SHOOT YOUR SHOT:

An Analysis on Recurve Target Archery and Factors that Influence Shooting Accuracy Success

Authors: Omar Hilal Shaban, Antonio Pano, Clay Burchell, Vivian Li, Connie Cui
Group 4

INTRODUCTION

Our team decided to observe and collect data on individuals' recurve target archery. In a recurve archery match, two archers take turns shooting an arrow at a 10-ring target. The outermost ring of the target is worth 1 point and the innermost circle (including the bull's eye) is worth 10 points. The match is divided by sets, in which each archer shoots three arrows and the archer with the most points from those three arrows wins the set and earns 2 set points. The first archer to earn 6 set points wins the match. While perusing archery match videos, we noticed that one sports commentator mentioned that the time it takes for one archer to shoot is essentially the amount of time their opponent has to load and prepare for their shot. Additionally, while initially researching the sport and potential factors that affect a shot, we came across an article about advice for archers. The article, along with other sources from a quick Google search, suggested keeping both eyes open while shooting. These initial findings made us curious about what other ways shooting time, coordination-related shooting preferences, and other factors could impact a shot. We determined that the measure of success for a shot would be the shot's accuracy (how close it is to the center of the target or bull's eye) and the shot's precision (how repeatable the shot is in terms of result). For this project, we will focus on accuracy. The first question that our project seeks to explore and answer is: What factors contribute most to predicting the accuracy of a shot? Secondly, we would like to examine the effects of stress on accuracy. More specifically, does the significance of a shot (whether it is low pressure or high pressure, as we will define later) affect the archer's accuracy?

To answer these questions, our team gathered data during the week of February 28, 2021. We did so by watching archery matches found on the World Archery, Olympics, and Competition Archery Media YouTube channels. Each of the team members was responsible for finding and viewing a match and collecting observations. We all made sure to select individuals' recurve archery matches to ensure consistency in the scoring-based variables. Our observations were compiled in a shared Google Sheets with a total of 127 rows and 24 columns. Each observation represents one shot, and we collected a total of 126 observations. For each observation, we recorded 24 variables, which included match details, player characteristics, and box-score statistics. Overall, the data represents a total of 5 matches from 5 different tournaments and 10 different archers (2 per match).

Obs	Archer Name	Shot	Target Distance (m)	Time to Shoot (s)	Eyes Open	Blinks	Points	Significant Shot	Desirable Outcome	MPDO
1	Valentina A. Giraldo	1	70	13	1	1	10	No	Tie	10

2	Ana P. Vazquez	1	70	12	2	3	9	No	Tie	10
3	Valentina A. Giraldo	2	70	10	1	2	10	No	Lead	10
4	Ana P. Vazquez	2	70	14	2	3	7	No	Tie	11
5	Valentina A. Giraldo	3	70	10	1	1	9	No	Lead	7
6	Ana P. Vazquez	3	70	12	2	0	9	No	Tie	13

While watching the matches, we were able to collect the data by referring to the scoreboard and countdown timer displayed on the screen, visually observing the archer, and gathering game/match information from the video descriptions. Target Distance was gathered from the video descriptions and did not change throughout the match. Time to Shoot was measured using the countdown displayed on the screen. Each archer has a maximum of 20 seconds to shoot so the Time to Shoot variable is equal to 20 minus the seconds displayed when the countdown stops. Eyes Open refers to whether the archer had both eyes open (2) when aiming and shooting or one eye closed or squinting (1). Next, we counted how many times the archer blinked while aiming (blinks during the time to shoot). The points are directly taken from the scoreboard. Significant Shot, Desirable Outcome, and Min Points to Desirable Outcome (MPDO) are related variables/metrics created using observed variables such as Order (which archer shoots first for the set). The other variables that are not included in this preview are fairly straightforward and self-explanatory.

