

Homework 1 Questions

What are three conclusions we can make about Kickstarter campaigns given the provided data?

One conclusion we can make about Kickstarter campaigns provided the data is that the majority, nearly 75%, of the campaigns are in the United States. Therefore, any other conclusions drawn from the data may be associated with American spending patterns and do not necessarily predict an international reaction to a given campaign.

Another conclusion that can be made from the data is that campaigns under the theater category have the most successful campaigns as well as the most campaigns in general. Campaigns under the category music, however, have the highest percentage of successful campaigns with a success rate of 77% opposed to a success rate of 60% for theater campaigns.

When looking at sub-category data, we find that the majority of campaigns are for plays. They have a total of 1066 compared to the next highest number of campaigns which is rock music with 260. The rock music sub-category, however, has a success rate of 100% while the play sub-category has a success rate of 65%.

What are some of the limitations of this dataset?

The most noticeable limitation of this dataset which is addressed in the bonus is the amount of each goal that was requested. It is reasonable to predict that projects with lower campaign goals will have a higher success rate since they will need fewer donors or donations have less value. The charts we have looked at do not include this variable so it may be that sub-categories have greater success rates due to lower campaign goals and not due to the sub-category itself.

One beneficial piece of data that was not included in this set was the percentage of viewers that donated to the cause. If we were attempting to use this data to create a successful campaign, it would be helpful to know which campaign pages were able to attract donations from the viewers of the page. It would also be helpful to know the number of viewers of each page so that we could model our own page after the most trafficked campaigns.

What are some other possible tables/graphs that we could create?

One graph that I would like to create would be a stacked chart of successful campaigns filtered by goal value. This way we could see how the category or sub-category related to the goal value. This could help address the problem discussed in the limitation section of the true cause of success for each campaign. I would also like to see a graph with average donation versus success to see if campaigns were more successful with a large amount of small donations or a small amount of large donations. To design a successful campaign or predict a campaign's future success it would be important to view both how success interacts with categories and sub-categories as well as the goal values.