Kickstarter and How to Predict a Successful Campaign

# Kevin AngottiWinter 2020

<https://github.com/kangotti>

<https://kangotti.github.io/Data_Science_Portfolio/>

# Which Domain?

Data for this project will come from the Kaggel.com site. Kaggle is a good source for data project ideas with submissions from people wanting to know the outcome of a potential project or current projects that have been attempted. The submission by Mickaël Mouillé seeks to find out what determines a successful campaign. Many of the data sets here are all organized and relatively easy to navigate for understanding.

1. Mouille, M. (2018). Kickstarter Projects. Datasets. Kaggle. Retrieved from <https://www.kaggle.com/kemical/kickstarter-projects>

Information on what and how the data is to be examined. The author also provides a good amount of information on how the data was compiled and what he would like out of any projects.

2. <https://www.kickstarter.com/>

This site s the main company page and provide adequate background on how a campaign gets started, and provides the mission goals and company outline.

3. Peters, D. (2013). 10 Tips I Wish I Knew Before I Launched My Kickstarter Campaign. Entrepreneur. Retrieved from <https://www.entrepreneur.com/article/229782>

This article gives a layout of other Kickstarter campaigns that people would have liked to know before they started out. This can help provide insights to results we ay see in the data.

4. Stimmel, G. (2019). 11 Proven Tips For Launching A Successful Kickstarter Project. Product hype. Retrieved from https://producthype.co/kickstarter-tips-2019/

This is another site that can be used to help make sense of the results and may even provide useful information in determining a target for the models.

5. Cood Backer. (2019). HOW TO TELL IF A KICKSTARTER CAMPAIGN IS LEGITIMATE. Cool Backer. Retrieved from <https://coolbacker.com/how-to-tell-if-a-kickstarter-campaign-is-legitimate/>

Looks into what could be a scam or legit Kickstarter campaign.

6. Albright, D. (2016). 3 Things to Consider Before Backing a Kickstarter Project. Make Use Of. Retrieved from <https://www.makeuseof.com/tag/3-things-consider-backing-kickstarter-project/>

Looks at what might help a Kickstarter become a project people will back.

7. Prondle, D. (2019). Don't get burned! How to back crowdfunding projects the smart way. Digital Trends. Retrieved from <https://www.digitaltrends.com/cool-tech/crowdfunding-tips-avoid-scams-kickstarter-indiegogo/>

Provides another aspect that could be used as helpful insights that might contribute to success in a campaign. This might be helpful in how people might pick a Kickstarter.

8. Benvides, N. (2017). What Makes a Successful Kickstarter Campaign? Towards Data Science. Medium. Retrieved from <https://towardsdatascience.com/what-makes-a-successful-kickstarter-campaign-ad36fb3eaf69>

Further information on what it takes for a Kickstarter to become successful. These types of articles can help provide reasons a campaign will be successful or fail.

9. James, T. (2021). 7 Keys for Successful Kickstarter Campaigns. Cover Kit. Retrieved from <https://convertkit.com/successful-kickstarter-campaigns>

More help t determine success.

10. Crockett, Z. (2019). What are your chances of successfully raising money on Kickstarter? The Hustle. Retrieved from <https://thehustle.co/crowdfunding-success-rate>

More success rate information helps understand the data, and the success rate is the main focus.

# Which Data?

I will examine the Kickstarter Projects data, which contains more than 300,000 Kickstarter projects.

<https://www.kaggle.com/kemical/kickstarter-projects>

This dataset provides 32 columns over two datasets, each containing 16 columns. Over 300 thousand rows in the datasets will provide a fair amount of data for model prediction. Within the dataset, key variables I will try and make correlations to are category, deadline, pledge, goal, and main category.

# Research Questions? Benefits? Why analyze these data?

* Can the funding goal of a project be used to predict if a project is successful?
* What factors contribute the most to a Kickstarter project?
* Is there a timeline that offers the best outcome for successful campaigns?

This project can provide answers to the submitted users' requests to understand or predict what contributes to a successful Kickstarter campaign. Other benefits point to helping future Kickstarter projects become more successful by understanding why campaigns succeed or fail.

I choose this data to analyze, mainly to help potential business enthusiasts succeed in their business ideas and goals. In the current COVID world, there is a need for new products and businesses to have a fighting chance. Many failed and canceled campaigns are shown in the data, so helping to determine the factors behind this could help Kickstarter and anyone who plans to start a campaign.

# What Method?

I will start this project with linear regression analysis for correlations of variables, then dive into either cluster analysis for deep learning if possible. By approaching the project to predict if a campaign is successful or not, there may be a need to model the project differently than my usual approaches. The main goal is to assist future Kickstarter campaigns so, the need for good accuracy in any of the models will be the primary measuring factors.

# Potential Issues?

So potential issue I can see is that there are only 16 variables per dataset. There seem to be more than enough rows of data to analyze, but the limited variables could become an issue when modeling occurs. In the description of the potential Kickstarter data project, it is described by the author, anyone creating the project should not use the pledged variable in any analysis as it is heavily tied to the goal variable and that it could provide a bias in the results. I will need to take a closer look at both of these variables and see how they might provide skewed results if used together in a model.

Modeling, I believe, provides the largest factor for if the project will become successful. If I cannot get any results from the models, I end up using, I am not sure there will be significant time to refocus the project on a different topic. With only four weeks from start to finish, I am worried about the success rate. Any potential delay could be costly.

# Concluding Remarks

I believe this project could provide a good amount of information to anyone looking to create a Kickstarter campaign. Kickstarter has become an excellent way for developers and businesses to get started, and a good amount has become very successful. Taking on a project like this will allow future data scientists and me with a good example of how to find insights and predictions on a dataset with limited variables. The potential results can provide Kickstarter with the ability to help aid their customers with a how-to guide or even a best practice type of guide. Anytime we can create a prediction on a dataset, the results can allow users to determine if the outcome of a campaign is worth the time and attention needed to become successful.