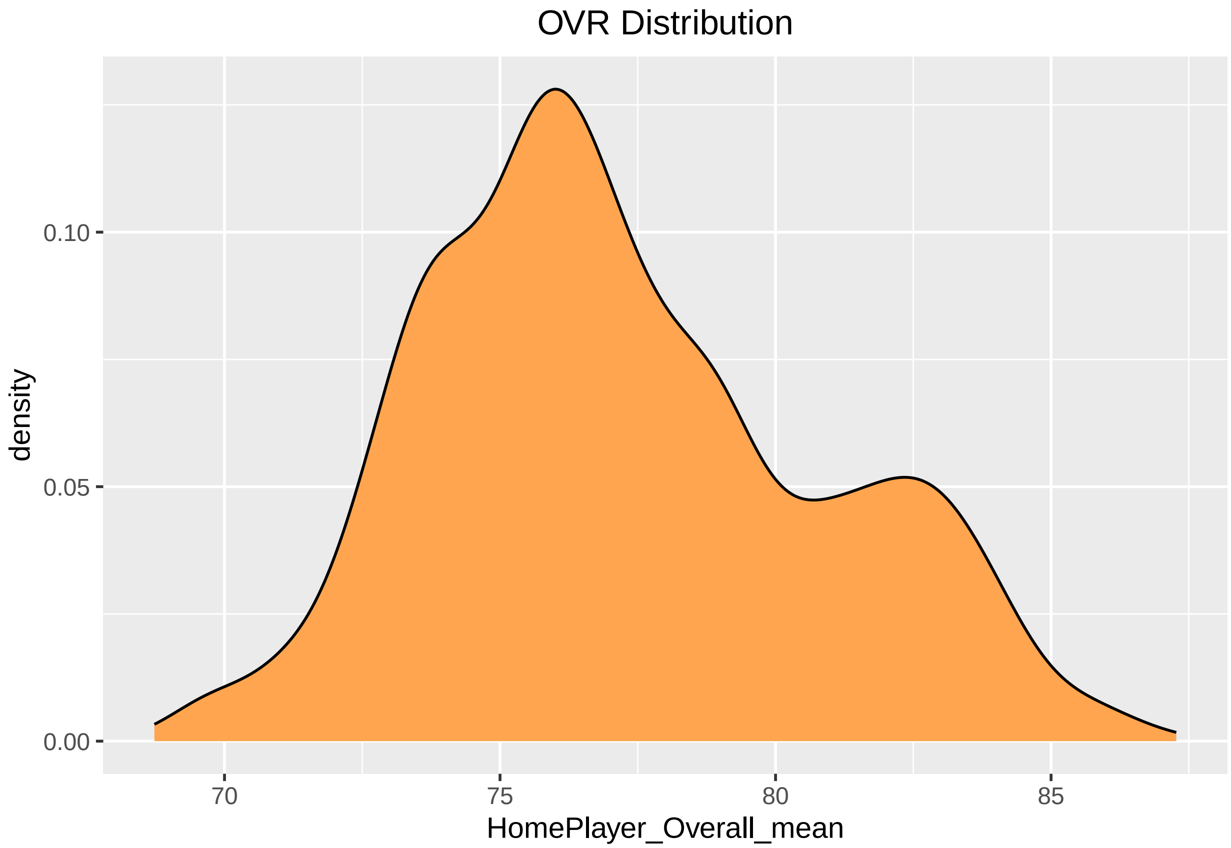
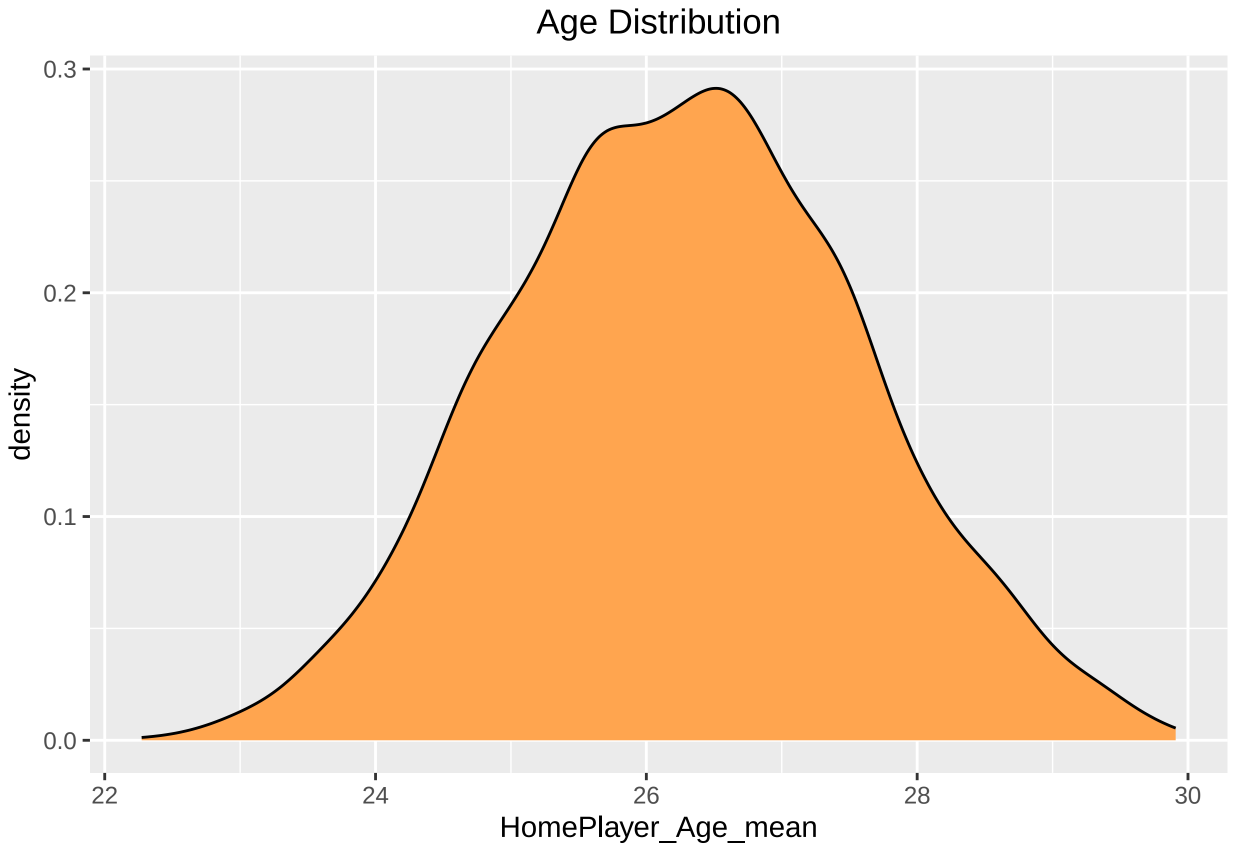
Exploring FIFA Ratings and motivation

2024-01-25

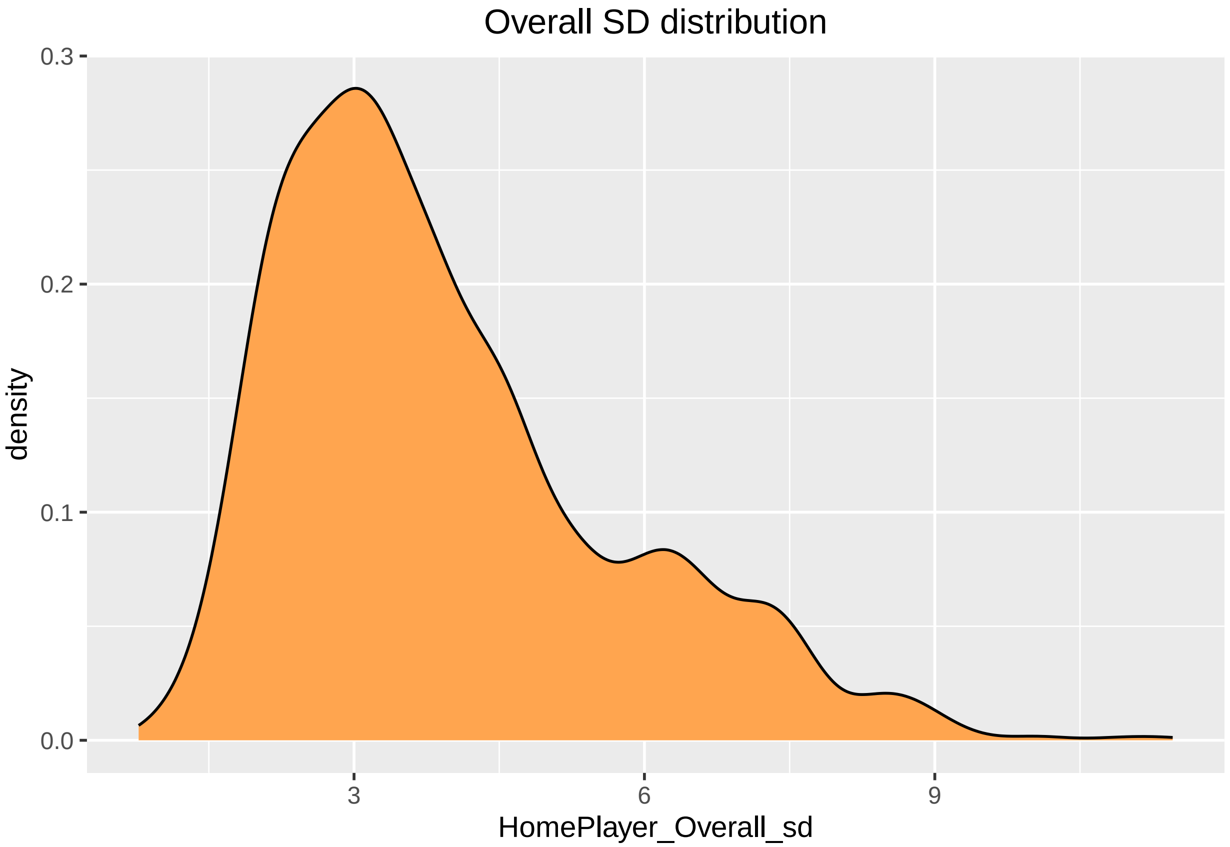
```r
ggplot(data=data, aes(x=HomePlayer_Overall_mean)) +
