How to Construct
A Winning Fantasy
Hockey Team

By Wil Moskal

Introduction

Fantasy Sports are a huge part of western culture. Over the past 10 years the number of fantasy sports users has practically tripled. In 2016, there were 57.4 million people in just the United states and Canada that participated in a fantasy sports league. Over 40 Million people (70%) of those who participated in fantasy sports had some sort of money going into their league. This paper will seek to explore how to construct a dominant fantasy hockey team, however, the process could be adapted to other fantasy sports. Whether you are just uber-competitive like me, you have money on the line, or you just want to be able to compete with some of your buddies, this paper seeks to show you how to win.

The Math

The Mathematics of Data Management will be used to deduce what the best draft strategy is. Specifically, the Mathematics of Single and double variable analysis will be used to Explain how to win.

Note*

Personally, my favorite stats to use for a fantasy hockey league are having 6 categories for Skater and 3 categories for goalies. The stats that I prefer to have for skaters are Goals(G), Assists(A), Plus-Minus(+/-),Penalties in minutes (PIM), shots on goal (SOG), and special teams points (STP). The stats that I prefer to have for goalies are Wins (W), Goals against average (GAA), and save percentage (SV %). The way I am analyzing these stats, however, can be used for any stats.

Constructing the Winning Team

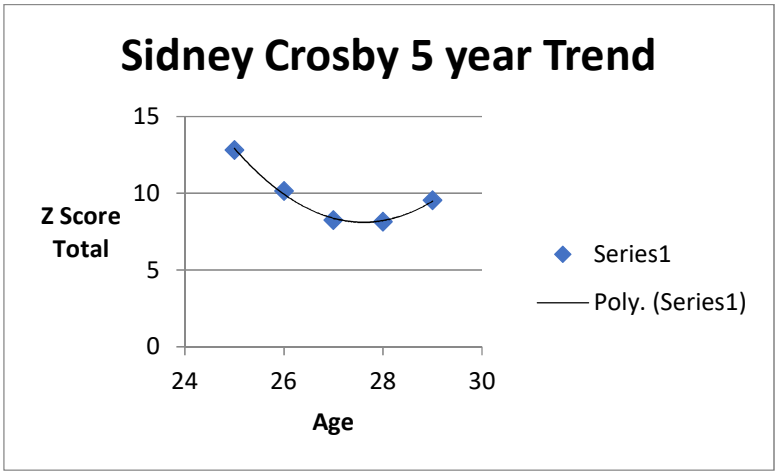
The First Step is Extraction. I took the top 100 players, with all their stats, at every position (50 for Goalie). Next, for each stat, for each position, I calculated the mean and standard deviation. With these numbers, I could know calculate the Z Scores for each player per stat. (A Z Score is a measure of how many standard deviations something is from the mean). Once this was completed for every player for every category, I could get a Z Score total for each player.
Next, I Calculated the Minimum, Maximum, Median, 1^{st} and 3^{rd} quartile values for each stat, and for the Z Score total. Now I was ready to move on to the next step. When you're in a fantasy league, your top pick can make or break your team. To make sure that you don't pick a bust on your top selection, some further analysis is necessary. I looked at the top 5 centers, however, this is because of a time constraint, and you could look at all positions. Of the top 5 centers, 4 had played 5 seasons or more, and 1 had only played 2 seasons. For the 4 top centers who had played in 5 or more seasons, I analyzed their 5-year Z Score trends, then created a scatter plot for each of them, and generated a trendline. Their Stats and graphs can be seen below

Sidney Crosby

Age	Z Total
25	12.83441171
26	10.16622059
27	8.260649379
28	8.145562695
29	9.545920287

