

# Crowdfunding Analysis

**Hypothesis:** The categories and sub-categories with the most funding and backers will have the most successful outcomes.

## Category and Country Analysis

In analyzing the Category and Analysis I found due to a data bias towards the US it would be important to analyze the US category trends solely. Also due to small sample sizes for the other countries, excluding the US, I have grouped the countries together in my analysis.

Three conclusions I can draw from the data for grouped countries:

- The largest number of backers and money pledged is going into Theater, Film & Video, and Technology.
- The categories that give the most weight to the overall total successful outcomes are Theater, Film & Video and Music. Although Music has a 15.50% weight Technology has a 14.73% weight on the overall success of the various category projects.
- Technology has just under half the number of observations but it does have larger funding and higher success rate. Music has 20 successful outcomes will Technology has 19, again showing why in other countries Technology may be the better investment versus Music.
- Had we grouped US and Other countries all together it would have seemed that Theater, Film & Video, and Music would be the best investments in all countries.

country	Grouped Countries			
	# of Successful Outcomes	Successful % of Total Outcome	% of Total Success	Total Outcomes
successful	Total Count of outcome			
Row Labels	Count of outcome	Count of outcome2	Count of outcome3	
film & video	26	54.17%	20.16%	48
food	5	45.45%	3.88%	11
games	7	63.64%	5.43%	11
music	20	43.48%	15.50%	46
photography	2	25.00%	1.55%	8
publishing	12	66.67%	9.30%	18
technology	19	79.17%	14.73%	24
theater	38	53.52%	29.46%	71
<b>Grand Total</b>	<b>129</b>	<b>54.43%</b>	<b>100.00%</b>	<b>237</b>

country	(Multiple Items)	Grouped Countries			
Row Labels	Sum of backers_count	Sum of % Backers/Goal2	Sum of Percent Funded	Sum of Average Donation	Sum of pledged
film & video	39448	117.80%	7454	3228	2334221
food	4377	40.02%	2751	642	331559
games	7008	60.17%	4071	752	385441
music	22524	123.39%	7037	2787	1422221
photography	784	13.56%	969	529	47303
publishing	18562	61.74%	4525	1290	1329775
technology	29257	80.30%	5159	1795	1984204
theater	59535	243.19%	14260	4643	3503995
<b>Grand Total</b>	<b>181495</b>	<b>740.17%</b>	<b>46227</b>	<b>15666</b>	<b>11338719</b>

### Categorical conclusion for data in US:

- The funding and number of backers for crowdfunding in the US is heavily weighted towards Theater, Film & Video and Music. After these three categories there is a huge drop in number outcomes in other categories.
- Same categories have the largest weight in total successful outcomes in the US.
- Although categories like Journalism and Photography have the highest success rates, the sample size is much smaller in comparison to the previous mentioned three.

country	US			
	# of Successful Outcomes	Successful % of Total Outcome	% of Total Success	Total Outcomes
	successful			Total Count of outcome
Row Labels	Count of outcome	Count of outcome2	Count of outcome3	
film & video	76	58.46%	17.43%	130
food	17	48.57%	3.90%	35
games	14	37.84%	3.21%	37
journalism	4	100.00%	0.92%	4
music	79	61.24%	18.12%	129
photography	24	70.59%	5.50%	34
publishing	28	57.14%	6.42%	49
technology	45	62.50%	10.32%	72
theater	149	54.58%	34.17%	273
Grand Total	436	57.14%	100.00%	763

country	US	-Y			
	Sum of backers_count	Sum of % Backers/Goal2	Sum of Percent Funded	Sum of Average Donation	Sum of pledged
film & video	82427	454.95%	27424	8780	5175855
food	24469	121.24%	7083	2500	1403620
games	30654	99.37%	6176	2642	1630376
journalism	1194	19.96%	603	122	36176
music	106478	461.75%	28979	8474	6057876
photography	23260	113.21%	6765	2191	1176628
publishing	34057	123.70%	8482	3692	1820052
technology	38237	278.31%	16501	4395	1849521
theater	204734	871.50%	52202	19086	12259232
Grand Total	545510	2543.98%	154215	51881	31409336

