

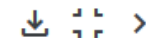
# Chess Patterns Through Data Visualization of 50,000 Lichess Games

---

Jeremy Neale

# Data: A. Revel from Kaggle

chess\_games.csv (4.38 GB)

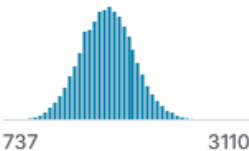
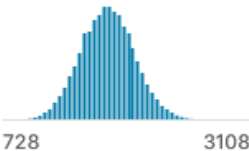


Detail Compact Column

8 of 15 columns ▾

## About this file

This csv file contains 6.25M chess games.

▲ Event Game type	▲ White White's ID	▲ Black Black's ID	▲ Result Game Result (1-0 White wins) (0-1 Black wins)	# WhiteElo White's ELO	# BlackElo Black's ELO	▲ Opening Opening name	▲ AN Movements in Movetext format.
Blitz 37% Classical 24% Other (2405818) 38%	118945 unique values	115946 unique values	1-0 50% 0-1 46% Other (240218) 4%			Van't Kruijs Opening 2% Scandinavian Defe... 2% Other (6010845) 96%	6188849 unique values
Classical	eisaaaa	HAMID449	1-0	1901	1896	Slav Defense	1. d4 d5 2. c4 c6 3. e3 a6 4. Nf3 e5 5. cxd5 e4 6. Ne5 cxd5 7. Qa4+ Bd7 8. Nxd7 Nxd7 9. Nc3 Nf6 10. ...
Blitz	go4jas	Sergei1973	0-1	1641	1627	King's Pawn Opening: 2.b3	1. e4 e5 2. b3 Nf6 3. Bb2 Nc6 4. Nf3 d6 5. d3 g6 6. Nbd2 Bg7 7. g3 Be6 8. Bg2 Qd7 9. O-O O-O 10. c3 ...
Blitz tournament	Evangelistaizac	kafune	1-0	1647	1688	Scandinavian Defense: Mieses-Kotroc Variation	1. e4 d5 2. exd5 Qxd5 3. Nf3 Bg4 4. Be2 Nf6 5. Nc3 Qd8 6. O-O Nc6 7. Bc4 e6 8. h3 Bh5 9. d3 Bd6

# Feature Engineering

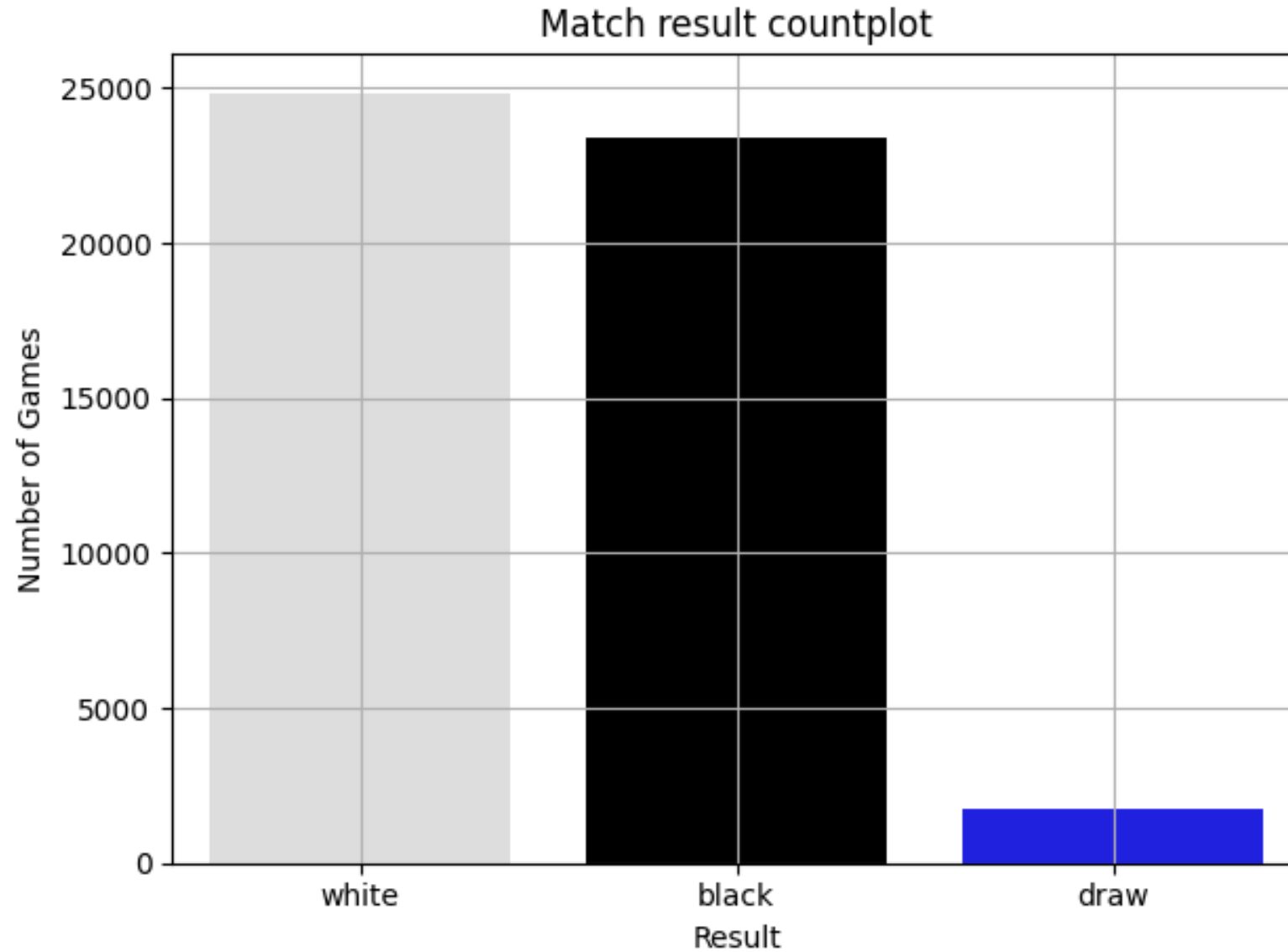
- Created 19 new columns
- Main ones:
  - First turn someone **castles** (minimum: 4 turns)
  - First turn someone **captures** a piece
  - First turn someone moves their **queen**
  - Type of castle each player performed (Kingside, Queenside, Neither)
  - Total checks
  - Total captures (maximum: 30)
- ‘1. e4 e5 2. Nf3 **Q**e7 3. Bb5 Nc6 4. **O-O** N**x**e4 ...’

