Heidi Landenberger

UC Berkeley Data Analytics Bootcamp

Homework Assignment #1

1. **What are three conclusions we can make about Kickstarter campaigns given the provided data?**
   1. Some general conclusions we can draw are:
      1. As the size of the goal gets larger, it appears that the campaigns tend to fail more.
      2. Approximately ¾ of Kickstarter’s campaigns have come from the United States and the United Kingdom.
      3. It appears that, by far, the most common subcategory of campaign on Kickstarter is Plays.
      4. It looks like campaigns are somewhat more likely to fail if they start in December, which makes sense, given that most people will likely be spending their money on the holidays, not some newfangled widget or someone’s play.
      5. It appears that people really don’t like food truck or restaurant campaigns, since nearly all of those fail.
2. **What are some of the limitations of this dataset?**
   1. We have no information on fraud. Chances are some of the cancelled campaigns were attempts at money laundering or fraud that were shut down by Kickstarter, but there’s nothing noted in the file on reasons for cancellation.
   2. There is no information on, if a campaign was successful, the rate of customers getting what the product they paid for.
3. **What are some other possible tables/graphs that we could create?**
   1. I would be interested in seeing the following graphs and/or tables:
      1. Graph showing country vs state across the different categories: How are campaigns in non-English-speaking countries performing?
      2. Table showing percent funded split into multiple buckets (less than 10% funded, 10-25%, 26-50%, 51-75%, 76-100%, 101%+) across the goal sizes, filtered by categories. Are certain sizes of campaigns more likely to be overfunded? Or does overfunding have anything to do with certain categories?