1. What are three conclusions we can make about Kickstarter campaigns given the provided data?
   1. Kickstarter project categories launched in the *US* that have a decent chance of being successful are\*:

1st = **music** with about a 79% success rate.

2nd = **film & video** with about a 62% success rate.

3rd = **theater** with about a 58% success rate.

\*these summary stats do not consider “live” projects.

*Global* Kickstarter stats indicate that the most successful project categories are\*:

1st = **music** with about a 79% success rate.  
 2nd = **film & video** with about a 58% success rate.  
 3rd = **photography** with about a 47% success rate.  
 \*these summary stats do not consider “live” projects.

Among those categories chosen, the most successful sub-category projects to launch in the *US* and *global* arenas are:

US:  
 music:  
 1st = **rock** with about a 37% success rate.  
 2nd = **indie rock** with about a 21% success rate.  
 film & video:  
 1st = **documentary** with about a 39% success rate.  
 2nd = **television** with about a 12% success rate.  
 theater:  
 1st = **plays** with about a 45% success rate.  
 2nd = **spaces** with about a 7% success rate.

Global:  
 music:  
 1st = **rock** with about a 37% success rate.  
 2nd = **indie rock** with a 20% success rate.  
 film & video:  
 1st = **documentary** with about a 35% success rate.  
 2nd tie = **shorts** and **television** with about a 12% success rate.  
 photography:  
 only sub-category is **photobooks** with about a 47% success rate.

* 1. On average the **best month** of the year to launch a project is **May** with a comparable success rate of 11% against all the other months. The **worst month** is **October** with a comparable fail rate of about 10% against all other months.
  2. Deciding the correct financial goal limit for the project can influence the probability that the project will be successful. The ranges that will produce a higher likely-hood of being successful are **less than 1k to 20k** and **35k to 40k**.

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

One limitation of this dataset is that it is very general in the sense that it only shows the country from where the project is launched. It would be beneficial to have some more detail about project locations, either city or region. Having that information may indicate that local networking and marketing can influence the success of that project.

Another limitation is that to understand individual donation amounts we need to aggregate two columns. It would be nice to have also the max and min amounts, that type of data would indicate if there are possible outliers in the amount donated.

Including the account user name that launched the project and how many projects that account has launched in the past would be great data. Having that data will provide a means to determine the probability of success for first time launchers.

Also, this dataset does not include whether a project’s Kickstarter homepage contains an embedded video that explains what the project is about. From previewing a few number of campaigns on kickstarter.com I noticed that there are many projects that do have a video that explains the project and rewards for backers, while others that do not have a video seem to be failing in achieving their goal. It would be great to verify this assumption with some data, it could be recorded as a binary data type.

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

* A pie chart could be utilized in the first task to indicate a distribution amongst category count.
* REGRESSION analysis (ANOVA table) can be applied to backers\_count vs pledged, with pledged being the dependent variable. If the p-value satisfies a set level of significance and the effect size is strong enough then future projects can estimate how much each backer, on average pledges. This information can help in determining what goal to set.
  + a scatterplot with these same variables can visual represent this relationship.