$y = 0.7091x^2 - 39.152x + 548.53$

$R^2 = 0.9944$

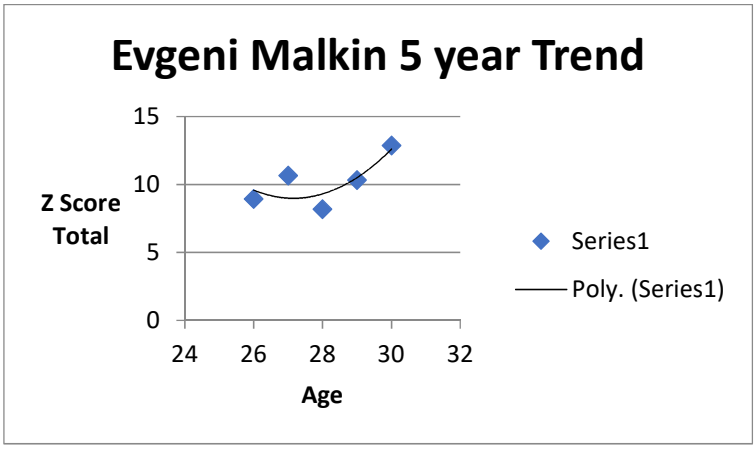


Evgeni Malkin

Age	Z Total
26	8.928754014
27	10.66538868
28	8.183364348
29	10.33927771
30	12.88084876

$y = 0.4463x^2 - 24.233x + 337.97$

$R^2 = 0.6509$

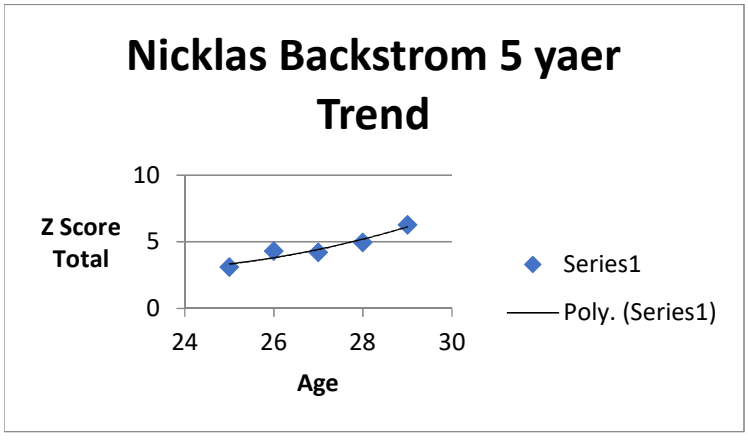


Nicklas Backstrom

Age	Z Total
25	3.112786894
26	4.30662094
27	4.21880491
28	4.951030958
29	6.267162495

$y = 0.076x^2 - 3.4111x + 41.083$

$R^2 = 0.9202$

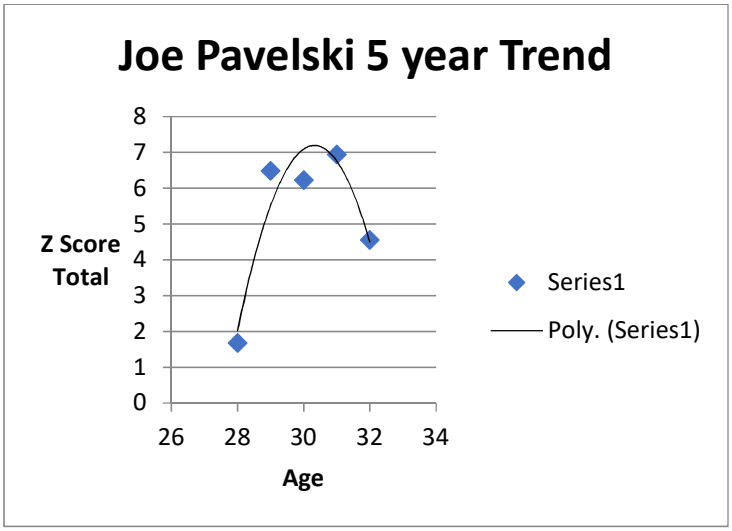


Joe Pavelski

Age	Z Total
28	1.677284854
29	6.481684697
30	6.22602669
31	6.94016933
32	4.554182083

