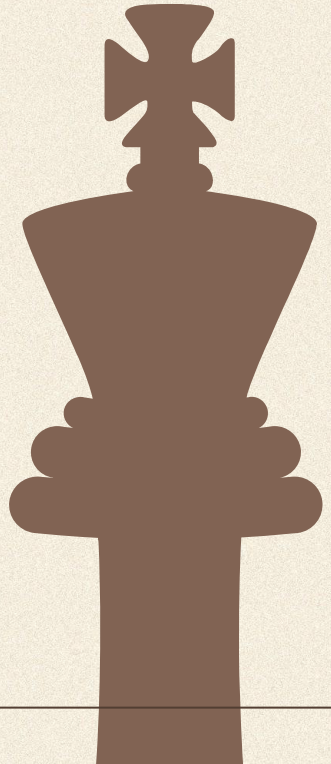
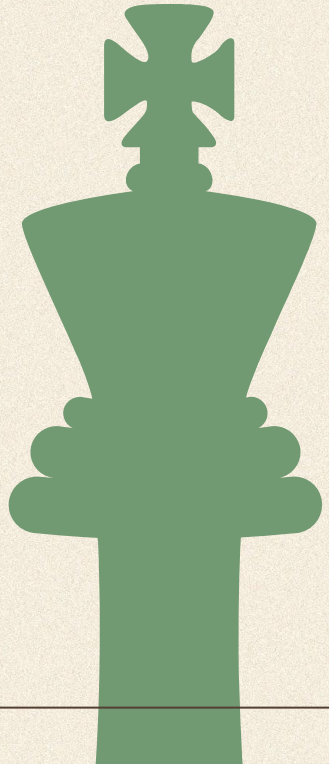
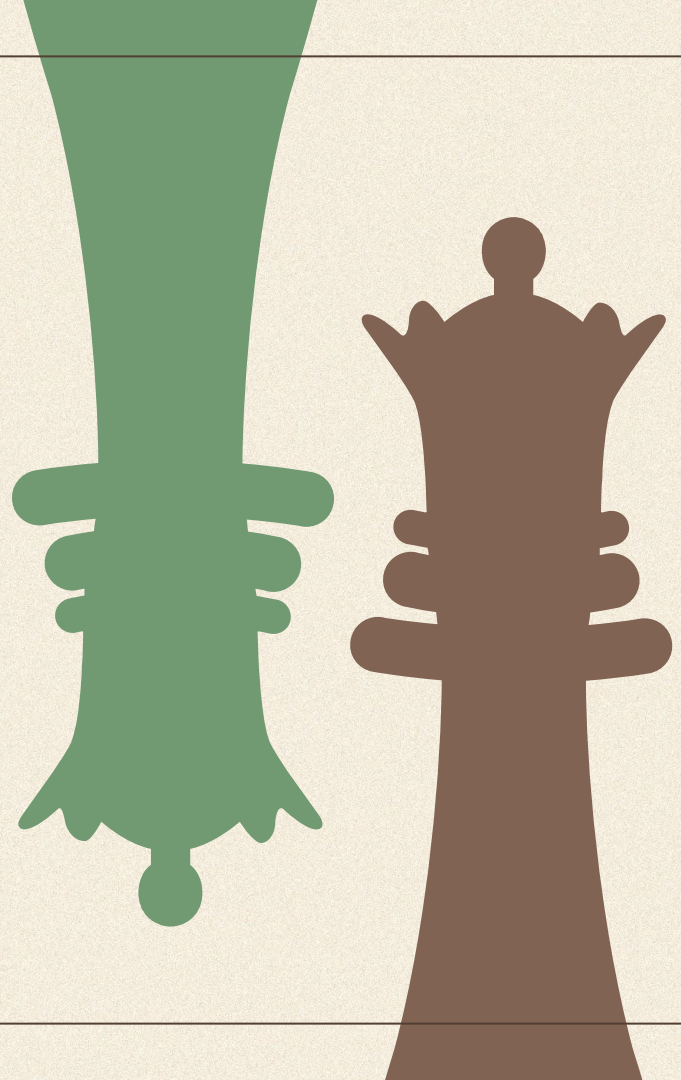
Investigating the Niemann Cheating Scandal

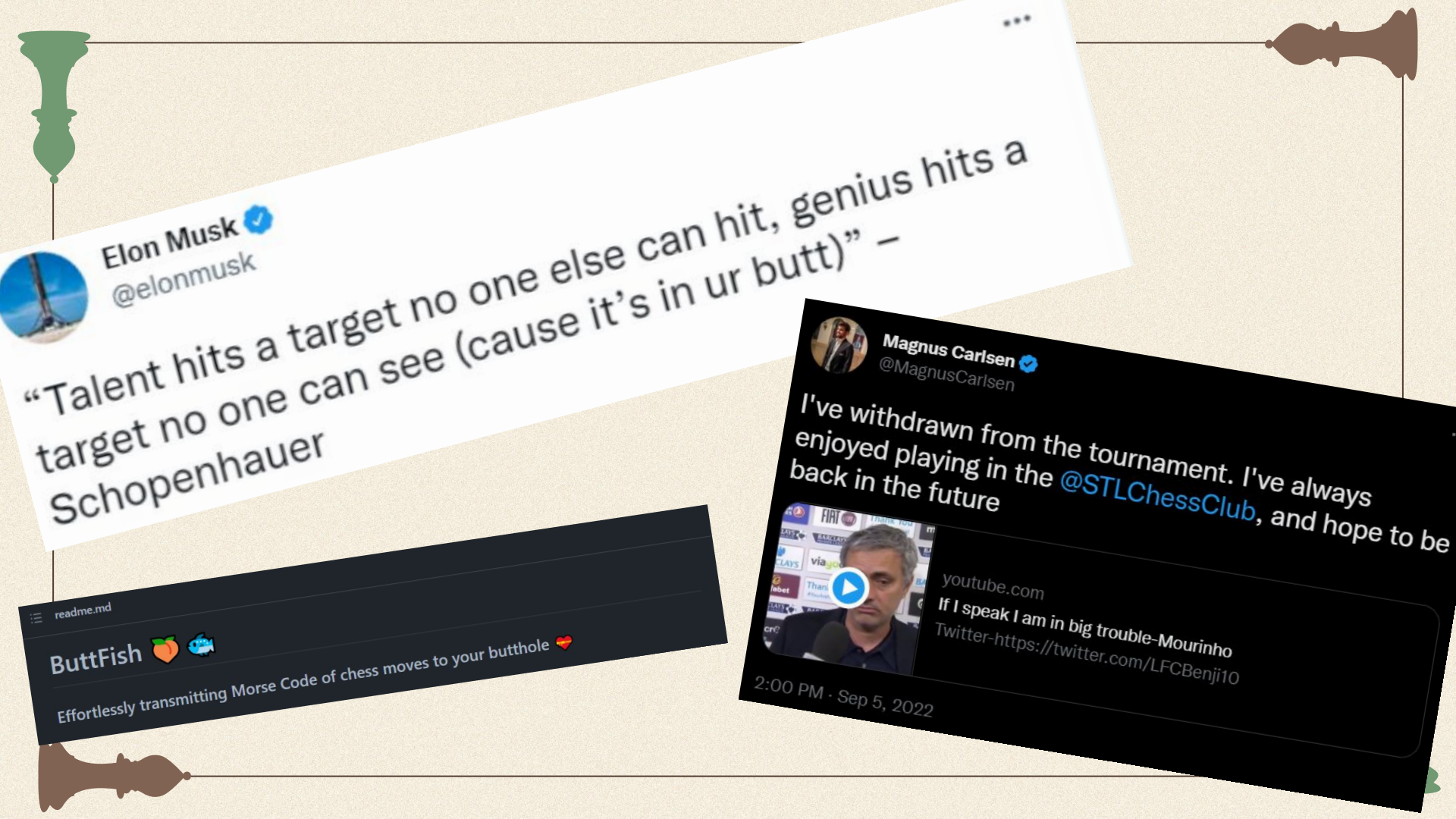
Predicting Performance in Chess via Regression and Statistical Inference Methodologies

