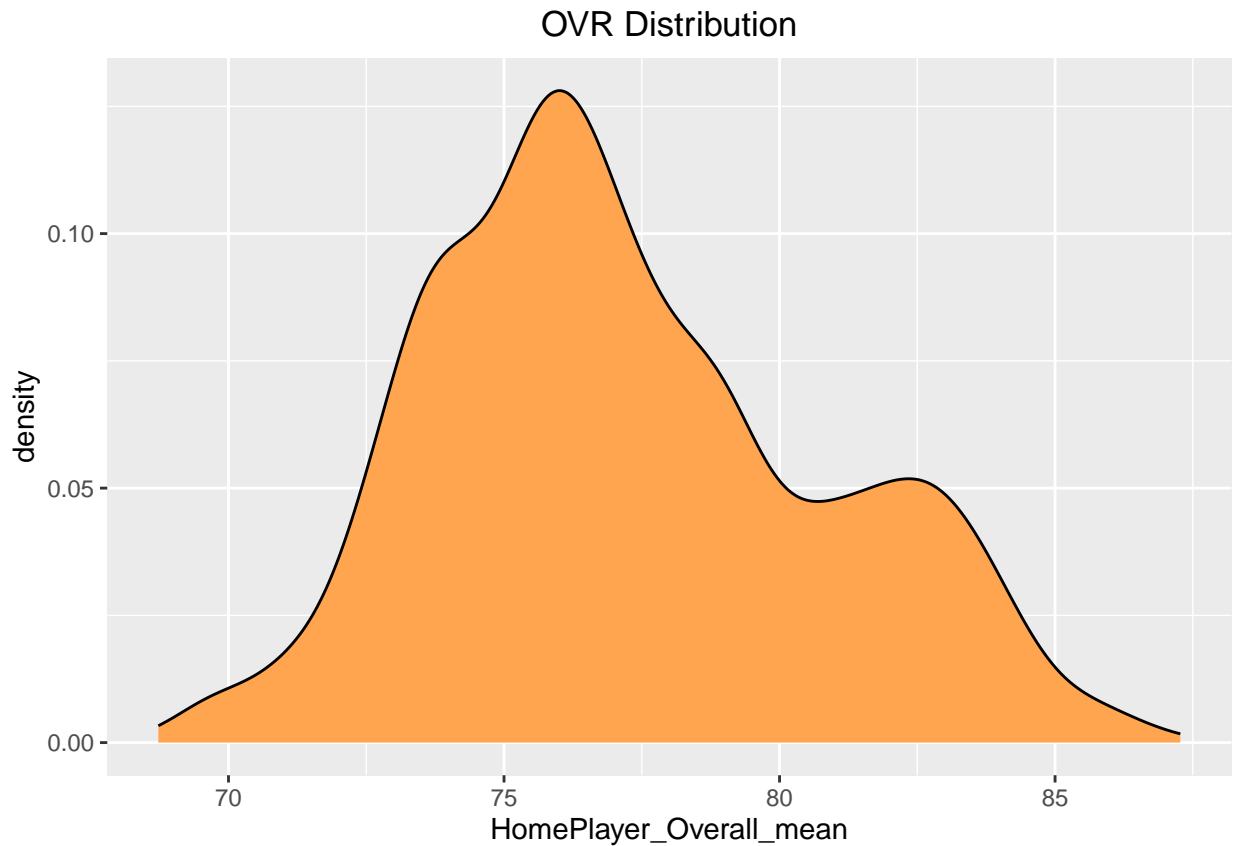


Exploring FIFA Ratings and motivation

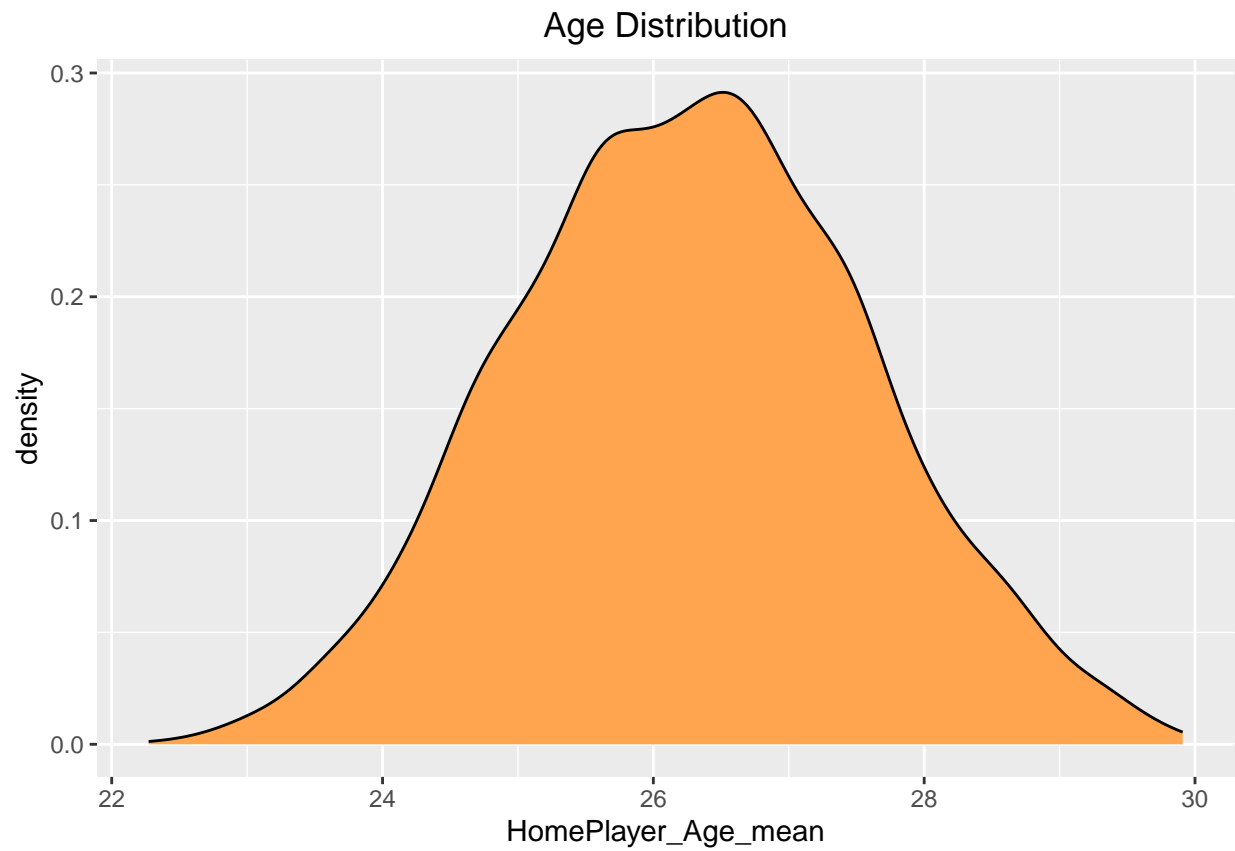
2024-01-25

Understanding distributions

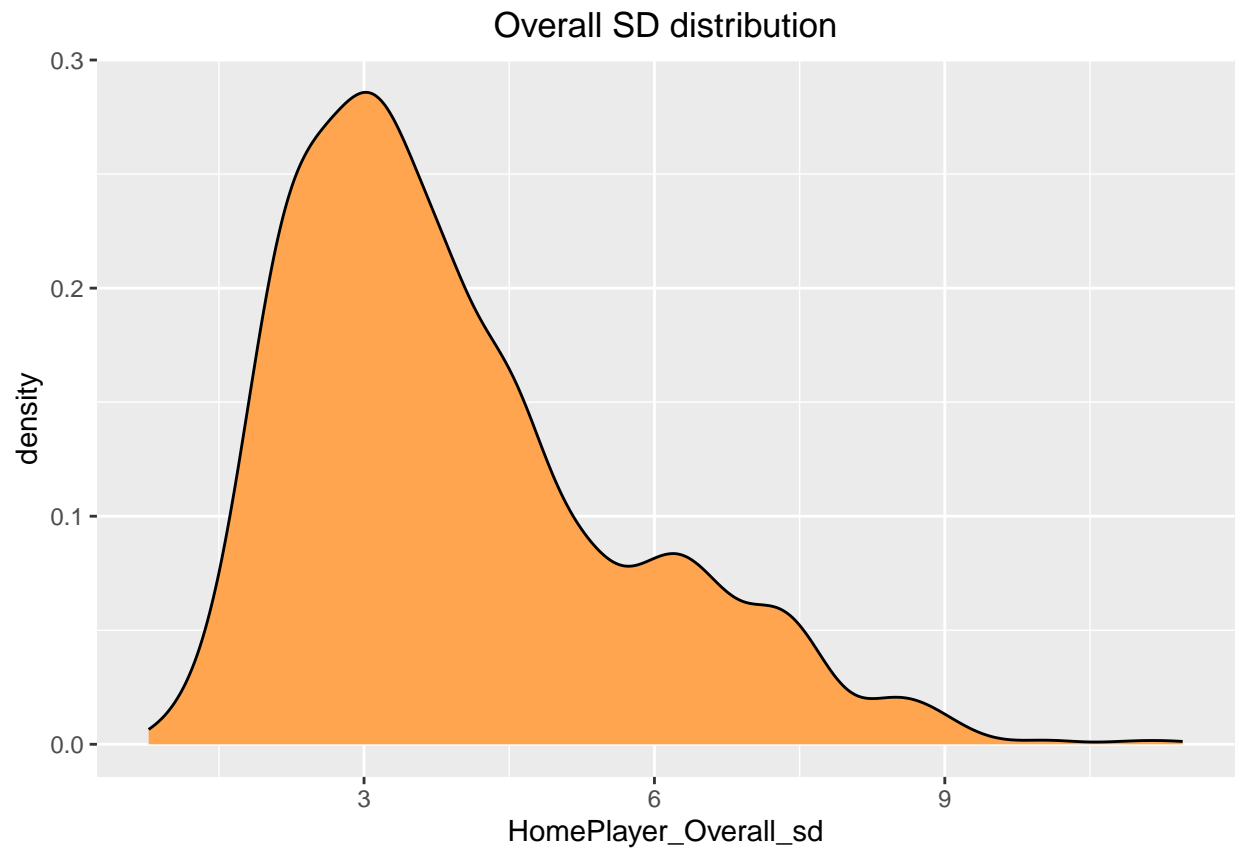
```
ggplot(data=data, aes(x=HomePlayer_Overall_mean)) +  
  geom_density(fill="tan1") + labs(title="OVR Distribution") + theme(plot.title = element_text(hjust = 0.5))
```



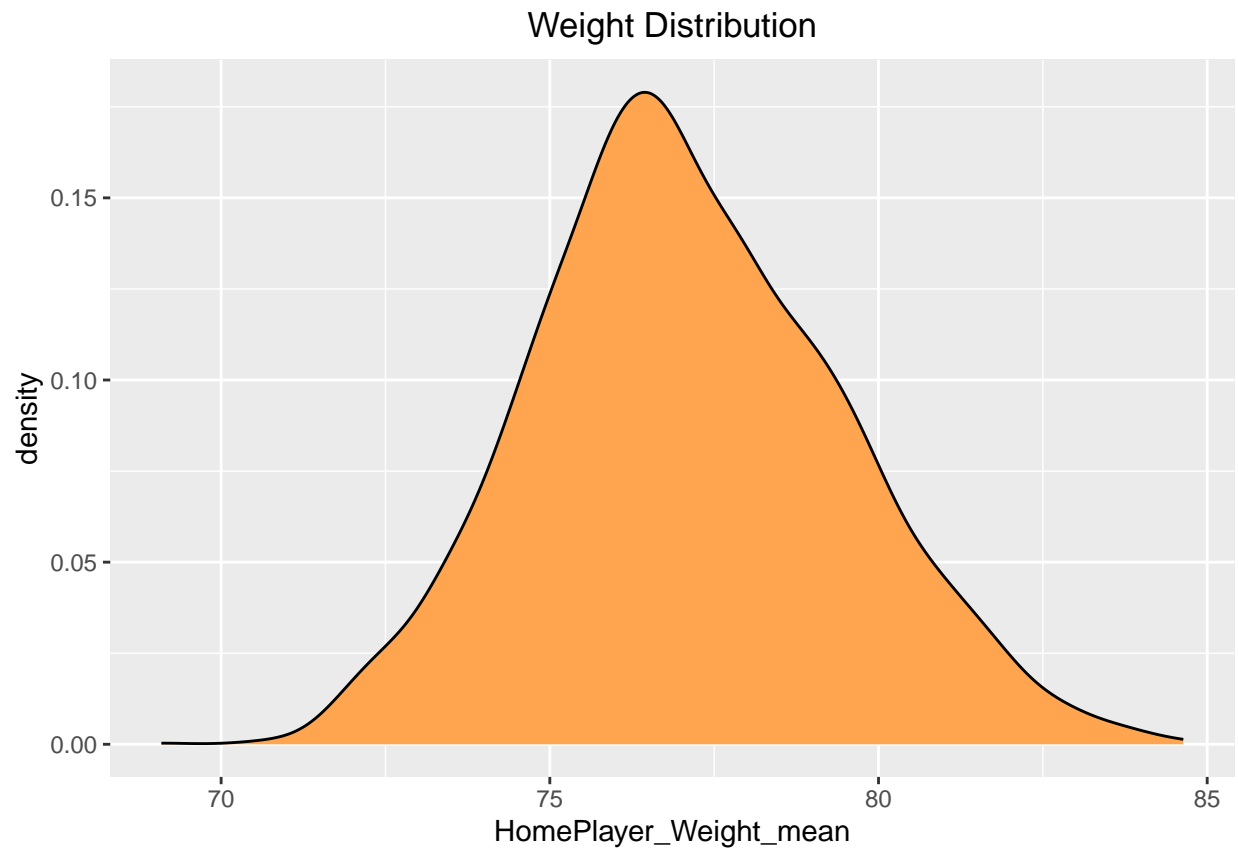
```
ggplot(data=data, aes(x=HomePlayer_Age_mean)) +  
  geom_density(fill="tan1") + labs(title="Age Distribution") + theme(plot.title = element_text(hjust = 0.5))
```



