



Oscar Winners Data Analysis

By: John Munger, Team Treehouse: Python Data Analysis - Project 8



Overview of Dataset

- The data set being presented includes information obtained from IMDB for 10 oscar winners from the year 2019
- It also includes 5 oscar winning movies from subsequent years.
- We set out to get a basic sense of the data set, and answer a few questions about relationships between certain variables.



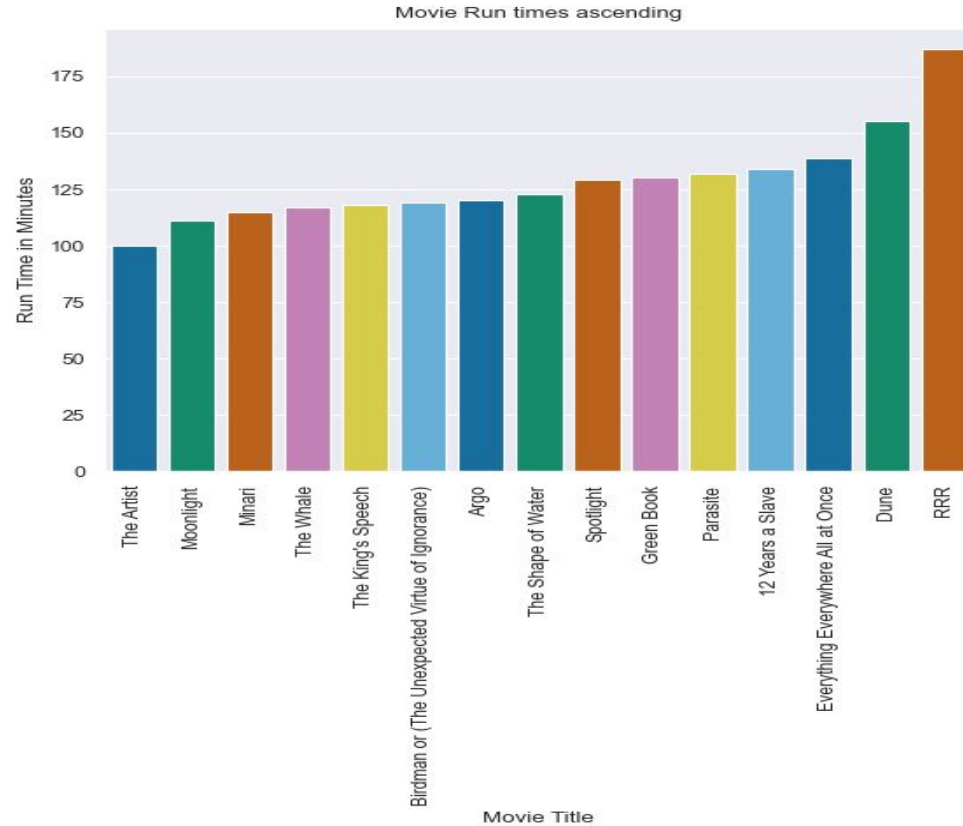
Questions Answered

What was the runtime for each movie and which movie had the longest runtime?

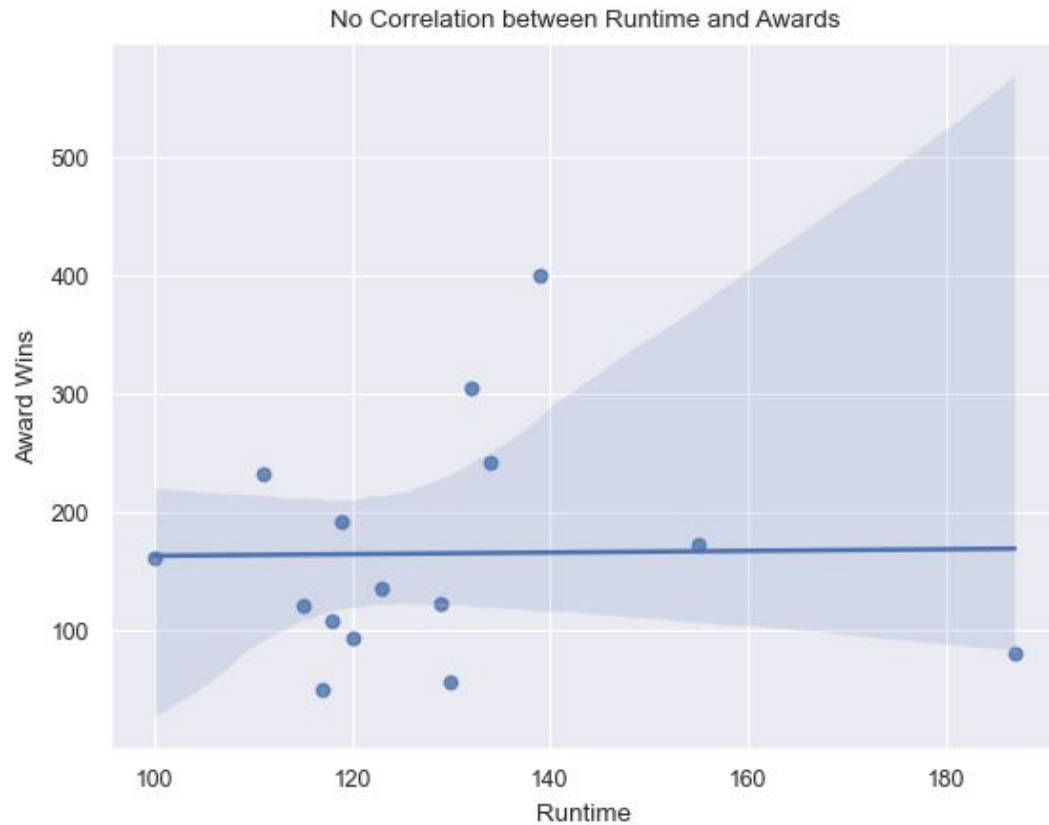
Is there any relationship between runtime and number of Awards won?

