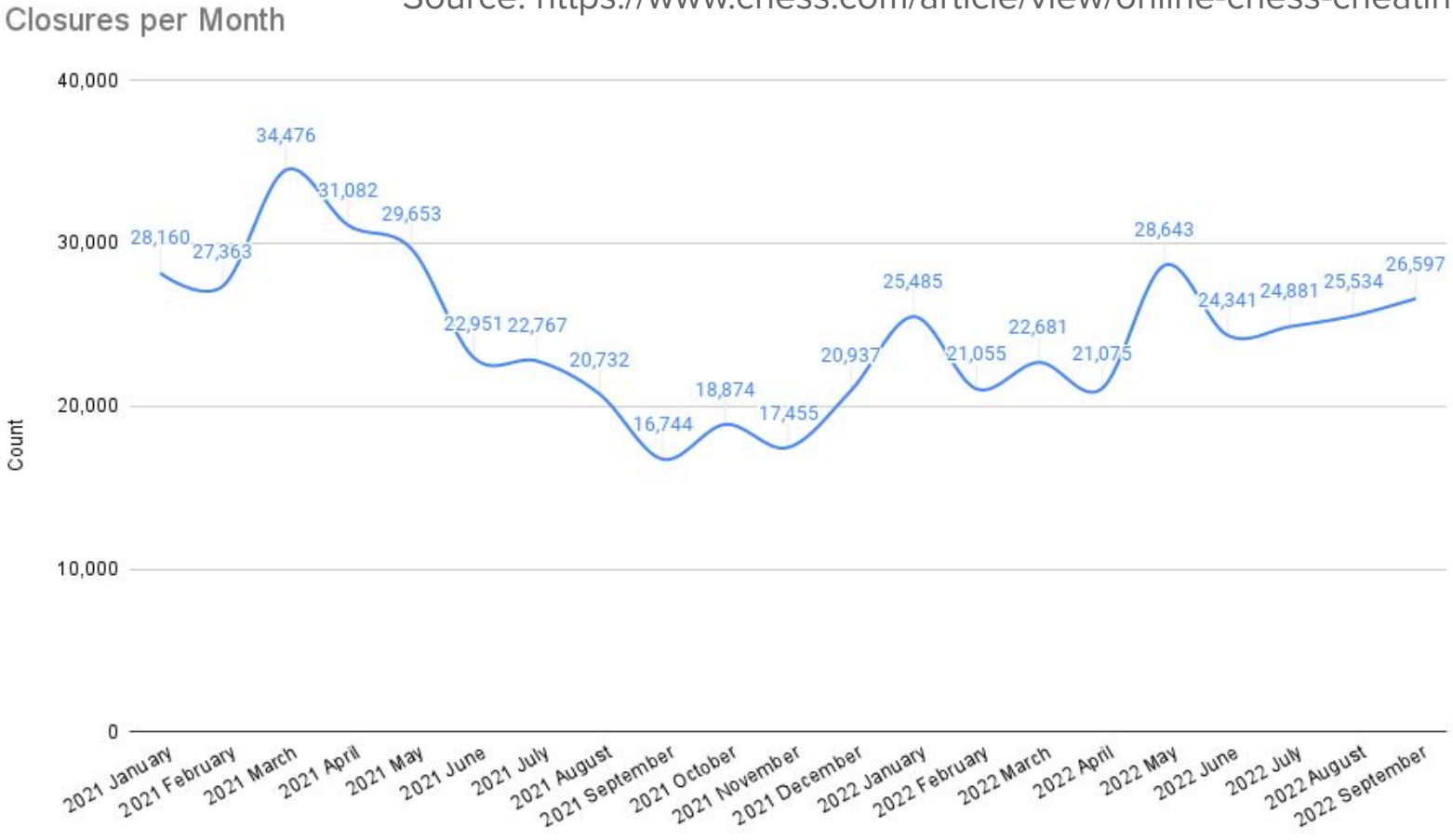
Interesting Results in Chess

Edward Wu
Project Advisor: Professor Fernandez-Granda

Source: https://www.chess.com/article/view/online-chess-cheating



Learning from AI

While AI can be used to cheat, it can also be used as a training partner. Players can use AI to:

- Learn patterns about what best moves might look like in complex positions,
- Perfect their play in openings and endgames, and
- Formulate more precise plans and counterattacks.

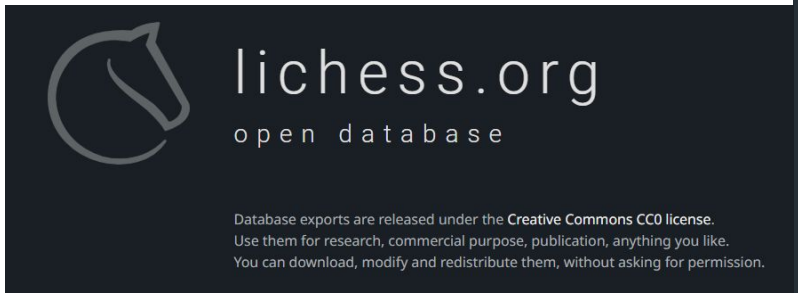
None of these examples constitute cheating, as players are allowed to prepare lines and materials ahead of game time. They cannot bring materials to the board, but anything they remember can be used.

Source: https://www.theguardian.com/sport/2023/aug/28/magnus-carlsen-hans-niemann-chess-cheating-allegations-settlement

Searching for Data

1. High-quality & Had lots of relevant features
2. Plentiful & Had lots of observations
3. Accessible

The Data



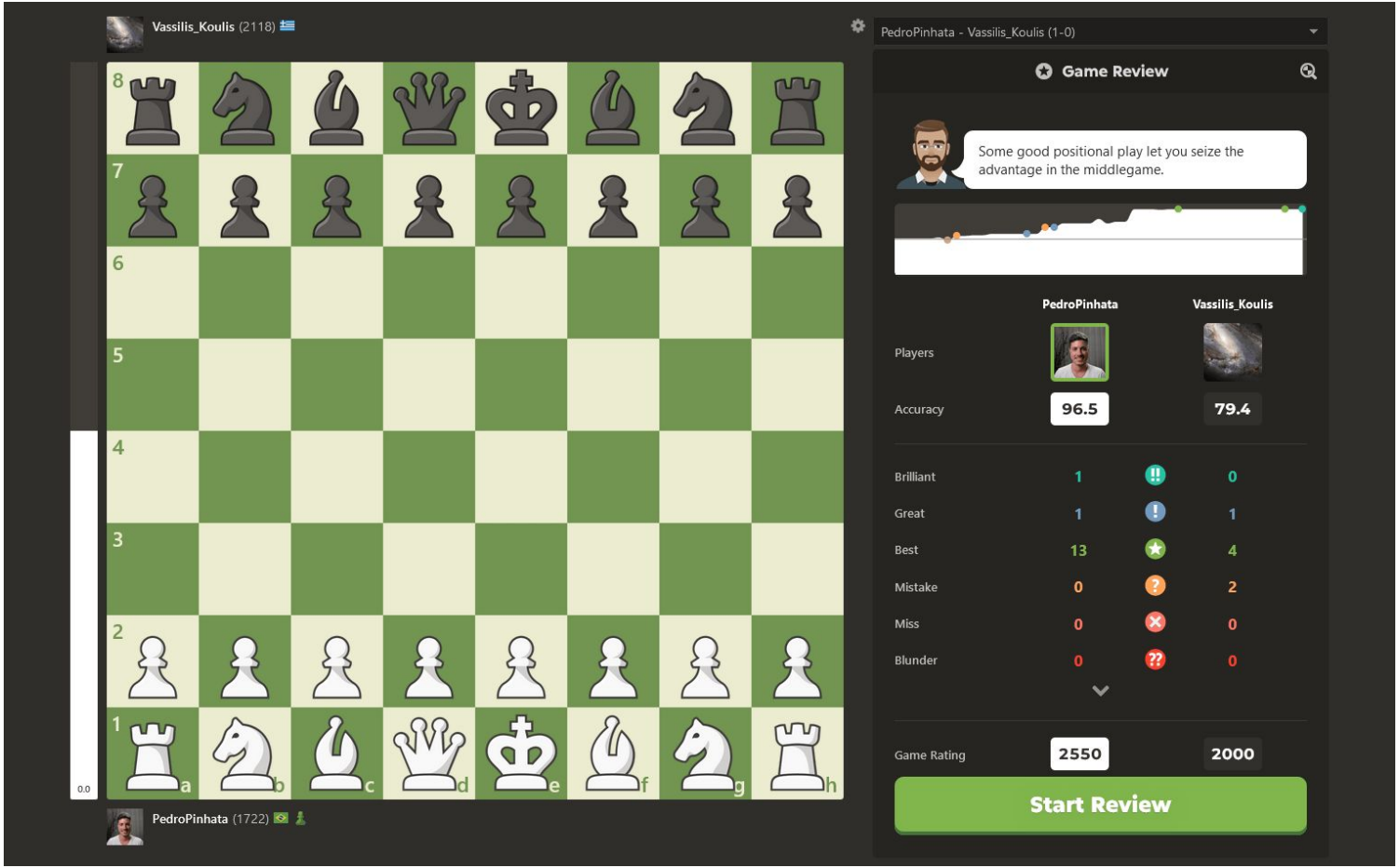
STANDARD CHESS | VARIANTS | PUZZLES | EVALUATIONS

6,111,210,071 standard rated games, played on lichess.org, in PGN format. Each file contains the games for one month only; they are not cumulative.

