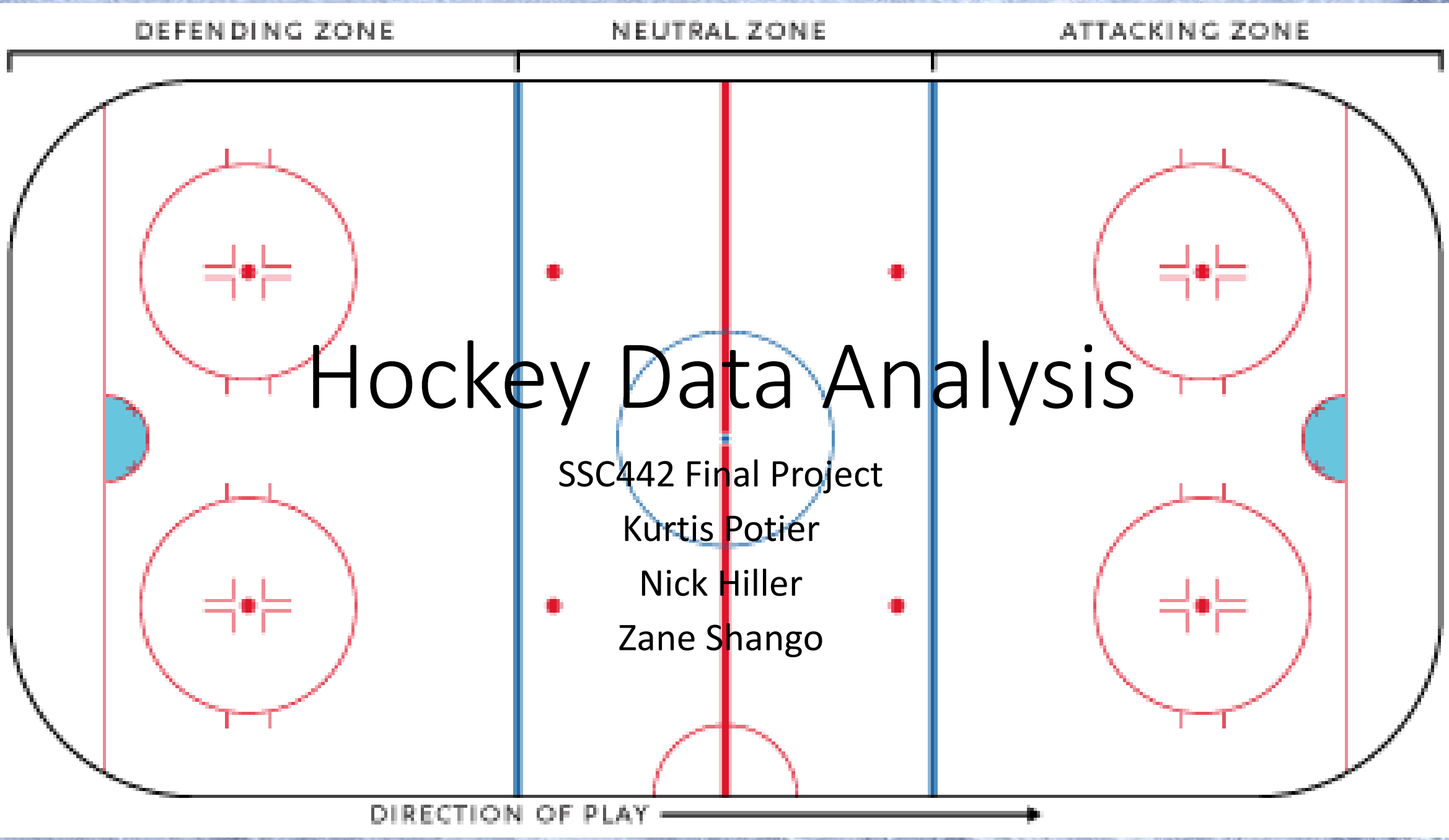


Hockey Data Analysis

SSC442 Final Project

Kurtis Potier

Nick Hiller

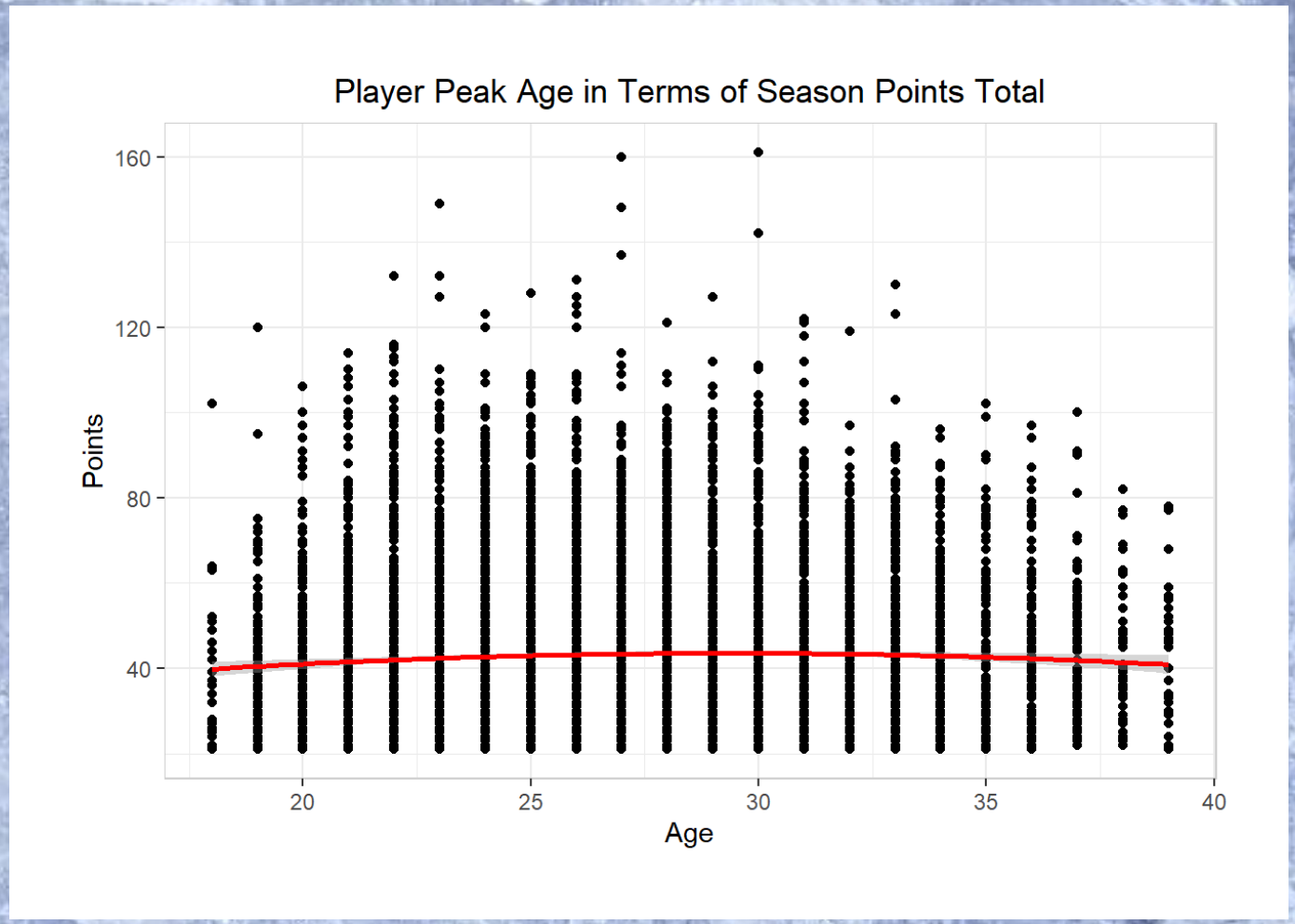
Zane Shango

Goals

- Answer useful general questions related to hockey player performance.
- Does tanking in the NHL lead to team success in the future?
- What is the value of a draft pick?

