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

The plot on sheet2 shows that there are more projects in theater, film and music categories than journalism.

On the other hand we can see than between music and theater categories, we can see percentage of successful projects in music is more than theater so we can come up with this conclusion that investing in music is less risky than theater or any other areas.

The bonus plot shows that, as the projects getting bigger the successful rate would decrease while the rate of being failed or canceled would increase. So we can conclude that investing in bigger projects is more risky.

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

I could say that we do not know the exact location of each projects if we knew them maybe we could come up with conclusion about each location.

The other limitation is that for example the word “play” is so wide and there are a lot of subcategory that could help us to find a more inform decision.

And if we knew the interest rate that each would offers to different investors we could understand how this fact could attract investors.

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

We could have done different tests for each variables to see if they are significant or not. To do that we need to do a T-Test for continuous variables and chi-square or fisher exact tests for categorical variables and at the end come up with the regression model with the significant variables.