Stats week 6 - Group Project

Applied Stats: Samanta, Ariana, Melisa, Felipe, Brendan

October 19, 2021

How to Win Awards

The client: Warner Bros. Film Studio The problem: Warner Bros. and their subsidiaries are having trouble winning best picture awards for their films. They want to know why that is - what do other studios have that they don't? What's different about their films? What can they do to maximise their chances of winning a best picture award? The task: Working in teams, you have 90 minutes to analyse the movies.csv dataset to provide insight for Warner Bros. into what makes an award-winning film. How you do that is up to your team, but you'll want to:

a) Spend five or ten minutes thinking about the problem, and how you will address it as a team.

b) Divide the workload appropriately between you. (Will you create smaller teams to focus on different aspects of the workflow? Will you have people focussing on technical assistance? Or visualisation? How will you coordinate activities? How will you share code/results?)

c) Think about the different skills you have learned, and how you could apply these to the case at hand. Remember that any recommendations you make will be stronger if they are supported by the concepts of inference which you have learned.

d) Make sure you have something completed within the deadline!

Analysis Conclusion

Our process: we plotted the correlation of pairs, as illustrated in the graphs below. We compared the means of the 'yes' and 'no' probabilities of our outcome variables - best pic nom, best pic win, best actor win, best actress win, best director win, when paired with our predictor variables - runtime, imdb rating, imdb num votes, critic score, audience score. From these graphs, we were able to come to a conclusion whether there was a significant correlative relationship between the individual variables. Our results are reported below in a table. A 'yes' signals that the mean of outcome variable success (ie - winning award) was two standard deviations away from the mean of outcome variable failure (ie - not winning an award). 'Maybe' signals that the means were observably different, but not of the same statistical significance as a 'yes'. A significant advantage of proceeding with our analysis in this way is the

variable	Runtime	IMDB Rating	IMDB number of votes	critic score	audience score
Best Picture win	yes	yes	yes	yes	yes
Best Picture Nom	maybe	yes	yes	yes	yes
best actor	maybe	no	no	no	no
best actress	no	no	no	no	no
best director	maybe	maybe	no	maybe	maybe

fact that we were also able to observe confounding variables. We found that there is a high correlation between imbd score, critic score and audience score. Additionally, there is also a light correlation between number of imdb votes and: imdb rating, runtime, audience score. ** Our findings were obtained by eye-balling our graphs (due to time constraints).

- ** Our findings were obtained by eye-balling our graphs (due to time constraints).

*** It is important to note that the sample size for 'best pic win' is quite small and so therefore, the obtained results in this analysis should be taken lightly and further analysis with a bigger sample size is needed.

**** We decided to analyse all 'award win' variables because we believe that there may be a relationship between pics winning general awards and then winning best pic.

**** This dataset may be inacurate. For example the hurt locker won best picture but wasn't nominated.

To conclude, our explanatory variables generally appear not to explain variability in our outcome variables for best actress and best actor win. Our explanatory variables generally appear to contribute to variability in our outcome variables for best pic nom and best pic win. There is an insufficient evidence to reach a conclusion on best director win.