SUMMARY

Table 1.1: This table summarizes the 5 main numeric variables utilized in our analyses. The minimum value, maximum value, mean, median, and standard deviation are calculated for the variables. Additionally, the number of unique values, most frequent value, and least frequent value are included to further give an idea of the shape of the data.

Variable	Min	Max	Mean	Median	Standard Deviation	# of Unique Values	Most Frequent Value	Least Frequent Value
Target	18	70	55.714	70	21.505	4	70	60

Distance								
Time to Shoot	3	17	9.460	10	3.209	15	10	3, 17
Blinks	0	4	0.669	0	1.071	6	0	4
Points	6	10	9.206	9	0.852	6	10	6
MPDO	7	13	10.039	10	1.024	8	10	13

Table 1.2: This table summarizes the main categorical variables recorded for the observations. These variables are binary, so the sum of the frequencies for each variable should add up to the total observation count of 126. The value with a higher frequency is listed first for each variable. Lastly, the relative frequency is calculated by dividing the frequency by 126 and converting to a percentage.

Variable	Value	Frequency (Counts)	Relative Frequency (%)
Setting	Outdoor	96	76.190
	Indoor	30	23.810
(Dominant) Hand	Right	111	88.095
	Left	15	11.905
Gender	Male	108	85.714
	Female	18	14.286
Eyes Open	2	78	61.905
	1	48	38.095
Desirable Outcome	Tie	76	60.317
	Lead	50	39.683

All other variables (Tournament, Archer Name, Opponent Name, Shot, Set, etc.) are match descriptors that provide context and background information about the shot.

Figure 1: The following figure is a correlation matrix between a subset of the variables described above. Our goal is to determine which factors of archery shooting might be related to one another. The results were somewhat underwhelming. The strongest correlation observed is between Points and Target Distance and is negative. This is not surprising since we expect accuracy to go down the further away a target is from the archer. Moreover, we find a positive correlation between Blinks and Time to Shoot due to the fact that a person will blink more as more time passes before the shot. Also, there is a slight positive correlation between MPDO (defined below) and Time to Shoot. A number of correlations are likely to be either artifacts of the sample size of our data or due to how some derived variables were calculated. Overall, the correlation matrix serves as a good sanity check for the variables and data we collected.