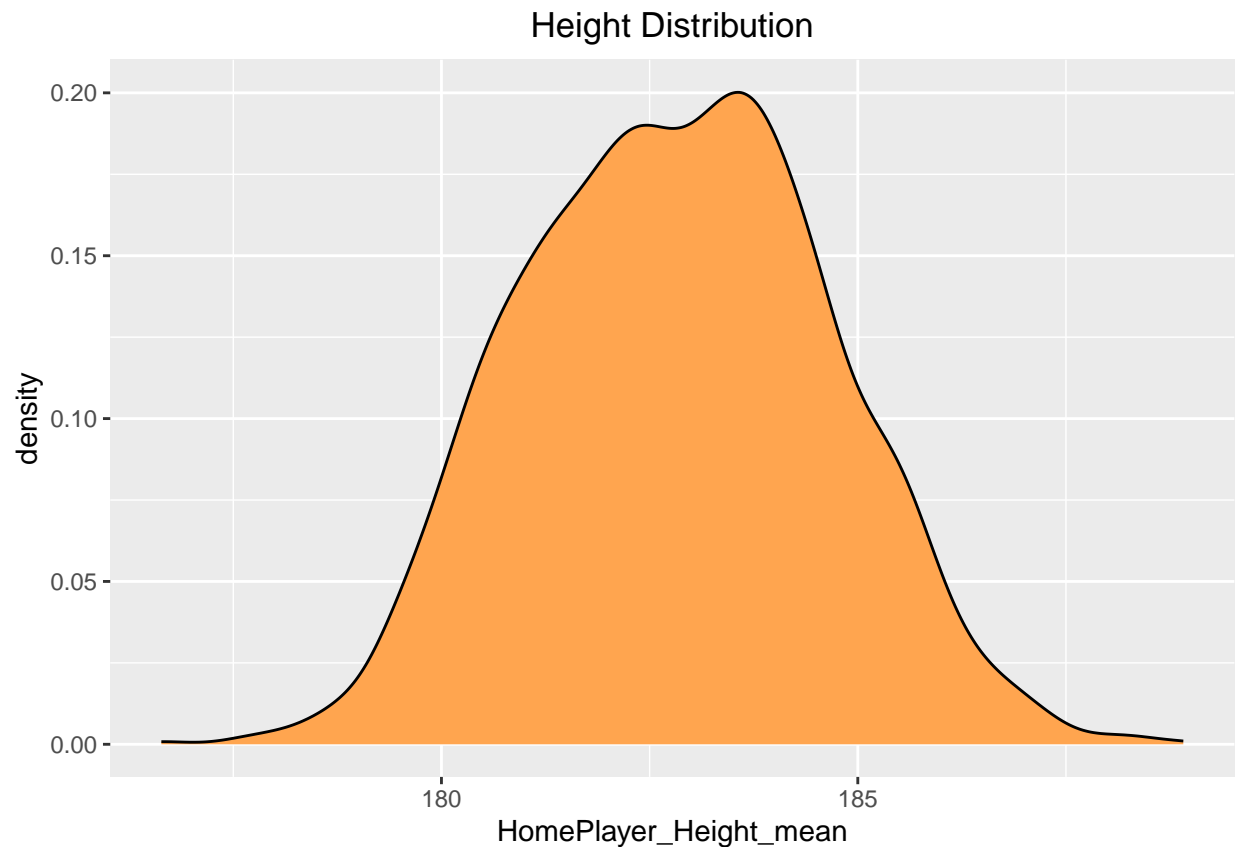
```
ggplot(data=data, aes(x=HomePlayer_Overall_sd)) +  
  geom_density(fill="tan1") + labs(title="Overall SD distribution") + theme(plot.title = element_text(h
```



```
ggplot(data=data, aes(x=HomePlayer_Weight_mean)) +  
  geom_density(fill="tan1") + labs(title="Weight Distribution") + theme(plot.title = element_text(hjust
```



```
ggplot(data=data, aes(x=HomePlayer_Height_mean)) +  
  geom_density(fill="tan1") + labs(title="Height Distribution") + theme(plot.title = element_text(hjust
```



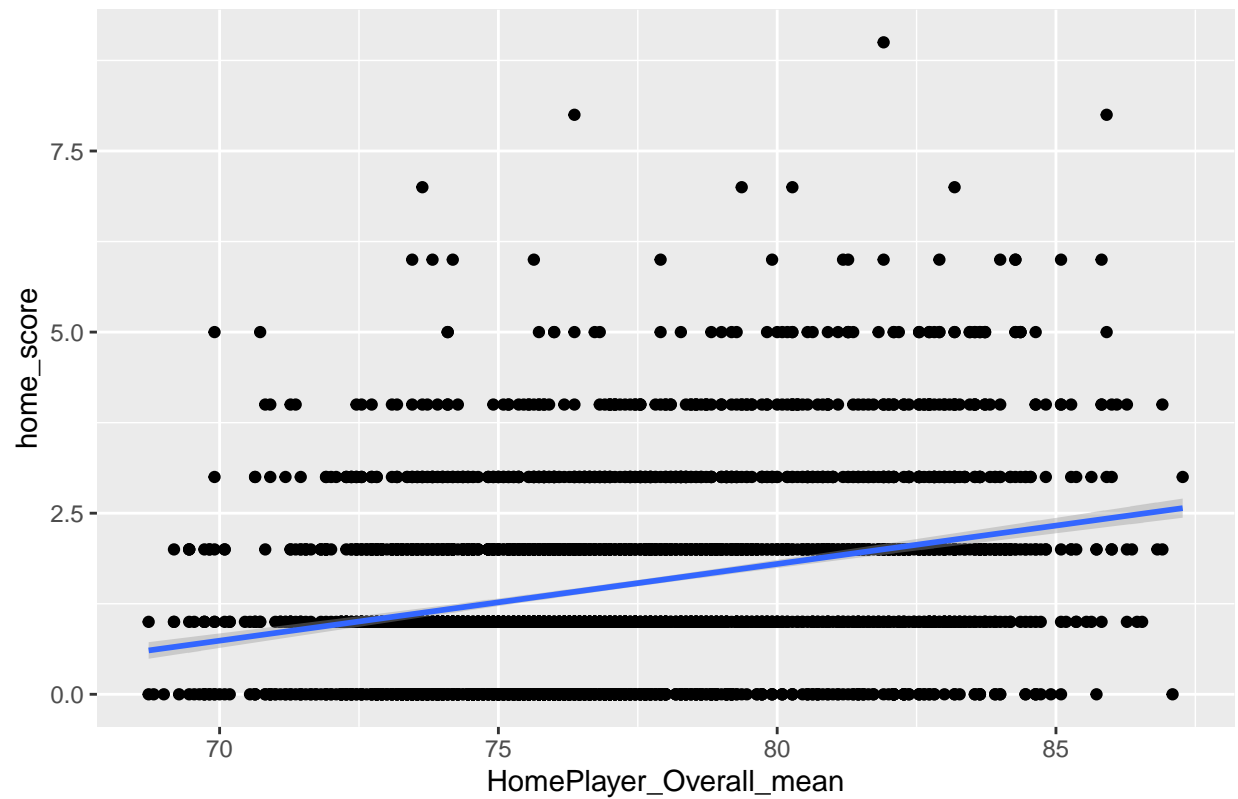
What we can see is that mean values distributions are close to normal and the sd reminds us of the χ^2 Distribution

home_score and each of the features. This will help us understand which values need the log transformation all graphs are for per starting lineup

```
ggplot(data = data, aes(x = HomePlayer_Overall_mean,  
                        y = home_score)) +  
  geom_point() + stat_smooth(method = "lm") +  
  ggtitle("Home Overall and the home_score")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

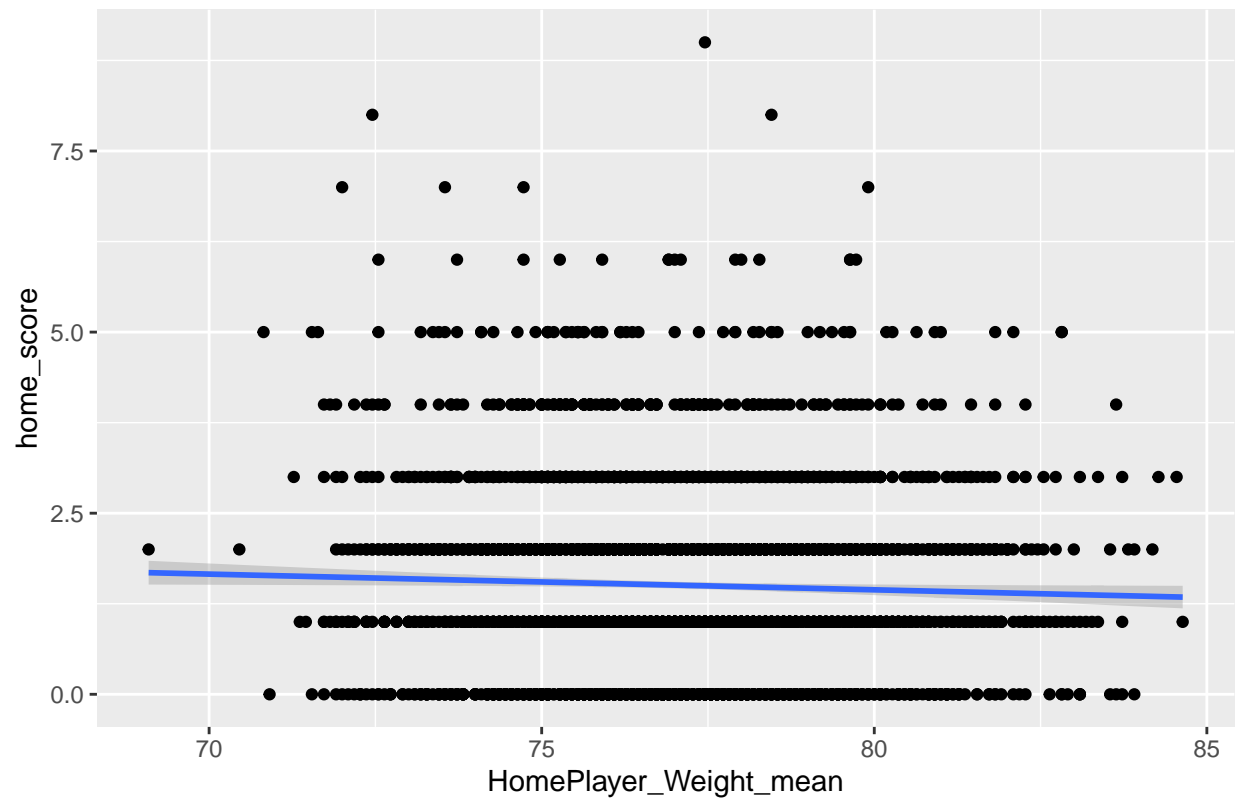
Home Overall and the home_score



```
ggplot(data = data, aes(x = HomePlayer_Weight_mean,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Weight and the home_score")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

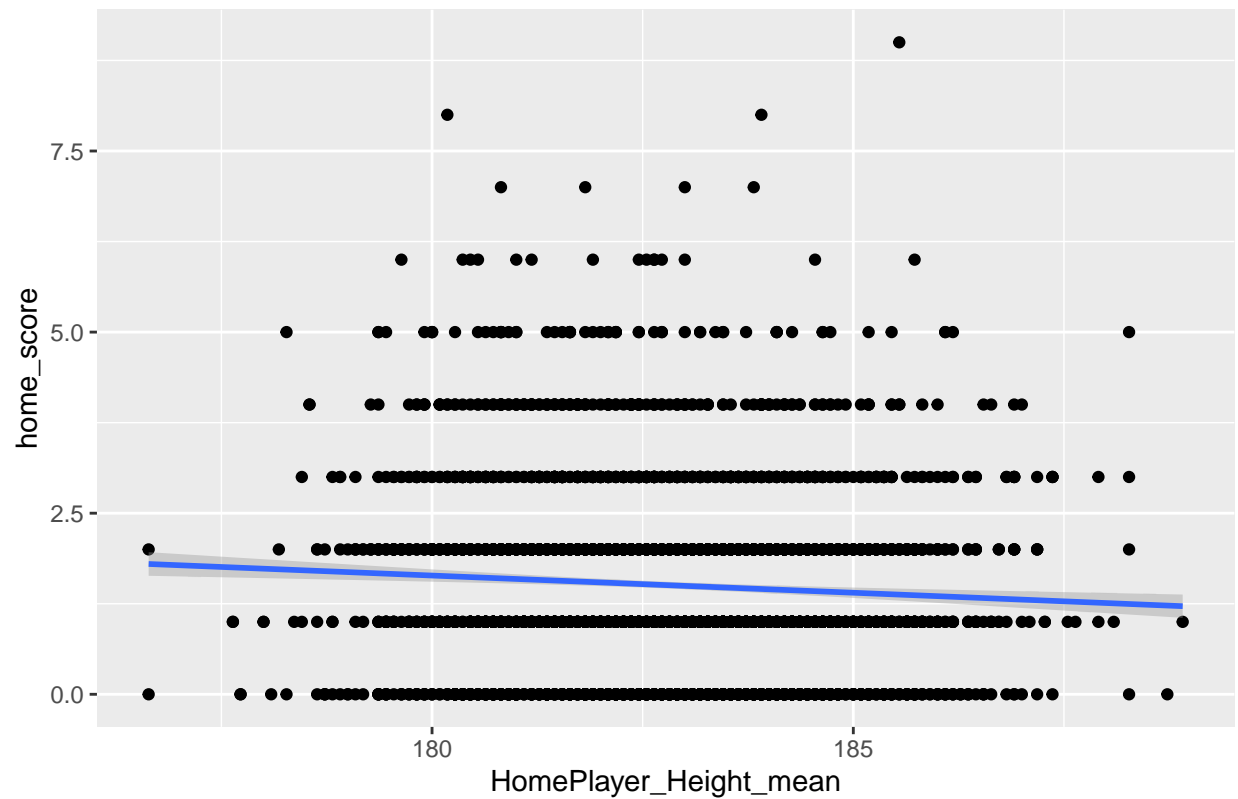
Home Weight and the home_score



```
ggplot(data = data, aes(x = HomePlayer_Height_mean,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Height and the home_score")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

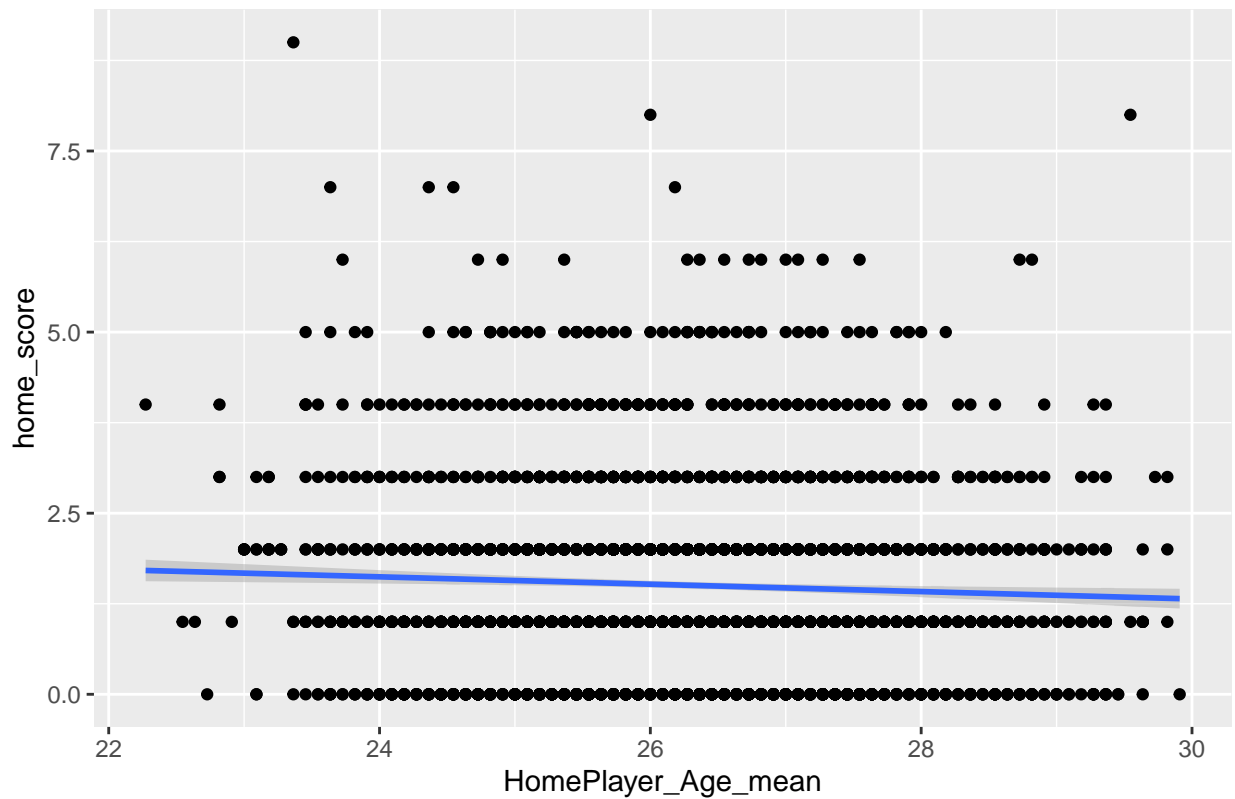
Home Height and the home_score



