|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns? 2. Overall the campaigns had fair outcome with more than 50% campaigns meeting their goal. There were little over 4,114 campaigns and 2,185 of them already met or exceeded their pledged goal. While 50 campaigns were still live, the rest were either cancelled or failed. 3. United states had the most highest number of campaigns with almost two third of campaigns. 4. Out of nine different categories the campaigns were run, category “Theater” had the highest number of total campaigns and successful campaigns followed by music and technology. Category Journalism had the least number of campaigns to begin with and all of them got cancelled eventually. In conclusion, “Theater”, “Music” and “Technology” were most favorable and Journalism was least favorable among fundraisers. 5. “3Doodler: The World's First 3D Printing Pen” (category technology) had the highest number of backers and pledged donation. And similarly some of the other technology campaigns had high number of backers and exceeded the funding goal by huge percentage. In conclusion, category “Technology” was able to attract more donors and more funding. 6. What are some limitations of this data set? 7. Incomplete information: given the information that some of the campaigns are still live, we cannot provide 100% accurate number of successful, failed or canceled campaigns. 8. Certain fields have very extreme outliers values, that makes it harder to come up with some of the statistical analysis. | | | | | | | | | |
|  | | | | | |  |  |  |  |
| 3. What are some other possible tables and/or graphs that we could create? | | | | | | | |  |  |
|  |  |  |  |  |  |  |  |  |  |

Other graphs that we could create are:

Graphs showing average donations per category.

Graphs showing Percent funded per category.

1. Bonus Statistical Analysis
2. Use your data to determine whether the mean or the median summarizes the data more meaningfully.

* The median better summarizes the data meaningfully because the number of backers per campaign is really skewed in the data. Since there are extremely small and large values (outliers), the value of the mean is distorted by the outliers.