# Feature Engineering

All new features created

- `first_castle_color` – The color of the first player to castle.
- `first_castle_turn` – The earliest castle turn in the game by either player.
- `white_castle` – The type of castle performed (long, short, or neither) by white.
- `black_castle` – The type of castle performed (long, short, or neither) by white.
- `first_white_capture` – The turn of the first white capture.
- `first_black_capture` – The turn of the first black capture.
- `first_capture_turn` – The earliest capture turn in the game by either player.
- `first_capture_color` – The color of the player who made the first capture.
- `num_white_captures` – Total number of captures by white.
- `num_black_captures` – Total number of captures by black.
- `total_captures` – Total number of captures in the game (white + black).
- `num_white_checks` – Total number of checks by white.
- `num_black_checks` – Total number of checks by black.
- `total_checks` – Total number of checks in the game (white + black).
- `first_white_queen_move` – Turn of white's first queen move.
- `first_black_queen_move` – Turn of black's first queen move.
- `first_queen_move_turn` – Earliest queen move turn in the game by either player.
- `first_queen_move_color` – The color of the player who moved the queen first.

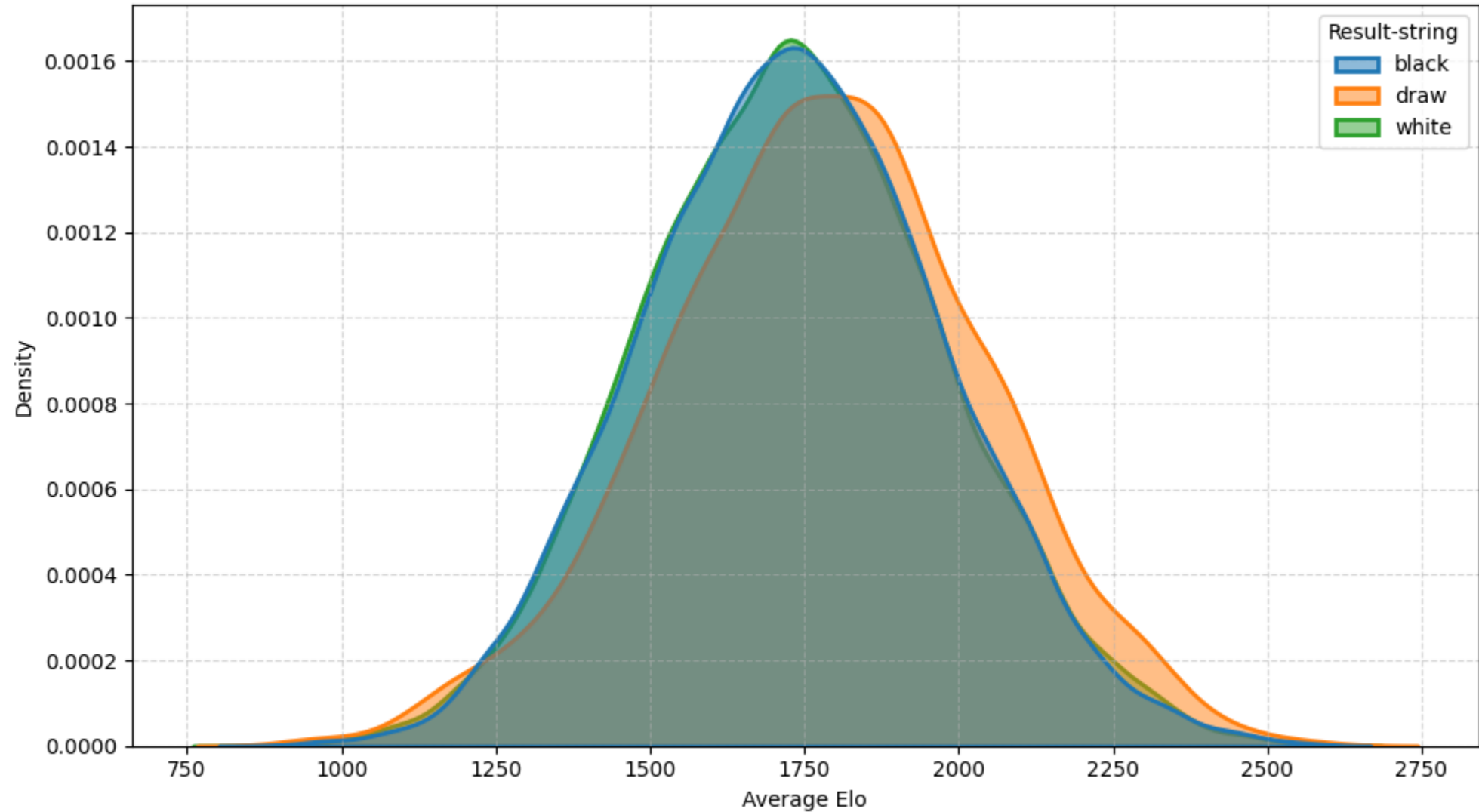
# Simple Bar plot



**Conclusion:** White wins at a higher rate than black, and draws are relatively uncommon.

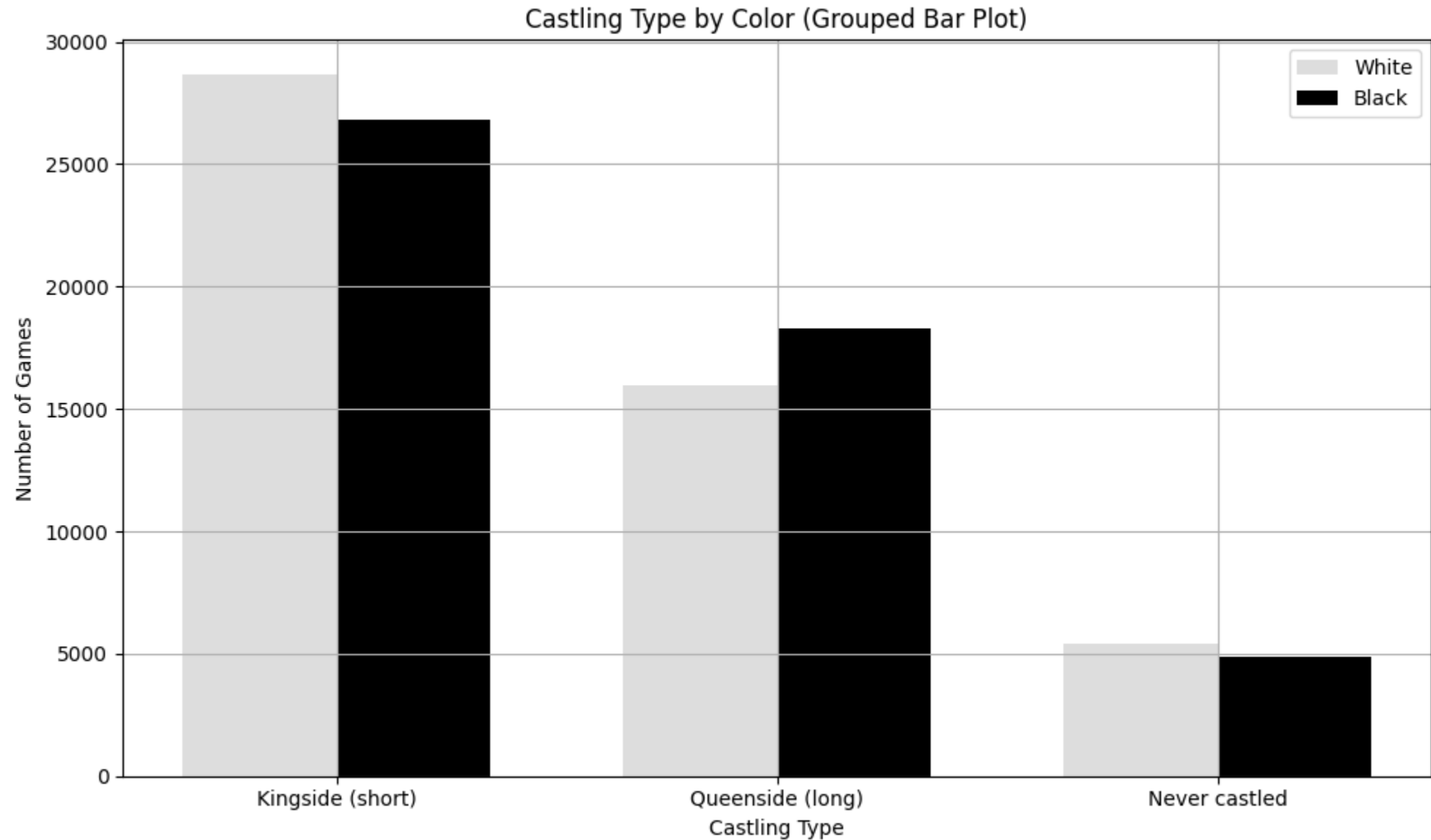
# Area plot

Elo Distribution by Game Result (Area Plot)



