Popular Chess Moves

"Is this allowed"???

NO

Sam Zandiasadabadi

Introduction

- **Using chess.com/explorer, identified the most popular moves**
- **Studied their effectiveness and repetition**
  - Exactly how many times a move has been played?
  - Why are they played in the first place?
  - What is the probability someone would play the move compared to the others?
- **Learning new skills to complete the project**
  - R Programming Language on Datacamp
  - Statistical identities and skills in SFSU
- **Used multiple IDEs to complete the project**

The Project

- **Most common moves are:**
  - e4 -> played 1,366,195 times
  - d4 -> played 1,050,651 times
  - Nf3 -> played 299,740 times
  - c4 -> played 211,935 times
  - g3 -> played 23,925 times

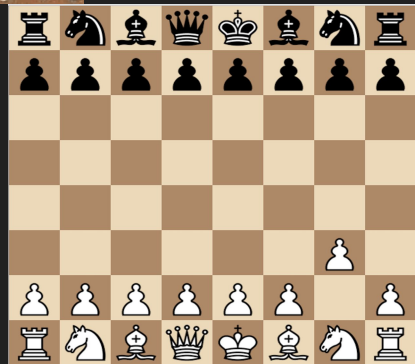
e4 →

d4 →

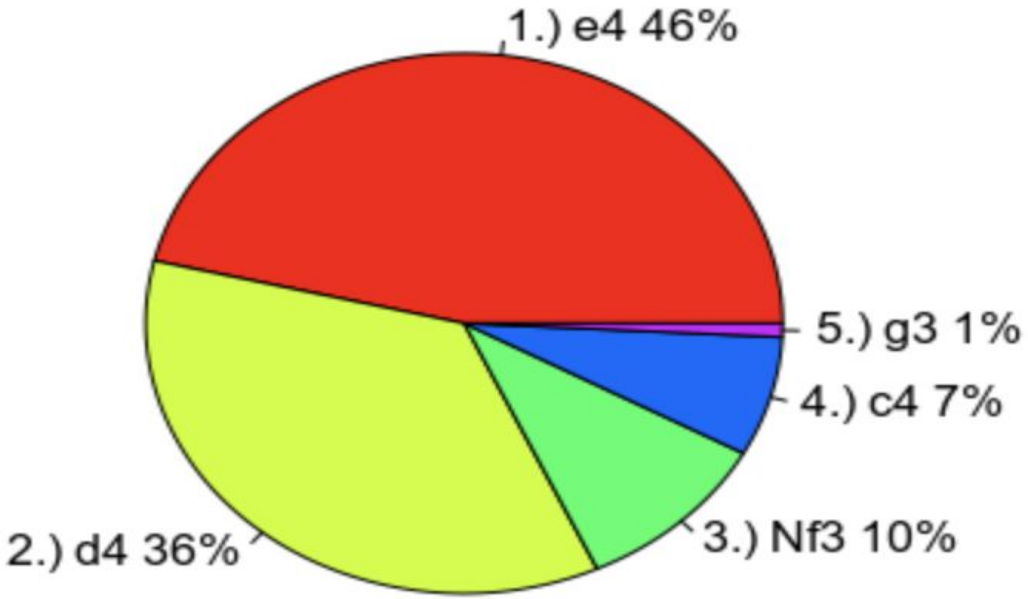
Nf3 →

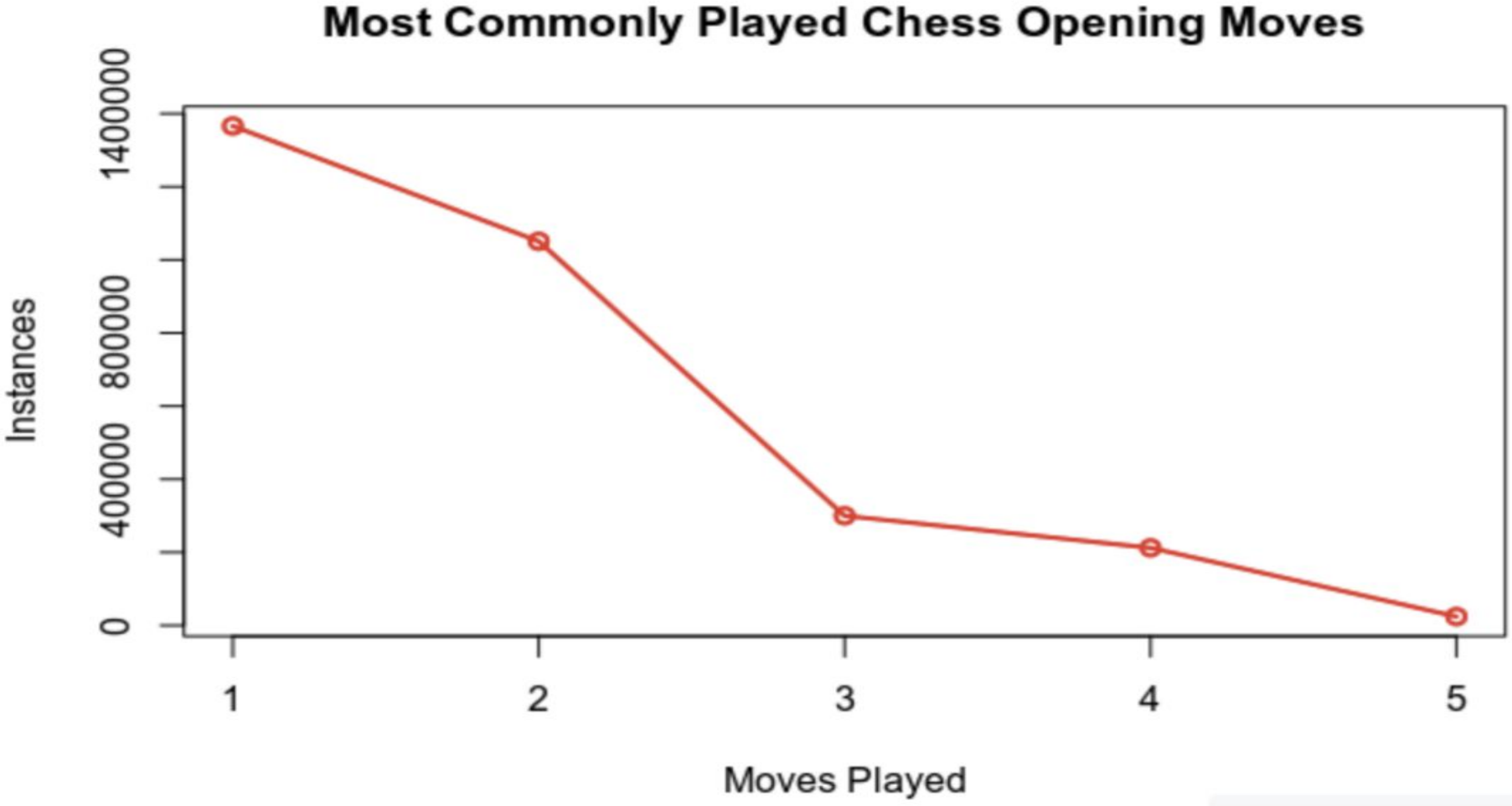
c4 →

g3 →



Most Commonly Played Chess Opening Moves




Most Commonly Played Chess Opening Moves
Instances
0
400000
800000
1400000
1
2
3
4
5
Moves Played

