Homework 1 Questions Section

1. Three conclusions that can be drawn are:
   1. The type of category that a Kickstarter is started under does not have an effect on the probability that the campaign will be successful or fail.
   2. The best time to start a Kickstarter fundraiser is the first quarter of the year if you want the greatest probability of your project being fully funded and thus deemed “successful”
   3. The price goal for a given project does not have an effect on whether or not it will be successful or fail, except in minimum and maximum price ranges.
2. Some limitations of this dataset are:
   1. The legitimateness of the projects that are captured within the data. Projects that are done on a whim or without any real credibility can skew the failed or canceled data.
      1. Some could be random projects that an individual is just doing for fun with no real intentions of producing anything
      2. Some could be legitimate projects that are used to make a difference within their respective categories.
   2. The variance of the length of the Kickstarter beginning and start time can play a huge roll in the percent that fail that may have been successful if the deadline was extended.
   3. The difference between the number of successful, failed, and cancelled makes it difficult to get an accurate reading for comparative statistics because of the larger number of successful projects within the data set.
3. Some other tables/graphs that we could create would be:
   1. A graph showing the length of projects from launch to deadline and their successfulness.
   2. A graph that shows the comparison of the number of backers and successfulness of the kickstarter.

Analysis Question

The mean summarizes the data more meaningfully because it encompasses all data points within the set. Median will not take into account outliers and the affect that the repetition of one number may have on the data.