rate in this cluster could be instrumental in shaping Kickstarter's marketing strategies and advising creators on best practices.

Cluster 3 — The Tech Trendsetters (2409 projects): Tech categories also lead in Cluster 3, which exhibits a notable balance in launch days but a preference for afternoon hours. Interestingly, projects here have a wider range of goal settings but maintain a moderate pledged amount, suggesting that while ambitions are high, the actual funding aligns with Kickstarter averages. The success rate is mixed, presenting an opportunity to delve deeper into what differentiates the successful campaigns from those that don't meet their goals.

Cluster 4 — The Niche Visionaries (5 projects): The projects have longer names and blurbs, which could imply a focus on niche, possibly innovative ideas that require more explanation. This cluster sees the highest goals set but does not necessarily correlate with a high pledge amount, indicating a potential misalignment with backer expectations or project feasibility. The insights from this cluster are particularly valuable as case studies due to their outlier nature, and

successes here could be a result of unique or disruptive project ideas.

Additional Business Insights:

1. Project Description Influence: Longer descriptions in Clusters 0 and 4 could either engage or overwhelm backers, affecting funding.

2. Campaign Duration and Success: Shorter, well-planned campaigns in Clusters 1 and 2 show higher success, emphasizing the importance of goal clarity and campaign brevity.

3. Seasonality and Launch Timing: Afternoon launches are common across clusters, with summer and fall being popular seasons, potentially reflecting peak backer activity.

4. Factors for Success or Failure: Strategic launch timing, realistic goal setting, and focused campaign lengths are common in successful projects, particularly in Clusters 1 and 2.