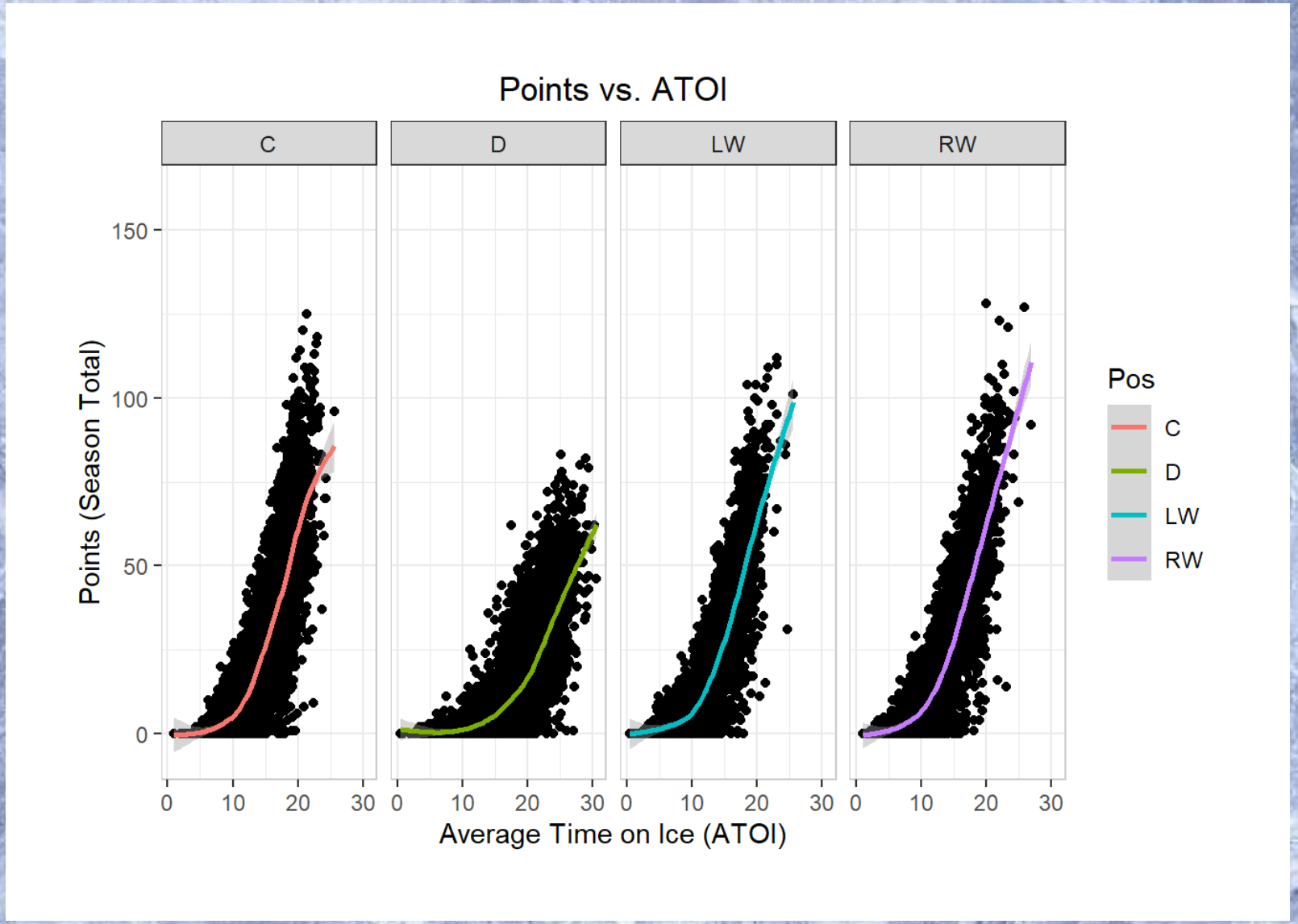
Peak Player Age Calculation

An often debated question in sports is whether to sign a free agent player. A key component of this is determining what that player's future production will be. To help determine this, points for a season were graphed against age. Additionally, a second degree polynomial was fitted to provide a numerical description of this curve. For this analysis, only players who had scored at least 20 points and were less than 40 years old were used in order to focus on career NHL players and filter out players that were not able to produce at the NHL level. After doing this, it was found that the peak age of a player is 29.



Points vs. Average Time on Ice

Points versus the average time on ice was analyzed to see if playing a player too much could be detrimental to their performance. The visualizations were separated by position as some positions play more than others.



