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

Based on the data below, all the proposal the success of all sectors were much higher than failed except food and technology. The goals of the sectors that failed might not be realistic.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Count of state** | **Column Labels** |  |  |  |  |  |
| **Row Labels** | **canceled** | **failed** | **live** | **successful** | **(blank)** | **Grand Total** |
| film & video | 40 | 180 |  | 300 |  | 520 |
| food | 40 | 100 |  | 60 |  | 200 |
| games |  | 100 |  | 120 |  | 220 |
| journalism |  |  | 20 | 4 |  | 24 |
| music | 87 | 304 |  | 309 |  | 700 |
| photography |  | 17 |  | 203 |  | 220 |
| publishing |  | 80 |  | 157 |  | 237 |
| technology | 125 | 256 | 6 | 213 |  | 600 |
| theater | 57 | 493 | 24 | 819 |  | 1393 |
| (blank) |  |  |  |  |  |  |
| **Grand Total** | **349** | **1530** | **50** | **2185** |  | **4114** |

The data of sub categories can give a better picture to analyze in which areas that they usually failed.

The month of the deadline for each categories also has some patterns. For instance, the categories that set their deadline for Jan, April, May and Octo are having a higher rate of success compared to other months.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sum of id** | **Column Labels** |  |  |  |  |
| **Row Labels** | **canceled** | **failed** | **live** | **successful** | **Grand Total** |
| Jan | 61907 | 269381 | 5848 | 364243 | 701379 |
| Feb | 47112 | 237199 | 42361 | 428810 | 755482 |
| Mar | 50111 | 210232 | 72660 | 347987 | 680990 |
| Apr | 35911 | 217472 |  | 402512 | 655895 |
| May | 36953 | 295702 |  | 546480 | 879135 |
| Jun | 47830 | 310907 |  | 466168 | 824905 |
| Jul | 63173 | 306171 |  | 421857 | 791201 |
| Aug | 50088 | 280810 |  | 388145 | 719043 |
| Sep | 44297 | 237094 |  | 319713 | 601104 |
| Oct | 25092 | 311822 |  | 379018 | 715932 |
| Nov | 63126 | 212863 |  | 364934 | 640923 |
| Dec | 44101 | 230079 |  | 220272 | 494452 |
| **Grand Total** | **569701** | **3119732** | **120869** | **4650139** | **8460441** |

1. What are some limitations of this dataset?

What I realized is that there were so many overlaps (higher dataset) in certain categories so that data can be significant but there are some categories had

Less numbers which it is why may not be accurate to compare or contrast. In order to compare different categories, datasets should be as close as possible to make meaningful comparison.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Count of state** | **Column Labels** |  |  |  |  |  |
| **Row Labels** | **canceled** | **failed** | **live** | **successful** | **(blank)** | **Grand Total** |
| animation |  |  |  | 100 |  | 100 |
| art books |  |  |  | 20 |  | 20 |
| audio |  |  | 20 | 4 |  | 24 |
| children's books |  | 40 |  |  |  | 40 |
| classical music | 40 |  |  |  |  | 40 |
| documentary | 40 | 80 |  | 60 |  | 180 |
| drama |  |  |  | 80 |  | 80 |
| electronic music |  |  |  | 40 |  | 40 |
| faith |  | 23 |  | 37 |  | 60 |
| fiction |  |  |  | 40 |  | 40 |
| food trucks | 20 | 80 |  | 40 |  | 140 |
| gadgets | 3 | 17 |  |  |  | 20 |
| hardware |  | 57 |  | 83 |  | 140 |
| indie rock | 27 | 113 |  | 20 |  | 160 |
| jazz |  | 60 |  |  |  | 60 |
| makerspaces | 20 |  |  |  |  | 20 |
| metal |  | 20 |  |  |  | 20 |
| mobile games |  | 20 |  | 20 |  | 40 |
| musical |  | 93 |  | 47 |  | 140 |
| nature |  |  |  | 20 |  | 20 |
| nonfiction |  | 40 |  | 20 |  | 60 |
| people |  | 3 |  | 17 |  | 20 |
| photobooks |  | 14 |  | 146 |  | 160 |
| places |  |  |  | 20 |  | 20 |
| plays | 37 | 340 | 5 | 684 |  | 1066 |
| pop | 20 | 17 |  | 3 |  | 40 |
| radio & podcasts |  |  |  | 20 |  | 20 |
| restaurants | 20 |  |  |  |  | 20 |
| rock |  | 71 |  | 189 |  | 260 |
| science fiction |  |  |  | 40 |  | 40 |
| shorts |  | 40 |  | 20 |  | 60 |
| small batch |  | 20 |  | 20 |  | 40 |
| space exploration | 40 | 20 |  |  |  | 60 |
| spaces | 20 | 60 | 19 | 88 |  | 187 |
| tabletop games |  |  |  | 80 |  | 80 |
| television |  | 60 |  |  |  | 60 |
| translations |  |  |  | 57 |  | 57 |
| video games |  | 80 |  | 20 |  | 100 |
| wearables | 24 | 60 | 6 | 110 |  | 200 |
| web | 38 | 102 |  | 20 |  | 160 |
| world music |  |  |  | 20 |  | 20 |
| (blank) |  |  |  |  |  |  |
| **Grand Total** | **349** | **1530** | **50** | **2185** |  | **4114** |

1. What are some other possible tables and/or graphs that we could create?

-Regression analysis to compare subgroups.

- The correlation between backer-counts and status might give an interesting outcomes