

THE GREAT ONE IS BORN: WAYNE GRETZKY'S MONUMENTAL SEASON

Tyler Ison

Submitted to the faculty of the University Graduate School  
in partial fulfillment of the requirements  
for the degree  
Master of Science  
in the Richard M. Fairbanks School of Biostatistics,  
Indiana University

May 2019

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

Master's Thesis Committee

---

William Fadel, PhD, Chair

---

Spencer Lourens, PhD

---

Ying Zhang, PhD

© 2019

Tyler Ison

## DEDICATION

To my parents (Scott and Holly Ison) and my siblings (Victoria and Elizabeth).

## ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my advisor Dr. William Fadel for all of the guidance and support he provided throughout my project. I am truly grateful for all the help and instruction that you have provided me over the course of this work. I would also like to thank Dr. Spencer Lourens for the time and help with the data scraping that he attributed to me and my project. I would also like to thank Dr. John Boardman for all his encouragement, support and guidance throughout my time spent at Franklin College.

Finally, I would like to thank my wonderful parents, Scott and Holly Ison, and my siblings, Victoria and Elizabeth, for their support throughout my entire academic career, and for their patience with me as I have been completing my degree. I could not have reached this point in my life without the guidance and support from all of you.

Tyler Ison

## THE GREAT ONE IS BORN: WAYNE GRETZKY'S MONUMENTAL SEASON

Statistics and athletic sports have always had a strong connection that many critics, fans and statisticians utilize to determine how successful a team or an individual player might be over an entire season or even throughout one's career. The success of a player or team is often characterized by investigating the consistency that has been shown throughout the season or career, which has led to more investigation of the streakiness of players. Studies have been done to examine great streaks, such as Joe DiMaggio's 56-game hitting streak or Tiger Woods' 142 consecutive cuts made streak, but what about the outstanding streak that occurred during the 1983-1984 NHL season? Wayne Gretzky, of the Edmonton Oilers, managed to showcase just how elite he was as a playmaker during that season. Gretzky produced a remarkable 51-game point streak, in which he recorded at least one goal or point in 51 consecutive games; a streak that has not received the recognition that it deserves. Using game-by-game data for the entire 1983-1984 NHL season for all players, the research looks at not only the evaluation of Gretzky's streak, but also compares his production and streak to the remainder of the league. Gretzky demonstrated why he is one of the greatest players to ever step foot on the ice, and his elite status is shown throughout this analysis. Comparing Gretzky's streak to that of DiMaggio's was shown to be a little challenging but, some general conclusions were made based on the comparison of analyses that were performed; but without the proper statistics being readily available, it is hard to adequately dictate which streak is ultimately more impressive or more rare.

William Fadel, PhD, Chair

## TABLE OF CONTENTS

List of Tables .....	viii
List of Figures .....	ix
List of Abbreviations .....	x
Introduction.....	1
Methods.....	5
Obtaining the Data .....	5
Analysis.....	5
Results.....	10
Conclusion .....	17
Appendix A.....	20
References.....	26
Curriculum Vitae	

## LIST OF TABLES

Table A.1: Wayne Gretzky's Game-by-Game Breakdown .....	20
Table A.2: Point Streaks Obtained by Wayne Gretzky .....	21
Table A.3: Scoring Droughts Obtained by Wayne Gretzky .....	22
Table A.4: Joe DiMaggio's Game-by-Game Breakdown.....	23
Table A.5: Hitting Streaks Obtained by Joe DiMaggio.....	24
Table A.6: Hitting Slumps Obtained by Joe DiMaggio.....	25



## LIST OF FIGURES

Figure 1: Wayne Gretzky Moving Average.....	10
Figure 2: Joe DiMaggio Moving Average .....	11
Figure 3: Simulations of Wayne Gretzky's 51-Game Point Streak.....	12
Figure 4: Simulations of Joe DiMaggio's 56-Game Hitting Streak .....	13
Figure 5: Distribution of Point Streaks Obtained for all NHL Players.....	14
Figure 6: Total Points Accumulated for all NHL Players.....	14
Figure 7: Point Streaks Obtained for all NHL Players .....	15
Figure 8: Distribution of Point Streak Lengths Obtained for 10,000 Simulations for Wayne Gretzky's 1983-1984 Season.....	16
Figure 9: Distribution of Hitting Streak Lengths Obtained for 10,000 Simulations for Joe DiMaggio's 1941 Season.....	16

LIST OF ABBREVIATIONS

NHL

National Hockey League

MLB

Major League Baseball

## INTRODUCTION

Athletic sports have always had a fascinating connection with statistics, regardless of the sport. Different sports utilize different statistics to determine the likelihoods of different outcomes; as well as to evaluate the effectiveness and efficiency of individual player or team performances. Sports fans and statisticians tend to utilize sports specific statistics to dictate the success of a team or individual player over a season or their entire career; often comparing players across the same season, as well as different generations. Player and team success are often considered by investigating streakiness, also referred to as consistency, to determine overall efficiency for an entire season, career, or length of time. Streaks are difficult to measure, because there are no rules defining what constitutes a meaningful streak; in certain ways, a streak can be related to a degree of variation that exists in the data. However, streaks are important to show that binary outcomes in sports can, on occasion, defy expectation (1).

Individual player performance and player streaks have more commonly been investigated in sports, such as baseball and golf. One of the most common sport streak investigations is Joe DiMaggio's 56-game hitting streak, that he obtained during the 1941 MLB season. Many fans, experts, and critics believe this is a feat that will never be beaten. Some believe that it should have never happened at all (2). Numerous studies involving simulations have been conducted to test the probability or likelihood of a hitting streak of this length to be replicated, exceeded, or miraculously occur. Statistics have been used to investigate several other streaks within the game of baseball over the years. These include: Ted Williams' 84-game on-base streak in 1949, Joe Sewell's 115-

game streak of not striking out in 1929, and the 8-game hitting streak of hitting at least one home run in each game shared by three players (3).

