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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Parent Category** | **canceled** | **failed** | **live** | **successful** | **Grand Total** | **% Success (excludes live in denominator)** |
| film & video | 40 | 180 |  | 300 | 520 | 58% |
| Food | 20 | 140 | 6 | 34 | 200 | 18% |
| Games |  | 140 |  | 80 | 220 | 36% |
| journalism | 24 |  |  |  | 24 | 0% |
| Music | 20 | 120 | 20 | 540 | 700 | 79% |
| photography |  | 117 |  | 103 | 220 | 47% |
| publishing | 30 | 127 |  | 80 | 237 | 34% |
| technology | 178 | 213 |  | 209 | 600 | 35% |
| theater | 37 | 493 | 24 | 839 | 1393 | 61% |

Campaigns in the parent category within performing arts (film& video, music, theatre) the greatest chances of success in comparison with other categories

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Sub-category** | **canceled** | | **failed** | **live** | **successful** | **Grand Total** | **%successful** | **successful share/total successful** |
| Plays | |  | 353 | 19 | 694 | 1066 | 66% | 32% |
| Rock | |  |  |  | 260 | 260 | 100% | 12% |

Within the sub-categories, plays account for highest share (32%) of all of successful campaigns while rock has the second highest share (12%) of all successful campaigns. However, all campaigns in the rock sub-category have been successful so far as compared to 66% percent for plays. So, if one has a choice of choosing to host one of the two types of Kickstarter campaigns, from a weighted average perspective, they may want to consider a rock campaign over plays as it has a higher chance of success.

c.

Web sub-category represents 3% (1 out of 38) of all the sub-categories and there are 160 campaigns in this category, which accounts for 4% of all campaigns launched in the dataset, which is virtually proportional to its share. However, all 160 campaigns have either failed or have been cancelled. This does not seem to be a very lucrative category so I might want to steer clear of this if I were considering a Kickstarter campaign.

1. **What are some limitations of this dataset?**
2. The dataset does not include other factors that may also contribute to success/failure rates such as rate or types of advertising, social media exposure, press alerts, etc.
3. The dataset does not include any information about location origin. Perhaps campaigns originating in the US may be better supported over other locations where Kickstarter may be less popular
4. The dataset does not cover the value or type of rewards that are being offered and the market value of such reward for each level of contribution. Perhaps, the higher the reward, the higher the level of support.
5. **What are some other possible tables and/or graphs that we could create?**
6. Histogram bar chart to determine the distribution of the data to better assess measures of centrality and outliers
7. Box and whisker plot to group data into quartiles to better display variation and the outliers
8. Scatterplot to see if there could be a correlation between successful campaigns and date ranges