

Final report title:

C1X Accuracy as a Predictor of Player Earnings in DGPT 2024

Group ID: B 143

Dataset number: DS350

Prepared by:

Ahmad Khalil 24092160

Srikanth Reddy 24091671

Nafeesh Ur Rehman 24092015

Ajay Bhukya 24091661

Table of Contents

Page numbers

1. Introduction	
1.1. Problem statement and research motivation	3
1.2. The data set	3
1.3. Research question	3
1.4. Null hypothesis and alternative hypothesis (H0/H1)	3
2. Background research	
2.1. Research papers (at least 3 relevant to your topic / DS)	3
2.2. Why RQ is of interest (research gap and future directions according to the literature)	4
3. Visualisation	
3.1. Appropriate plot for the RQ output of an R script (NOT a screenshot) and required supplementary graph/table	4
3.2. Additional information relating to understanding the data	5
3.3. Useful information for the data understanding	5
4. Analysis	
4.1. Statistical test used to test the hypotheses and output	5
4.2. The null hypothesis is rejected based on the p-value	5
5. Evaluation – group’s experience at 7COM1079	
5.1. What went well	6
5.2. Points for improvement	6
5.3. Group’s time management	6
5.4. Project’s overall judgement	6
5.5. Comment on GitHub log output	6
6. Conclusions	
6.1. Results explained.	7
6.2. Interpretation of the results	7
6.3. Reasons and/or implications for future work, limitations	8
7. Reference list	
<i>Harvard (author, date) format</i>	8
8. Appendices	
A. R code used for analysis and visualisation.	8
B. GitHub log output.	9

1. Introduction

1.1. Problem statement and research motivation (100 words)

Professional disc golf performance is driven by a mix of throwing power, accuracy, approach of play, and putting, however it is unclear which skills most significantly predict financial success on the Disc Golf Pro Tour. In traditional golf, strokes-gained and putting metrics have been shown to explain ordinary variation in outcomes and earnings.

According to Broadie M. 2014 (a book), Every shot counts: Using the revolutionary strokes gained approach to improve your golf performance and strategy.

1.2. The data set (75 words)

The dataset contains performance statistics for 97 professional disc golfers from the first 11 events of “Disc Golf Pro Tour season in 2024”. It also includes 25 columns like scoring averages, driving accuracy, approach consistency, putting performance, strokes gained metrics, rounds, and total earnings. Each row represents one player who finished above the cash line in at least one event. For this study, the key variables examined are Circle 1 putting accuracy (C1X) and total earnings.

1.3. Research question (50 words)

“Is there a correlation between Circle 1 putting accuracy (C1X) and total earnings among professional disc golfers in Disc Golf Pro Tour season 2024?”

1.4. Null hypothesis and alternative hypothesis (H0/H1) (100 words)

- **Null hypothesis (H₀)**

There is no correlation between Circle 1 putting accuracy (C1X) and total earnings among professional disc golfers in Disc Golf Pro Tour season 2024?

- **Alternative hypothesis (H₁)**

There is a correlation between Circle 1 putting accuracy (C1X) and total earnings among professional disc golfers in Disc Golf Pro Tour season 2024?

2. Background research

2.1. Research papers (at least 3 relevant to your topic / DS) (200 words)

- I. Broadie, M. (2014). Every Shot Counts. Penguin.

- In this book Author Broadie said Using the revolutionary strokes gained approach to improve your golf performance and strategy.

- II. Do kinematic study assessments improve accuracy & precision in golf putting? A comparison between elite and amateur golfers: A systematic review and meta-analysis (Merry et al., 2022)
 - This paper reviews 25 studies in which comparing putting performance (like accuracy and precision) between elite and amateur golfers.
- III. Assessing Golfer Performance on the PGA Tour (Broadie, 2011).
 - In this paper Mark Broadie (Author) from Columbia university wrote that Strokes Gained method which is a system that decomposes golf performance into components (like driving, approach, putting, etc.) and quantifies each shot's contribution compared to the field. This method was highly adopted by most of professional tours to evaluate which aspects of the game most affect scoring and earnings

