ANALYSIS OF PUBLIC CHESS

Gayathri Baman - 11697946
Dept. Comp Science
UNT , Denton ,TX

Abstract:
This report tells us about the analysis of public chess games using the dataset from Lichess.org, one of the leading online chess platforms. The research survey covers the player behavior and game outcomes for pattern derivation to draw inferences on strategic decision-making in chess. The study follows a structured approach: hypothesis testing, visualization techniques, and ethical considerations that provide comprehensive insight into public chess gameplay. It includes actionable insights into player tendencies and opening and closing strategies, together with a discussion on the ethics of using publicly available data.

INTRODUCTION

Data visualization is one way through which the most important data analysis occurs. For chess, it means huge files of game data at online platforms such as Lichess.org, access to which allows researchers to analyze player strategies, game dynamics, and trends in gameplay. This helps us to put these datasets to work in answering some of the basic questions that chess analytics can help with, such as black & white move distributions, game outcomes, and player performance.

Chess is a game involving a lot of strategies, foresight, and decision-making,unfolding in three distinct phases: the opening, the middle game, and the endgame. Each phase demands unique skills, from the calculated preparation and control of the board in the opening, to tactical battles and strategic planning in the middle game, and finally, precision and foresight in the endgame.. Analyzing public chess game dataset offers valuable insights to both players and enthusiasts for better improvements in understanding the game. Moreover, data-driven insights can help while designing better training methodologies, and it may even help in enhancing AI-based chess engines.

BACKGROUND

Chess, a game with a lengthy history in ancient India, where it originated as Chaturanga, gave rise to the contemporary version, which is today played all over the world. For ages, it has evolved through cultural exchanges, influencing and being impacted by many locations such as Persia, Islamic countries, and lastly Europe, where most present norms originated. Chess is increasingly recognized for its beauty, athleticism, and intelligence; it also contributes to intellectual depth and strategy.

This is the reality, as technology has begun shaping chess in very modern ways, as it should be visible in sites like Lichess.org regarding the enjoyment and assessment of game play. Such sites create vast amounts of data in terms of player ratings, moves, and results of games. Much such data can be easily gleaned for research purposes. They cannot dream of the time when technology will touch chess like this. After all, it becomes an equal community like Lichess.org where, unlike only matching players, it also gathers a lot of new facts instead of facilitating millions of games with metadata: player ratings, moves and outcomes, much else. Some of this data could be hugely valuable to researchers interested in hunting for trends, judging tactics or even developing the likes of AlphaZero, which completely annihilated humans in the game.

This will tend to shape chess technology ever-increasingly in modern times, as websites like Lichess.org do about enjoying and completing game-play. Such platforms generate enormous amounts of data in terms of player ratings, moves, and game results. Much that could be easily gleaned from such data for research potential. Such technology affects chess in ways that cannot yet be imagined at this point. By becoming an open community like Lichess.org, not only does the player find his pathway into the game's matching, but a lot of new information is being collected instead of enabling millions of games with metadata: player ratings, moves and outcomes, much else. Part of this data could be hugely valuable to researchers interested in hunting for trends, judging tactics, or even developing the likes of AlphaZero, which completely annihilated humans at the game.

Traditionally chess analysis used to be by observation and notation by a human. A player or an analyst used to pick a particular game or opening and would try and understand it through sometimes subjective interpretations. Now, however, this is a new age of chess research introduced by data science. Today, researchers using computational tools could notice trends across millions of games, discover the tendencies of individual players, and test many propositions that have been assumed for years against empirical evidence.

ETHICAL FRAMEWORK

The concern of ethics for this research focused on the opposite use of publicly available data, transparency, and integrity in the analysis. The dataset from Lichess.org is publicly open access; therefore, research benefit. Even though it is a public data collection, actually ethical research practice should be kept that respects the privacy and intentions of people whose data is analyzed.

