Premier League Multi-Level Modeling Project

Dylan Li, Liam Quach, Brendan Callender

Background

League Structure

- 20 Clubs/Teams
- 38 game season
  - Face each team twice - Both home and away

League Point System:

- Win - 3 points
- Draw - 1 point
- Loss - 0 points

Pos	Club	P	GD	Pts
1	Man City	37	60	88
2	Arsenal	37	61	86
3	Liverpool	37	43	79
4	Aston Villa	37	20	68
5	Spurs	37	10	63
6	Chelsea	37	13	60
7	Newcastle	37	21	57
8	Man Utd	37	-3	57
9	West Ham	37	-12	52
10	Brighton	37	-5	48
11	Bournemouth	37	-12	48
12	Crystal Palace	37	-6	46
13	Wolves	37	-13	46
14	Fulham	37	-8	44
15	Everton	37	-10	40
16	Brentford	37	-7	39
17	Nott'm Forest	37	-19	29
18	Luton	37	-31	26
19	Burnley	37	-36	24
20	Sheffield Utd	37	-66	16

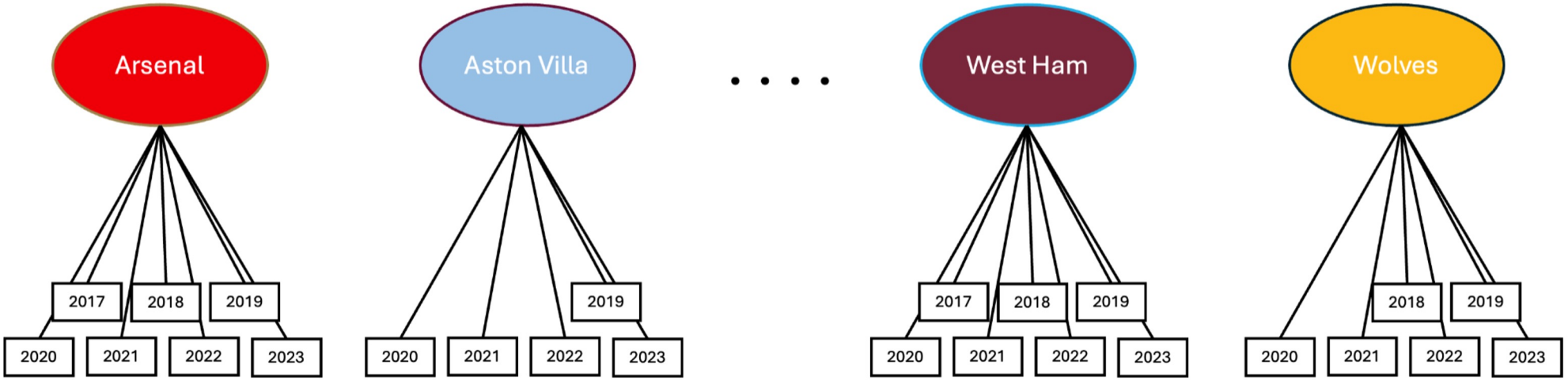
*Everton deducted six points following a breach of the Premier League's Profitability and Sustainability Rules
*Nottingham Forest deducted four points following a breach of the Premier League's Profitability and Sustainability Rules

Research questions

1. What **predictors** are **most associated** with **higher** or **lower** point totals in the Premier League?

1. Is **higher spending** in the transfer market associated with **higher** points totals in the Premier League?

Multi-level Structure of the Data



Data from 2017-2018 season up to 2023-2024 season

Project Data

Variable Description	Variable Type
Points	Response
Goals/90	L1 Predictor
Goals conceded/90	L1 Predictor
Expected goals/90	L1 Predictor
Expected goals conceded/90	L1 Predictor
Net Spend	L1 Predictor
Average Net Spend	L2 Predictor

1 row = A Singular Season for an Individual Club

Null Model

$$Points_{ij} = \beta_{00} + u_j + \epsilon_{ij} \quad \text{where } u \sim N(0, \tau_0^2) \text{ and } \epsilon \sim N(0, \sigma^2)$$