Month	Size	Games	Download
2024 - October	30.7 GB	94,254,891	.pgn.zst / .torrent
2024 - September	28.6 GB	87,713,219	.pgn.zst / .torrent
2024 - August	30 GB	92,198,878	.pgn.zst / .torrent

	game_id	type	start_time	increment	winner	num_ply	move_ply	move	player	cp	cp_rel	cp_loss	blunder	time_spent	clock
35	tvycmBtA	Rapid	600	10	1	40	35	c8d7	-1	-1.15	1.15	0.45	False	37.0	520
36	tvycmBtA	Rapid	600	10	1	40	36	a1c1	1	-0.7	-0.70	0.29	False	67.0	321
37	tvycmBtA	Rapid	600	10	1	40	37	a8c8	-1	-0.99	0.99	2.07	True	13.0	517
38	tvycmBtA	Rapid	600	10	1	40	38	d2e4	1	1.08	1.08	0.01	False	33.0	298
39	tvycmBtA	Rapid	600	10	1	40	39	d7g4	-1	1.07	-1.07	1.55	False	61.0	466
40	gTyqnWXd	Bullet	60	0	1	55	0	e2e4	1	0.1	0.10	-0.02	False	NaN	60
41	gTyqnWXd	Bullet	60	0	1	55	1	e7e5	-1	0.12	-0.12	0.25	False	NaN	60
42	gTyqnWXd	Bullet	60	0	1	55	2	g1f3	1	0.37	0.37	0.14	False	1.0	59
43	gTyqnWXd	Bullet	60	0	1	55	3	d7d6	-1	0.23	-0.23	0.38	False	1.0	59
44	gTyqnWXd	Bullet	60	0	1	55	4	d2d4	1	0.61	0.61	0.21	False	1.0	58
45	gTyqnWXd	Bullet	60	0	1	55	5	c8g4	-1	0.4	-0.40	0.52	False	1.0	58

How can we Detect Cheating?



Source: https://www.chess.com/news/view/game-review-design-update

What can we use this data for?

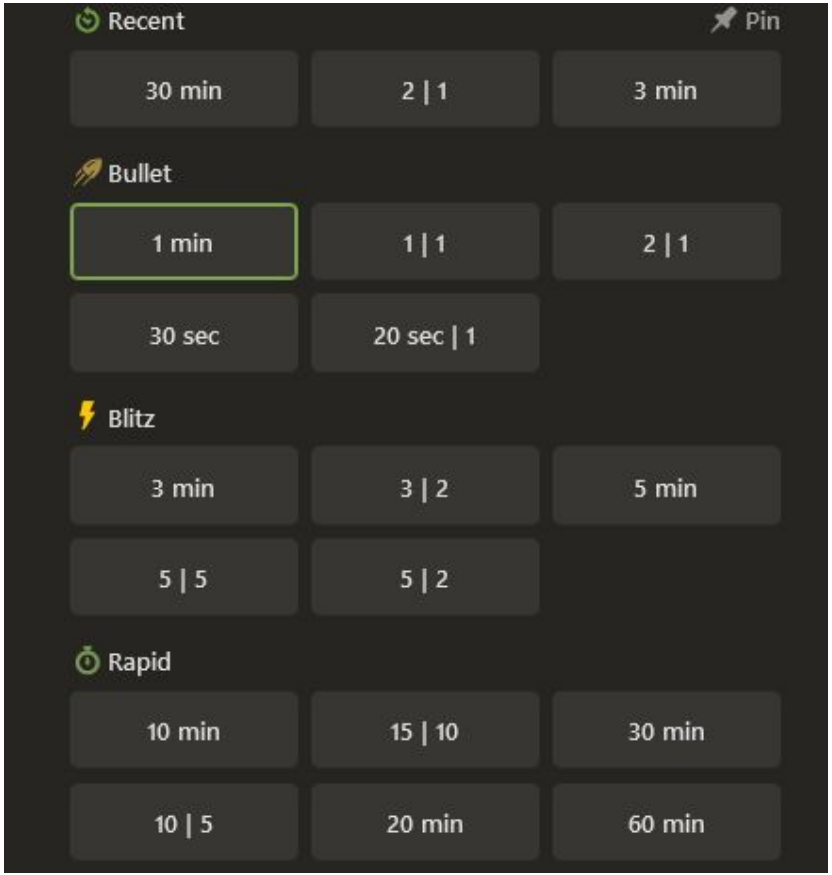
To answer questions:

1. Do blunders end up following longer periods of thinking?
2. Does thinking longer tend to lead to better moves?
3. Do more experienced players spend more time thinking?
4. How often do players who make the first check win the game?
5. Are more experienced players more cautious with their pieces?
6. Is the amount of time that a player thinks correlated with the number of possible moves?

Time Control (Definition)

The time control of a game dictates how long players have to think on their moves.

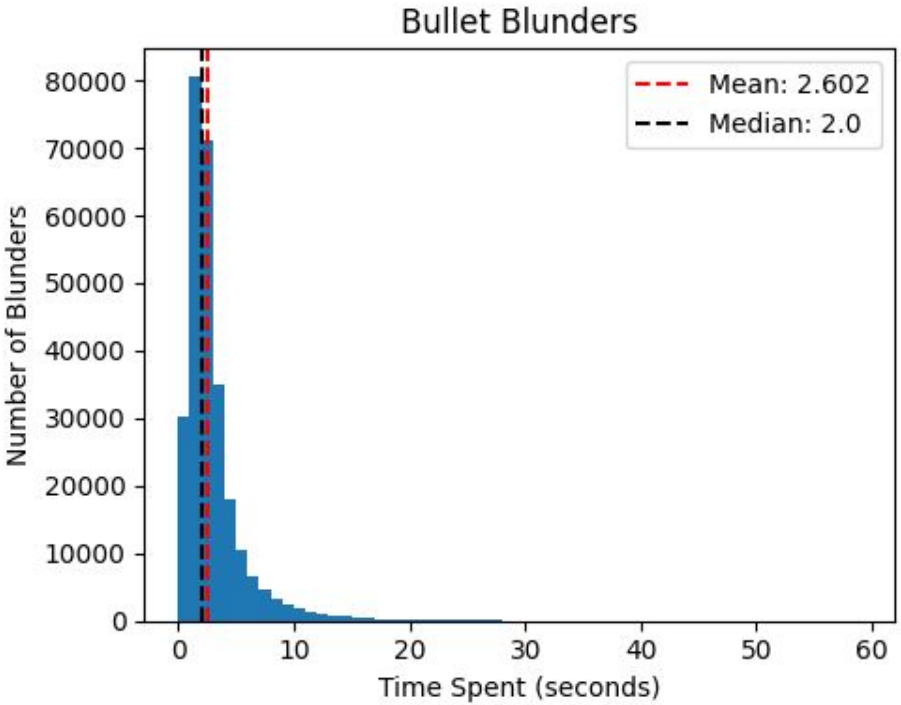
Time control has two parts, the base time and the increment. The base time is the number of minutes that each player starts with. The increment is the number of seconds that each player gains after making a move, which helps avoid timeouts. For example, 5 | 2 means a base time of 5 minutes and an increment of 2 seconds.

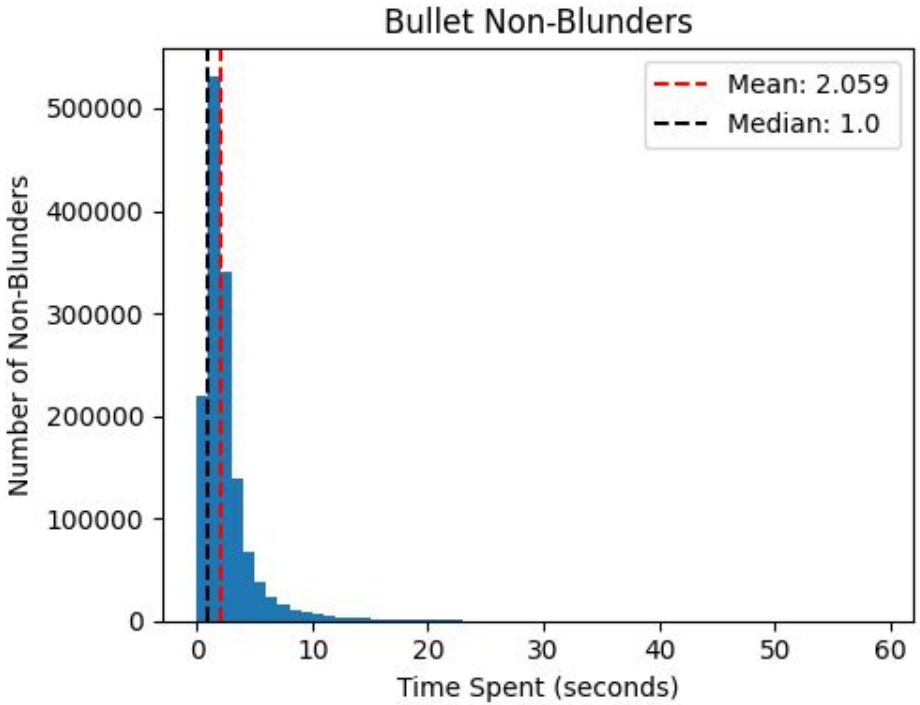


Question 1:

