Darts: An Analysis of Strategy and Factors Contributing to Success

Crystal Baker
Fab Joseph
Lex Varani
Simon Weisenhorn
Advaith Deo

STOR 538

Mario Giacomazzo

March 12, 2021

Introduction

Although often dismissed as a pub game, our group decided to examine professional darts for this sports analytics project. In professional darts, players start with 501 points. During each turn, a player throws three darts. The total points scored from each turn are then subtracted from the total score. Points scored is determined by where the dart landed on the dartboard. A traditional dartboard is divided equally into twenty segments numbered 1 through 20, as well as the bullseye. The bullseye is the red dot in the middle of the dartboard and the outer bull is its surrounding green circle, earning 50 and 25 points respectively. There are sections of the dartboard colored red and green that are multipliers to a player's score. The outermost alternating red and green circle on the dartboard represents the double ring, earning twice the value of the number landed on. There is another red and green circle in between the double ring and the bullseye, known as the treble ring, earning three times the number landed on. If a dart lands on the black or white section of a dartboard, there is no multiplier and the player earns the number landed on. The first player to reach exactly zero points wins that leg. Meaning, if a player needs to throw 60 points to get to 0 and they throw 80, then their score remains at 60 and must play another turn. However, the last dart that takes a player to zero must land on the double multiplier ring. The number of legs needed to win the match depends on the tournament at which it's being played.

The goal of our analysis is to identify strategies players employ to win games and if any variables significantly contribute to a player winning. In particular, our group wanted to know which locations on the dartboard are targeted in pivotal situations over the course of a leg. We hypothesized that a player will most likely target T20 (treble 20) in the first couple turns of the game, and aim for numbers with little or no multipliers towards the end of a game. We also hypothesized that the time players took to throw their darts would have an impact on whether or not they won the leg. Originally, our group recorded 17 different variables to answer these questions and discovered compelling patterns of dart player behavior. After collecting data, we added even more variables. Below is a preview of our data.

Player Name	Opponent Name	Turn Number	Throw 1	Throw 2	Throw 3	Points Left	Time per Turn	Point Diff
Gerwyn Price	Gary Anderson	1	T20	20	20	401	7.81	100
Gary Anderson	Gerwyn Price	1	T20	T20	20	361	6.91	40
Gerwyn Price	Gary Anderson	2	T20	T20	19	262	9.41	99
Gary Anderson	Gerwyn Price	2	20	19	20	302	8.09	-40
Gerwyn Price	Gary Anderson	3	20	4	T20	178	8.62	124

Gary Anderson	Gerwyn Price	3	18	18	T20	206	7.31	-28
Gerwyn Price	Gary Anderson	4	20	20	18	120	8.02	86
Gary Anderson	Gerwyn Price	4	T20	T20	T18	32	7.41	88

From February 26th to March 6th, each group member was tasked with watching 3 legs of dart games. The professional dart games we watched came from national competitions including the 2019/2020 Darts World Championship, 2013 Grand Slam of Darts, 2017 World Matchplay Final, 2019 US Darts Masters, 2016 World Darts Championship, 2018 Grand Slam of Darts, WDF Virtual Cup 2020, German Darts Open 2018, PDC Super Series 1, and individual games hosted by USA Darts. These games were either uploaded onto Youtube or found on PDCTV. Our data collection process was a simple undertaking since we equally delegated responsibilities. We aggregated our data into a Google Sheets file. There are 469 observations in our dataset, each representing a single turn of a dart player.

To get a better understanding of our dataset, we will explain 5 variables seen in the preview. The variable, Throw 1, represents the location of the first dart thrown and was measured by finding where the first dart landed on the dartboard. For example, T20 means the dart landed on the number 20 in the treble ring. The variable, Throw 2, represents the location of the second dart thrown and was measured by finding where the second dart landed on the dartboard. The variable, Points Left, represents how many points a dart player has left to reach zero and was measured by subtracting points scored from 501. The variable, Time per Turn, is the number of seconds it took for a player to throw all three of their darts and was measured by timing a dart player's turn. The variable, Point Diff, represents the score difference between the player throwing and their opponent. Point Diff was computed by subtracting the total points of the opponent from the total points of the throwing player.