  geom_density(fill="tan1") + labs(title="OVR Distribution") + theme(plot.title = element_text(hjust =
```



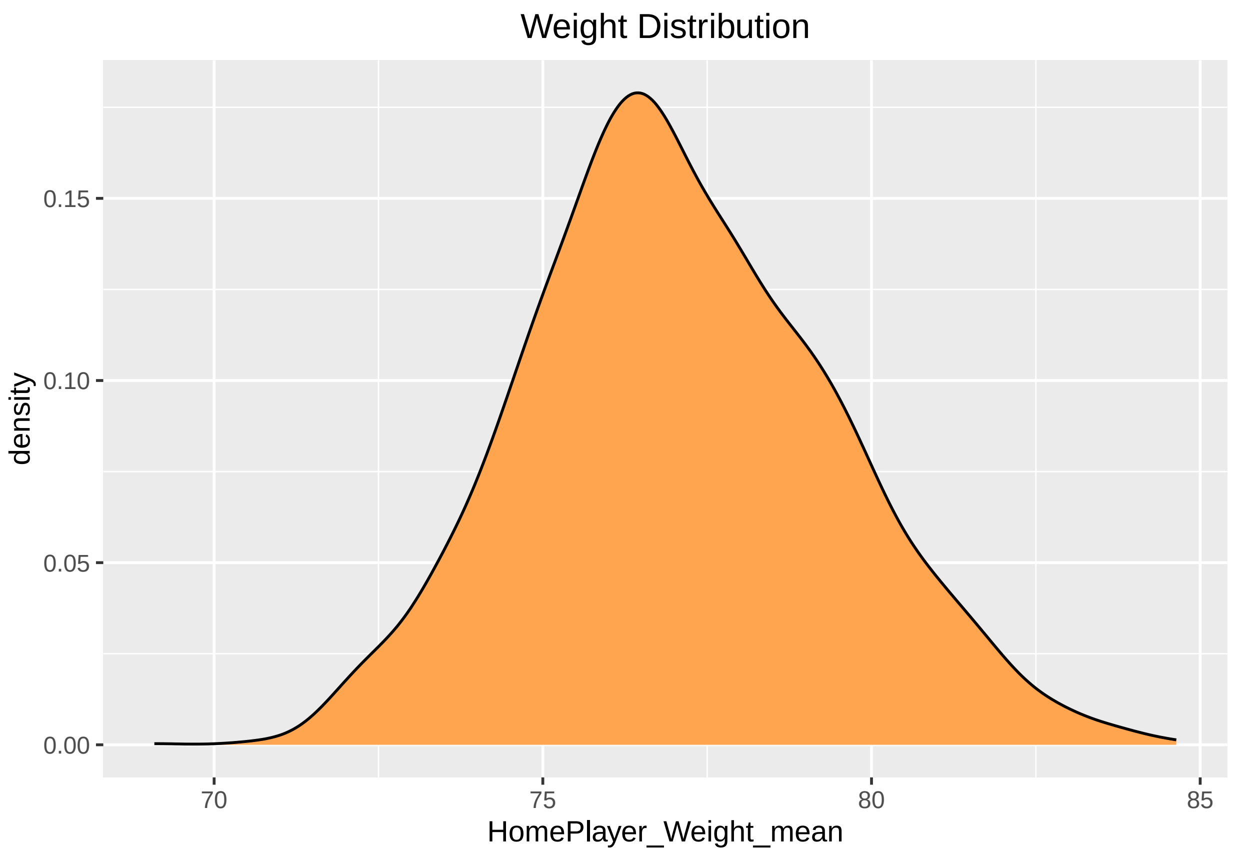
```r
ggplot(data=data, aes(x=HomePlayer_Age_mean)) +
