Michelle Fitzpatrick

**1 Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**

(1) Journalism start-ups seem to have a hard time flourishing, as we can see that all start-ups under this category were canceled. On the other hand, start-ups in the music and theatre industries seem to do better, with over half the start-ups in these categories succeeding.

(2) Looking by subcategory, we can clearly see that plays seem most successful, with a bit over a 65% success rate. On the other hand, all projects under art books, audio, science fiction or world music were canceled. Additionally, all projects under animation, children’s books, drama, fiction, gadgets, jazz, mobile games, nature, people, places, restaurants, video games failed.

(3) Finally, using the Pivot Table by Date, we can see that more Kickstarter projects were successful (2,185) than failed (1,530) or were canceled (349). Put together, failed and canceled projects accounted for 46% of all projects (1,879 projects).

**2 What are some limitations of this dataset?**

While we can parse apart variables such as number of backers, amount invested, category and subcategory, it is more difficult to analyze the data in the blurbs, which may provide abstract insight into why one project may have been more successful than another. Is one project more novel, creative, or able to reach a wider audience? Is there a specific need for some projects over others? Additionally, we also lack data regarding the number of backers a project garnered during certain time points. It would be interesting to see if certain projects pick up investment momentum as time goes on, or whether projects that are highly invested in to begin with do better than those who get more backers with time. Finally, we lack knowledge as to why certain projects were canceled. What separates them from those that failed?

**3 What are some other possible tables and/or graphs that we could create?**

One could correlate the amount invested (by average or total) to the state (success, failed, canceled, live) using a scatter plot to determine the relationship between these two variables. This could serve as a good indication as to whether more heavily founded projects led to success more often, or perhaps with greater ease, than those less-funded projects. One could also create a scatter plot delineating the relationship between state and year, to uncover any specific correlation or trends between these two variables. This would serve as an indication as to whether start-ups were more likely to thrive or fail during certain years. This way one could go as far to correlate a projects’ state to the how the economy is performing, to see if this has any effect on budding businesses.

One could also create a stacked column chart that analyzes the relationship between the project’s origin country and the amount of successful, canceled, failed or live projects coming from each country. This could help us determine any regional or cultural variables that may be contributing to a project’s likelihood to succeed or fail. Finally, we could also inspect the relationship between date created and date ended to see if the projects’ timelines may be contributing to their respective successes or failures. Do projects with a tighter timeline do better or worse? Do projects with lengthy time lines meet their goals at the same rate as those with tighter timelines?