```
ggplot(data = data, aes(x = HomePlayer_Age_mean,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Age and the home_score")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

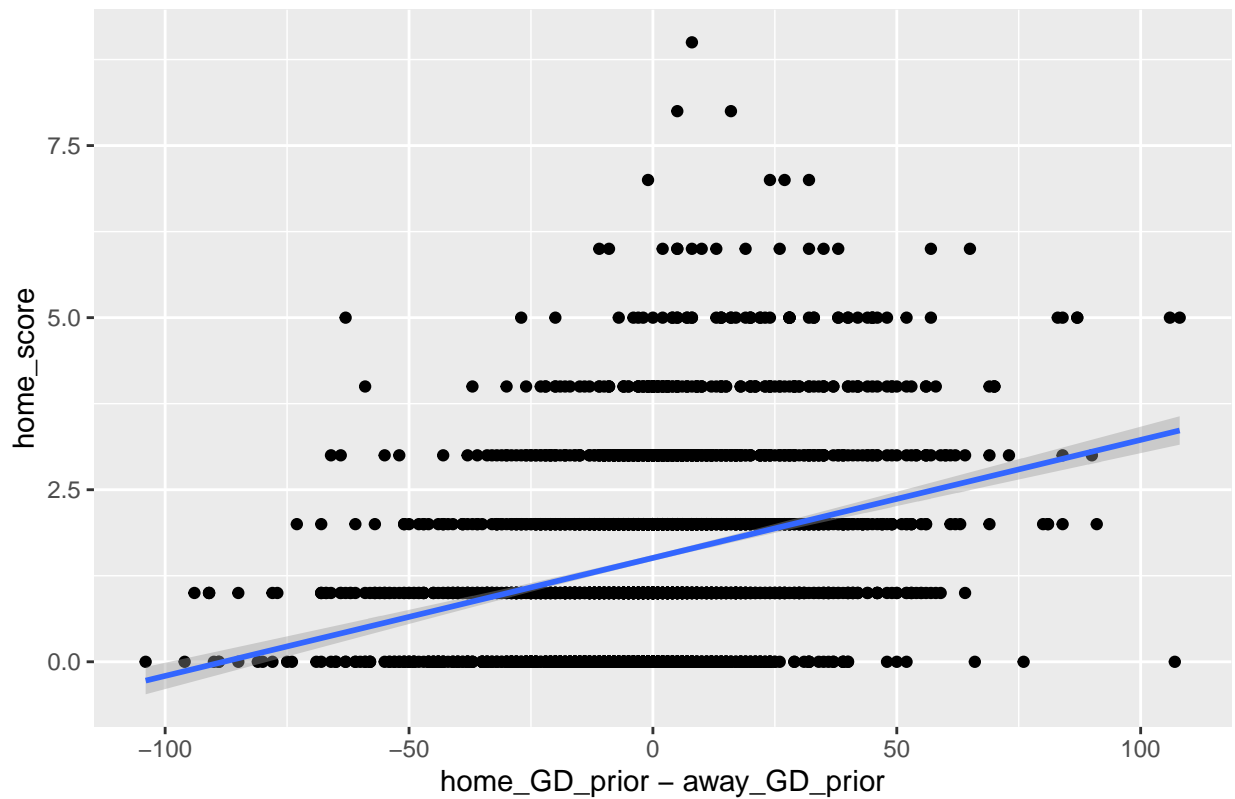

Home Age and the home_score



```
ggplot(data = data, aes(x = home_GD_prior - away_GD_prior,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Age and the home_score")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

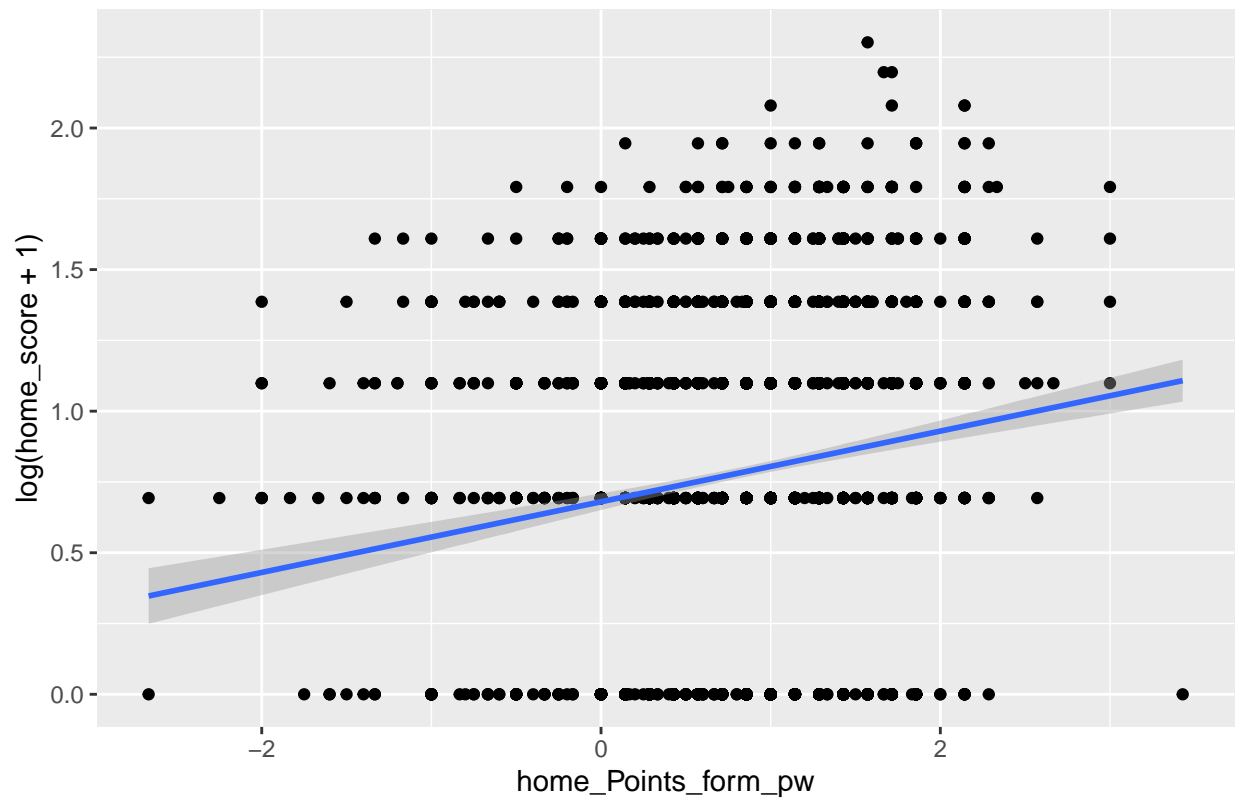
Home Age and the home_score



```
ggplot(data = data, aes(x = home_Points_form_pw,  
  y = log(home_score+1))) +  
  geom_point() + stat_smooth(method = "lm") +  
  ggtitle("Home Age and the home_score")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

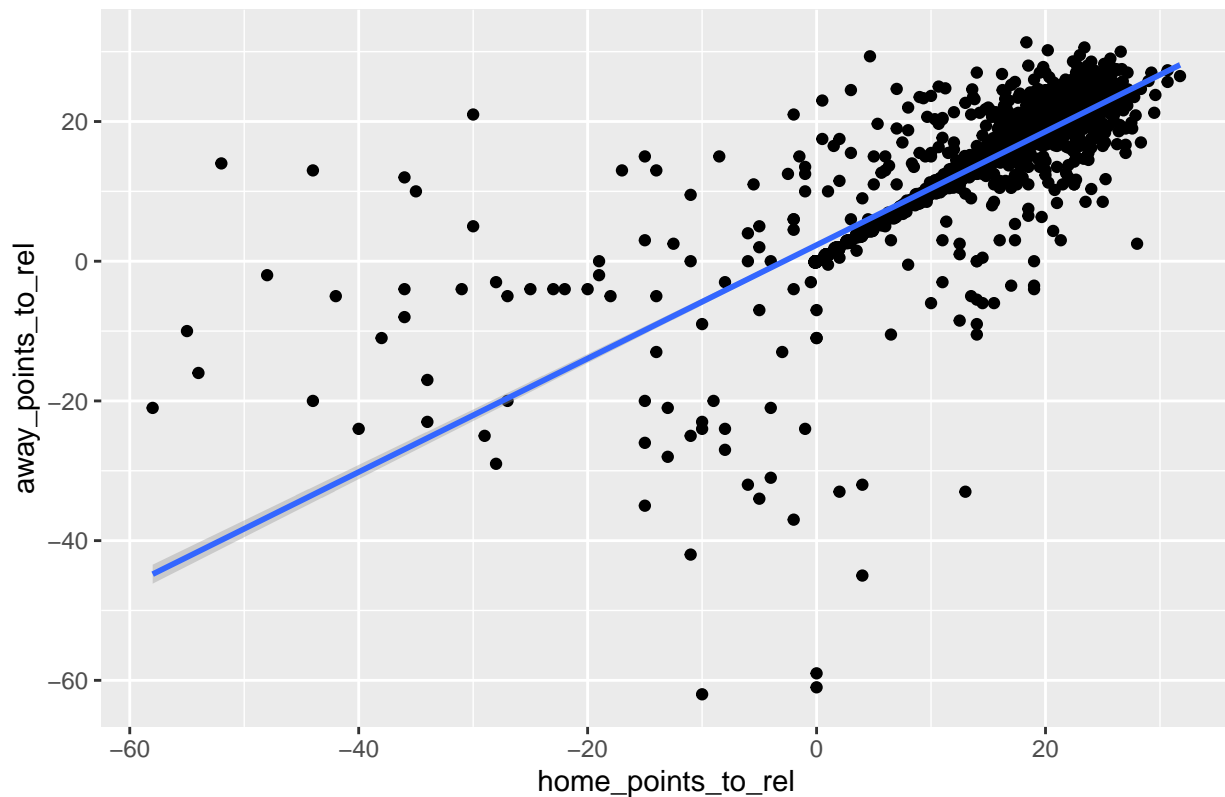
Home Age and the home_score



```
ggplot(data = data, aes(x = home_points_to_rel,
                        y = away_points_to_rel)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Age and the home_score")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

Home Age and the home_score



We can see only for the average rating of the starting 11, that it has positive effect, and for the others its pretty close to even 0.

Why is it?

Maybe because for some areas of the age, raising the average could be beneficial, for example: A team who is full of very young and inexperienced players could be worse than a team full of 26-30 year old players. It could also be true for both height and weight. Weight is self explanatory, but height could give us the idea, that for example there is connection between player height and speed / agility.

How can we show it?

Maybe we can try when running our models, keep in mind adding a squared version of height weight and age

```
teams_lm = lm(data=data, formula = I(home_score-away_score) ~
              home_team_name+away_team_name)
summary(teams_lm)
```

```
##
## Call:
## lm(formula = I(home_score - away_score) ~ home_team_name + away_team_name,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.9372 -1.0813 -0.0153  1.0246  7.9065
##
```

