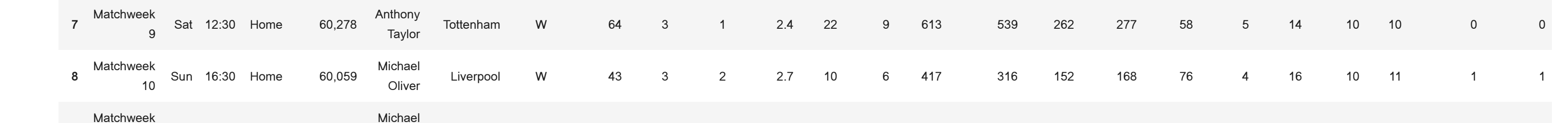


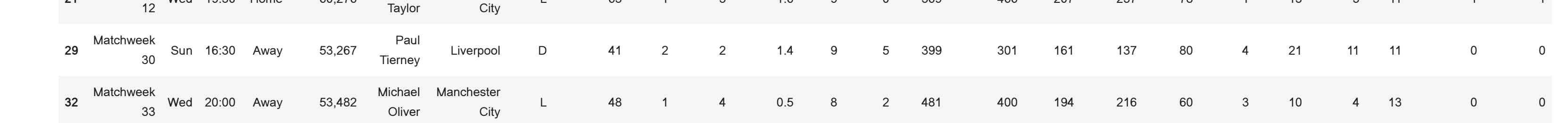
The dataframe contains all the matchweeks for Arsenal but we only need the data for the top 6 teams. We will first look at how the top six teams are registered in the dataframe (their name can be found in the 'Opponent' column). Then, we create a list and pass it into a filter where it displays only the data containing the names of those teams inside the previously mentioned list.

Finally, we create a new abridged dataframe containing only the data versus the top teams in the Premier League.



Now that we have made a filter. Let's create a new dataframe containing only the data we need

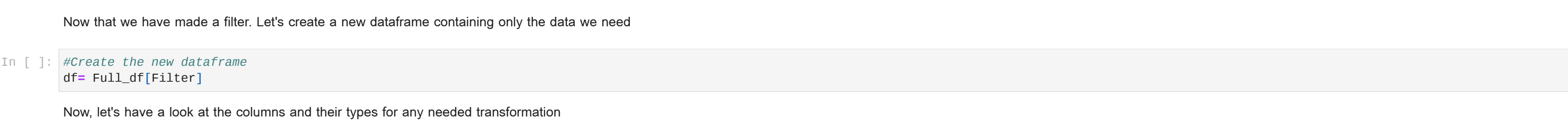
Now, let's have a look at the columns and their types for any needed transformation



From the following output, we can classify the changes that need to be made as the following:

- The columns Day and Result can be transformed from object to category to improve and optimize the performance.
- The column Attendance is more appropriate being an integer rather than an object to perform numerical aggregations.

Further changes can be made to make the performance better. For example, the integer columns can be changed from 64 bit to 16 or 8 since they don't contain a large range of numbers but since this is already a small dataset, it's not necessary, yet it's worth noting that it's a beneficial habit to embed in your workflow



Now that the dataframe is ready for analysis. We are going to perform some comparison between the performance vs the top 6. The following insights that I'm aiming to extract are:

- The number of goals Arsenal scored vs the top six
- The number of goals Arsenal conceded vs the top six
- The number of wins, draws and losses vs the top six
- The total amount of points gained vs the top six teams
- The total number of attendances
- The average possession percentage vs the top six
- The average passes made vs the top six
- The average expected goals vs the top six
- The match result based on the opponents
- The match result based on the venue
- The percentage of goals for against top 6 from the total goals scored
- The percentage of goals against against top 6 from the total goals conceded
- The average possession vs the top six compared to the overall average possession
- The average of goals scored and conceded against top six compared to the total average
- The shot accuracy percentage
- The passing accuracy percentage
- The percentage of attendances for the top 6 teams compared to the total number of attendances

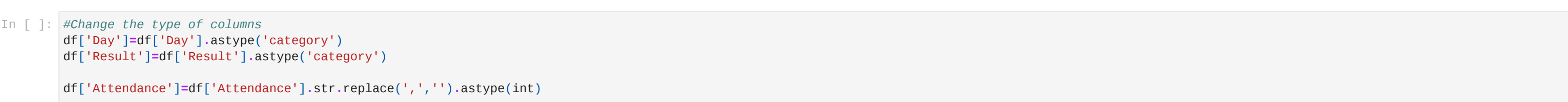
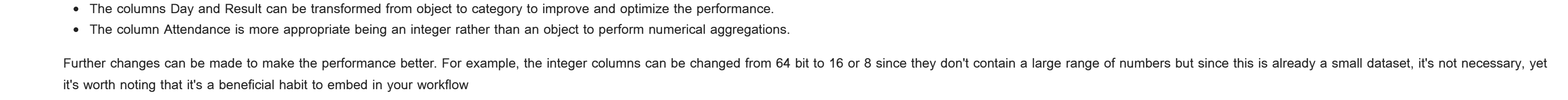
I already have data available for the overall stats from a previous project I have made so all that is left is comparing the data and see any useful insights.

First off, let's create a filter for home matches and away matches since it's going to be used on a number of insights we listed



A portion of the questions we listed above fall in the bin of either sum (total) or mean (average). Let's have a look at the following output:

- Arsenal managed to score 20 goals versus the top six teams and conceded 18, resulting in a goal difference of a positive 2.
- Arsenal culminated a total attendance record of 583k spectators in the 10 matches against the top teams.
- Arsenal had an average possession of 53.5, meaning that they had a small edge in ball controlling terms.
- On average, Arsenal had an xG performance of 1.8, suggesting that the team had a solid offensive effort against the top 6 teams from various variables like attacking tactics, player positioning, and overall gameplay were effective in generating high quality scoring chances.



Based on the following output we can conclude to:

- Arsenal managed to do the double on Chelsea and Tottenham, winning both in the home and away matches.
- Manchester City did the double on Arsenal, winning against the Gunners in both scenarios which was a vital point to win the league.
- Arsenal managed to win against Liverpool in the home fixture and also scrapped out a 2-2 draw in the away fixture.
- Arsenal and Manchester United both got 3 points from each other with Arsenal losing the away fixture 3-1 then making a comeback in the Emirates to win 3-2