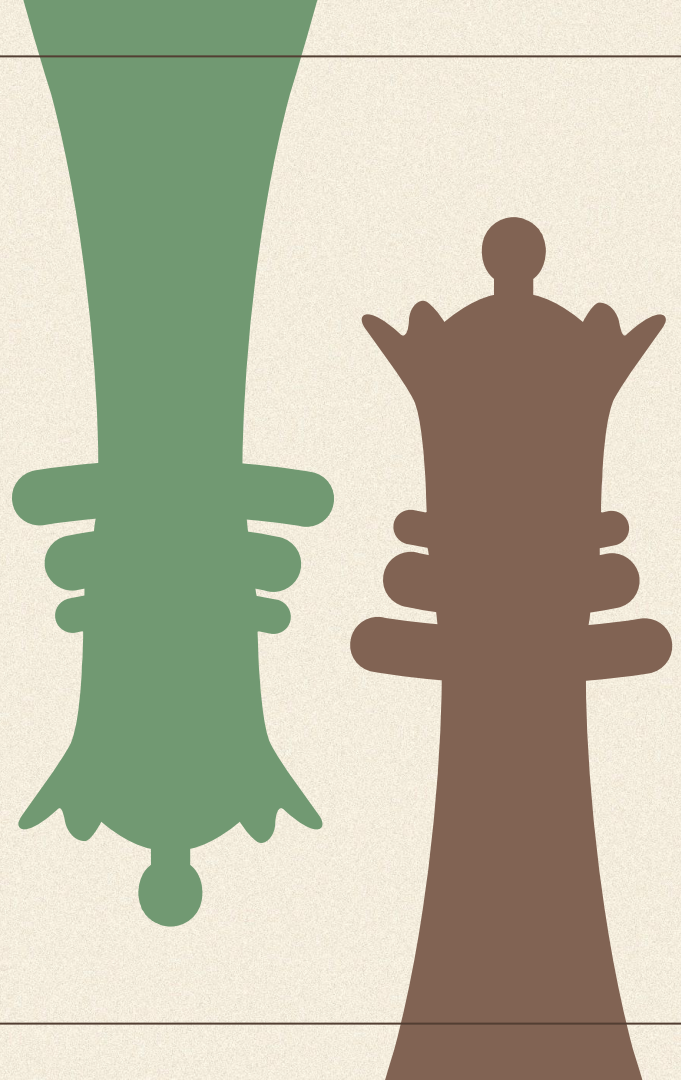






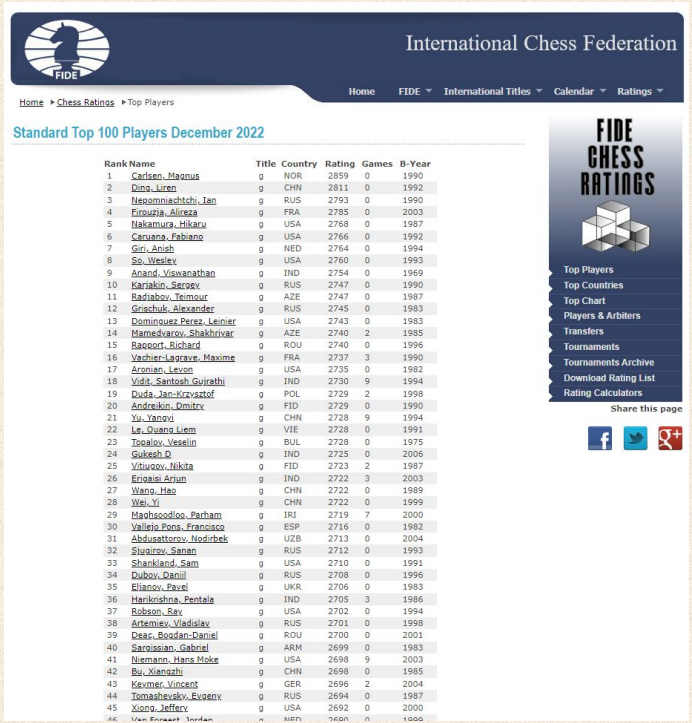
Background


Elon Musk
@elonmusk
“Talent hits a target no one else can hit, genius hits a target no one can see (cause it’s in ur butt)” – Schopenhauer
readme.md
ButtFish
Effortlessly transmitting Morse Code of chess moves to your butthole
Magnus Carlsen
@MagnusCarlsen
I've withdrawn from the tournament. I've always enjoyed playing in the @STLChessClub, and hope to be back in the future
youtube.com
If I speak I am in big trouble-Mourinho
Twitter-https://twitter.com/LFCBenji10
2:00 PM · Sep 5, 2022