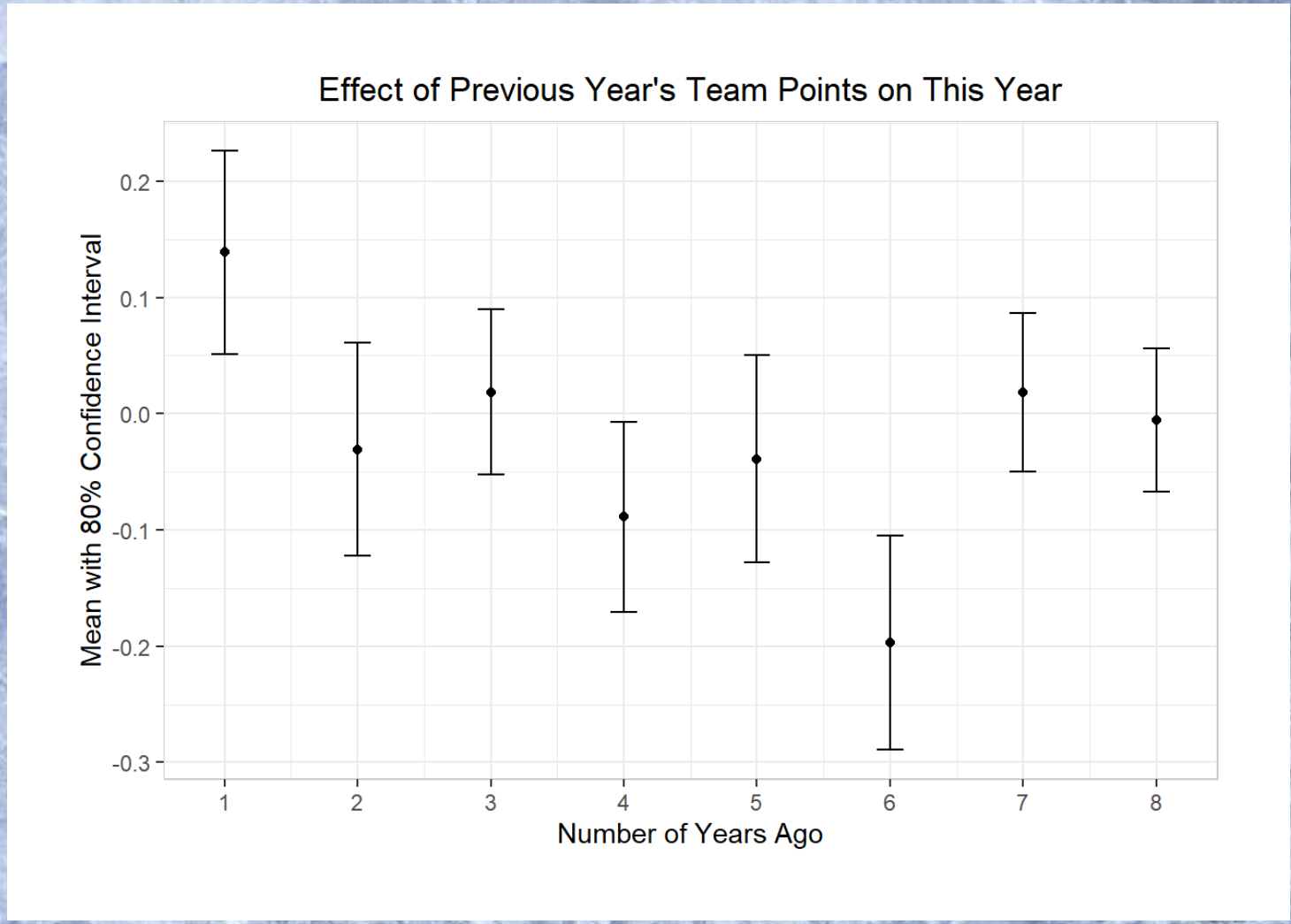
Does Tanking Work?

Tanking is the idea that a team can improve its future years performance by doing poorly this year. The reasoning behind this is that doing poorly results in a team receiving higher overall draft selections. A counterpoint to this is that a losing culture does not attract talented free agents who want to win a Stanley Cup and or compete in the playoffs. To determine if tanking is actually an effective way to build a team, a linear regression was performed on every NHL team over its last 30 years. This regression predicted team points using the team points obtained in each of the previous 8 seasons. The coefficients of these regressions were then averaged and an equation for predicted team points was determined:

$$Points = 102.931 + 0.139(lag1) - 0.031(lag2) + 0.019(lag3) - 0.089(lag4) - 0.039(lag5) - 0.197(lag6) + 0.0199(lag7) - 0.005(lag8)$$