Points for **Club j** in **Season i**

Random Effect for each **Club**

Model Summary

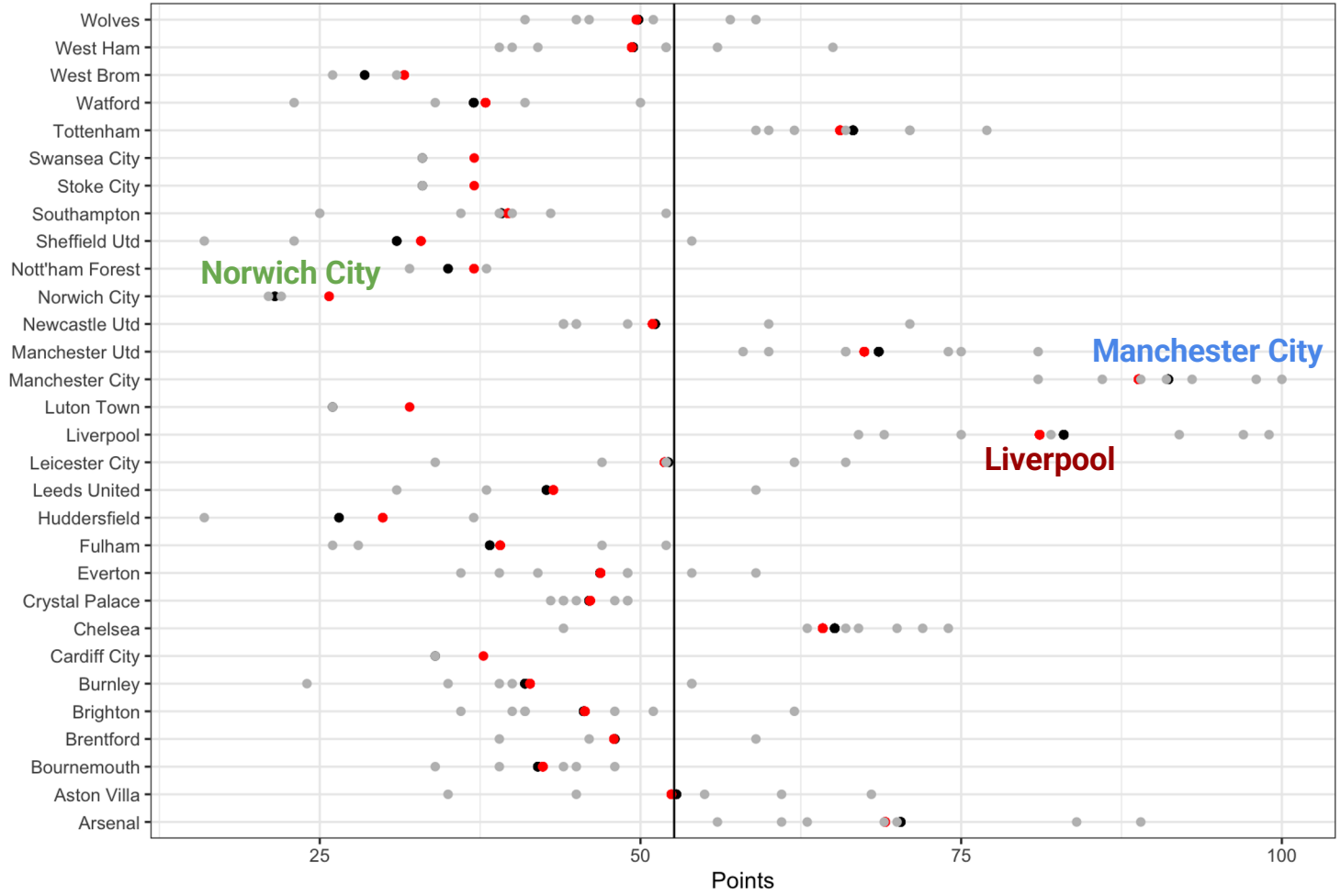
$\hat{\sigma}^2 = 255.3$

$\hat{\tau_0}^2 = 99.6$

$ICC = 0.72$ ⟶ Season point totals highly correlated within Clubs

95% CI for $\tau_0 = (11.99, 21.18)$ ⟶ Significant club-to-club variability

Random Effect Shrinkage



Final Model

$$Points_{ij} = \beta_{00} + u_j + \beta_1(G/90)_{ij} + \beta_2(GA/90)_{ij} + \beta_3(\overline{NetSpend})_j + \epsilon_{ij}$$