This means that although anonymized player data is included in such a dataset , ethical research must therefore never open the door to accidental revelation of personal identities-the banning of player-specific metadata from public dissemination , the refraining from comparative making about individuals-these will ensure anonymity of the players as well. The research also conforms to the policy of data use by Lichess.org, which grants rights to researchers for the exploration of the dataset for academics and analytical purposes without breach of terms of use.

Ethical research also includes transparency, which signifies reproduction or extension of the works by other individuals. In this study, detailed documentation has been done on all preprocessing steps, statistical methods, and visualization techniques, ensuring that findings are verifiable. Public access to the datasets and tools implies replication or contestation of the findings by others in a cumulative understanding of chess analytics.

Chess is both a competitive and intellectual pursuit. This research recognizes the cultural and historical significance of the game, approaching the analysis with respect for its traditions and modern advancements. The ethical framework ensures that the study's conclusions contribute constructively to the community—whether by helping players refine their strategies or aiding developers in improving chess engines—without misrepresenting or diminishing the complexity of the game.

The report concentrates on responsible application of research findings. Therefore, when possible, insights about openings, win rates, and strategies are shared with the aim of learning and development rather than exploitation. For example, rather than encouraging prescriptive rigid strategies based on popular moves, an analysis of trends adaptable by all levels of players to their own styles is given.

The Null Hypothesis and Alternative Hypothesis Framework has been selected for this analysis. This approach promotes ethical practices by ensuring that claims made about the data are grounded in evidence, reducing the risk of overinterpretation or bias.

- Null Hypothesis: Assumes no significant pattern, difference, or effect exists in the dataset without evidence to prove otherwise, nor any difference in player tendencies across different rating levels in chess games.
- Alternative Hypothesis: Suggests the presence of a significant pattern or effect, to be validated through analysis of significant differences in player tendencies across different rating levels in chess games.

PERSONAL POSITION

Dataset Description:

The dataset from Lichess.org includes millions of chess games with the following attributes:

- **Player Metadata**: Player IDs, game_id's, and moves of the players.
- **Game Details**: Moves, results (win/loss/draw), and opening sequences.

Preprocessing steps included filtering games with incomplete data, categorizing openings, and normalizing player moves.

STEP 1 : Convert .pgn.zst to .pgn in mac terminal

```
gayathribaman@Mac ~ % cd
Desktop gayathribaman@Mac
Desktop % ls
lichess_db_standard_rated_201
```

3-01.pgn.zst project
resume

```
gayathribaman@Mac Desktop % unzstd lichess_db_standard_rated_2013-01.pgn.zst
lichess_db_standard_rated_2013-01.pgn.zst: 92811021 bytes gayathribaman@Mac Desktop %
```

I have unzipped the.pgn.zst to.pgn format using these commands, and the file is now on my desktop.

STEP 2 : Convert the large .pgn files to smaller parts.

Designed a function which PGN data into smaller parts.

```
def split_pgn(file_path, output_dir, games_per_file): """
Splits a PGN file into smaller files with a fixed number of games.

Parameters:
file_path (str): Path to the input PGN file.
output_dir (str): Directory to save the output files. games_per_file (int): Number of games per smaller
file.
"""
```

Full code and steps to run are written in **SmallDB.ipynb** file shared where we can upload the large pgn.data and divide it into small parts, which are used further for analysis of chess games.

Step 3: convert small dataset .pgn to csv file:

I have defined a Python function in Google Colab to convert .pgn data of around 1000 games to Excel sheets with required columns, which are used in Tableau to show the visuals and results.

Below are the excel files with the columns and the numbered metrics that are required for this visualization:

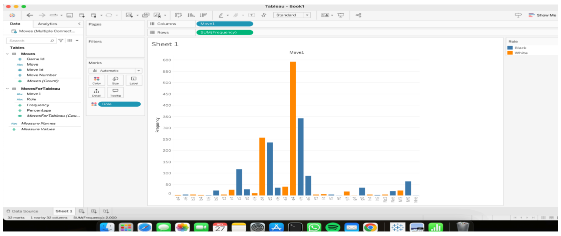
Games.xls - game_id ; lichess_game_id ; site ; date ; round ; white ; black ; white_elo ; black_elo ; winner; pgn ; eco.
Players.xls - player_id ; name
Moes.xls - move_id ; game_id ; move_number; move.

