***Given the data provided, whatare some conclusions we can draw about crowdfunding campaigns?***

* Success/failure/cancel rates for campaigns is somewhat consistent across category, even though the number of campaigns across categories can vary greatly.
  + Success: 44-67%
  + Failure: 26-48%
  + Canceled: 2-10%
  + These all disregard the category of “journalism,” likely due to the low sample number for that category (n=4).
  + The only categories to have an above average success rate (57%) are technology (67%), photography (62%), and publishing (60%).
* The months with the greatest number of campaigns are January and July.
  + However, the highest campaign success rates are found in early summer (June and July). This peak in performance, however, suffers a regression right after this window, leading to the month of August having the lowest success rates, and highest cancel/failure rates.
* The 55% of campaigns fall in the range of $1000-$9999, with success rates being highest for projects seeking between $1000-$4999 (23% of all campaigns, 83% success rate), and significantly lower for projects seeking between $5000-$9999 (32% of all campaigns, 52% success rate). This is not surprising—campaigns that need less money to be successful tend to be more successful.
  + However, it is perhaps a little surprising that the set of campaigns seeking the smallest amount of money, less than $1000 (5% of all campaigns) only has a success rate of 59%.
  + The only other type of campaign charactarazing a significant portion of all observed campaigns is the highest of goal campaigns: those seeking $50,000 or more (31% of all campaigns). Unsurprisingly, these were the least successful campaigns, and they had the lowest success rate (37%).
* An over-generalization: in order to maximize the chance that your campaign gets funded, you should launch it in late June or early July, and you should seek between $5000-$9999. The category of your campaign doesn’t have a super significant impact on the success rate, however if you’re seeking funding for technology, that might help your odds of success.

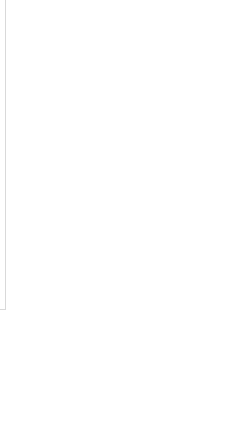
***What are some limitations of this dataset?***

* This geographic scope of this dataset is limited to a pretty narrow set of only 7 countries with large economic bases (US, GB, CA, etc.) It’s likely that the conditions for campaign success vary greatly as national economic context changes. For example, a country with a higher GDP could be expected to have more cash to spend on charity like crowdfunding. Furthermore, in countries with more equitable wealth distributions, the set of people with capacity to donate to a crowdfund would grow, and we’d expect the variance among donation sizes to increase.
* This dataset largely considers campaigns in the entertainment sector. However, some of the biggest campaign categories that are typically crowdfunded are not represented here (education, healthcare, for example).
* This dataset does not reveal any information about the nature or strategy of the campaign itself. It could be improved by including share-rate data. What portion of people sharing the campaign also donated to it?
* This dataset is restricted to 2010-2020, both endpoints narrowly avoiding major global finacial events (2008 Fincancial Crisis and 2020-21 COVID-19 Outbreak). Having crowdfunding peformance data for these time periods would improve our ability to contextualize our projections with respect to more complicated, chaotic, and reactive financial contexts.

***What are some other possible tables and/or graphs that we could create, and what additional value would they provide?***

* Tracking campaign success and campaign rates year-over-year could tell us if the market for crowdfunding trending up or down in the near future.
* We could look at campaign duration, which would reveal that the duration of the campaign has negligble association on the success of the campaign. Successful campaigns average 15.1 days, and failed campaigns average 15.7 days, and all campaigns average 15.4 days. This could help campaign leaders project success or failure early and throughout the campaign.
* We could look at average donate size with respect to category. This could be used to help campaign leaders develop strategy: how many backers are they looking for? What’s the size of donation they’re expecting? That would inform how many people you ask to donate how much money.
* How impactful is the “spotlight” or “staff pick” on a campaign on the number of backers, and the size of their contributions? Does it depend on the size or category of the campaign? If this is an extra campaign cost, what is it’s expected value add?

***What better summarizes our data: mean or median?***

* The count of backers among both successful and failed campaigns is skewed right, with a significant tail extending to the right. This implies that there is an impactful number of campaigns with an “abnormally large” number of backers. Here we can define “abnormally large” as having a number of backers greater than two standard deviations above the mean number of backers.
  + Among **successful** campaigns (std dev = 1267.4, mean = 851.1). There are 30 such campaigns, of the 565 total.
  + Among **failed** campaigns (std dev = 961.3, mean = 585.6). There are 19 such campaigns, of the 364 total.
  + It is worth noting, that these abnormally highly-backed campaigns make up approximately 5% of both successful ***and*** failed campaigns, and so does not seem to impact the success rate of campaigns.
* The impact of having this many campaigns with an abnormally large number of backers is that is biases the ***mean*** up, overrepresenting how many backers there are in the average campaign.
* This being the case, the ***median*** as a measure of center is more resistant to these abnormal data points, and so is a more accurate indicator of the average number of backers a campaign would have. For both successful and failed campaigns, the median is significantly lower.
  + Successful: mean = 851.1, median = 201
  + Failed: mean = 585.6, median = 114.5
* Unsurprisingly perhaps, the median number of backers in a successful campaign is 76% higher than among unsuccessful campaigns. But note the very large difference between the mean and median, among both successful and failed campaigns, with the median for both being 4-5x as high as the mean.

***What dataset has more variability in the number of backers: successful or failed campaigns?***

* Among the successful campaigns analyzed, the number of backers has a standard deviation of 1267.4, variance = 1 606 216.6.
* Among the failed campaigns analyzed, the number of backers has a standard deviation of 961.3, variance = 682 103.1.
* Based on comparison of these measures of variability, variability among successfully backed campaigns is greater, even though both successful and failed campaigns have proportional levels of campaigns with an abnormally high number of backers.
  + Failed campaigns have fewer backers donating (median = 114.5, compared to 201 among successful campaigns). There is a floor to how many backers we can have (none), and so given that unsuccessful campaigns tend to have fewer backers, this leads to a greater number of campaigns falling in the range of “no backers” to “not that many” backers.
  + If a campaign has a high variance, that would imply that there is a greater chance of having a number of backers far from the median number of backers. The minimum possible number of backers (none) is much closer to the median number of backers (114.5 for failed, 201 for successful campaigns), however the maximum number of backers for this set is significantly further from the median (6080 for failed, 7295 for successful campaigns). So if we have more campaigns with a greater number of backers, you’re more likely to collect more contributions, resulting in a greater likelihood of being successful.