**Conclusion:** Higher rated players tend to draw games more often.

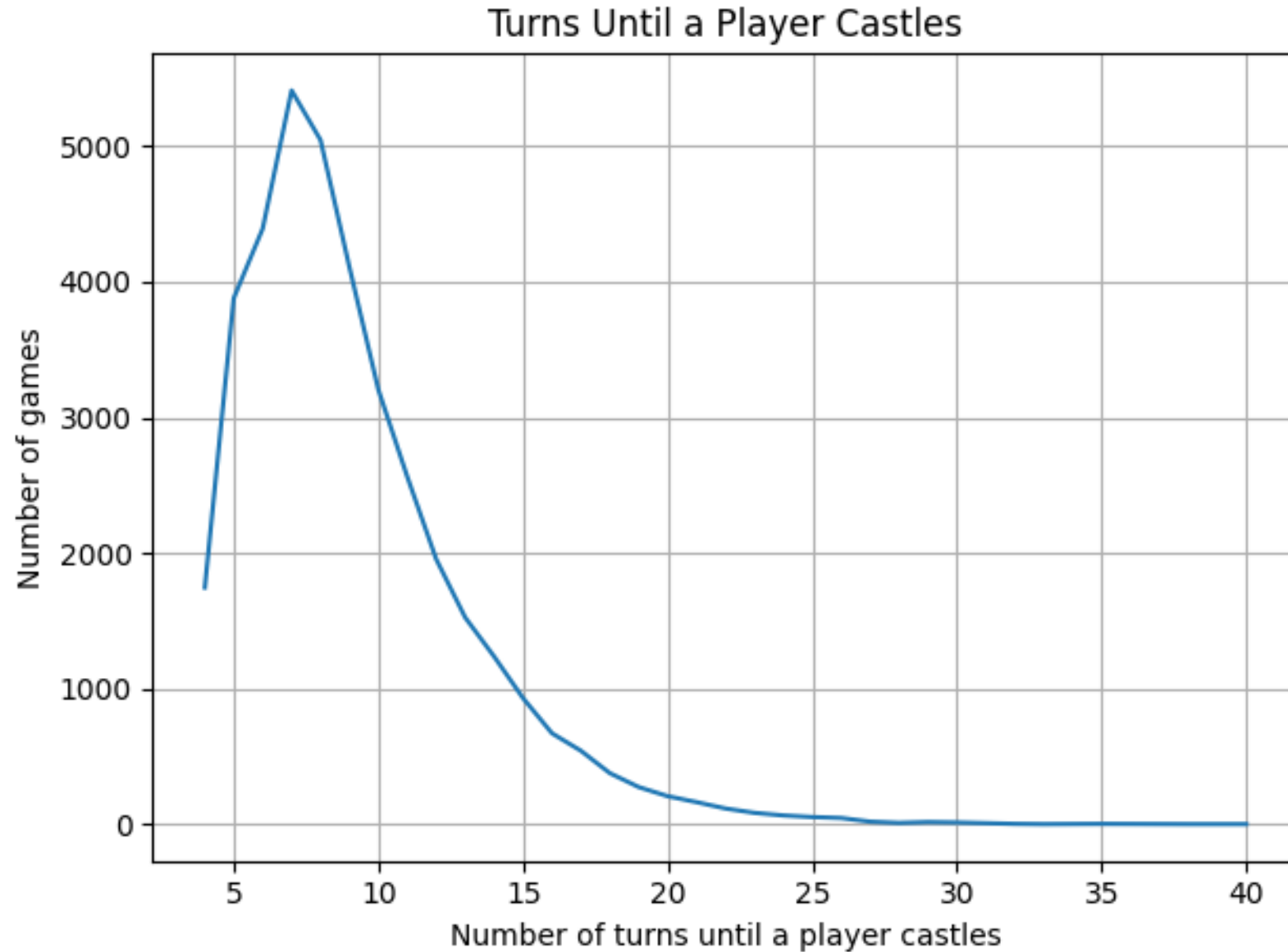
# Grouped bar plot



**Conclusion:** Black tends to castle queenside more often than white.



# Line plot

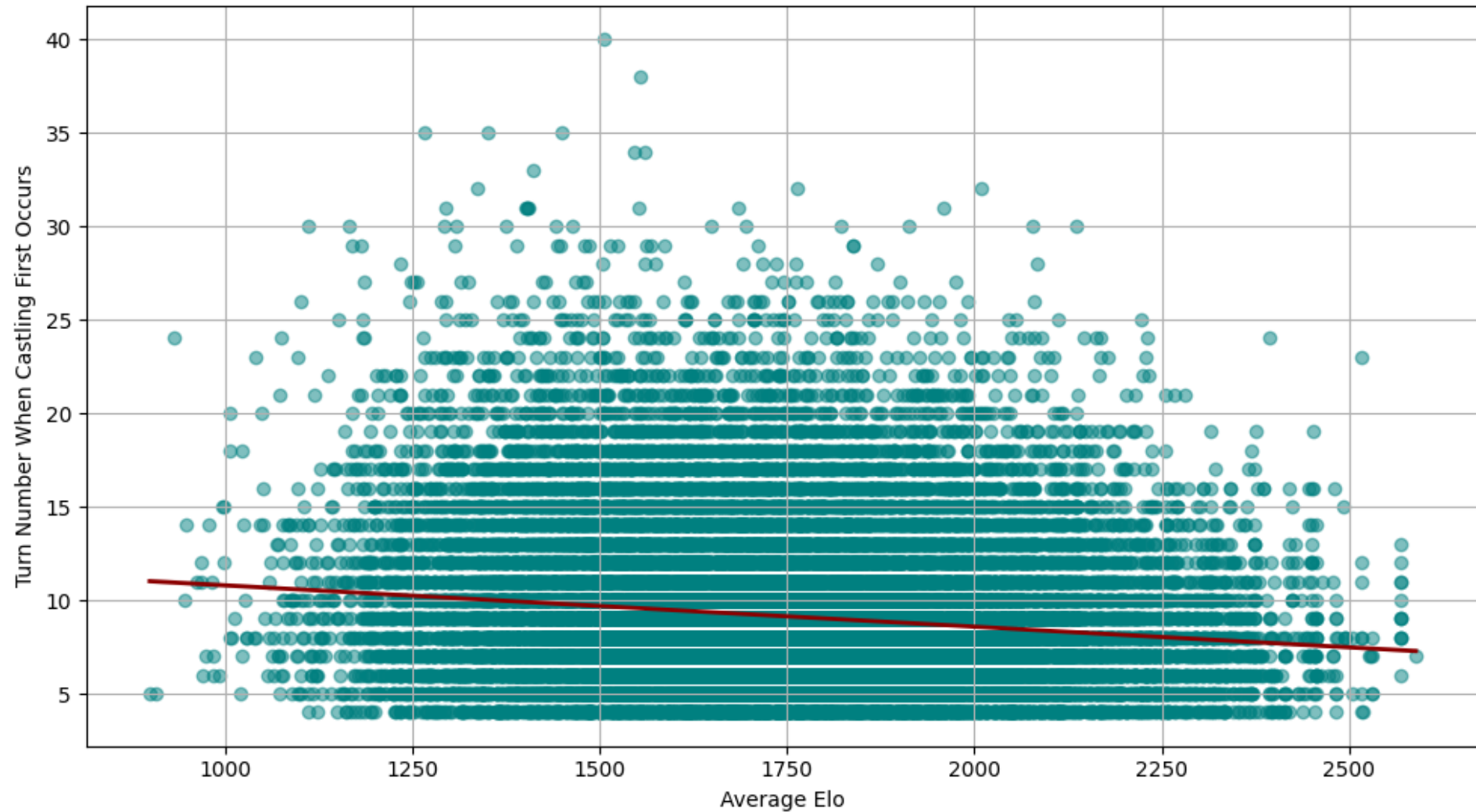