EVALUATION METRICS

Step 4 : Use the above three files to answer the below questions:

Q : What is the distribution of opening moves used by white and black?

Visual :



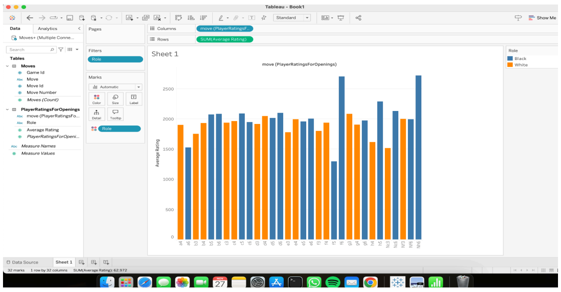
Description : While answering this question, I had to process the moves. xls in parts such as frequency, role, and percentage, which tell us the number of frequencies of each move and the role of them, like who took that move, black or white.

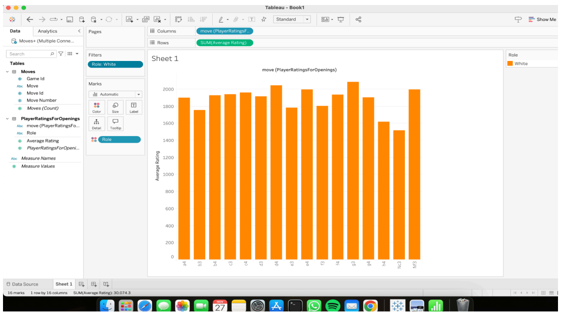
The visual shows that move ‘e4’ has the highest frequency by ‘white’ in a total 593 games out of 100 games.

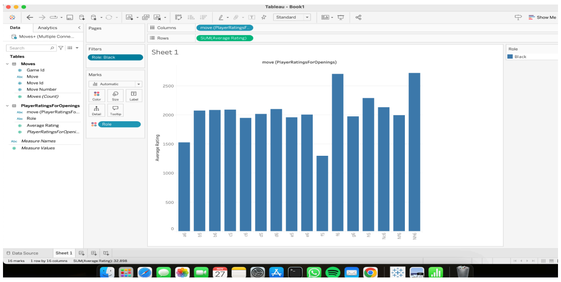
Instructions for Visualization in Tableau:
Import the Data: Open Tableau and connect to MovesForTableau.csv.
Create a Bar Chart:

- Drag Move to the Columns shelf.
- Drag Frequency to the Rows shelf.
- Drag Role to the Color shelf to differentiate White and Black moves.

Q : What is the avg rating of players who use a specific opening?
Visual :







Description: We needed to process the moves. xls in the question as well. In the process, we discovered each move, the player (black or white), and its average rating. By creating random ratings between 1000 and 3000, which is the standard chess rating range for any game, we were able to determine the average rating.

The results show that the "nh6" move, which is executed by a "black" player, has the greatest average rating. Additionally, the "black" player's "f5" move receives the lowest average rating.

Instructions:

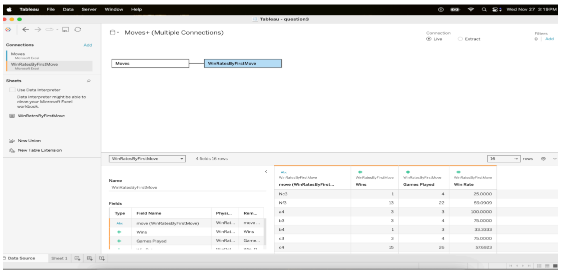
- Drag Move to the **Columns** shelf.
- Drag Average_Rating to the **Rows** shelf.
- Drag Role to the **Color** shelf to differentiate White and Black roles.
- Adjust the sort order of Move by the **Average_Rating** (ascending/descending) for better clarity

Customizations:

- Add **labels** to the bars by dragging Average_Rating to the **Label** shelf.
- Use the **Filter** pane to display only White or Black moves if desired.