```

## Coefficients:
##
## (Intercept)
## home_team_nameArsenal
## home_team_nameAston Villa
## home_team_nameBrentford
## home_team_nameBrighton & Hove Albion
## home_team_nameBurnley
## home_team_nameCardiff City
## home_team_nameChelsea
## home_team_nameCrystal Palace
## home_team_nameEverton
## home_team_nameFulham
## home_team_nameHuddersfield Town
## home_team_nameHull City
## home_team_nameLeeds United
## home_team_nameLeicester City
## home_team_nameLiverpool
## home_team_nameManchester City
## home_team_nameManchester United
## home_team_nameMiddlesbrough
## home_team_nameNewcastle United
## home_team_nameNorwich City
## home_team_nameQueens Park Rangers
## home_team_nameSheffield United
## home_team_nameSouthampton
## home_team_nameStoke City
## home_team_nameSunderland
## home_team_nameSwansea City
## home_team_nameTottenham Hotspur
## home_team_nameWatford
## home_team_nameWest Bromwich Albion
## home_team_nameWest Ham United
## home_team_nameWolverhampton Wanderers
## away_team_nameArsenal
## away_team_nameAston Villa
## away_team_nameBrentford
## away_team_nameBrighton & Hove Albion
## away_team_nameBurnley
## away_team_nameCardiff City
## away_team_nameChelsea
## away_team_nameCrystal Palace
## away_team_nameEverton
## away_team_nameFulham
## away_team_nameHuddersfield Town
## away_team_nameHull City
## away_team_nameLeeds United
## away_team_nameLeicester City
## away_team_nameLiverpool
## away_team_nameManchester City
## away_team_nameManchester United
## away_team_nameMiddlesbrough
## away_team_nameNewcastle United
## away_team_nameNorwich City

```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	0.39994	0.24238	1.650	0.09904	.
home_team_nameArsenal	1.16656	0.21680	5.381	7.99e-08	***
home_team_nameAston Villa	-0.21954	0.24059	-0.912	0.36158	
home_team_nameBrentford	0.23046	0.41668	0.553	0.58025	
home_team_nameBrighton & Hove Albion	-0.07717	0.24067	-0.321	0.74849	
home_team_nameBurnley	-0.08383	0.22266	-0.376	0.70659	
home_team_nameCardiff City	-0.75766	0.41668	-1.818	0.06911	.
home_team_nameChelsea	1.05962	0.21680	4.887	1.08e-06	***
home_team_nameCrystal Palace	-0.01399	0.21680	-0.065	0.94855	
home_team_nameEverton	0.47247	0.21680	2.179	0.02939	*
home_team_nameFulham	-0.69267	0.31829	-2.176	0.02962	*
home_team_nameHuddersfield Town	-0.62791	0.31818	-1.973	0.04854	*
home_team_nameHull City	-0.19213	0.31840	-0.603	0.54628	
home_team_nameLeeds United	-0.13889	0.31839	-0.436	0.66270	
home_team_nameLeicester City	0.55719	0.21680	2.570	0.01022	*
home_team_nameLiverpool	1.46761	0.21680	6.769	1.55e-11	***
home_team_nameManchester City	1.88913	0.21680	8.714	< 2e-16	***
home_team_nameManchester United	0.99469	0.21680	4.588	4.66e-06	***
home_team_nameMiddlesbrough	-0.19454	0.41663	-0.467	0.64058	
home_team_nameNewcastle United	0.15347	0.22270	0.689	0.49080	
home_team_nameNorwich City	-0.74436	0.27783	-2.679	0.00742	**
home_team_nameQueens Park Rangers	0.07695	0.41659	0.185	0.85346	
home_team_nameSheffield United	0.02391	0.31832	0.075	0.94013	
home_team_nameSouthampton	0.22906	0.21680	1.057	0.29080	
home_team_nameStoke City	0.08075	0.25524	0.316	0.75175	
home_team_nameSunderland	-0.35036	0.27794	-1.261	0.20756	
home_team_nameSwansea City	-0.01161	0.25524	-0.045	0.96372	
home_team_nameTottenham Hotspur	1.12767	0.21680	5.201	2.11e-07	***
home_team_nameWatford	-0.23983	0.23029	-1.041	0.29776	
home_team_nameWest Bromwich Albion	-0.24760	0.24051	-1.029	0.30334	
home_team_nameWest Ham United	0.24191	0.21680	1.116	0.26459	
home_team_nameWolverhampton Wanderers	0.21557	0.25538	0.844	0.39867	
away_team_nameArsenal	-0.86365	0.21680	-3.984	6.95e-05	***
away_team_nameAston Villa	-0.20775	0.24059	-0.863	0.38794	
away_team_nameBrentford	-0.35033	0.41668	-0.841	0.40055	
away_team_nameBrighton & Hove Albion	-0.32877	0.24067	-1.366	0.17203	
away_team_nameBurnley	-0.18032	0.22266	-0.810	0.41809	
away_team_nameCardiff City	0.13482	0.41668	0.324	0.74629	
away_team_nameChelsea	-1.27060	0.21680	-5.861	5.12e-09	***
away_team_nameCrystal Palace	-0.41921	0.21680	-1.934	0.05326	.
away_team_nameEverton	-0.30150	0.21680	-1.391	0.16443	
away_team_nameFulham	0.21605	0.31829	0.679	0.49733	
away_team_nameHuddersfield Town	0.58754	0.31818	1.847	0.06491	.
away_team_nameHull City	0.44914	0.31840	1.411	0.15846	
away_team_nameLeeds United	-0.39180	0.31839	-1.231	0.21857	
away_team_nameLeicester City	-0.69178	0.21680	-3.191	0.00143	**
away_team_nameLiverpool	-1.39386	0.21680	-6.429	1.49e-10	***
away_team_nameManchester City	-1.88483	0.21680	-8.694	< 2e-16	***
away_team_nameManchester United	-1.00428	0.21680	-4.632	3.77e-06	***
away_team_nameMiddlesbrough	0.22389	0.41663	0.537	0.59105	
away_team_nameNewcastle United	0.03773	0.22270	0.169	0.86547	
away_team_nameNorwich City	0.62894	0.27783	2.264	0.02366	*

```
## away_team_nameQueens Park Rangers      0.74106      0.41659      1.779      0.07536 .
## away_team_nameSheffield United          0.14486      0.31832      0.455      0.64908
## away_team_nameSouthampton              -0.30115      0.21680     -1.389      0.16492
## away_team_nameStoke City                -0.06808      0.25524     -0.267      0.78969
## away_team_nameSunderland                0.03840      0.27794      0.138      0.89012
## away_team_nameSwansea City              -0.07294      0.25524     -0.286      0.77507
## away_team_nameTottenham Hotspur        -1.22754      0.21680     -5.662 1.64e-08 ***
## away_team_nameWatford                  -0.01163      0.23029     -0.050      0.95973
## away_team_nameWest Bromwich Albion     -0.14168      0.24051     -0.589      0.55587
## away_team_nameWest Ham United          -0.55705      0.21680     -2.569      0.01024 *
## away_team_nameWolverhampton Wanderers -0.60853      0.25538     -2.383      0.01724 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.655 on 2977 degrees of freedom
## Multiple R-squared:  0.2346, Adjusted R-squared:  0.2187
## F-statistic: 14.72 on 62 and 2977 DF, p-value: < 2.2e-16
```

```
data <- data%>%
  mutate(win = ifelse(home_score>away_score, 1, 0))

# and one without interactions
teams_lr = glm(data=data, formula = win ~ home_team_name + away_team_name, family='binomial')
summary(teams_lr)
```

```
##
## Call:
## glm(formula = win ~ home_team_name + away_team_name, family = "binomial",
##      data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0994  -0.9750  -0.5887   1.0367   2.2513
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.008801   0.314743  -0.028  0.977693
## home_team_nameArsenal      1.252440   0.285108   4.393  1.12e-05 ***
## home_team_nameAston Villa  -0.350083   0.325136  -1.077  0.281602
## home_team_nameBrentford     0.082698   0.543860   0.152  0.879141
## home_team_nameBrighton & Hove Albion -0.355900   0.324659  -1.096  0.272980
## home_team_nameBurnley      -0.097320   0.293490  -0.332  0.740196
## home_team_nameCardiff City  -0.177032   0.558020  -0.317  0.751054
## home_team_nameChelsea       0.946085   0.281523   3.361  0.000778 ***
## home_team_nameCrystal Palace -0.195571   0.286937  -0.682  0.495504
## home_team_nameEverton       0.418948   0.280714   1.492  0.135585
## home_team_nameFulham        -0.730944   0.465707  -1.570  0.116523
## home_team_nameHuddersfield Town -0.747296   0.465304  -1.606  0.108266
## home_team_nameHull City     -0.082715   0.421020  -0.196  0.844247
## home_team_nameLeeds United  -0.150988   0.424887  -0.355  0.722320
## home_team_nameLeicester City  0.552021   0.280383   1.969  0.048975 *
## home_team_nameLiverpool     1.326937   0.286187   4.637  3.54e-06 ***
## home_team_nameManchester City  1.729719   0.295317   5.857  4.71e-09 ***
## home_team_nameManchester United  0.920615   0.281370   3.272  0.001068 **
```

```

## home_team_nameMiddlesbrough      -0.779690    0.619268   -1.259 0.208012
## home_team_nameNewcastle United    0.091237    0.290912    0.314 0.753806
## home_team_nameNorwich City        -0.618751    0.393031   -1.574 0.115417
## home_team_nameQueens Park Rangers -0.189577    0.556887   -0.340 0.733538
## home_team_nameSheffield United     0.219975    0.410871    0.535 0.592383
## home_team_nameSouthampton         0.061982    0.283360    0.219 0.826854
## home_team_nameStoke City           0.127544    0.331913    0.384 0.700778
## home_team_nameSunderland          -0.728786    0.397284   -1.834 0.066592 .
## home_team_nameSwansea City         0.212017    0.330757    0.641 0.521519
## home_team_nameTottenham Hotspur    1.196940    0.284387    4.209 2.57e-05 ***
## home_team_nameWatford              -0.142647    0.304687   -0.468 0.639658
## home_team_nameWest Bromwich Albion -0.325225    0.323097   -1.007 0.314133
## home_team_nameWest Ham United       0.326108    0.281107    1.160 0.246015
## home_team_nameWolverhampton Wanderers 0.280475    0.329864    0.850 0.395172
## away_team_nameArsenal              -0.847819    0.280645   -3.021 0.002520 **
## away_team_nameAston Villa           0.064718    0.309698    0.209 0.834472
## away_team_nameBrentford            -0.476078    0.530949   -0.897 0.369903
## away_team_nameBrighton & Hove Albion -0.495892    0.307579   -1.612 0.106909
## away_team_nameBurnley               -0.379602    0.284891   -1.332 0.182714
## away_team_nameCardiff City          0.509583    0.561404    0.908 0.364040
## away_team_nameChelsea              -1.401203    0.291057   -4.814 1.48e-06 ***
## away_team_nameCrystal Palace        -0.599969    0.278158   -2.157 0.031011 *
## away_team_nameEverton              -0.477566    0.277956   -1.718 0.085772 .
## away_team_nameFulham                0.084513    0.410135    0.206 0.836744
## away_team_nameHuddersfield Town     0.323086    0.417520    0.774 0.439036
## away_team_nameHull City             0.365954    0.419914    0.871 0.383482
## away_team_nameLeeds United          -0.485634    0.406191   -1.196 0.231861
## away_team_nameLeicester City        -0.677129    0.279125   -2.426 0.015271 *
## away_team_nameLiverpool            -1.684769    0.299909   -5.618 1.94e-08 ***
## away_team_nameManchester City       -1.823951    0.305391   -5.973 2.34e-09 ***
## away_team_nameManchester United     -1.330719    0.289312   -4.600 4.23e-06 ***
## away_team_nameMiddlesbrough        -0.027907    0.533499   -0.052 0.958282
## away_team_nameNewcastle United     -0.102461    0.285531   -0.359 0.719712
## away_team_nameNorwich City          0.718658    0.380136    1.891 0.058688 .
## away_team_nameQueens Park Rangers   1.431203    0.680965    2.102 0.035577 *
## away_team_nameSheffield United     -0.005525    0.408636   -0.014 0.989213
## away_team_nameSouthampton          -0.498120    0.277917   -1.792 0.073080 .
## away_team_nameStoke City            -0.461473    0.326694   -1.413 0.157787
## away_team_nameSunderland            -0.113308    0.355103   -0.319 0.749662
## away_team_nameSwansea City          -0.046443    0.327558   -0.142 0.887249
## away_team_nameTottenham Hotspur    -1.241431    0.287381   -4.320 1.56e-05 ***
## away_team_nameWatford               0.108879    0.297018    0.367 0.713937
## away_team_nameWest Bromwich Albion  -0.347325    0.307262   -1.130 0.258313
## away_team_nameWest Ham United       -0.571372    0.278370   -2.053 0.040115 *
## away_team_nameWolverhampton Wanderers -0.817503    0.330690   -2.472 0.013432 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 4175.8 on 3039 degrees of freedom
## Residual deviance: 3661.6 on 2977 degrees of freedom
## AIC: 3787.6
##

```

```
## Number of Fisher Scoring iterations: 4
```

```
lm_model <- felm(data=data, formula = home_score ~
  HomePlayer_Overall_mean +
  HomePlayer_Overall_min +
  HomePlayer_Overall_max +
  HomePlayer_Overall_sd +
  HomePlayer_Height_mean + I(HomePlayer_Height_mean^2) +
  HomePlayer_Weight_mean + I(HomePlayer_Weight_mean^2) +
  home_Points_prior + away_Points_prior +
  HomePlayer_Age_mean + I(HomePlayer_Age_mean^2) |
  home_team_name + away_team_name + fixture | 0 | 0)

summary(lm_model)
```

```
##
```

```
## Call:
```

```
##      felm(formula = home_score ~ HomePlayer_Overall_mean + HomePlayer_Overall_min +      HomePlayer_Overall_max +
```

```
##
```

```
## Residuals:
```

```
##      Min      1Q  Median      3Q      Max
## -3.3514 -0.6892 -0.0005  0.5631  6.9363
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## HomePlayer_Overall_mean      0.047293   0.027896   1.695 0.090159 .
## HomePlayer_Overall_min     -0.011649   0.017930  -0.650 0.515950
## HomePlayer_Overall_max      0.007701   0.027004   0.285 0.775538
## HomePlayer_Overall_sd       0.011310   0.063335   0.179 0.858284
## HomePlayer_Height_mean      2.735895   2.498745   1.095 0.273682
## I(HomePlayer_Height_mean^2) -0.007373   0.006834  -1.079 0.280779
## HomePlayer_Weight_mean     -0.737419   0.601897  -1.225 0.220652
## I(HomePlayer_Weight_mean^2)  0.004702   0.003887   1.210 0.226567
## home_Points_prior           0.009393   0.002498   3.760 0.000174 ***
## away_Points_prior          -0.007559   0.002482  -3.046 0.002348 **
## HomePlayer_Age_mean         0.169022   0.650196   0.260 0.794923
## I(HomePlayer_Age_mean^2)    -0.003596   0.012355  -0.291 0.771040
```