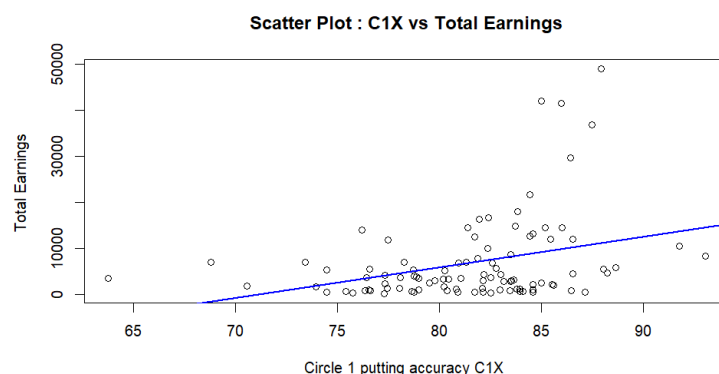
2.2. Why RQ is of interest (research gap and future directions according to the literature) (100 words)

Although putting accuracy is widely recognised as a key performance indicator in traditional golf (Broadie, 2014), This literature shows us that disc golf was significantly under-researched, with some studies and analysing the performance or competitive results. Existing disc golf research are examining biomechanics, throwing distance, or gender base performance differences, but there is no studies which directly investigate whether Circle 1 putting accuracy predicts financial success on the Disc Golf Pro Tour. This creates a clear research gap. Understanding this relationship can help the players and prioritise training, support data-driven coaching strategies, and guide future analytical research in emerging disc golf performance.

3. Visualisation

3.1. Appropriate graphs for the RQ (50 words)

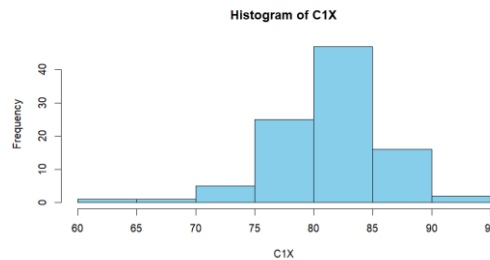
We do check the normality of our data before making a choice about the plot as the result of our test our data independent variable circle 1 putting accuracy (C1X) was normal but total earning was not normal, so we selected the scatter plot for our visualisation.



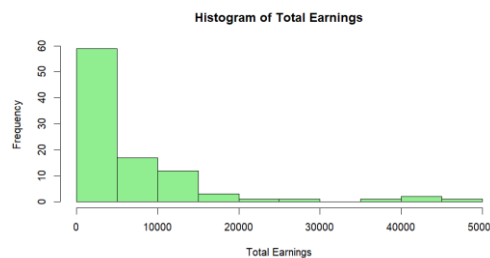
3.2. Additional information relating to understanding the data (optional) (50 words)

We make two extra plots to understand the data

- First is to check the normality of Circle 1 putting accuracy (C1X) which is totally normal.



- The second plot is to check the normality of total earnings which is not normal.



3.3. Useful information for the data understanding (50 words)

From our scatter plot we can easily analyse a weak but noticeable relationship between circle 1 putting accuracy (C1X) and total earnings. Those players who have higher putting accuracy in circle 1 earns more although the spread of points indicates variability. This conclude that other performance factors also make influence on earnings.

4. Analysis

4.1. Statistical test used to test the hypotheses and output (75 words)

As the type of our research question is correlation so we got two options to test which are Spearman which is used for data which is not normal and the second is Pearson for the normal data. As circle 1 putting accuracy (C1X) column has normal data and Total earning column has not normal values so we will use Spearman test.

4.2. The null hypothesis is rejected based on the p-value (100 words)

From our statistical test by using Spearman correlation, we got correlation coefficient $\rho = 0.3056$ with a p-value of 0.0023. Since our p-value is below than 0.05 threshold, so we reject the null hypothesis and conclude that there is

a statistically significant relationship between Circle 1 putting accuracy (C1X) and total earnings of the players in Disc Golf Pro Season 2024. The positive correlation indicates that players with higher C1 putting accuracy tend to earn more money during the season. On the other hand, the strength of this relationship is moderate, that means C1X contributes to earnings, but it is not the sole predictor of financial success in disc golf.

5. Evaluation – group’s experience at 7COM1079

5.1. What went well (75 words)

Approximately every this goes well but when you are working with different students with different backgrounds, it brings some interesting challenges which helps us to understand each other and work on clock and put our best to complete the tasks.