Data

Data Gathering



International Chess Federation

Home ▸ Chess Ratings ▸ Top Players

Home | FIDE | International Titles | Calendar | Ratings

Standard Top 100 Players December 2022

Rank	Name	Title	Country	Rating	Games	B-Year
1	Carlsen, Magnus	g	NOR	2859	0	1990
2	Ding, Liren	g	CHN	2811	0	1992
3	Nepomniachtchi, Ian	g	RUS	2793	0	1990
4	Firouzja, Alireza	g	FRA	2785	0	2003
5	Nakamura, Hikaru	g	USA	2768	0	1987
6	Caruana, Fabiano	g	USA	2766	0	1992
7	Giri, Anish	g	NED	2764	0	1994
8	So, Wesley	g	USA	2760	0	1993
9	Anand, Viswanathan	g	IND	2754	0	1969
10	Karjakin, Sergey	g	RUS	2747	0	1990
11	Radjabov, Teimour	g	AZE	2747	0	1987
12	Grischuk, Alexander	g	RUS	2745	0	1983
13	Dominguez Perez, Leinier	g	USA	2743	0	1983
14	Mamedyarov, Shakhriyar	g	AZE	2740	2	1985
15	Rapoport, Richard	g	ROU	2740	0	1996
16	Vachier-Lagrave, Maxime	g	FRA	2737	3	1990
17	Aronian, Levon	g	USA	2735	0	1982
18	Vidit, Santosh Gujrathi	g	IND	2730	9	1994
19	Duda, Jan-Krzysztof	g	POL	2729	2	1998
20	Andreikin, Dmitry	g	FID	2729	0	1990
21	Yu, Yangyi	g	CHN	2728	9	1994
22	Le, Quang Liem	g	VIE	2728	0	1991
23	Topalov, Veselin	g	BUL	2728	0	1975
24	Gukesh D	g	IND	2725	0	2006
25	Vitiugov, Nikita	g	FID	2723	2	1987
26	Erigaisi Arjun	g	IND	2722	3	2003
27	Wang, Hao	g	CHN	2722	0	1989
28	Wei, Yi	g	CHN	2722	0	1999
29	Maghsoodloo, Parham	g	IRI	2719	7	2000
30	Vallejo Pons, Francisco	g	ESP	2716	0	1982
31	Abdusattorov, Nodirbek	g	UZB	2713	0	2004
32	Sjugirov, Sanan	g	RUS	2712	0	1993
33	Shankland, Sam	g	USA	2710	0	1991
34	Dubov, Daniil	g	RUS	2708	0	1996
35	Eljanov, Pavel	g	UKR	2706	0	1983
36	Harikrishna, Pentala	g	IND	2705	3	1986
37	Robson, Ray	g	USA	2702	0	1994
38	Artemiev, Vladislav	g	RUS	2701	0	1998
39	Deac, Bogdan-Daniel	g	ROU	2700	0	2001
40	Sargissian, Gabriel	g	ARM	2699	0	1983
41	Niemann, Hans Moke	g	USA	2698	9	2003
42	Bu, Xiangzhi	g	CHN	2698	0	1985
43	Keymer, Vincent	g	GER	2696	2	2004
44	Tomashevsky, Evgeny	g	RUS	2694	0	1987
45	Xiong, Jeffery	g	USA	2692	0	2000
46	Van Foreest, Jorden	g	NED	2690	0	1999

FIDE CHESS RATINGS

- Top Players
- Top Countries
- Top Chart
- Players & Arbiters
- Transfers
- Tournaments
- Tournaments Archive
- Download Rating List
- Rating Calculators

Share this page

Download PGN Files

Download | Purchase | PGN Mentor | Contact Us

View the chess games from these files with PGN Mentor, the very best program available for working with PGN files.

Portable Game Notation (PGN) is the most popular standard for the representation of chess games. PGN is designed for ease of reading and writing by humans as well as computer programs. The files below are available for download, completely free. Enjoy!

Multiple file downloads are available with the registered version of PGN Mentor.

Players updated: August 2022
Openings updated: January 2022

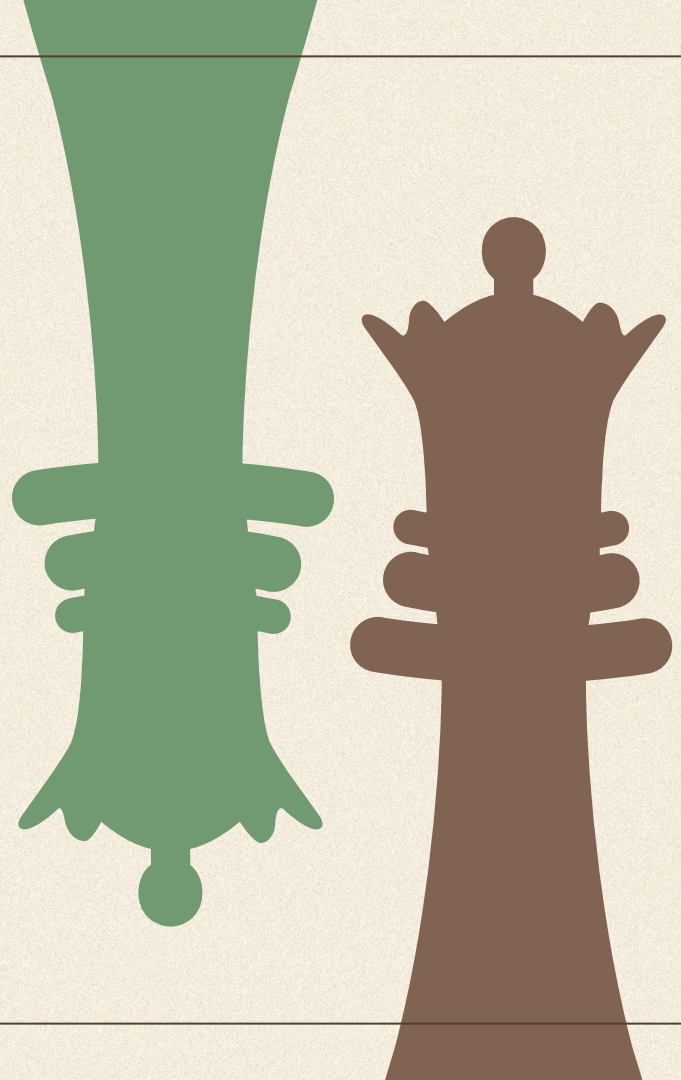
Players | Openings | Events

Players

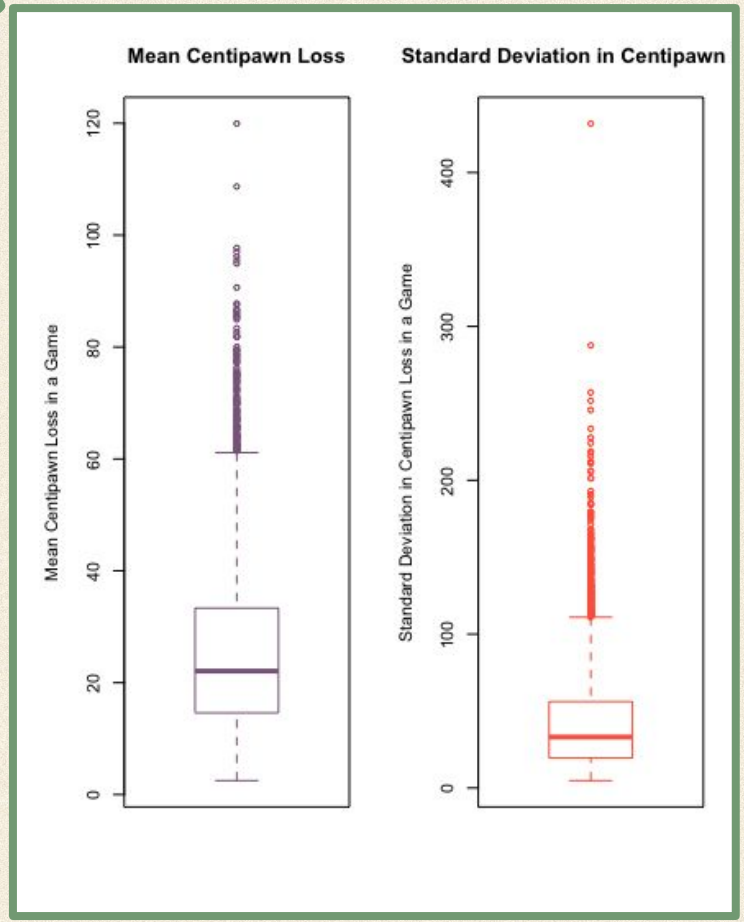
A-L		M-Z	
Adams.pgn Download View	Michael Adams, 3380 games	MacKenzie.pgn Download View	George MacKenzie, 198 games
Akobian.pgn Download View	Varuzhan Akobian, 1429 games	Malakhov.pgn Download View	Vladimir Malakhov, 1973 games
Akopian.pgn Download View	Vladimir Akopian, 1957 games	Mamedyarov.pgn Download View	Shakhriyar Mamedyarov, 4029 games

