

Kickstarter Campaign Analysis

Jessica Chen, Jake Flores, Andrew Mike, Matthew Mutammara
STAT 605: R For Data Science | Spring 2019



Introduction

Kickstarter is a crowdfunding platform for creative projects. Ideas for films, art, technology, and more can be pitched as Kickstarter campaigns in hopes of gathering enough community support to fund the project. **Kickstarter's "All or Nothing" rule** states that a creator can only collect the funds if their funding goal has been reached by the deadline. If the goal is not reached in time, money is returned to backers. Our research was motivated by a simple question: **What factors contribute to Kickstarter campaign success?**

Data Overview

Primary dataset [1]:

Number of rows: 2,553,392

Number of columns: 37

Timespan of Data: January 2018 - December 2018

Important columns:

- Campaign category
- Success/failure
- Campaign description
- URL for campaign webpage

Secondary dataset [2]:

Our secondary dataset introduces a new topic: reward levels. Backers can contribute to campaigns in different dollar amounts, each of which is associated with a different reward from the campaign. This dataset gives us the number of reward levels offered by each campaign.

Text Analysis

Each Kickstarter campaign includes a short description of the project. We searched for keywords in the descriptions of successful and failed campaigns and discovered the following results:

"First" is found in 5.6% of successful campaign descriptions, compared to 3.4% of failed campaigns.

Exclamation Marks ("!") are found in 26.7% of successful campaign descriptions, compared to 21.8% of failed campaigns.

While these are notable numbers for single words, the overall impact of any given word in the description seems to be negligible.

Conclusions

Category: Music and comics campaigns perform well. Art and journalism perform poorly.

Reward Levels: More possible reward levels is associated with higher rates of success. Campaigns that offer more donation options make it easier to donate and, as a result, see higher levels of success.

Comments: More comments is associated with higher rates of success. Campaigns should encourage and facilitate discussions about their project to increase community interest.

Description: Some words and punctuation were associated with small changes in campaign success. Overall, keywords don't seem to have much impact. Future analyses may consider sentiment analysis as a more significant predictor of success.

References

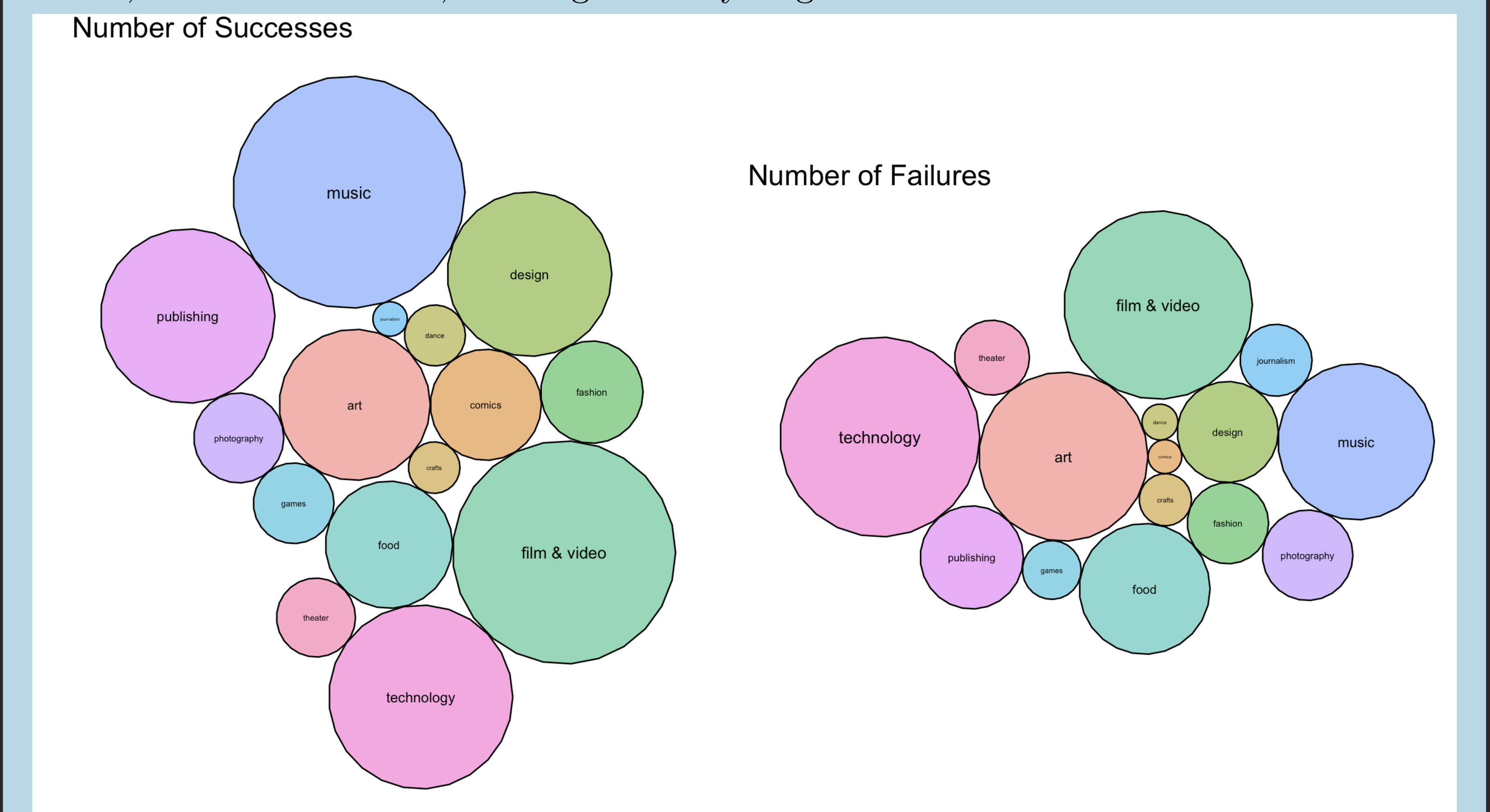
- [1] <https://webrobots.io/kickstarter-datasets/>
- [2] <https://github.com/nalamidi/Kickstarter-Predictive-Analysis/tree/master/Kickstarter>

Exploratory Data Analysis

Kickstarter campaigns are separated into categories based on the content of the project being funded. The bubble plot below includes the 15 most common categories. Each bubble and color corresponds to a category, while the size of each bubble corresponds to the number of successes or failures.

