

General UEFA Soccer Performance



Anusha Bhat
Nandini Neralagi
Noelani Phillips
Marcos Pi Marrero

UEFA Dataset

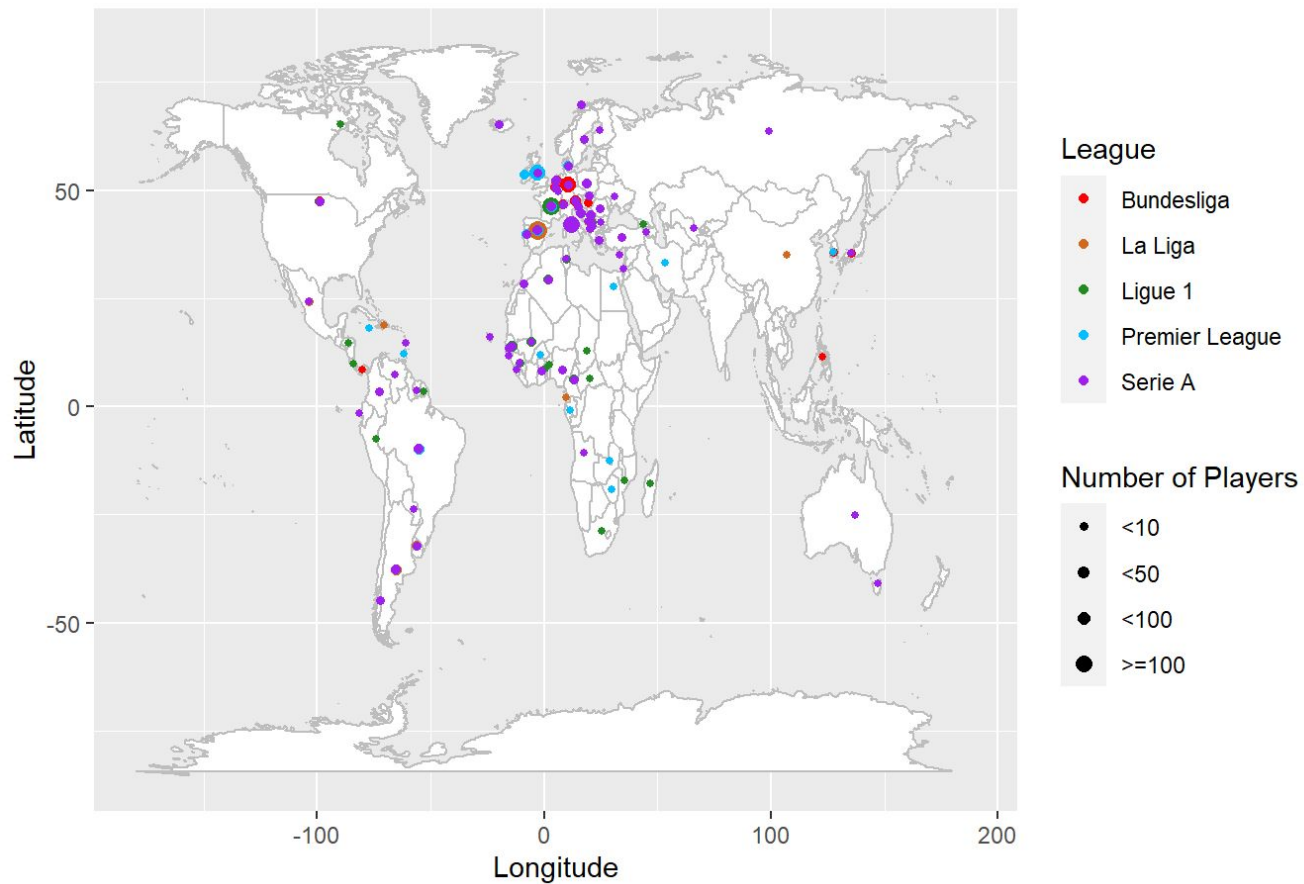
- 2,921 male soccer players across prominent UEFA leagues,
 - including the Premier League, Ligue 1, Bundesliga, Serie A, and La Liga
- 143 variables
 - player demographic information
 - passing accuracy
 - offensive and defensive actions
 - other key aspects of individual and team performance
- Includes all UEFA games from 2021-2022