So - we believe that Warner Bros. Studios may be more likely to obtain a win in the best pic category from longer running times, and by obtaining higher critic scores, higher imdb scores, imdb number votes, and audience scores.
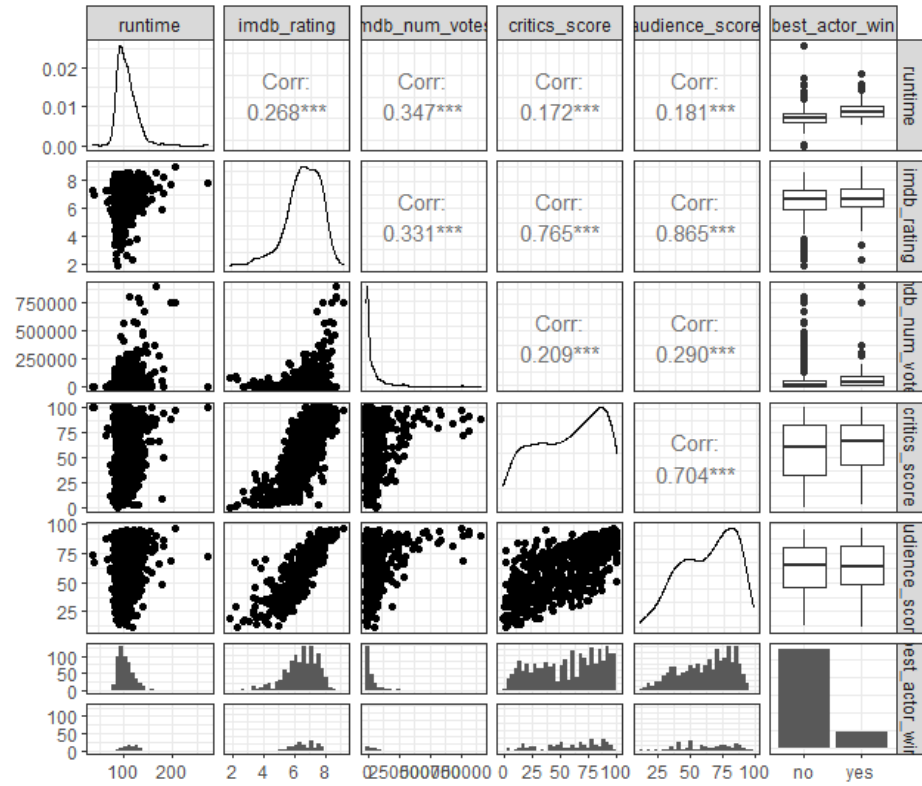
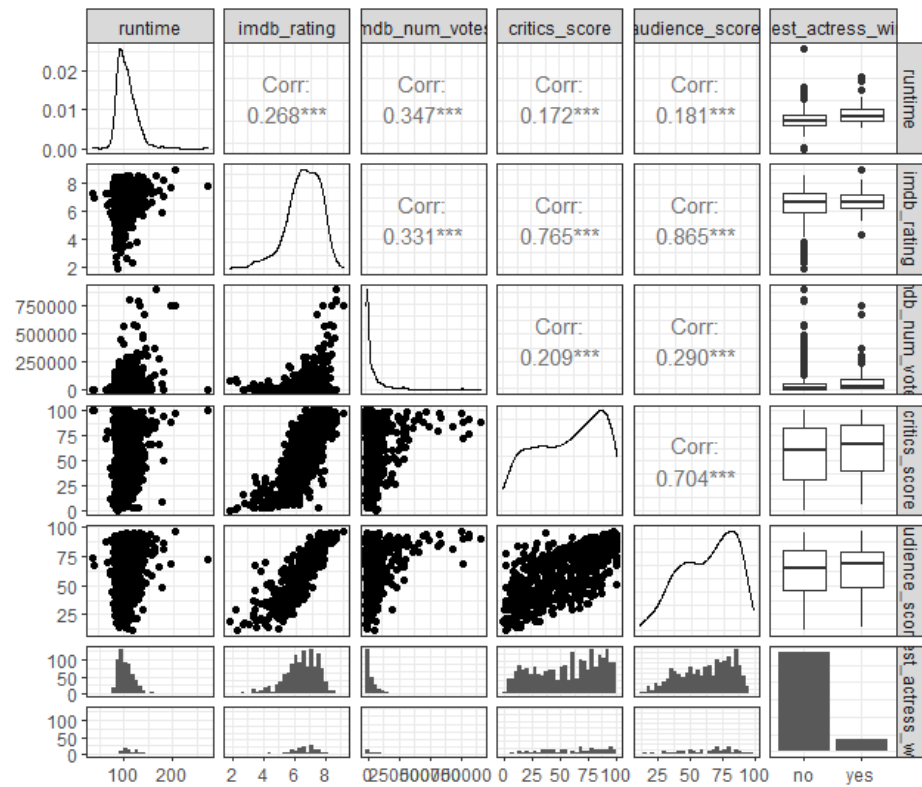# Best Picture Nomination

