Matthew Copello

* **Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?**

1. There are more theater crowdfunding campaigns than there are any other.
2. Journalism, however, has the highest rates of successful crowdfunding (100%), compared to Theater’s 71%.
3. Crowdfunded Food & Video game campaigns have a less than half chance of succeeding (48% & 43%, respectively.)

|  |  |
| --- | --- |
|  | **rates of success** |
| theater | 0.705882353 |
| technology | 0.666666667 |
| publishing | 0.597014925 |
| photography | 0.619047619 |
| music | 0.565714286 |
| journalism | 1 |
| games | 0.4375 |
| food | 0.47826087 |
| film & video | 0.573033708 |

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

The most recent data in this dataset is 1/27/2020 which is 4 months before the COVID lockdown. The lack of more modern data makes generalizing the current data set to a post-covid era extremely fallible.

Some of the data represented in this set was incomplete and needed to be removed.

Although Journalism showed a 100% success rate, only 4 campaigns were reported. The low number of reported campaigns relative to other categories could have been grounds to omit the category entirely.

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

An additional table & graphic would be similar to the “rates of success” table I included above.

Diving deeper into the likelihood of success (and failure) would provide valuable information for the types of categories that are be worth investing and pursuing as an entrepreneur.

* **Use your data to determine whether the mean or the median better summarizes the data.**

When summarizing the successful/failed data independently, I believe the use of mean is appropriate. However, if we are summarizing the successful data relative to the failed data (or visa-versa), I believe the median would be more representative due to the large difference in sample size (there is about half as many failed campaigns as there are successful campaigns.)

* **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

The successful campaigns have more variability, and it does make sense. Primarily because there are more successful campaigns and therefore, more backers.

In addition, they could have more variability within their number of backers (i.e., many small donations or a few large donations). I believe this will have a large impact on the variability of our dataset.