Final Report

Dylan Li, Liam Quach, Brendan Callender

I. Introduction

The English Premier League (EPL) is the top tier of professional football (soccer) in England and is considered one of the most popular and competitive leagues in the world. The league is made up of twenty clubs (teams) that compete over a season for the Premier League title with new clubs added each year via a system of promotion and relegation. Each year, three new clubs are promoted from the second division based on the previous year's results with these promoted teams replacing the bottom three teams from the previous year's Premier League.

Over the course of a season, each team plays a total of 38 matches, facing every other team twice—once at home and once away. Teams are rewarded points from each game as follows: 3 points for a win, 1 points for a draw, and 0 points for a loss. The team with the most points at the end of the 38-game season is crowned as the Premier League Champions.

For our project, we are interested in exploring the following research questions:

1. **What factors are associated with higher or lower point totals in the English Premier league?**
2. **Is spending more money in the off-season associated with earning more points the following season?**
3. **How do differences in expected metrics compared tk actual metrics impact clubs point totals?**

II. Data Source & Methods

To answer our research questions, we collected English Premier League season-level data spanning from the 2017-2018 season up to the most recently completed 2023-2024 season. Data was collected from two sites: rbref.com and transfermarkt.com. The data collected from fbref includes performance related metrics for each team over the season as predictors as well as point totals for each team at the end of the season for our response variable. The performance metrics include total goals scored, total goals conceded, expected goals scored, expected goals

conceded, average % possession, shooting metrics and more. The data collected from transfermarkt includes data relating to each teams expenditure and sales when buying or selling players in the transfer market. This data includes total money spent, total money from sales, net spend, number of players bought, number of players sold, and more. Money related variables are measured in millions of euros.

Predictors relating to season totals such as goals scored and goals conceded were scaled down to per 90, or per game. This was done by dividing these metrics by the total games played which is 38. See *Table 1* below for descriptions of key variables in our dataset. See *Table 2* below for a sample of rows from the data.

Table 1: Description of Dataset Variables

Variable	**Role**	**Range of Values**
Points	Response	(16, 100)
Goals/90	L1	(0.52, 2.78)
Goals Against/90	L1	(0.58, 2.74)
Average Possession of Ball (%)	L1	(35.4, 71.0)
...	...	...
Net Spend (in €1,000,000)	L1	(-118.07, 562.39)
Club Average Net Spend (in €1,000,000)	L2	(-7.72, 139.39)
Actual vs Expected Goals/90 Difference	L1	(-0.37, 0.72)
Actual vs Expected Goals/90 Against Difference	L1	(-0.37, 0.72)

Table 2: Example Rows from Dataset

Club	Season	Pts	GF	GA	Poss	...	NetSpend	Mean_NetSpend
Chelsea	2017	70	1.63	1.0	55.6	...	65.9	139.0
Arsenal	2017	63	1.95	1.34	61.4	...	-9.55	100.0
Everton	2017	49	1.16	1.53	45.5	...	76.8	26.0

To analyze the data, we will employ multi-level regression models, also known as hierarchical linear models. This approach is well-suited for the structure of the dataset, in which we have repeat observations for different clubs over several seasons. This structure can be visualized as seen in *Figure 1* below.

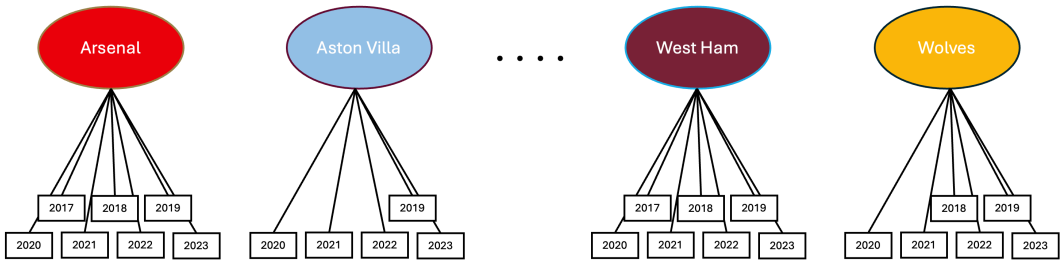


Figure 1: Multi-level Structure of Data

III. Results

Exploratory Data Analysis

This section presents the exploratory data analysis conducted to understand the key relationships between variables in the dataset. This exploratory data analysis was conducted before the model fitting process to gain an initial understanding of our research questions.

Figure 2 below shows the join distribution of goals scored per game and goals conceded per game, colored by season point totals. We there is a strong, negative correlation between goals scored per game and goals conceded per game. This means that teams who tend to score more, also tend to conceded less as well. When considering the season point totals, we see that decreasing the number of goals conceded per game is associated with higher point totals holding goals scored constant. Additionally, increasing the number of goals scored per game is assoicated with higher point totals holding goals conceded constant. Lastly, we see that jointly decreasing the number of goals conceded per game and increasing the number of goals scored per game is associated with the largest increase in season point totals.

