

NHL PLAYERS SALARIES PREDICTED... GOOD OR BAD VALUES?

By: Vincent Ditmore, Cory Swift, Bobby Mullins

Goal of this project:

- To make a model that could accurately predict the value of an NHL player during a chosen season.
- We then compare that value to their actual contract to determine whether the player was a actually good value or not.
- Doing this to see if each player is worth their salary.

Variables used in the Model

Response Variable:

→ Salary

Predictor Variables:

→ Age

→ Plus/Minus

→ Games Played

→ Goals

→ Assists

→ Shots

→ Penalty Minutes

→ Ice time per game

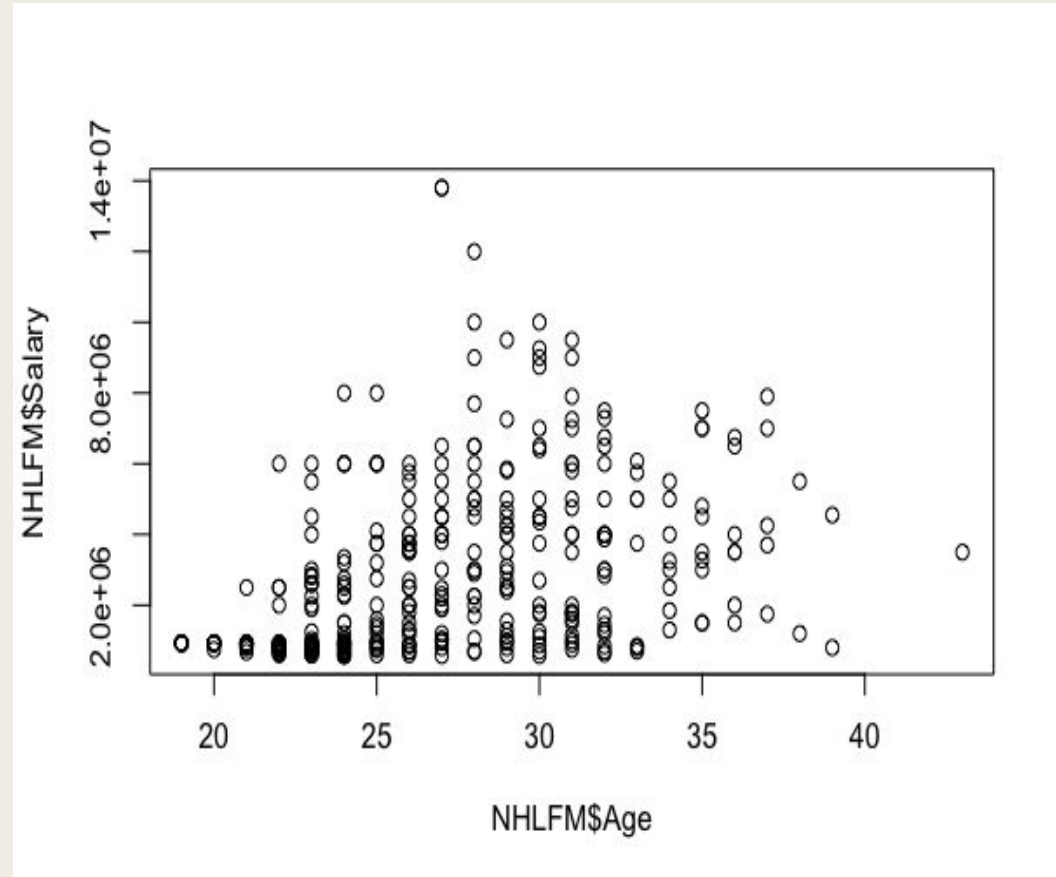
→ Points

Full Model (Starting Point)

- This model only works for Forwards since we used offensive based scoring stats.
-(defenseman/goalies excluded)
- $Fmodel = lm(\text{Salary} \sim \text{PlusMinus} + \text{Age} + \text{GamesPlayed} + \text{Goals} + \text{Assists} + \text{Shots} + \text{PenaltyMinutes} + \text{TOIPerGame} + \text{Points})$
- R^2 value of .4274