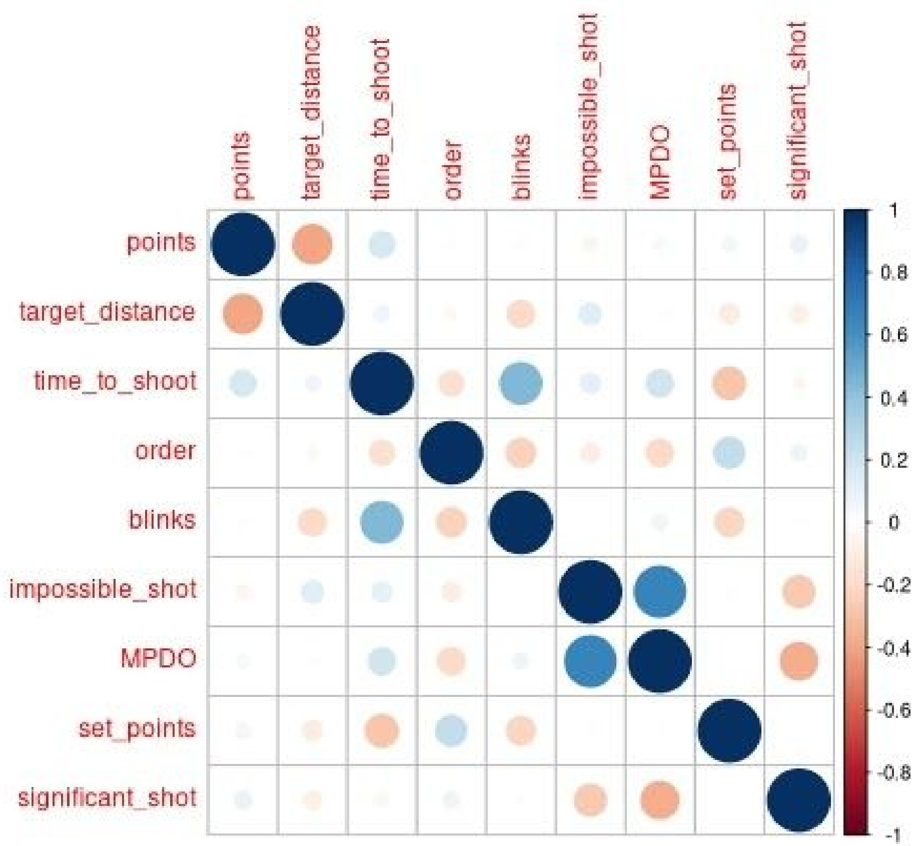
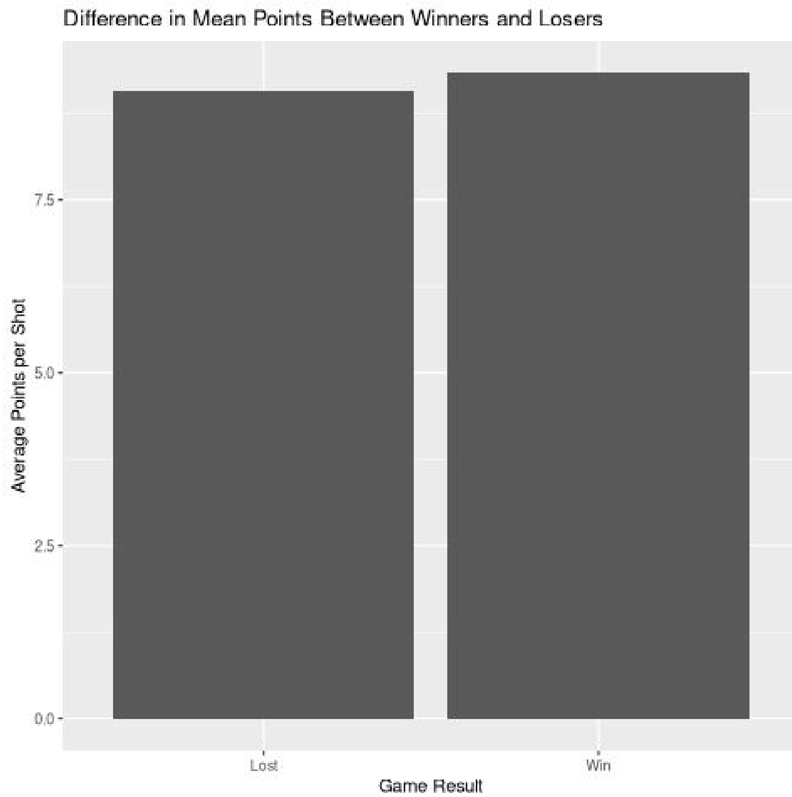