Along with baseball, streaks and player performances have been compared within the game of golf. One of the most incredible being Tiger Woods' 142 consecutive cuts made streak from 1998 to 2005. This streak may be considered the most difficult and most impressive in all of sports by many mathematicians, experts, and fans of the game due to the challenges that golfers face. Golf is a sport with a much different aspect; it is a game where each golfer is competing against more than one hundred additional golfers in any given tournament. Following the first two rounds, which consists of 36-holes, roughly half of the golfers are cut and sent home from the tournament, so golfer consistency is critical. For Tiger Woods to have such consistency to continuously make cut after cut is impressive, but for Woods to make 142 consecutive cuts is remarkable. Not only was Tiger Woods simply making the cuts during his streak, he was winning tournaments or finishing high up on the leaderboard week after week.

Both baseball and golf include some of the most remarkable streaks in the entire sports industry, but there is one sport that includes a specific streak that has not received the recognition and investigation that it deserves. In 1983, Wayne Gretzky of the Edmonton Oilers showcased his talent and demonstrated why he is one of the greatest centers to ever play the game of hockey. Gretzky had a 51-game point streak during the 1983-1984 season, where he recorded either a goal or an assist in each game. Gretzky's streak is of great interest and will be thoroughly investigated both individually and compared to the rest of the players in the National Hockey League for the 1983-1984 season.

Ice hockey is a “stick-and-ball” game that developed from field hockey in the 1870s (4). It is considered the fastest paced and most violent team sport in the entire sports industry (5). It is a game that is often characterized by high intensity intermittent skating filled with quick changes in tempo and duration with excessive amounts of body contact (6). Each game is broken into three 20-minute periods, where players rotate on and off the ice in terms of shifts. These shift lengths vary depending on the player or game situation. Each team battles for puck control while trying to generate offensive advantages that lead to goals, and ultimately win the game.

Hockey uses specific statistics known as Corsi and Fenwick statistics to determine how successful teams are, along with how efficient individual players are, throughout the season or in a given game. Jim Corsi, a former goaltending coach for the Buffalo Sabres, introduced the most advanced statistic in all of hockey (7). The Corsi statistic refers to all shots that are directed towards the goal regardless of whether it enters the net, gets blocked by the goalie or another skater, or misses the net completely (8). This advanced statistic is used to determine the impact that an individual player has on their team in any given game. The Fenwick statistic, created by Battle of Alberta writer Matt Fenwick, does not take into account blocked shots, simply because many believe blocking shots is a skill, rather than a series of random events (7). Unfortunately, these advanced statistics were not recorded until the early 2000s, thus resulting in the use of certain assumptions for our analysis.

In 1983, one of the most remarkable accomplishments in the sports industry took place. Wayne Gretzky, of the Edmonton Oilers, scored a goal or recorded an assist in 51 consecutive games; a task that has yet to be matched. A point streak is commonly

defined as the number of consecutive games in which a player receives at least one goal scored or assist. Unlike many other streaks in sports, this streak lacks the recognition and investigation that it deserves. For the analysis of Gretzky's streak, the current research will be looking at the entire 1983-1984 National Hockey League game-by-game data for each individual player and team obtained from NHL.com (9).

Within this research, we will investigate the change in Gretzky's scoring average over shorter lengths, as well as comparing his streak to the rest of the league and see how it relates. Due to the lack of advanced statistics being readily available, some assumptions must be made in order to perform a more in-depth analysis of interest. One major assumption is made when calculating the scoring average for Gretzky and every other skater in the League for the entire 1983-1984 season, which is that every assist that a player received were the only passes into the scoring zone that each player made. Due to the lack of shooting statistics and time on ice for each player, we cannot adequately calculate the Corsi or Fenwick statistics. Instead, research will rely on the assumption that is being made to perform an analysis.

Analyzing the streakiness of a player, whether it be a positive streak or a slumping streak, can be done in two ways: find the longest slump or streak for a given player or run a simulation to predict a streak (10). Current research will investigate both of these methods in this paper to evaluate both the likelihood of this streak occurring and the comparison of Wayne Gretzky's outstanding 51-game point streak during the 1983-1984 NHL season.

## METHODS

### **Obtaining the Data**

For the analysis of Gretzky's streak, the data were obtained from NHL.com with the following criteria: game by game data from October 01, 1983 to April 30, 1984 for all skaters (9). This criterion provided the entire game by game statistics for each player that participated in any game in the 1983-1984 season. The data were scraped from the internet using the package RSelenium in R. The game by game data for Joe DiMaggio's 1941 MLB season was obtained manually from Baseball-Reference.com (11) and used for the analysis and comparison of DiMaggio's 56-game hitting streak to Gretzky's 51-game point streak. A hitting streak is commonly described as the number of consecutive games in which a player receives at least one hit (12).

### **Analysis**

Prior to advanced statistics being recorded by sports team statisticians, certain statistics were not recorded – shots blocked, pass attempts, time on ice, shifts, and more. Due to the lack of available statistics, one major assumption was needed in order to calculate the scoring averages for Wayne Gretzky and every other player in the league. Due to the lack of available data, the general assumption was made that each assist that was encountered was off of all of the passes that were made into the scoring zone; thus we simply add the total number of assists received to the total number of shots taken to obtain the total number of attempts, or opportunities, to score a point. This became the denominator value that was used to calculate the scoring average for each player where the total number of points was divided by the total number of attempts. This is not a fully accurate assumption due to the cause of inflating the scoring average since not every pass

is resulting in an assist. Underestimating the number of passing attempts results in a smaller denominator value than what we would actually expect to see, but with the limitations of the recorded data at the time of this streak, it is one that can be made to evaluate each player's efficiency as a scorer or playmaker for their team since a player can receive a point by either scoring a goal or passing the puck to a player that scores a goal, or receiving an assist. Based on this assumption, the scoring average for Gretzky and all other players was calculated using equation 1.

$$\text{Scoring Average} = \frac{\text{Total Points Scored}}{\text{Total Shots Taken} + \text{Total Assists}}$$