**Conclusion:** Turn 7 is the most common turn for either player to be the first one to castle.



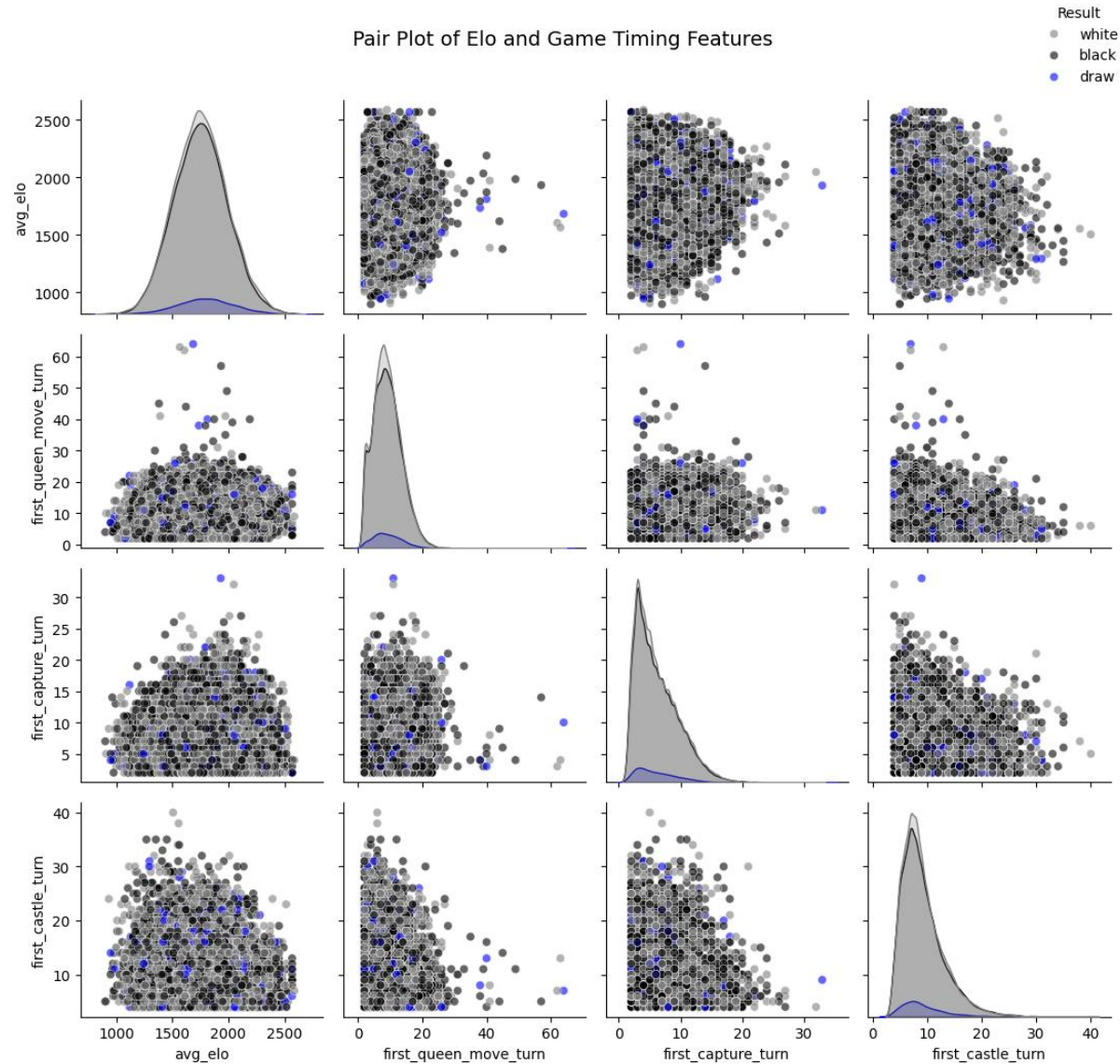
# Reg plot

Avg Elo vs. First Castling Turn



**Conclusion:** Higher rated players tend to castle earlier.