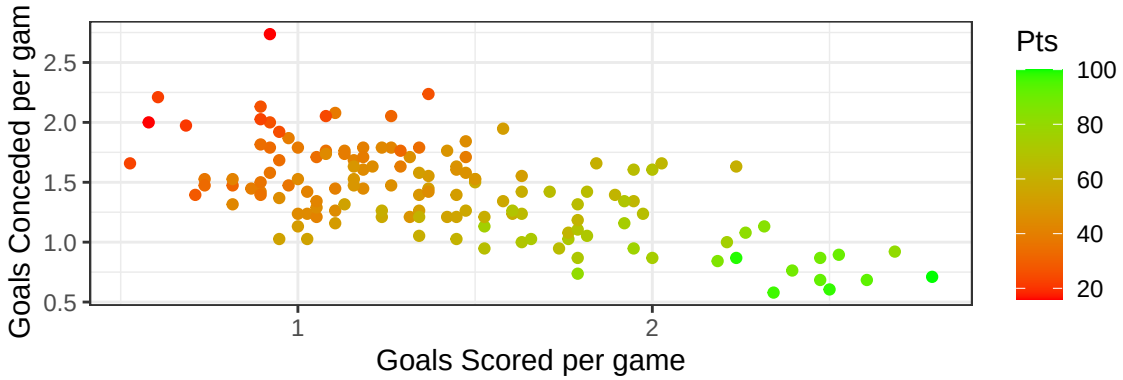


Figure 2: Impact of Goals Scored/90 and Goals Conceded/90 on Season Point Totals

Figure 3 below shows the relationship between average % possession of the ball and season point totals. The plot shows a strong, positive associated between % possession and points with higher values for % possession associated with higher point totals. This makes sense intuitively because teams with more possession tend to have the ball more which reduces the chances of the opposing team scoring and gives your team more chances to score goals.

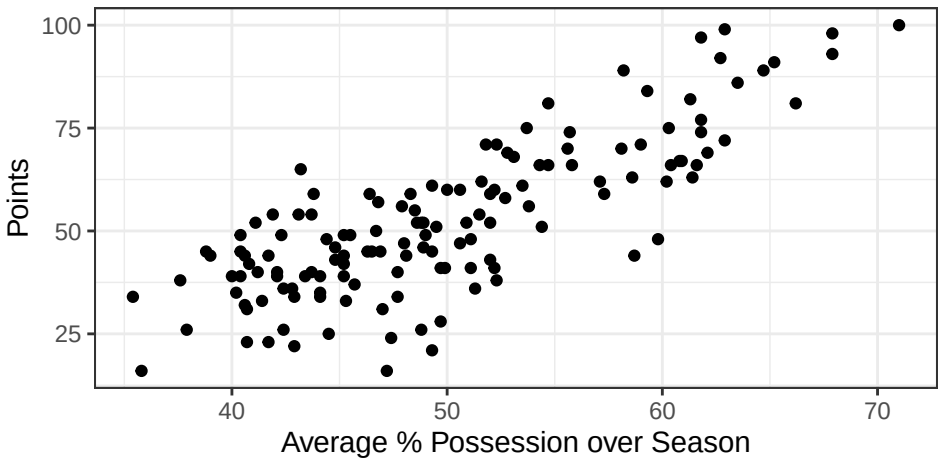


Figure 3: Season Point Totals by Season Average % Possession

Lastly, the plots below in *Figure 4* and *Figure 5* show the impact of spending on season point totals. *Figure 4* shows the relationship between season point totals and the net spend of the club for individual seasons. Larger values for net spend represent a club spending more money on new players while smaller values indicate a club spending less money with negative values indicating a team made profit selling players in the market. From the plot, we see a weak positive association with teams who spend more money being associated with higher point totals. We also notice a major outlier in the data with Chelsea in the 2023-2024 season. This is a valid data point and represents the season in which Chelsea had new owners invest large amounts of money into the team. This is not normal behavior for when teams get new owners and can serve as an example of how making too many changes to a team can have a negative impact on performance.

In *Figure 5*, we see a much stronger positive association between net spend and points when aggregated for each club. This demonstrates that consistent investment into a team over many seasons is more strongly associated with higher point totals than just a single season of large investment. (As they say… Rome wasn't built in a day)

