Jennifer Fajardo

Post-Pandemic Play: Evaluating Home Advantage in European Soccer After COVID-19

Checkpoint 7

Brief Overview

This analysis examines whether home field advantage within the Premier League declined during the COVID-19 pandemic and if it rebounded afterward. Home field advantage is evaluated using home points per match (Pts/MP) and its relationship with average home stadium attendance across pre COVID, COVID, and post COVID seasons.

Data Used

I used the same cleaned Premier League dataset from Checkpoint 4 **(Clean Data PL)**. However, to support this extended analysis, I reorganized the data and added two variables: **COVID Season** and **Stable Team.** A COVID season indicator flags the 2020-21 season. A stable team indicator identifies teams that played in all three seasons. This allows the analysis to account for promotion and relegation effects while still including all 20 teams that competed each season.

Squad	Season	Home Pts/MP	COVID Season	Average Home Attendance	Stable Team
Arsenal	2018-19	2.37	0	59,899	1
Arsenal	2020-21	1.47	1	632	1
Arsenal	2023-24	2.47	0	60,236	1
Aston Villa	2020-21	1.32	1	526	0
Aston Villa	2023-24	2.11	0	41,858	0
Bournemouth	2018-19	1.53	0	10,532	0
Bournemouth	2023-24	1.42	0	11,103	0
Brentford	2023-24	1.16	0	17,082	0
Brighton	2018-19	1.21	0	30,426	1
Brighton	2020-21	1.11	1	523	1

CP6 Baseline

In Checkpoint 6, I used a **group mean comparison across seasons** for the 14 consistent teams only. Results were interpreted relative to the pre COVID baseline to quantify the magnitude and direction of change during and after the pandemic. During the 2020-21 season (COVID), teams earned an average of **0.39** fewer home points per match. This decline was not uniform across teams as some clubs experienced little change or even improved home performance. Checkpoint 6 relied on season level averages and masks team level variation, limiting its ability to explain why some teams were more affected than others.

Upgrades in CP7

1. First, I added **average home stadium attendance** as a key feature to test whether crowd presence helps explain team level differences in home field advantage.
2. Second, I expanded the sample to **all 20** teams for each season to assess whether the patterns observed among stable teams also hold across the full league.

Why? My exploratory data analysis in Checkpoint 5 showed a clear positive relationship between attendance and home performance and substantial variation in home advantage across teams.

CP7 Analysis Spec

Component	Specification
Outcome	Home Points Per Match (Pts/MP)
Inputs	Average Home Attendance, COVID Season Indicator, Stable Team Indicator
Sample	All 20 Premier League teams
Row Definition	One team observed in a given season
Rule	Estimate the relationship between attendance and home points per match while controlling for season effects
Expected Direction	Higher attendance associated with higher home performance

CP6 vs CP7 Analysis Spec

	Question	Method	Sample	Outcome
CP6	What happened on average?	Home Pts/MP group mean comparison	14 consistent teams	Home Pts/MP fell by 0.39 during COVID
CP7	Why did it happen & for which teams?	Regression with average attendance and controls	All 20 teams	Reduced attendance explains much of the COVID season decline

Results & Comparison

To extend the Checkpoint 6 baseline, I ran a **linear regression** with home points per match as the outcome with COVID season, average home stadium attendance, and team stability as predictors. This model moves beyond descriptive averages and tests whether attendance helps explain differences in home field advantage across teams.

In Checkpoint 6, a group mean comparison showed that home points per match declined by 0.39 during the COVID season among teams that played all three seasons. In Checkpoint 7, once attendance is included in the regression, the COVID season effect is substantially reduced indicating that lower attendance explains much of the decline observed in the baseline comparison. The regression explains approximately **41%** of the variation in home Pts/MP across team seasons, suggesting a meaningful improvement in explanatory power relative to the CP6 approach.

As a basic check, the attendance relationship remains positive when restricting the sample to teams that competed in all three seasons. This indicates that the result is not driven solely by promotion or relegation effects.

Interpretation

- When attendance dropped during COVID, home field advantage weakened.
- Much of the COVID season decline in home performance is explained by reduced attendance rather than the season itself.

Limits

Average attendance during the COVID season is extremely low across teams. While expected, this reduces the ability to account for differences between matches or individual capacity restriction within stadiums.

Next Steps

(1) I will keep the current model with attendance as the main explanatory variable.
(2) Re-run the model using performance gap (home goal differential - away goal differential) as a check.
    - This should confirm that the attendance result is not driven by the points metric alone.
(3) Write a short, clear executive summary with my findings.