5.2. Points for improvement (75 words)

Well, everyone is not good with GitHub so there is a lot to improve on it. This is the reason that only one person is doing all the commits. On the other hand, English is not our first language so sometimes it is a bit difficult to convey and understand each other. But everyone did their part properly and we completed this task on time and try our best to do it as good as we can.

5.3. Group’s time management (50 words)

To manage our time and do discussions we do weekly meetings on zoom. Most of time it was 1 meeting every week but if someone encountered any problem then we do more sessions as required and discuss what we did, what everyone is doing now and what is left to do in further.

5.4. Project’s overall judgement (50 words)

Firstly, overall judgement is that it goes very well and secondly it took some time to understand the dataset because the data set is about Disc golf and we did not know about that game. It takes couple of days to understand the game, methods and terminologies of game.

5.5. Comment on the GitHub log output (50 words)

The GitHub log output shown in Appendix B provides clear evidence of consistent version control throughout the development of this project. The commit history demonstrates regular updates, including data analysis steps, code refinements, report writing, and corrections based on feedback.

1. **Commit Message:** [As we used the spearman method for statistical testing and we got value of $p=0.002331$ and correlation ρ is 0.3056 so it is clear that our null hypothesis is rejected]

Explanation: This commit documents the key finding of our report. By applying the Spearman correlation test, we established that there is a statistically significant relationship between C1X putting accuracy and total earnings. Rejecting the null hypothesis provides the analytical foundation for the rest of the research, guiding conclusions and supporting the project's overall claim about the performance–earnings relationship.

2. **Commit Message:** [Checking the types of our Column (i.e vector, numeric or integer etc) which are used in our Research Question. As our RQ is no Correlation so we can only use numeric and integer type.]

Explanation: This commit ensures that the dataset is suitable for performing valid statistical analysis. Since correlation tests require numeric or integer variables, confirming the data types prevents errors, avoids incorrect test results, and maintains the integrity of the research process. This step strengthens the reliability of the entire analysis by guaranteeing that only appropriate variables are included.

3. **Commit Message:** [checking the data is normal or not. In our case one variable C1X is normal But total_earning is not normal. So, we use Spearman method.]

Explanation: This commit records an essential diagnostic step that determines the appropriate statistical method for the project. By evaluating the normality of both variables, we ensured that the chosen correlation test aligns with the assumptions of the data. Since total earnings were not normally distributed, using the Spearman method provides a more accurate and reliable measure of association. This decision improves the validity of the final research findings.

6. Conclusions

6.1. Results explained (75 words)

The results we got from Spearman correlation test showed us statistically significant positive relationship between Circle 1 putting accuracy (C1X) and total earnings, with correlation coefficient $\rho = 0.3056$ and $p = 0.0023$. This indicates that the players who putt more accurately from Circle 1 tend to earn higher prize money during the 2024 Disc Golf Pro Tour season. However, the correlation is weak to moderate, which is suggesting that while putting accuracy contributes to financial success, additional performance factors also play an important role.

6.2. Interpretation of the results (75 words)

The results indicate that Circle 1 putting accuracy plays a meaningful, though not dominant, influence on a player's financial success in professional disc golf. It also suggests that improving short-range putting could enhance a player's earnings potential across a season. The findings reinforce the importance of

developing consistent putting skills as part of competitive training. However, the modest correlation also highlights that success depends on multiple performance factors beyond putting alone.

6.3. Reasons and/or implications for future work, limitations of your study (50 words)

The study is limited by because of a single season and two variables, which means that other important performance factors were not analysed. Future work could explore multivariate models across multiple seasons to identify stronger factors of earnings for players. Additional research may also examine psychological, environmental, and course-specific factors on putting performance.

7. Reference list (*not included in the word count*)

Broadie, M. (2014) Every Shot Counts: Using the Revolutionary Strokes Gained Approach to Improve Your Golf Performance and Strategy. New York: Penguin.

Merry, B., Baker, I. and Dutheil, F. (2022) 'Do kinematic study assessments improve accuracy and precision in golf putting? A systematic review and meta-analysis', *Performance Analytics in Applied Human Movement*, 7(2), pp. 115–132.

Giljarhus, K.E., Gooding, C. and Njærheim, P. (2022) 'Disc golf trajectory modelling combining computational fluid dynamics and rigid body dynamics', *Sports Engineering*, 25, Article 9.