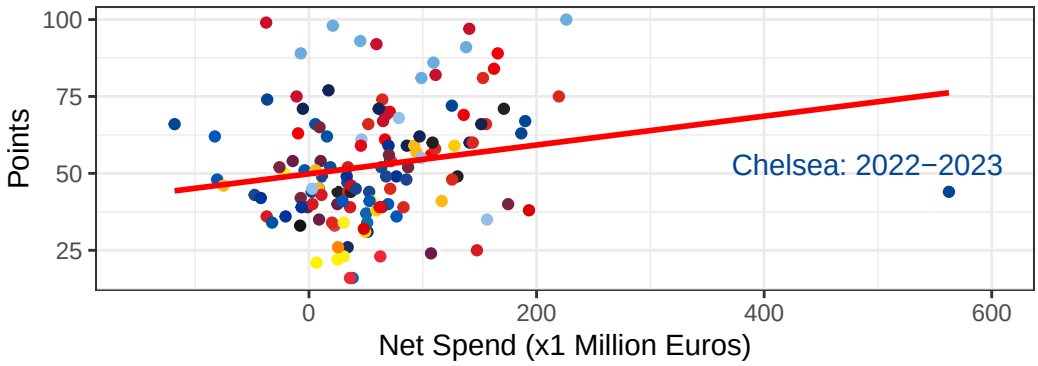


Figure 4: Season Point Totals by Single Season Net Spend

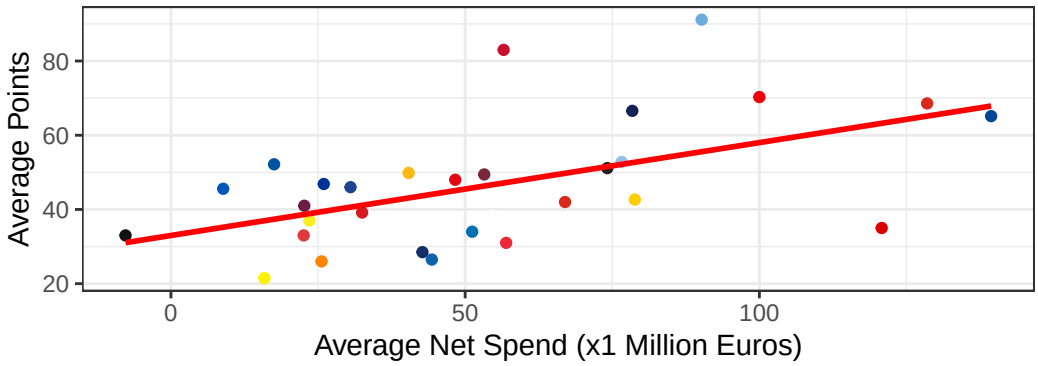


Figure 5: Club Average Points by Club Average Net Spend

ANOVA

After performing aour exploratory data analysis, we conducted an initial Analysis of Variance (ANOVA) test to explore whether there is significant club-to-club variability in season point totals. See results in *Table 3* below. Looking at the p-value resulting from the ANOVA, we have significant evidence that at least 2 clubs have different mean point totals. This is supported by *Figure 6* below which shows the distribution of point totals by each club. We see clubs like Manchester City have very high point totals while clubs like West Brom and Norwich City have very low point totals.

Table 3: *ANOVA for Significance of Club-to-Club Variability*

	df	**SSE**	**MSE**	**F-Statistic**	**P-Value**
Club	29	37233	1283.89	12.848	$<$ 0.0001
Residuals	110	10992	99.93		

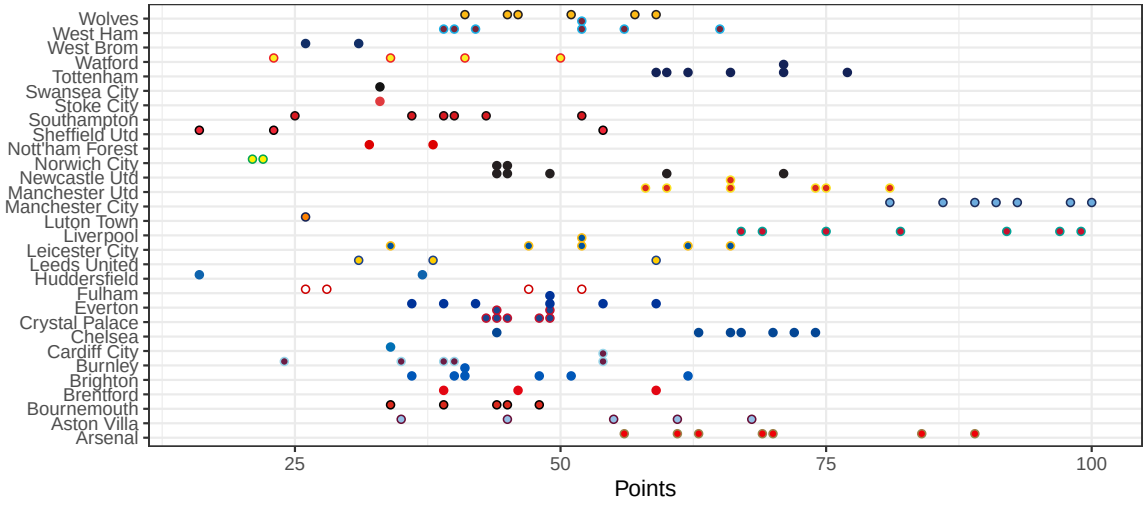


Figure 6: Distribution of Point Totals by Team

Null Model

After finding significant club-to-club variability in the season point totals, we fit an initial null model which includes no predictors and random effects for each club. The model can be written out as seen below:

$$Points_{ij} = \beta_{00} + u_j + \epsilon_{ij}$$

where $u_j \sim N(0, \tau_0^2)$ and $\epsilon_{ij} \sim N(0, \sigma^2)$

The summary of the null model output can be found below in *Table 4*. Looking at the ICC of the model, we see that approximately 72% of the variation of season point totals is at the club level while 28% of the variation is within each club. This matches what we saw in *Figure 6* above with clubs point totals being similar across the seasons and more different across clubs. Additionally, we see the model supports the fact that there is significant club-to-club variation

with the confidence interval for τ_0 not containing 0. Lastly, we see the resulting predictions from the null model in *Figure 7 below*. The black points represent the mean points for each club while the red points represent the predicted points for each club. The distance from the red and black points represent the shrinkage that occurs when using a multi-level model. We can see a club like Luton Town has large shrinkage towards the mean because there is only one season of data for Luton in the EPL.