$y = -0.9579x^2 + 58.097x - 873.68$

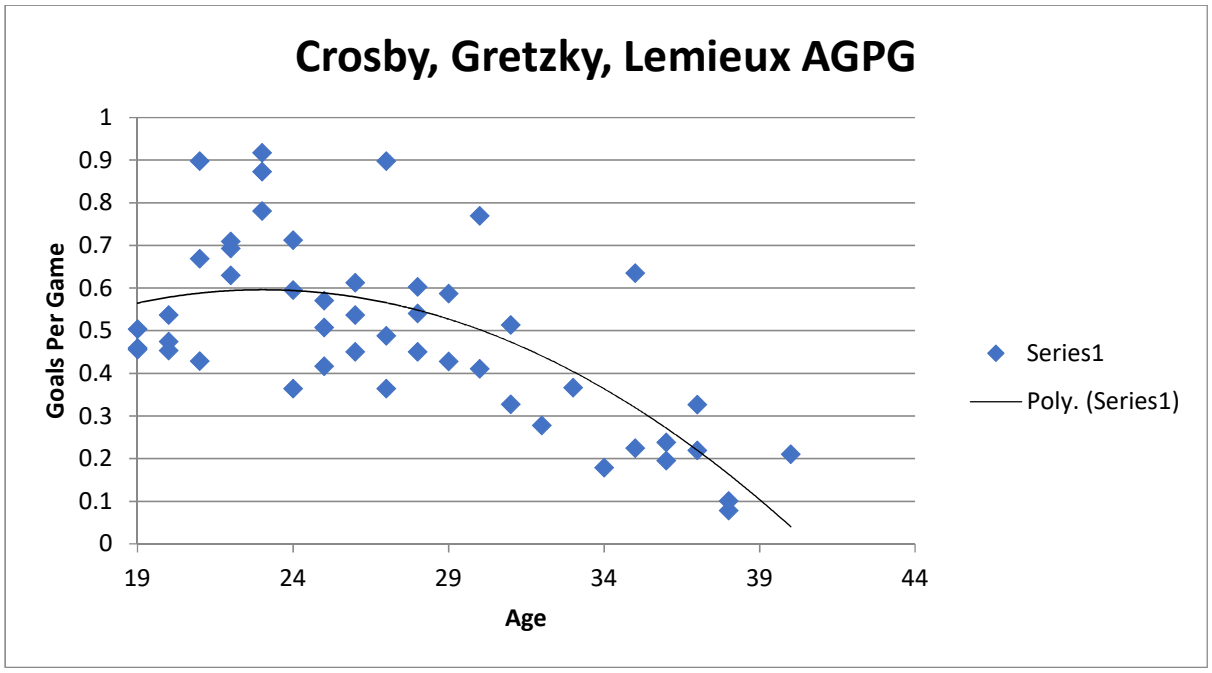
$R^2 = 0.9007$



By generating these graphs for each of the players, we can tell their approximate trajectory for next year, However, the trends will not always hold true. The R^2 value gives a representation of how good of a fit the trend line is. Overall, players generally follow a downward parabolic curve throughout their career. Because of this, we can tell that Sidney Crosby, Evgeni Malkin, and Nicklas Backstrom will not continue their current trajectory, even though they are trending upward now. Because we know this, we must consider if we think their current age is the age when they will start trending downward. Based on this analysis, we can also tell that joe Pavelski appears to be starting his downward trend, and is a player who would be risky to take. Connor McDavid, the other of the top 5 centers, has only played 2 seasons in the NHL. Because of this, there is not enough NHL data to create a trendline, and Other ways of projecting his trend for next year must be used. Connor McDavid has been described as a generational talent, and has been Compared to Sidney Crosby, and even Wayne Gretzky. Because of this, we can try to analyze the careers of Sidney Crosby, Wayne Gretzky, and Mario Lemieux to try to determine the Trend that Connor McDavid will take next year (Because of time restrictions, I only calculated trends for Goals per Game and Assists per Game). When Analyzing Wayne Gretzky and Mario Lemieux, I Adjusted their stats because Scoring happened at a much higher rate than it does now. I adjusted by multiplying their goals and Assists per game by a factor of 0.8. Once I had these numbers calculated, I created scatter plots for each player, and an overall Scatter Plot. Below are the Scatter plots for Adjusted Goals Per Game, and Adjusted Assists Per Game.

