1. Given the provided data, what are 3 conclusions we can draw about crowdfunding campaigns?  
   1. Theater is by far the most popular category type of crowd funding project on this platform as it has the greatest grand total of projects (344). The second most popular category type is Film & Video, which only has a grand total of 178 projects.
   2. Journalism is the least popular category type as it has the lowest grand total of projects (4). However, it should be noted that the Journalism category has a 100% success rate as all 4 projects successfully reached their funding target.
   3. Projects started between late May and early July have the highest chance of success compared to the rest of the months in the year as there is a spike in the success of campaigns during that period according to the Line Chart that is counting the outcomes (success, failed, canceled) of each project based on the month.
2. What are some limitations of the dataset?  
   1. Success is measured by weather or not the target goal for raising the capital was reached. It does not indicate whether the project itself was successful. Example, if the goal was to reach $20k in funding to make a theater play and they received $25k in funding that would be considered a success. However, if they make the play and no one shows up to it, 0 tickets are sold, it is a commercial failure, but is still considered a success by the crowdfunding measure.
3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?  
   1. Row: Parent Category  
      Values: Sum of Average Donation  
        
      This would give us an idea of which categories have the highest spending per donor. Which is Theater @ $23,728. These are the projects that people are interested in donating money towards. It would be prudent to advertise the platform as funding the arts as this will lead to more dollars spent on the platform.
   2. Row: Country  
      Values: Sum of Pledged, Sum of Backers Count  
        
      This will tell us which countries are providing the most people as well as dollars to the platform. By far and away this is the United States, which makes it an outlier. Filtering out the United States shows that Canada, followed by Italy contribute the 2nd and 3rd most people as well as total dollars to the platform. However, if we further analyze this data, it we can show how much money everyone pledged on per person in each country. In this way we see that Great Britain is actually the true second-best contributor to the platform on a per capita basis (65 pounds per person). Which could mean that any advertising spent in the UK may see a greater return of money to the platform.
   3. Row: Spotlight  
      Values: Count of Outcome

Columns: Outcome

Non-Spotlight 418 successful / 732 total = 57%

Spotlight 147 successful / 268 total = 54%  
  
Unfortunately this shows that projects that are spotlighted have a lower chance of being successful. Thus, Non-Spotlight projects have a better chance of being successful, which means upselling a “spotlight” service wouldn’t help customers.

* 1. Row: Staff\_Pick  
     Values: Count of Outcome

Columns: Outcome

This would show us the number of success’ / failures of the “staff pick”. This will allow us to do more analysis. Of the 49 “Staff Picks” 28 were successful (57.14%). Of the 951 “non staff picks” 537 were successful (56.46%). Therefore, it could be concluded that if your project is a “staff pick” it might be slightly more successful in achieving its goal, although it is within the margin of error.

Bonus – “Sheet: Bonus 2”

1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.
   1. I believe that the mean of the number of backers provides a better insight into the success / failure of a project. The average number of backers of a successful project is 851, the average number of backers of a failed project is 586. Therefore we can conclude that the more backers that one has the more chance of success the project has.

Now, if we were to look at the median number of backers on successful projects, we see that the number is 201. Which just means that the number of backers directly in the middle of the entire “successful” dataset is 201. Which does not provide any real insight to draw from.

1. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
   1. It appears that successful campaigns do have more variability. This makes sense as the unsuccessful campaigns dataset will have more individual projects with 0 backers in total. Additionally, more of the failed projects will have backers closer to 0 as well, projects with 2-3 backers only. Which would mean that the variance of the data set would be lower than successful campaigns.