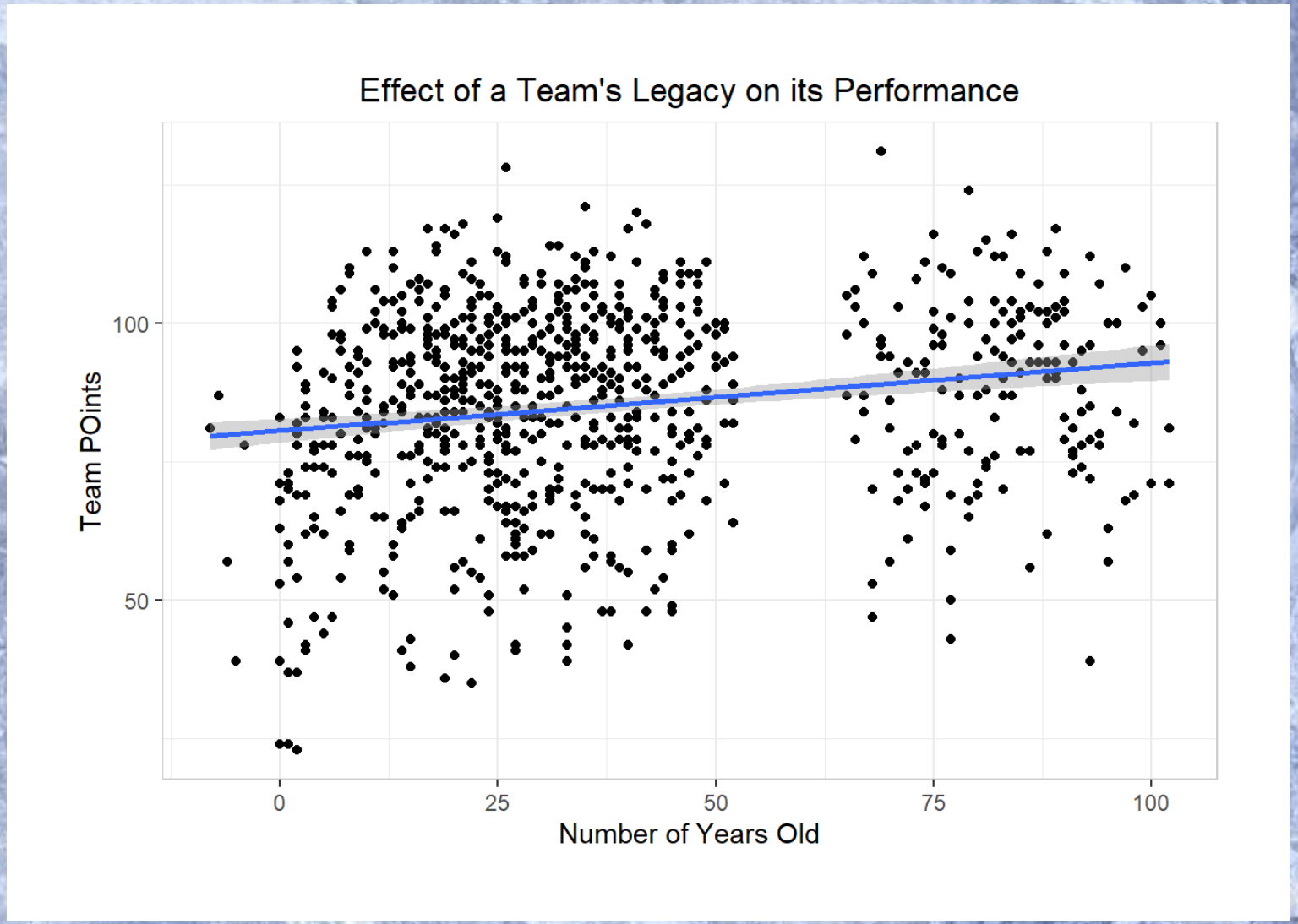
Does Tanking Work (continued)?

A visualization was then produced that shows the effect of each previous year on the current year's team points. Positive values show that it is helpful to do better in that year to increase points in the current year. Negative values show that it is helpful to do worse in that year to increase points in the current year. What the visualization shows is that tanking does work, but not immediately. A cycle of about 6 years seems to be what it takes for a "rebuilding" process that involves using tanking as a strategy. The 7th and 8th year had no significant effect and broke from the general trendline.