Table 4: Null Model Summary

Parameter/Statistic	**Estimate**
σ^2	255.3
τ_0^2	99.6
95% CI for τ_0	(11.99, 21.18)
ICC	0.72

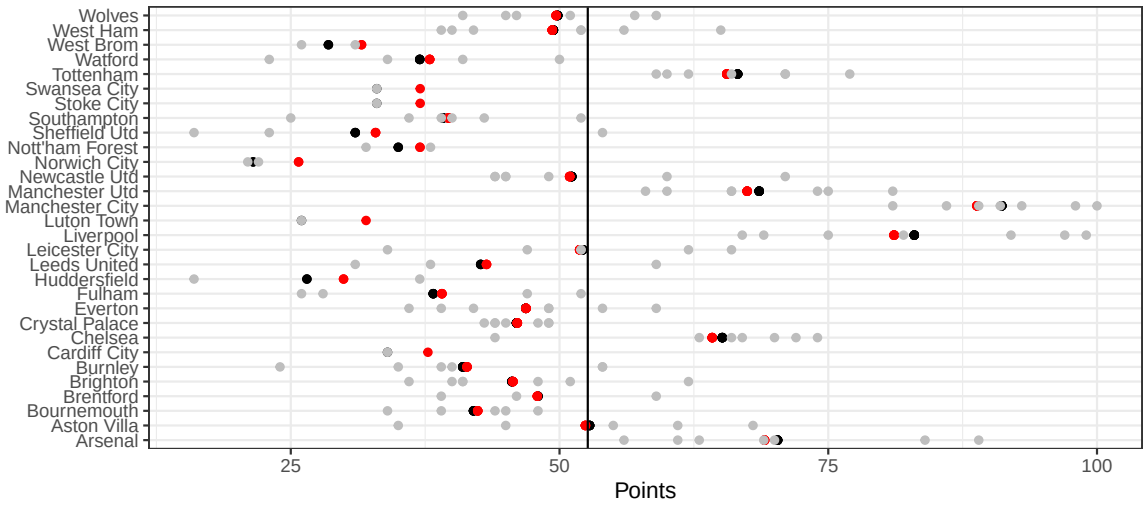


Figure 7: Null Model Predictions

Model Fitting Process

For our model fitting process, we first began by including goals scored per game and goals conceded per game due to our EDA showing a strong joint association with points. We found these variables to be extremely predictive of points so they were included in each subsequent model. From there we continued to explore new models by adding different level 1 (L1) predictors. If a predictor was significant, it was left in the model. We found no additional L1 predictors to be significant after adding goals scored and conceded. After exploring L1 predictors, we added our only level 2 (L2) predictor to the model which is average net spend

for each club. We found this to be significant and moved on to adding random slopes to the model. Adding random slopes for goals scored and goals conceded did not significantly improve the fit of the model so no random slopes were included in the final model. The complete model fitting process with output can be found below in the *Appendix*.

Final Model

After the model fitting process we finished with a final model that can be written as follows:

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

where $u_j \sim N(0, \tau_0^2)$ and $\epsilon_{ij} \sim N(0, \sigma^2)$

Note: All predictors included in model have been grand-mean-centered

A summary table of the output from the final model can be found below in *Table 5*. We see that the coefficients for the predictors included in the model match what we would expect after conducting our EDA. However, we do notice that there is only a small effect associated with clubs spending more money over many seasons after accounting for goals scored and goals conceded per game. Lastly, we notice that all the club-to-club variability in point totals can be explained by our model including goals scored, goals conceded, and average net spend. This matches our intuition due to how teams are awarded points: 3 for a win, 1 for a draw, and 0 for a loss. Simply put, more goals scored and less goals conceded means more wins which means more points.

Table 5: Final Model Summary

Parameter	**Estimate**	**Interpretation**
σ^2	19.96	92% of Level 1 variability explained when compared to null model
τ_0^2	0	100% of club-to-club variability in point totals explained by predictors
β_{00}	52.63	Predicted point total for club with average goals scored, goals conceded, and net spend
β_1	23.04	Associated increase in predicted points with each 1 increase in goals scored per game after adjusting for goals conceded, net spend, and club.

Parameter	Estimate	Interpretation
β_2	-21.87	Associated increase in predicted points with each 1 increase in goals conceded per game after adjusting for goals scored, net spend, and club.
$50\beta_3$	1.55	Associated increase in predicted points with each €50,000,000 increase in club average net spend after adjusting for goals scored, goals conceded, and club.

Model Diagnostics

Looking at the diagnostic plots for our final model, we see that the linearity, normality, and equal variance assumptions are not violated. Firstly, in *Figure 8* below, we can see a random scatter both above and below the horizontal line which indicates linearity is not violated. In the same plot, we also see that there is no obvious pattern of fanning in the data which means the equal variance assumption is not violated. Lastly, in *Figure 9* we see the points in the plot follow the diagonal line indicating the normality assumption is not violated. In these plots, we do notice potential outliers with both positive and negative residuals. however, these points are completely valid observations and have no reason to be removed from the data.