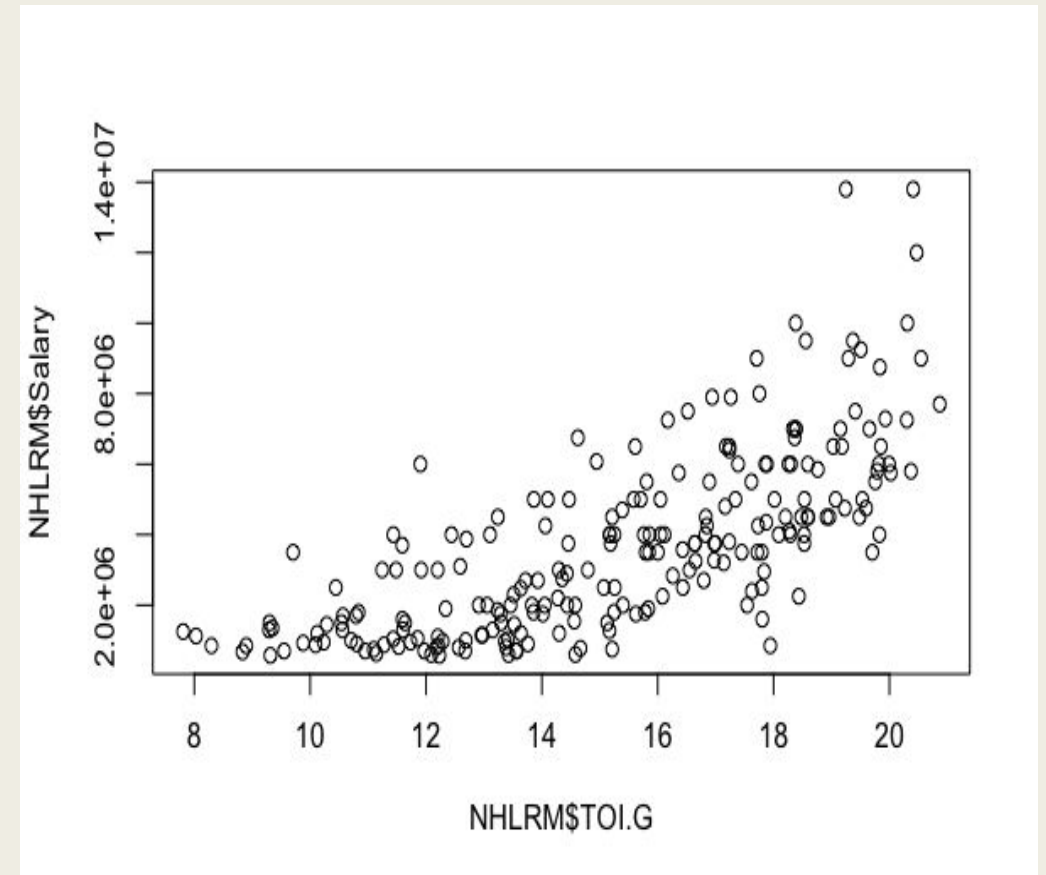
Problems Encountered: Age

- Age Range
 - *Entry Level Contracts*
 - *Constant Variance*
- Quick Fix:
(Cut under 24 & over 38)
- Remove Variable



Problems Encountered: TOI/G

- A little more tricky
 - *We wanted to distinguish the pay grade between different ice times*
 - *4th line, Top 9, Elite Players*
 - *Top 6 vs Bottom 6*



Problems Encountered: GP

- Negative Coefficient
 - *Why?*
- Boolean Variable
 - *Fail*
- Taken out of model
 - *Justified*

Coefficients:

| | Estimate |
|-----------------|----------|
| (Intercept) | 3255563 |
| PlusMinus3 | 11653 |
| GamesPlayed3 | -45228 |
| Goals3 | 23886 |
| Assists3 | 74284 |
| Shots3 | 11888 |
| PenaltyMinutes3 | -3204 |
| TOIPerGame3 | 889240 |

Problems Encountered: Weighting Goals and Assists

- We did not like how it was weighted
- Resulted to weighing them the same by using Points

Coefficients:

| | Estimate |
|-----------------|----------|
| (Intercept) | 560250 |
| PlusMinus6 | 2974 |
| Goals6 | 36788 |
| Assists6 | 77742 |
| Shots6 | 4986 |
| PenaltyMinutes6 | -4547 |
| TOIPerGame6 | 944628 |

Final Model:

- Interpreting Predictor Variables
- Multiple R-squared value
- Performance Based
 - *Defensive Forwards undervalued*
 - *Jonathan Toews*
 - *Salary6 ~ PlusMinus6 + Points6 + Shots6 + PenaltyMinutes6 + TOIPerGame6*
 -

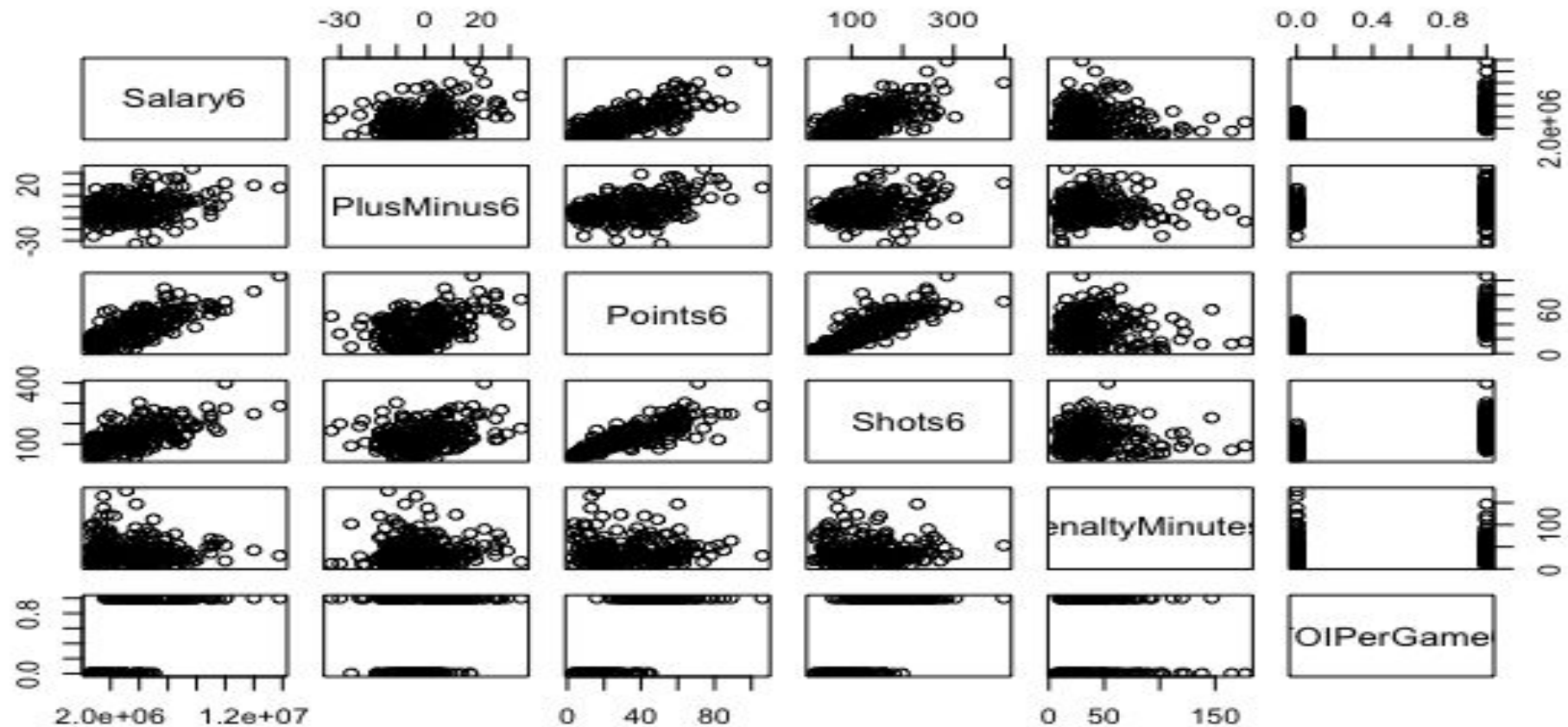
Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) | |
|-----------------|----------|------------|---------|----------|-----|
| (Intercept) | 662412 | 292805 | 2.262 | 0.02468 | * |
| PlusMinus6 | 1697 | 10649 | 0.159 | 0.87353 | |
| Points6 | 66363 | 11985 | 5.537 | 8.92e-08 | *** |
| Shots6 | 2638 | 2914 | 0.905 | 0.36637 | |
| PenaltyMinutes6 | -5018 | 3536 | -1.419 | 0.15733 | |
| TOIPerGame6 | 1007850 | 327614 | 3.076 | 0.00237 | ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