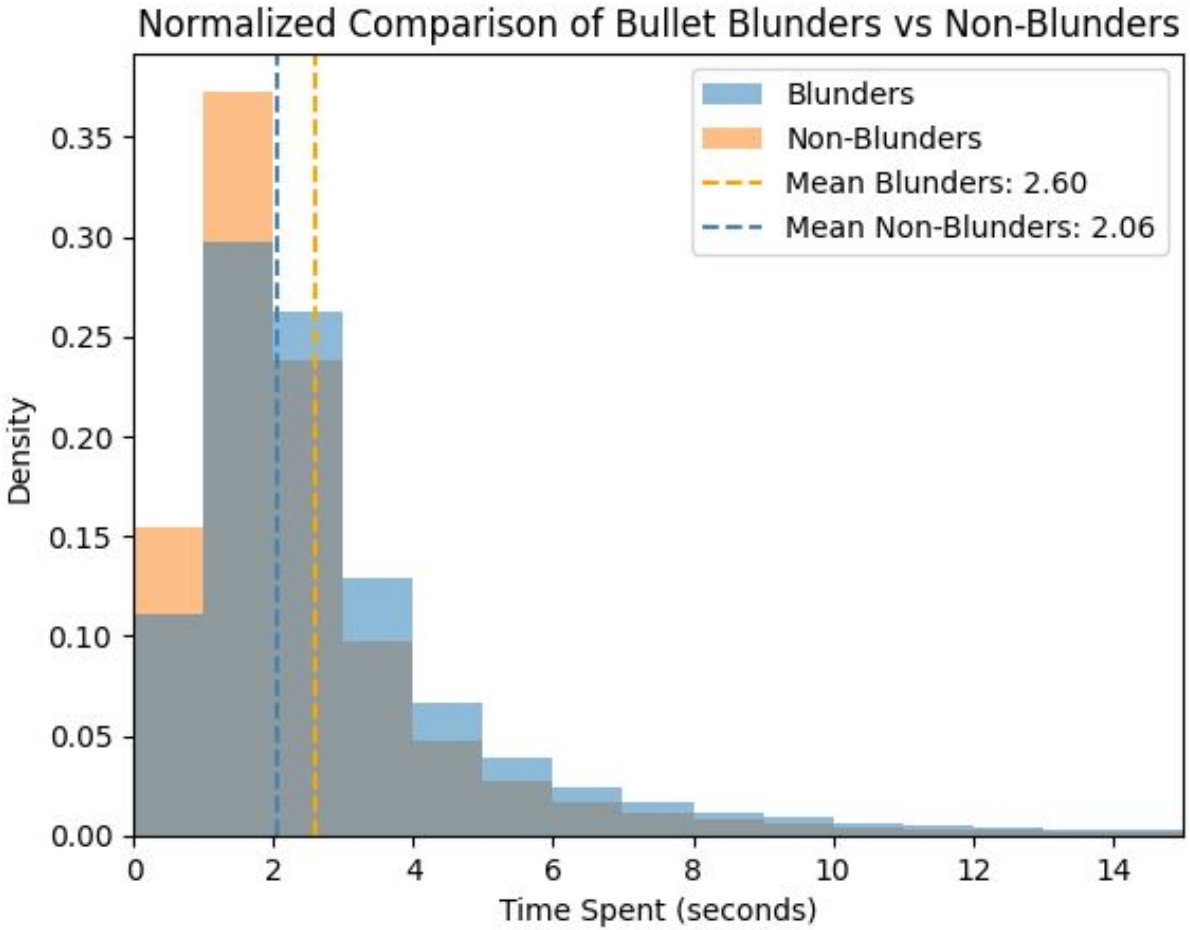
How do the distributions of thinking time differ between moves that are blunders and moves that are not?

Question 1 - How do the distributions of thinking time differ between moves that are blunders and moves that are not?

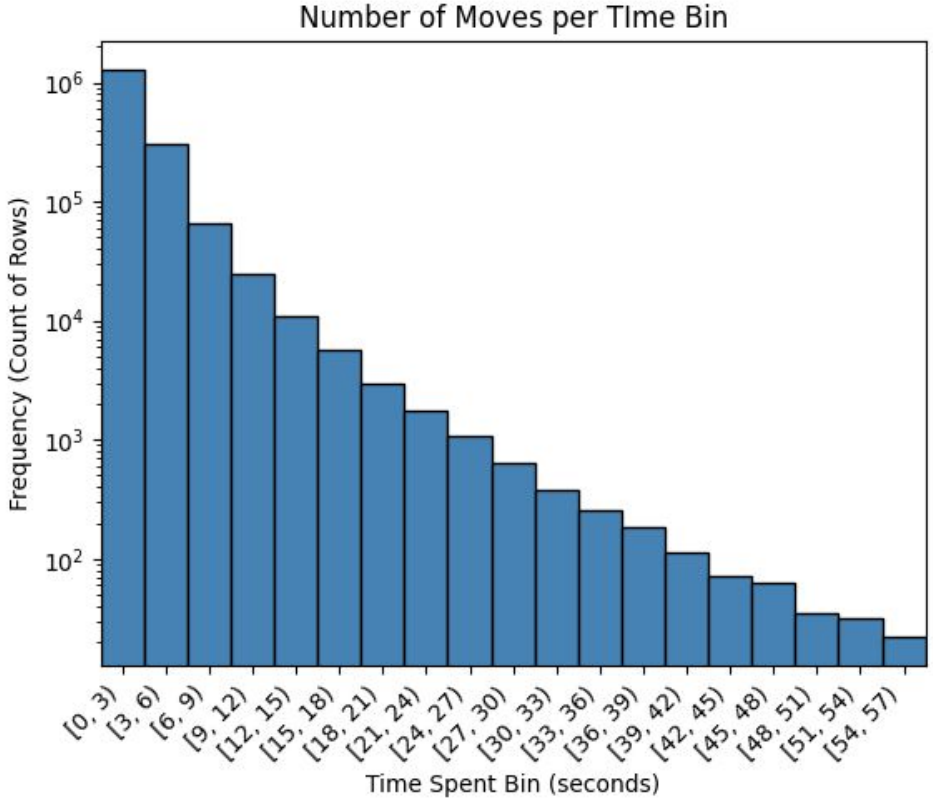


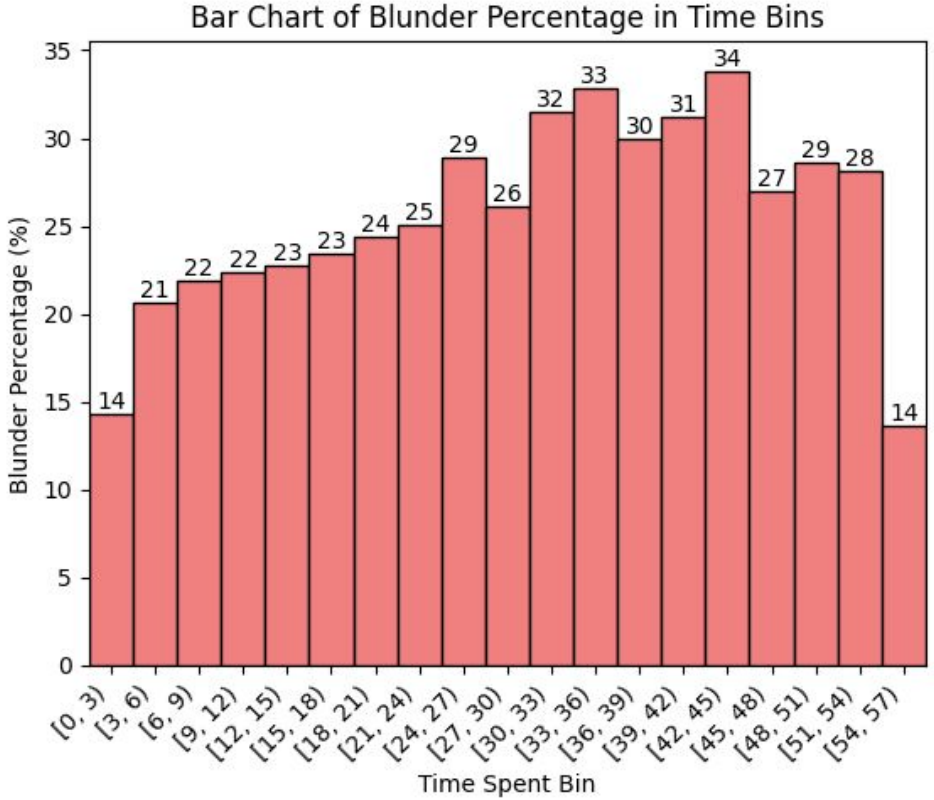


Question 1 - How do the distributions of thinking time differ between moves that are blunders and moves that are not?



Question 1.1 - Given a time spent thinking, what is the likelihood that the resulting move is a blunder?





Some Possible Explanations

- Positions that are more complex cause players to think more, but they are also harder to find the best move for.
- Lower elo players make up the majority of the data, and they are responsible for this correlation.

Question 3:

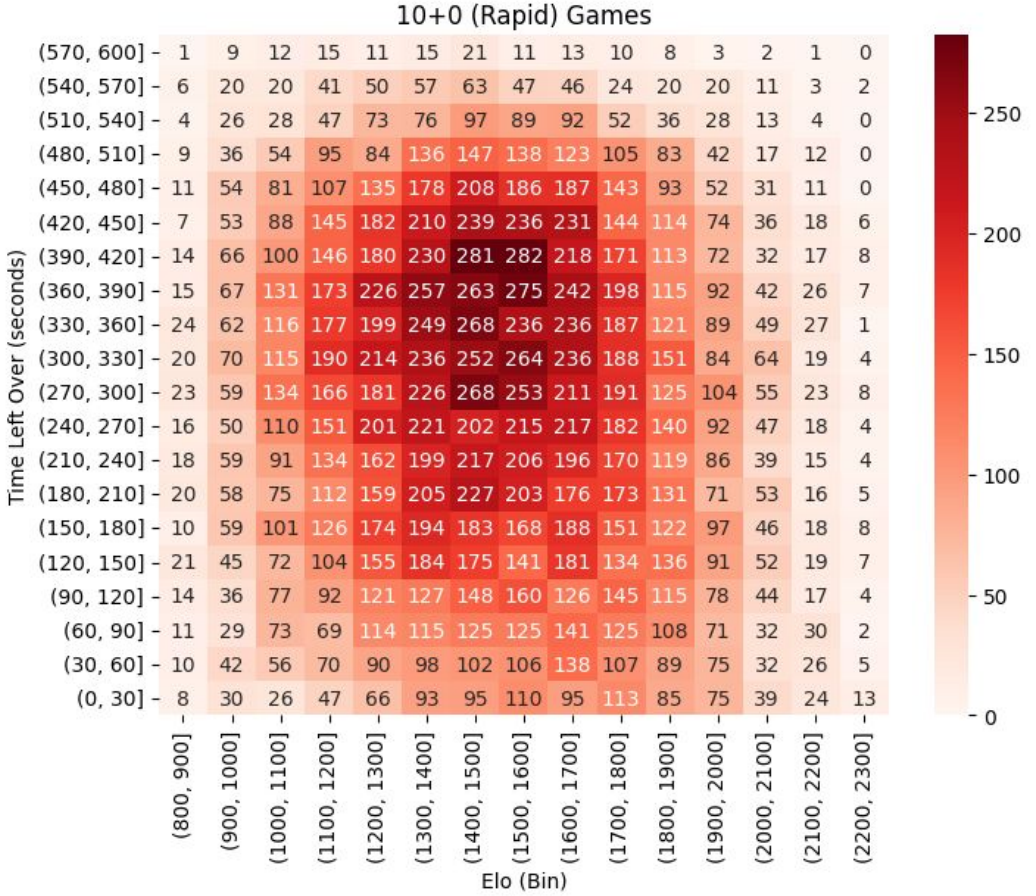
Do more experienced players spend more time thinking?