where $u \sim N(0, \tau_0^2)$ and $\epsilon \sim N(0, \sigma^2)$

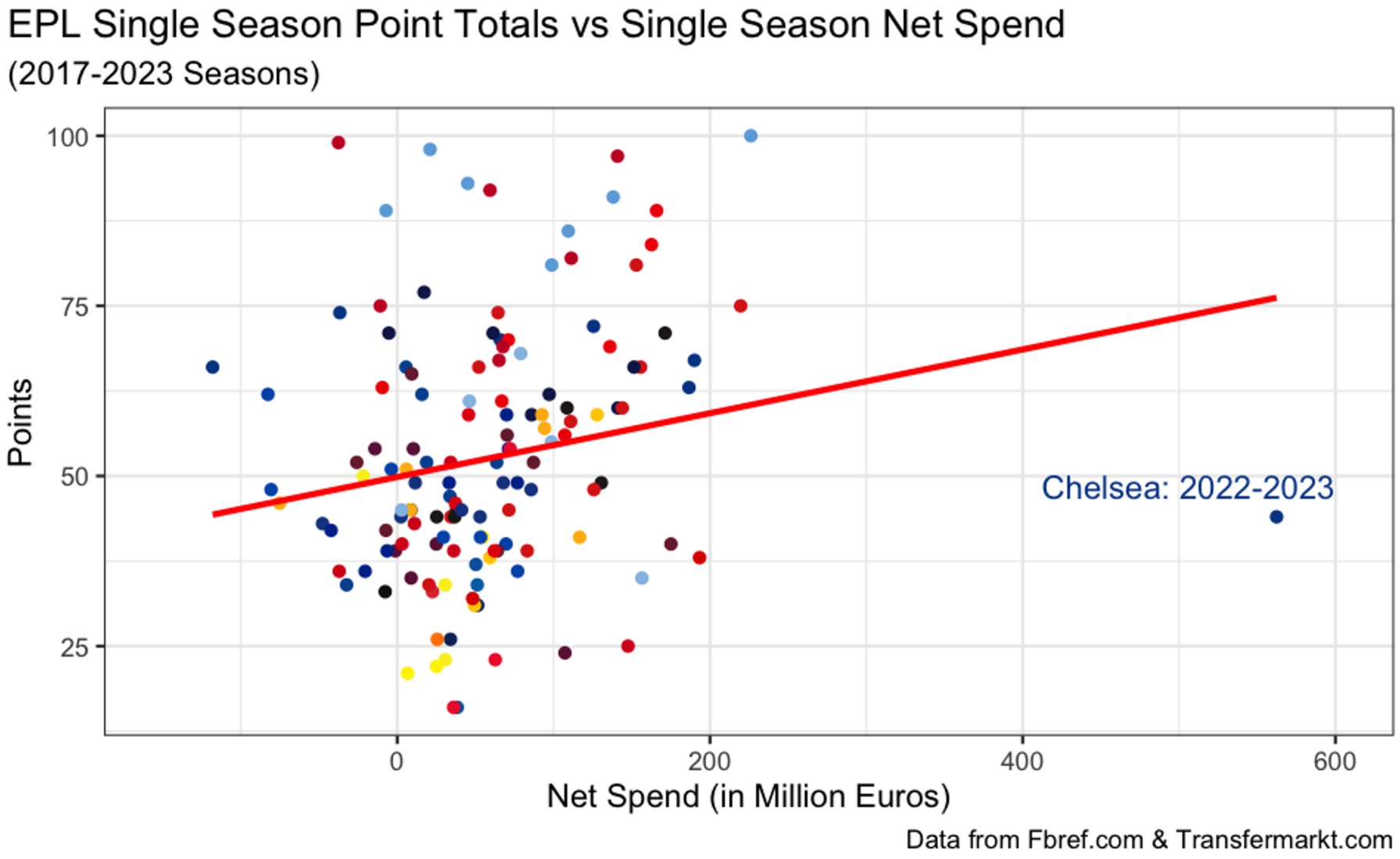
Final Model vs Null Model:

$$\hat{\sigma}^2 = 19.96$$

$$\hat{\tau_0}^2 = 0$$

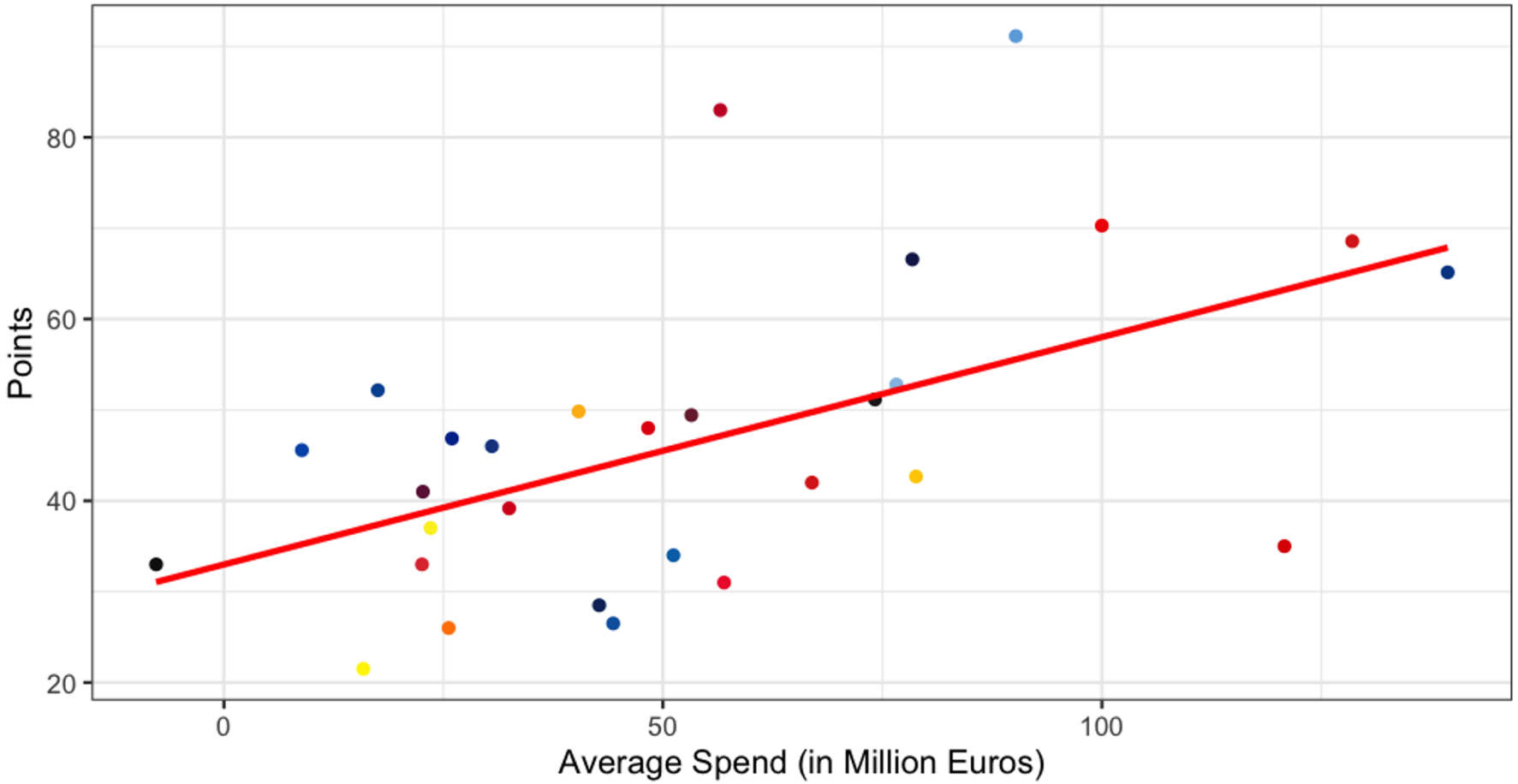
$$\text{Level 2 Variability Explained} = 100\%$$