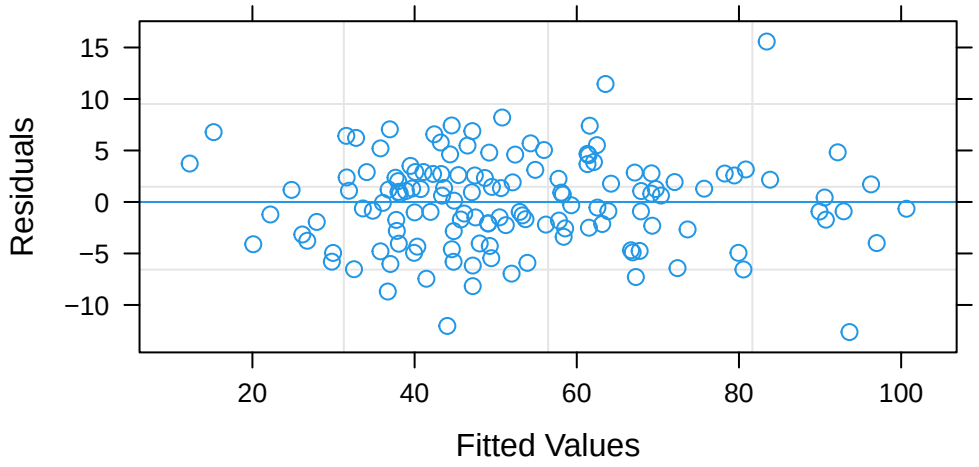


Figure 8: Residuals vs Fitted Plot

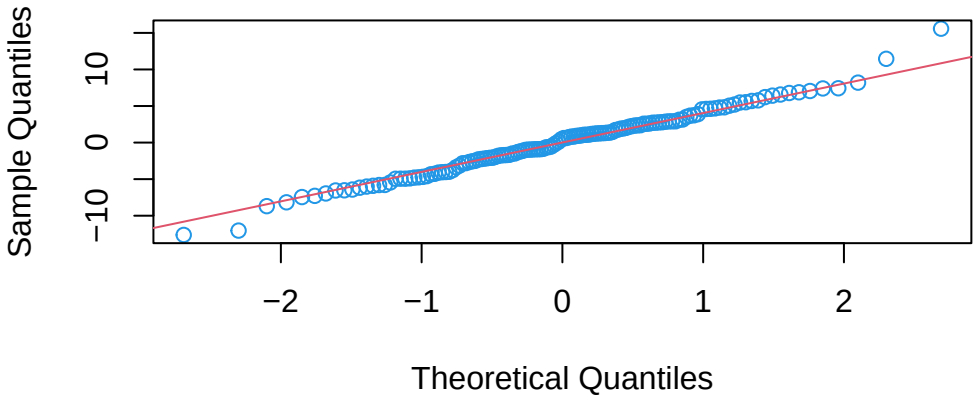


Figure 9: Normal-QQ Plot

V. Discussion

Research Questions

From our model fitting process, we were able to answer our proposed research questions.

1. **What factors are associated with higher or lower point totals in the English Premier league?**

We found that the most important factors for teams scoring higher EPL point totals were scoring more goals per game and goals conceding less goals per game. More specifically, we found scoring more goals had a larger impact than conceding less goals on point totals when looking at the magnitude of the coefficients. This could relate to the fact that conceding less goals without scoring more may lead to more ties while scoring more is a more sure way to win games which earns more points. This may support the hypothesis that offensive teams earn more points than defensive teams. We also found teams that invest larger amounts of money in their team also achieve higher point totals however this effect was much smaller than the effects of scoring more goals and conceding less.

2. **Is spending more money in the off-season associated with earning more points the following season?**

We found that consistent spending has a larger effect on clubs achieving higher point totals than single season spending. Our model found clubs who receive consistent investment over

many seasons are associated with higher point totals while teams who receive larger investment over a single season when accounting for goals scored and conceded. Perhaps if we were to remove the goals scored and goals conceded predictors we would see a larger associated effect.

3. **How do differences in expected metrics to actual metrics impact a clubs point totals?**

We found that differences in goals scored and expected goals scored as well as goals conceded and expected goals conceded do not significantly predict EPL club point totals. This makes sense because a clubs point total at the end of the season should reflect thier actual performance and not their expected performance. Thus, expected metrics can be used to predict how many points a club should have scored if a club performed according to their expected metrics.. This can be useful in accessing if a team is over-performing or under-performing according to their expected metrics.

Implications

Our model reveals clubs should focus their energy towards setting focus towards scoring more goals per game while limiting the number of goals they concede per game. This can inform where clubs should focus their spending, how they should set up tactically, and more to maximize their point totals in the EPL. An example of this is Arsenal this season who are maximizing their goals scored per game by using corner kick routines which can be practiced to score additional non-open-play goals each game.

Limitations and Next Steps

An interesting future step we can consider with this data is to exclude goals scored and goals conceded from our model entirely, and examine if the other variables are significant predictors of season point totals. The benefit to this analysis, if the predictions are accurate, will be the ability to more specifically inform how clubs can achieve higher point totals. As our current model stands, it can be used to predict season results after estimating each teams goals scored per game and goals conceded per game if we simplify the model. However, if we were to only use variables that we can acquire before a season starts, our model will struggle to predict results for a season before it begins.

