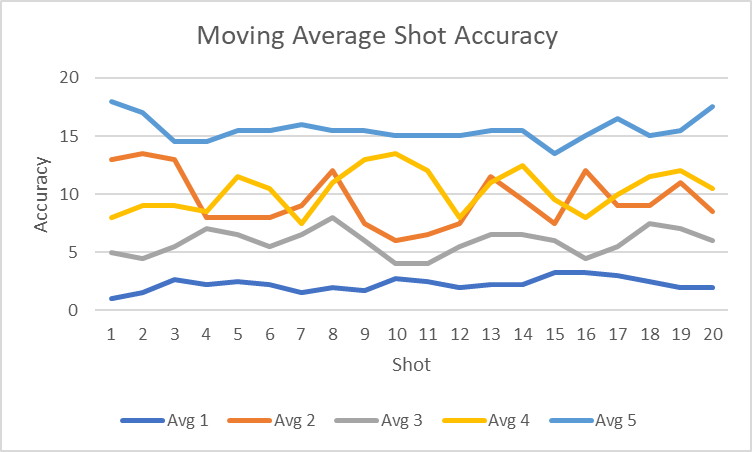
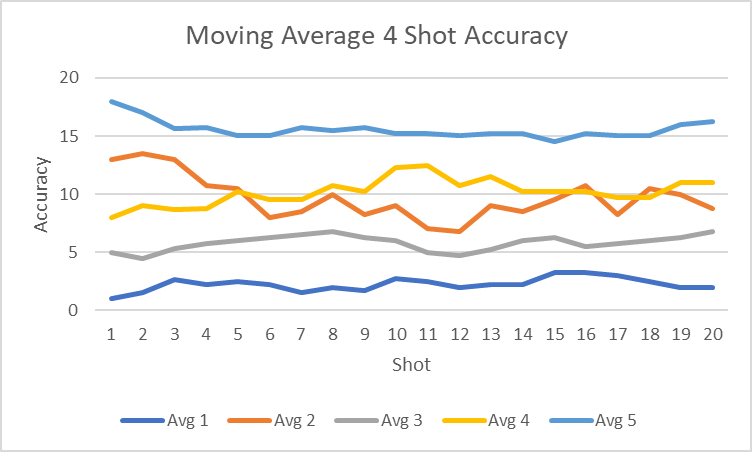
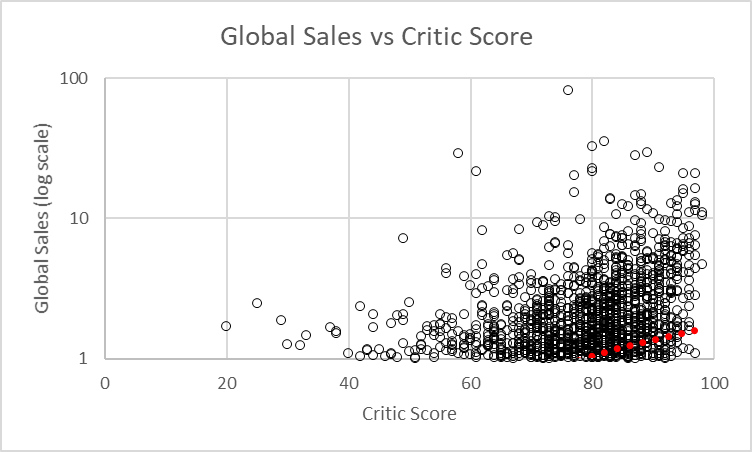
Will Seymour

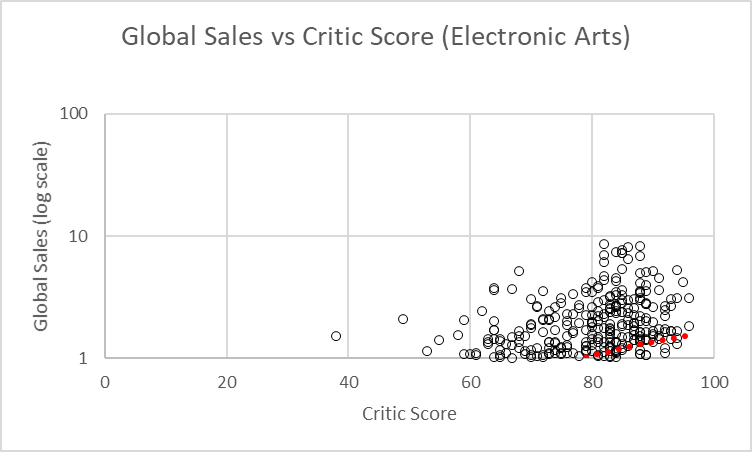
HW1

Part1.

1. 
2. 

Part 2. 1.



2. 

3.

4. After you have run regression, explain why your choice was or was not a good choice.

I think using minutes played was a good choice. There was a strong correlation of 0.9 between these two variables. There were other variables like field goals (FG), 2-point scores (2PT) and 3-point scores (3PT) but I didn’t want to chose them since they contribute directly to the player’s points.

\* Do the results have a good `R Square` value?

Yes, r2 = 0.82 indicating 82% of the variance observed in PTS is explained by minutes per game

\* Is your choice statistically reliable?

Yes, the p value was 3.5x10-222. Pretty low by most standards.

\* Explain what your coefficients mean.

|  |  |
| --- | --- |
|  | *Coefficients* |
| Intercept | -80.81635175 |
| Minutes Played | 0.509360016 |

The intercept is where the regression line crosses the x axis and represents the baseline number of point a player would be expected to score. In this case, since we were looking at the total minutes played for each team over a season, (most players played hundreds if not thousands of minutes and no players played zero minutes), the model is not accurate at the lower extreme. This is common in regression analysis.

The coefficient of Minutes Played (0.51) means that for every minute played a player is predicted to score 0.51 points.

\* Finally produce a scatter plot and explain your overall analysis.

Next page

