Module 1 Challenge

# Questions:

1. Given the provided data, what are the three conclusions that we can draw about crowdfunding campaigns?
2. What are some limitations of this dataset?
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

# Answers:

## Question 1:

The provided data shows us that crowdfunding campaigns in the “Photography” category which are conducted outside of the USA have a low rate of success, whereas the same campaigns conducted in the USA have a high rate of success. Of the 34 photography crowdfunding campaigns conducted within the USA during the period in question, 24 of them were successful: a success rate of 71%. But of the 8 total photography crowdfunding campaigns conducted outside of the USA, only 2 were successful: a success rate of 25%. This finding implies that photography crowdfunding campaigns have a materially higher chance of succeeding if they are undertaken in the USA.

Secondly, it is notable from the dataset that the success rates of crowdfunding campaigns in the “Theatre” category, when broken down by month, appear to be seasonal in nature. Campaigns created in some months of the year experience persistently lower rates of success than those created in other months. Of the 344 total theatre-related crowdfunding campaigns conducted, the dataset shows that 187 of these were successful: an overall success rate of 54%. But if we analyse the success rate of campaigns created in the month of May we find a success rate of only 37%, and for campaigns created in the month of August the success rate is only 38%. This compares to success rates of more than 60% for theatre-related crowdfunding campaigns created in February, April, June, or September. The month in which a theatre-related crowdfunding campaign is created appears to have a material impact on its chance of success.

Another takeaway from the data is that crowdfunding campaigns in both the “Audio” and “World Music” sub-categories had a 100% success rate. This is an indication that campaigns in these sub-categories are well-received by funders and that the demand to donate to campaigns in these categories is high relative to the number of campaigns created. Whilst the total number of campaigns for these sub-categories are low (with 7 campaigns between them), the data indicates that future campaigns undertaken in these categories will have a high likelihood of success.

## Question 2:

The background information mentions that the dataset is a sample (or are sample projects). There is no information given as to whether the data is an accurate representation of all crowdfunding activity (i.e. whether the same is representative). If the sample data is not representative of the population, analysis of the data may result in flawed inferences.

Secondly, for a number of sub-categories in the dataset the number of sample campaigns is very low, particularly when filtered by country. For example, when looking at campaigns conducted in Australia, the sub-categories of Animation, Indie Rock, Jazz, Mobile Games, Non-Fiction, Radio & Podcasts, and Wearables only have one campaign each over the entire sample period. Asserting that the rate of campaign success for one of these sub-categories maybe 100%, or 0%, is not particularly meaningful and is unlikely to be an accurate representation of chances of campaign success for future campaigns undertaken in that country and sub-category combination.

## Question 3:

A table showing the success rates of “Staff Pick” and “Spotlight” campaigns relative to those not assigned with these monikers could be valuable. If the data was able to show that those assigned with the above monikers had higher rates of success, then those running campaigns may be able to improve their chances of a successful campaign by purchasing “Spotlight” status, or consulting with Kickstarter and Indiegogo staff to increase the campaign’s appeal.

A graph that shows the proportion of successful campaigns relative to total campaigns by year (rather than aggregating multiple years by month as the “Outcome by Date” sheet does). This would enable analysis of the overall trends in the success rates of crowdfunding campaigns over time. For example, are successful campaign rates steadily increasing over time as crowdfunding becomes more mainstream? Or are the campaign success or failure rates increasing year after year in certain categories? This would allow those seeking crowdfunding to target categories or sub-categories that are increasing in popularity with funders.

A graph analysing the relationship between duration of a campaign and whether it was successful. When aggregated, this would show whether longer duration campaigns have a higher probability of being successful. This may not only prevent some campaign operators from cancelling their campaign too early, but also provide data to better inform deadline dates for those planning campaigns.

By including a table that filters for what we might call “very successful” campaigns, campaigns that are more than 200% funded, we could ascertain the categories and sub-categories that have very high popularity amongst funders. Doing this would highlight pockets of high fundraising opportunity for people or organisations operating in those categories and sub-categories.