VI. Appendix

ANOVA

```r
model00 <- lm(Pts ~ Club, data = prem)
broom::tidy(anova(model00))
```

```
# A tibble: 2 x 6
  term         df  sumsq meansq statistic   p.value
  <chr>     <int>  <dbl>  <dbl>     <dbl>     <dbl>
1 Club         29 37233. 1284.       12.8  1.10e-23
2 Residuals   110 10992.   99.9      NA    NA
```

Null Model

```r
model0 <- lmer(Pts ~ 1 + (1 | Club), data = prem)
summary(model0)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Pts ~ 1 + (1 | Club)
   Data: prem

REML criterion at convergence: 1108.9

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.02395 -0.60329 -0.08726  0.65915  2.11491

Random effects:
 Groups   Name        Variance Std.Dev.
 Club     (Intercept) 255.32   15.979
 Residual              99.66    9.983
Number of obs: 140, groups:  Club, 30

Fixed effects:
            Estimate Std. Error t value
(Intercept)   47.388      3.091   15.33
```

```r
confint(model0)
```

```
Computing profile confidence intervals ...
```

```
                2.5 %   97.5 %
.sig01      11.987270 21.17922
.sigma       8.798443 11.45651
(Intercept) 41.151433 53.48127
```

Model Fitting Process

```r
model1 <- lmer(Pts ~ GF + GA + (1 | Club), data = prem)
summary(model1)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Pts ~ GF + GA + (1 | Club)
   Data: prem

REML criterion at convergence: 815.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.7861 -0.6433  0.0348  0.6699  3.2020

Random effects:
 Groups   Name        Variance Std.Dev.
 Club     (Intercept)  0.9551  0.9773
 Residual             19.9221  4.4634
Number of obs: 140, groups:  Club, 30

Fixed effects:
            Estimate Std. Error t value
(Intercept)   49.474      3.024   16.36
GF            24.099      1.042   23.14
GA           -21.903      1.325  -16.54

Correlation of Fixed Effects:
   (Intr) GF
GF -0.848
GA -0.909  0.582
```

Add Level 1 Predictors

```r
model2_1 <- lmer(Pts ~ GF + GA + xG_diff + xGA_diff + (1 | Club), data = prem)
summary(model2_1)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Pts ~ GF + GA + xG_diff + xGA_diff + (1 | Club)
   Data: prem

REML criterion at convergence: 803.9

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.64241 -0.67011  0.06865  0.63657  3.05612

Random effects:
 Groups   Name        Variance Std.Dev.
 Club     (Intercept)  1.119   1.058
 Residual             19.622   4.430
Number of obs: 140, groups:  Club, 30

Fixed effects:
            Estimate Std. Error t value
(Intercept)   49.586      3.825  12.965
GF            23.015      1.484  15.510
GA           -20.942      1.713 -12.229
xG_diff        4.072      3.011   1.352
xGA_diff      -2.996      2.847  -1.052

Correlation of Fixed Effects:
         (Intr) GF     GA     xG_dff
GF       -0.835
GA       -0.887  0.510
xG_diff   0.420 -0.691 -0.134
xGA_diff  0.483 -0.227 -0.632  0.082
```

```r
model2_2 <- lmer(Pts ~ GF + GA + NetSpend + (1 | Club), data = prem)
summary(model2_2)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Pts ~ GF + GA + NetSpend + (1 | Club)
   Data: prem
```

```
REML criterion at convergence: 823.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.8045 -0.6651  0.0632  0.6753  3.2813

Random effects:
 Groups   Name        Variance Std.Dev.
 Club     (Intercept)  0.7151  0.8456
 Residual             20.1985  4.4943
Number of obs: 140, groups:  Club, 30

Fixed effects:
             Estimate Std. Error t value
(Intercept)  49.312092   3.046415  16.187
GF           24.041964   1.038091  23.160
GA          -21.871705   1.330367 -16.440
NetSpend      0.003499   0.005168   0.677

Correlation of Fixed Effects:
         (Intr) GF     GA
GF       -0.837
GA       -0.911  0.583
NetSpend -0.114 -0.068  0.074
```

```
model2_3 <- lmer(Pts ~ GF + GA + Poss + (1 | Club), data = prem)
summary(model2_3)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Pts ~ GF + GA + Poss + (1 | Club)
   Data: prem

REML criterion at convergence: 817.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.7761 -0.6278  0.0249  0.6584  3.1837

Random effects:
 Groups   Name        Variance Std.Dev.
 Club     (Intercept)  1.006   1.003
 Residual             20.024   4.475
```

```
Number of obs: 140, groups:  Club, 30

Fixed effects:
             Estimate Std. Error t value
(Intercept)  48.39335    5.02915   9.623
GF           23.85014    1.40678  16.954
GA          -21.76906    1.41327 -15.403
Poss          0.02484    0.09327   0.266

Correlation of Fixed Effects:
     (Intr) GF     GA
GF    0.152
GA   -0.786  0.179
Poss -0.797 -0.668  0.340
```

```
model2_4 <- lmer(Pts ~ GF + GA + Age + (1 | Club), data = prem)
summary(model2_4)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Pts ~ GF + GA + Age + (1 | Club)
   Data: prem

REML criterion at convergence: 814.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.8025 -0.6074  0.0163  0.6629  3.1979

Random effects:
 Groups   Name        Variance Std.Dev.
 Club     (Intercept)  1.331   1.154
 Residual             19.710   4.440
Number of obs: 140, groups:  Club, 30

Fixed effects:
            Estimate Std. Error t value
(Intercept)  40.3656    11.8084   3.418
GF           24.2571     1.0710  22.650
GA          -21.7253     1.3373 -16.246
Age           0.3235     0.4098   0.789

Correlation of Fixed Effects:
```

```
    (Intr) GF     GA
GF  -0.381
GA  -0.348  0.576
Age -0.966  0.173  0.121
```