  geom_density(fill="tan1") + labs(title="Age Distribution") + theme(plot.title = element_text(hjust =
```



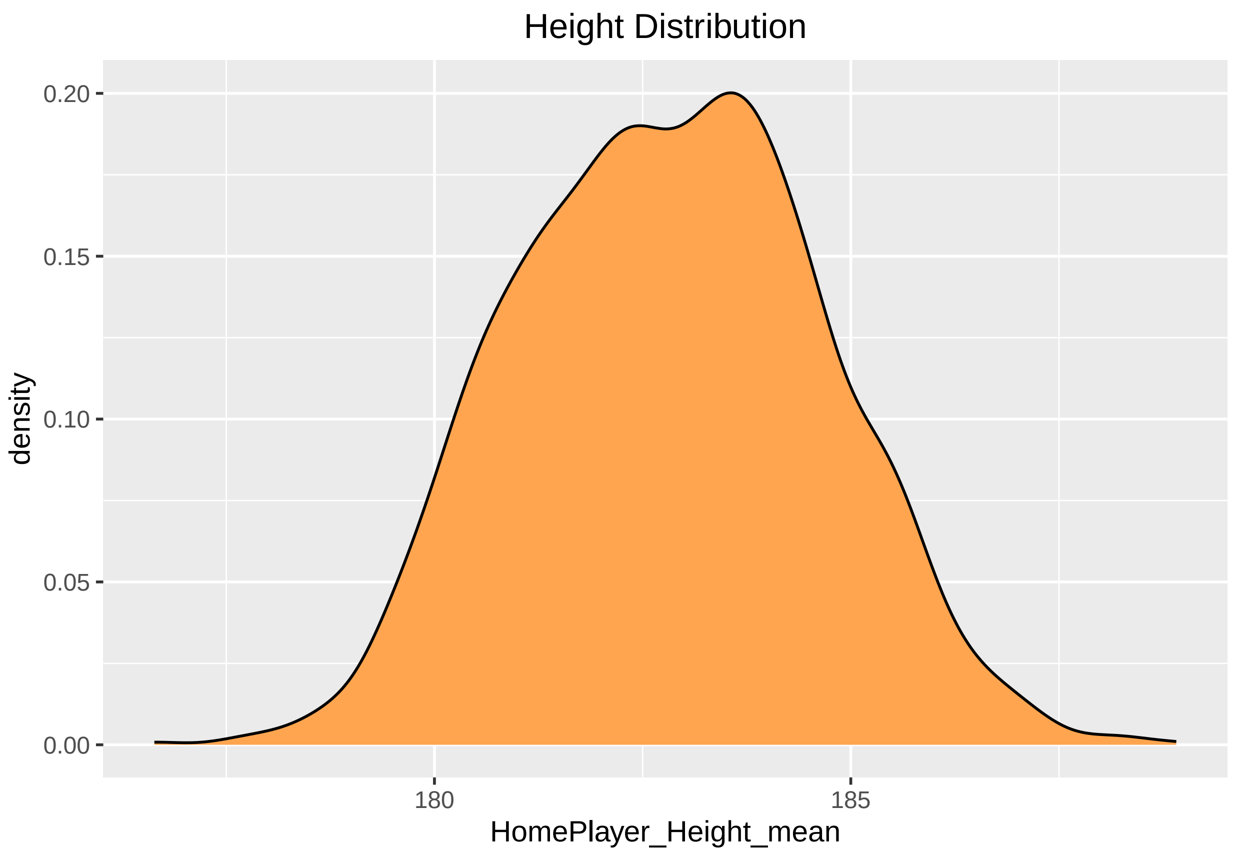
```
ggplot(data=data, aes(x=HomePlayer_Overall_sd)) +