### Sub-Category Analysis Grouped Countries

- For the other countries Technology had better success rate but had a smaller number of observations. When we dive into the sub-category the tech is just two categories and they both show high success percentages. So again, reinforcing the decision to choose Technology in the top 3 investments over Music.
- Furthermore Music, when broken down by sub-category, the individual categories only show one sub-category with over a 50% success rate. That category is Jazz and there are less than 10 observations. The Music category over all is less than 50% success rate for other countries.

A	F	G	H	country	Grouped Countries		
country	Grouped Countries			Parent Category			
Parent Category							
	successful			Row Labels	Count of Sub Category	Count of outcome	Total Count of Sub Category
	Count of Sub Category	Count of outcome		electric music		0.00%	2
	9	81.82%	11	indie rock	5	45.45%	11
	10	76.92%	13	jazz	4	66.67%	6
				metal	1	25.00%	4
				rock	10	43.48%	23
Grand Total	19	79.17%	24	Grand Total	20	43.48%	46

### **Sub-Category Analysis US**

- Within the sub-categories for the US of Theater, Music, and Film & Video there are still some specific sub-categories that may have a larger probability of success when looking at the percentage of total success, for example Rock in Music.

### **Seasonality Analysis:**

- The seasonality analysis showed that even in the more successful categories such as Theater there were some cyclical. In the US there was a drop in success rates in Theater during May and August. Whereas in the other countries Jan through March and October and November had success rates below 50%.
  - Also, in the US the Film & Video, another of its more successful categories, has the lowest success rate in Dec, this remains relatively consistent when you break it down by year.
- 

### **Limitations to the Data**

- The two biggest limitations to the data are the small sample sizes for the other countries excluding the US. Which lead us to the second limitation which is the US bias in the data, Category, Sub-category, and Seasonality. The heavy bias in the US would lead to a conclusion that all countries have more successful crowdfunding projects in Theater, Film & Video, and Music.
  - Another limitation is the lack of that consistent data over a single year, and multiple years. The duration for certain categories has a short duration.
  - In the seasonal analysis the highly cyclical categories highly affect the success for the month, again with US bias.
  - In the seasonality analysis there is a limitation in a year-to-year comparison, as there can be different scenarios that can occur affecting the success of the category in that month or year.
- 

### **Possible tables and graphs**

- In my analysis I added some addition views to the pivot tables that showed the percent success of the total outcomes, each categories & sub-categories success as a percent of total success, the total amount of outcomes overall, and number of backers per goal amount. This gives insight into how much weight the success in that category has on the overall success for each total's successful outcomes. It also shows where the most backers are and where the most funding is going.
  - I think a histogram would be helpful to show the skew and bias in the data.
  - I would like to see which categories and sub-categories had the most percentage funded over 100%. This would show what specific categories that may be getting the most funding. This goes along with my overall hypothesis that those with the most funding and backers are more likely to be successful.
-

## **Statistical Analysis**

- In evaluating the successful backer count and failed backer count the median better represents the data. Based on the Mean being much larger than the Median there is a skew in the data. After creating a histogram, it doesn't show a normal distribution where a mean would be a good summarizing tool; it shows a right skew making the median a better option to summarize the data.
- Based on the variance calculated for both the successful and failed backer count, the successful backer count shows a much higher variance. Thus, there is more variability in the successful backers, meaning much more backer counts above the mean and below the mean. Yes, this does make sense, because the successful outcomes have more backers and will have a wider range of number of backers. Total number of successful backer count is 480,898 versus the total number of failed backers count which is 213,164, that is over double the number of backers, leading to larger possible variance and distribution.