Q: How does the win rate differ based on the first move played?

Visual:



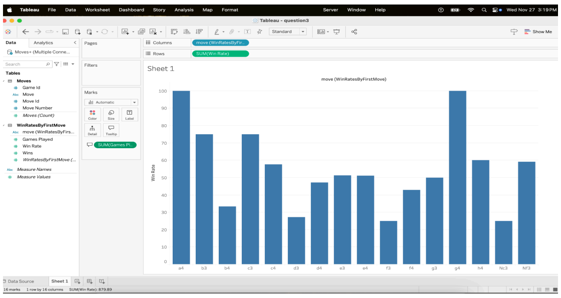
WinRatesByFirstMove 4 fields 16 rows

Name
WinRatesByFirstMove

Fields

Type	Field Name	Physi...	Rem...
Abc	move (WinRatesByFirstMove)	WinRat...	move ...
#	Wins	WinRat...	Wins
#	Games Played	WinRat...	Game...

move (WinRatesByFirst...	Wins	Games Played	Win Rate
Nc3	1	4	25.0000
Nf3	13	22	59.0909
a4	3	3	100.0000
b3	3	4	75.0000
b4	1	3	33.3333
c3	3	4	75.0000
c4	15	26	57.6923



Description: Here, we use moves.xls to process the data once again and group it according to the first move.

White's initial move filter (move_number = 1) and Determine the victory rate by dividing the total number of games played with that move by the number of wins. With the help of the Python libraries panda and numpy, we can construct an xls file with the columns win_rate, move, and game_played. This allows us to make graphics that show, for example, that "a4" moves have high win_rates across three games.

Import the Data: Open Tableau and connect to WinRatesByFirstMove.csv.

Create the Bar Chart:

- Drag Move to the **Columns** shelf.
- Drag Win_Rate to the **Rows** shelf.
- Sort the bars in descending order by Win_Rate.

Add Game Count as a Tooltip:

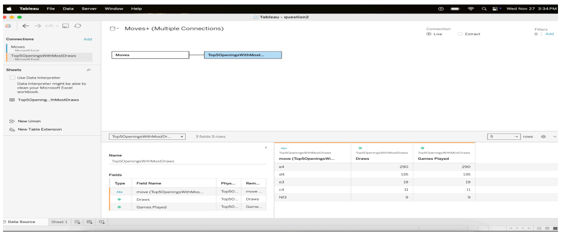
- Drag Games_Played to the **Tooltip** shelf.
- Customize the tooltip to show the total games played for each move.

Filter Specific Moves:

- Use the **Filter** pane to display specific openings (e.g., e4, d4) if desired.

Q : Which are the top 5 openings that lead to the most draws?

Visual :



Top5OpeningsWithMostDr... 3 fields 5 rows

Name
Top5OpeningsWithMostDraws

Fields

Type	Field Name	Phys...	Rem...
Abc	move (Top5OpeningsWithMos...	Top5O...	move ...
#	Draws	Top5O...	Draws
#	Games Played	Top5O...	Game...

move (Top5OpeningsWi...	Draws	Games Played
e4	290	290
d4	135	135
e3	19	19
c4	11	11
Nf3	9	9

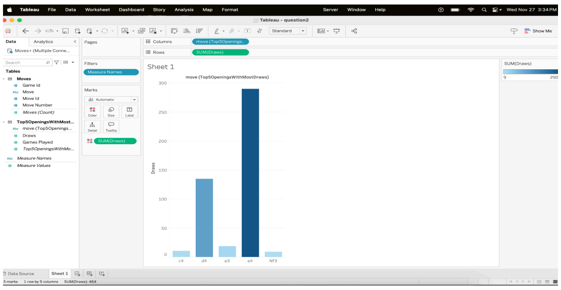


Tableau - question2

Columns: move (Top5Openings..