Question 3 - Do more experienced players spend more time thinking?

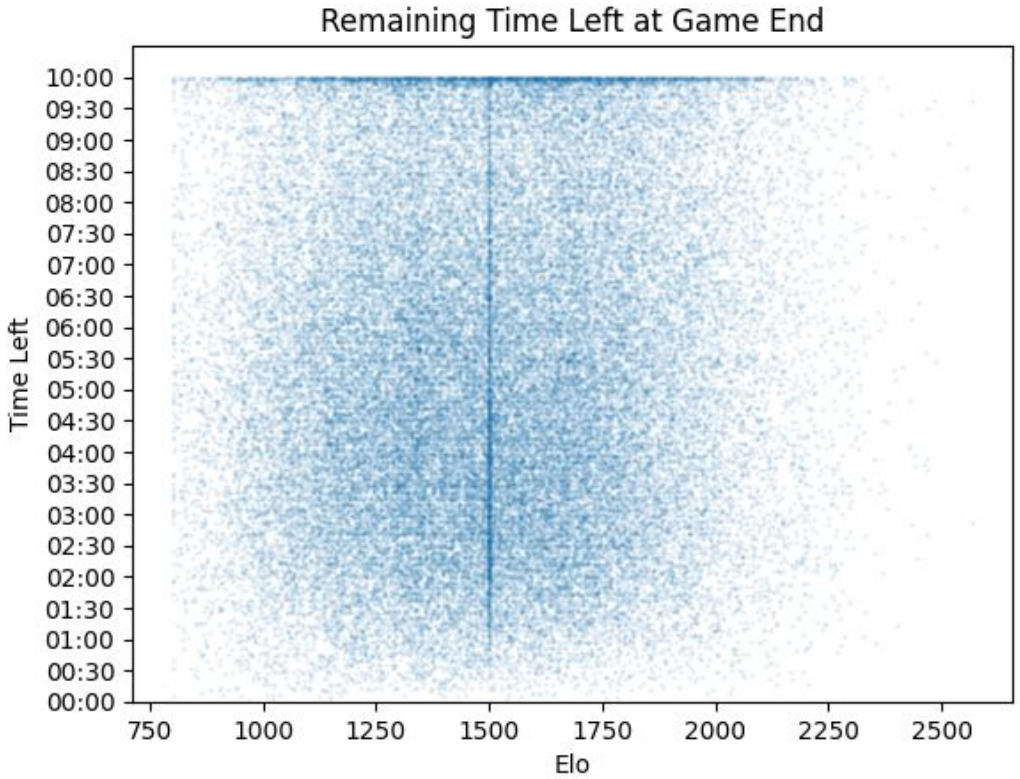
To measure how much time players spend thinking, we instead consider the complement: how much time is leftover on their clocks when the game ends.

I filtered specifically for Rapid games that were 10+0. Each player starts with 10 minutes and receives 0 seconds of additional time per move made. This second point allows this calculation to be made independent of the length of the game in plies.

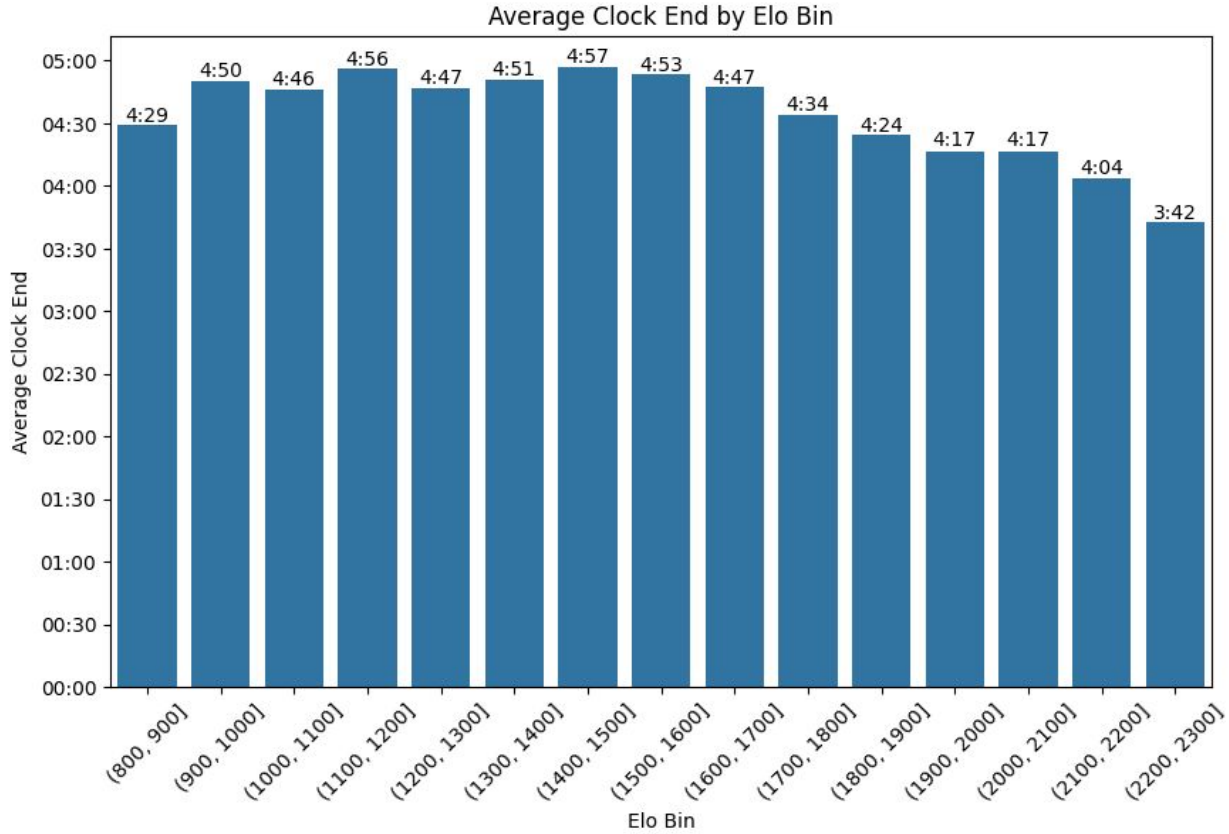
We observe a negative correlation between average elo and the clock end of -0.067.



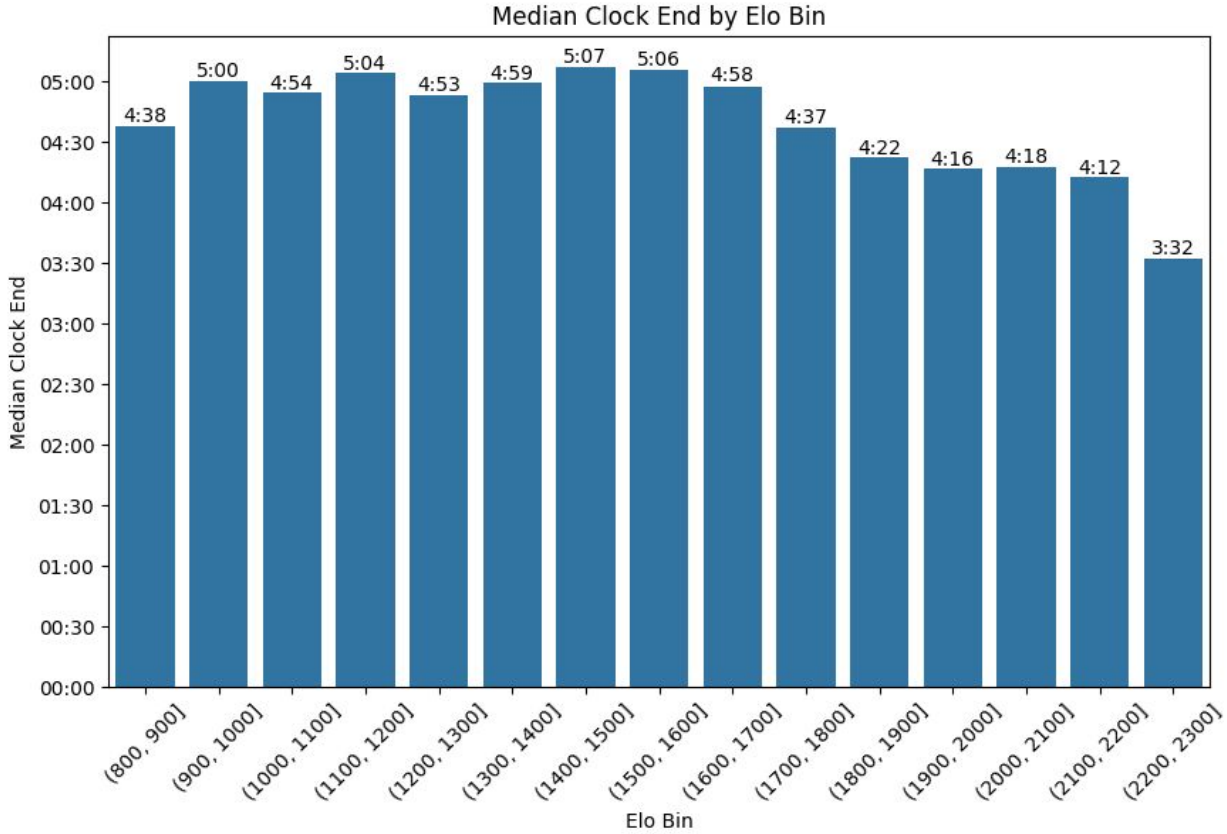
Question 3 - Do more experienced players spend more time thinking?



Question 3 - Do more experienced players spend more time thinking?