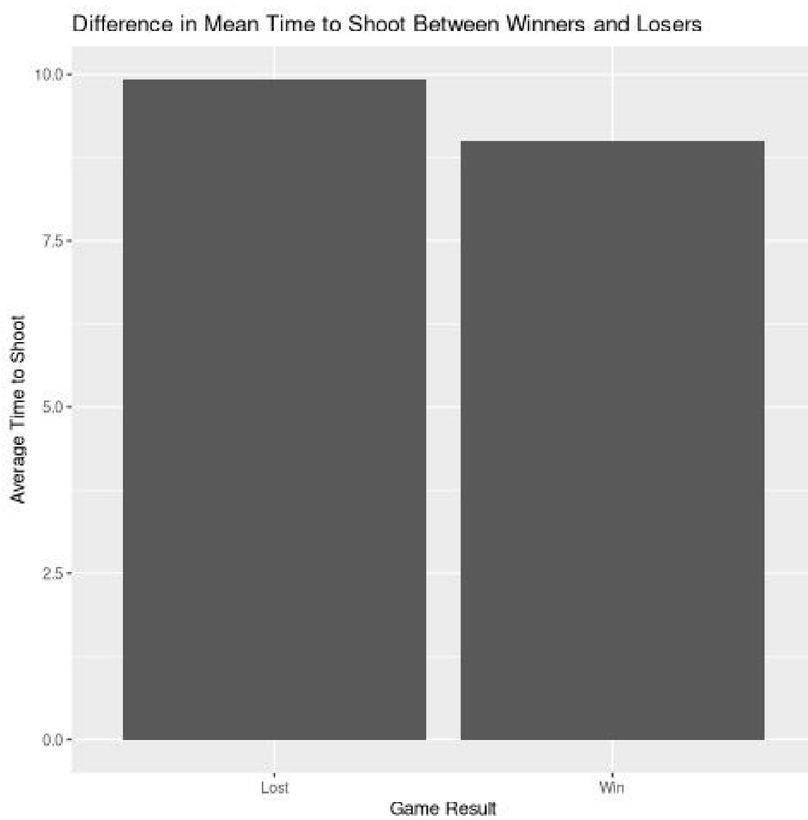
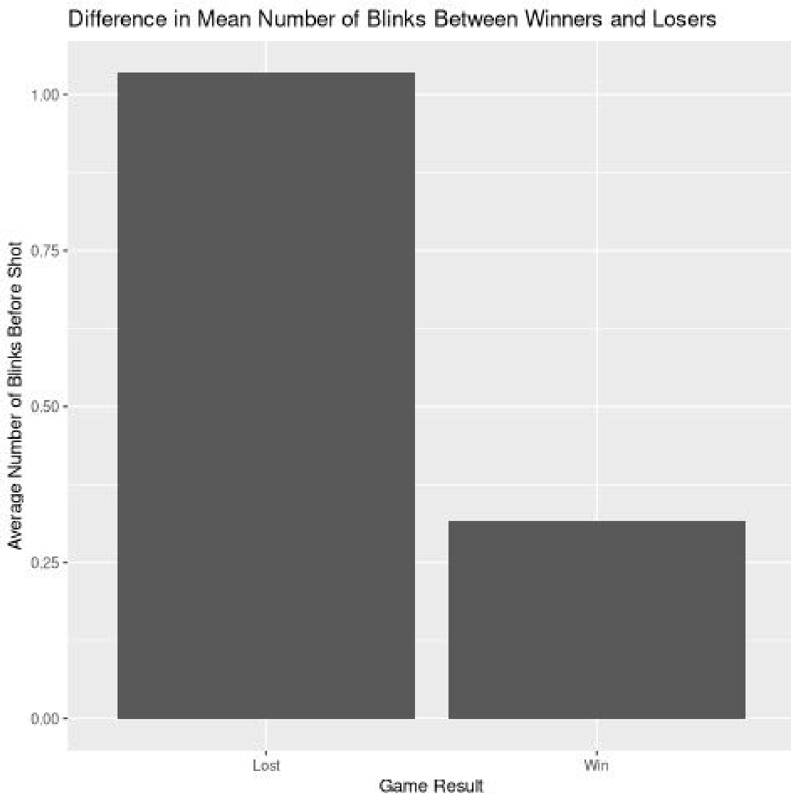