Resulting Dataset

- **Data**
  - Contains information on over-the-board chess games played by 31 players
- **Key Variables**
  - Elo - player's Elo (ranking) before the game
  - Opponent Elo - the Elo of the opponent the player is playing against
  - Mean Centipawn Loss - number of hundredths of a pawn by which a player deviated from the most accurate move calculated by a computer averaged over all moves in the game
  - Standard Deviation in Centipawn Loss - standard deviation in the number of hundredths of a pawn by which a player deviated from the most accurate move calculated by a computer for all moves in the game

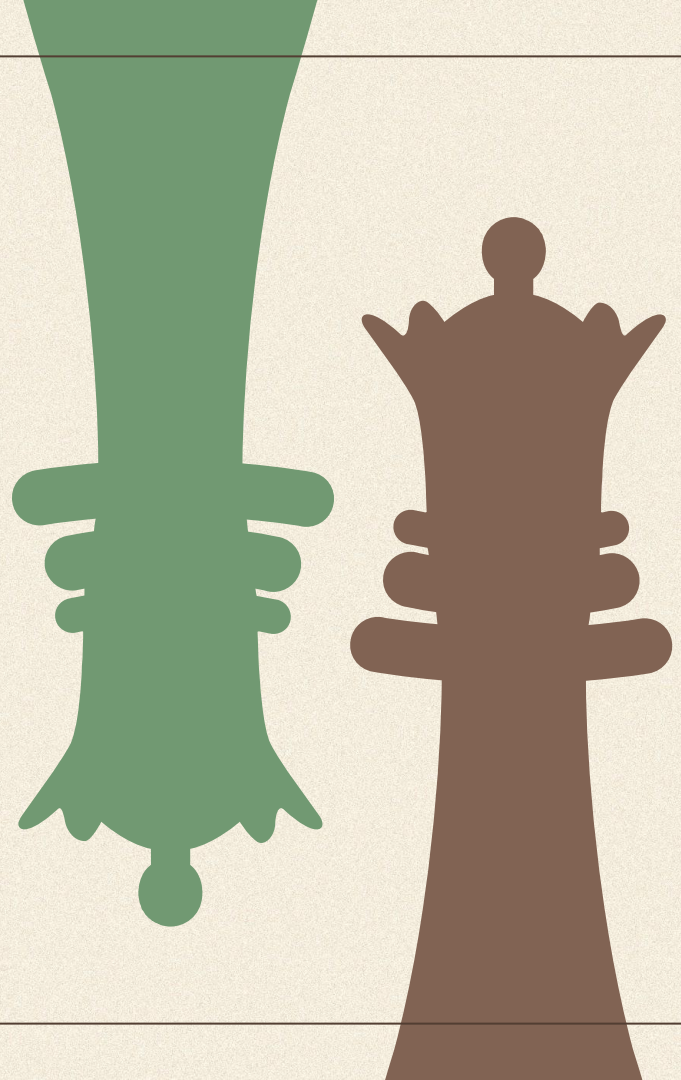


EDA



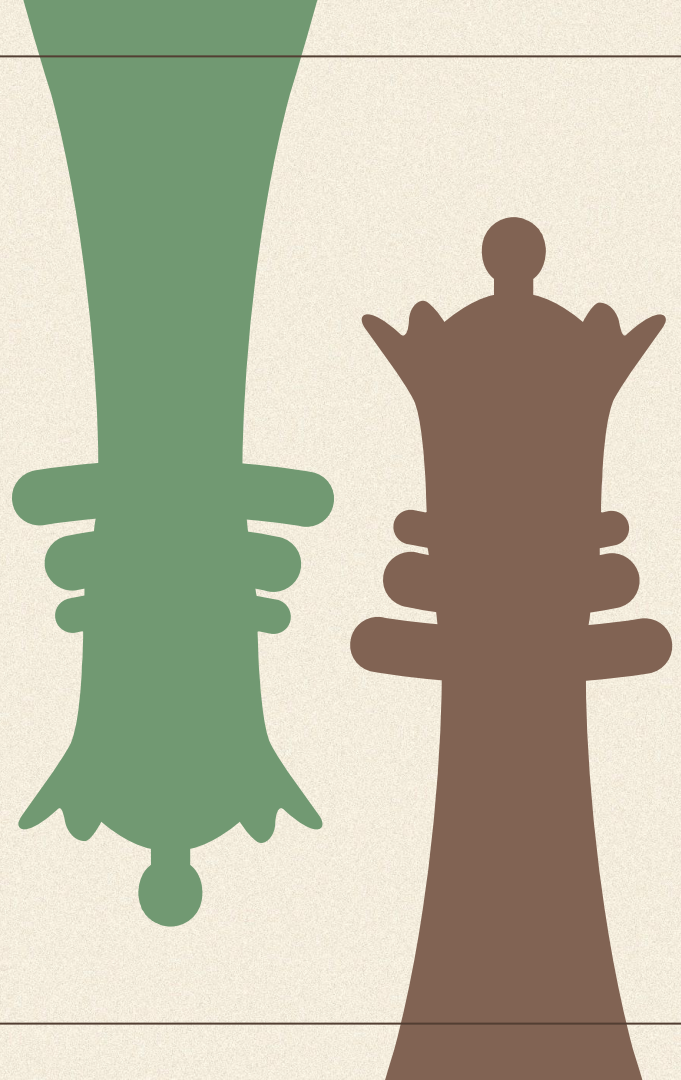
Exploring Centipawn Loss

- Mean Centipawn Loss ranges from 2 CP to 119 CP
  - Centered around CP of 14 to 33
- Standard Deviation in Centipawn Loss ranges from 4.77 CP to 431.85 CP
  - Centered around CP of 19.52 to 56.15



01

Question 1



What are the best predictors for how well and consistently a chess player performs?