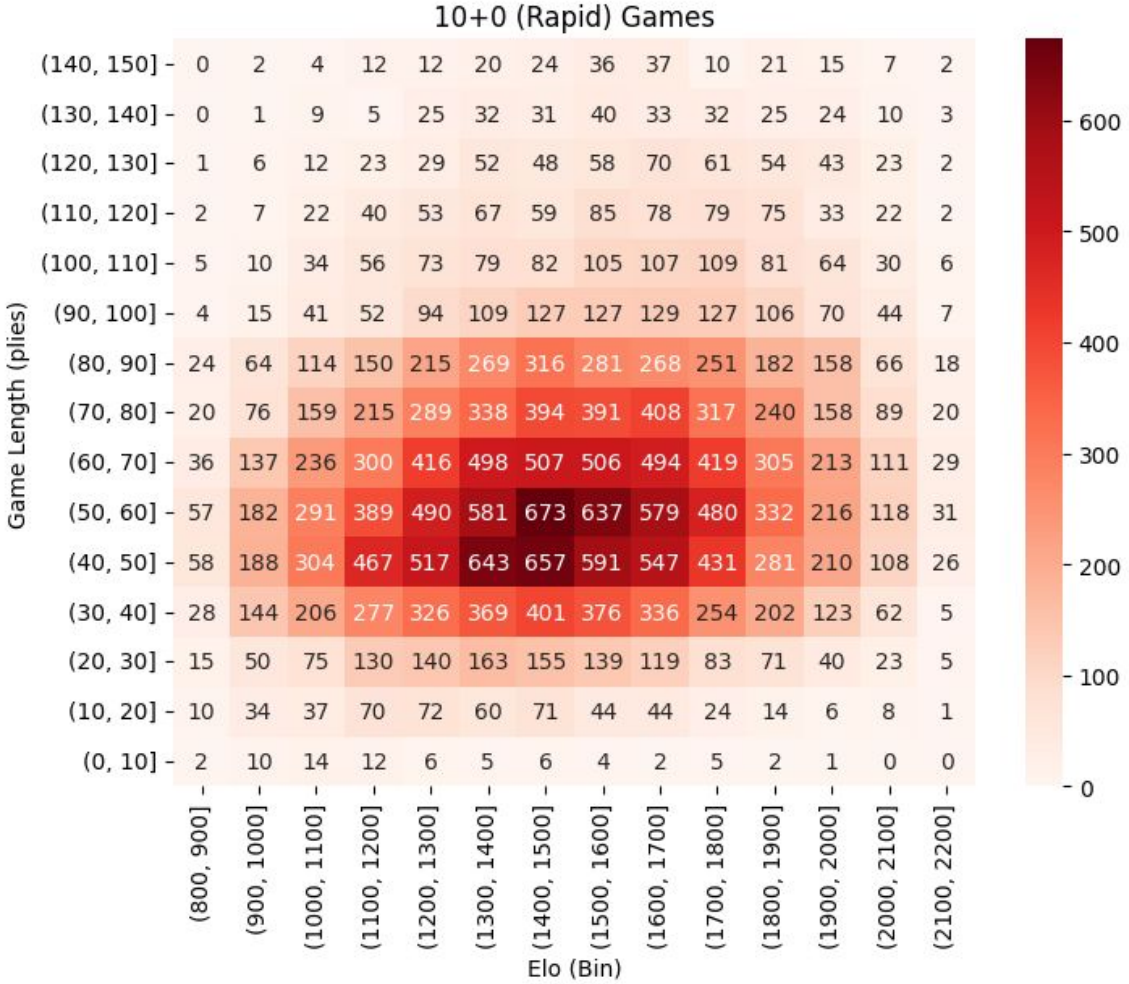


Question 3 - Do more experienced players spend more time thinking?

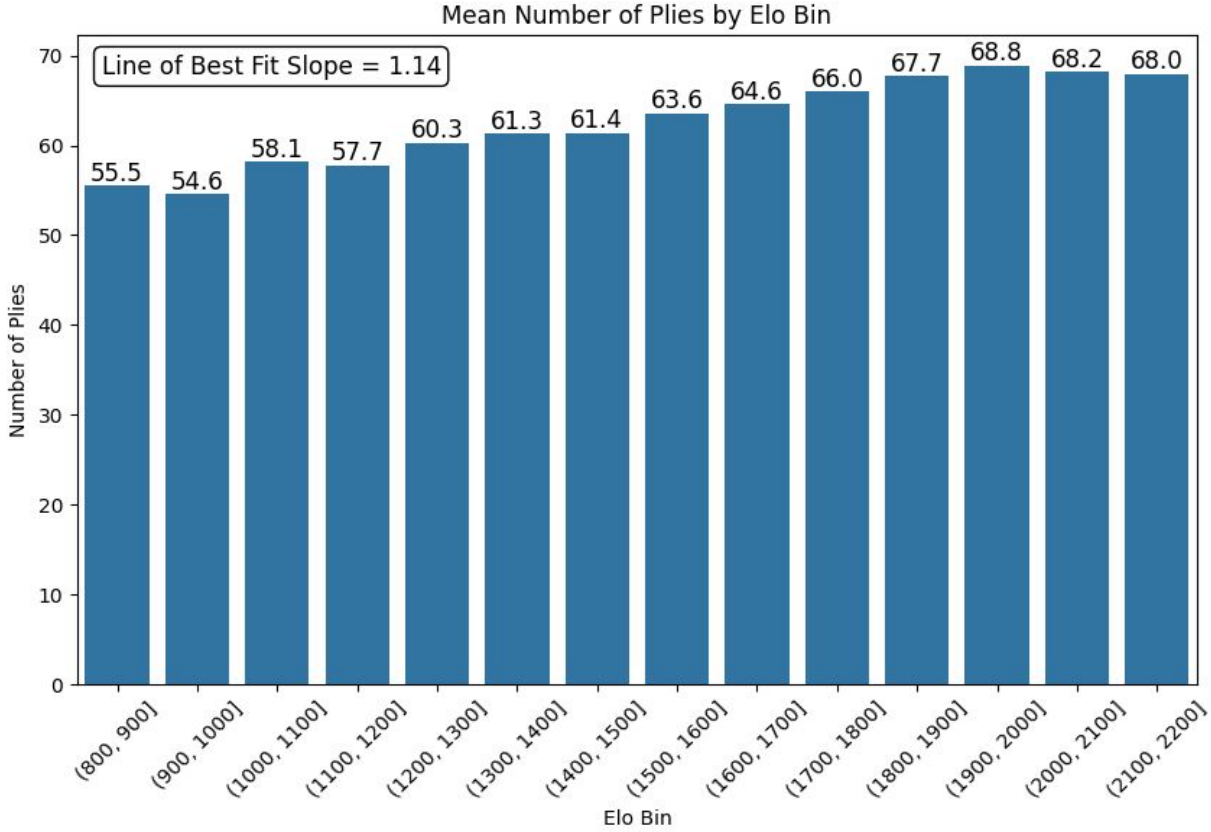


Question 3.1 - Do more experienced players play games that last more moves?

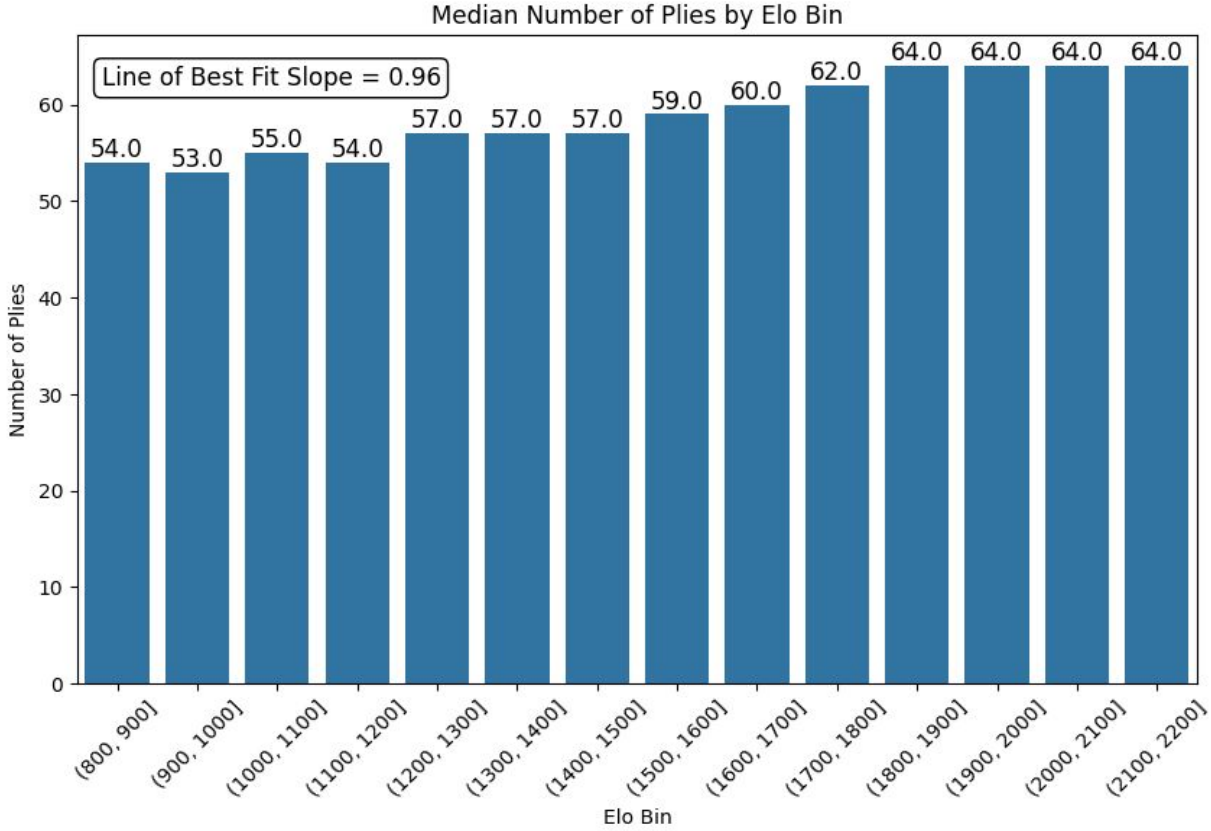
We observe a correlation of 0.1519.



Question 3.1 - Do more experienced players play games that last more moves?