In our raw data, we have two variables called Checkout Turn and Checkout Opportunity. The Checkout Turn variable is whether or not the respective turn ended the game. Checkout Opportunity is if a player has an opportunity to end a game with a throw. A checkout opportunity occurs whenever a player has 170 or fewer points left. Although if a player has 169, 168, 166, 165, 163, 162, or 159 points left, a player cannot checkout in one turn. These numbers are called bogey numbers because it requires more than one turn to get these numbers to zero. For both of these checkout variables a 1 represents a yes and a 0 represents a no. The rest of the variables are easy to understand from their name alone. We created another dataset called “Legs’ which gives summary statistics for each leg of a game, so we could get a broad understanding of dart player behavior on a leg to leg basis.

Summary

Table 1

Statistics	Mean	Standard Deviation	Min	Q1	Q2	Q3	Max
Darts per Turn	2.89	0.422	0	3.00	3.00	3.00	3.00
Points per Turn	92.79	43.45	0	59.00	96.00	135.00	180.00
Point Difference	50.20	90.06	-205	-6.00	51.00	119.00	272.00
Checkout Opportunity	0.449	0.50	0	0	0	1	1
Time per Turn (sec)	8.43	2.95	1.58	6.43	7.88	10.23	25.91
Turn Number	3.07	1.54	1.00	2.00	3.00	4.00	7.00
Points Left	197.97	142.93	0.00	56.00	201.00	321.00	492.00

Table 1 summarizes 6 of the numeric statistics that we recorded from the professional darts games. For each of the statistics in this table, we calculated the mean, standard deviation, first, second, and third quartiles, and the maximum and minimum numbers. We noticed that the average leg consisted of 3 turns, sometimes stooping below 2 and above 4. The amount of time taken per turn to throw the three darts averaged at 8.43 seconds, showing that the darts players we watched often shot the three darts in quick succession, taking 1 to 2 seconds between throws.

Table 2

Variable	Potential Values	Count	Relative Frequencies
Throw 1	20	185	39.44%
	T20	152	32.40%
	5	21	4.47%
	19	17	3.62%
	T5	10	2.13%
Throw 2	20	146	31.13%
	T20	141	30.06%
	19	47	10.02%
	T19	21	4.47%
	No Throw	17	3.62%
Throw 3	20	123	26.22%
	T20	99	21.10%
	T19	40	8.52%
	19	37	7.89%
	No Throw	31	6.61%
Player Name	Gary Anderson	44	9.38%
	Peter Wright	35	7.46%
	Phil Taylor	33	7.03%
	Michael van Gerwen	30	6.39%
	Adrian Lewis	29	6.18%
Opponent Name	Gary Anderson	44	9.38%
	Peter Wright	36	7.67%
	Phil Taylor	32	6.82%
	Adrian Lewis	28	5.97%
	Michael van Gerwen	28	5.97%

Table 2 presents 5 of our categorical statistics, including the names of the players and their opponents, as well as the values obtained from each of the throws. For the first throw, the highest frequency was 20 points with no double or triple multiplier, followed closely by 20 with a triple multiplier (60 points). The highest frequency for the second and third throws were also 20 followed closely by treble 20. The player that appeared most in our data is Gary Anderson, who has had a very successful professional darts career, formerly being ranked number 1 and a two-time Professional Darts Corporation World Champion.