A point streak is commonly defined as the number of consecutive games in which a player records at least one point via goal or assist (12). In order to track the scoring streaks of each individual player, a binary variable was created to express whether a player received a point for a given game, where a player that recorded at least one point for the given game would receive a numerical value of 1 or receive a numerical value of 0 if a point was not recorded, which is represented in equation 2.

$$f(\text{Score}) = \begin{cases} 1, & \text{if point} > 0 \\ 0, & \text{otherwise} \end{cases}$$

Following the creation of the score binary variable, a string of 1s and 0s was created to visually show the games in which Wayne Gretzky received or did not receive a point. Equation 3 expresses how a variable was created to represent each game that a player participated in during the season in chronological order as a running sequence –  $x_{1_i}, x_{2_i}, \dots, x_{n_i}$ , where  $x_n$  is the number of the played game for each player,  $i$ .

$$\text{Game Number} = (x_{n_i})_1^{n_i}$$

Streak lengths were computed to create a list of the tallied number of consecutive games where a point was recorded, as well as creating a similar list of the tallied number of



consecutive games in which a point was not recorded for Wayne Gretzky during the entire 1983-1984 season.

An alternative way of investigating the streaky scoring performances of players is to use the scoring average, which is computed over short time intervals to visually represent the fluctuation of scoring efficiency throughout the entire season (12). In this analysis, we looked at repeatedly computing Wayne Gretzky's scoring average every ten games to evaluate the fluctuation of his scoring productivity. A moving average function was created using equation 4, where  $n_i$  represents the game of interest and  $w$  represents the window length.

$$\text{Moving Average} = \frac{\sum_{n_i}^{n_i+(w-1)} \text{Scoring Average}}{w}$$

The created moving average function was then applied the dataset to create a plot of Wayne Gretzky's moving scoring average, measured every ten games, throughout the entire 1983-1984 season.

In addition, we wanted to see the top scoring streaks for the entire league and see how those streaks compared to Gretzky's, and we wanted to investigate the top scoring droughts for the entire league compared to Gretzky's. Following the calculation of maximum scoring streak lengths for each individual player, a simulation was created to see whether Gretzky's 51-game scoring streak was unusual or not by running ten thousand replications of the streakiness, or clumpiness, statistic to compute ten thousand values of the clumpiness measure of the simulation procedure (10, 12).

$$\text{Sum of Squares of Gaps} = \sum_{i=0}^n (X_i - \bar{X})^2$$

In order to test the clumpiness statistic and determine whether Gretzky's streak would be considered normal for his play during the 1983-1984 season, the sum of squares of the

gaps between successful games in which a point was scored was taken to represent the clumpiness measure or observed value, where  $X_i$  is the  $i^{th}$  game in the set and  $\bar{X}$  is the mean of all games in the set (Equation 5). In order to test this clumpiness measure, we created a function to perform one iteration of the simulation experiment where the input is a vector of 0s and 1s (12). We then created another function in order to find random arrangements of the input vector to find the gaps between games in which a score occurred. The sum of squares of the gaps was computed. This simulation was repeated ten thousand times.

Following the initial representation of the maximum streak lengths for each player in the entire league, a visual representation was desired to see how Gretzky's streak stacked up to the rest of the league. The first method used to visually represent this was to create a histogram of the distribution of all the scoring streaks obtained by every player during the 1983-1984 season. Additionally, we wanted to see how Gretzky's total points scored game by game compared to other players in the league that participated in the same number of games as Gretzky by creating a step plot to show each time a point was obtained. We also wanted to see how Gretzky's streak compared to some additional players' streaks by creating a plot to show each time a point streak was extended. Finally, we decided to run a simulation on Gretzky's season scoring statistics ten thousand times using a Poisson distribution, where  $x$  represents the total number of games played and  $\lambda$  is equal to his season long points per game value (Equation 6).

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

To compare Joe DiMaggio's 56-game hitting streak to Wayne Gretzky's 51-game point streak, the same analysis as mentioned above was used to statistically and visually

represent the hitting streaks and droughts DiMaggio obtained, his moving average during the 1941 season, his clumpiness statistic and simulation of his 1941 season through ten thousand simulations using the Poisson distribution to further investigate his hitting streak.

## RESULTS

Wayne Gretzky started the 1983-1984 season by recording at least one point in 51 consecutive games, then he had one game where he did not record a point, followed by a 5-game point streak, and so on. Throughout the entire season, Gretzky obtained point streaks of 51 games, 5 games, 10 games and 5 games during the 74-game season, however, he encountered three games in which he did not record a point, which ended his point streaks on three separate occasions. Joe DiMaggio started the 1941 season with a hit in eight consecutive games, followed by a span of three games in which he did not record a single hit, and so on. DiMaggio encountered five hitting streaks of shorter lengths leading up to his record 56-game hitting streak that he followed up with another 16-game hitting streak after going hitless in one game. Additional tables can be found in Appendix A.

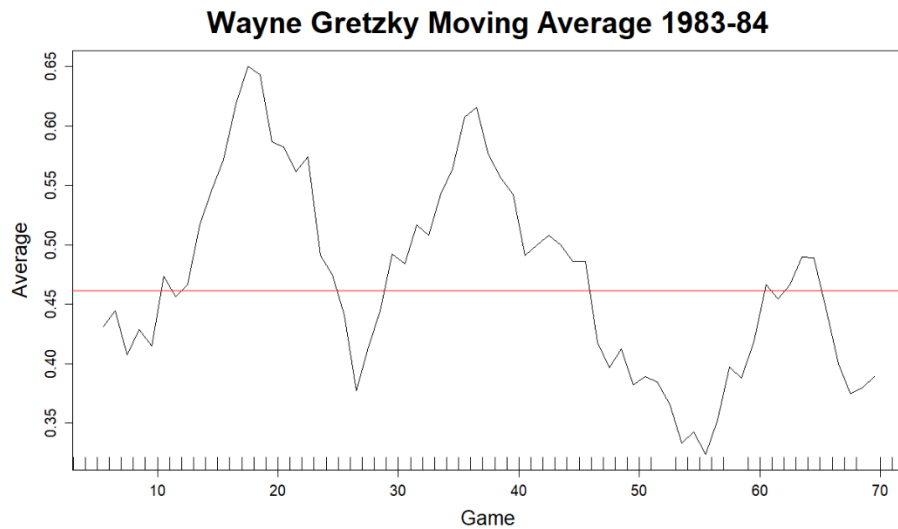


Figure 1: Wayne Gretzky Moving Average. Moving average plot of Gretzky's scoring average for the 1983-84 season using a window of 10 games. The red horizontal line show Gretzky's season scoring average. The games where Gretzky recorded at least one point are shows on the horizontal axis.

