UT DATA ANALYSIS & VISUALIZATION: EXCEL HOMEWORK

Without further analysis, any conclusions drawn here cannot be quantified. However, some comparisons can be made between different categories, subcategories, and launch dates.

One type of comparison is the number of campaigns in each category. From the Category Stats analysis, it can be observed that *theater* has the most campaigns out of any category, followed by *music* and *technology*. However, a significant number of technology campaigns were canceled; if canceled campaigns are ignored, the category with the third-most campaigns is *film & video* instead of *technology*. Further data on why campaigns are canceled could provide insight into these patterns and determine when Kickstarter is most helpful.

Turning to subcategories, *plays* were by far the most highly represented—both within the *theater* parent category and overall. However, the second- and third-most highly represented subcategories were outside of the *theater* category, being *rock* (*music*) and *documentary* (*film & video*), respectively. Interestingly, both of these subcategories consisted entirely of successful campaigns, whereas a significant proportion of *plays* campaigns were unsuccessful. It would be interesting to investigate whether frequent success within a subcategory might make people working in that subcategory more likely to consider Kickstarter than in other subcategories.

Start dates do not have a clear effect on success or overall number of campaigns. However, there are some differences between different times of the year. For example, more successful campaigns were started in May than any other month. This observation alone is likely not useful, but it can be contrasted with the fact that there was no spike in failed campaigns the same month. It is possible that starting a campaign in May has some positive effect. It is also possible that something occurs in May, such as a Kickstarter convention, that helps campaigns gain the knowledge and energy they need for a good head-start. Without further data, the exact cause cannot be determined. However, a stacked area chart, rather than the raw line chart that was instructed, can be used to show that May was also the most popular month to start campaigns. This could support the hypothesis that there was an event that prompted people to kick-start their Kickstarters around May. In contrast, December was the least popular month.

Another useful type of chart for this dataset would be normalized (“100%”) charts to show relative rates of success among categories, subcategories, and start dates. Using these charts uncovers that the category of *music* was the most successful, followed by *theater* and then *film & video*. The data are too granular to glean reliable conclusions among subcategories in this regard. For start dates, it can be observed that the months of February through May have higher success rates than others. December had the lowest success rate of any month. There appears to be a correlation between success rate and popularity of start dates. However, only speculation can tell whether one of these effects causes the other.