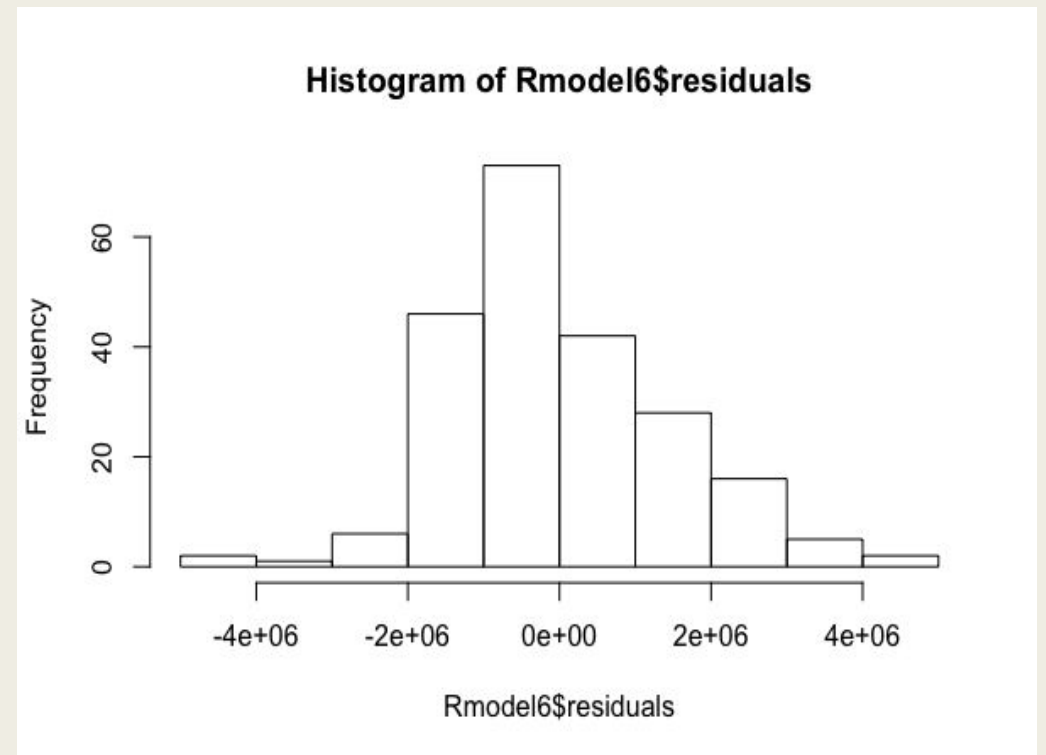
Residual standard error: 1496000 on 215 degrees of freedom
Multiple R-squared: 0.6281, Adjusted R-squared: 0.6195
F-statistic: 72.64 on 5 and 215 DF, p-value: < 2.2e-16