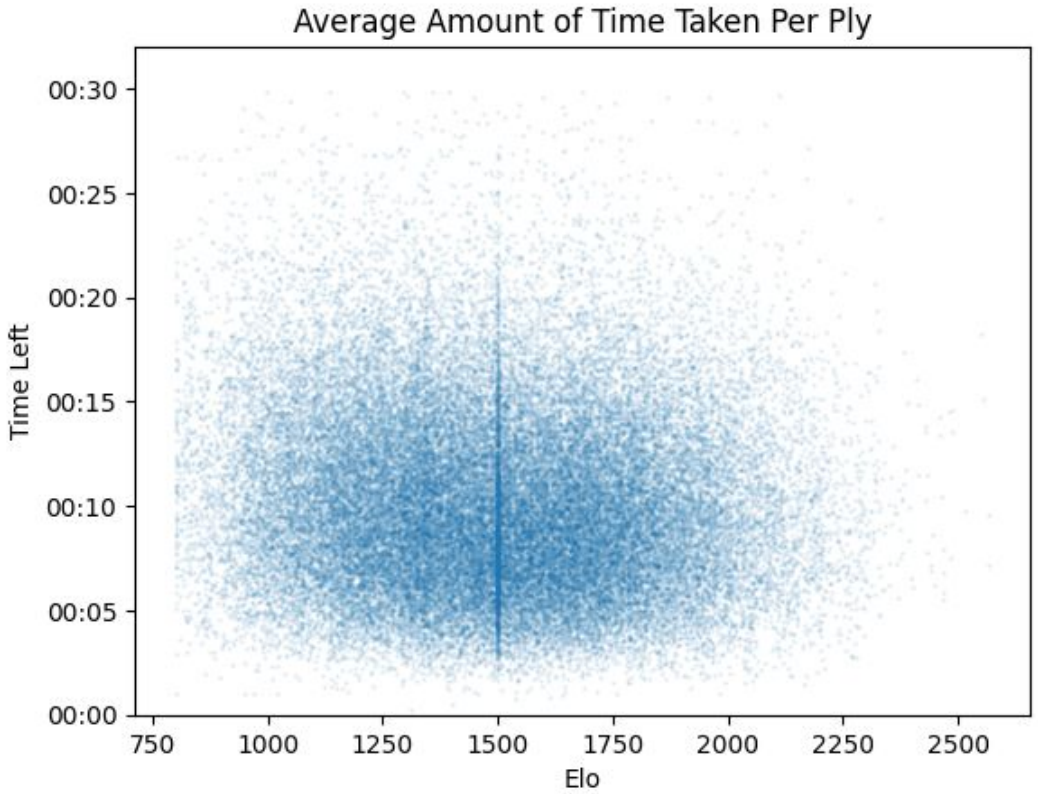


Question 3.1 - Do more experienced players play games that last more moves?



Question 3.2 - Do more experienced players play moves faster?

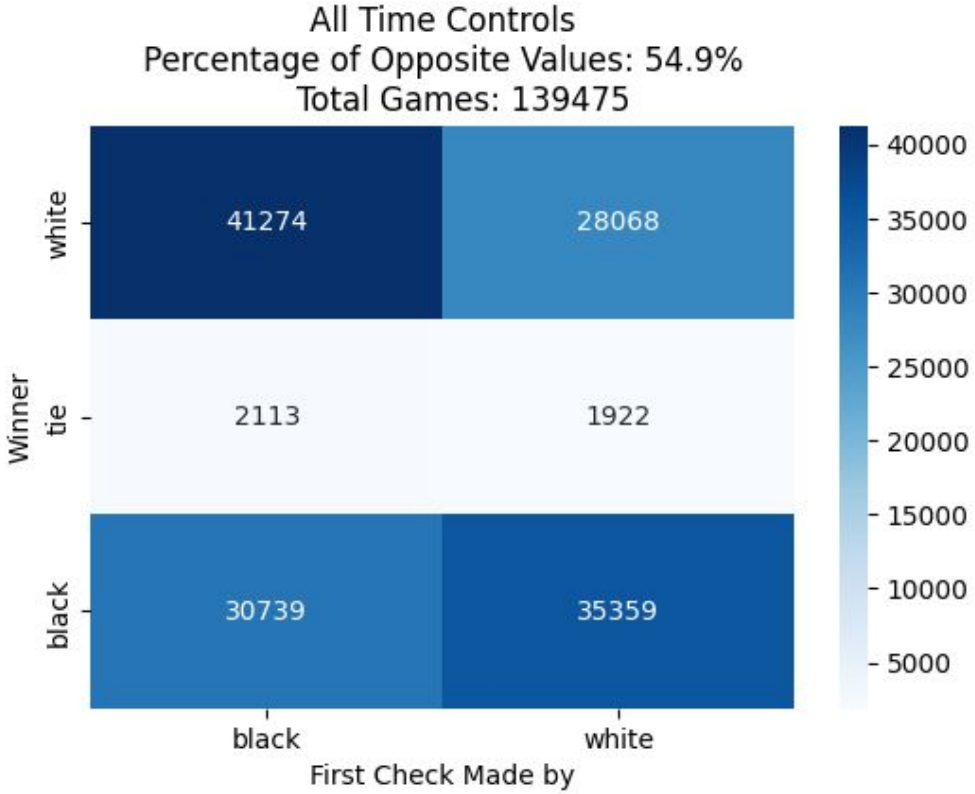
We observe a correlation of -0.0851.



Question 4:

How often do players who make the first check win the game?

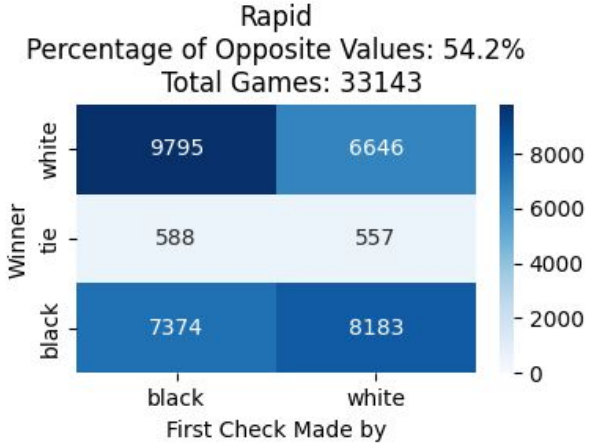
Question 4 - How often do players who make the first check win the game?

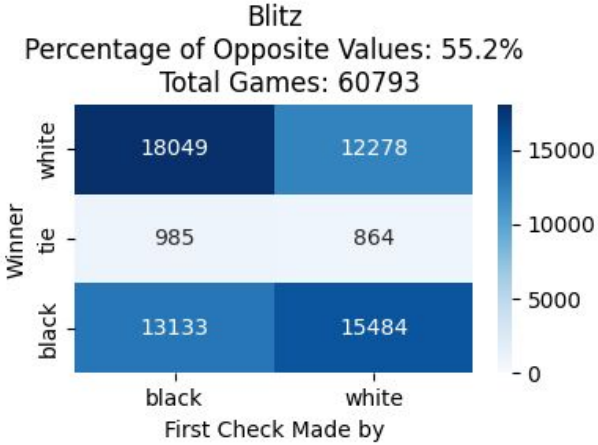


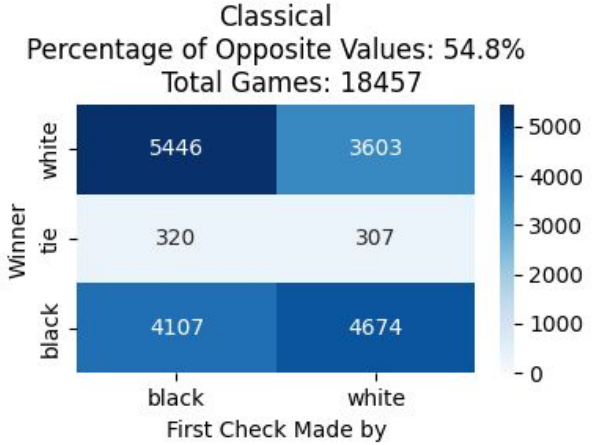
Possible Explanations

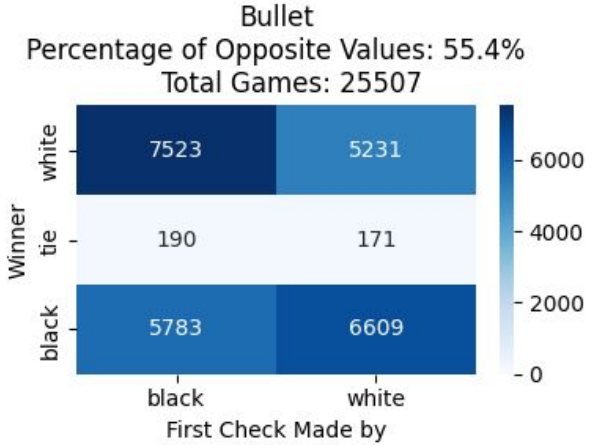
- This result is surprising because the check is one step towards checkmate, the goal of the game
- Maybe the check is used as a defensive measure. Chess engines will often play a check if they see that they will be mated as an attempt to extend the game by a move or two, or distract the opponent.
- Maybe players are under time pressure.
- Maybe the dataset is dominated by lower elo players, who are more likely to check their opponent just because they can.

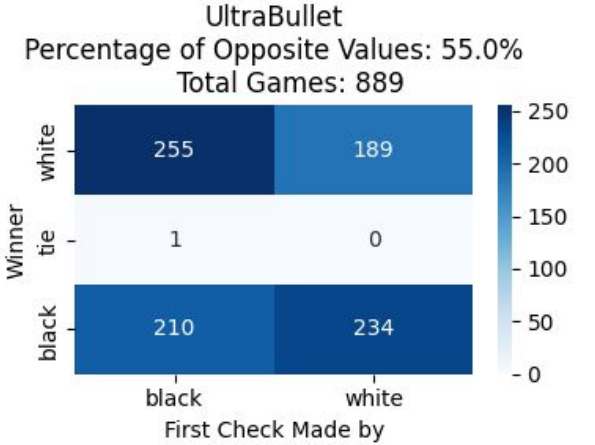
Question 4.1.1 - Is the result consistent upon stratification on time control?