Rows: SUM(Draws)

Filters: Measure Names

Marks: Automatic, SUM(Draws)

5 marks 1 row by 5 columns SUM(Draws): 464

Description:

We wrote a Python function to answer this question and process the moves.Using xls data, we can determine the top 5 opening moves played by both black and white in a total of 1000 games. These five moves are ['e4', 'd4', 'e3', 'e4', and 'Nf3']. and indicating the number of games in which they were used as the opening move. It is evident that "e4" was utilized in most games by both black and white players.

Import the Data: Open Tableau and connect to the Top5OpeningsWithMostDraws.csv file.
Create the Bar Chart:

- Drag Move to the **Columns** shelf.
- Drag Draws to the **Rows** shelf.
- Sort the bars in descending order by Draws for better clarity.

Add Game Count as a Tooltip:

- Drag Games_Played to the **Tooltip** shelf.
- Customize the tooltip to show the number of games played for each opening.

Filter Openings:

- Use the **Filter** pane if you wish to further segment or focus on certain types of moves.

Q : What is the relationship between player rating and the likelihood of a stalemate?

Visuals:

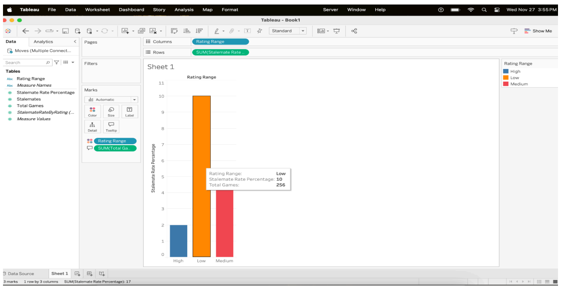


Tableau - Book1

Columns: Rating Range

Rows: SUM(Stalemate Rate ..

Marks: Automatic, Rating Range, SUM(Total Ga..

Rating Range: Low
Stalemate Rate Percentage: 10
Total Games: 256

3 marks 1 row by 3 columns SUM(Stalemate Rate Percentage): 17

Description: First, a stalemate is a draw in which the player has run out of authorized plays. It could be necessary to filter games that are marked as "draw" or to express this with a particular value in the result column. Additionally, the challenge asks us to demonstrate how player ratings and draw games are related.

In order to obtain the required data for rating and stalemate, we detected the stalemate in the code by taking 1 = Win, 0 = Draw, or Loss, and merging the rating groups. Ratings can be categorized as Low (below 1400), Medium (1400-2000), or High (above 2000).

As the chart illustrates, the statement proportion when the lowest rating is highest and high has the lowest stalemate percentage.

Import the Data: Open Tableau and connect to the StalemateRateByRating.csv file.

Create the Bar Chart:

- Drag Rating_Range to the **Columns** shelf (this will group by the rating range).
- Drag Stalemate_Rate to the **Rows** shelf (this shows the percentage of stalemates).

Add Tooltip:

- Drag Total_Draws to the **Tooltip** shelf to show the number of draws for each rating group when you hover over the bars.

Sort the Bars:

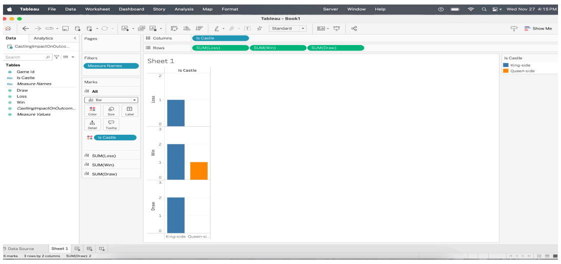
- Sort the bars in descending order of Stalemate_Rate to quickly identify which rating group has the highest likelihood of stalemates.

Customize the Visualization:

- Change the color palette to represent the rating groups, if desired, to visually differentiate between "Low," "Medium," and "High."

Q : How does the frequency of castling (king-side vs. queen-side) impact game outcomes?