Linear Regression w/ Outliers

Normalizing response variables by applying a natural log transformation:

Mean CP = **ln(**Mean CP**)**

SD CP = **ln(**Std CP**)**

With the OLS step forward regression method (provided by the olsrr library), we can select the subset of variables to create the best linear regression models predicting Mean CP and SD CP from the following full models:

$$\text{MeanCP}_i = \beta_0 + \beta_1 \cdot \text{StdCP}_i + \beta_2 \cdot \text{Elo}_i + \beta_3 \cdot \text{OppElo}_i + \beta_4 \cdot \text{Age OR Time}_i + \varepsilon_i$$

where $\varepsilon_i \in N(0, \sigma^2)$ iid

$$\text{StdCP}_i = \beta_0 + \beta_1 \cdot \text{MeanCP}_i + \beta_2 \cdot \text{Elo}_i + \beta_3 \cdot \text{OppElo}_i + \beta_4 \cdot \text{Age OR Time}_i + \varepsilon_i$$

where $\varepsilon_i \in N(0, \sigma^2)$ iid

Linear Regression w/ Outliers

Optimal regression equation with step forward variable selection (best model chosen based on R^2, Predicted R^2, Adjusted R^2, AIC, Mallow's Cp):

$$\text{MeanCP} = 0.996 + 0.745 \cdot \text{StdCP} - 0.0002 \cdot \text{Elo} - 3.7 \cdot 10^{-5} \cdot \text{OppElo}$$

Adjusted-R^2 = 0.8562

F-statistic p-value < $2.2 \cdot 10^{-16}$

t-value=9.247
$p < 2 \cdot 10^{-16}$ ***

t-value=-3.514
p = 0.000447 **

NOT SIGNIFICANT

INTERPRETATION:

On average, given identical player Elos, with every 1% increase in Std CP, Mean CP increases by $(1.01^{0.745} - 1) * 100$, or 0.744%, and on average, given identical Std CPs, with every point increase in player Elo, Mean CP increases by $(e^{-0.0002} - 1) * 100$, or -0.016%.

Linear Regression w/o Outliers

Repeating the same model selection process after removing Mean CP and Std CP outliers from the dataset:

$$\text{MeanCP} = 0.97 + 0.74 \cdot \text{StdCP} - 1.59 \cdot 10^{-4} \cdot \text{Elo} - 2.35 \cdot 10^{-5} \cdot \text{OppElo}$$

Adjusted-R^2 = 0.8568

F-statistic p-value < $2.2 \cdot 10^{-16}$

t-value = 156.523
p < $2 \cdot 10^{-16}$ ***

t-value = -3.503
p = 0.00046 ***

NOT SIGNIFICANT

INTERPRETATION:

On average, given identical player Elos, with every 1% increase in Std CP, Mean CP increases by $(1.01^{0.74} - 1) * 100$, or 0.742%, and on average, given identical Std CPs, with every 1 point increase in player Elo, Mean CP increases by $(e^{-0.000159} - 1) * 100$, or -0.016%.

Linear Regression w/ Outliers

Optimal regression equation with step forward variable selection (best model chosen based on R^2, Predicted R^2, Adjusted R^2, AIC, Mallow's Cp):

$$\text{StdCP} = -0.42 + 1.14 \cdot \text{MeanCP} - 1.20 \cdot 10^{-4} \cdot \text{Elo} - 3.27 \cdot 10^{-5} \cdot \text{OppElo}$$

Adjusted-R^2 = 0.8564

F-statistic p-value < $2.2 \cdot 10^{-16}$

MeanCP: t-value = 156.622, $p < 2 \cdot 10^{-16}$ ***

Elo: t-value = 2.175, p = 0.0297 *

OppElo: NOT SIGNIFICANT

INTERPRETATION:

On average, given identical player Elos, with every 1% increase in Mean CP, Std CP increases by $(1.01^{1.14} - 1) * 100$, or 1.141%, and on average, given identical Mean CPs, with every 1 point increase in player Elo, Std CP increases by $(e^{-0.0002} - 1) * 100$, or 0.012%.

Linear Regression w/o Outliers

Repeating the same model selection process after removing Mean CP and Std CP outliers from the dataset:

$$\text{StdCP} = -0.404 + 1.146 \cdot \text{MeanCP} - 1.03 \cdot 10^{-4} \cdot \text{Elo} - 3.85 \cdot 10^{-5} \cdot \text{OppElo}$$

Adjusted-R^2 = 0.8577

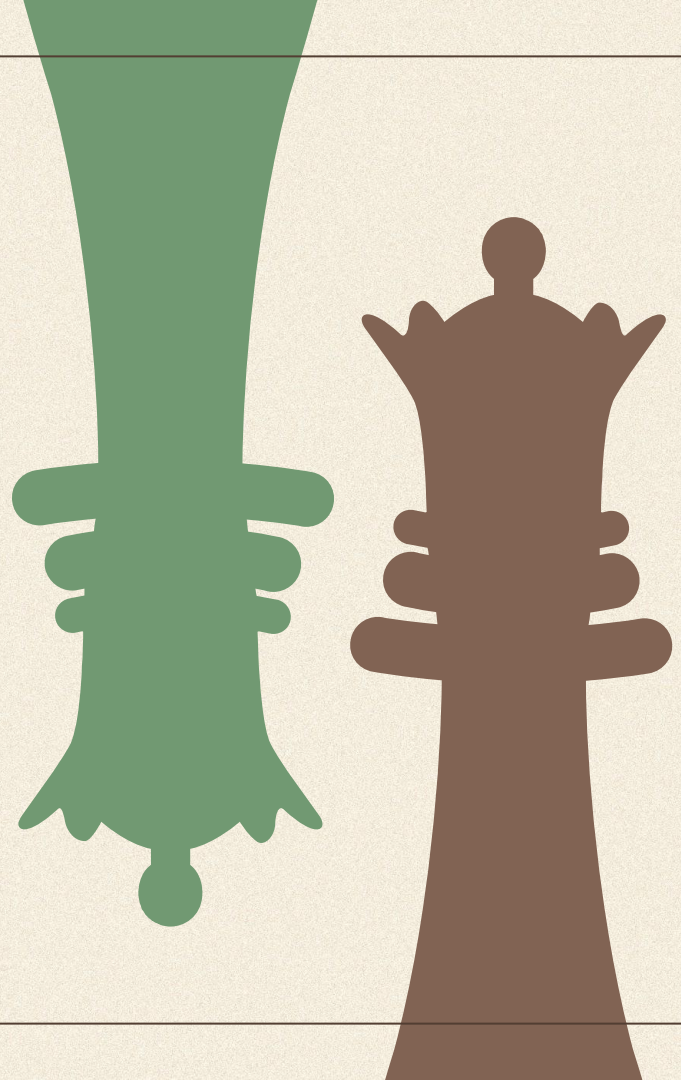
F-statistic p-value < $2.2 \cdot 10^{-16}$

t-value = 156.622
$p < 2 \cdot 10^{-16}$ ***

NOT SIGNIFICANT

NOT SIGNIFICANT

INTERPRETATION:
On average, with every 1% increase in Mean CP, Std CP increases by $(1.01^{1.146} - 1) * 100$, or 1.147%.



