Benjamin Roseburrough

Homework #1

Conclusions About Kickstarter Campaigns

Based on the available data, there are multiple conclusions that can be drawn about Kickstarter campaigns. Here are some of them.

1. The highest amount of Kickstarter campaigns by far are for theater-this is driven by initiatives for plays.
2. Journalism has the fewest Kickstarter campaigns-this is likely in part, because it only possesses one sub-category (audio).
3. The amount of Kickstarter campaigns by ‘State’ vary considerably from year-to-year, but as an aggregate, two observations can be made:
   1. The amount of successful Kickstarter campaigns tends to peak in May and then generally go down for the remainder of the year.
   2. The amount of failed and canceled Kickstarter campaigns tends to remain consistent over the course of a year.

Limitations of the Dataset

Here are some limitations of the dataset.

1. There are little more than 4,000 rows of data, which means that this is likely only a very small portion of all the Kickstarter campaigns that have been run since 2009.
   1. The conclusions drawn here are likely only estimates of the true trends involved with Kickstarter campaigns.
2. Kickstarter has a number of categories such as ‘Art’, ‘Comics’, ‘Crafts’, and ‘Dance’ that are not represented here in the data.
   1. This omission could lead to conclusions that are not fully representative of trends.
3. Based on a Time Series Plot, there was a sharp upward trend in the number of Kickstarter campaigns of all ‘States’ from 2009 to 2015, but particularly with respect to successful and failed campaigns. After 2015, there was a precipitous drop in campaigns.
   1. Assuming the data is representative of Kickstarter campaigns at large, this could indicate that there was some controversy or failing associated with Kickstarter.
   2. Alternatively, it could indicate that people found a more suitable substitute for crowdfunding campaigns.

Alternative Tables & Graphs

There are other possible tables & graphs that could be created to convey more information.

* Control Charts could be created to show trending & outliers in the yearly or monthly ‘State’ counts.
* Distribution Identification Charts could be made to determine the probability distributions of ‘State’ counts by ‘Category’ from year-to-year to see whether the data is Normal, or whether there are any categories that consistently impart special-cause variation.
* ANOVA or Mood’s Median Analyses could be made to compare yearly ‘State’ counts with ‘Category’ data to see whether the differences between years are statistically significant or not.