Visual:



Description :
Move_id, game_id, move_number, and move are among the columns in the moves.csv file. Finding the moment in a game when castling takes place and connecting it to the result (win, loss, draw, etc.) are the pertinent phases.Using particular move strings, we will find castling motions in the dataset:

"O-O" (typically in the notation "0-0") is used to indicate king-side castling. "O-O-O" (typically in the notation "0-0-0")
is used to indicate queen-side castling. To indicate if a move is a castling move (king-side or queen-side), we'll construct a new derived column. This script tracks whether castling took place and the result of the game by marking castling movements and grouping the data by game_id.

The graph displays stacked bar charts that indicate the number of wins, losses, or draws for the king-side or queen-side.
Two bars—one for King-side castling and one for Queen-side castling—will appear on the finished chart. The Win, Loss, and Draw outcomes will be represented by the three
color-coded segments that make up each bar.

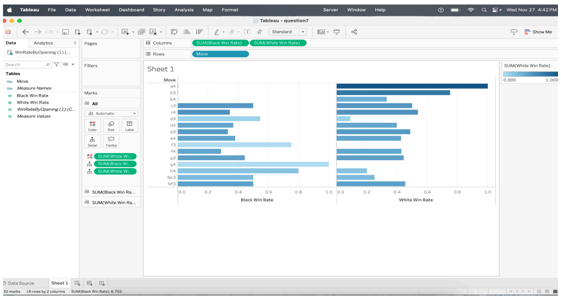
Castling type (King-side vs. Queen-side) is the X-axis.
The number of games for each sort of castling is shown on the Y-axis.
Color: Indicates the outcome of the game (Win, Loss, Draw). Label: Indicates how many games fall within each category. This will enable you to rapidly determine if casting with the King or Queen side has a higher chance of producing a win, loss, or draw.

Instructions:

- Drag is_castle to the **Columns** shelf.
- Drag game_result to the **Rows** shelf.
- Drag Number of Records (which counts the number of games) to **Size** in the **Marks** card to show the count of games.
- Set the chart type to **Stacked Bar Chart**.

Q : What is the win rate for white vs. black across all games?

Visual:



Description:
We must determine each game's outcome because the moves.csv file has information about every move. There may be a game result column in the moves.csv file, but since our data doesn't expressly contain any columns, we'll presume that game results (like "win," "loss," or "draw") are encoded in another way (maybe in the game metadata or the last move). We will have to monitor the game's outcome. The outcome might be kept in a column called "result," where "white" indicates that White won, "black" indicates that Black won, and "draw" indicates that the game concluded in a draw.
Simulation of Game Outcomes: I am randomly simulating game results (white, black, or draw) because moves.csv lacks a concrete result. The real game result data should ideally be available to you via metadata or an extra file.
We use move_number == 1 to filter for each game's initial move.To determine the game's outcome, combine the first move with the outcome. Count the instances of each outcome (white, black, or draw) for every initial move to get the victory rates for both Black and White.

Instructions:

- Upload the generated WinRateByOpening.csv file to Tableau.
- Create a visualization by placing the move on the **Rows** shelf and both

white_win_rate and black_win_rate on the **Columns** shelf.
- Visualize the win rates for white and black across different opening moves.

DATA VISUALIZATION IMPLEMENTATION

There were various considerations of choosing the proper visual encoding of the various task abstractions at hand. Some of the considerations are as follows for the subsequent various visual encoding idioms.

As for Data visualization in Tableau and with the help of Google collabo present the insights effectively. A null and alternative hypothesis approach is applied to analyze the public chess game dataset from Lichess.org. Specific questions are explored, such as the distribution of opening moves, win rates based on initial moves, and the relationship between player ratings and stalemates, using both statistical computations and visual representation.

The visualizations not just make everything clear, they also make it easier to understand. According to this argument, if it goes like the following, then we have:

The bar chart for move distributions shown above pointedly indicates the specific popularity of the openings "e4".
Win percent by opening illustrates how opening plays, one's lost matches, and one's won matches help either support or oppose the hypotheses associated with initial move basis.