Add Level 2 Predictors

```
model2_5 <- lmer(Pts ~ GF + GA + Mean_NetSpend + (1 | Club), data = prem)
```

```
boundary (singular) fit: see help('isSingular')
```

```
summary(model2_5)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Pts ~ GF + GA + Mean_NetSpend + (1 | Club)
   Data: prem

REML criterion at convergence: 815.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.8271 -0.6048  0.1171  0.6113  3.4867

Random effects:
 Groups   Name        Variance Std.Dev.
 Club     (Intercept)  0.00    0.000
 Residual             19.96    4.467
Number of obs: 140, groups:  Club, 30

Fixed effects:
               Estimate Std. Error t value
(Intercept)    49.11412    2.95766  16.606
GF             23.03892    1.05930  21.749
GA            -21.86604    1.29747 -16.853
Mean_NetSpend   0.03120    0.01181   2.643

Correlation of Fixed Effects:
            (Intr) GF     GA
GF          -0.769
GA          -0.919  0.557
Mean_NtSpnd -0.087 -0.362  0.056
```

```
optimizer (nloptwrap) convergence code: 0 (OK)
boundary (singular) fit: see help('isSingular')
```

```
model2_6 <- lmer(Pts ~ GF + GA + NetSpend + Mean_NetSpend + (1 | Club), data = prem)
```

```
boundary (singular) fit: see help('isSingular')
```

```
anova(model2_5, model2_6)
```

```
refitting model(s) with ML (instead of REML)
```

```
Data: prem
Models:
model2_5: Pts ~ GF + GA + Mean_NetSpend + (1 | Club)
model2_6: Pts ~ GF + GA + NetSpend + Mean_NetSpend + (1 | Club)
         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
model2_5    6 824.35 842.00 -406.17   812.35
model2_6    7 826.20 846.79 -406.10   812.20 0.1519  1     0.6968
```

Random Slopes

```
model3_1 <- lmer(Pts ~ GF + GA + Mean_NetSpend + (1 + GF | Club), data = prem)
```

```
boundary (singular) fit: see help('isSingular')
```

```
anova(model2_5, model3_1)
```

```
refitting model(s) with ML (instead of REML)
```

```
Data: prem
Models:
model2_5: Pts ~ GF + GA + Mean_NetSpend + (1 | Club)
model3_1: Pts ~ GF + GA + Mean_NetSpend + (1 + GF | Club)
         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
model2_5    6 824.35 842.00 -406.17   812.35
model3_1    8 828.11 851.65 -406.06   812.11 0.2357  2     0.8888
```

```
model3_2 <- lmer(Pts ~ GF + GA + Mean_NetSpend + (1 + GA | Club), data = prem)
```

```
boundary (singular) fit: see help('isSingular')
```

```
anova(model2_5, model3_2)
```

```
refitting model(s) with ML (instead of REML)

Data: prem
Models:
model2_5: Pts ~ GF + GA + Mean_NetSpend + (1 | Club)
model3_2: Pts ~ GF + GA + Mean_NetSpend + (1 + GA | Club)
         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
model2_5    6 824.35 842.00 -406.17   812.35
model3_2    8 828.30 851.83 -406.15   812.30 0.0507  2      0.975
```

```
model3_3 <- lmer(Pts ~ GF + GA + Mean_NetSpend + (1 + GA + GF | Club), data = prem)
```

```
boundary (singular) fit: see help('isSingular')
```

```
anova(model2_5, model3_2)
```

```
refitting model(s) with ML (instead of REML)

Data: prem
Models:
model2_5: Pts ~ GF + GA + Mean_NetSpend + (1 | Club)
model3_2: Pts ~ GF + GA + Mean_NetSpend + (1 + GA | Club)
         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
model2_5    6 824.35 842.00 -406.17   812.35
model3_2    8 828.30 851.83 -406.15   812.30 0.0507  2      0.975
```

Final Model

```
prem$GF_c <- scale(prem$GF, scale = FALSE)
prem$GA_c <- scale(prem$GA, scale = FALSE)
prem$Mean_NetSpend_c <- scale(prem$Mean_NetSpend, scale = FALSE)

final_model <- lmer(Pts ~ GF_c + GA_c + Mean_NetSpend_c + (1 | Club), data = prem)
```