```
## ---
```

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

```
##
```

```
## Residual standard error: 1.241 on 2112 degrees of freedom
```

```
## Multiple R-squared(full model): 0.3643   Adjusted R-squared: 0.08529
```

```
## Multiple R-squared(proj model): 0.01716   Adjusted R-squared: -0.4142
```

```
## F-statistic(full model):1.306 on 927 and 2112 DF, p-value: 5.541e-07
```

```
## F-statistic(proj model): 3.073 on 12 and 2112 DF, p-value: 0.0002549
```

```
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

```
lm_model <- felm(data=data, formula = home_score ~
  HomePlayer_Overall_mean +
  HomePlayer_Overall_min +
  HomePlayer_Overall_max +
  HomePlayer_Overall_sd +
  HomePlayer_Height_mean + I(HomePlayer_Height_mean^2) +
```



```

      HomePlayer_Weight_mean + I(HomePlayer_Weight_mean^2) +
      home_Points_prior + away_Points_prior+
      HomePlayer_Age_mean + I(HomePlayer_Age_mean^2) |
      home_team_name+ away_team_name + fixture| 0 | 0)

summary(lm_model)

##
## Call:
##   felm(formula = home_score ~ HomePlayer_Overall_mean + HomePlayer_Overall_min +      HomePlayer_Ov
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.3514 -0.6892 -0.0005  0.5631  6.9363
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## HomePlayer_Overall_mean      0.047293   0.027896   1.695 0.090159 .
## HomePlayer_Overall_min     -0.011649   0.017930  -0.650 0.515950
## HomePlayer_Overall_max      0.007701   0.027004   0.285 0.775538
## HomePlayer_Overall_sd      0.011310   0.063335   0.179 0.858284
## HomePlayer_Height_mean      2.735895   2.498745   1.095 0.273682
## I(HomePlayer_Height_mean^2) -0.007373   0.006834  -1.079 0.280779
## HomePlayer_Weight_mean     -0.737419   0.601897  -1.225 0.220652
## I(HomePlayer_Weight_mean^2)  0.004702   0.003887   1.210 0.226567
## home_Points_prior          0.009393   0.002498   3.760 0.000174 ***
## away_Points_prior         -0.007559   0.002482  -3.046 0.002348 **
## HomePlayer_Age_mean        0.169022   0.650196   0.260 0.794923
## I(HomePlayer_Age_mean^2)   -0.003596   0.012355  -0.291 0.771040
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.241 on 2112 degrees of freedom
## Multiple R-squared(full model): 0.3643   Adjusted R-squared: 0.08529
## Multiple R-squared(proj model): 0.01716   Adjusted R-squared: -0.4142
## F-statistic(full model):1.306 on 927 and 2112 DF, p-value: 5.541e-07
## F-statistic(proj model): 3.073 on 12 and 2112 DF, p-value: 0.0002549
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE

```

We can also use form instead of just the points prior why can this help? beacuse a teams form can change for certain periods in a season which can absolutely affect the result of a match.

```

data <- data%>%
  mutate(GD = home_score-away_score)
lm_model <- felm(data=data, formula = I(home_score-away_score) ~
      HomePlayer_Overall_mean +
      HomePlayer_Overall_min*
      HomePlayer_Overall_max +

      AwayPlayer_Overall_mean +
      AwayPlayer_Overall_min*
      AwayPlayer_Overall_max +
      home_Points_form_pw +away_Points_form_pw

```

```
| home_team_name + away_team_name + fixture | 0 | home_team_name)

summary(lm_model)
```

```
##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      HomePlayer_Overall_min
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.0472 -0.9096  0.0000  0.8995  7.7276
##
## Coefficients:
##                                     Estimate Cluster s.e. t value
## HomePlayer_Overall_mean              6.789e-02   3.242e-02   2.094
## HomePlayer_Overall_min             -2.465e-01   1.595e-01  -1.546
## HomePlayer_Overall_max             -1.827e-01   1.371e-01  -1.333
## AwayPlayer_Overall_mean            -5.510e-02   3.321e-02  -1.659
## AwayPlayer_Overall_min              1.602e-02   1.171e-01   0.137
## AwayPlayer_Overall_max              2.896e-03   1.048e-01   0.028
## home_Points_form_pw                1.311e-01   5.371e-02   2.441
## away_Points_form_pw               -1.571e-01   9.635e-02  -1.630
## HomePlayer_Overall_min:HomePlayer_Overall_max  2.642e-03   1.905e-03   1.387
## AwayPlayer_Overall_min:AwayPlayer_Overall_max -6.177e-05   1.396e-03  -0.044
##                                     Pr(>|t|)
## HomePlayer_Overall_mean              0.0445 *
## HomePlayer_Overall_min              0.1323
## HomePlayer_Overall_max              0.1923
## AwayPlayer_Overall_mean              0.1071
## AwayPlayer_Overall_min              0.8921
## AwayPlayer_Overall_max              0.9781
## home_Points_form_pw                 0.0206 *
## away_Points_form_pw                 0.1131
## HomePlayer_Overall_min:HomePlayer_Overall_max  0.1754
## AwayPlayer_Overall_min:AwayPlayer_Overall_max  0.9650
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.718 on 2114 degrees of freedom
## Multiple R-squared(full model): 0.415   Adjusted R-squared: 0.159
## Multiple R-squared(proj model): 0.01285   Adjusted R-squared: -0.4191
## F-statistic(full model, *iid*):1.621 on 925 and 2114 DF, p-value: < 2.2e-16
## F-statistic(proj model):  4.08 on 10 and 31 DF, p-value: 0.001216
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

```
data <- data%>%
  mutate(GD = home_score-away_score)
lm_model <- felm(data=data, formula = I(home_score-away_score) ~
  HomePlayer_Overall_mean +
  HomePlayer_Overall_min*
  HomePlayer_Overall_max +

  AwayPlayer_Overall_mean +
```

```

      AwayPlayer_Overall_min*
      AwayPlayer_Overall_max +
      home_GD_form + away_GD_form
    | home_team_name + away_team_name + fixture | 0 | home_team_name)

summary(lm_model)

```

```

##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      HomePlayer_Overall_min
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.8530 -0.9222  0.0000  0.9114  7.6789
##
## Coefficients:
##                                     Estimate Cluster s.e. t value
## HomePlayer_Overall_mean              6.732e-02   3.364e-02   2.001
## HomePlayer_Overall_min             -2.522e-01   1.594e-01  -1.582
## HomePlayer_Overall_max             -1.883e-01   1.356e-01  -1.389
## AwayPlayer_Overall_mean            -5.400e-02   3.341e-02  -1.616
## AwayPlayer_Overall_min              1.773e-02   1.196e-01   0.148
## AwayPlayer_Overall_max              6.699e-03   1.063e-01   0.063
## home_GD_form                       1.678e-02   8.452e-03   1.986
## away_GD_form                      -3.560e-02   9.738e-03  -3.656
## HomePlayer_Overall_min:HomePlayer_Overall_max  2.702e-03   1.900e-03   1.422
## AwayPlayer_Overall_min:AwayPlayer_Overall_max -9.878e-05   1.421e-03  -0.070
##                                     Pr(>|t|)
## HomePlayer_Overall_mean              0.05420 .
## HomePlayer_Overall_min              0.12377
## HomePlayer_Overall_max              0.17480
## AwayPlayer_Overall_mean             0.11619
## AwayPlayer_Overall_min              0.88311
## AwayPlayer_Overall_max              0.95017
## home_GD_form                       0.05597 .
## away_GD_form                       0.00094 ***
## HomePlayer_Overall_min:HomePlayer_Overall_max  0.16493
## AwayPlayer_Overall_min:AwayPlayer_Overall_max  0.94503
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.713 on 2114 degrees of freedom
## Multiple R-squared(full model): 0.4179   Adjusted R-squared: 0.1633
## Multiple R-squared(proj model): 0.0179   Adjusted R-squared: -0.4118
## F-statistic(full model, *iid*):1.641 on 925 and 2114 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.148 on 10 and 31 DF, p-value: 0.0001981
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE

```

```

data <- data%>%
  mutate(GD = home_score-away_score)
lm_model <- felm(data=data, formula = GD ~
  HomePlayer_Overall_mean+
  HomePlayer_Overall_sd +

```

```

      AwayPlayer_Overall_mean+
      AwayPlayer_Overall_sd+
      home_GD_form_pw +
      away_GD_form_pw
    | home_team_name + away_team_name + fixture | 0 | home_team_name)

summary(lm_model)

```

```

##
## Call:
##   felm(formula = GD ~ HomePlayer_Overall_mean + HomePlayer_Overall_sd +      AwayPlayer_Overall_mean
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.7324 -0.9209  0.0000  0.8991  7.6967
##
## Coefficients:
##              Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.040284      0.017987   2.240 0.032430 *
## HomePlayer_Overall_sd    0.075266      0.029016   2.594 0.014359 *
## AwayPlayer_Overall_mean -0.039145      0.024335  -1.609 0.117846
## AwayPlayer_Overall_sd   -0.009143      0.024990  -0.366 0.716959
## home_GD_form_pw         0.110400      0.051156   2.158 0.038782 *
## away_GD_form_pw        -0.249601      0.067475  -3.699 0.000836 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.713 on 2118 degrees of freedom
## Multiple R-squared(full model): 0.4169   Adjusted R-squared: 0.1634
## Multiple R-squared(proj model): 0.01619   Adjusted R-squared: -0.4116
## F-statistic(full model, *iid*):1.644 on 921 and 2118 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.756 on 6 and 31 DF, p-value: 0.0004082
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE

```

```

lm_model <- felm(data=data, formula = I(home_score-away_score) ~
      HomePlayer_Overall_mean
      +AwayPlayer_Overall_mean+
      HomePlayer_Overall_sd+
      AwayPlayer_Overall_sd+
      home_Points_form_pw +
      away_Points_form_pw
    | home_team_name + away_team_name + fixture | 0 | home_team_name)

summary(lm_model)

```

```

##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      AwayPlayer_Overall_mean
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.9142 -0.9253  0.0000  0.9045  7.7366
##

```

```
## Coefficients:
##               Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.04308      0.01822   2.364   0.0245 *
## AwayPlayer_Overall_mean -0.04090      0.02450  -1.669   0.1051
## HomePlayer_Overall_sd    0.07338      0.02975   2.467   0.0194 *
## AwayPlayer_Overall_sd   -0.01581      0.02410  -0.656   0.5167
## home_Points_form_pw      0.14074      0.05435   2.590   0.0145 *
## away_Points_form_pw     -0.16246      0.09703  -1.674   0.1041
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.718 on 2118 degrees of freedom
## Multiple R-squared(full model): 0.4136   Adjusted R-squared: 0.1585
## Multiple R-squared(proj model): 0.0105   Adjusted R-squared: -0.4198
## F-statistic(full model, *iid*): 1.622 on 921 and 2118 DF, p-value: < 2.2e-16
## F-statistic(proj model): 3.878 on 6 and 31 DF, p-value: 0.005298
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

```
lm_model <- felm(data=data, formula = I(home_score-away_score) ~
  HomePlayer_Overall_mean_ln + AwayPlayer_Overall_mean_ln+
  HomePlayer_Overall_sd+
  AwayPlayer_Overall_sd+
  home_Points_form_pw +
  away_Points_form_pw
  | home_team_name + away_team_name + fixture | 0 | home_team_name)

summary(lm_model)
```

```
##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean_ln +      AwayPlayer_Overall_r
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.9215 -0.9261  0.0000  0.9114  7.7347
##
## Coefficients:
##               Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean_ln  3.04323      1.42645   2.133   0.0409 *
## AwayPlayer_Overall_mean_ln -2.92730      1.90052  -1.540   0.1336
## HomePlayer_Overall_sd      0.07434      0.02968   2.505   0.0177 *
## AwayPlayer_Overall_sd     -0.01699      0.02492  -0.682   0.5005
## home_Points_form_pw        0.14166      0.05472   2.589   0.0145 *
## away_Points_form_pw       -0.16300      0.09701  -1.680   0.1030
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.718 on 2118 degrees of freedom
## Multiple R-squared(full model): 0.4133   Adjusted R-squared: 0.1582
## Multiple R-squared(proj model): 0.01009   Adjusted R-squared: -0.4204
## F-statistic(full model, *iid*): 1.62 on 921 and 2118 DF, p-value: < 2.2e-16
## F-statistic(proj model): 3.696 on 6 and 31 DF, p-value: 0.006919
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

```
lm_model <- felm(data=data, formula = I(home_score-away_score) ~
  HomePlayer_Overall_mean_sqrt + AwayPlayer_Overall_mean_sqrt+
  HomePlayer_Overall_sd+
  AwayPlayer_Overall_sd+
  home_GD_form_pw + home_points_to_ucl+home_points_to_rel+
  away_GD_form_pw + away_points_to_ucl+away_points_to_rel
  | home_team_name + away_team_name + fixture | 0 | home_team_name)
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
summary(lm_model)
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
##
```

```
## Call:
```

```
## felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean_sqrt + AwayPlayer_Overall
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -7.7231 -0.9212  0.0000   0.9002   7.7193
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean_sqrt  0.682708    0.320403   2.131  0.04115 *
## AwayPlayer_Overall_mean_sqrt -0.638989    0.420888  -1.518  0.13910
## HomePlayer_Overall_sd        0.070574    0.028326   2.492  0.01828 *
## AwayPlayer_Overall_sd       -0.013431    0.026014  -0.516  0.60932
## home_GD_form_pw              0.105108    0.053445   1.967  0.05823 .
## home_points_to_ucl          -0.008394    0.007041  -1.192  0.24226
## home_points_to_rel              NaN      0.000000      NaN      NaN
## away_GD_form_pw             -0.246990    0.069052  -3.577  0.00117 **
## away_points_to_ucl           0.014855    0.008846   1.679  0.10317
## away_points_to_rel          -0.010193    0.010046  -1.015  0.31815
```

```
## ---
```

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

```
##
```

```
## Residual standard error: 1.713 on 2115 degrees of freedom
```

```
## Multiple R-squared(full model): 0.4176    Adjusted R-squared: 0.1632
```

```
## Multiple R-squared(proj model): 0.01731    Adjusted R-squared: -0.412
```

```
## F-statistic(full model, *iid*):1.641 on 924 and 2115 DF, p-value: < 2.2e-16
```