Figure 2: As mentioned in the introduction, one of the main focuses of this study is to understand the factors that define success in the game of recurve archery. Thus, we studied a number of shot metrics and how they differ between players who won and players who lost in the data that we collected. The bar plots shown below include some relevant results:


Difference in Mean Points Between Winners and Losers
Average Points per Shot
7.5
5.0
2.5
0.0
Lost
Win
Game Result


Difference in Mean Time to Shoot Between Winners and Losers
Average Time to Shoot
10.0
7.5
5.0
2.5
0.0
Lost
Win
Game Result



As we can see, players that won had a higher average Points per Shot, a shorter average Time to Shoot, and a much lower average Number of Blinks.

Figure 3.1: This figure displays R summary output of a linear regression on Time To Shoot, Eyes Open, and Blinks.

```
Call:
lm(formula = Points ~ `Time to Shoot (s)` + `Eyes Open` + Blinks +
    Gender, data = DATA)

Residuals:
    Min      1Q  Median      3Q     Max
-2.7744 -0.6636  0.1898  0.4830  1.2745

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          8.58407    0.36784  23.336  < 2e-16 ***
`Time to Shoot (s)`  0.04887    0.02572   1.900   0.0599 .
`Eyes Open`2        -0.64490    0.15761  -4.092 8.06e-05 ***
Blinks               0.06697    0.08714   0.769   0.4438
GenderMale           0.59084    0.24331   2.428   0.0167 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7902 on 113 degrees of freedom
Multiple R-squared:  0.1843,    Adjusted R-squared:  0.1555
F-statistic: 6.385 on 4 and 113 DF,  p-value: 0.0001142
```

Figure 3.2: This figure displays R summary output of the updated linear model after the stepwise process was applied in order to find out which variables are worth keeping while minimizing the utility of the original.

```
Call:
lm(formula = Points ~ `Time to Shoot (s)` + `Eyes Open` + Gender,
    data = DATA)

Residuals:
    Min      1Q  Median      3Q     Max
-2.7883 -0.6226  0.1539  0.4892  1.2676

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          8.61006    0.36563  23.549  < 2e-16 ***
`Time to Shoot (s)`  0.05588    0.02400   2.328   0.0217 *
`Eyes Open`2        -0.61080    0.15097  -4.046 9.52e-05 ***
GenderMale           0.50965    0.21879   2.329   0.0216 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7888 on 114 degrees of freedom
Multiple R-squared:  0.1801,	Adjusted R-squared:  0.1585
F-statistic: 8.346 on 3 and 114 DF,  p-value: 4.592e-05
```

Figure 4.1: This figure displays R summary output for a multiple linear regression performed on the metrics (Order, MPDO, Impossible Shot, and Significant Shot) as predictors for Points.

```
lm(formula = points ~ order + MPDO + impossible_shot + significant_shot,
    data = archery)

Residuals:
    Min      1Q  Median      3Q     Max
-3.2339 -0.3126 -0.0850  0.7661  1.0639

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       7.71019    1.11141   6.937 2.12e-10 ***
order             0.01733    0.15578   0.111    0.912
MPDO              0.14890    0.10642   1.399    0.164
impossible_shot  -0.35067    0.26394  -1.329    0.186
significant_shot  0.20622    0.19179   1.075    0.284
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8544 on 121 degrees of freedom
Multiple R-squared:  0.02536,	Adjusted R-squared:  -0.006861
F-statistic: 0.7871 on 4 and 121 DF,  p-value: 0.5357
```

Figure 4.2: The figure displays a two-sample t-test done in R.