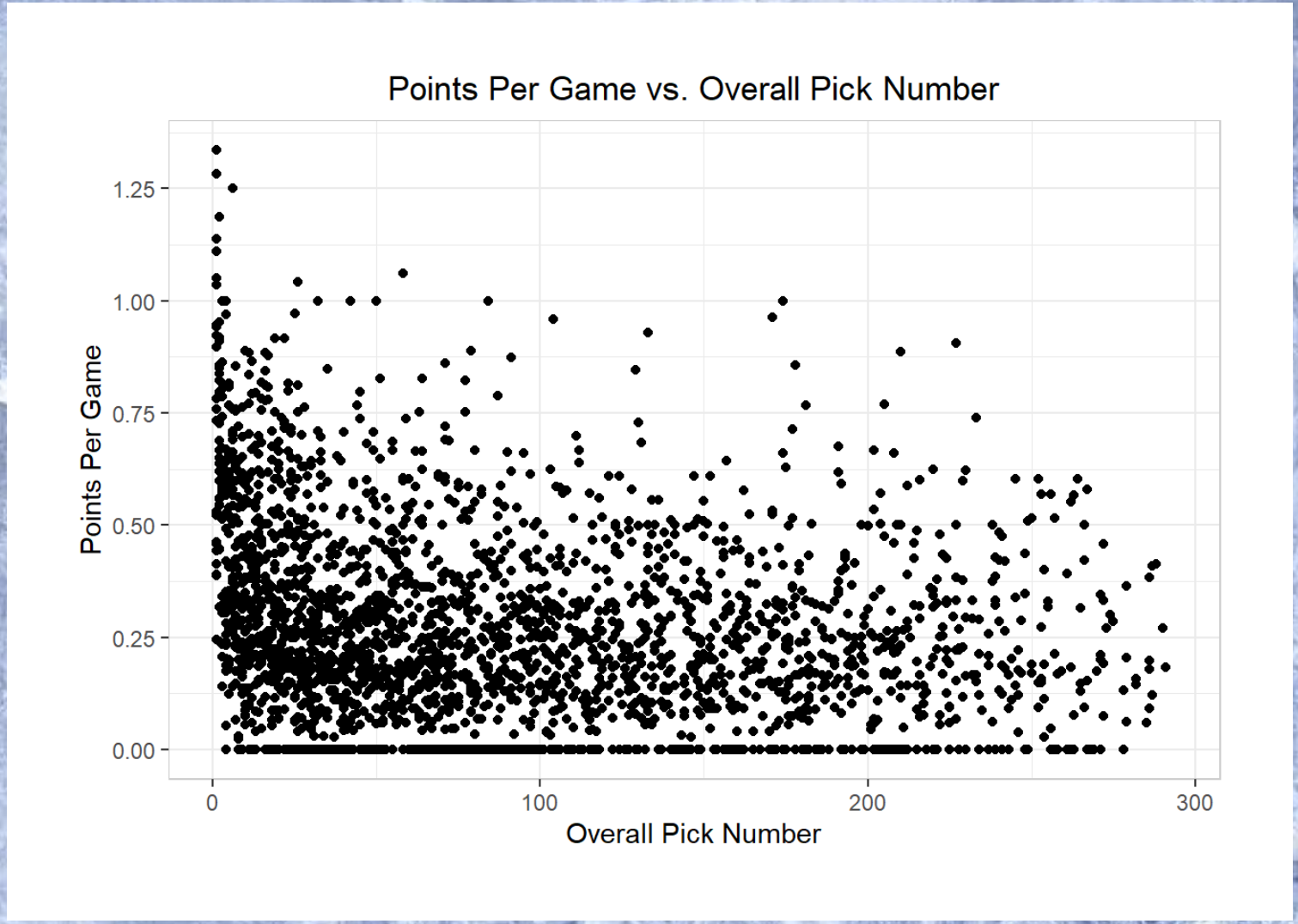
Does City Temperature and Organization Age Matter?

In order to determine if there were other effects that influence a team's performance, the temperature of the city in which the team plays and the number of years old a team is were used as predictors of team points. The results found that temperature was not a significant effect, but the number of years old was. This relationship was then visualized.



Predicting Skater Points Using Draft Position

In order to visualize the correlation between a player's overall draft pick spot and their points per game productivity, these variables were shown in a scatter plot. Additionally, a regression was ran that predicts the number of points a player will score per game with a two degree polynomial of their overall draft position and categorical variables of their position.