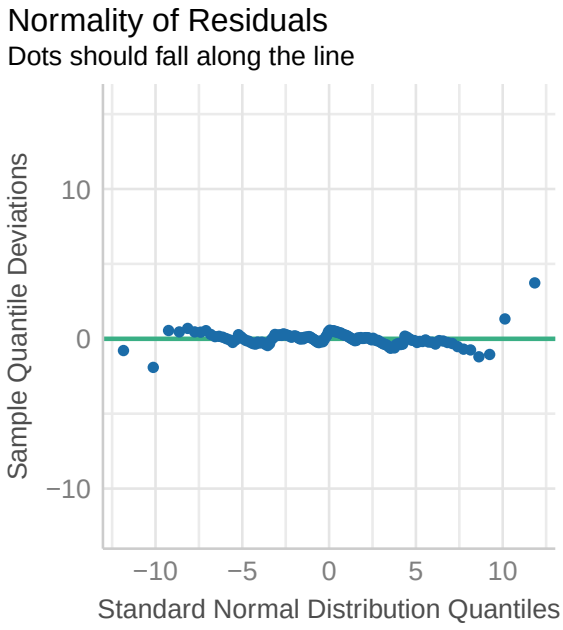
```
boundary (singular) fit: see help('isSingular')
```

```
summary(final_model)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Pts ~ GF_c + GA_c + Mean_NetSpend_c + (1 | Club)
   Data: prem

REML criterion at convergence: 815.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.8271 -0.6048  0.1171  0.6113  3.4867

Random effects:
 Groups   Name        Variance Std.Dev.
 Club     (Intercept)  0.00    0.000
 Residual             19.96    4.467
Number of obs: 140, groups:  Club, 30

Fixed effects:
                 Estimate Std. Error t value
(Intercept)      52.62857    0.37756 139.390
GF_c             23.03892    1.05930  21.749
GA_c            -21.86604    1.29747 -16.853
Mean_NetSpend_c   0.03120    0.01181   2.643

Correlation of Fixed Effects:
            (Intr) GF_c   GA_c
GF_c         0.000
GA_c         0.000  0.557
Mn_NtSpnd_c  0.000 -0.362  0.056
```

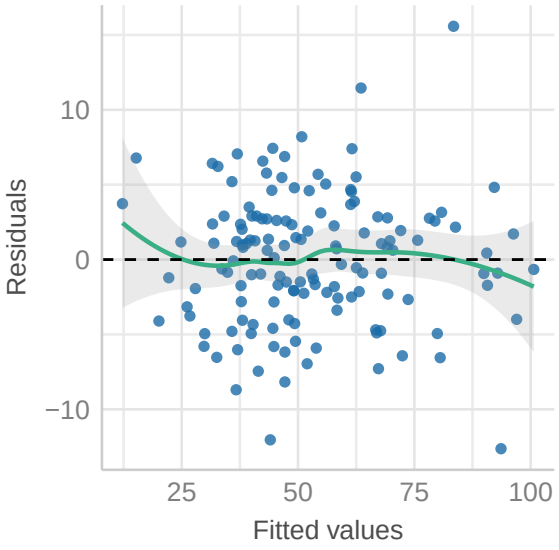
```
optimizer (nloptwrap) convergence code: 0 (OK)
boundary (singular) fit: see help('isSingular')
```

Model Diagnostics



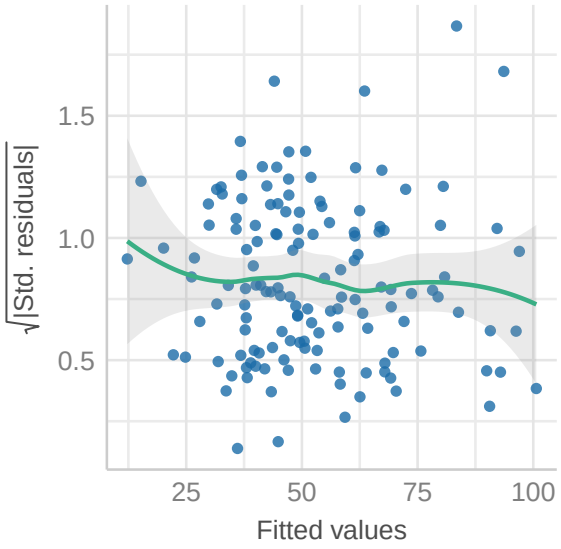
Linearity

Reference line should be flat and horizontal



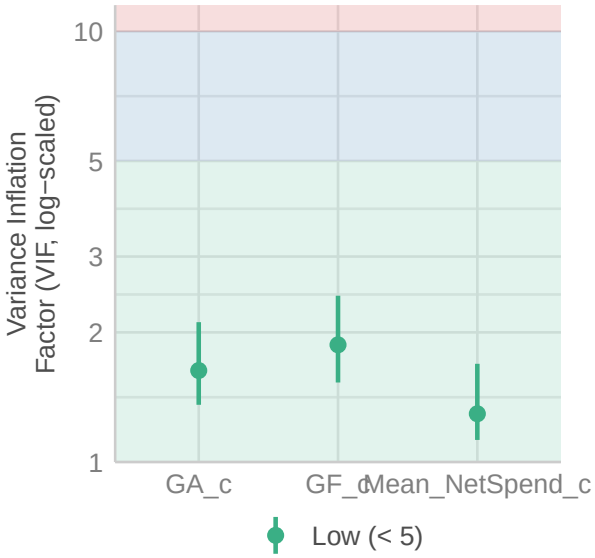
Homogeneity of Variance

Reference line should be flat and horizontal



Collinearity

High collinearity (VIF) may inflate parameter uncertainty



Influential Observations

Points should be inside the contour lines