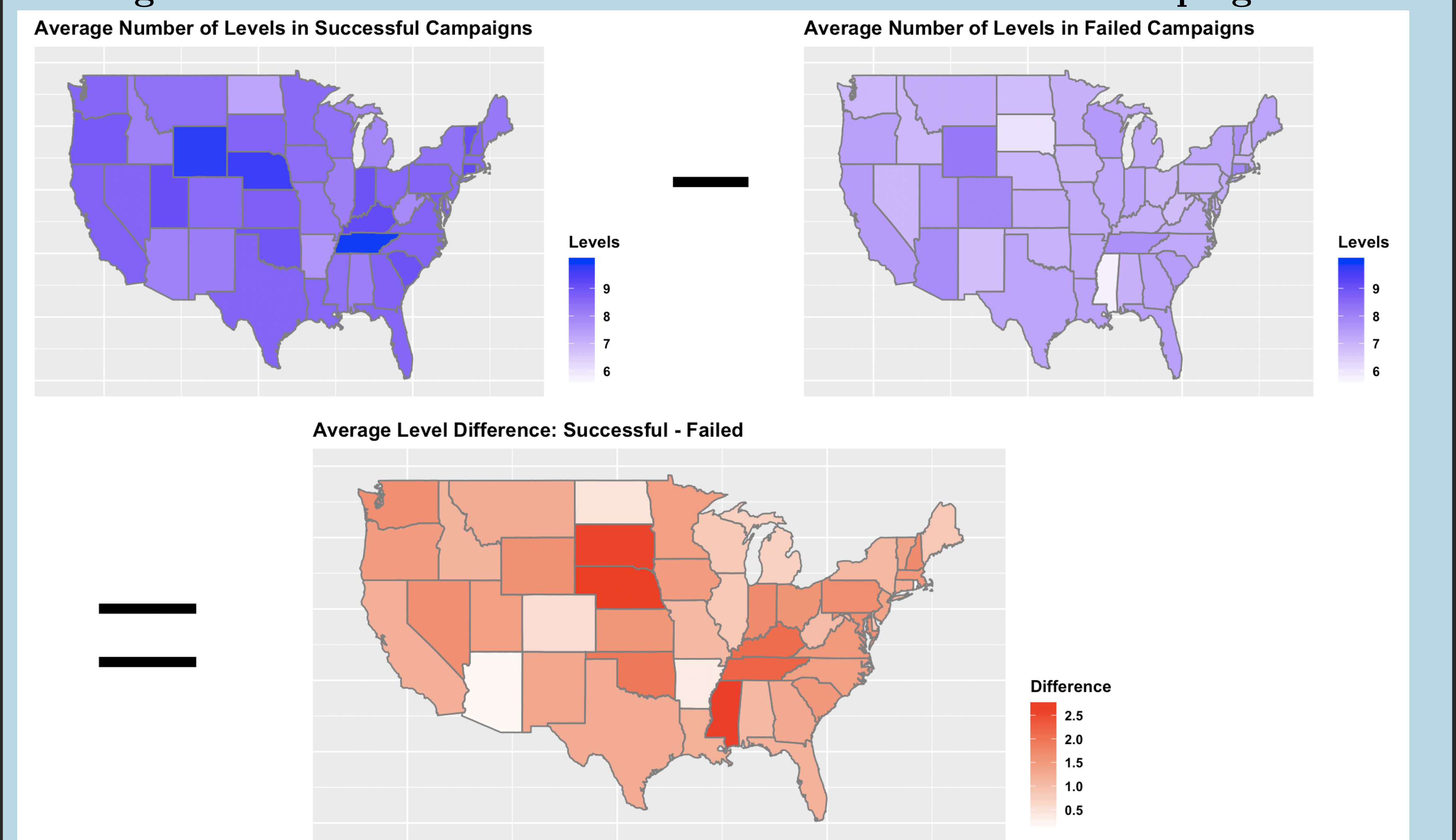
First, we can infer the overall distribution of campaigns among the different categories. It seems that film & video (19%), music (16%), and technology (15%) are the most common categories.

We are also able to visualize the ratio of success to failure for each category. The technology bubbles are roughly the same size, indicating a 50% success rate. This is lower than the average success rate of about 60%. Higher performing categories like comics, on the other hand, have significantly larger success bubbles than failure bubbles.



Analyzing Success: Reward Levels and Comments

Below we visualize the impact of reward levels with color indicating the average number of reward levels. By subtracting the failed campaigns from the successful campaigns, we arrive at our difference map: **The average difference in reward levels between successful campaigns and failed campaigns.** While the difference in levels between successful and failed campaigns does not follow any notable geographic pattern, we do see that the entire map shows positive values. This means that, **across every state, the successful campaigns tend to offer more levels than the failed campaigns.** So overall, **offering more reward levels tends to be associated with campaign success.**



We also found a relationship between success and the number of comments on a campaign's webpage. **On average, successful campaigns have 26.94 comments versus only 0.94 for failed campaigns.** The violin plot to the right shows a more detailed distribution of comments. Overall, it seems that **comments are a good measure of community interest and, as a result, likelihood of campaign success.**