Research Questions

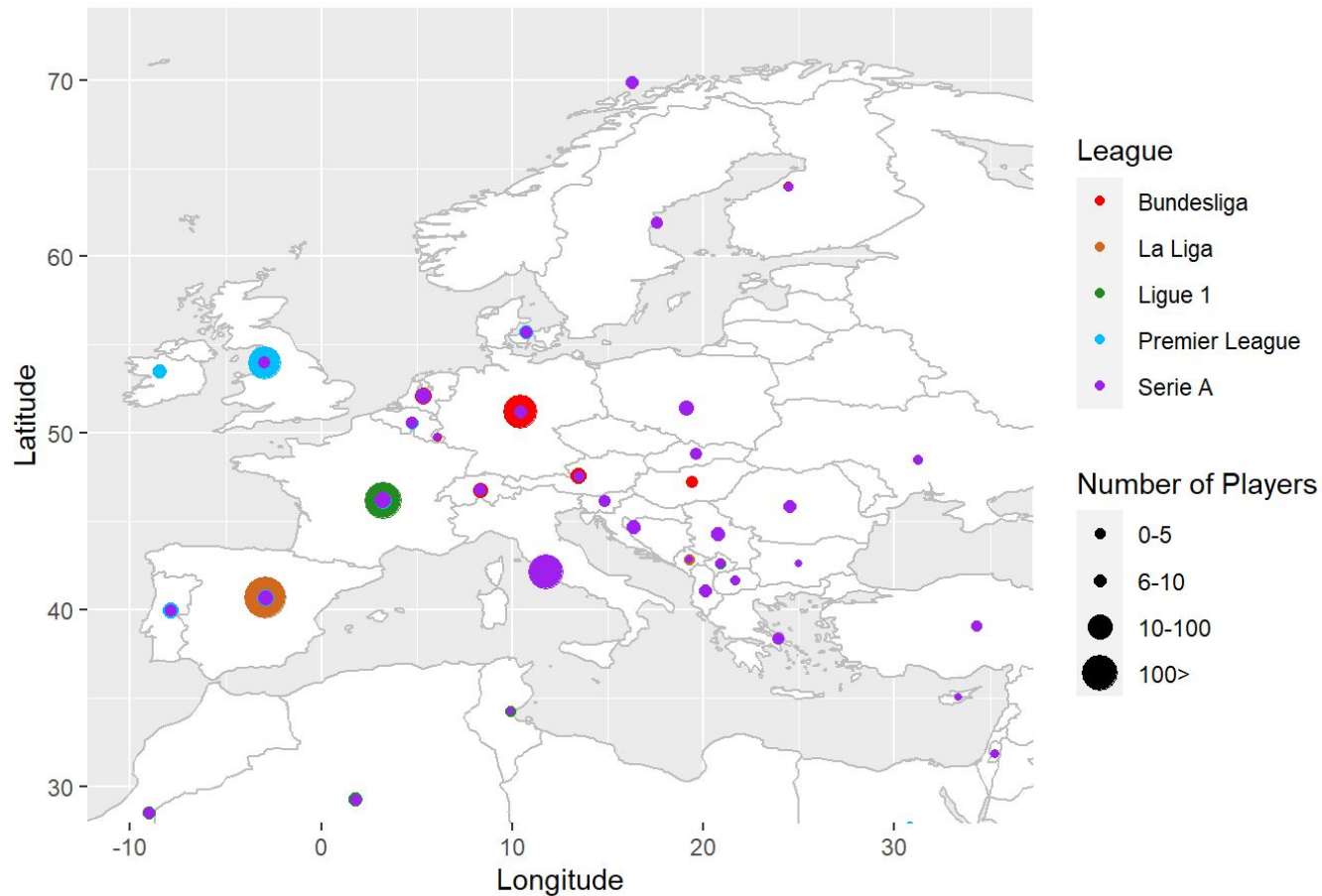
1. Which geographical areas are under and over-represented in the UEFA European Leagues, and how do these representations break down across different positions?
2. Do any trends of higher passing performance emerge within a specific position and league?
3. What defensive factors are predictors of whether a player is able to make a goal, and does this potential relationship vary by position?

Which geographical areas are under and over-represented in the UEFA European Leagues, and how do these representations break down across different positions?

