Project 2 Write Up:

1) **What is/are the question(s)? A question is a question about the world which generated your data. Questions end with question marks. Questions that essentially rephrase "what's in my data?" don't count.**

   The question is, would a higher budget for a film's production lead to a better outcome in terms of higher rating, more revenue, and more popularity? If so, does it have to do with the release year of a film and its genre?

2) Why is this a good design to answer your question(s)? Include a list (bullet points ok) of ways your vis is dynamic and/or interactive, and indicate whether or how each one helps answer your question(s).

- Scatterplot uses position as an encoding method to accurately depict the quantitative relations between budget and a selected outcome.
- Tooltip is used to give specific information of a film when the user moves the mouse over a point in the plot.
- A slider bar allows the user to select the release year of films and enquire if the release years play a role in the relationship between a film's budget and outcome.
- A dropdown selector allows the user to choose among the three measures of outcome to investigate and this helps to answer the question of whether a higher budget of a film leads to a better rating, revenue, or popularity.
- checkboxes allow the user to select the genre of films to filter out films in the plot. Each film is colored with its corresponding genre color and therefore answers whether the relationship we are investigating between budget and investment is also influenced by genres of films.
-

3) What's the answer to your question(s)? What's a conclusion the user can draw that would have been harder to see with a lesser vis?

A higher budget overall does lead to a better revenue, but the same would not really apply for ratings and popularity. There is not a strong correlation between popularity and budget or budget and rating. For instance, for drama movies, a higher budget can sometimes even lead to a lower rating (for example in years from around 1985 to 1995)! This would be hard to see if we do not allow the user to select genres and years.

4) How has the class content informed your design? Also, what do you wish had been taught so that you would have had an easier time doing p2?

Using different colors to represent different categorical genres and position to encode 2 dimensional values are choices that class content informed us. The lessons on the interactivity of JS React adds more dynamics to out visualization. We also take algebraic visualization principles into the consideration of our design to make sure that meaningful changes in the real world would lead to meaningful changes in the visualization as can been seen by the changes in the visualization with respect to changes in year and genre. We are also able to pinpoint the deficiencies in our visualization (overlapping data as hallucinators) and compensate by using tooltips to show more information for a specific film.

Because of lack of knowledge in statistics and how to implement and visualize statistical analysis on data, we are unable to draw more specific and meaningful conclusions. For example, we don't know how to implement a regression analysis to show the linear regression fit. It would be nice if the class had touched more on the preprocessing of datasets.