02

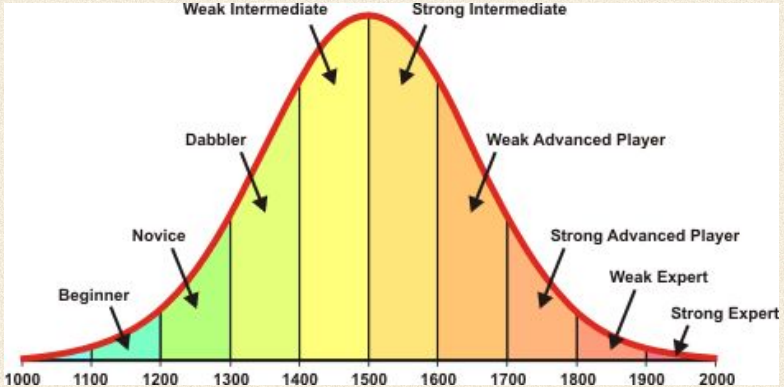
Question 2

Methods for Question 2

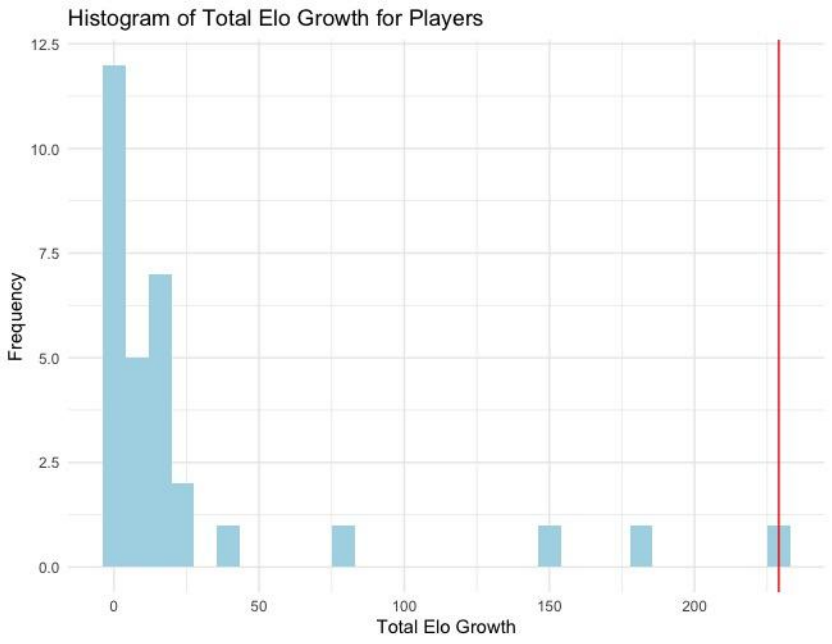
Question: How does Niemann's growth compare to other Grandmasters?

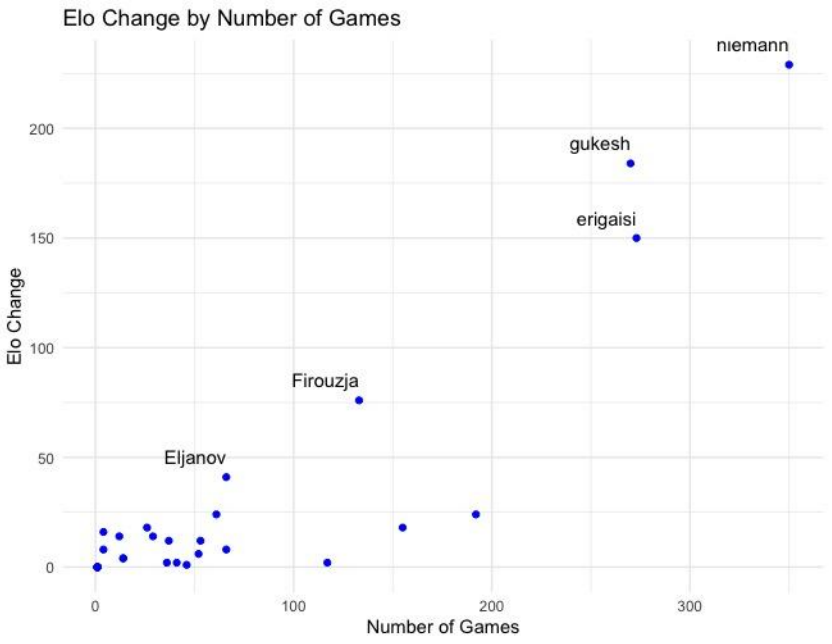
Methods

- Elo as a measure of performance
- Take Maximum Elo vs First Elo
- Compare Niemann's Elo Change to Other Players