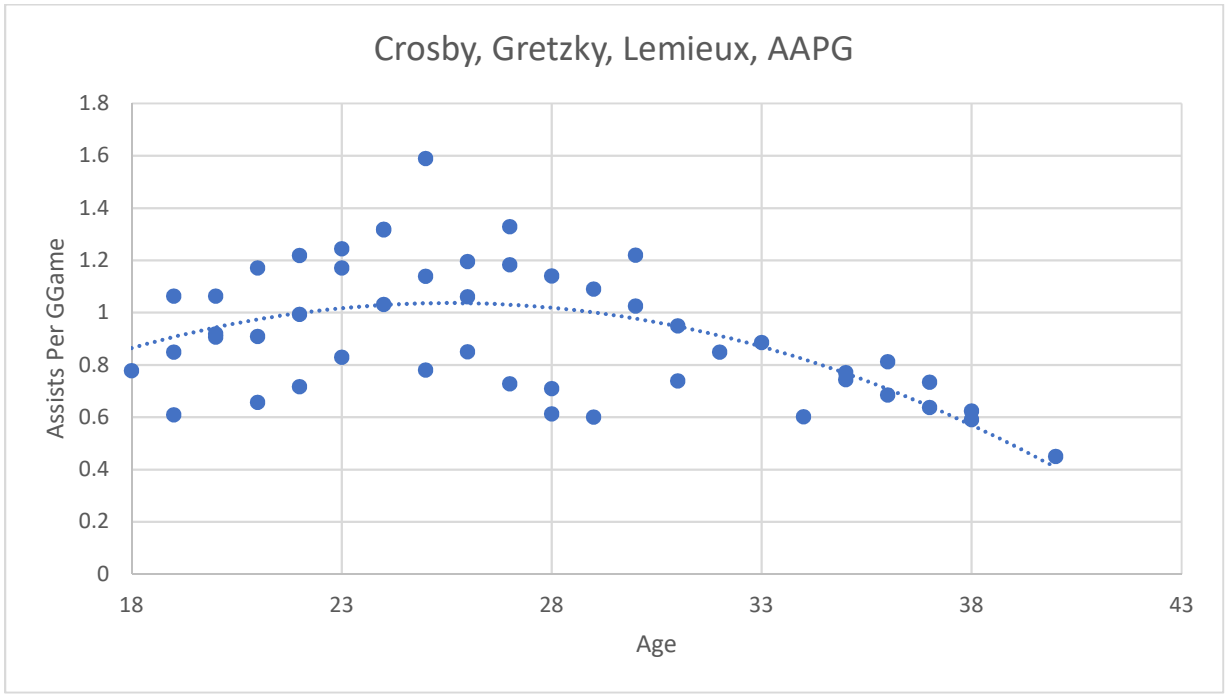
Trend Of Superstars Goal Per Game

$y = -0.0019x^2 + 0.0888x - 0.4263$
$R^2 = 0.4839$



Trends Of Superstars Assists Per Game

$y = -0.003x^2 + 0.1541x - 0.9332$
$R^2 = 0.3538$



Obviously, Connor McDavid will not exactly follow the trends of these players. For one thing, The Adjusted Stats May be off. Also, the players that they are playing with are different, and the players skill levels are different. However, the trend of Connor McDavid, likely will be like that of These three superstars. From the graphs, you can tell that the overall trend is that they increase for a small amount of time, then they start to decline. This is likely due to other players learning how to defend against them. For this reason, Connor McDavid is likely to still be increasing, and would be an excellent selection for your team.