Figure 1

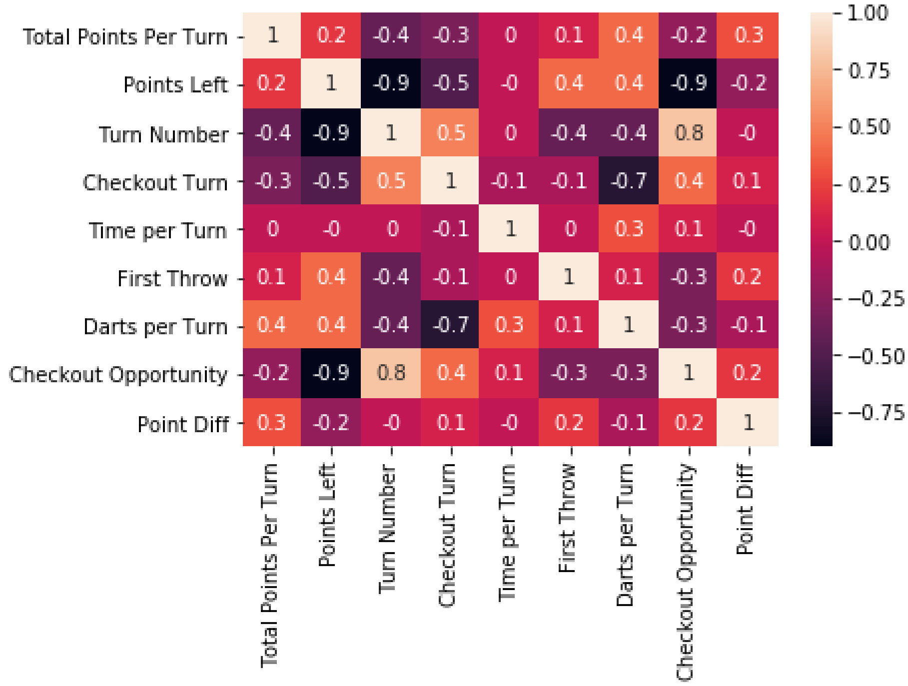


Figure 1 is a correlation matrix of all of the variables we used in our analysis. Many of the stronger correlations are easily explainable and not very significant. For example, Points Left and Turn Number are strongly negatively correlated. This is because as turns go on, a player accumulates points and has less points remaining. But otherwise, there were not many strong correlations that surprised us.

Figure 2

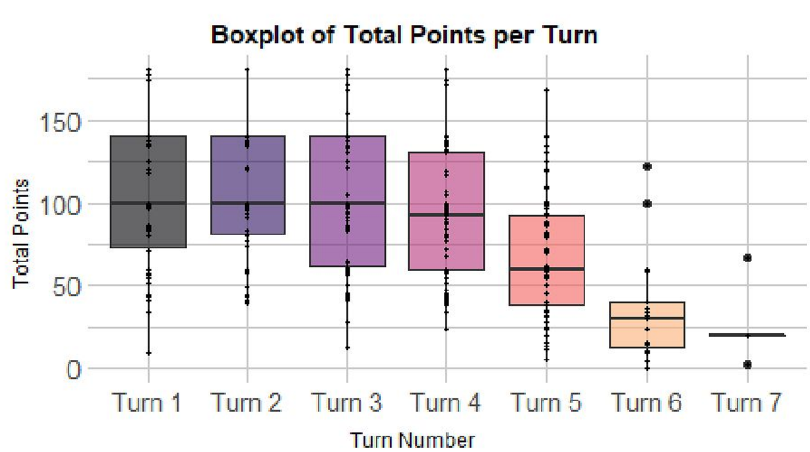


Figure 2 shows a boxplot of points scored for each turn. This visual summary identifies the relationship between total points for each respective turn number by including the median value and dispersion of data. From this figure it is clear that dart players target lower points when they are nearing the end of the game.

Insights

Time and Checkout Opportunities

From figure 1, we can see that the variable 'Time per Turn' does not have a strong correlation with any of the other variables, the strongest correlation being 0.3 with the variable 'Darts per Turn.' However, we initially hypothesized that the time taken to throw the darts would play a role in a players' success, and we are still interested to see if this can be further evaluated. From figure 2, we see that there are less points per turn at turns 5, 6, and 7, as these turns have higher potential for checkout opportunities, or opportunities to win the game on that turn. We felt as though during opportunities for checkout, players may potentially take more time, which would impact their success in completing the checkout. Thus, we chose to further investigate whether the checkout time has a correlation with the amount of legs a player won in our data. In order to accomplish this, we took the number of checkout opportunities and divided it by the number of turns to get a checkout rate for each of the players we took data for.