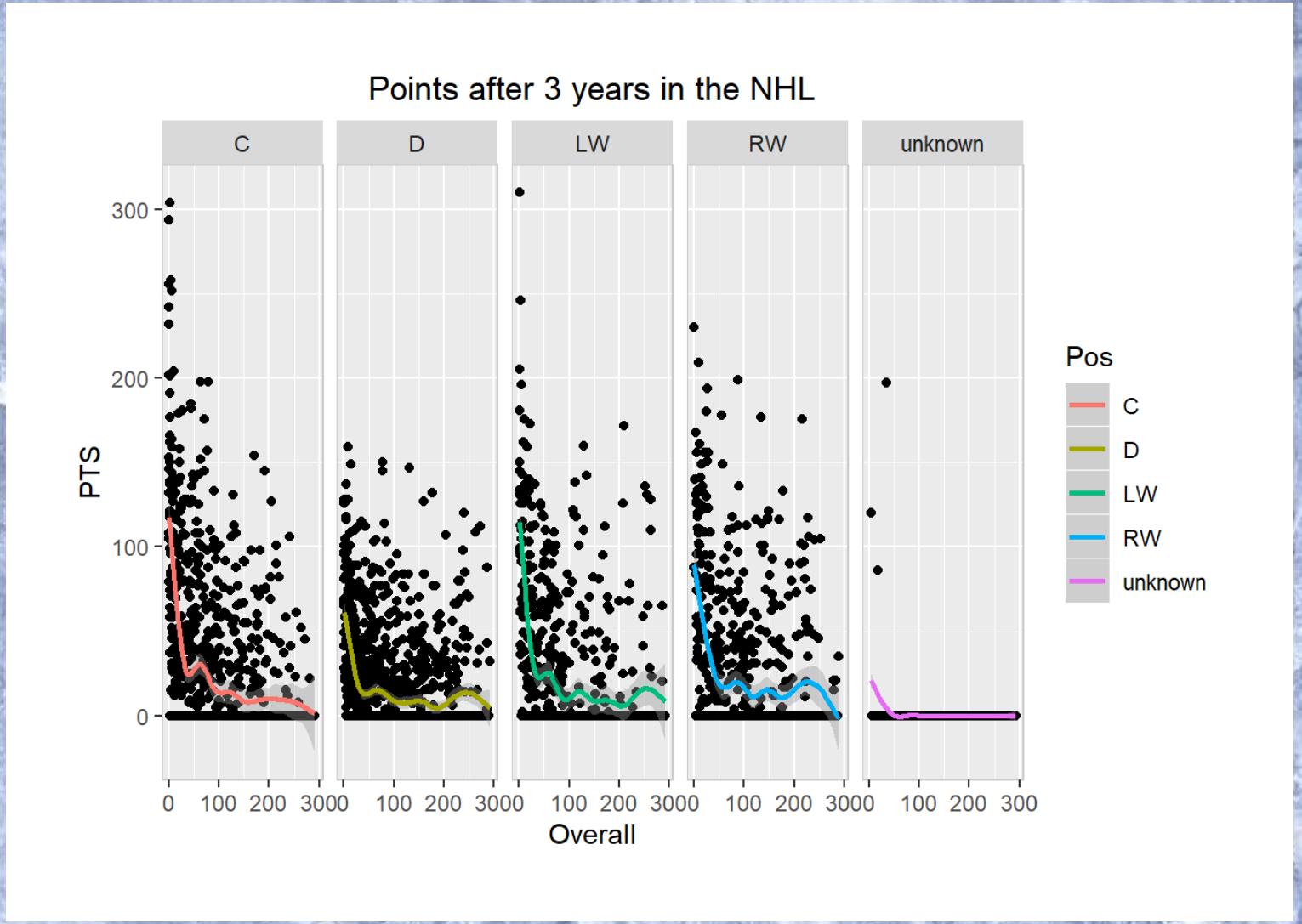
Prediction of Points in First Three Years in the NHL Based on Draft Position

Valuing draft picks is complicated. For this analysis it was chosen to look at their production during their entry level contracts. This contract activates once a player plays at least 9 games in an NHL season. A player can be kept in minor leagues for a couple years and then brough up to the NHL and the 3 years begins at this time. To see what production can be expected from a draft pick, a regression using the 4th degree polynomial of overall pick number was used as a predictor for points. A knn regression was also tried. Root mean squared error for both types of regressions was similar, so the results of the polynomial regression are shown.

$$Points = 83.71 - 2.236(overall) + 0.0239(overall)^2 - 0.0000991(overall)^3 + 0.00000149(overall)^4$$