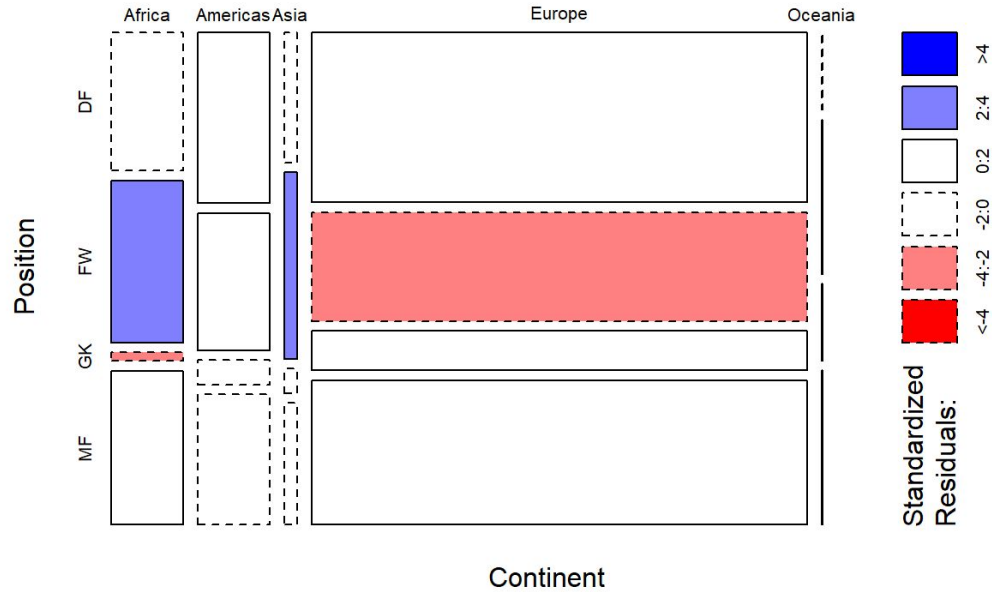
Players' Nationalities by League



Players' Nationalities by League (Europe)