  geom_density(fill="tan1") + labs(title="Overall SD distribution") + theme(plot.title = element_text(h
```



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

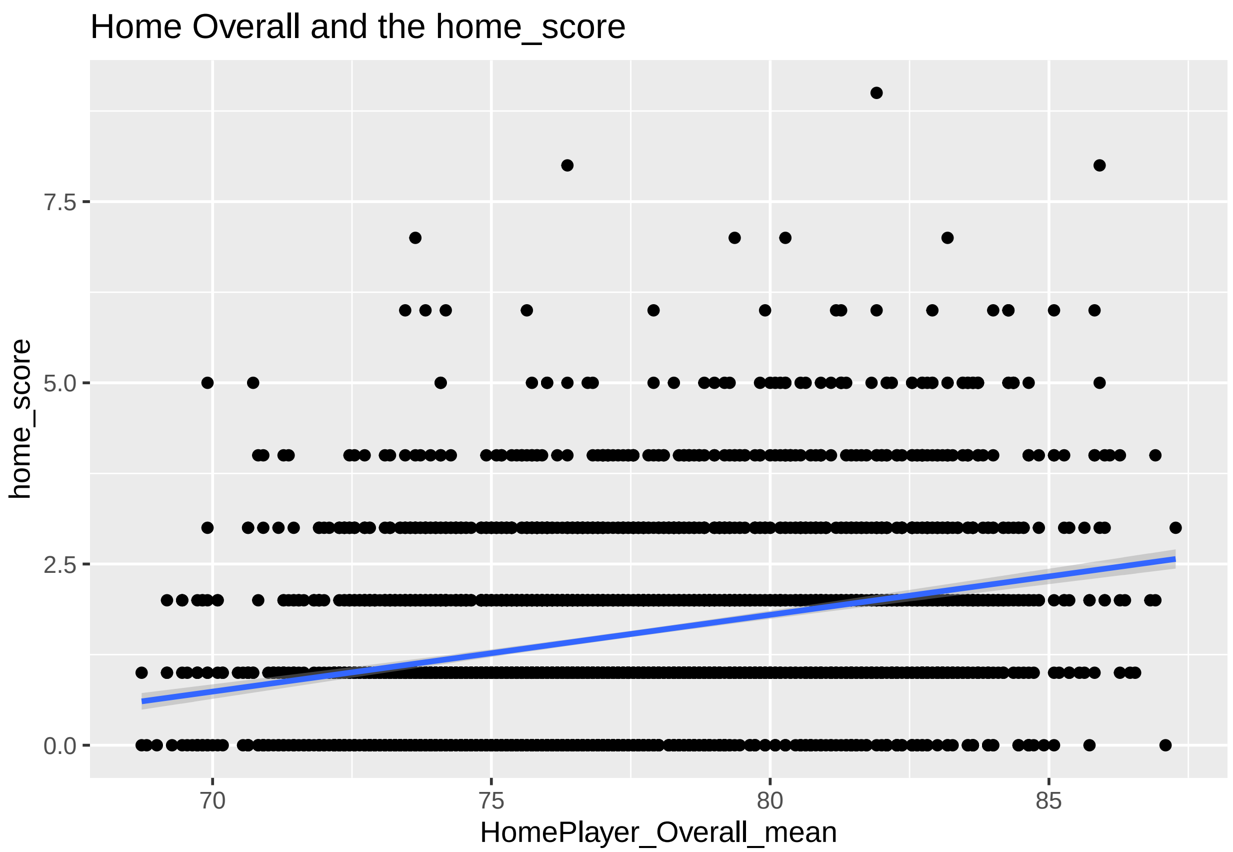


```