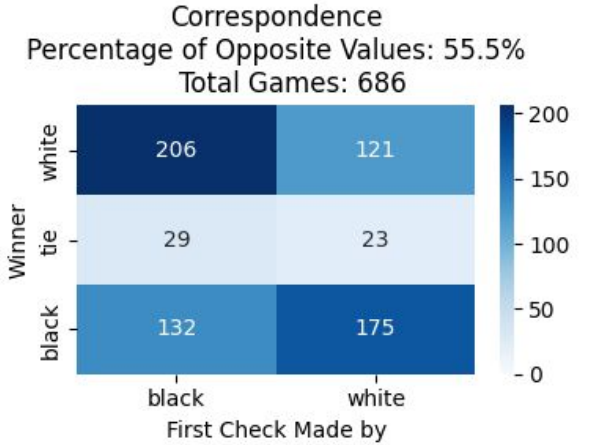






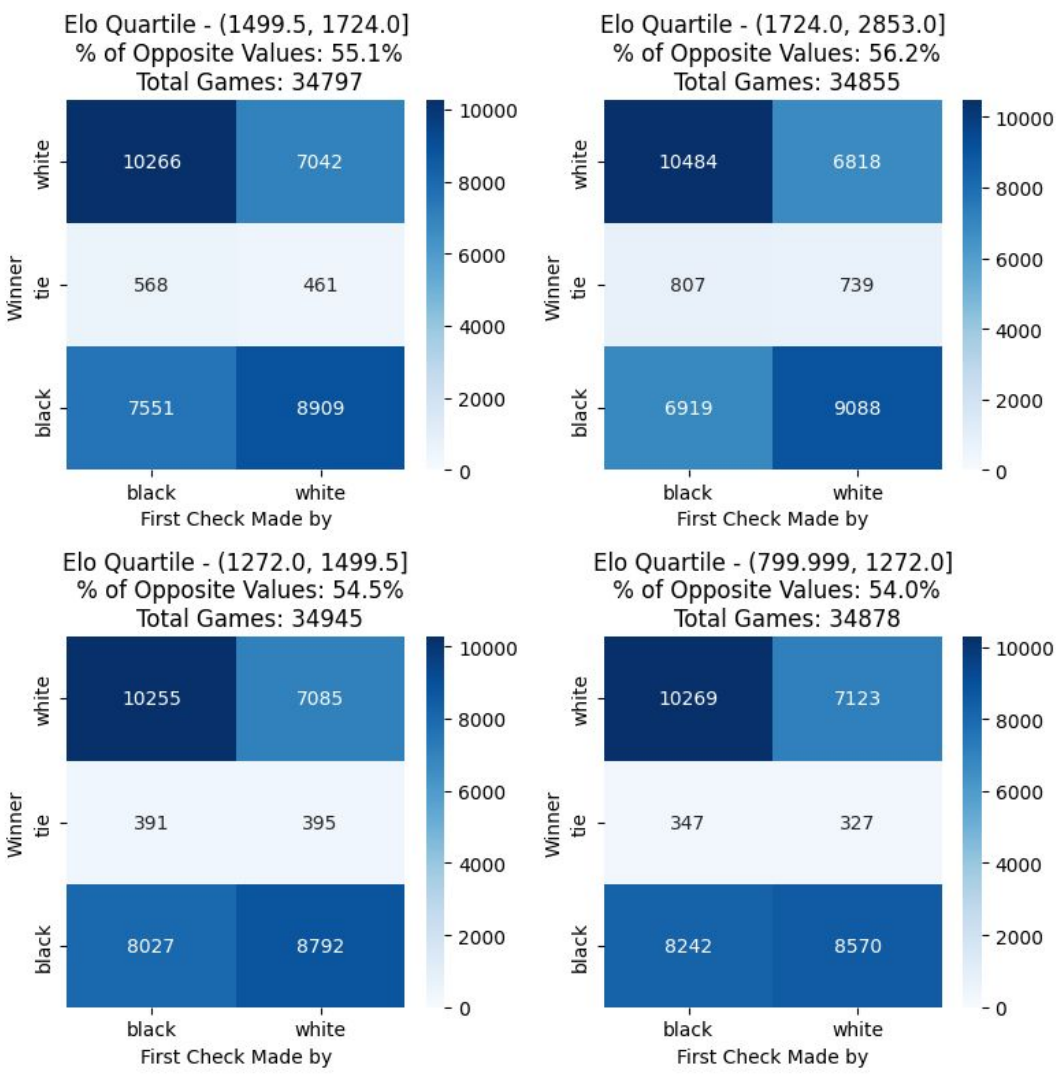




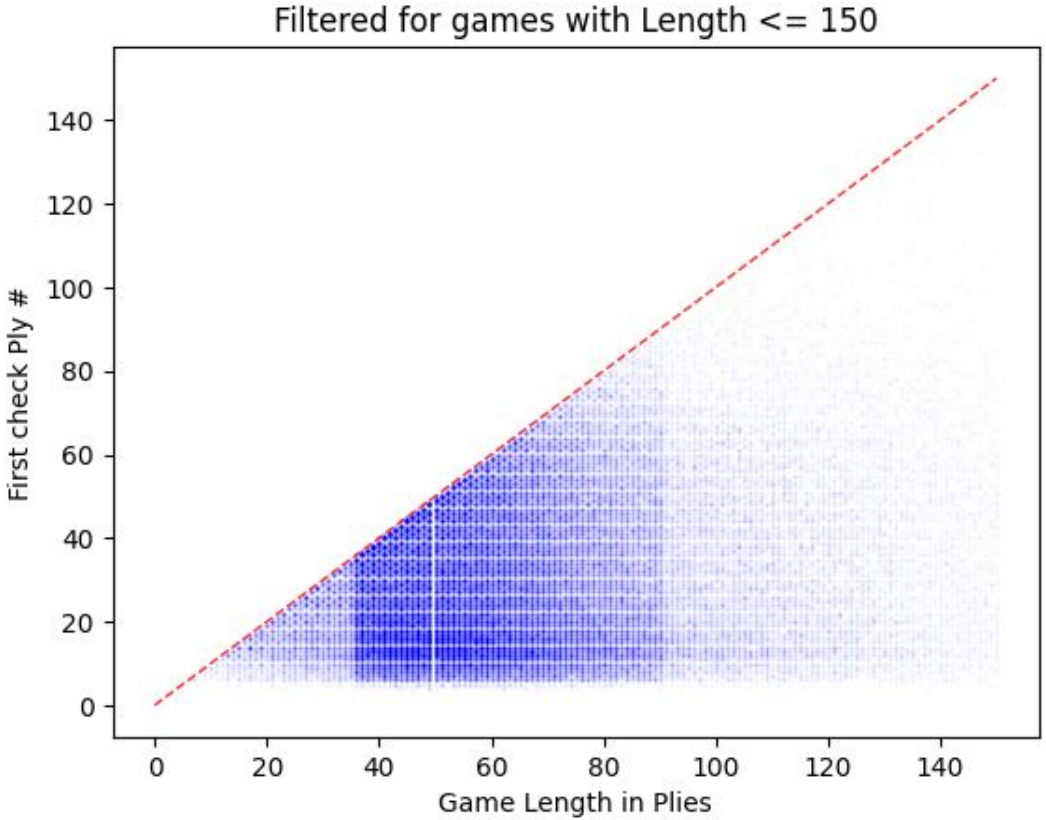


Question 4.1.2 - Is the result consistent upon stratification on elo?

The reality is that the percentage is higher for higher elo players!



Question 4.2.1 - How often is the first check also linked with the end of the game?



Board Evaluation (Definition)

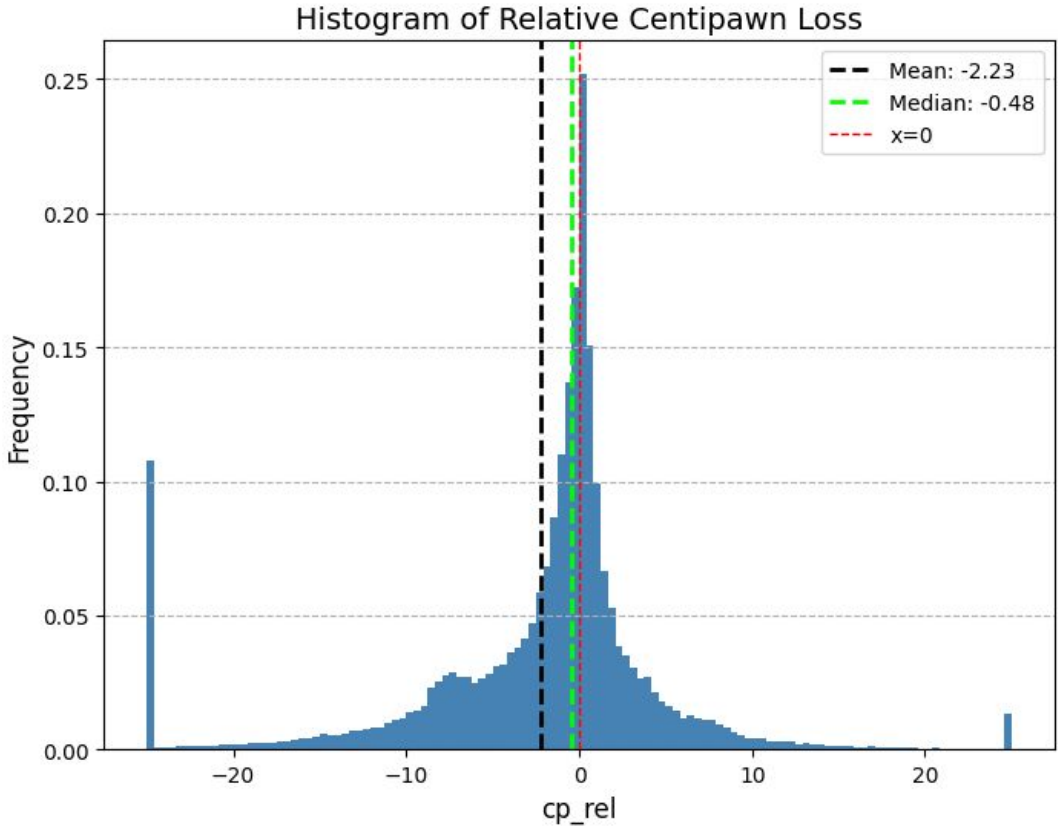
Speaking of the complexity of the board, the board evaluation is the engine's evaluation of how much the board state favors either player. An evaluation of 0 means the board is a tie, while a positive evaluation means white has the advantage.