Checkout Rate = Number of Checkout Opportunities / Number of Turns

Next, we needed to find the average time taken by players in turns where they had an opportunity for a checkout.

Average Time Taken at Checkout =
Σ Time Taken During Checkout Opportunity / Number of Checkout Opportunities

Using these two statistics, we created a metric for the times taken during checkout opportunities, known as "Time During Checkout."

*Time During Checkout = Checkout Rate * Average Time Taken at Checkout*

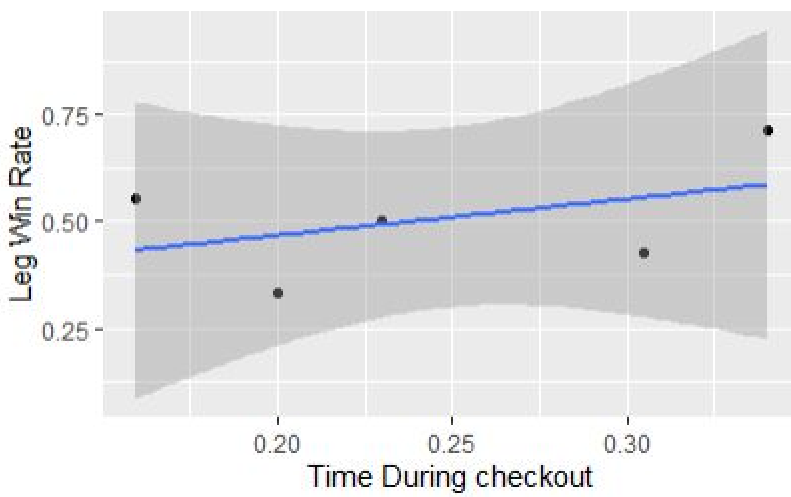
The idea behind this is to compare the "Time During Checkout" with another simple statistic we derived from our data, which is just the rate of Legs won, computed by taking the Number of Legs Won divided by the Number of Legs Played per player. This also ensures that this rate will be adjusted for players who we recorded more games for.

Leg Win Rate = Number of Legs Won / Number of Legs Played

For the 5 players with the highest frequency in our data, we calculated their "Time During Checkout," giving an indication of who used more time when they came to checkout opportunities.

Player Name	Time During Checkout
Phil Taylor	0.3406
Peter Wright	0.305
Adrian Lewis	0.23
Michael van Gerwen	0.1998
Gary Anderson	0.1592

After compiling the Leg Win Rate for each of the five players, we created a figure to demonstrate the relationship and also ran a t-test to evaluate its significance.