ggplot(data=data, aes(x=HomePlayer_Height_mean)) +
  geom_density(fill="tan1") + labs(title="Height Distribution") + theme(plot.title = element_text(hjust
```



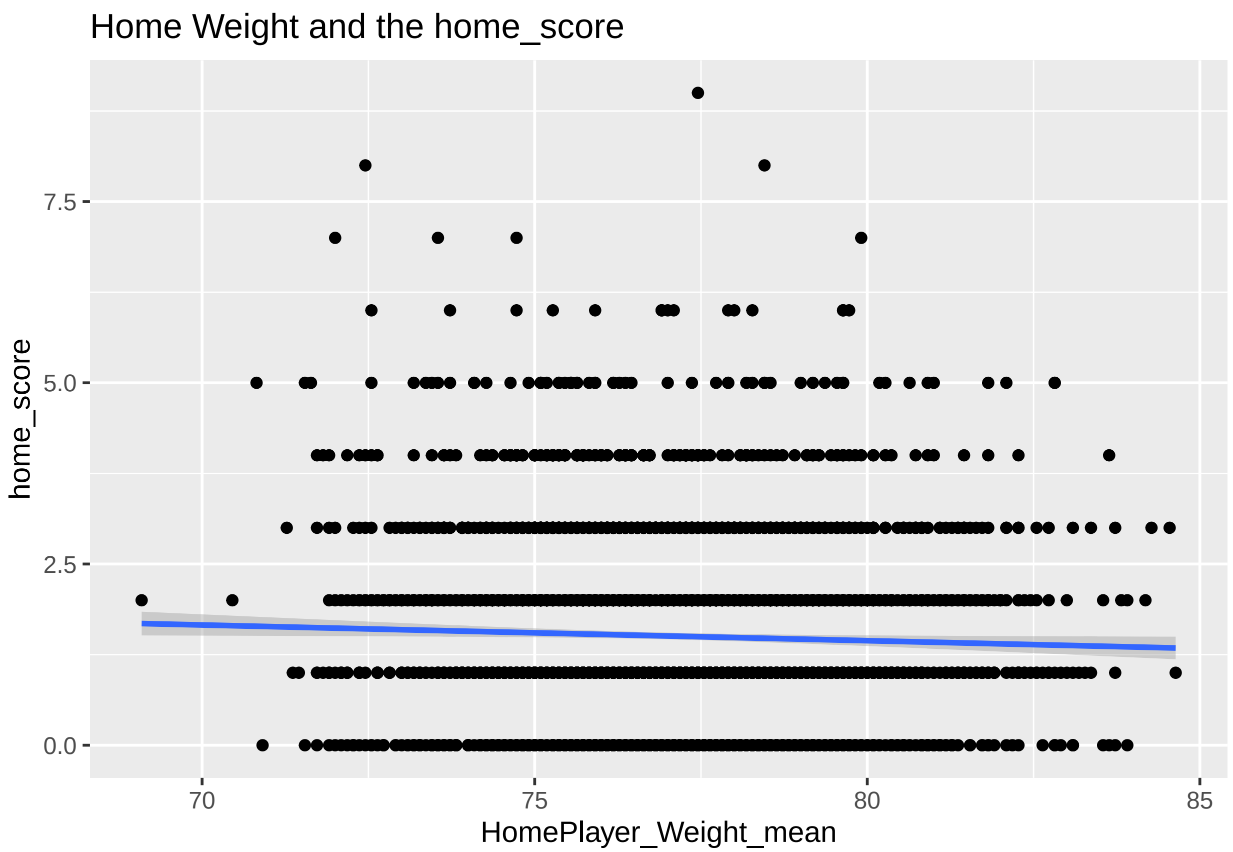
```