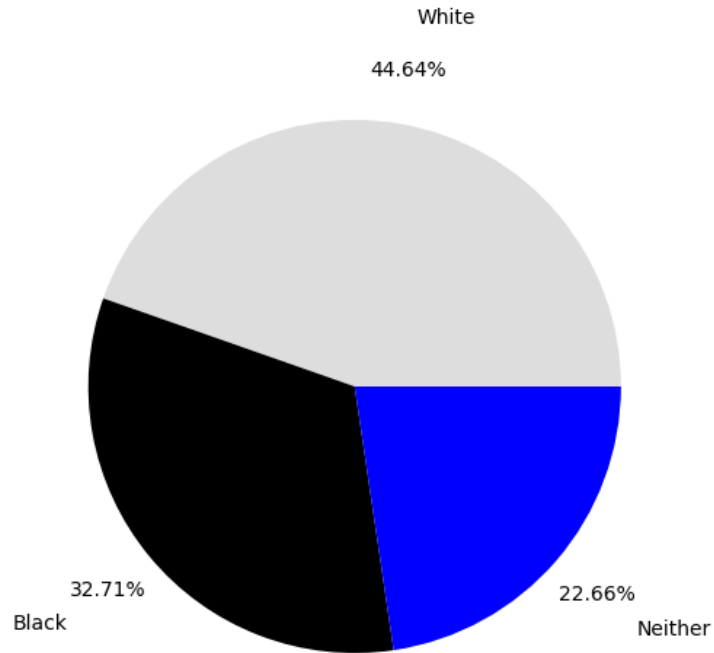
# Pair plot



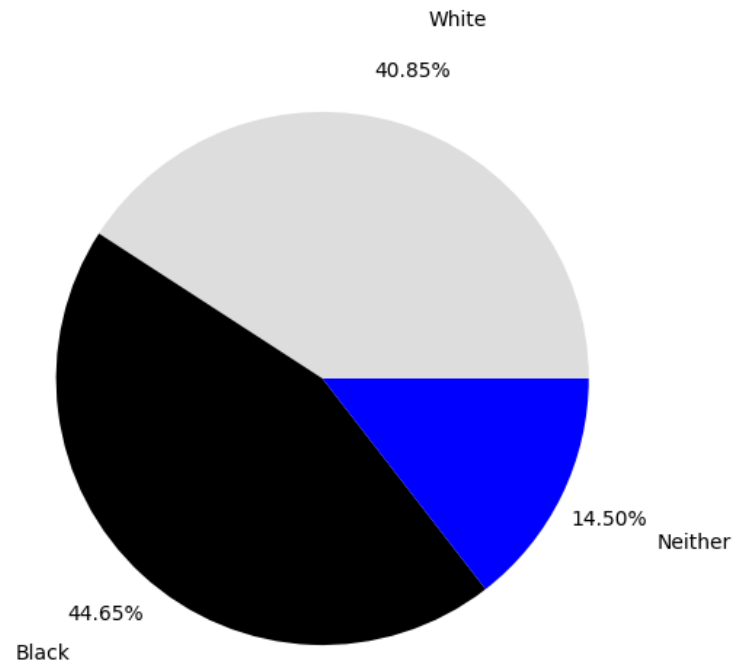
**Conclusion:** First capture and first castle are negatively related, and first queen move and first castle are negatively related.