Distribution of Player Position by Continent of Nationality

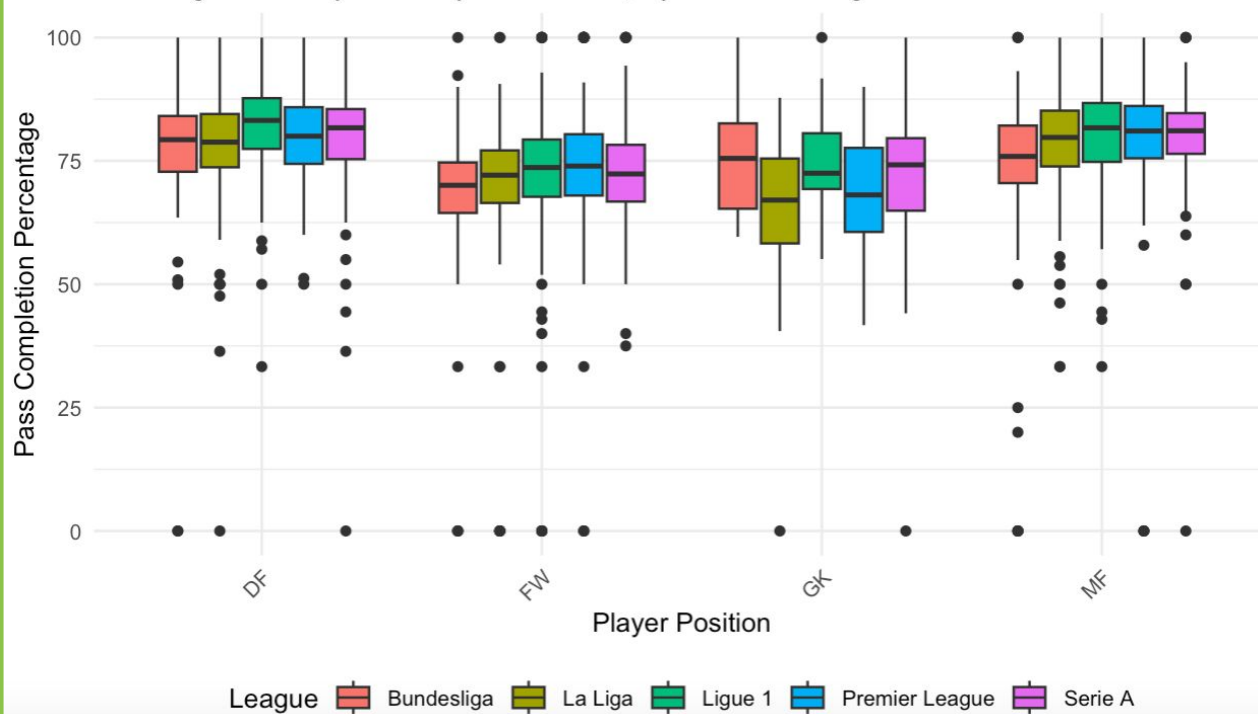


Null Hypothesis:

- The Continent and Position Variables are independent.

Do any trends of higher passing performance emerge within a specific position and league?

Passing Accuracy vs. Player Position, by Soccer Leagues



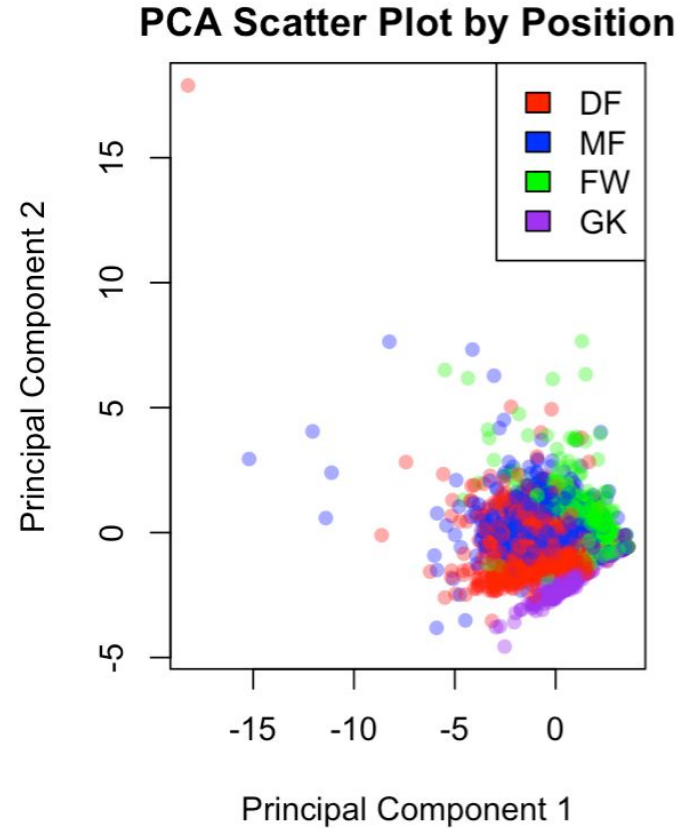
PasTotCmp% ~ Pos_simplified * Comp

ANOVA Result

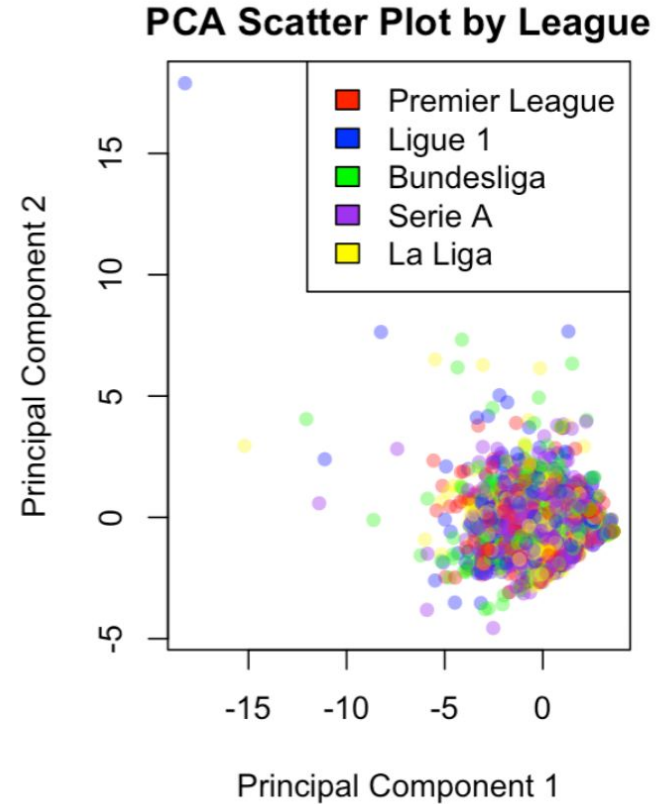
```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Pos_simplified    3  41498   13833   94.623 < 2e-16 ***
## Comp              4   6916    1729   11.827 1.57e-09 ***
## Pos_simplified:Comp 12   4830     403    2.754 0.000999 ***
## Residuals        2901 424093     146
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Intercept)              77.5778      0.8687  89.300 < 2e-16 ***
factor(Pos_simplified)FW  -9.0877      1.3781  -6.594 5.17e-11 ***
factor(Pos_simplified)GK  -3.0144      2.0396  -1.478 0.139548
factor(Pos_simplified)MF  -3.8478      1.2983  -2.964 0.003066 **
```