Investigating Gretzky's scoring average, computed over short time intervals, we can see that he must have been a more dominant scorer during his 51-game point streak and slightly less dominant during the period following his big streak, even though he immediately followed up his 51-game point streak with a 5-game point streak. Figure 1 shows the moving average of Gretzky's 1983-1984 season while using a horizontal line to represent his season long scoring average as a reference point to see where Gretzky experienced greater dominance as a scorer throughout the season. A similar analysis was done to investigate and replicate Joe DiMaggio's moving average during his 1941 season that was previously expressed in *Analyzing Baseball Data with R*.

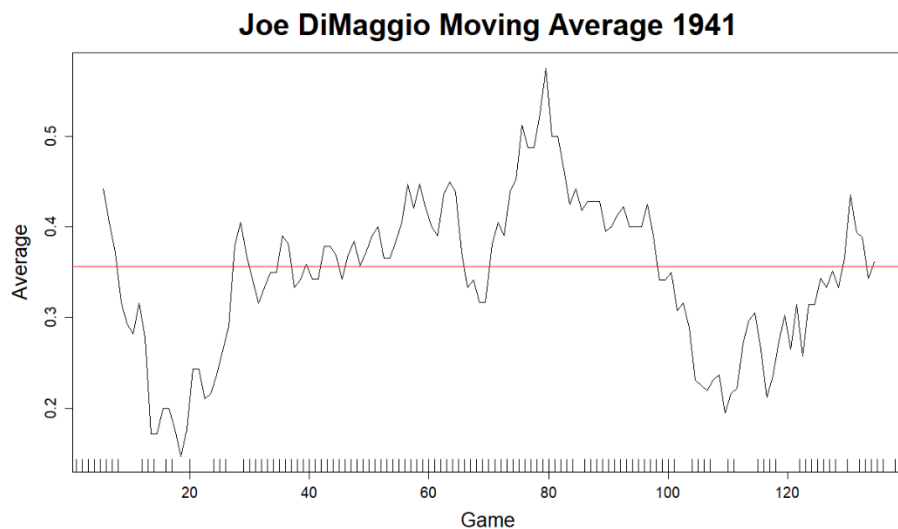


Figure 2: Joe DiMaggio Moving Average. Moving average plot of DiMaggio's batting average for the 1941 season using a window of 10 games. The red horizontal line shows DiMaggio's season batting average and the games where a hit was recorded are marked.

Figure 2 shows how DiMaggio's hitting performance improved consistently during his 56-game hitting streak, followed by a range of games in which DiMaggio encountered a slight hitting slump towards the end of the season (12). Examining Gretzky's streakiness, based on the sequence of receiving a point in a game or not, was done by running a simulation ten thousand times based on the gaps between successive games where a point

was scored. The results of these ten thousand simulations are presented in the histogram (see Figure 3).

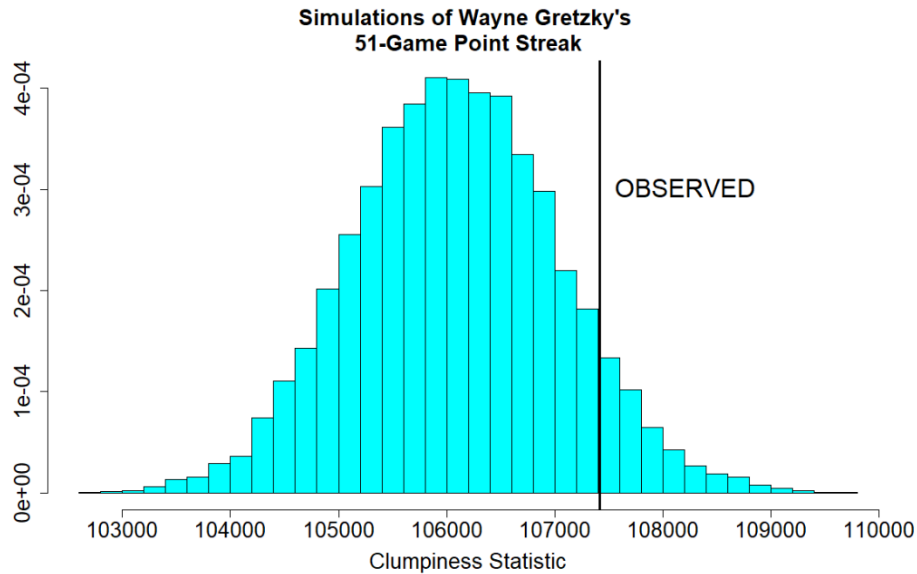


Figure 3: Simulations of Wayne Gretzky’s 51-Game Point Streak. Histogram of ten thousand values of the clumpiness statistic assuming all arrangements of receiving a point scored in a given game. The solid black vertical line is the observed value of the clumpiness statistic for Gretzky.

We can see that the Gretzky’s clumpiness measure falls slightly into the right tail of the distribution, which indicates that he did in fact display slightly more streakiness than one would expect to see for Gretzky, but it was not unusual for his caliber of play during the 1983-1984 season. Alternatively, for DiMaggio, we can see that the clumpiness measure for his 56-game hitting streak drastically falls into the right tail of the distribution of replications, which indicates that his hitting streak was in fact an abnormality (see Figure 4).