```
## F-statistic(proj model): 4.726 on 10 and 31 DF, p-value: 0.0003968
```

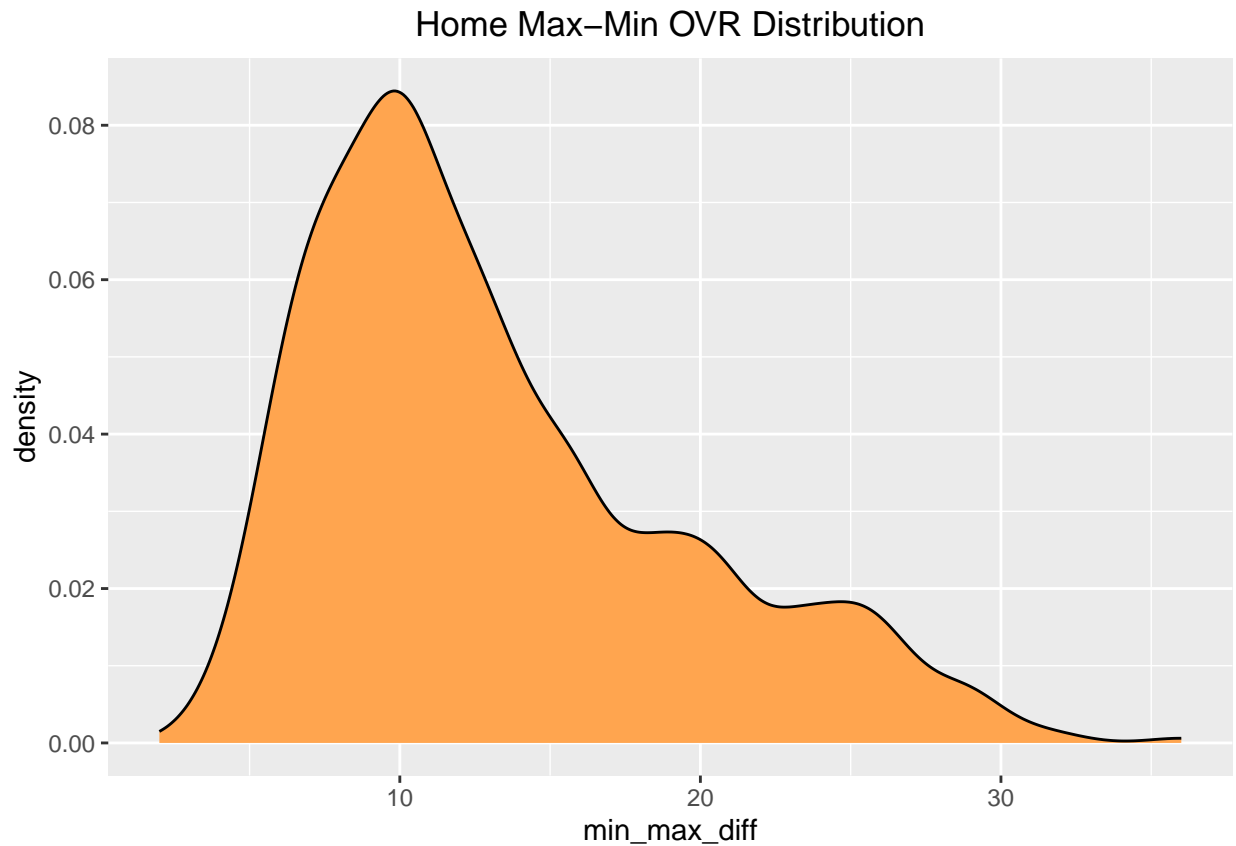
```
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

```
data$min_max_diff = data$HomePlayer_Overall_max - data$HomePlayer_Overall_min
```

```
summary(data$min_max_diff)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      2.00   9.00   12.00   13.27   17.00   36.00
```

```
ggplot(data=data, aes(x=min_max_diff)) +  
  geom_density(fill="tan1") + labs(title="Home Max-Min OVR Distribution", xlab="HomePlayer_Overall_Max-I
```



What is interesting to see, is that we can calculate derivatives for marginal utility of increasing the min and the max, and we get there are certain fields, where improving the best player is worth more than improving our worse player! this shows the increasing marginal utility of player talent (in certain areas).

An interesting result in our data shows that the Standard deviation has a positive effect on the teams performance: this can show us that there might be no diminishing returns on talent in our football team, since if there was, the standard deviation would cause harm to our team. Another proof of having

Now lets see another thing, whether or not the points prior to the match have a quadratic effect, intuition: if a team already has the championship, they won't take the match seriously, and same if the team is relegated, however there will be a similar motivation for near relegation or top 4 matches

```
lm_model <- felm(data=data, formula = I(home_score-away_score) ~  
  HomePlayer_Overall_mean + AwayPlayer_Overall_mean+  
  HomePlayer_Overall_sd+  
  AwayPlayer_Overall_sd+  
  home_GD_form_pw +  
  away_GD_form_pw  
  | home_team_name + away_team_name + fixture | 0 | home_team_name)  
  
staying_league_model <- felm(data=data, formula = I(home_score-away_score) ~  
  HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
```

```

      HomePlayer_Overall_sd+
      AwayPlayer_Overall_sd+
      +home_GD_form_pw + away_GD_form_pw+
      home_points_to_rel + away_points_to_rel
      | home_team_name + away_team_name + fixture | 0 |home_team_name)

tr_model <- felm(data=data, formula = I(home_score-away_score) ~
      HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
      HomePlayer_Overall_sd+
      AwayPlayer_Overall_sd+
      +home_GD_form_pw + away_GD_form_pw+
      home_points_to_championship + away_points_to_championship
      | home_team_name + away_team_name + fixture | 0 | home_team_name)
summary(lm_model)

```

```

##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      AwayPlayer_Overall_mean
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.7324 -0.9209  0.0000  0.8991  7.6967
##
## Coefficients:
##              Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.040284      0.017987   2.240 0.032430 *
## AwayPlayer_Overall_mean -0.039145      0.024335  -1.609 0.117846
## HomePlayer_Overall_sd    0.075266      0.029016   2.594 0.014359 *
## AwayPlayer_Overall_sd   -0.009143      0.024990  -0.366 0.716959
## home_GD_form_pw          0.110400      0.051156   2.158 0.038782 *
## away_GD_form_pw         -0.249601      0.067475  -3.699 0.000836 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.713 on 2118 degrees of freedom
## Multiple R-squared(full model): 0.4169   Adjusted R-squared: 0.1634
## Multiple R-squared(proj model): 0.01619   Adjusted R-squared: -0.4116
## F-statistic(full model, *iid*):1.644 on 921 and 2118 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.756 on 6 and 31 DF, p-value: 0.0004082
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE

```

```
summary(tr_model)
```

```

##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      AwayPlayer_Overall_mean
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.7301 -0.9177  0.0000  0.8986  7.7115
##
## Coefficients:

```



```
##               Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean    0.040841    0.017927  2.278 0.02976 *
## AwayPlayer_Overall_mean   -0.038349    0.024032 -1.596 0.12069
## HomePlayer_Overall_sd      0.072465    0.029001  2.499 0.01797 *
## AwayPlayer_Overall_sd     -0.012026    0.025582 -0.470 0.64158
## home_GD_form_pw           0.105807    0.053165  1.990 0.05546 .
## away_GD_form_pw           -0.244987    0.068511 -3.576 0.00117 **
## home_points_to_championship -0.006089    0.007600 -0.801 0.42913
## away_points_to_championship 0.007350    0.007899  0.931 0.35927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.713 on 2116 degrees of freedom
## Multiple R-squared(full model): 0.4173    Adjusted R-squared: 0.1631
## Multiple R-squared(proj model): 0.01681    Adjusted R-squared: -0.4121
## F-statistic(full model, *iid*):1.642 on 923 and 2116 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.695 on 8 and 31 DF, p-value: 0.0001746
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

```
summary(staying_league_model)
```

```
##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      AwayPlayer_Overall_mean
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.7453 -0.9212  0.0000  0.8982  7.7024
##
## Coefficients:
##               Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.040505    0.018194  2.226 0.03340 *
## AwayPlayer_Overall_mean -0.039248    0.024077 -1.630 0.11319
## HomePlayer_Overall_sd    0.075911    0.028960  2.621 0.01345 *
## AwayPlayer_Overall_sd   -0.008613    0.025953 -0.332 0.74223
## home_GD_form_pw          0.103058    0.053156  1.939 0.06168 .
## away_GD_form_pw         -0.243739    0.068847 -3.540 0.00129 **
## home_points_to_rel       -0.006879    0.007202 -0.955 0.34689
## away_points_to_rel        0.006501    0.008993  0.723 0.47517
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.713 on 2116 degrees of freedom
## Multiple R-squared(full model): 0.4172    Adjusted R-squared: 0.1629
## Multiple R-squared(proj model): 0.01656    Adjusted R-squared: -0.4124
## F-statistic(full model, *iid*):1.641 on 923 and 2116 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.571 on 8 and 31 DF, p-value: 0.0002084
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

Now lets keep in mind:

```

sqrt_model <- felm(data=data, formula = I(home_score-away_score) ~
  HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
  HomePlayer_Overall_mean_sqrt+AwayPlayer_Overall_mean_sqrt+
  HomePlayer_Overall_sd+
  AwayPlayer_Overall_sd+
  home_Points_prior +
  away_Points_form
  | home_team_name + away_team_name + fixture| 0 | home_team_name)

ln_model <- felm(data=data, formula = I(home_score-away_score) ~
  HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
  HomePlayer_Overall_mean_ln+AwayPlayer_Overall_mean_ln+
  HomePlayer_Overall_sd+
  AwayPlayer_Overall_sd+
  home_GD_form_pw +
  away_GD_form_pw
  | home_team_name + away_team_name + fixture | 0 | home_team_name)

both_model <- felm(data=data, formula = I(home_score-away_score) ~
  HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
  HomePlayer_Overall_mean_ln+AwayPlayer_Overall_mean_ln+
  HomePlayer_Overall_mean_sqrt+AwayPlayer_Overall_mean_sqrt+
  HomePlayer_Overall_sd+
  AwayPlayer_Overall_sd+
  home_GD_form_pw +
  away_GD_form_pw
  | home_team_name + away_team_name + fixture| 0 | home_team_name)
summary(sqrt_model)

```

```

##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      AwayPlayer_Overall_mean
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.8689 -0.9244  0.0000  0.9117  7.7798
##
## Coefficients:
##              Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean      2.260397      0.595171   3.798 0.000638 ***
## AwayPlayer_Overall_mean     -1.396959      0.910634  -1.534 0.135161
## HomePlayer_Overall_mean_sqrt -38.885888     10.529523  -3.693 0.000851 ***
## AwayPlayer_Overall_mean_sqrt  23.826927     16.003324   1.489 0.146624
## HomePlayer_Overall_sd        0.001796      0.028166   0.064 0.949558
## AwayPlayer_Overall_sd        0.019953      0.040415   0.494 0.624994
## home_Points_prior            0.006168      0.002031   3.037 0.004814 **
## away_Points_form            -0.023355      0.013044  -1.791 0.083133 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.714 on 2116 degrees of freedom
## Multiple R-squared(full model): 0.4167    Adjusted R-squared: 0.1623

```

```
## Multiple R-squared(proj model): 0.01582    Adjusted R-squared: -0.4135
## F-statistic(full model, *iid*):1.638 on 923 and 2116 DF, p-value: < 2.2e-16
## F-statistic(proj model): 8.989 on 8 and 31 DF, p-value: 2.752e-06
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

```
summary(ln_model)
```

```
##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      AwayPlayer_Overall_mean
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.7371 -0.9036  0.0000  0.8989  7.7171
##
## Coefficients:
##              Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean      1.118762    0.284353   3.934 0.000438 ***
## AwayPlayer_Overall_mean     -0.626410    0.445476  -1.406 0.169622
## HomePlayer_Overall_mean_ln -82.828361   22.195576  -3.732 0.000765 ***
## AwayPlayer_Overall_mean_ln  45.109112   34.401946   1.311 0.199407
## HomePlayer_Overall_sd       0.009019    0.028442   0.317 0.753301
## AwayPlayer_Overall_sd       0.026739    0.040615   0.658 0.515173
## home_GD_form_pw            0.099388    0.050095   1.984 0.056170 .
## away_GD_form_pw            -0.241276    0.065401  -3.689 0.000859 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.71 on 2116 degrees of freedom
## Multiple R-squared(full model): 0.4194    Adjusted R-squared: 0.1662
## Multiple R-squared(proj model): 0.02042    Adjusted R-squared: -0.4069
## F-statistic(full model, *iid*):1.656 on 923 and 2116 DF, p-value: < 2.2e-16
## F-statistic(proj model): 10.33 on 8 and 31 DF, p-value: 6.604e-07
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

```
summary(both_model)
```

```
##
## Call:
##   felm(formula = I(home_score - away_score) ~ HomePlayer_Overall_mean +      AwayPlayer_Overall_mean
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.6922 -0.9044  0.0000  0.9049  7.6685
##
## Coefficients:
##              Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean     -8.921e+00    8.667e+00  -1.029 0.31129
## AwayPlayer_Overall_mean       7.407e+00    1.175e+01   0.631 0.53293
## HomePlayer_Overall_mean_ln  -8.366e+02    6.485e+02  -1.290 0.20656
## AwayPlayer_Overall_mean_ln   6.499e+02    8.626e+02   0.753 0.45686
## HomePlayer_Overall_mean_sqrt  3.482e+02    2.999e+02   1.161 0.25446
## AwayPlayer_Overall_mean_sqrt -2.790e+02    4.027e+02  -0.693 0.49347
```

```
## HomePlayer_Overall_sd      -3.084e-03    3.329e-02  -0.093  0.92677
## AwayPlayer_Overall_sd      3.766e-02    4.258e-02   0.884  0.38330
## home_GD_form_pw            9.850e-02    5.004e-02   1.968  0.05802 .
## away_GD_form_pw           -2.384e-01    6.658e-02  -3.580  0.00116 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.71 on 2114 degrees of freedom
## Multiple R-squared(full model): 0.4198    Adjusted R-squared: 0.1659
## Multiple R-squared(proj model): 0.02103    Adjusted R-squared: -0.4073
## F-statistic(full model, *iid*):1.654 on 925 and 2114 DF, p-value: < 2.2e-16
## F-statistic(proj model): 8.588 on 10 and 31 DF, p-value: 1.611e-06
## *** Standard errors may be too high due to more than 2 groups and exactDOF=FALSE
```

