5/28/23

Ryan Lund

Module 1 Homework Answers

Data Analytics Bootcamp 2023

1. Three conclusions that can be drawn from the Crowdfunding Data are: 1) Successful campaigns appear to have peaked in the month of July while failed campaigns peaked in March and August. Canceled campaigns also peaked in August. 2) Plays accounted for the largest portion of campaigns at 34.4% and therefore had the highest number of successful campaigns (187) and highest number of failed campaigns (132). 3) Film and Video had the second highest number of outcomes at 178 compared to the highest, theater, at 344. In film and video campaigns, the U.S. had 130 outcomes and therefore accounted for the majority of campaigns out of all the countries that had campaigns.
2. Some limitations of the dataset include the projected number of backers and city of the country that the backers where from.
3. Additional analysis that could be performed with the data set is number of days that the campaign lasted. By calculating the number of campaign days, a graph comparing the number of campaign days to the pledged amounts could be created to analyze the relationship between how long a campaign lasted and the amount of pledged raised over the campaign.

**Statistical Analysis Answers** (note, these answers are also in the excel workbook on the ‘Statistical Analysis’ tab).

1. There is more variability in successful campaigns than failed campaigns, indicating that there were more different types of successful campaigns than different types of failed campaigns. I believe this makes sense because the appeared to be a negative relationship between the number of pledged and failed campaigns: Failed campaigns appeared to have less number of pledges compared to successful campaigns.
2. Distribution of successful backers: The data is skewed to the right. The mean is a better measure of the data because it is near the center and the median is small and near the first quartile of data. Also, there are no outliers that would skew the mean upward or downward.
3. Distribution of failed backers: The data is skewed to the right. The median is a better measure of central tendency than the mean because there are two outliers beyond the upper whisker that are skewing the mean upward.