Following the test to determine whether Gretzky’s streak was an anomaly, we wanted to see just how his streak and scoring compared to every other player in the league. Figure 5 gives a visual representation of all point scoring streaks that were obtained for the entire league throughout the season. Gretzky’s streak has been labeled to

show just how remarkable his streak was compared to the rest of the league. Looking at the histogram of the distributions of all point streaks for the league, we can see that Gretzky's streak falls in the right tail of the distribution, which indicates that his streak is abnormal.

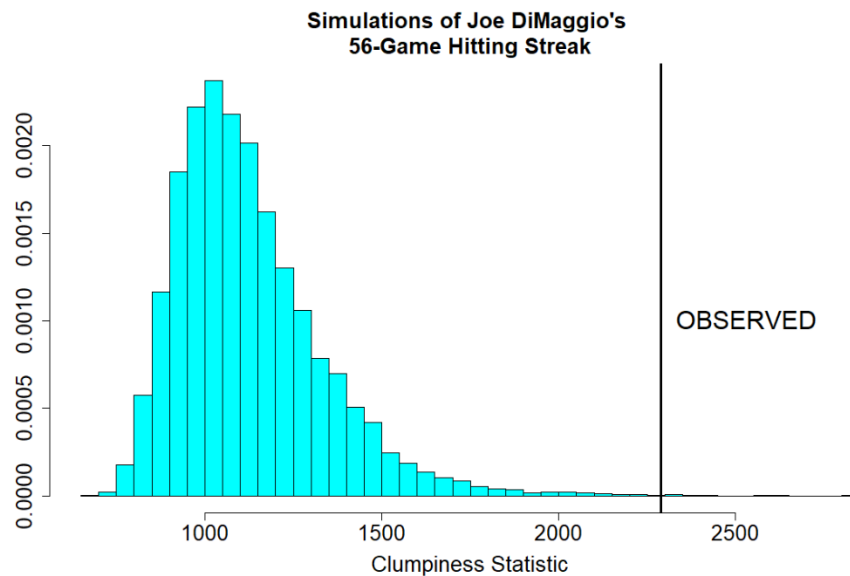


Figure 4: Simulations of Joe DiMaggio's 56-Game Hitting Streak. Histogram of ten thousand values of the clumpiness statistic assuming all arrangements of hits and outs are equally likely for Joe DiMaggio. The solid black vertical line is the observed value of the clumpiness statistic for DiMaggio.

To further understand how impressive Gretzky was during this season, we wanted to see how his point accumulation stacked up against the remainder of the league for the entire season. We can see in Figure 6 that Gretzky begins to pull away from the rest of the league, in terms of points scored per game throughout the entire season, which also shows how dominant he was as a scorer compared to every other player in the league. Gretzky accumulated a total of 205 points (87 goals and 118 assists) in only 74 games during the 1983-1984 season compared to 126 points scored for the next closest player, which was teammate Paul Coffey.

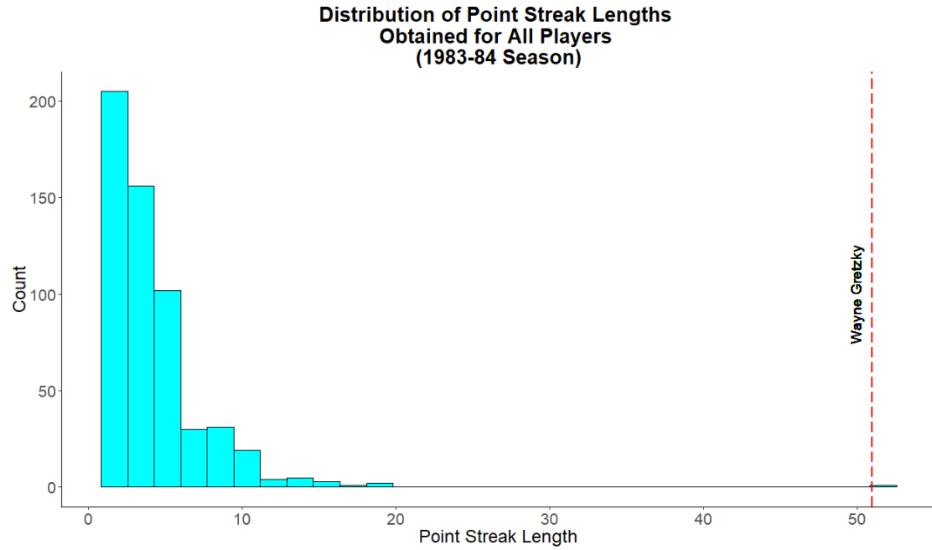


Figure 5: Distribution of Point Streaks Obtained for all NHL Players. Histogram of the distribution of all point streaks obtained by all players during the 1983-84 season. Wayne Gretzky’s 51-game point streak is shown by the dashed red line.

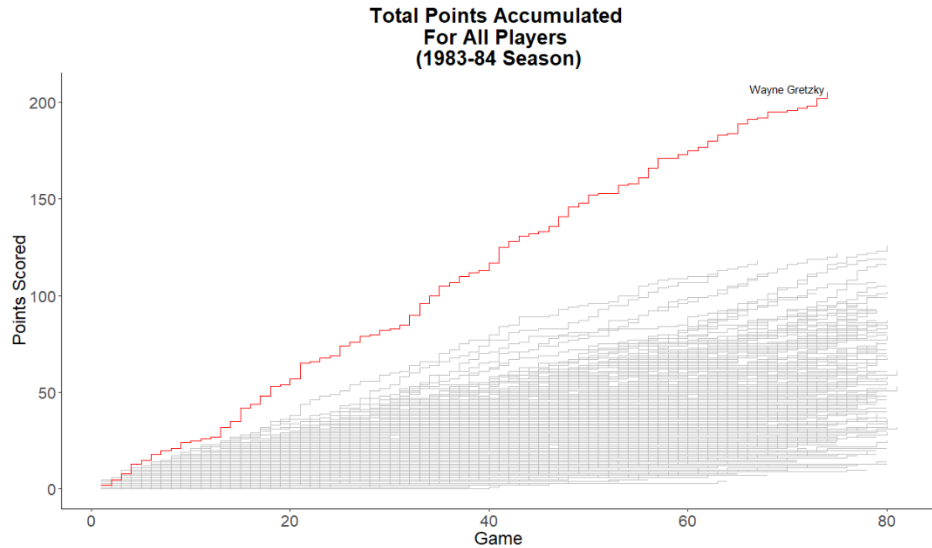


Figure 6: Total Points Accumulated for all NHL Players. Step plot of total points accumulated for each player during the 1983-84 season. The red step plot represents the points accumulated for Wayne Gretzky.

