**Data Analysis HW 1**

**Kickstarter – Report**

1. **What are three conclusions we can make about Kickstarter campaigns given the provided data?**
   1. Most campaigns on Kickstarter are from arts and creative industry: Film, Music, Photography, Journalism, Publishing,

and theater.

* 1. By Category: Theater(overall) has highest number of campaign, with highest successful outcome. If we just focus on US, theater still leads the number of campaign but music category has highest successful outcome.
  2. By Sub-Category: Play has the highest number of campaign with highest successful outcome. There were few sub-categories like: classical music, electronic music, hardware, rock, shorts, television with 100% successful outcome.
  3. By Monthly: Campaign started in first 6 months (January – June) are more likely to success as compare to those started in last six months (July – December) of the year.

* 1. By Goal: Campaign with goal of less than 10,000 are more likely (more than 50%) to be successful. Campaign with goal of 45,000 or more are more likely (more than 50%) to fail.

1. **What are some of the limitations of this dataset?**
   1. Some of the data (Goal column) are not normalized into uniform currency
   2. I found ambiguity around “canceled” and “failed” outcome. What are the definitions used to define the outcome?
2. **What are some other possible tables/graphs that we could create?**
   1. We can create very basic pie chart to see the overall outcome of Kickstarter campaign.
   2. On the monthly-outcome summary sheet we can add more filters: by category and by sub-category, in either case we can compare outcome by year and month.
   3. We can create a graph to find the relationship between pledge and campaign (days open) or any other data columns.