

Script

I think I successfully answered my initial question with my analysis of this dataset.

Fractional score and opponent Lichess rating have a correlation coefficient of -0.95. This means that there is a really strong, negative association between fractional score and opponent rating. This finding makes a lot of sense since a person would win a lot less against stronger opponents and win a lot more against weaker opponents. Since this association is so strong, it's hard to make conclusions based on just the fractional scores of other variables. For example, Magnus had a 0.731 fractional score in 2018 and a fractional score of 0.707 in 2020. Does this mean that Magnus played worse in 2020 or he fought stronger opponents in 2020? It's hard to tell with just a bar chart that graphs the fractional scores of variables like year, color, first move, and time control. To deal with this issue, I tried to use a metric called [linear performance rating](#). It is a simple way of getting someone's rating using the average rating of their opponents and their fractional score against those opponents. These two factors vary a lot and ideally, the linear performance rating will reduce that variance. The equation is $(\text{average rating of opponents}) + 800 * (\text{fractional score}) - 400$. Unfortunately, there are issues with this equation that I did not realize until after I graphed the linear performance rating of each variable. The linear performance rating equation makes the assumption that the fractional score will approximately change from 1 to 0 when the average opponent rating increases by 800. This is a fair assumption, but it only makes sense when someone fights against both stronger and weaker players in that rating range of 800. For example, let's say I have a fractional score of 0.39 against 1800 rated players and a fractional score of 0.87 against 1400 rated players. This means my linear performance rating ranges from 1696 to 1712. This range is a lot smaller compared to my opponent rating range of 400. In this case, I could successfully use linear performance rating as a metric to judge other variables like year and color without average opponent rating determining

Script

anything. However, a person that only fights against stronger opponents will almost never have a fractional score above 0.5. A person that only fights against weaker opponents will almost never have a fractional score below 0.5. This person is Magnus Carlsen since he is the best player in the world. Since Magnus's fractional score ranges from 0.64 to 0.84 depending on the variables I have analyzed, the linear performance rating ends up being heavily positively correlated with average opponent rating. To prove this, I created a scatter plot to show the correlation between linear performance rating and average opponent rating. The correlation coefficient between those two ratings is 0.997. This ultimately means that there is no point in using linear performance rating as a way to isolate average opponent rating from other variables. With this knowledge, I ended up creating a scatter plot for fractional score vs average opponent rating to use as my main analysis to draw conclusions from. I plotted 22 points based on variables such as year (2018, 2019, 2020, 2021), color (Black, White), opponent rating range (<2700, 2700s, 2800s, 2900s, 3000s, 3100s), time control (1 minute+0 second games or not), first move as black (d5, Nf6, c5, g6), first move as white (e4, d4, c4, Nf3) and another point based on the dataset in its entirety. Then I created a residual plot to clearly see which variables led to a higher than expected fractional score. According to this plot, Magnus plays better as white than as black. This makes a lot of sense since white moves first and in general, white wins more games than black. Magnus's best first move as white is d4 and Magnus's best first move as black is g6. There is a huge 5 point difference between g6 and the 2nd best move as black, c5. I am not sure why this is the case, but this significant difference also exists in [Lichess's game database](#). For example, when white plays e4, black wins 26% of the time with c5, and 30% of the time with g6. In addition, the plot shows that Magnus's best year was 2019 and his worst year was 2018. This could be because Magnus started playing on Lichess from December 2017. In 2019, maybe he had a skill

Script

gap against most other players, and they slowly caught up in 2020 and 2021. I cannot be 100% sure. To add on, Magnus plays a lot worse when he's not playing 1 minute+0 second time control. The 'not 1+0' time control has the lowest residual at -0.046. Magnus plays better than expected against opponents rated less than 2700 and 2900-2999 rated opponents. It makes sense that Magnus plays better than expected against opponents rated less than 2700 because at some point, the relationship between fractional score and opponent rating stops being linear. For example, Magnus would probably have at least a 0.99 fractional score against opponents rated between 0 and 2000. I don't know how to explain why he plays better than expected against specifically 2900-2999 rated opponents. He plays worse than expected against opponents rated more than 3100. This makes sense for a similar reason to why he plays better than expected against really low rated opponents. The correlation coefficient for the fractional scores and average opponent ratings from each variable value is -0.703. R^2 is 0.494. This means that about half of the variance of fractional score shown in the scatter plot can be explained by average opponent rating. That means that the other half of the variance is due to other factors including time control, color, first move, and year. In addition, Magnus's mood, time of day, specific opponents can also influence this model.