**Challenge 1**

Given the data we can draw three conclusions about crowdfunding campaigns.

1. The majority of campaigns are successful. This is true across all categories except “games” and within games, it is only “mobile games” that fail more often. The only other subcategory with more failed campaigns is science-fiction books.
2. The overall best time to start a campaign is in the early summer months (June or July). However, this varies for some categories and other categories have too few data points to draw specific conclusions.
3. The United States has the vast majority of campaigns (>75%). This is so large that campaign categories that are more common in non-US countries are hidden by the large number of US campaigns unless specifically filtered to. For example, Film & Video and Music are more popular in Great Brittan (GB) than Theater, which is the most popular overall and in the US by a large margin.

**Limitations**

One limitations of this dataset are that we do not know the range or exact amounts of donations. Some campaigns may have only succeeded due to a very small number of large value donations. Without this data, we cannot do a statistical analysis to determine how many donation outliers without all donations for each campaign.

This data does not give provide any information about outside influences that would drive people to make back the campaign. The best we have in number of backers. We do not have a value for hits on the campaign webpage to measure how many views the campaign got, and we do not know if any advertising and marketing was done that could have led to more donations. This would likely have a major effect on the success of a campaign.

Another limitation is that the donation amounts and totals are in different currencies. While in this activity we do not make any charts using any of the currency values, if we were to do so, we would need to convert to a single currency and preferably with a conversion value unique to the start or end date for each campaign.

**Other Tables/Charts**

One possible useful value could be calculating the total duration of a crowdfunding campaign. This column would be straightforward to calculate by subtracting the ending timestamp from the beginning timestamp and converting the value to number of days. The created pivot table could also be filtered by category. It would be useful to know if leaving a campaign going for longer could lead to a successful outcome.

Another useful set of charts would be to use the 100% Stacked Column charts with pivot tables for “staff\_pick” and “spotlight” determine if they had a meaningful effect on the success of a campaign.

A third useful chart could be to graph the number of successful campaigns vs the goal amount. Using this chart, we could see if there is a trend where less campaigns are successful if they have a higher goal.

**Bonus**

In Excel, Sheet “Bonus1”

**Bonus Statistical Analysis**

In Excel, Sheet “Bonus2”

The median summarizes the data more meaningfully because there are very few campaigns with thousands of backers, but their numbers are large enough to significantly skew the mean higher. This is obvious by seeing that the max value for both successful and failed campaigns are over 5 standard deviations away from the mean, but the min value is within 1 standard deviation. I created two histogram plots to show this more clearly.

There is more variability with successful campaigns. This makes sense because at some point, with enough backers, every campaign will succeed. After a campaign has passed the successful threshold, more backers can still be added, leading to large variability in the number of backers.