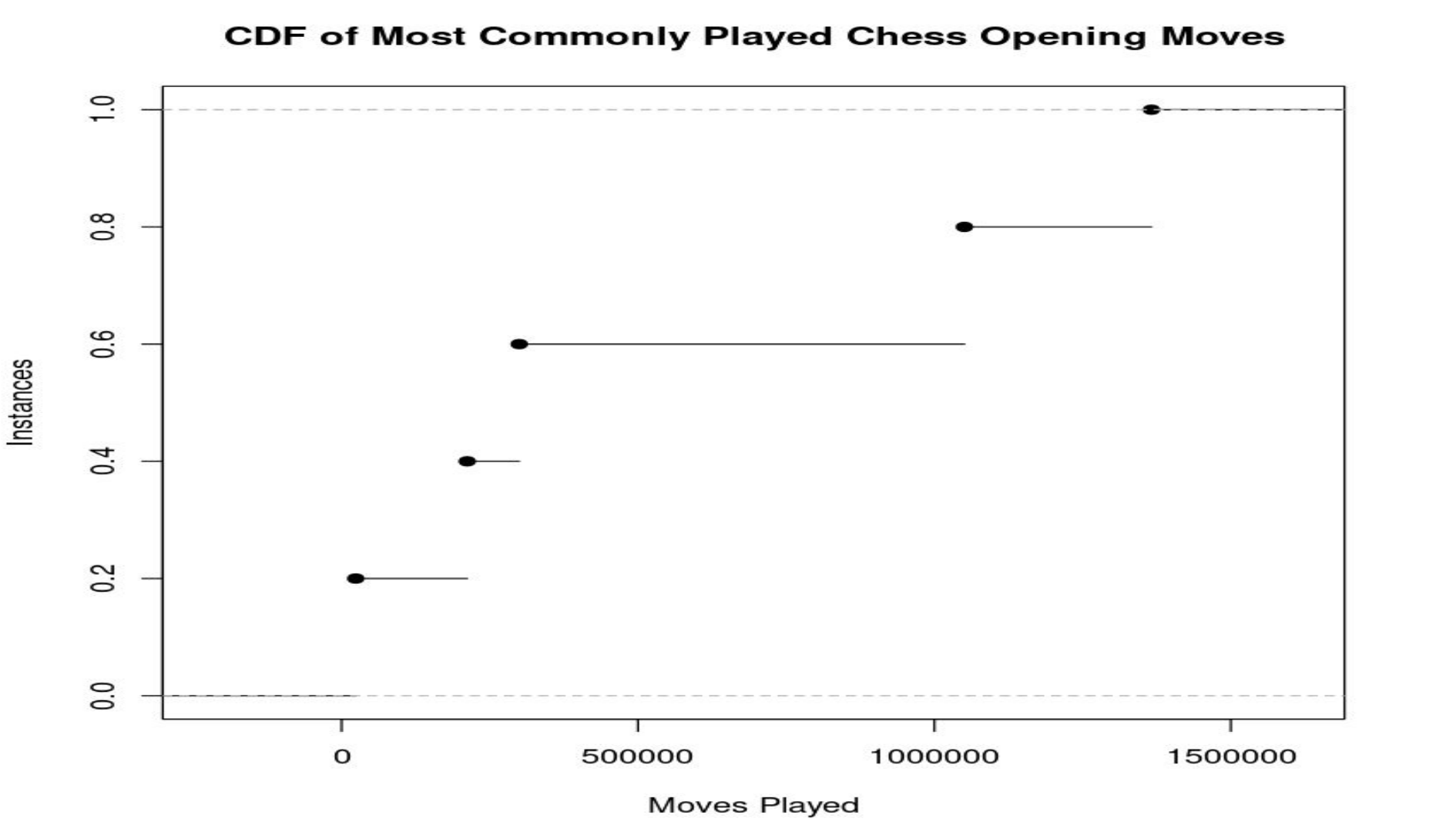
The Project

- ★ The cdf of data suggests:
  - ➢ the popularity of a certain move
  - ➢ the likelihood that the move will fall within the frequency distribution
- ★ Gives us an opportunity to look at the data
- ★ Most challenging aspect of the entire project

Result

executed in 0.822 sec(s)

```
[1] "1.) e4 46%"  "2.) d4 36%"  "3.) Nf3 10%" "4.) c4 7%"   "5.) g3 1%"
Empirical CDF
Call: ecdf(whitePcs)
 x[1:5] =  23925, 2.1194e+05, 2.9974e+05, 1.0507e+06, 1.3662e+06
```


CDF of Most Commonly Played Chess Opening Moves
Instances
0.0
0.2
0.4
0.6
0.8
1.0
0
500000
1000000
1500000
Moves Played

Future Direction

- **Would like to expand on the number of moves**
- **Deeper study of the reasoning behind moves**
  - Are there two separate moves that have been played the same amount of times?
  - Is there a way to predict the move that will be played?
- **To learn even more and strengthen my knowledge both in:**
  - R Programming Language
  - Statistical identities and skills

Bibliography

"Chess Opening Explorer & Database." *Chess.com*, https://www.chess.com/explorer.

"Data Manipulation with Dplyr." *DataCamp*, https://app.datacamp.com/learn/courses/data-manipulation-with-dplyr.

"Data Visualization in R." *DataCamp*, https://app.datacamp.com/learn/courses/data-visualization-in-r.

"Intermediate R." *DataCamp*, https://app.datacamp.com/learn/courses/intermediate-r.

"Introduction to Data Visualization with ggplot2." *DataCamp*, https://app.datacamp.com/learn/courses/introduction-to-data-visualization-with-ggplot2.

"Introduction to R." *DataCamp*, https://app.datacamp.com/learn/courses/free-introduction-to-r.

"Introduction to Statistics in R." *DataCamp*, https://app.datacamp.com/learn/courses/introduction-to-statistics-in-r.

"Introduction to the Tidyverse." *DataCamp*, https://app.datacamp.com/learn/courses/introduction-to-the-tidyverse.

Appendix

```
whitePcs <- c(1366195, 1050651, 299740, 211935, 23925)

label <- c("1.) e4", "2.) d4", "3.) Nf3", "4.) c4", "5.) g3")

percentage <- round(whitePcs / sum(whitePcs) * 100)

label <- paste(label, percentage)

label <- paste(label,"%",sep = "")

pie(whitePcs, label = label, col = rainbow(length(label)),

   main = "Most Commonly Played Chess Opening Moves")

plot(whitePcs, xlab = "Moves Played", ylab = "Instances", main = "Most Commonly Played Chess Opening
        Moves", col = "red", type = "o", lwd = 2)

label

CDF <- ecdf(whitePcs)

plot(CDF, xlab = "Moves Played", ylab = "Instances", main = "CDF of Most Commonly Played Chess
        Opening Moves")
```