Additionally, we wanted to see each point streak for each player, during the 1983-1984 season, by creating a plot where there would be an increase every time there was a point scored in a game, otherwise the plot would restart at zero when a point was not recorded (see Figure 7). We can see that Gretzky streak is much greater than anyone



else's in the league, and the next closest streak obtained to Gretzky's 51-game point streak was only a 19-game point streak that was obtained on two occasions. These figures all demonstrate how talented Gretzky was as a scorer and playmaker for his team, and how remarkable he was as a player altogether.

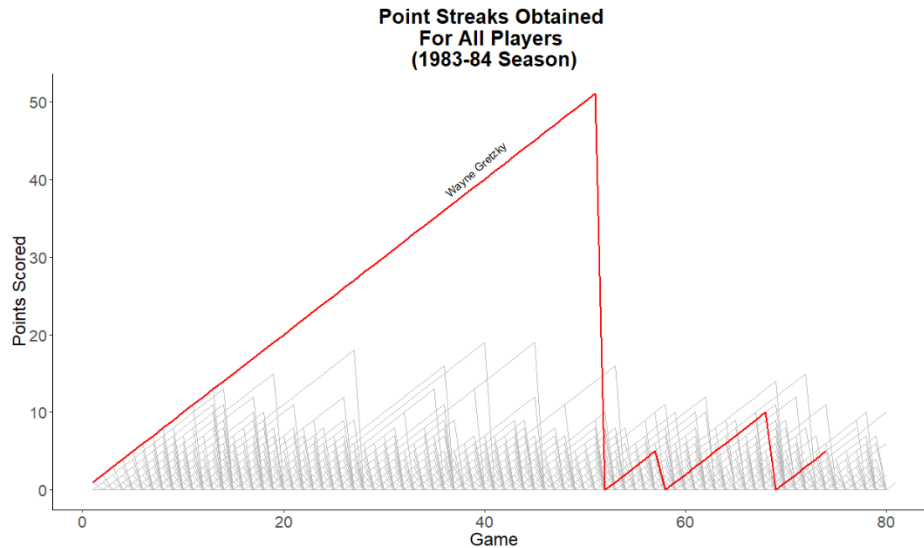


Figure 7: Point Streaks Obtained for all NHL Players. Plot of the point streaks obtained by each player during the 1983-84 season. An increase in the plot was experienced when a given player recorded a goal in a given game, thus extending their corresponding point streak. If a point was not recorded, the streak dropped down to 0 and started over. The red line represents the point streaks obtained for Wayne Gretzky.

Finally, by looking at the simulation of ten thousand replications of Wayne Gretzky's 1983-1984 season, using a Poisson distribution, we can see that Gretzky is predicted to replicate his 51-game point streak numerous times, which can be seen in Figure 8. Based on the output, Gretzky is predicted to obtain a streak of 74 games, which would imply that he recorded at least one point in every single game that he participated in throughout the season. However, after running the same simulation for Joe DiMaggio's 1941 season, the simulation does not show any predictions that he would

duplicate his streak or even obtain it to begin with. Figure 9 shows that the closest DiMaggio even gets to matching his 56-game point streak is only about 30 games.

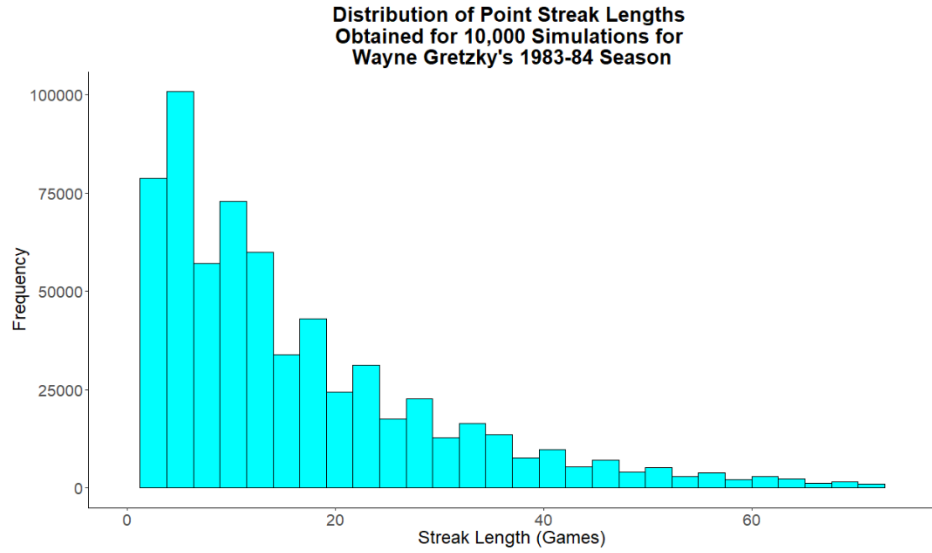


Figure 8: Distribution of Point Streak Lengths Obtained for 10,000 Simulations for Wayne Gretzky's 1983-1984 Season. Histogram of 10,000 simulations of Wayne Gretzky's 1983-1984 season point streaks following a Poisson distribution.