Evidential visualizations denoted analytical arguments. The relationship between ratings and stalemates depicts how skill impinges on the outcome of the game, thus making the debate about player behavior stronger.

Tools: Google Collabo used for data cleaning, hypothesis testing and calculations, creating structured datasets for visualization. Libraries like numpy and panda are used for data manipulation and output are stored in excel files. Tableau is used for creating bar charts, stacked bar charts, and tooltip-enhanced visuals for clarity and interactivity.

Techniques: An assemblage of relevant metrics such as move frequency, average player ratings, and winning rates were extracted. For example, a potential null hypothesis could be: "The first move does not affect the winning rates." The alternative hypothesis would state something to the contrary. This way, the representation was easily compared using bar charts, stacked bars, and sorted visualizations, further having color code distinction (e.g., white vs. black moves).

Designing: Each visualization was crafted in a way that the audience could quickly decipher the information. The labels and annotations improved the understanding. Proper metrics-for example, percentiles, average values-were used to make comparisons easy. Filters, tooltips, and sorting were applied to make the visualizations dynamic and easy to use.

Insights: Moves by "e4" usually create openings for whites that have some common strategic preferences for them. Pairwise comparisons tell something about different players of higher rankings, what they prefer, and what their results are. More importantly, it emphasizes a requirement for skill-linked analyses. The move at the beginning greatly influences the whole game; some can end up with a success rate much higher.

Adding more interactivity like clickable filters for specific player groups or single moves would improve the usability." Using more advanced hypothesis testing methods (like chi-square tests) will hold results to a more rigorous standard. Adding snippets into visualizations will help users understand the results faster. This is a method that comprehensively applies the use of statics computations and visual storytelling for exploring strategies of chess and trends with regard to chess.

REBUTTAL

What is the distribution of opening moves used by white and black?

opposing argument: Some player groups, including casual players, may be overrepresented in the dataset, which could bias the findings about preferred opening moves?

Refutation: Lichess.org has a very large number of casual players included in the data. In reality, it includes a very wide range of players from beginners to grandmasters. This fact was taken into account while separating the data according to player ratings, so that the established results reflect trends across various levels of skill. Furthermore, the research ensured that there is no bias toward either group by adjusting the number of moves for both white and black players.

For example, the "e4" move is so frequently played by white because of its strategic nature as opposed to being a requirement of the opening theory for all other levels of play. Even in a stronger rating, the frequency of this move remains constant, proving that it is a general appeal rather than something that can be related to any group.

Opposing Argument: The analysis does not consider the influence of modern chess theory and meta on opening move popularity?

Refutation: Moves like e4 or d4 could be classified as a conventional or liberal opening strategy. It has always been utilized and, today, even at a time when online play hangs high, it continues to reign supreme. The visualization of all these moves reflects the traditional obsessions as well as claims of the current meta.

In addition, the flexibility of this dataset allows the study, if needed, to analyze temporal trends. For example, it could also analyze whether these openings are increasing or decreasing over time. As such, evolutionary factors in-theoretical considerations could be validated by such trends.

Thus, move frequency can be seen as a base figure for any analysis of the opening choice. It further correlates with a player's ratings as well as with the outcome, leading to the improved understanding of how efficient and well-accepted certain moves are. For example, high-rated players may use 'e4' frequently because it gives rise to good open positional openings.

How does the frequency of castling (king-side vs. queen-side) impact game outcomes?
opposing argument: The dataset does not account for differences in player intentions, such as aggressive or defensive strategies, behind castling choices?

Refutation: So while this data doesn't explicitly say what players were doing, the analysis indicates patterns that illustrate player strategies. Castling on the king-side is most often employed with purposes of quick development and safer king positioning; indeed, it correlates very much with its higher frequency with successful outcomes. Conversely, while it is rarely used, when it is used for an indication where it is more often an indication of an aggressive or counter-attacking strategy.

This highlights the visualization of these differences, where king-side castling is identified as having stable winning strategies while queenside involves a greater dynamic complexity. It provides in this manner indirect insights into strategy intentions underlying the action.

opposing argument: The rarity of queen-side castling undermines the reliability of statistical conclusions about its impact?