Now the weird part is that we see teams with a home “disadvantage”, this is odd, this might be due to them being biased compared to the causal effect, and the fact that they might be affected from stuff like budget for example.

```
hg_reg = glm(data=data, formula = home_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
             HomePlayer_Overall_sd+
             AwayPlayer_Overall_sd+
             +home_GD_form_pw + away_GD_form_pw + home_team_name + away_team_name, family='poisson')

ag_reg = glm(data=data, formula = away_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
             HomePlayer_Overall_sd+
             AwayPlayer_Overall_sd+
             +home_Points_prior + away_GD_form_pw + home_team_name + away_team_name, family='poisson')

summary(ag_reg)
```

```
##
## Call:
## glm(formula = away_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
##      HomePlayer_Overall_sd + AwayPlayer_Overall_sd + +home_Points_prior +
##      away_GD_form_pw + home_team_name + away_team_name, family = "poisson",
##      data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3355  -1.2091  -0.1062   0.5528   3.9625
##
## Coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.472926    0.962271  -1.531 0.125849
## HomePlayer_Overall_mean    0.001866    0.010475   0.178 0.858599
## AwayPlayer_Overall_mean    0.023564    0.009947   2.369 0.017836 *
## HomePlayer_Overall_sd   -0.013438    0.011994  -1.120 0.262565
## AwayPlayer_Overall_sd    0.013250    0.011478   1.154 0.248313
## home_Points_prior   -0.001963    0.001090  -1.800 0.071797 .
## away_GD_form_pw     0.070619    0.026886   2.627 0.008624 **
## home_team_nameArsenal  -0.494077    0.137463  -3.594 0.000325 ***
## home_team_nameAston Villa -0.025562    0.117326  -0.218 0.827526
## home_team_nameBrentford -0.429480    0.234891  -1.828 0.067487 .
## home_team_nameBrighton & Hove Albion -0.204312    0.122102  -1.673 0.094270 .
```

## home_team_nameBurnley	-0.247796	0.113504	-2.183	0.029025	*
## home_team_nameCardiff City	0.193685	0.183735	1.054	0.291812	
## home_team_nameChelsea	-0.473884	0.147094	-3.222	0.001275	**
## home_team_nameCrystal Palace	-0.210611	0.110259	-1.910	0.056114	.
## home_team_nameEverton	-0.244956	0.119270	-2.054	0.039996	*
## home_team_nameFulham	0.036347	0.151216	0.240	0.810048	
## home_team_nameHuddersfield Town	-0.069810	0.159531	-0.438	0.661679	
## home_team_nameHull City	0.015822	0.155485	0.102	0.918950	
## home_team_nameLeeds United	-0.029472	0.154548	-0.191	0.848762	
## home_team_nameLeicester City	-0.268575	0.114996	-2.336	0.019516	*
## home_team_nameLiverpool	-0.591545	0.148968	-3.971	7.16e-05	***
## home_team_nameManchester City	-0.583960	0.157649	-3.704	0.000212	***
## home_team_nameManchester United	-0.508660	0.145173	-3.504	0.000459	***
## home_team_nameMiddlesbrough	-0.285003	0.224711	-1.268	0.204688	
## home_team_nameNewcastle United	-0.192599	0.112473	-1.712	0.086822	.
## home_team_nameNorwich City	0.162869	0.127648	1.276	0.201984	
## home_team_nameQueens Park Rangers	-0.168625	0.221670	-0.761	0.446834	
## home_team_nameSheffield United	-0.371972	0.175723	-2.117	0.034276	*
## home_team_nameSouthampton	-0.196477	0.110368	-1.780	0.075045	.
## home_team_nameStoke City	-0.151731	0.132111	-1.149	0.250760	
## home_team_nameSunderland	-0.083943	0.139215	-0.603	0.546526	
## home_team_nameSwansea City	-0.152106	0.131251	-1.159	0.246501	
## home_team_nameTottenham Hotspur	-0.507375	0.140480	-3.612	0.000304	***
## home_team_nameWatford	-0.015086	0.112041	-0.135	0.892892	
## home_team_nameWest Bromwich Albion	-0.039258	0.118521	-0.331	0.740468	
## home_team_nameWest Ham United	-0.114463	0.111259	-1.029	0.303574	
## home_team_nameWolverhampton Wanderers	-0.282948	0.139200	-2.033	0.042085	*
## away_team_nameArsenal	0.092199	0.136439	0.676	0.499201	
## away_team_nameAston Villa	-0.117900	0.142782	-0.826	0.408953	
## away_team_nameBrentford	0.224447	0.220691	1.017	0.309145	
## away_team_nameBrighton & Hove Albion	-0.215209	0.145898	-1.475	0.140196	
## away_team_nameBurnley	-0.194909	0.133364	-1.461	0.143881	
## away_team_nameCardiff City	-0.404522	0.295197	-1.370	0.170579	
## away_team_nameChelsea	0.143678	0.144025	0.998	0.318479	
## away_team_nameCrystal Palace	0.026872	0.123822	0.217	0.828191	
## away_team_nameEverton	-0.155583	0.133077	-1.169	0.242357	
## away_team_nameFulham	-0.316200	0.208778	-1.515	0.129893	
## away_team_nameHuddersfield Town	-0.504875	0.228602	-2.209	0.027207	*
## away_team_nameHull City	-0.541391	0.231282	-2.341	0.019241	*
## away_team_nameLeeds United	0.330400	0.165000	2.002	0.045239	*
## away_team_nameLeicester City	0.170496	0.122262	1.395	0.163160	
## away_team_nameLiverpool	0.208112	0.142225	1.463	0.143396	
## away_team_nameManchester City	0.287095	0.149266	1.923	0.054432	.
## away_team_nameManchester United	-0.000161	0.143782	-0.001	0.999106	
## away_team_nameMiddlesbrough	-0.754113	0.331448	-2.275	0.022893	*
## away_team_nameNewcastle United	-0.239701	0.134640	-1.780	0.075025	.
## away_team_nameNorwich City	-0.619443	0.205671	-3.012	0.002597	**
## away_team_nameQueens Park Rangers	-0.066213	0.250845	-0.264	0.791812	
## away_team_nameSheffield United	-0.572104	0.230782	-2.479	0.013176	*
## away_team_nameSouthampton	-0.068568	0.126163	-0.543	0.586793	
## away_team_nameStoke City	-0.285085	0.158685	-1.797	0.072406	.
## away_team_nameSunderland	-0.163157	0.169399	-0.963	0.335470	
## away_team_nameSwansea City	-0.218558	0.156117	-1.400	0.161525	
## away_team_nameTottenham Hotspur	0.171882	0.137542	1.250	0.211420	

```
## away_team_nameWatford -0.163097 0.137704 -1.184 0.236252
## away_team_nameWest Bromwich Albion -0.347983 0.152757 -2.278 0.022726 *
## away_team_nameWest Ham United 0.087383 0.123953 0.705 0.480831
## away_team_nameWolverhampton Wanderers -0.208041 0.155422 -1.339 0.180715
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 3966.0 on 3039 degrees of freedom
## Residual deviance: 3416.8 on 2971 degrees of freedom
## AIC: 8491.5
##
## Number of Fisher Scoring iterations: 5
```

```
summary(hg_reg)
```

```
##
## Call:
## glm(formula = home_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
##      HomePlayer_Overall_sd + AwayPlayer_Overall_sd + +home_GD_form_pw +
##      away_GD_form_pw + home_team_name + away_team_name, family = "poisson",
##      data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6001  -1.0479  -0.1213   0.5494   3.9348
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.049740   0.859029  -1.222 0.221704
## HomePlayer_Overall_mean    0.027563   0.009045   3.048 0.002308 **
## AwayPlayer_Overall_mean   -0.006608   0.009286  -0.712 0.476739
## HomePlayer_Overall_sd     0.031701   0.009813   3.231 0.001235 **
## AwayPlayer_Overall_sd    -0.004933   0.010775  -0.458 0.647053
## home_GD_form_pw     0.027049   0.024010   1.127 0.259920
## away_GD_form_pw    -0.086119   0.024841  -3.467 0.000527 ***
## home_team_nameArsenal     0.076855   0.120543   0.638 0.523754
## home_team_nameAston Villa -0.233617   0.128289  -1.821 0.068603 .
## home_team_nameBrentford  -0.142746   0.231271  -0.617 0.537089
## home_team_nameBrighton & Hove Albion -0.303159   0.130793  -2.318 0.020457 *
## home_team_nameBurnley    -0.358171   0.121561  -2.946 0.003215 **
## home_team_nameCardiff City -0.192817   0.235808  -0.818 0.413537
## home_team_nameChelsea    -0.022410   0.129303  -0.173 0.862405
## home_team_nameCrystal Palace -0.276711   0.115363  -2.399 0.016457 *
## home_team_nameEverton    -0.082803   0.114609  -0.722 0.469997
## home_team_nameFulham     -0.572798   0.199909  -2.865 0.004166 **
## home_team_nameHuddersfield Town -0.732131   0.215550  -3.397 0.000682 ***
## home_team_nameHull City  -0.148742   0.170436  -0.873 0.382819
## home_team_nameLeeds United -0.130008   0.169735  -0.766 0.443708
## home_team_nameLeicester City  0.011881   0.110345   0.108 0.914254
## home_team_nameLiverpool   0.150134   0.126796   1.184 0.236389
## home_team_nameManchester City  0.267139   0.132473   2.017 0.043741 *
## home_team_nameManchester United -0.063795   0.127924  -0.499 0.617995
```

```

## home_team_nameMiddlesbrough      -0.499035    0.257696   -1.937  0.052804 .
## home_team_nameNewcastle United    -0.118229    0.114898   -1.029  0.303482 .
## home_team_nameNorwich City        -0.303283    0.159062   -1.907  0.056560 .
## home_team_nameQueens Park Rangers -0.170302    0.226778   -0.751  0.452675 .
## home_team_nameSheffield United    -0.402245    0.188118   -2.138  0.032495 *
## home_team_nameSouthampton         -0.089632    0.111118   -0.807  0.419874 .
## home_team_nameStoke City           -0.203294    0.134697   -1.509  0.131230 .
## home_team_nameSunderland           -0.420198    0.160456   -2.619  0.008825 **
## home_team_nameSwansea City         -0.262182    0.136861   -1.916  0.055406 .
## home_team_nameTottenham Hotspur    0.040249    0.122648    0.328  0.742783 .
## home_team_nameWatford              -0.210098    0.121829   -1.725  0.084613 .
## home_team_nameWest Bromwich Albion -0.300316    0.130082   -2.309  0.020963 *
## home_team_nameWest Ham United      -0.053857    0.111213   -0.484  0.628196 .
## home_team_nameWolverhampton Wanderers -0.279753    0.137679   -2.032  0.042161 *
## away_team_nameArsenal              -0.183670    0.119472   -1.537  0.124208 .
## away_team_nameAston Villa          -0.158882    0.110069   -1.443  0.148887 .
## away_team_nameBrentford            -0.080439    0.185898   -0.433  0.665229 .
## away_team_nameBrighton & Hove Albion -0.333379    0.114576   -2.910  0.003618 **
## away_team_nameBurnley              -0.212755    0.101723   -2.092  0.036481 *
## away_team_nameCardiff City         -0.220346    0.195564   -1.127  0.259861 .
## away_team_nameChelsea              -0.344091    0.132753   -2.592  0.009543 **
## away_team_nameCrystal Palace        -0.183614    0.099320   -1.849  0.064501 .
## away_team_nameEverton              -0.129312    0.107577   -1.202  0.229348 .
## away_team_nameFulham               -0.070632    0.142780   -0.495  0.620819 .
## away_team_nameHuddersfield Town     0.000905    0.138880    0.007  0.994801 .
## away_team_nameHull City             -0.011734    0.140546   -0.083  0.933461 .
## away_team_nameLeeds United          0.002190    0.138239    0.016  0.987363 .
## away_team_nameLeicester City        -0.150905    0.102207   -1.476  0.139817 .
## away_team_nameLiverpool            -0.334548    0.131239   -2.549  0.010798 *
## away_team_nameManchester City       -0.525767    0.146142   -3.598  0.000321 ***
## away_team_nameManchester United     -0.326446    0.129911   -2.513  0.011976 *
## away_team_nameMiddlesbrough         -0.223330    0.197546   -1.131  0.258256 .
## away_team_nameNewcastle United      -0.092682    0.099655   -0.930  0.352356 .
## away_team_nameNorwich City          -0.012885    0.120563   -0.107  0.914886 .
## away_team_nameQueens Park Rangers   0.330219    0.162519    2.032  0.042165 *
## away_team_nameSheffield United      -0.230708    0.149472   -1.543  0.122712 .
## away_team_nameSouthampton           -0.163373    0.099602   -1.640  0.100952 .
## away_team_nameStoke City            -0.125795    0.119151   -1.056  0.291076 .
## away_team_nameSunderland            -0.054113    0.124002   -0.436  0.662553 .
## away_team_nameSwansea City          -0.115154    0.117765   -0.978  0.328162 .
## away_team_nameTottenham Hotspur     -0.312437    0.125730   -2.485  0.012956 *
## away_team_nameWatford               -0.095075    0.103270   -0.921  0.357236 .
## away_team_nameWest Bromwich Albion  -0.267645    0.113356   -2.361  0.018221 *
## away_team_nameWest Ham United       -0.169775    0.102387   -1.658  0.097283 .
## away_team_nameWolverhampton Wanderers -0.434379    0.133773   -3.247  0.001166 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 3842.7 on 3039 degrees of freedom
## Residual deviance: 3280.2 on 2971 degrees of freedom
## AIC: 9159.2
##

```

```
## Number of Fisher Scoring iterations: 5
```

```
# Stepwise selection for hg_reg model  
step_hg_reg <- stepAIC(hg_reg, direction = "both")
```

```
## Start: AIC=9159.25
```

```
## home_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +  
##   HomePlayer_Overall_sd + AwayPlayer_Overall_sd + +home_GD_form_pw +  
##   away_GD_form_pw + home_team_name + away_team_name  
##
```