- PasTotCmp
- PasTotDist
- PasTotPrgDist
- PasAss
- PasProg
- PasInt
- PasBlocks

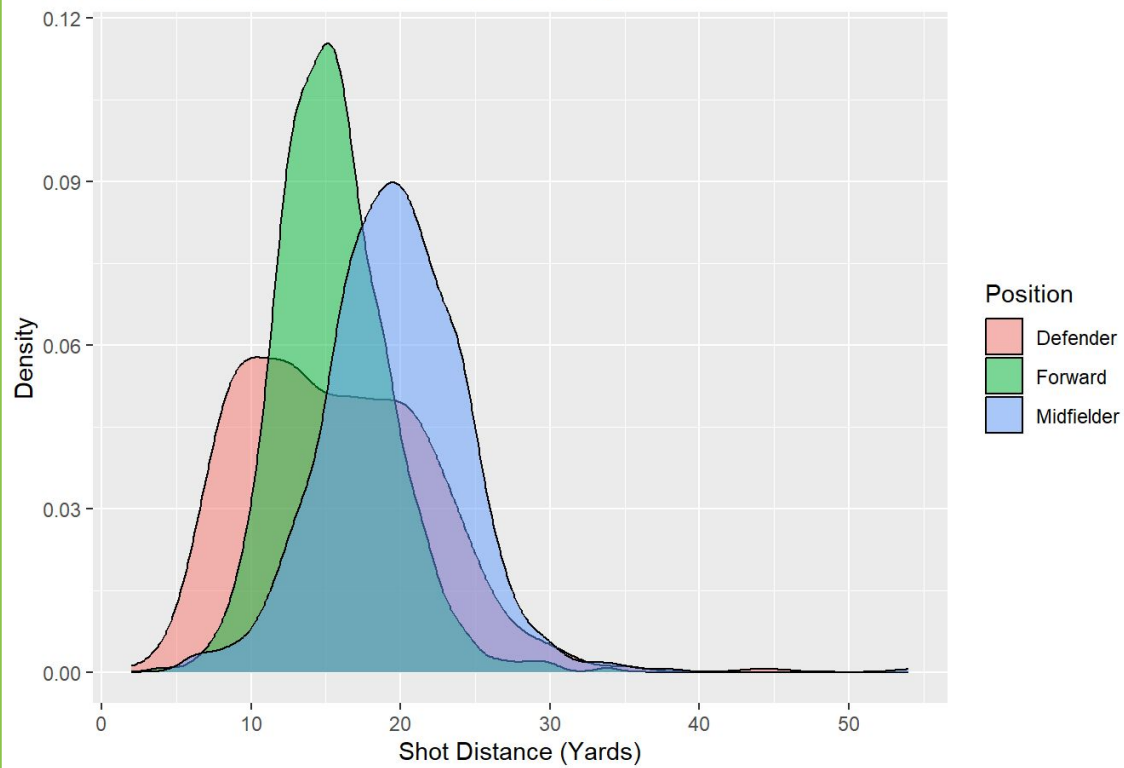


- PasTotCmp
- PasTotDist
- PasTotPrgDist
- PasAss
- PasProg
- PasInt
- PasBlocks

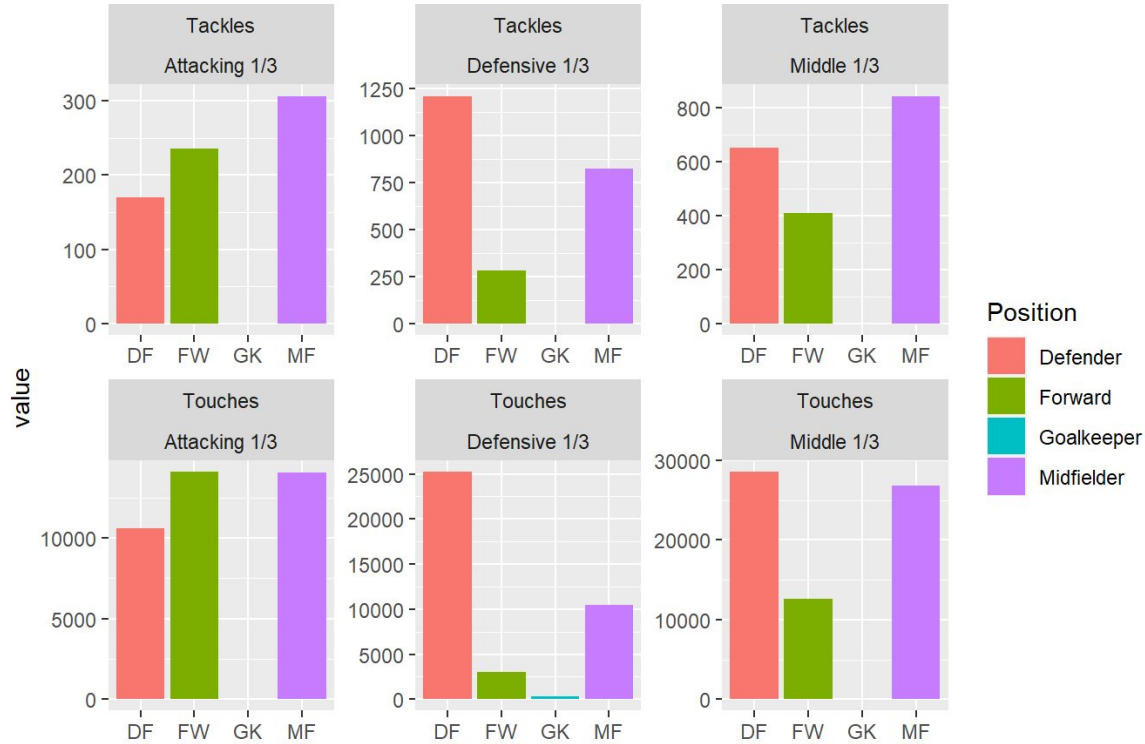


What defensive factors are predictors of whether a player is able to make a goal, and does this potential relationship vary by position?

