**Module 1 Questions**

1. Three conclusions that can be drawn about crowdfunding campaigns based on this data:

* On the whole, successful campaigns recruited more backers (a mean of 851 and a median of 201) when compared to failed campaigns (a mean of 585 and a median of 115).
* There is a significant increase in chances of launching a successful crowdfunding campaign in mid-Summer (58% in July vs 46% in May)
* The chances of success of a crowdfunding campaign are in general higher when the fundraising goal is somewhere between $15,000 and $34,999 backers.
  + Aiming to raise more than $50,000 yielded a 53% failure rate, which makes sense if we think about needing to raise a large amount of money being more difficult
  + Aiming to raise less than $1,000 yielded a 39% failure rate. It might be interesting to dig into why.

1. Some limitations of this dataset include:

* The range of countries that we are seeing data from. This might mean we are comparing apples and oranges depending on cultural differences in crowdfunding, but we can filter that out. A larger issue is that we are comparing a range of currencies, if we don’t calculate all donations into USD.

1. Other potential tables or graphs we could create could be:

* Success rate of crowdfunding projects by year, which would potentially reveal the effect of economic conditions on crowdfunding, especially since this data ranges from 2010 to 2019 and the economy broadly became stronger in that time range.
* Success rate of crowdfunding projects by length of fundraising window. By using the date launched and the deadline, you could see if the success rate varied by the amount of time given to raise funds.

1. I think the median better summarizes this data because there are some campaigns with an unusually high number of backers, which skews the mean towards a higher number.

* Looking only at the successful campaigns, there 387 successful campaigns with fewer than 500 backers and 176 successful campaigns with more than 500 backers. The median better represents this data by giving less weight to the few campaigns with extremely high numbers of backers.

1. There is more variability within successful campaigns. The variance for successful outcomes is 1,603,374 and for failed outcomes is 921,575. I was initially surprised by this as I felt there were probably more ways to fail than to succeed.

However, I can see that this makes some sense since there are such a wide range of project sizes that could be successful with wide ranges in how much money they needed to raise (and therefore how many backers they could have required). On the other hand, if we make an assumption that the most likely reason for not being able to meet fundraising goals is not enough backers, it makes sense that these would cluster closer together around a smaller number.