	Df	Deviance	AIC
## - away_team_name	31	3327.9	9144.9
## - AwayPlayer_Overall_sd	1	3280.4	9157.5
## - AwayPlayer_Overall_mean	1	3280.7	9157.8
## - home_GD_form_pw	1	3281.5	9158.5
## <none>		3280.2	9159.2
## - HomePlayer_Overall_mean	1	3289.6	9166.6
## - HomePlayer_Overall_sd	1	3290.5	9167.6
## - away_GD_form_pw	1	3292.2	9169.2
## - home_team_name	31	3355.1	9172.2

```
##
```

```
## Step: AIC=9144.92
```

```
## home_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +  
##   HomePlayer_Overall_sd + AwayPlayer_Overall_sd + home_GD_form_pw +  
##   away_GD_form_pw + home_team_name  
##
```

	Df	Deviance	AIC
## - home_GD_form_pw	1	3328.7	9143.7
## <none>		3327.9	9144.9
## - AwayPlayer_Overall_sd	1	3331.2	9146.2
## - HomePlayer_Overall_mean	1	3338.1	9153.2
## - HomePlayer_Overall_sd	1	3338.3	9153.4
## + away_team_name	31	3280.2	9159.2
## - home_team_name	31	3404.4	9159.4
## - away_GD_form_pw	1	3350.3	9165.3
## - AwayPlayer_Overall_mean	1	3356.7	9171.7

```
##
```

```
## Step: AIC=9143.73
```

```
## home_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +  
##   HomePlayer_Overall_sd + AwayPlayer_Overall_sd + away_GD_form_pw +  
##   home_team_name  
##
```

```
##
```

	Df	Deviance	AIC
## <none>		3328.7	9143.7
## + home_GD_form_pw	1	3327.9	9144.9
## - AwayPlayer_Overall_sd	1	3332.0	9145.0
## - HomePlayer_Overall_mean	1	3339.5	9152.6
## - HomePlayer_Overall_sd	1	3339.6	9152.6
## + away_team_name	31	3281.5	9158.5
## - away_GD_form_pw	1	3351.0	9164.0
## - home_team_name	31	3413.5	9166.5
## - AwayPlayer_Overall_mean	1	3357.3	9170.3


```
# Stepwise selection for ag_reg model
step_ag_reg <- stepAIC(ag_reg, direction = "both")
```

```
## Start: AIC=8491.5
## away_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
##   HomePlayer_Overall_sd + AwayPlayer_Overall_sd + +home_Points_prior +
##   away_GD_form_pw + home_team_name + away_team_name
##
##           Df Deviance    AIC
## - home_team_name      31   3472.8 8485.5
## - HomePlayer_Overall_mean  1   3416.8 8489.5
## - HomePlayer_Overall_sd   1   3418.0 8490.8
## - AwayPlayer_Overall_sd   1   3418.1 8490.8
## <none>                  3416.8 8491.5
## - home_Points_prior      1   3420.0 8492.8
## - AwayPlayer_Overall_mean  1   3422.4 8495.2
## - away_GD_form_pw        1   3423.7 8496.4
## - away_team_name         31   3508.7 8521.4
##
## Step: AIC=8485.54
## away_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
##   HomePlayer_Overall_sd + AwayPlayer_Overall_sd + home_Points_prior +
##   away_GD_form_pw + away_team_name
##
##           Df Deviance    AIC
## - AwayPlayer_Overall_sd   1   3474.6 8485.4
## <none>                  3472.8 8485.5
## - away_GD_form_pw        1   3480.2 8490.9
## + home_team_name         31   3416.8 8491.5
## - home_Points_prior      1   3481.6 8492.3
## - HomePlayer_Overall_sd   1   3483.2 8493.9
## - AwayPlayer_Overall_mean  1   3484.5 8495.2
## - away_team_name         31   3561.9 8512.6
## - HomePlayer_Overall_mean  1   3534.0 8544.8
##
## Step: AIC=8485.39
## away_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
##   HomePlayer_Overall_sd + home_Points_prior + away_GD_form_pw +
##   away_team_name
##
##           Df Deviance    AIC
## <none>                  3474.6 8485.4
## + AwayPlayer_Overall_sd   1   3472.8 8485.5
## + home_team_name         31   3418.1 8490.8
## - away_GD_form_pw        1   3482.3 8491.1
## - home_Points_prior      1   3482.6 8491.4
## - AwayPlayer_Overall_mean  1   3484.5 8493.3
## - HomePlayer_Overall_sd   1   3484.9 8493.7
## - away_team_name         31   3564.7 8513.4
## - HomePlayer_Overall_mean  1   3535.8 8544.6
```

```
# Summaries of the models after stepwise selection
summary(step_hg_reg)
```

```
##
## Call:
## glm(formula = home_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
##      HomePlayer_Overall_sd + AwayPlayer_Overall_sd + away_GD_form_pw +
##      home_team_name, family = "poisson", data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5347  -1.0422  -0.1189   0.5656   4.1045
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.283099   0.714878   0.396 0.692098
## HomePlayer_Overall_mean      0.028125   0.008583   3.277 0.001050 **
## AwayPlayer_Overall_mean     -0.026629   0.004995  -5.331 9.76e-08 ***
## HomePlayer_Overall_sd       0.031901   0.009613   3.318 0.000905 ***
## AwayPlayer_Overall_sd     -0.016187   0.008943  -1.810 0.070296 .
## away_GD_form_pw     -0.111153   0.023520  -4.726 2.29e-06 ***
## home_team_nameArsenal       0.094205   0.118341   0.796 0.426006
## home_team_nameAston Villa  -0.227143   0.127967  -1.775 0.075895 .
## home_team_nameBrentford    -0.123517   0.230521  -0.536 0.592086
## home_team_nameBrighton & Hove Albion -0.287684   0.130499  -2.204 0.027490 *
## home_team_nameBurnley     -0.349398   0.121378  -2.879 0.003995 **
## home_team_nameCardiff City -0.189798   0.235149  -0.807 0.419585
## home_team_nameChelsea       0.005823   0.126084   0.046 0.963167
## home_team_nameCrystal Palace -0.266578   0.115049  -2.317 0.020499 *
## home_team_nameEverton     -0.075908   0.113685  -0.668 0.504323
## home_team_nameFulham      -0.586450   0.199547  -2.939 0.003294 **
## home_team_nameHuddersfield Town -0.736534   0.215134  -3.424 0.000618 ***
## home_team_nameHull City    -0.148328   0.170069  -0.872 0.383119
## home_team_nameLeeds United -0.144553   0.169302  -0.854 0.393206
## home_team_nameLeicester City  0.029325   0.109458   0.268 0.788765
## home_team_nameLiverpool     0.181817   0.123280   1.475 0.140258
## home_team_nameManchester City  0.312325   0.127057   2.458 0.013965 *
## home_team_nameManchester United -0.041253   0.125116  -0.330 0.741613
## home_team_nameMiddlesbrough -0.470095   0.257363  -1.827 0.067762 .
## home_team_nameNewcastle United -0.122218   0.114610  -1.066 0.286250
## home_team_nameNorwich City  -0.315226   0.158470  -1.989 0.046680 *
## home_team_nameQueens Park Rangers -0.213902   0.225826  -0.947 0.343539
## home_team_nameSheffield United -0.411093   0.187762  -2.189 0.028565 *
## home_team_nameSouthampton  -0.079599   0.110736  -0.719 0.472252
## home_team_nameStoke City    -0.198660   0.133809  -1.485 0.137634
## home_team_nameSunderland    -0.417140   0.159969  -2.608 0.009117 **
## home_team_nameSwansea City  -0.252906   0.136285  -1.856 0.063493 .
## home_team_nameTottenham Hotspur  0.066157   0.119980   0.551 0.581359
## home_team_nameWatford      -0.206496   0.121623  -1.698 0.089537 .
## home_team_nameWest Bromwich Albion -0.286770   0.129726  -2.211 0.027065 *
## home_team_nameWest Ham United -0.042957   0.110623  -0.388 0.697783
## home_team_nameWolverhampton Wanderers -0.264588   0.137268  -1.928 0.053913 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
```

```
## Null deviance: 3842.7 on 3039 degrees of freedom
## Residual deviance: 3328.7 on 3003 degrees of freedom
## AIC: 9143.7
##
## Number of Fisher Scoring iterations: 5
```

```
summary(step_ag_reg)
```

```
##
## Call:
## glm(formula = away_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
## HomePlayer_Overall_sd + home_Points_prior + away_GD_form_pw +
## away_team_name, family = "poisson", data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4443  -1.2363  -0.1062   0.5716   3.8461
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.280302   0.726747   1.762  0.07812 .
## HomePlayer_Overall_mean -0.039414   0.005089  -7.745 9.57e-15 ***
## AwayPlayer_Overall_mean  0.028074   0.008967   3.131  0.00174 **
## HomePlayer_Overall_sd  -0.032608   0.010239  -3.185  0.00145 **
## home_Points_prior  -0.002918   0.001036  -2.817  0.00485 **
## away_GD_form_pw      0.074025   0.026727   2.770  0.00561 **
## away_team_nameArsenal    0.080719   0.129034   0.626  0.53160
## away_team_nameAston Villa -0.146524   0.142087  -1.031  0.30243
## away_team_nameBrentford   0.245734   0.219465   1.120  0.26285
## away_team_nameBrighton & Hove Albion -0.209099   0.145629  -1.436  0.15105
## away_team_nameBurnley    -0.203069   0.133186  -1.525  0.12733
## away_team_nameCardiff City -0.395323   0.294397  -1.343  0.17933
## away_team_nameChelsea     0.115790   0.135611   0.854  0.39319
## away_team_nameCrystal Palace  0.015014   0.123445   0.122  0.90320
## away_team_nameEverton    -0.175475   0.129516  -1.355  0.17546
## away_team_nameFulham     -0.296267   0.207378  -1.429  0.15311
## away_team_nameHuddersfield Town -0.449373   0.227459  -1.976  0.04820 *
## away_team_nameHull City  -0.574339   0.230931  -2.487  0.01288 *
## away_team_nameLeeds United  0.320968   0.164559   1.950  0.05112 .
## away_team_nameLeicester City  0.162323   0.120773   1.344  0.17894
## away_team_nameLiverpool   0.177763   0.135650   1.310  0.19004
## away_team_nameManchester City  0.254581   0.140812   1.808  0.07061 .
## away_team_nameManchester United -0.015857   0.135738  -0.117  0.90700
## away_team_nameMiddlesbrough -0.730072   0.330980  -2.206  0.02740 *
## away_team_nameNewcastle United -0.248401   0.134255  -1.850  0.06428 .
## away_team_nameNorwich City -0.628900   0.205266  -3.064  0.00219 **
## away_team_nameQueens Park Rangers -0.173197   0.249648  -0.694  0.48783
## away_team_nameSheffield United -0.548339   0.230435  -2.380  0.01733 *
## away_team_nameSouthampton -0.081077   0.125644  -0.645  0.51874
## away_team_nameStoke City  -0.314071   0.157120  -1.999  0.04562 *
## away_team_nameSunderland -0.202119   0.168671  -1.198  0.23080
## away_team_nameSwansea City -0.232667   0.154466  -1.506  0.13200
## away_team_nameTottenham Hotspur  0.154749   0.130749   1.184  0.23659
## away_team_nameWatford    -0.180975   0.137490  -1.316  0.18808
```

```
## away_team_nameWest Bromwich Albion    -0.362063    0.151874   -2.384    0.01713 *
## away_team_nameWest Ham United          0.065997    0.122638    0.538    0.59048
## away_team_nameWolverhampton Wanderers -0.181564    0.152196   -1.193    0.23289
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 3966.0  on 3039  degrees of freedom
## Residual deviance: 3474.6  on 3003  degrees of freedom
## AIC: 8485.4
##
## Number of Fisher Scoring iterations: 5
```

```
hg_reg = glm(data=data, formula = home_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
             HomePlayer_Overall_sd+
             AwayPlayer_Overall_sd+
             +home_GD_form_pw + away_GD_form_pw, family='poisson')

ag_reg = glm(data=data, formula = away_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
             HomePlayer_Overall_sd+
             AwayPlayer_Overall_sd+
             +home_Points_prior + away_GD_form_pw, family='poisson')

summary(ag_reg)
```

```
##
## Call:
## glm(formula = away_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
##      HomePlayer_Overall_sd + AwayPlayer_Overall_sd + +home_Points_prior +
##      away_GD_form_pw, family = "poisson", data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3438  -1.2803  -0.1147   0.5681   4.1171
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.941619   0.552701  -1.704 0.088443 .
## HomePlayer_Overall_mean -0.041148   0.005021  -8.196 2.49e-16 ***
## AwayPlayer_Overall_mean  0.057045   0.005312  10.739 < 2e-16 ***
## HomePlayer_Overall_sd  -0.034467   0.010183  -3.385 0.000712 ***
## AwayPlayer_Overall_sd   0.016572   0.009858   1.681 0.092762 .
## home_Points_prior    -0.002731   0.001037  -2.632 0.008483 **
## away_GD_form_pw       0.118982   0.025584   4.651 3.31e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 3966.0  on 3039  degrees of freedom
## Residual deviance: 3561.9  on 3033  degrees of freedom
## AIC: 8512.6
##
```

```
## Number of Fisher Scoring iterations: 5
```

```
summary(hg_reg)
```

```
##
## Call:
## glm(formula = home_score ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +
##      HomePlayer_Overall_sd + AwayPlayer_Overall_sd + +home_GD_form_pw +
##      away_GD_form_pw, family = "poisson", data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4004  -1.0332  -0.1481   0.5664   3.9653
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.923722    0.529166  -3.635 0.000278 ***
## HomePlayer_Overall_mean  0.058466    0.004796  12.189 < 2e-16 ***
## AwayPlayer_Overall_mean -0.030025    0.004925  -6.096 1.08e-09 ***
## HomePlayer_Overall_sd   0.039546    0.008432   4.690 2.73e-06 ***
## AwayPlayer_Overall_sd  -0.016714    0.008890  -1.880 0.060100 .
## home_GD_form_pw         0.068752    0.022757   3.021 0.002519 **
## away_GD_form_pw        -0.105583    0.023300  -4.532 5.86e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 3842.7  on 3039  degrees of freedom
## Residual deviance: 3404.4  on 3033  degrees of freedom
## AIC: 9159.4
##
## Number of Fisher Scoring iterations: 5
```