UC Berkeley Data Analytics Bootcamp

Martin Gutierrez

Homework 1

Kickstart My Chart

Create a report in Microsoft Word and answer the following questions...

1. What are three conclusions we can make about Kickstarter campaigns given the provided data?

a) Roughly 68% (2833 of 4114) of all Kickstarter campaigns have been for arts (theater, music, film & video and photography) and approximately 14.4% for technology and approx.. 6.3% for publishing and journalism. Games and Food have been the least popular campaigns with approx. 5.3% and 4.9% of campaigns, respectively.

b) It is no surprise then that the highest success rates are also in arts, (from high to low in) music 77%, film & video 57.8% and theater 60%. Likewise, the lowest success rates are in journalism 0% success rate, food 17% and games 36%.

c) The success rates of the campaigns have been highest when the goals are smallest (from approx. 70% from goals of less than $1000) and they go down in linear fashion as the goals increase reaching a low point of 19% for campaigns of more than $50,000.

In contrast, the lowest failure and cancellation rates are in the campaigns with smallest goals (from approx. 25% failed and 4% canceled for campaigns with goals of less than $1000) and they grow in almost linear fashion as the goals increase; the highest failure rates (57% failed and 22% canceled) are found in campaigns with goals of over $50,000

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

a) There is no way of knowing the range of donations for each campaign. Some campaigns may have been funded by a few big sponsors while others could have been made by grassroots actions.

b) There is no way to know the actual ROI of each campaign

c) There is no way to know the rate of donations for each campaign

d) Since there are no exchange rates we cannot compare the actual funds obtained in different countries

e) The data does not explain what spot-light or staff pick means

f) Unless we add new columns with percentages it is difficult for the pivot tables to give us results with % and not with raw numbers

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

We could create tables and graphs of:

a) state using percentages

b) funds raised or goal attainment by length of the campaign

c) length of campaign vs success and failure rates

c) success rate, funds raised by subcategory

d) goals by country and category and subcategory

e) success rate by category by month

f) length of campaign by category and subcategory

g) length of campaign or success of campaign by month of year and by country