Is there a relationship between Box Office sales and awards?

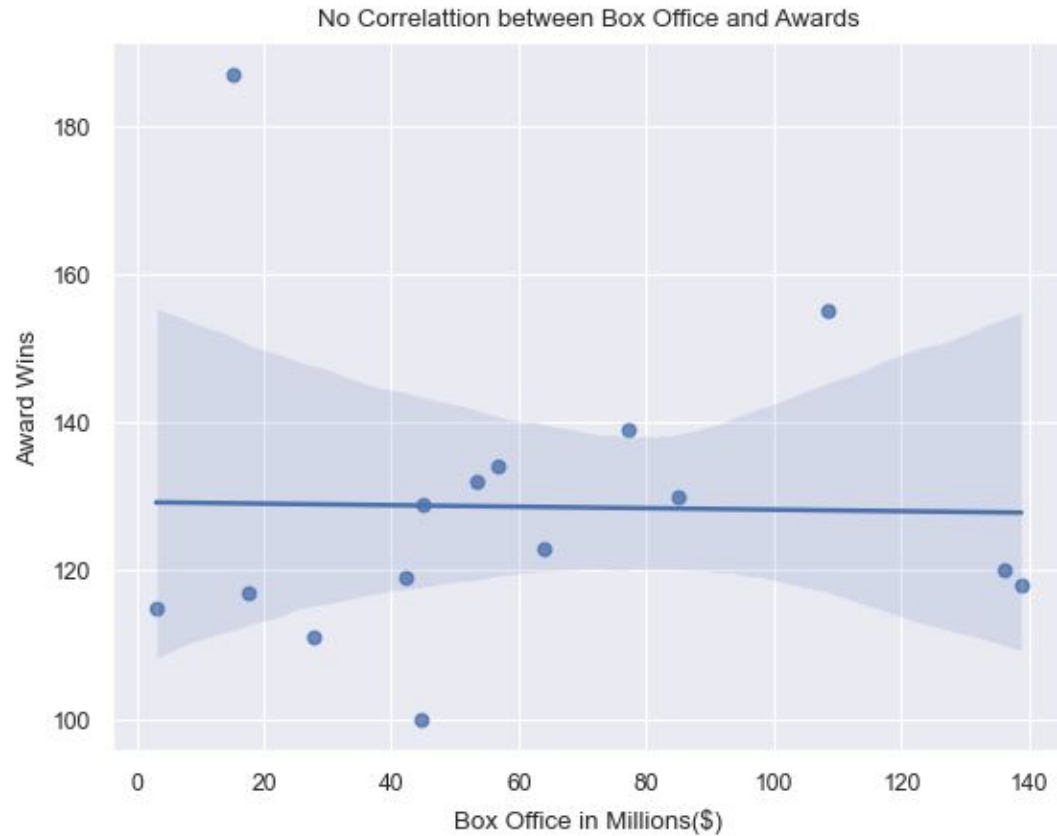
Is there a relationship between box office earnings and movie runtimes?



As we can see, the longest runtime in our dataset was RRR with a runtime of 187 minutes



We found a correlation of 0.104 which is not statistically significant and amounts to no correlation.



We found a correlation of 0.066 which is not statistically significant and amounts to no correlation.

	Awards	Run Time	Box Office
Awards	1.000000	0.123810	0.066667
Run Time	0.123810	1.000000	0.219048
Box Office	0.066667	0.219048	1.000000

Correlations table: There does seem to be very weak correlation between runtime and Box Office Sales.



Conclusions

I strongly recommend resist drawing strong conclusions from this very no representative data set. Given that we did not include any non winners in the dataset it is hard to have any confidence in mathematical relationships(or lack thereof) found in this data set.

Despite strong recommendations against drawing conclusions, we find this data to be provocative, and we recommend further study. We find the relationships being studied to be of interest. We recommend that future researchers look at the relationship between metacritic scores and winning awards