Hellström, J. (2009) 'Competitive golf: A review of the most relevant performance factors', *International Journal of Sports Science & Coaching*, 4(3), pp. 295–308.

8. Appendices

A. R code used for analysis and visualisation

```
#Now check that the data is normal or not
hist(DGPT24$C1X, main="Histogram of C1X", xlab="C1X", col="green",
border="black")
hist(DGPT24$total_earnings, main="Histogram of Total Earnings",
xlab="Total Earnings", col="yellow", border="black")

#Now we will do statistical test by using spearman method
cor.test(DGPT24$C1X, DGPT24$total_earnings, method="spearman")

#we will make a scatter plot with a trend line
plot(DGPT24$C1X,
+   DGPT24$total_earnings,
+   main = "Scatter Plot : C1X vs Total Earnings",
+   xlab = "Circle 1 putting accuracy C1X ",
+   ylab = "Total Earnings ",)

#for trend line
abline(lm(total_earnings ~ C1X, data = DGPT24), col = "blue", )
```


B. GitHub log output.

commit f2c778358b6468df29ee2aa09b887c05db1e81b9 (HEAD -> main, origin/main)

Author: Ahmad Khalil <ak24akr@herts.ac.uk>

Date: Thu Dec 11 22:19:06 2025 +0000

adding contributor.csv file

commit 56aa15631c47f719b0088502de423fe3cb438d00

Author: Ahmad Khalil <ak24akr@herts.ac.uk>

Date: Thu Dec 11 00:22:47 2025 +0000

for Visualisation we made a scatter plot with trend line to represent the behaviour of our data

commit e89129278f80af6ed3b704b74c65325f0f130c9f

Author: Ahmad Khalil <ak24akr@herts.ac.uk>

Date: Thu Dec 11 00:00:02 2025 +0000

as we used the spearman method for statistical testing and we got value of $p=0.002331$ and correlation ρ is 0.3056 so it is clear that our null hypothesis is rejected

commit 22d8f5e63b09d32d104b2381e5ec40f8578b040d

Author: Ahmad Khalil <ak24akr@herts.ac.uk>

Date: Wed Dec 10 00:54:47 2025 +0000

checking the data is normal or not. In our case one variable C1X is normal But total_earning is not normal. So, we use Spearman method.

commit dc2f2939c00e40d8b37141c69cf7d5275fb48171

Author: Ahmad Khalil <ak24akr@herts.ac.uk>

Date: Tue Dec 9 23:35:56 2025 +0000

Add files via upload

commit 60d26acba4a5464301217627918659e628d761e0

Author: Ahmad Khalil <ak24akr@herts.ac.uk>

Date: Tue Dec 9 23:34:22 2025 +0000

Removing un realistic data from our columns and also test the data that is it normal or not

commit a7fe5c77f14c0f5ce488dcfd638549e5727c2ddd

Author: Ahmad Khalil <ak24akr@herts.ac.uk>

Date: Tue Dec 9 23:31:24 2025 +0000

Describe your changes here

commit 20ab72d5c2f4263615aed5f596b8c82a54254f02
Author: Ahmad Khalil <ak24akr@herts.ac.uk>
Date: Sun Dec 7 19:55:21 2025 +0000

OUR first Column C1X which is numeric and Our Second column
total_earnings are integer type.
commit d26550204ccdfce59288eacde7803939c9806327
Author: Ahmad Khalil <ak24akr@herts.ac.uk>
Date: Sun Dec 7 19:49:55 2025 +0000

Checking the types of our Column (i.e vector, numeric or integer etc) which
are used in our Research Question. As our RQ is on Correlation so we can
only use numeric and integer type.

commit e7dd6bb019181e836224cc18e2255c060c0fc35c
Author: Ahmad Khalil <ak24akr@herts.ac.uk>
Date: Thu Dec 4 00:07:33 2025 +0000

change column names which are possible to get better understanding of data

commit a6b35bb878f9a5ce5643dfcb58566980d3d476bc
Author: Ahmad Khalil <ak24akr@herts.ac.uk>
Date: Tue Dec 2 22:16:24 2025 +0000

Add files via upload

commit bbe62cb43c0292cf697fb4d5a39570714967a022
Author: Ahmad Khalil <ak24akr@herts.ac.uk>
Date: Tue Dec 2 22:12:21 2025 +0000

This is our dataset with all the values
(END)