Extrapolation

Using the equation of the line of the curve of best fit, we can Extrapolate as to what Connor McDavid's goals and Assists per game will be.

Goals: $-0.0019*(20)^2 + 0.0888*20 - 0.4263$

Projected goals per game: 0.5897

Assists: $y = -0.003*(20)^2 + 0.1541*20 - 0.9332$

Projected Assists per Game: 0.9488

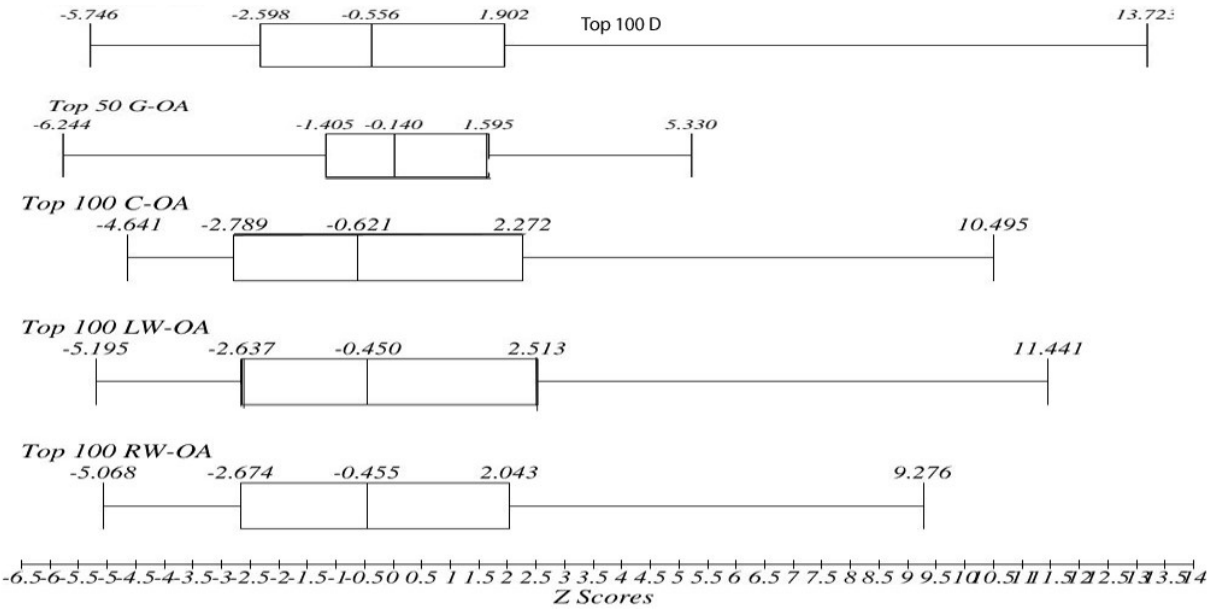
IF Connor McDavid plays 82 games, then he is projected to score 48 goals, and projected to get 78 assists. However, as mentioned before, the correlation is not an exact fit, and this is not likely to happen as there is simply less scoring now.

The Verdict

After analyzing the top Centers (Due to time constraint, in actual application, more positions should be considered), I have deduced that selecting any of Sidney Crosby, Evgeni Malkin, or Connor McDavid would be best for your team. The reason being that Connor McDavid is still trending upward, even after his phenomenal season this previous year. Evgeni Malkin will help you in all 5 categories, and had the highest Z total of any Center. Sidney Crosby has long been claimed the best player in the world, Barring injury, he is set to have another great year. The reason Nicklas Backstrom was not recommended was because, while he is great, he is a tier bellow Crosby, Malkin and McDavid. The Reason joe Pavelski was not recommended was because he appears to be on the downward part of his career.