Evaluation is measured in centipawn loss. One centipawn is 1/100 of a pawn.

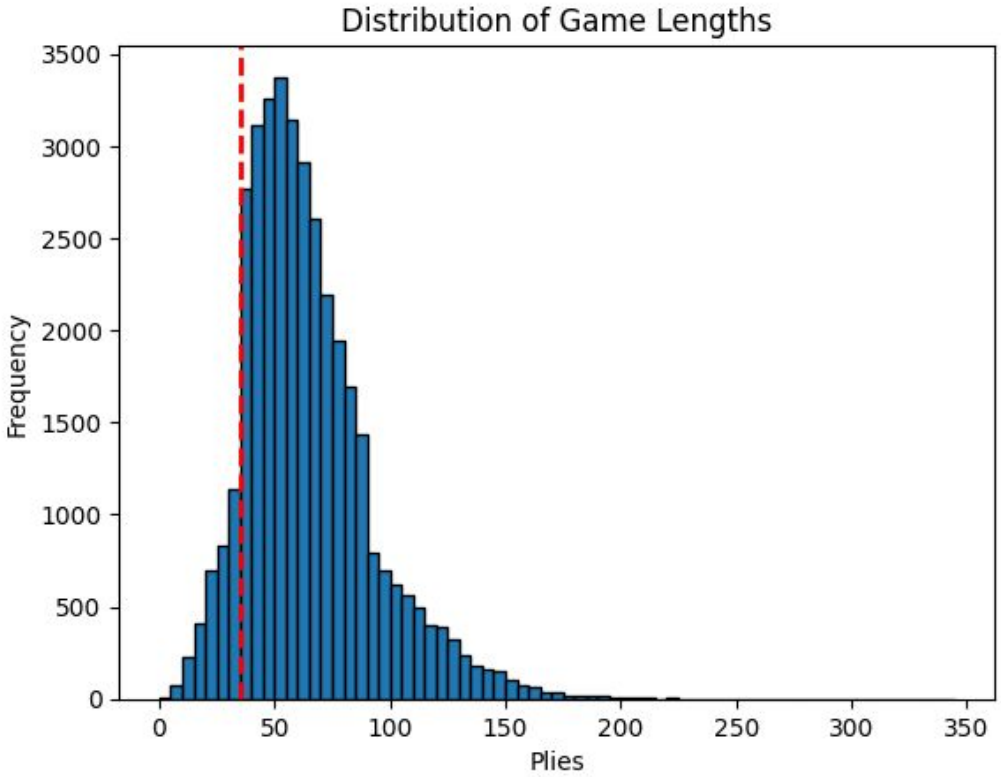
Question 4.2.2 - How often is the first check used as a last ditch effort?

There is an imbalance in the relative evaluation of the board when the first check is made.

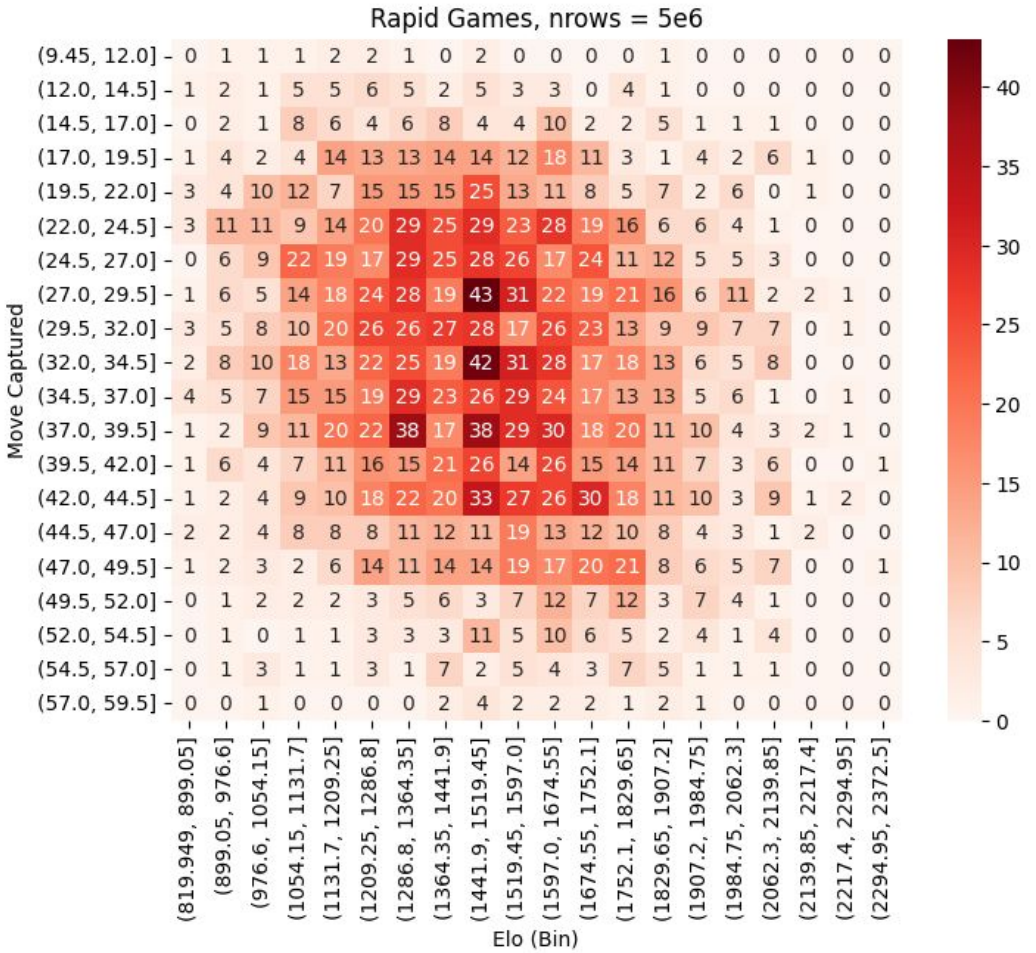
The mean and median are both negative, indicating the person making the check tends to be losing (and likely believes they are losing).



Question 5 - Do more experienced players keep their queens on the board for longer?



Question 5 - Do more experienced players keep their queens on the board for longer?



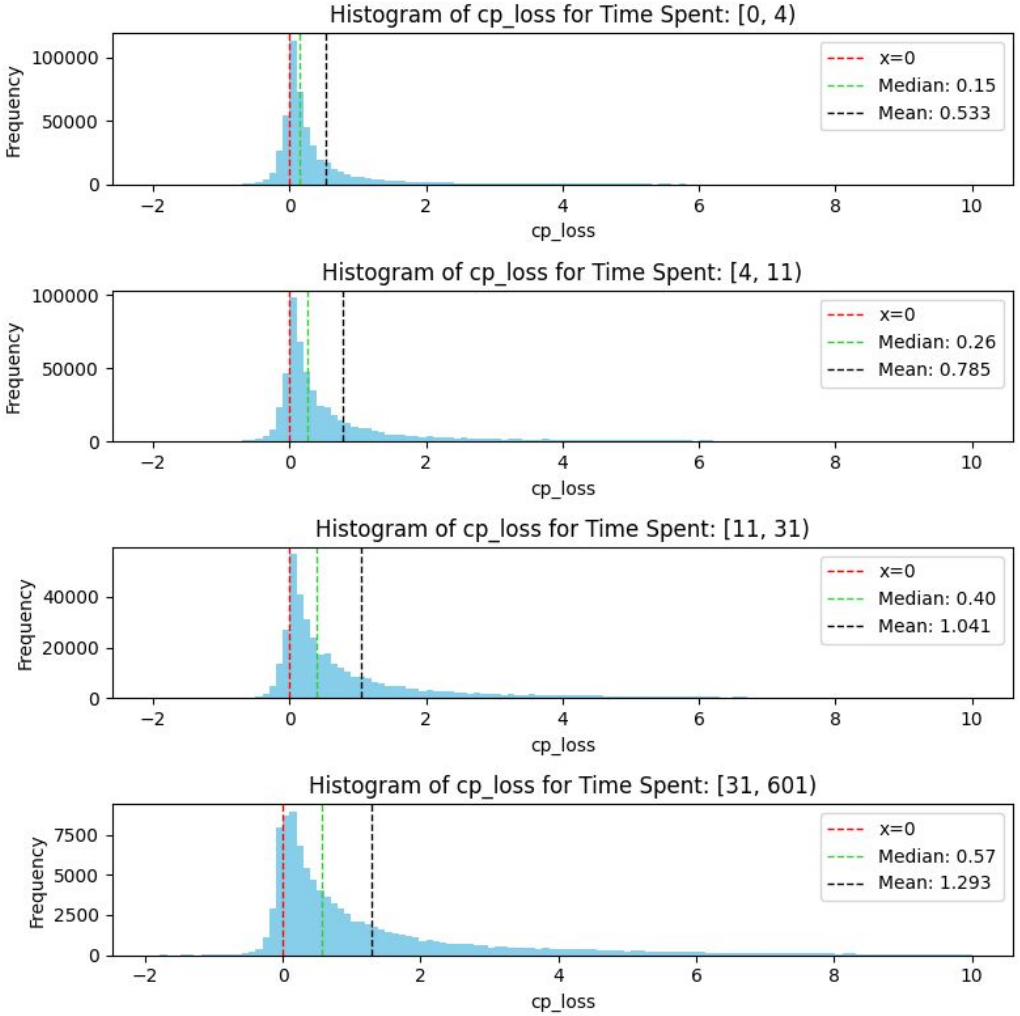
Conclusion

I found the following three interesting observations:

1. In bullet (1+0) games, thinking longer correlates with a higher probability of making a blunder.
2. In rapid (10+0) games, more experienced players use up more of their clock time.
3. The player who makes the first check more often than not ends up losing the game.

Thank you!

Question 2 - What is the relationship between time spent thinking and the change in evaluation caused by the move?



Question 2 - What is the relationship between time spent thinking and the change in evaluation caused by the move?