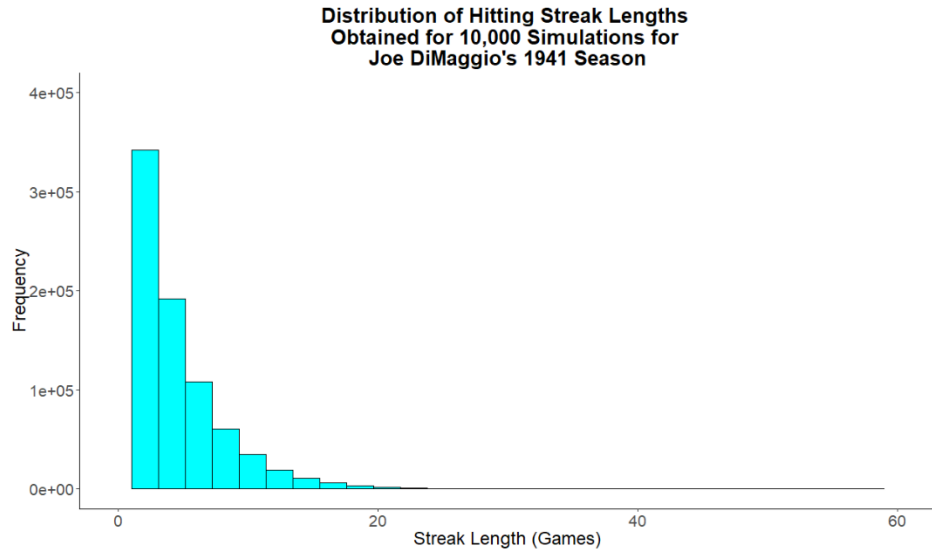


Figure 9: Distribution of Hitting Streak Lengths Obtained for 10,000 Simulations for Joe DiMaggio's 1941 Season. Histogram of 10,000 simulations of Joe DiMaggio's 1941 season hitting streaks following a Poisson distribution.

## CONCLUSION

This comprehensive analysis on the scoring and point streak of Wayne Gretzky highlights how dominant he was as a player and how remarkable his 1983-1984 season was compared to the remainder of the league. Gretzky displayed an exceptional talent level throughout the entire season, which is supported by his outstanding 205 points scored in only 74 games. Gretzky proved his dominance by contributing points in a rapid and consistent fashion, which played a major role into the success of his teams' season as they went on to win the Stanley Cup Playoffs. In terms of points scored throughout the season, no other player could contend with Gretzky. Gretzky showcased his playmaking ability by creating scoring opportunities for himself and his team more consistently and frequently than any other player in the league. Gretzky began to separate himself from the remainder of the league in terms of point accumulation for the entire season within the first twenty games and he never looked back. He continued his monumental season until the very last game when he and his teammates hoisted that historic trophy as champions.

Scoring this record amount of points in a substantially low amount of games is impressive, but to record a goal or assist in 51-games straight is monumental. Much like Gretzky's ability to separate himself from the rest of the league in terms of points scored, he was also able to separate himself in a historic fashion when he took the ice on the first game of the season; this was the beginning of one of the most remarkable moments in sports history. Gretzky proved his greatness by finding ways to create scoring advantages for his entire team that enabled him to obtain such a remarkable record and streak, and it is evident that no other player shared his level of greatness.

Gretzky started out the season in a similar fashion as many other players by recording at least one point in a few consecutive games, but it did not stop there. Gretzky continued this until he had scored at least one point in 51 consecutive games; the next closest player was just below a 20-game streak. This is something that many fans, statisticians and critics believe will likely never be broken. Based on the elite level of play that Gretzky performed routinely, it is hard to not believe that he was capable of something so historic and magnificent at any given time. Although Gretzky's streak is shown to express some streakiness and resemble some slight abnormality for his caliber of play throughout the season, we can still see how monumental and historic it ultimately was.

But in comparison to Joe DiMaggio's 56-game hitting streak, one can say it is not as impressive. DiMaggio's hitting streak is another streak that many believe should never have taken place to begin with, but for DiMaggio to obtain that streak and then immediately follow it up with another hitting streak of 16 games shows how truly dominant he was in his sport at the time. DiMaggio's streak is shown to express a great amount of abnormality which resembles that a great amount of streakiness must have occurred for Joe to record at least one hit in 56 consecutive games. In terms of abnormality, it is evident that, based on the season data of Gretzky and DiMaggio, DiMaggio resembled the most streakiness for his streak and unlikelihood of replication based on the ten thousand simulations of corresponding season data. Thus, one could conclude that DiMaggio's streak is far more impressive than that of Gretzky. However, it is hard to adequately determine which streak was the most impressive or unlikely to occur without the proper readily available statistics for the NHL prior to the 2000s.

Therefore we can only admire the greatness of Wayne Gretzky and Joe DiMaggio and acknowledge that both Gretzky's 51-game point streak and DiMaggio's 56-game hitting streak are two records in sports that are monumental and likely to never be broken.

## APPENDIX A

Table A.1: Wayne Gretzky's Game-by-Game Breakdown. Table of all games in which Wayne Gretzky recorded or did not record a point during the 1983-1984 NHL season. A value of 1 represents an event in which a point was scored, whereas a value of 0 represents an event in which a point was not scored.

Game	Score	Game	Score	Game	Score
1	1	26	1	51	1
2	1	27	1	52	0
3	1	28	1	53	1
4	1	29	1	54	1
5	1	30	1	55	1
6	1	31	1	56	1
7	1	32	1	57	1
8	1	33	1	58	0
9	1	34	1	59	1
10	1	35	1	60	1
11	1	36	1	61	1
12	1	37	1	62	1
13	1	38	1	63	1
14	1	39	1	64	1
15	1	40	1	65	1
16	1	41	1	66	1
17	1	42	1	67	1
18	1	43	1	68	1
19	1	44	1	69	0
20	1	45	1	70	1
21	1	46	1	71	1
22	1	47	1	72	1
23	1	48	1	73	1
24	1	49	1	74	1
25	1	50	1		

Table A.2: Point Streaks Obtained by Wayne Gretzky. Table of all the point per game streaks that Wayne Gretzky obtained during the 1983-1984 NHL season.