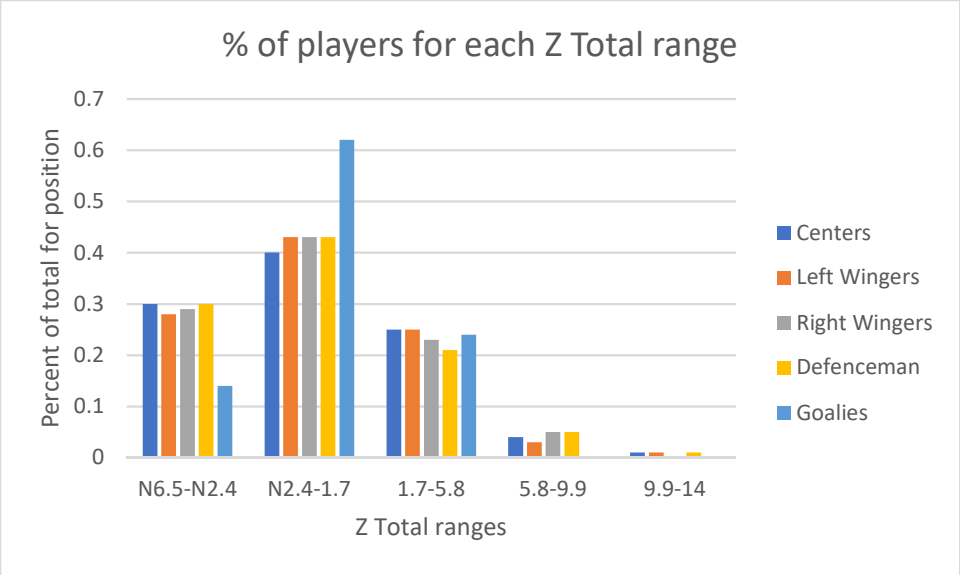
After the First Round

While the first round can be the most important pick, it is also important to pick good players in later rounds too. Based on the calculations of Minimum, Maximum, Median, 1^{st} and 3^{rd} Quartile, I created a box and whisker plot, which is a visual representation of the spread of the data.



By analyzing this graph, we can see that for defenseman, there is a really big difference between the top couple and the median. For Goalies, we can tell that there is a much smaller difference between the best couple and the middle of the pack. For Centers, there is a fairly large difference between the top centers and the middle centers, however, between the median and 3rd quartile, there is a large selection of players. We can also tell that the worst Centers are close to the 1st quartile centers. For Left wingers, there is a very large gap between the top couple and the 3rd quartile. For right wingers, there is a relatively small difference between the best and the 3rd quartile. Taking all this knowledge into account, we can understand the importance of each position. For Defenseman, it is likely worth it to take one with one of your top picks, because the talent will drop off quickly. For goalies, it is not essential to pick one with one of your top picks, but you do not want to wait until the last couple goalies. For Centers, left wingers and right wingers, it is beneficial to take them early, but it is not bad if you don't, because the middle is bunched relatively tightly.

In addition to selecting which position it is best to take early, you should also look at the Z totals of each player, and select players who are listed low, but have Z Score totals relatively high to the players around them.



*Note, This is a histogram, I recognize that there is gap, even though it is a continuous data set. This was done for readability, but the data set *is* continuous.

This Histogram shows that There is no elite goalies, and no elite right wingers, so generally you should not be looking to take one of these positions with your top picks. It also shows that most of the goalies are in the median Z total range, with only a few being in the top, so it is important to take a goalie early, though generally not 1st.

Extraneous Variables

Some of the extraneous variables that there may be with this project include, but are not limited to, Player Injury, players team (strength of team), Off ice things (player harassment charges, players' wives giving birth), incorrect adjusting of stats, and Human Error. All extraneous variables must be considered, even though it may not be possible to compensate for all of them.

Future Exploration

Given more time, I would consider more than just the top Centers, go more into depth on analysis of players in later rounds, and more time analyzing the trend most likely for Connor McDavid's career. I would Also look more into when players of each position should be drafted. Finally, in the future I would like to analyze who the ideal pick is for each round.

Works Cited

http://fsta.org/research/industry-demographics/

www.fantrax.com