*Q1. What are three conclusions we can make about Kickstarter campaigns given the provided data?*

A1. A couple of conclusions we can make about the Kickstarter campaign given the provided data are:

1. Categories like file & video, music and theaters have more than a 50% success rate whereas other categories like journalism and food have less than 25% success rate. Breaking this up into sub-categories, we can observe that sub-categories like documentary, hardware, indie rock, rock & tabletop games have an extremely high success rate whereas others like animation, food trucks, video games, jazz, and drama have an extremely high failure rate
2. The success rate is higher for projects having a lower Kickstarter goal e.g. projects with a goal of less than $10,000 has a success rate of 71% whereas projects with a goal of >=$50,000 have a success rate of 19%
3. The average cancelation rate for projects having a goal of greater than or equal to $25,000 is more than 2 times of those having a goal of less than $25,000 e.g. projects with a goal of >=$25,000 have an average cancelation rate of 16.3% whereas projects with goal of less than $25,000 have an average cancelation rate of 7.2%
4. Project with creation date in the first half of the year (Jan – July) seems to have a higher success rate that projects with creation date in the second half of the year (Aug - Dec)

*Q2. What are some of the limitations of this dataset?*

A2. Some of the limitations of the dataset are

* Although we are able to calculate the average contribution based on the pledged amount and no of backers, we don’t have clarity on how much was contributed by each backer. This is important is 90% of contribution can come from one user or it could be more evenly distributed. This gives us an insight into the weightage for the backing
* We don’t have details on why a project was canceled and we also don’t have the date of cancelation. If projects were canceled closer to the deadline as they are not able to raise the goal amount, technically they are failed projects and not canceled. Also, insight into the reason for cancelation can give a lot of valuable information
* The data has a launched date and deadline date but does not indicate what date the project reached its goal. Having this data can help in identifying the time it took for the project to reach it’s goal and we can view this information at a country, category or sub-category level and identify useful patterns
* No clarity on how some data is being set. e.g. spotlight value seems to be true only for successful projects but its not clear if that is the only factor that affects the spotlight

*Q3. What are some other possible tables/graphs that we could create?*

A3. Other possible tables/graphs that we could create are

* A graph to view the relation between % of projects that met the goal (i.e. successful) vs. country to see if projects in a certain country become more successful than others
* A graph to view the relation between projects picked by staff (Staff picked as True) vs. the ones that were not picked by staff and see how effective the staff was in picking projects that became a success
* A graph to view the goal of a project vs. the number of backers for successful projects to see projects with larger goals just had more backers or if they had a similar number of backers as that of a smaller goal project but had larger average contribution per backer.
* A graph to view the impact of the popularity of a project based on the number of backers vs. success/failure of the projects to see how popularity impacts its success
* A graph to view no of backers and their average contribution for successful projects to see if it were few people that contributed most of the money for successful projects or many people that contributed smaller amounts to result in the success of a project