Figure 1: Best Picture Nomination.
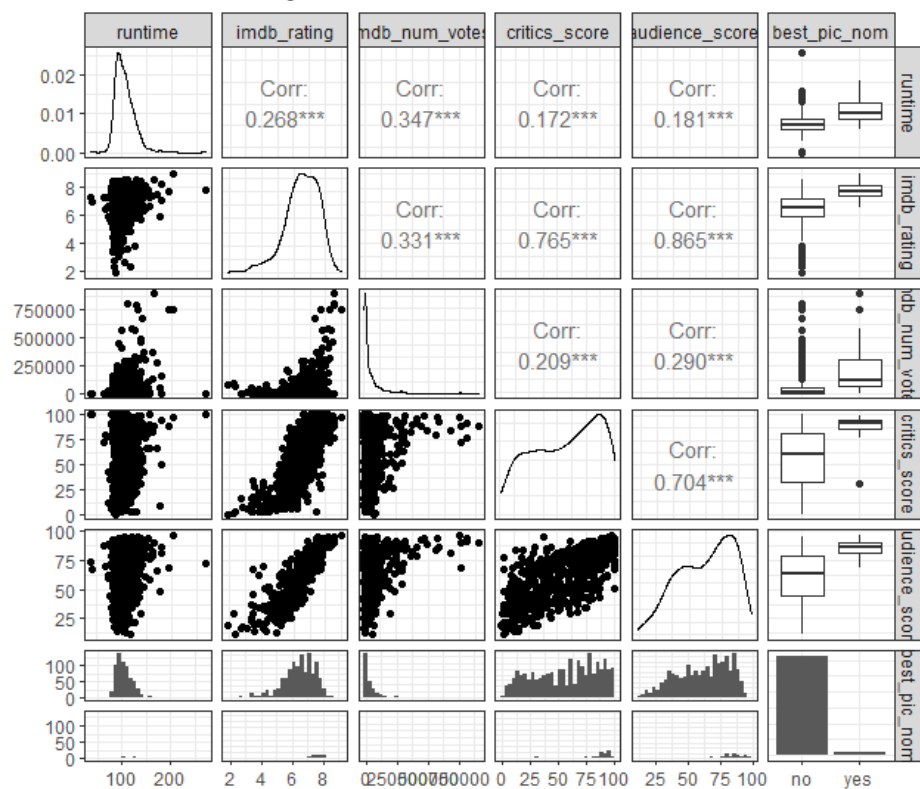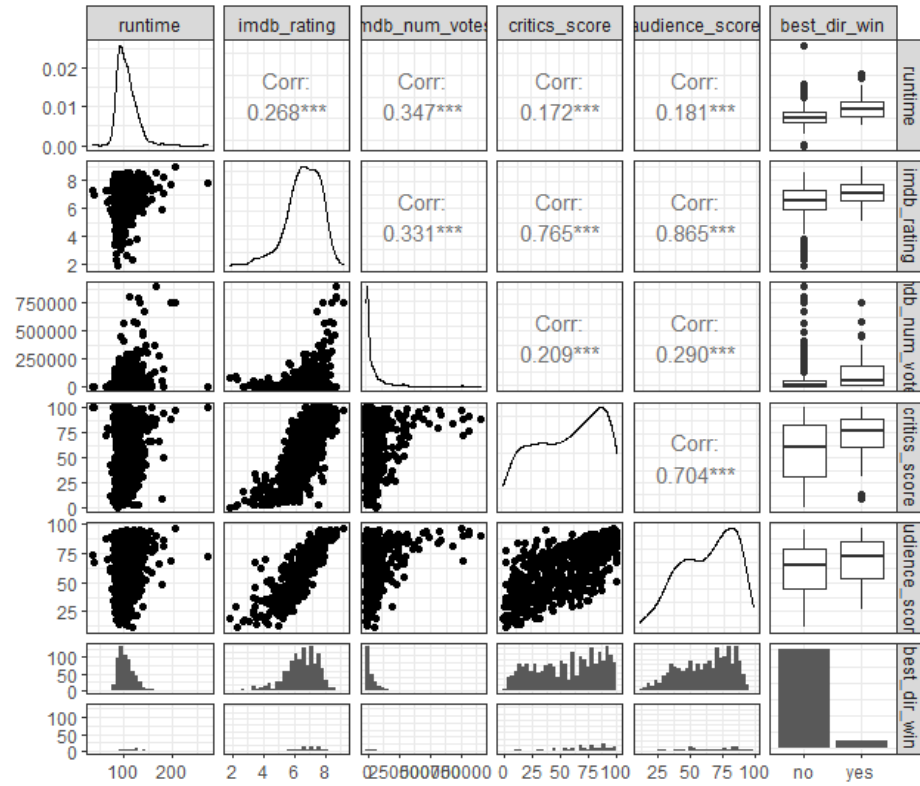
Best Actor Win Plot

Figure 2: Best Actor Win.

Best Actress Win Plot

Figure 3: Best Actress Win.Plot

Best Director Win

Figure 4: Best Director Win.Plot

Best Pic Win

Figure 5: Best Pic Win