ggplot(data = data, aes(x = HomePlayer_Overall_mean,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Overall and the home_score")
```

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



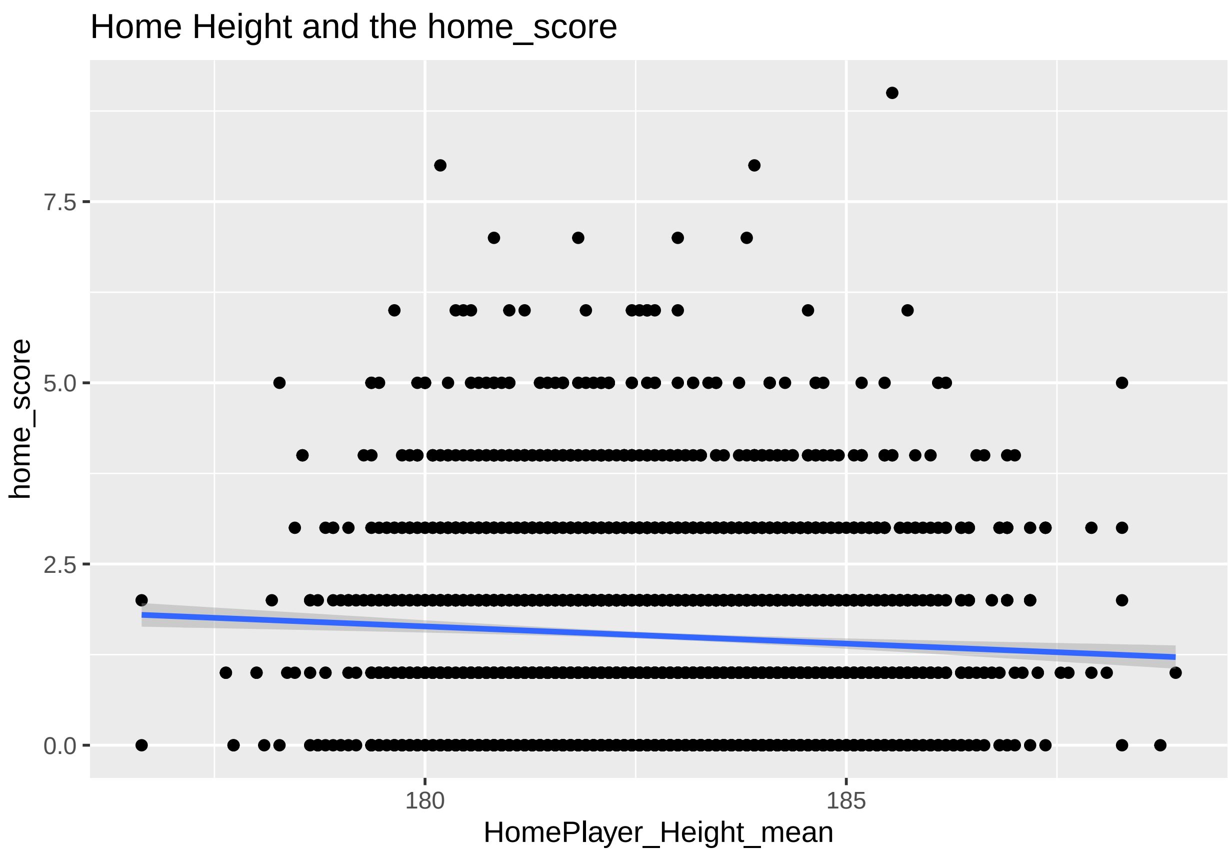
```