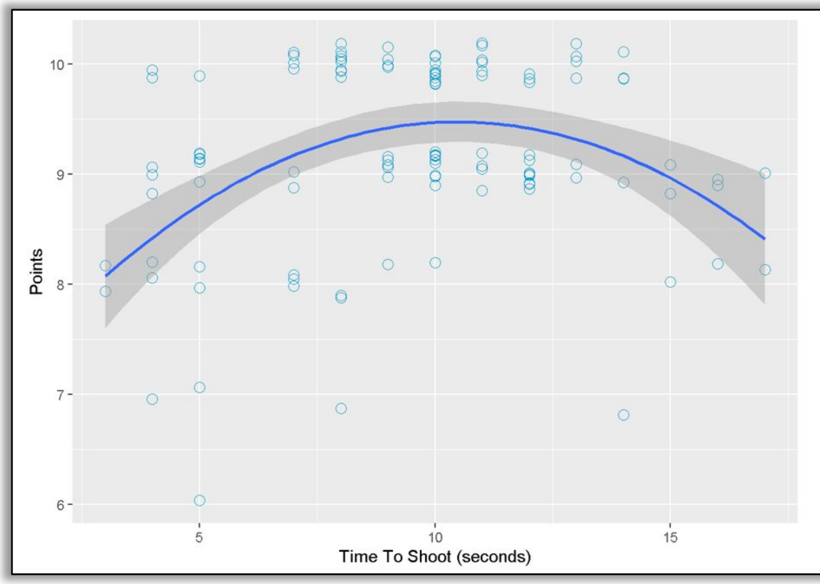
```
	Welch Two Sample t-test

data:  filter(archery, order == 1)$points and filter(archery, order == 2)$MPDO
t = -3.271, df = 116.35, p-value = 0.00141
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.9174256 -0.2254315
sample estimates:
mean of x mean of y
 9.206349  9.777778
```

INSIGHTS

Factors for Predicting Shots:

Apart from basic derivation from the dataset shown in Table 1, like the average point of all the shots being 9.21, we dive into finding out what exactly contributes to the points of every shot. There are two ways of going about this: Simple Linear Regression with Points on Time to Shoot, Eyes Open and Blinks, or Multiple Linear Regression with those same variables all taken into account. The reason for the many Simple Linear Regressions is due to the nature of Multiple Linear Regressions, which is the desire to make use of all variables whether they were actually helpful or not. To start this process off, we began with making the 3 SLRs for Time to Shoot, Eyes Open, Blinks and getting p-values of 0.049, 6.3e-05, and 0.889, respectively. This leads us to conclude that using the number of blinks by themselves to gauge a shot's point is not enough, especially with a negative adjusted r-squared. Although this might be helpful, we need to remember that the game of archery isn't based solely on these factors but the combination of them. So we move on to do a multi-linear regression. Deriving an adequate p-value and a model that can confidently make up roughly 12% of actual results, we realize that it is only when the independent variables are used together that they are *truly* significant. We can read in this summary that having both eyes open actually contributes to a lower point average! Although even after this analysis, the independent variable Blinks still has a p-value too high to prove that blinks are worthy of being included. To tackle this, we use stepwise selection in R to simplify the current model which leads us to our second output.

This second model (Figure 3.2) constitutes the better picture and not only simplifies our model (Figure 3.1) but also improves it ever so slightly by encompassing 0.0030 more of results that would be to come in the future. It is surprising that despite many expert's opinions on shooting with one/two eyes, analysis of this data seems to tell us that shooting with one eye leads to a better shot. At the end of that day, it is critical to remember that all sports, including archery, have varying levels in the skill of their players. Not having a large amount of data on both varying skill levels and gender ultimately might lead us to think that males are better at archery even though this could be far from the truth. We also conclude that with every second that goes by, an archer has a higher likelihood of getting a larger point in his or her next shot with this data's preference having that second fall in between 8-12. This is portrayed in the next figure.



Minimum Points to Desirable Outcome (MPDO):

In this section, we take a deeper dive into the game of archery from the players' perspective. We surmise that archery is a game of mental aptitude more so than physical talent. As such, we set out to find a way to quantify different factors that affect a player's mental state from simple distractions to more intricate stress factors. From this line of reasoning, we defined the following new metrics:

1. Desirable Outcome: Binary variable for an individual shot that can either be *tie* or *lead* depending on the player's current total points compared to their opponent's within a set. Note that it is sometimes desirable to tie the set rather than lead in it because that is simply the best that a player can do in that situation.
2. Min Points to Desirable Outcome (MPDO): Numeric variable that indicates how many points are required to reach the Desirable Outcome for a given shot.
3. Impossible Shot: Binary variable for a shot that indicates whether or not it is possible to achieve the MPDO. This metric follows from the fact that MPDO could be greater than 10 (max possible points) in which case it would indicate a game state that is impossible to bounce back from.
4. Significant Shot: Binary variable for a shot that indicates whether or not a shot is particularly significant in the set or game. A shot is categorized as significant if it occurs last in a set and is possible.

