* Three conclusions that we can draw about Kickstarter:
  + The state of the sub-categories of previous campaigns can provide valuable insight to predict whether a future campaign will succeed.
    - For example, campaigns that are categorized as Music have the highest success rate, however when broken down into sub-categories it becomes clear that Jazz has a 100% failure rate. On the other hand, the Food category is the least successful, but 100% of sub-category Small Batch campaigns have succeeded.
  + The timing of the start of the campaign can affect the likelihood of success.
    - Campaigns that begin in May have the highest success rate, while campaigns that start in December are more likely to fail than succeed.
  + The Spotlight designation is a better predictor of success than Staff Pick.
    - All campaigns that had the Spotlight succeeded, even those that were not labeled as Staff Picks. Campaigns that were Staff Picks but did not have the Spotlight label all failed.
* Some of the limitations of this data set are:
  + Having data on the donation amount per backer on each campaign would be more useful than looking at the average donation per campaign. The average donation on successful campaigns ranges from $3.25 to $2500.97. Being able to calculate the median donation amount would provide better data because any outliers can significantly skew the average donation – one or two extremely large donations could be the reason that certain campaigns succeeded.
  + There could be additional factors that contribute to the success of campaigns – some campaigns create donation tiers where the backer receives a certain gift as a thank you once the campaign has succeeded.
  + No information as to why campaigns were canceled. It would be interesting to see if there are any common factors that resulted in the campaign being canceled. Can cancelations be predicted and avoided?
* What are some other possible tables and/or graphs we could create?
  + Line graph that shows state of sub-categories over time
    - This would provide data on whether sub-categories are more likely to succeed during specific months of the year. For example, a campaign for Small Batch ice cream is likely to be more successful in the summer than in the winter.
  + Pivot table that shows the state of sub-categories per goal amount
    - This would provide data on whether there is a particular goal amount where a specific sub-category is no longer successful.
  + Box chart that shows the state of sub-categories for both Staff Picks and non-Staff Picks
    - This would show whether certain campaign types are more likely to be chosen as a Staff Pick. It would also show the amount of successful vs unsuccessful campaigns for non-Staff Picks
* Use your data to determine whether the mean or median summarizes the data more meaningfully:
  + Median summarizes the data more meaningfully. There is a large variance in the number of backers in both the successful and unsuccessful campaigns which indicates there are outliers that are being calculated into the mean. Using the median value provides a more accurate representation of how many backers a successful campaign has because it does not take into account the few campaigns that had an extremely high number of backers (maximum number of backers was 26,457).
* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
  + There is more variability with successful campaigns. I do think this makes sense – a campaign doesn’t necessarily need a lot of backers to be successful. It could have a small number backers who each donate a large amount, many backers who each donate a small amount, or something in between. Without looking at the data, my assumption would be that unsuccessful campaigns typically have a relatively low number of backers. The data shows that the median number of backers is only 4 with a standard deviation of 61 – compared to a median of 62 and standard deviation of 844 of successful campaigns.