# Pie charts

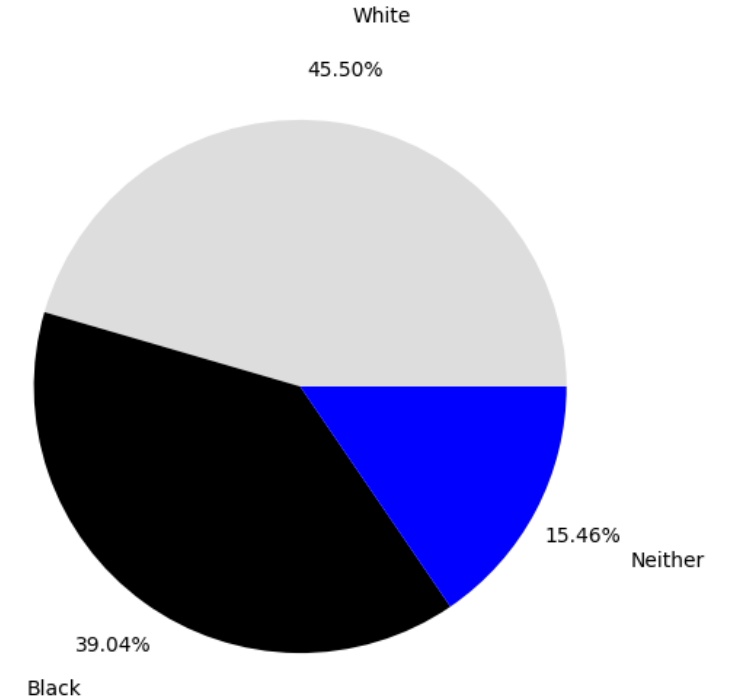
Who Castled First



Who Captured First

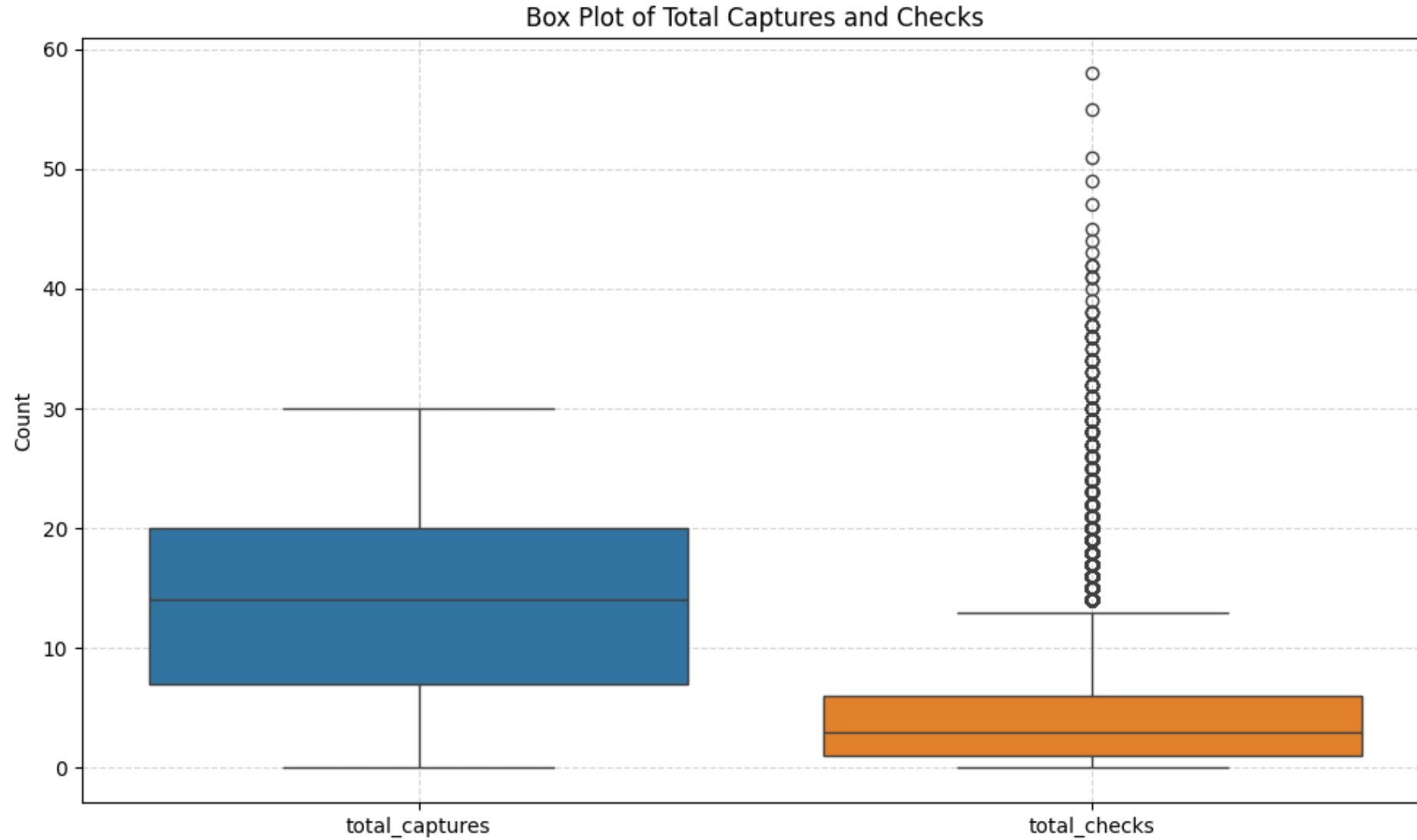


Who Moved Queen First



**Conclusion:** White tends to castle and move their queen first, but black tends to capture the first piece.

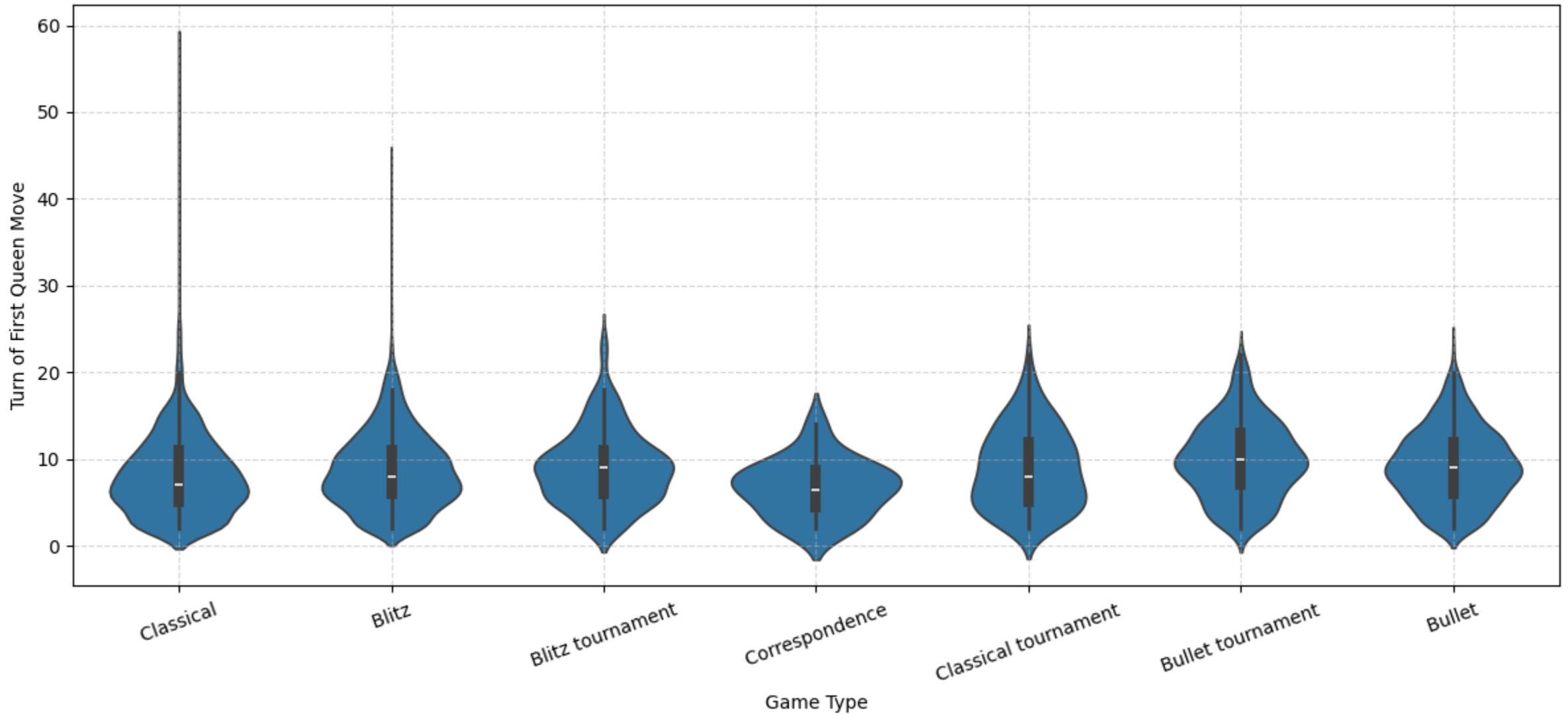
# Box plot



**Conclusion:** Total captures is roughly evenly distributed, while total checks are skewed right with outliers.

# Violin plot

Turns Until a Queen is Moved by Game Type



**Conclusion:** A queen is brought out earlier in longer games like Classical compared to short games like Blitz and Bullet.