In addition to the aforementioned metrics, we also took into consideration the Order in which players shoot in a set (first or second to go). The premise behind this analysis is to determine whether or not Desirable Outcome, MPDO, Impossible Shot, Significant Shot, and Order are adequate metrics for measuring the pressure that a shot puts on an archer and whether that affects the number of points earned.

Our analysis showed that there is in fact no statistically significant evidence that these metrics have an influence on Points. Figure 4.1 displays the results of a multiple linear regression performed on the metrics as predictors for Points.

Furthermore, we learned by performing two-sample t-tests that there is no significant difference between Average Points when comparing whether or not a shot is impossible or significant. While the results may not be the most exciting, they tell us that these measures are not an effective way to measure how much pressure a shot puts on an archer. Perhaps the reason is that players are aware of the mental battle they must engage in when shooting and therefore take any necessary actions to avoid being impacted by being down in points or having to shoot a significant shot. However, our analysis did produce the interesting result found in Figure 4.2.

The two-sample t-test shows a statistically significant difference between the MPDO observed for a shot in which the archer is shooting first versus shooting second. This might be an aftermath of the way MPDO was defined; we assume that the person who shoots first hopes to get more than an opponent who will get a bull's eye when they shoot, i.e. they assume the worst case every time.

CONCLUSION

While there are many interesting insights and patterns that can be drawn from our analysis, it goes without saying that there are also certain aspects of this project that our team believes could have been improved in one way or another. One of these includes having more accurate data. In terms of accuracy, while our data was taken by observations of games provided online, certain variables such as Blinks may not be as concrete as the videos we watched for our observations did not always pan to the archer's face leading up to each shot while they draw their bow and the presence of the archers' hats may have also interfered with this variable by obscuring the archer's face for part of the time leading up to each shot; thus it is impossible to determine whether or not they may have blinked outside of what our team was able to discern from the video of each game.

In addition to this, we wished to be able to gather another variable that many members of our team believe would have proved to be a potential confounding variable and potential factor in shooting accuracy in the form of wind speed. Due to the limitation of only being able to observe what each archery match video provided us, while there may have been a wind speed statistic available at the outdoor tournaments itself, this information was not available to us through the video. We had originally hoped to be able to utilize wind speed as a variable to compare indoor and outdoor performance to determine if wind speed may play an important role in the success of archers at outdoor tournaments but were unable to do so due to these issues.

In order to limit any confounding variables, we had to focus on only certain aspects of the sport. We decided to focus on recurve bow matches instead of matches with compound bows due to the differences in scoring and consistency of the data. Choosing to take down observations of both bow types would have required use to find a method to properly scale or categorize the points earned by each arrow, and would have negatively impacted the analysis of the players attributes and skill. In addition, we found that matches at the professional level were far more available, and more consistent to the professional field. Of the available amerateur matches, most had a large skill gap that interfered with the data analysis.

Our team postulated that, in the future, it would be interesting to analyze and compare the performance of archers by including more variables such as arrow speed, wind speed at outdoor tournaments, target distance, as well as potentially dividing up the target into regions smaller than just the rings so our observations would not only include the number of points each shot earned, but also coordinates where on the target it is located. Adding a variable such as the latter would allow us to better measure the archer's precision throughout a match to see whether their shots regularly land in a specific region grouped together, or tend to have more variability and are spread across the target.

In addition, while our team collected our data at each arrow shot level as one observation, it would also prove insightful to explore data that is collected at a player or even team level in order to draw more meaningful conclusions regarding specific players and what they may be doing well or not as well when compared to other players. This sort of analysis may also prove to be more marketable if we were to consider the archery community as our audience or customer.