$$\text{Level 1 Variability Explained} = 92.2\%$$

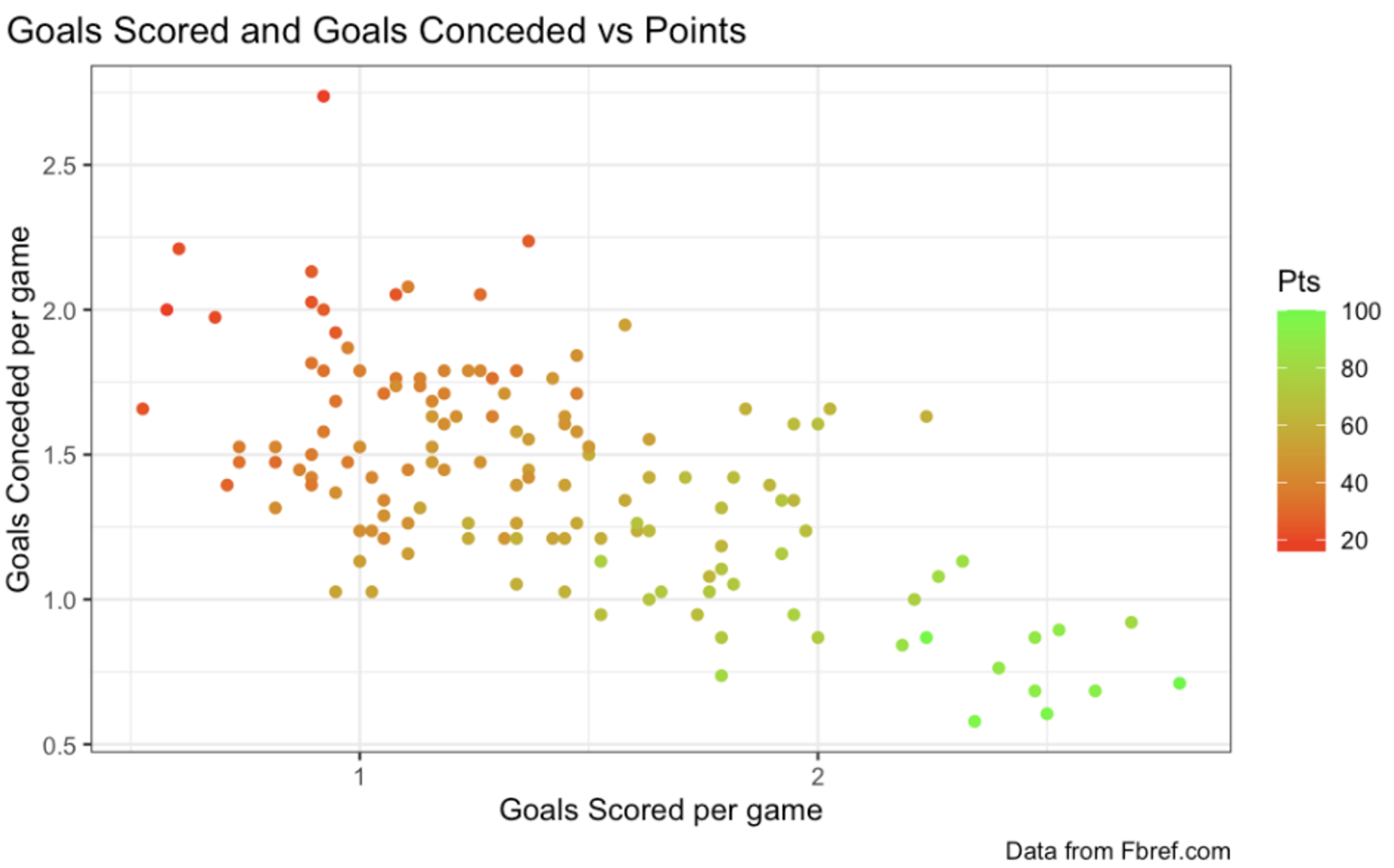

EPL Single Season Point Totals vs Single Season Net Spend
(2017-2023 Seasons)
100
75
50
25
Points
0
200
400
600
Net Spend (in Million Euros)
Chelsea: 2022-2023
Data from Fbref.com & Transfermarkt.com

Premier League Point Totals by Average Net Spend

(2017-2023 Seasons)



Data from Fbref.com & Transfermarkt.com


Goals Scored and Goals Conceded vs Points
Goals Conceded per game
2.5
2.0
1.5
1.0
0.5
1
2
Goals Scored per game
Pts
100
80
60
40
20
Data from Fbref.com

Model Coefficients

For every 1 additional goal a club scores per game, there is an associated **increase** in the predicted points of **23 points** after adjusting for club, goals conceded and net spend.

For every 1 additional goal a club concedes per game, there is an associated **decrease** in the predicted points of **21.8 points** after adjusting for…

For every €100,000 increase in a clubs average net spend, there is an associated **increase** in the predicted points of **3.12 points** after adjusting for…

Thank You!