1. **Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?**
2. Campaign Success Rate and Categories:

About 56% of crowdfunding campaigns are successful.

Campaigns under the "Plays" sub-category have the highest success rate, while campaigns in the "Theater" parent category also exhibit high success rates.

Both these categories are most popular in the United States, suggesting a potential correlation between campaign category, geographical location, and success.

1. Influence of "Staff Picks":

Campaigns designated as "staff picks" have a success rate of approximately 57.14%.

In contrast, campaigns not designated as "staff picks" have a success rate of about 56.47%.

The data suggests that being a "staff pick" offers only a marginal advantage regarding campaign success.

1. Success rate by Parent Category:

* Journalism: 100% (This unusually high success rate might indicate a limited number of campaigns in this category or other factors)
* Technology: 66.67%
* Photography: 61.90%
* Publishing: 59.70%
* Film & Video: 57.30%
* Music: 56.57%
* Theater: 54.36%
* Food: 47.83%
* Games: 43.75%

Despite journalism’s high success rate (100%), it has a limited number of campaigns(4). This, combined with the low backer count and potentially shorter project time frames, suggests that this 100% success rate might not be wholly representative and should be interpreted carefully or require additional information on its successability.

2. **What are some limitations of this dataset?**

1. **​​Sample Size Limitation:** The dataset has limited entries in certain categories, like "Journalism," with only four campaigns. This can lead to potentially unrepresentative statistics, making conclusions for such categories less reliable. This means that the success rate of 100% for this category might not be representative due to the limited sample size.
2. **Absence of Detailed Information:** The dataset needs more information on rewards or specific project features. This omission restricts our understanding of potential factors influencing campaign success, as the nature and appeal of rewards can significantly impact backers' decisions.
3. **The subjectivity of "Staff Picks":** While we have a column indicating whether a campaign was a "staff pick,” the criteria or reasons for these selections are unknown. This subjectivity can introduce bias.
4. **Geographical Bias:** Most campaigns in the dataset originate from the United States, with fewer campaigns from other countries. This skewness can limit the generalizability of conclusions to a global context.

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

We can create possible tables based on the relationship between geographical location and outcome.

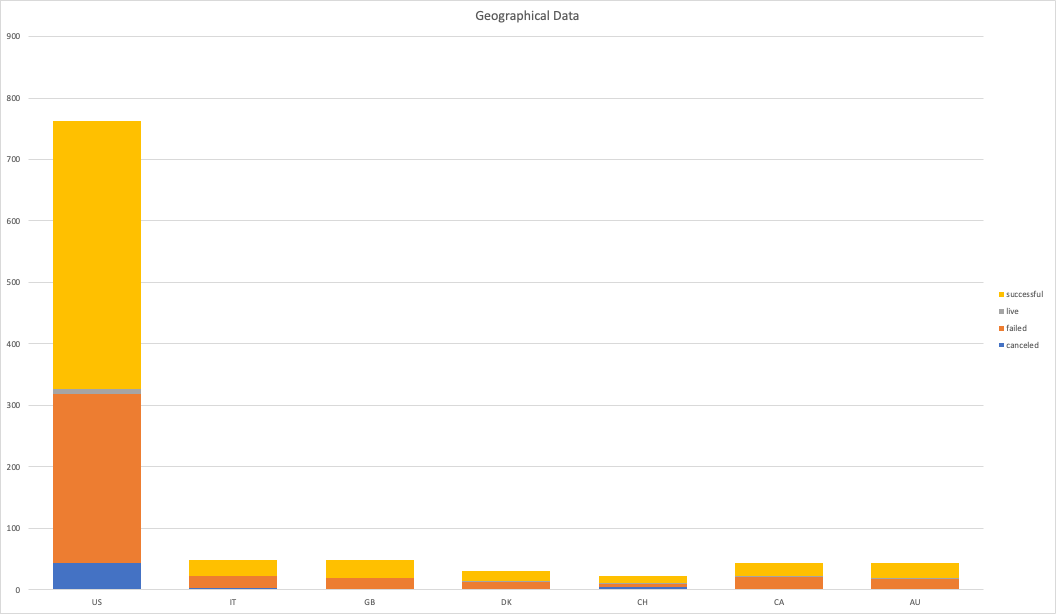
**Additional Value They Provided: Insights**

Shows Limitations-Geographical Bias: The dataset has a significant bias towards campaigns from the United States, with 763 campaigns. In comparison, other countries have fewer campaigns, which might limit our ability to draw globally representative conclusions about crowdfunding.

Here is the table:

| **Count of id** | **Column Labels** |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Row Labels** | **canceled** | **failed** | **live** | **successful** | **Grand Total** |
| US | 44 | 274 | 9 | 436 | 763 |
| IT | 3 | 19 |  | 26 | 48 |
| GB | 1 | 18 | 1 | 28 | 48 |
| DK | 1 | 12 | 1 | 17 | 31 |
| CH | 4 | 6 | 1 | 12 | 23 |
| CA | 2 | 19 | 1 | 22 | 44 |
| AU | 2 | 16 | 1 | 24 | 43 |
| **Grand Total** | **57** | **364** | **14** | **565** | **1000** |

**Graph**

****

**Diversity of the dataset.**

The distribution of campaigns across countries is as follows:

* US (United States): 763 campaigns
* GB (United Kingdom): 48 campaigns
* IT (Italy): 48 campaigns
* CA (Canada): 44 campaigns
* AU (Australia): 43 campaigns
* DK (Denmark): 31 campaigns
* CH (Switzerland): 23 campaigns