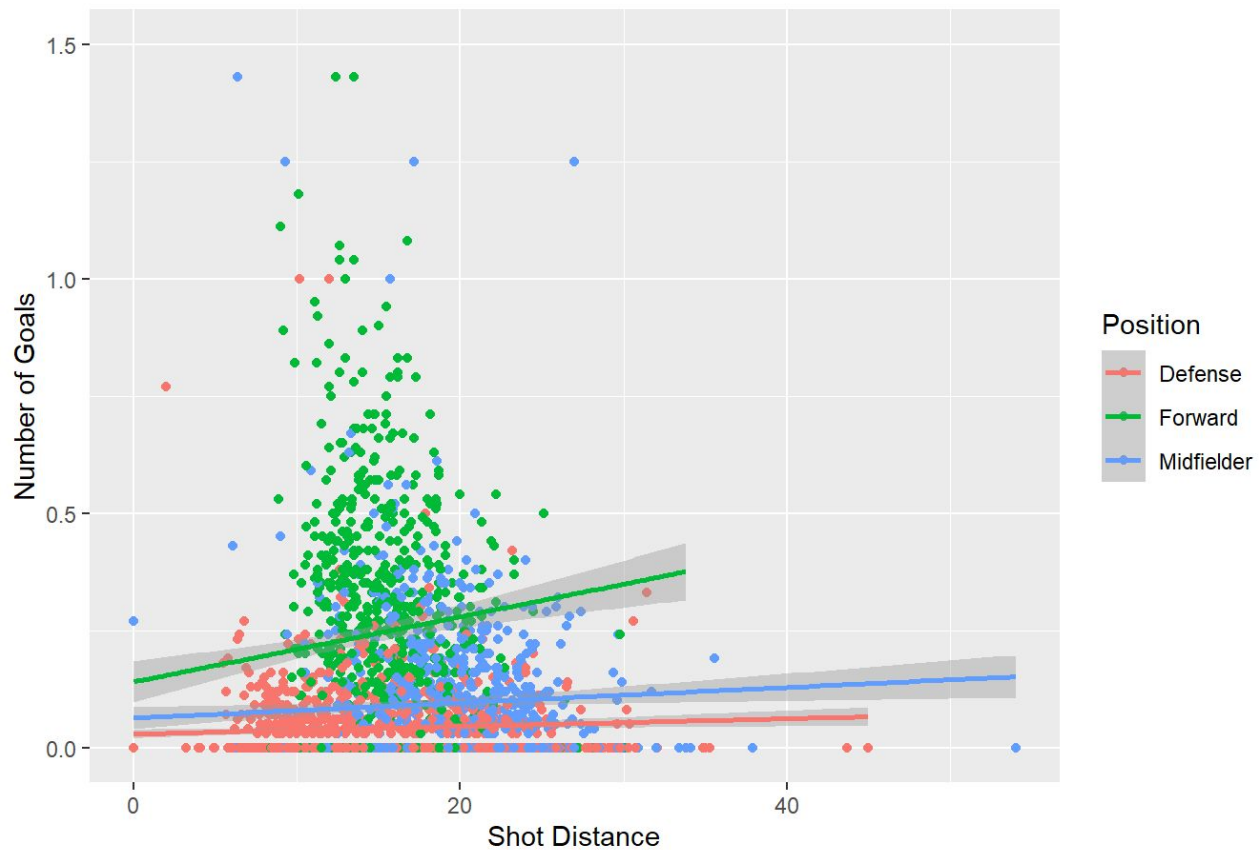
Distribution of Shot Distance across Player Position



Total Player Touches and Tackles by Zone and Position



Shot Distance v.s. Number of Goals by Player Position



Analysis of Variance Table

Model 1: Goals ~ Int + Fls

Model 2: Goals ~ ShoDist * Pos_simplified + Int + Tkl + Fls

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	2701	150.56				
2	2695	132.63	6	17.933	60.732	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null hypothesis:

- Reduced model is correct

Alternative hypothesis:

- Full model is correct

```
##
## Call:
## lm(formula = Goals ~ ShoDist * Pos_simplified + Int + Tkl + Fls,
##     data = noGoalie)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.3680 -0.0849 -0.0347  0.0412  4.7572
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.0519446   0.0162431   3.198   0.0014 **
## ShoDist                    0.0009145   0.0008670   1.055   0.2916
## Pos_simplifiedFW           0.1494937   0.0256989   5.817 6.69e-09 ***
## Pos_simplifiedMF           0.0329268   0.0228894   1.439   0.1504
## Int                        -0.0071958   0.0038362  -1.876   0.0608 .
## Tkl                        -0.0078002   0.0032996  -2.364   0.0181 *
## Fls                        0.0043884   0.0031099   1.411   0.1583
## ShoDist:Pos_simplifiedFW    0.0039499   0.0016270   2.428   0.0153 *
## ShoDist:Pos_simplifiedMF    0.0007627   0.0012928   0.590   0.5553
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2218 on 2695 degrees of freedom
## Multiple R-squared:  0.1528, Adjusted R-squared:  0.1503
## F-statistic: 60.77 on 8 and 2695 DF, p-value: < 2.2e-16
```

Conclusions

- Europe and Africa are more commonly represented areas in UEFA Leagues
- Forwards are over-represented in Africa and underrepresented in Europe
- Goalkeepers from Africa are underrepresented
- Defenders tend to exhibit higher passing performance
- PCA shows within-group similarity for positions in passing metrics, but no noticeable trends for leagues.
- Player position, shooting distance, and tackles are significant predictors for number of goals
- Shooting distance has a positive, linear association with goals, especially for forwards

Future Outlook

Potential areas to explore in the future include:

1. How team tactics influence individual passing performance and whether certain approaches enhance passing capabilities of players in specific positions.
2. How player fatigue affects their performance
3. Which teams have the most effective set-piece strategies