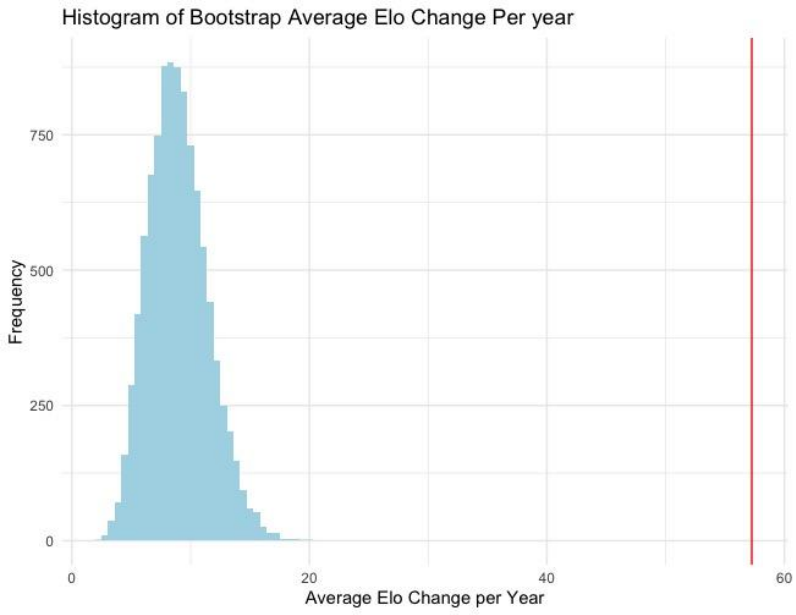


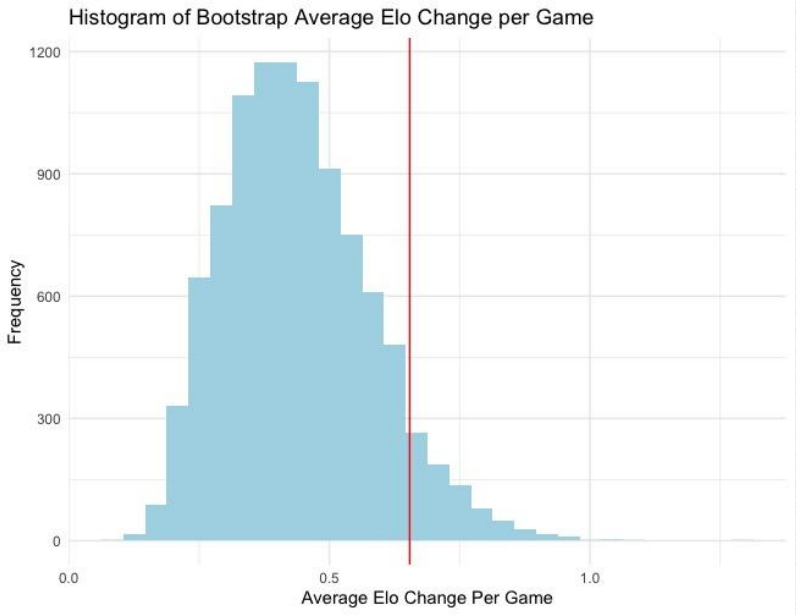
Visualization





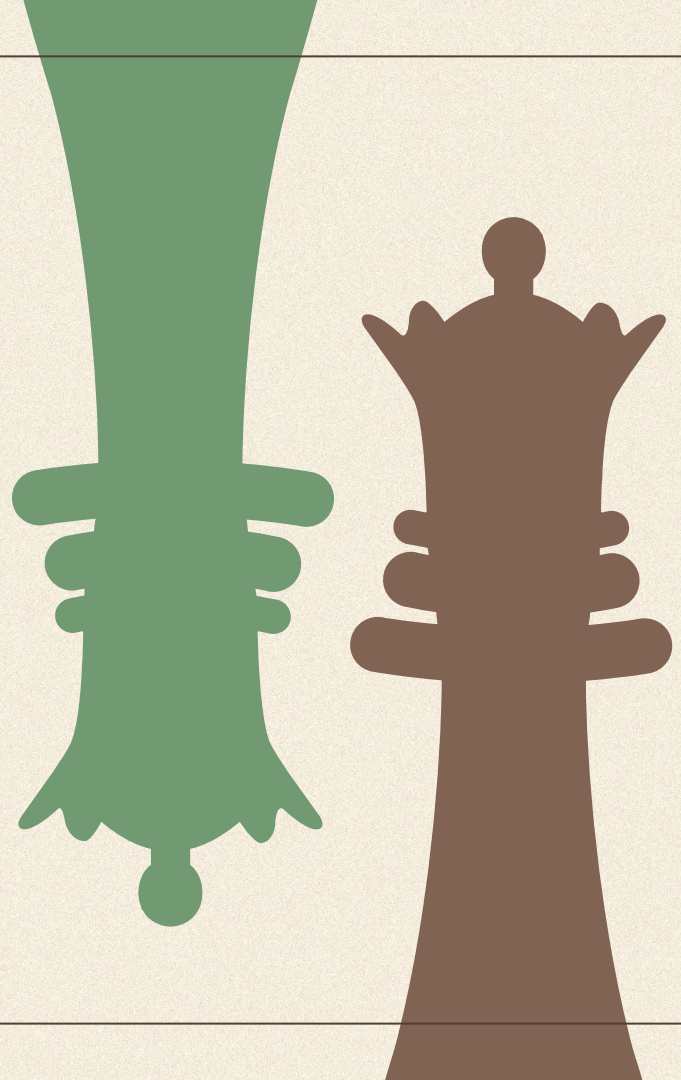
Results





95% Confidence Interval: [4.55 : 14.24]

95% Confidence Interval: [0.211 : 0.745]



03

Question 3

How do other GMs perform when compared to the current Champion Magnus Carlsen?

Mean Centipawn Loss T-Test

Bu Xiangzhi	Mean CP	0.0003076189
Ding Liren	Mean CP	0.0001958873
Ian Nepo	Mean CP	0.0001694836
Wei Yi	Mean CP	0.0159982436
Jennifer Yu	Mean CP	0.0371047003

How do other GMs perform when compared to the current Champion Magnus Carlsen?

Std Centipawn Loss T-Test

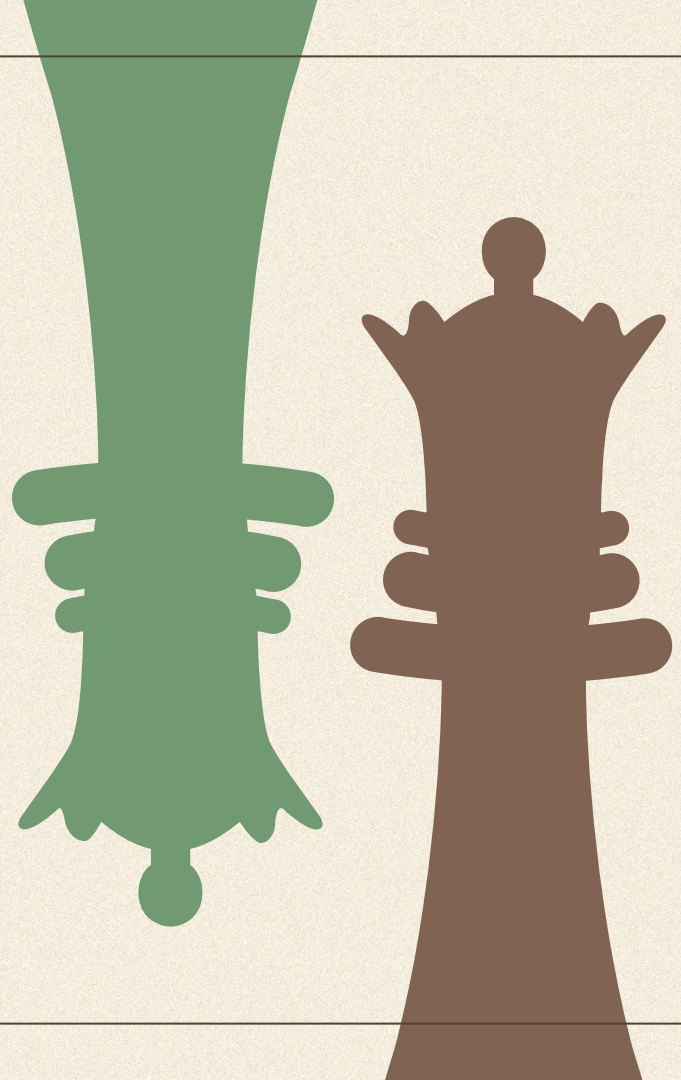
Dmitry Andreikin	Std CP	2.524653e-02
Bu Xiangzhi	Std CP	6.122359e-04
Ding Liren	Std CP	3.393064e-05
Ian Nepo	Std CP	5.687054e-06
Wei Yi	Std CP	6.322617e-03
Jennifer Yu	Std CP	7.677776e-03

How do other GMs perform when compared to the current Champion Magnus Carlsen?

