Carré Perkins Homework Assignment #1

1. What are three conclusions we can make about Kickstarter campaigns given the provided data?
   1. Looking at the combined past 7 years, successful campaigns peak in May. The only month in which more campaigns failed than were successful occurs in the month of December. Overall, campaigns have the highest rate of success in the first half of the year, with the worst month occurring in March. Failed campaigns are relatively low between Feb and Apr. Based on time of the year alone, I would recommend running a campaign in the months between Feb and May.
   2. Looking strictly at Campaigns by Category, “Music” has the greatest success rate at 77% followed by “Theater” at 60% and “Film and Video” at 58% (ie 58% of all Film and Video campaigns were successful). “Food” had the greatest failed rate (70%) followed by “Games” at 63%. “Photography” and “Publishing” each had a 53% failure rate. I would recommend a campaign in “Music”, “Theater” or “Film & Video”
   3. Of campaigns categorized as “Theater”, Plays is the most successful sub-category. For “Music”, Rock is the most successful sub-category and for “Film & Video”, Documentary is the most successful subcategory.

I would recommend one starts a kickstarter centered around Documentaries, Rock, or Plays that run during Feb – May for optimal chances of success.

1. What are some of the limitations of this dataset?

This data does not take “curbside” appeal of the campaign into account. How long the blurb or story on the site is, number of updates made, number of comments, number of pictures, etc.

1. What are some other possible tables/graphs that we could create?

I think it would be worthwhile to dive into what categories have been successful only for the past 2-3 years. As we know, life on the internet is fast paced and trends come and go quickly. Maybe successful campaign months are beginning to shift; maybe successful categories are beginning to shift. Although looking at 7 years of data provides many more data points, perhaps the older sets aren’t relevant and might be adding noise to the analysis.