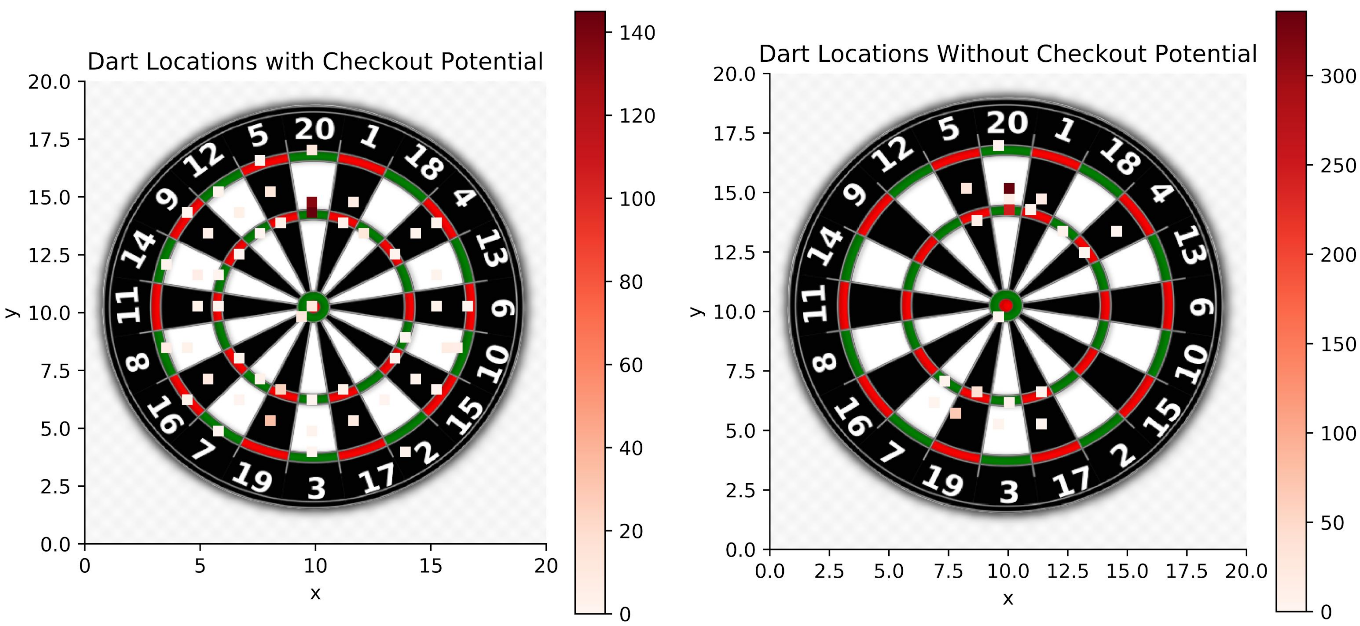



```
        Welch Two Sample t-test

data:  graph$`Time During Checkout` and graph$`Leg Win Rate`
t = -3.5966, df = 6.0355, p-value = 0.0113
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.43560364 -0.08316436
sample estimates:
mean of x mean of y
 0.246946  0.506330
```

The t-test provided a p-value of 0.0113 for the relationship between Time During Checkout and Leg Win Rate, indicating that the correlation between these two variables is strong and that we can say with 95% confidence that the relationship is statistically significant. In our introduction, our thought process was that a player may take more time when they are in need of a specific amount of points, which would result in them being successful in winning the leg. In examining the relationship between the amount of time a player takes when they have an opportunity for a checkout and how successful they are in the amount of legs they win, we found that our initial hypothesis was correct, and that there is a competitive advantage to the amount of time taken to shoot your darts.

Pivotal Situation Analysis



Here we developed two diagrams that map the locations of the dart from the data collected. The diagram on the left displays where the players threw the dart when there was potential for a checkout turn and the diagram on the right displays where the players threw the dart when there was not a potential for a checkout turn. Each square represents the frequency of the dart thrown and is colored based on whether it was recorded often or less frequently. The darker the square means that players threw there more often in comparison to the lighter squares.

A checkout turn is a crucial situation in darts, since this a turn where the game could be won. It is most evident that the single 20 and the triple 20 are the most popular locations to throw the dart whether there was a checkout available or not. It is also worth noting that the single 19 is another popular location to throw. This is most likely because both of those locations are the two highest number of points available relative to the size of the target location, maximizing the expected amount of points earned each throw. Thus, the dart players are trying to score the most amount of points possible on the early turns when a checkout is not available and then aim for smaller numbers when a checkout is available. This is also the reason why there are more dart throw locations when there is a potential for a checkout versus when there is not a potential for a checkout.

Critiques and Room for Improvement

Looking back at the work we did with analyzing darts, there are a few changes we would have liked to make. The first being the surplus of variables in our dataset. Before analyzing darts, most of our group members had little knowledge about the sport. This lack of knowledge caused us to collect unimportant variables that did not contribute to the analysis. This is reinforced in our correlation matrix where many variables have little correlation.

A direct example of an insignificant variable was our binary lean variable. Since most of us had not watched a significant amount of darts, we were unfamiliar with the throwing form. We assumed that some players may incorporate leaning forward while others might not. Our assumption was quickly debunked when we observed most dart players lean forward while throwing. Therefore, it was an insignificant variable that was not worth measuring its effect on performance because of the lack of diversity in our data.

Another change that we would have made would be collecting our data from the same league. After our analysis of different games, it is evident that some players were a lot better than others, potentially skewing our sample data and making it more difficult to make generalizations of the entire sport.