1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

* When we look at the amount of Kickstarter campaigns by parent category, we can appreciate that theater campaigns are the most ubiquitous, with 1393 projects in total. Music campaigns come in a distant second place, with 700, and technology campaigns are in third place, with 600. The least explored categories are journalism (with 24 campaigns), followed by food with 194 campaigns and a tie between photography and games campaigns, both with 220 campaigns. Looking at the number of successful campaigns, we can see that theater campaigns were the most successful at 839, followed by music campaigns at 540 and film & video campaigns at 300 successful campaigns. The least successful campaign categories were journalism, with no successful campaigns, then food campaigns with only 34 successes, and a tie between games and publishing campaigns with 80 successes.

From this we can conclude that if someone were to invest in a Kickstarter campaign, the safest bet would be to invest in a music campaign, since they are the most prone to succeed.

* If we break down the projects by subcategory, we can notice most Kickstart campaigns are labeled under theater/plays, with a total of 1047 campaigns. Music/rock campaigns follow with 260, and technology/wearables campaigns come in third place with 200. The campaigns that are least sought after are a tie of several subcategories (music/metal, music/world music, publishing/art books, photography/nature, technology/gadgets, photography/people, food/restaurants, photography/places, technology/makerspaces, and publishing/radio & podcasts) with 20 campaigns, then journalism/audio campaigns with 24 campaigns, and food/small batch campaigns with 34. In terms of successful campaigns, theater/plays are also in first place, with 694 successful campaigns, followed by music/rock campaigns at 260 successful campaigns and film & video/documentary campaigns at 180 successful campaigns. The least successful campaigns are   
    
  We can conclude that the safest bet on a Kickstarter campaign would be investing on a “music/rock” campaign, since they are the most successful.
* If we look at the outcome of campaigns per month of creation (expressed in a line graph), we can see that the months where campaigns are created the most are July with 387 campaigns, followed by May with 386 and June with 385. The months in which campaigns are least created are December (252 campaigns created), then September (298 campaigns created) and March (316 campaigns created). As far as successes go, the most successful month to create a Kickstart campaign is May (234 successful campaigns), followed by June (211 successful campaigns) and February (202 campaigns). The least successful months are December (111 successful campaigns), September (147 successful campaigns) and August (166 successful campaigns).   
    
  Looking at the monthly data, if someone were to create a Kickstarter campaign, it would be best for that person to create a campaign in either May or June, since those are the months that have the greatest number of successful campaigns.

1. What are some limitations to the data?

* Not every country was represented in the dataset. A greater number of countries would make a more robust dataset.
* The dataset doesn’t provide specific reasons why a project was cancelled.
* There is no geographical data (beyond country) in the dataset.

1. What are some other possible tables and/or graphs that we could create?

* We could create tables that assess the impact of the country of origin on a project’s outcome.
* We could also create a table that measures the impact of “Editor’s pick” or “Spotlight” endorsements on campaign’s outcome.
* We could take a look at the effect of higher or lower fundraising goals over campaign’s outcomes. (This was covered under the Bonus section of the homework assignment)
* Another possible study we could do is to find, among successful campaigns, which categories or subcategories are able to exceed their campaign goal.

1. (Statistical Analysis Bonus) Use your data to determine whether the mean or the median summarizes the data more meaningfully.

* For the successful campaigns, there is a wide range of number of backers per campaign, with the minimum number of backers being 1 and the maximum being 26457. For this subset, the mean would be a better summary of the data than the median. For the unsuccessful campaigns’ subset, the median makes more sense because there is an important amount of campaign with zero backers. This is reflected by the median being closer to zero (4) than the mean (20).

1. (Statistical Analysis Bonus) 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 the number of backers of successful campaigns than the number of backers of unsuccessful campaigns because the range of backers per campaign is much wider in the successful campaigns (minimum: 1, maximum: 26457) than in unsuccessful campaigns (minimum: 0, maximum: 975). As a consequence of this, the variance (and standard deviation) of the “successful” dataset is much higher than the “unsuccessful” dataset.