Point Streaks Obtained	Consecutive Games			
	51	5	10	5

Table A.3: Scoring Droughts Obtained by Wayne Gretzky. Table of all streaks of consecutive games in which Wayne Gretzky did not record a point during the 1983-1984 NHL season.

Scoring Droughts Obtained	Consecutive Games		
	1	1	1



Table A.4: Joe DiMaggio's Game-by-Game Breakdown. Table of all games in which Joe DiMaggio recorded or did not record a hit during the 1941 MLB season. A value of 1 represents an event in which a hit was obtained, whereas a value of 0 represents an event in which a hit was not obtained.

Game	Hit	Game	Hit	Game	Hit	Game	Hit	Game	Hit
1	1	29	1	57	1	85	0	113	0
2	1	30	1	58	1	86	1	114	0
3	1	31	1	59	1	87	1	115	1
4	1	32	1	60	1	88	1	116	1
5	1	33	1	61	1	89	1	117	1
6	1	34	1	62	1	90	1	118	1
7	1	35	1	63	1	91	1	119	0
8	1	36	1	64	1	92	1	120	0
9	0	37	1	65	1	93	1	121	0
10	0	38	1	66	1	94	1	122	1
11	0	39	1	67	1	95	1	123	1
12	1	40	1	68	1	96	1	124	1
13	1	41	1	69	1	97	1	125	1
14	1	42	1	70	1	98	1	126	1
15	0	43	1	71	1	99	1	127	1
16	1	44	1	72	1	100	1	128	1
17	1	45	1	73	1	101	1	129	0
18	0	46	1	74	1	102	0	130	1
19	0	47	1	75	1	103	0	131	0
20	1	48	1	76	1	104	1	132	1
21	0	49	1	77	1	105	1	133	1
22	0	50	1	78	1	106	1	134	1
23	0	51	1	79	1	107	1	135	1
24	1	52	1	80	1	108	0	136	1
25	1	53	1	81	1	109	0	137	0
26	1	54	1	82	1	110	1	138	1
27	0	55	1	83	1	111	1	139	1
28	0	56	1	84	1	112	0		

Table A.5: Hitting Streaks Obtained by Joe DiMaggio. Table of all the hitting streaks that Joe DiMaggio obtained during the 1941 MLB season.

Hitting Streaks Obtained	Consecutive Games													
	8	3	2	1	3	56	16	4	2	4	7	1	5	2

Table A.6: Hitting Slumps Obtained by Joe DiMaggio. Table of all streaks of consecutive games in which Joe DiMaggio did not receive a hit during the 1941 MLB season.

	Consecutive Games												
Hitting Droughts Obtained	3	1	2	3	2	1	2	2	3	3	1	1	1

## REFERENCES

1. McMullen, P.R. (2017). Standardization of Winning Streaks in Sports. *Applied Mathematics*, **8**, 344-357. Retrieved from [https://file.scirp.org/pdf/AM\\_2017032716133787.pdf](https://file.scirp.org/pdf/AM_2017032716133787.pdf)
2. Arbesman, S, and Strogatz, S.H. (2008). A Monte Carlo Approach to Joe DiMaggio and Streaks in Baseball. Retrieved from <https://arxiv.org/ftp/arxiv/papers/0807/0807.5082.pdf>
3. Rothman, S., & Le, Q. (2010). Streaking: Finding the Probability for a Batting Streak. In Gallian J. (Ed.), *Mathematics and Sports* (pp. 31-52). Mathematical Association of America. Retrieved from <http://www.jstor.org/stable/10.4169/j.ctt6wpwsw.7>
4. Martel, J., Origins of Ice Hockey (2019). In *The Canadian Encyclopedia*. Retrieved from <https://www.thecanadianencyclopedia.ca/en/article/origins-of-ice-hockey>
5. Sim, F. H., Simonet, W. T., Melton, L. J., & Lehn, T. A. (1988). Ice Hockey Injuries. *The American Journal of Sports Medicine*, 16(1\_suppl), S-86-S-96. Retrieved from <https://doi.org/10.1177/03635465880160S119>
6. Montgomery, D.L. *Sports Medicine* (1988) 5: 99. Retrieved from <https://doi.org/10.2165/00007256-198805020-00003>
7. Hammond, R. (2011). Understanding Advanced Stats, Part One: Corsi and Fenwick. Retrieved from <https://www.matchsticksandgasoline.com/2011/7/29/2290643/understanding-advanced-stats-part-one-corsi-fenwick>
8. Laws, K. (2015). Hockey 201: What the Heck is Corsi? *The Hockey Writers*. Retrieved from <https://thehockeywriters.com/hockey-201-what-the-heck-is-corsi/>
9. NHL.com. Statistics. Retrieved from <http://www.nhl.com/stats/player?reportType=game&dateFrom=1983-10-01&dateTo=1984-04-30&gameType=2&filter=gamesPlayed,gte,1&sort=gameDate&ferer=https://www.google.com/&httpsredir=1&article=1067&context=statsp>
10. Sison, C. (2017) Analyzing Baseball Data with R (pp. 64-68). Retrieved from <https://www.google.com/&httpsredir=1&article=1067&context=statsp>
11. Baseball Reference. Joe DiMaggio: 1941 Batting Game Log. Retrieved from <https://www.baseball-reference.com/players/gl.fcgi?id=dimagjo01&t=b&year=1941>
12. Marchi, M, Albert, J, and Baumer, B.S. (2019). *Analyzing Baseball Data with R: Second Edition* (pp. 229-242).

# CURRICULUM VITAE

Tyler Ison

## **Education**

- M.S. in Biostatistics, Indiana University degree, earned at Indiana University-Purdue University Indianapolis, IN 2019
- B.A. in Biology, Minors in Applied Mathematics and Quantitative Analysis, Franklin College, IN 2016

## **Professional Experience**

- Data Scientist, B2S Life Sciences, Franklin, IN, July 2019 – present
- Bioinformatics Intern, Eli Lilly & Company, Indianapolis, IN, January 2016