Kruskal-Wallis Test

- Compared variance of players for both mean and std centipawn loss
- Done with and without Carlsen

	Test Statistic	P Value
Before	187.25	2.2e-16
After	182.02	2.2e-16

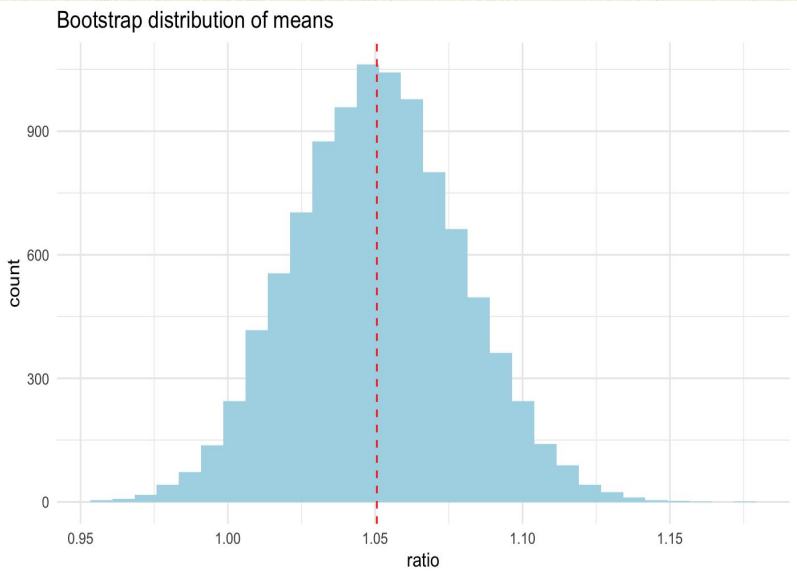


04

Question 4

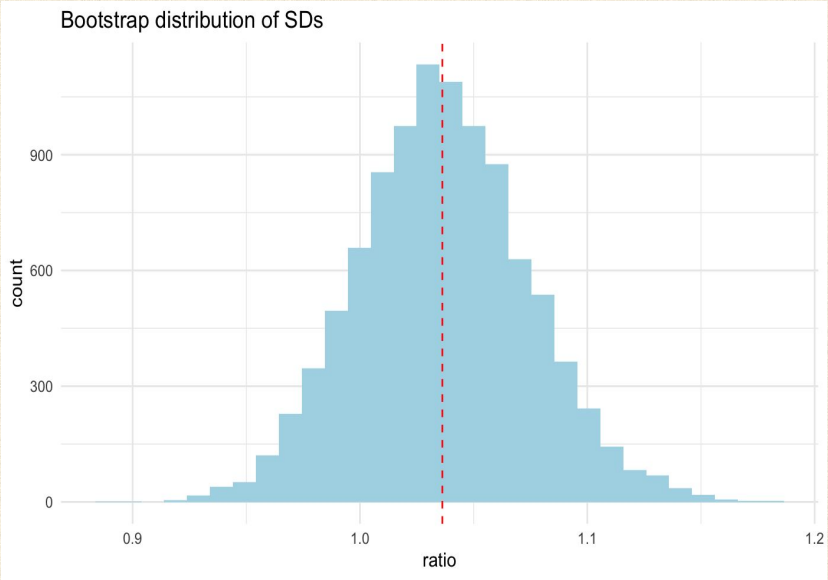
How do other GMs perform when compared to the rising chess prodigy Niemann?

Bootstrapping of Mean_cp



95% Confidence Interval: [1.00 : 1.11]

Bootstrapping of Std_cp



95% Confidence Interval: [0.97 : 1.11]

How do other GMs perform when compared to the rising chess prodigy Niemann?

T-test of Mean_cp

t=1.8045,

p-value=0.03593

95% confidence interval:

0.1115292 ~ Inf

T-test of Std_cp

t=0.96464,

p-value=0.1676

95% confidence interval:

-1.106044 ~ Inf

INTERPRETATION:

Niemann shows a statistically significant higher mean centipawn loss in comparison with other players, thus we reject the theory of Niemann cheating by using a computer helper.

Conclusions

What are the best predictors for how well and consistent a player performs?

- As player's **accuracy** tends to **decrease** as their **consistency decreases** (and vice versa).
- A player's **accuracy** tends to **increase** as their **strength increases**.

How does Niemann's growth compare to other GMs?

- Niemann has a **statistically significant** higher **yearly ELO growth** than the average player, but there is no evidence to suggest that his ELO **growth per game** is different from the average player. Mixed results that call for more research but **no definitive proof** of cheating.

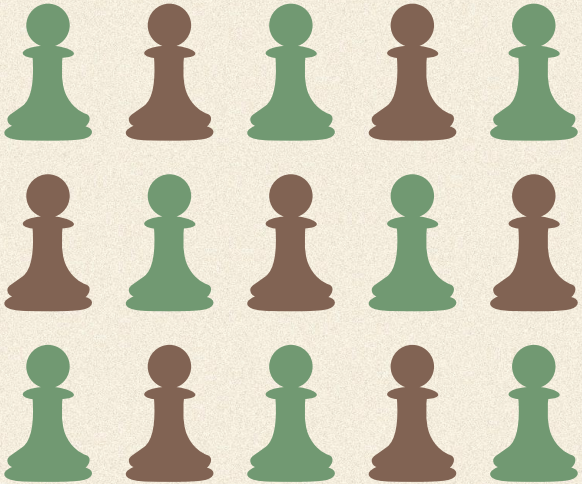
How do other GMs perform when compared to the current World Champion Magnus Carlsen?

- Out of the 31 grandmasters in our dataset, Magnus Carlsen performed with **overall less accuracy and consistency** than only 5-6 other grandmasters.

How do other GMs perform when compared to the rising chess prodigy Niemann?

- Niemann performed **less accurately** than other GMs but played with the **same consistency.** Mixed results that call for more research but **no definitive proof** of cheating.

Discussion and Limitations



- Consideration of **dependencies within the data**
  - Elo is inherently correlated w/ the player
  - Centipawn loss will be correlated w/ game time format
  - Centipawn **calculations will differ** based on the version of Stockfish used
- Our dataset **only considers over-the-board games and 31 professional players**
- When playing chess at grandmaster levels, there **tends to be many outliers and influential points**.
  - This could be because that at this level, the players tend to play variably (e.g. more risky / out-of-theory moves, time constraints).

Future Studies and Implications

- There may be other analyses like Time Series Analysis that can provide better insight into whether, players like Niemann are cheating by considering performance over time.
- Our dataset stops before the game when Niemann was accused of cheating against Carlsen. Future studies might try fitting the Mean CP and Std CP from that game into our regression models.