Atilla Toy

Data Science Bootcamp HW#1

1. Conclusions:

A. Theater generates the most kickstarter campaigns, the majority of which are for plays

B. Music is most likely to have successful campaign, having 5 subcategories with a 100% success rate.

C. May-Jun-Jul is the most active time of year for KickStarter campaigns in sheer quantity

2. Data limitations:

This data lacks specific information on the donors. If each donation was logged there would be more potential for analysis:

While you can determine the country of origin for the campaign, you cannot tell where the donors are from. This limits the ability to track global appeal, (i.e. a UK Kickstarter receiving donations from the US)

It would also be useful to know the dates of specific donations. This could tell us the rates at which donations come at various stages of a campaign.

Finally, subcategories can be more granular, using keywords from the title and blurbs.

Example:

|  |  |
| --- | --- |
| Jane Don't Date - TV pilot (sitcom) | Cursed with attracting odd men, an independent woman takes on the Chicago dating scene again with the help of her offbeat friends. |

This Kickstarter could be tagged with “sitcom”, “dating”, “Chicago”

3. Additional Analysis:

A table and chart comparing campaign length (start date subtracted from end date) to campaign state could tell us about the lifespan of a campaign and be compared across categories and years. Campaigns finishing faster could indicate popularity, and may differ by year and category.

A table and chart comparing average donation to completion rate could show us if big time donors or sheer numbers of smaller donors are more likely to generate a successful campaign. Also comparing categories could tell us which categories court the biggest donations from the average donor.