Alison McCondichie

Challenge 1 – Excel

1. What is the data set? What is the purpose of the project?
   a. The data set provides information from 1,000 crowdfunding campaigns. The purpose of this project is to analyze and data and present visuals and create additional statistical information to predict how one could create a successful crowdfunding campaign. The information can be organized and analyzed to point out differences between successful and failed campaigns that we can use to highlight helpful insights.
2. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
   a. The average donation for a campaign is not as influential to success as one might expect. How *much* each backer donates is not as contributing as how *many* backers a campaign has. Campaign ID 33 was successful with an average donation of $33, while Campaign ID 11 was a failure with an almost double average donation of $112.22.
   b. Out of all the sub-groups, Plays have the most successes AND the most failures. This could simply be because Plays have the most total campaigns overall. It also means that the data sets specifically within Plays campaigns could give us a lot of useful information because they have a "constant", they all share a common sub-group. This would also mean that the sub-group of a campaign is not a major contributing factor to success.
   c. The data tells us that one of the clearest indicators for whether or not a campaign will be successful is the initial funding goal. Out of 1,000 campaigns, we found that 3 specific goal ranges had a 100% success rate. Is this only because those goal ranges between $15,000-$19,999 and $30,000-$34,999 are easier to obtain? This may be connected to a limitation of the data set. It is an interesting conclusion to notice.

3. What are some limitations of this dataset?
   a. What constitutes a "failed" campaign vs a "successful"? If a campaign were to be $1 short of goal, would it still be considered a failure? Let's say for Campaign A has a goal of $100 and it raises $120, that would technically be a success. If Campaign B has a goal of $10,000 and it raises $9,750 then it would register as a failure, even though it was more *successful* than Campaign A in terms of how much money was actually raised. This spreadsheet constitutes a failure if the campaign did not reach its funding goal. There could be an argument that the word "success" subjectively could mean how much money a campaign raised, even if they didn't quite reach their goal. Raising $1,000 more than another campaign could be a subjective success.

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

a. A Line Chart that compared the length of time a campaign ran versus the outcome would be interesting to see. We have a graph that shows Launch date of a campaign versus the outcome, but I'd like to know if how long a campaign runs has a strong positive or negative impact on the outcome.

5. Statistical Analysis: Is there more variability in successful campaigns or failed campaigns? Does this make sense? Why or why not?
   a. There is a greater variance in successful campaigns. This is because while there are outliers in both data sets, successful campaigns have outliers that are bigger numbers. It makes sense considering the amount of backers contributes more to a campaign's chance of success than how much each backer donates. As mentioned above, how many backers you get is more important, so successful campaigns tend to have more backers so the higher numbers in the successful data set contribute to a higher variance.

	Outcome	Backers_Count		Outcome	Backers_Count				
1	Outcome	Backers_Count		Outcome	Backers_Count				
2	successful	158		failed	0				
3	successful	1425		failed	24		Successful Campaigns		
4	successful	174		failed	53		Mean:	851	
5	successful	227		failed	18		Median:	201	
6	successful	220		failed	44		Minimum:	16	
7	successful	98		failed	27		Maximum:	7295	
8	successful	100		failed	55		Variance:	1603374	
9	successful	1249		failed	200		Standard Deviation:	1266	
10	successful	1396		failed	452				
11	successful	890		failed	674				
12	successful	142		failed	558				
13	successful	2673		failed	15				
14	successful	163		failed	2307				
15	successful	2220		failed	88				
16	successful	1606		failed	48				
17	successful	129		failed	1		Failed Campaigns:		
18	successful	226		failed	1467		Mean:	586	
19	successful	5419		failed	75		Median:	115	
20	successful	165		failed	120		Minimum:	0	
21	successful	1965		failed	2253		Maximum:	6080	
22	successful	16		failed	5		Variance:	921575	
23	successful	107		failed	38		Standard Deviation:	960	
24	successful	134		failed	12				
25	successful	198		failed	1684				

Parent Category (All)
country (All)

Count of outcome	Column Labels				
Row Labels	canceled	failed	live	successful	Grand Total
animation	1	10	2	21	34
audio				4	4
documentary	4	21	1	34	60
drama	2	12	1	22	37
electric music		8		10	18
fiction	1	7		9	17
food trucks	4	20		22	46
indie rock	3	19		23	45
jazz	1	6		10	17
metal		3		4	7
mobile games		8	1	4	13
nonfiction	1	6	1	13	21
photography books	4	11	1	26	42
plays	23	132	2	187	344
radio & podcasts		4		4	8
rock	6	30		49	85
science fiction		9		5	14
shorts	1	5	1	9	16
television	3	3		11	17
translations		7		14	21

Campaign Outcome Per Sub-Category

Sheet tabs: Goal Analysis | Statistical Analysis | ParentCategory_Pivot1 | SubCategory_Pivot2 | DateCreated_Pivot3

	A	B	C	D	E	F	G	H
3	Greater than or equal to 50000	114	163	28	305	37%	53%	9%

Outcomes Based on Goal

Legend: Percentage Successful — Percentage Failed — Percentage Canceled

Goal in Dollars

Sheet tabs: Goal Analysis | Statistical Analysis | ParentCategory_Pivot1 | SubCategory_Pivot2 | DateCreated_Pivot3