Checking Assumptions: Linear Model



Checking Assumptions: Error

- Error Independently and Identically distributed

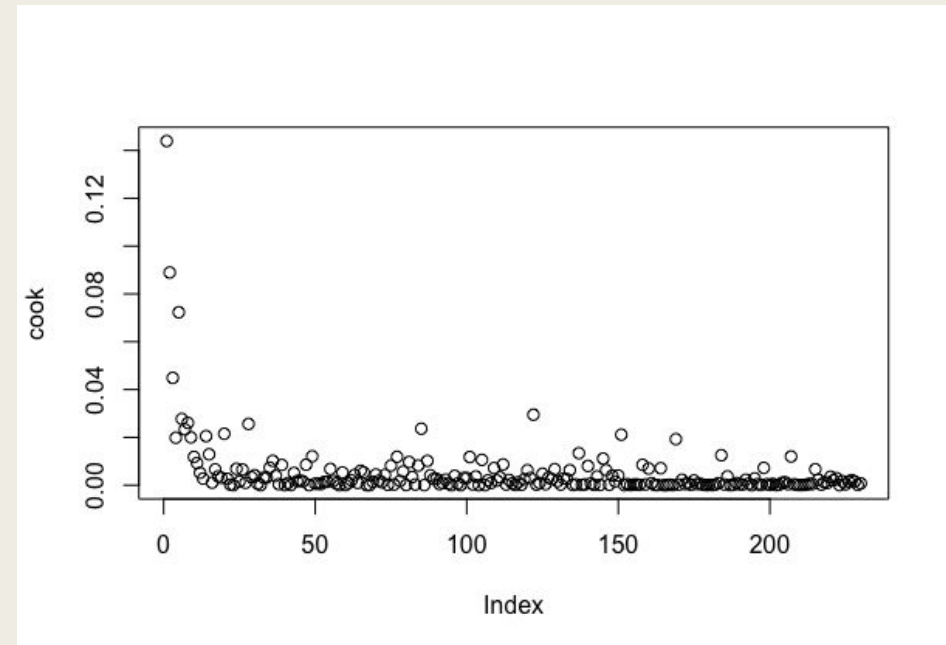
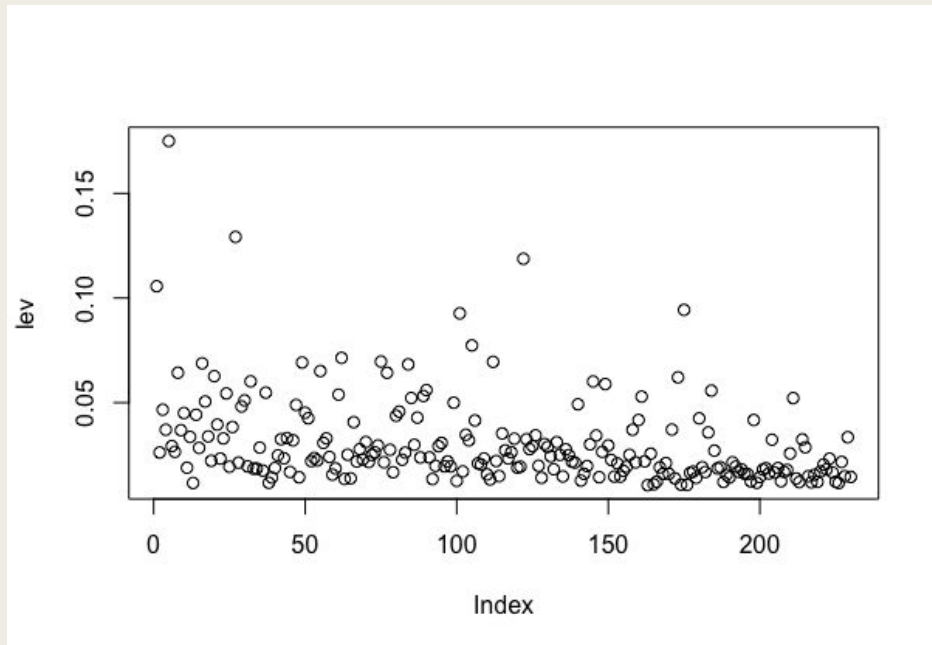


Checking Assumptions: Predictor Variables

- Obviously some collinearity as expected, however, not a problem
 - *Under 10*

```
> vif(Rmodel6)
      PlusMinus6      Points6      Shots6 PenaltyMinutes6      TOIPerGame6
      1.281219      5.752199      3.382735      1.004770      2.648284
>
```

Checking Assumptions: Reliable Observations



How We Determined Value

- We divided Predicted Salary over Actual Salary
- Bad Values $< 100\%$
- Good Values $> 100\%$

The Best Values

| | |
|-----------------|-----------|
| ■ Lee Stepniak | ■ 618.26% |
| ■ Kyle Palmieri | ■ 365.50% |
| ■ Mike Hoffman | ■ 306.77% |
| ■ Cam Atkinson | ■ 226.39% |

The Worst Values

| | |
|-------------------|----------|
| ■ Joffrey Lupul | ■ 26.42% |
| ■ Jonathan Toews* | ■ 41.36% |
| ■ Matt Moulson | ■ 37.67% |
| ■ Marian Gaborik | ■ 39.28% |

People you might know: Popular Players

| | |
|-------------------|-----------|
| ■ Patrick Kane | ■ 67.68% |
| ■ Sidney Crosby | ■ 64.89% |
| ■ Connor McDavid* | ■ 544.94% |