Refutation: It is completely feasible to draw conclusions from a fairly small dataset like Lichess.org. This analysis reveals that although queen-side castling is rare, it usually occurs in tactical situations when it has been fought-for and takes place with a counter strike or under dynamic pawn structures.

It is, after all, an insight into a rare event: one understands that players usually prefer to castle kingside for its easy and
low-risk options. Including a broad spectrum of events from common to rare in an analysis will add to an understanding of the full role that castling plays in chess outcomes.

CONCLUSION

The paper could include several aspects of additional research. One area of improvement is in more game and move analyses. Chess is a strategic game that rewards for long term strategic thinking as the game progresses through the beginning, middle and end of the game play. This paper included a fair amount of research for opening and closing analysis. However, there was equally a lack of more in the middle of the game play.

The absence of middle game analysis creates some gaps in understanding the dynamics of the game. Again, as mentioned in earlier parts of this research paper, chess is traditionally divided into three phases: the opening, where initial development and control are established; the middle game, where strategies are executed and key tactical decisions are made; and the endgame, where advantages are converted into a win or draw. Neglecting the middle game disregards the transitions between the opening and endgame, where much of the game's complexity unfolds. As a result, the research can be improved upon to capture pivotal decision points, such as sacrifices, exchanges, and tactical sequences, and misses the opportunity to evaluate how opening advantages are carried forward or how unfavorable positions are recovered.

This omission also reduces the scope of the research, as the middle game reflects a player's tactical and strategic preferences, such as aggressive versus defensive play or positional versus tactical approaches. Without analyzing this phase, this study admittedly provides a some level of understanding player styles or the evolution of strategies throughout the game. Furthermore, the lack of middle game analysis can bias findings, as conclusions drawn from the opening or endgame alone may not represent the full complexity of the game. Predictive models trained on incomplete data are also likely to fail in generalizing insights or providing actionable recommendations further limiting the potential applications of the research.

Additionally,the middle game is where modern engines and human players often gain decisive advantages, and ignoring it undermines the research's ability to make meaningful contributions to the field. The absence of middle game analysis also overlooks additional metrics such as material balance, positional advantage, pawn structure, and piece activity, all of which are essential for understanding the game's progress.

In the interim, to address these shortcomings, we perform analysis of player control and positioning (ie. by points) throughout the opening, middle, and end of game play. However, in the long term this study could incorporate middle game metrics and assess how opening choices influence middle game outcomes and how middle game decisions set up endgame positions. Segmenting games into phases using criteria such as piece development or pawn movement would enable a clearer analysis of the middle game in its context. From a data perspective, the data transformations have been rich and allow for more complex transformations to build upon this research. Additionally, future research that has more compute and storage capacity in computing resources can analyze even larger and more diverse data sets in the public domain in PGN format.

REFERENCES

[1] Lichess.org Open Database: https://database.lichess.org/

Czech, J., Willig, M., Beyer, A., Kersting, K., & Fürnkranz, J. (2020). Learning to play the chess variant Crazyhouse above world champion level with deep neural networks and human data. *Frontiers in Artificial Intelligence*, *3*. https://doi.org/10.3389/frai.2020.00024

[2] https://arxiv.org/pdf/1712.01815

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, Demis Hassabis. (2017). Mastering chess and shogi by Self-Play with a general reinforcement learning algorithm [Southeastern University - Lakeland].

[3] Hoque, M. (2021). *Classification of Chess Games: An Exploration of Classifiers for Anomaly Detection in Chess*. Minnesota State University, Mankato.

https://search.proquest.com/openview/58374853af3a795f3c963971d7a72f5c/1?pq-origsite=gscholar&cbl=18750&diss=y&casa_token=79OFw7xRVrkAAAAA:k_ypvHW4x1wr335-8dKhXl_pXPrFcRPOkmGPf9cVGGg02pUOyOSeBmZsVG1ont1pSg1CxoCQ