1. Certain project categories are more likely to have failed projects. For example, food projects have less than a 25% chance of succeeding whereas music projects have higher than a 75% success rate. Also it seems that the month a project is created may have some correlation to whether it fails or not. For example, December was the only month where more projects failed than succeeded (when created in that month). Lastly, projects in the technology and journalism category were the most likely to be canceled. So it may be riskier to fund projects of these types.
2. We don’t know if this dataset is representative of all Kickstarter projects. There are only about 4000 records and we’ve been told that there have been over 300,000 projects launched. If projects of certain category types are over-represented in the sample versus the actual, some of our conclusions may be wrong. For example, it seems that “plays” are the most popular sub-category. But if they are over-represented in this sample then that insight may be wrong.
3. We could create a chart to see if “staff picks” or “spotlighted” projects were more likely to fail or succeed. We could also look at percent funded versus the project category to see if certain categories got way over the funding goal.
4. Statistical Analysis: The mean summarizes the data better because the large number of zero person backed projects skews the median too low.