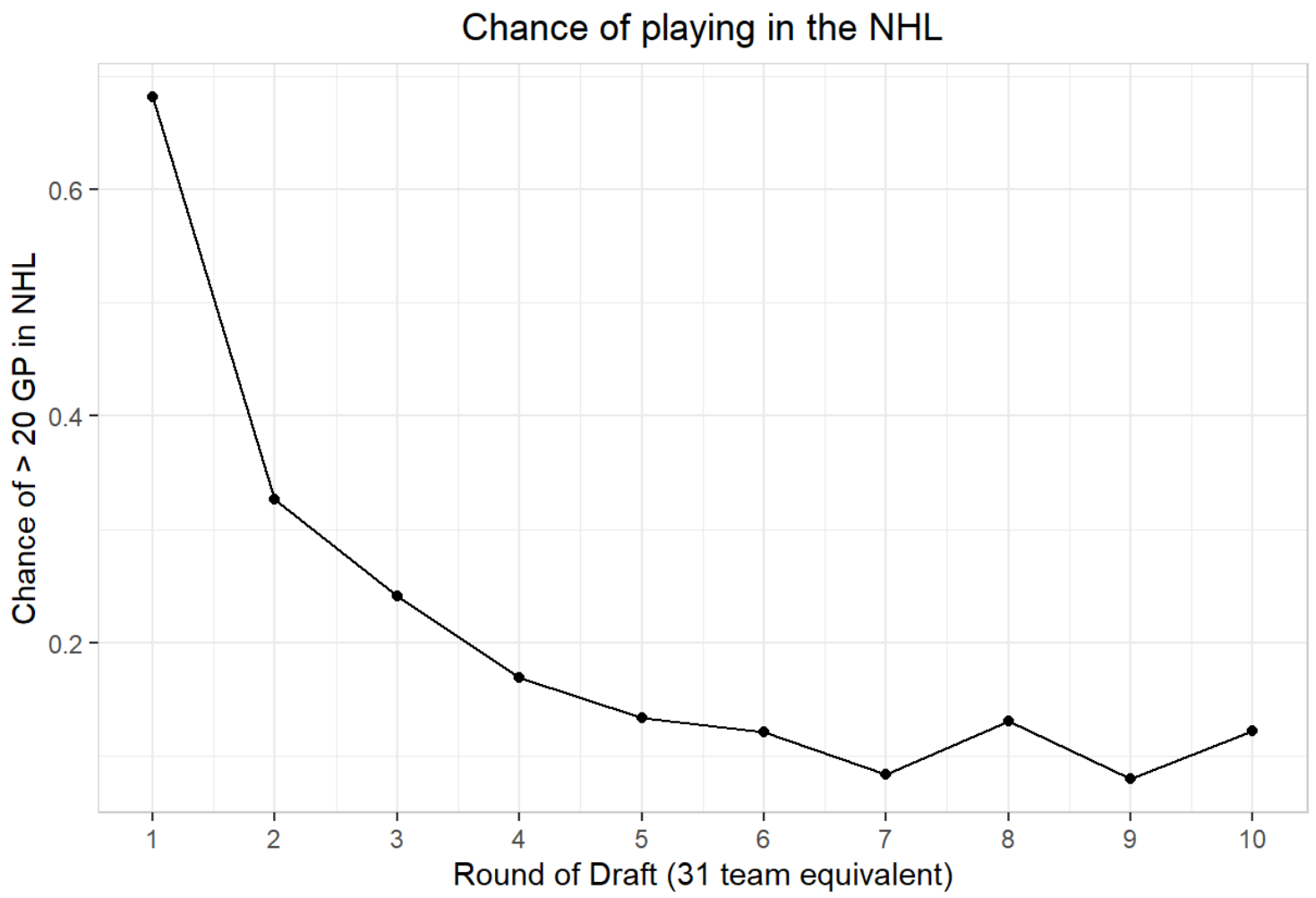
Prediction of Points in First Three Years in the NHL Based on Draft Position

This was then visualized by position. The unknown plot generally refers to a player that never made it to the NHL but does not include all players that did not make it to the NHL. For this reason, position was not used as a predictor here.



Probability of Making NHL by Draft Round

The probability of making the NHL based upon the round a player is drafted in. The threshold for this was playing at least 20 games in the NHL.



Conclusion

Overall, our analysis provided some useful general statistics about NHL player performance as well as showing that tanking can be a useful strategy to set a team up for future success. The expected production from a draft pick was determined.

References

- https://www.hockey-reference.com/
- http://www.nhl.com/stats/