ggplot(data = data, aes(x = HomePlayer_Weight_mean,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Weight and the home_score")
```

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



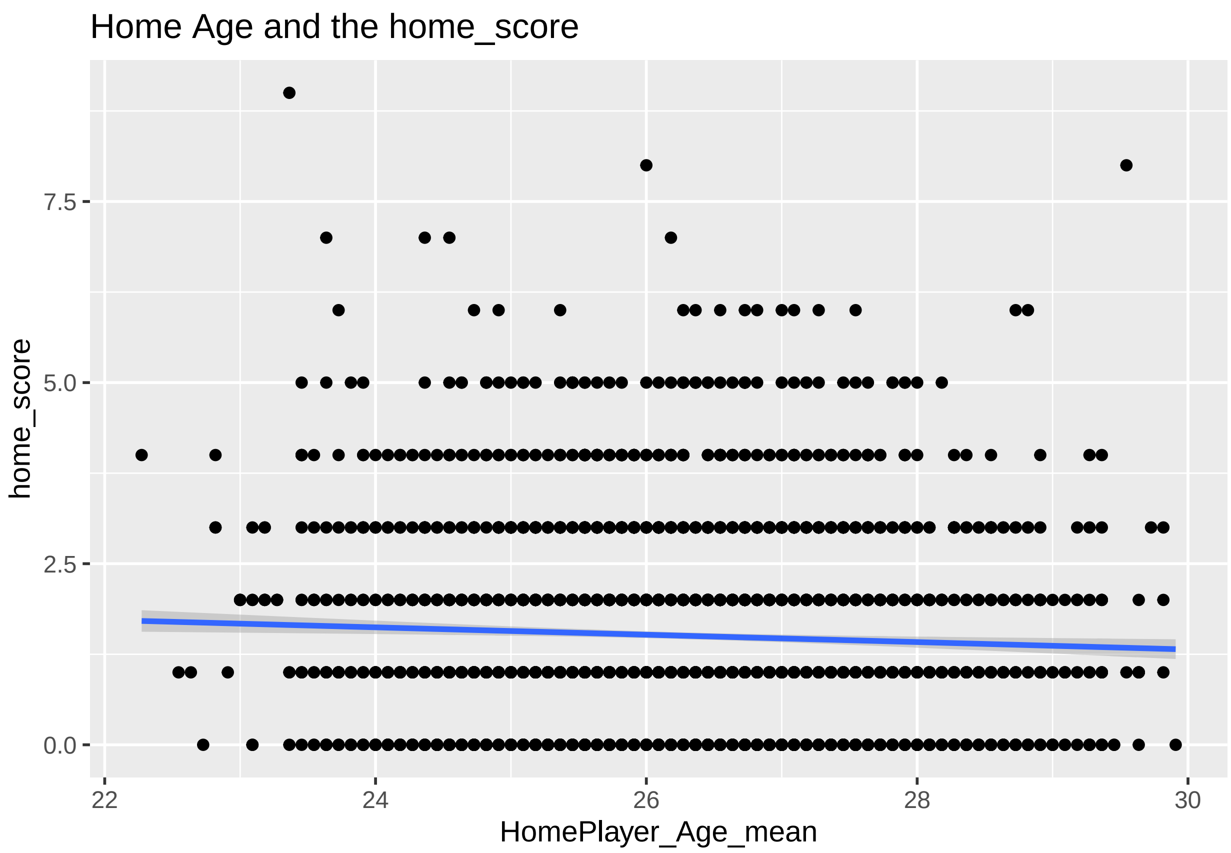
```
ggplot(data = data, aes(x = HomePlayer_Height_mean,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Height and the home_score")
```

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



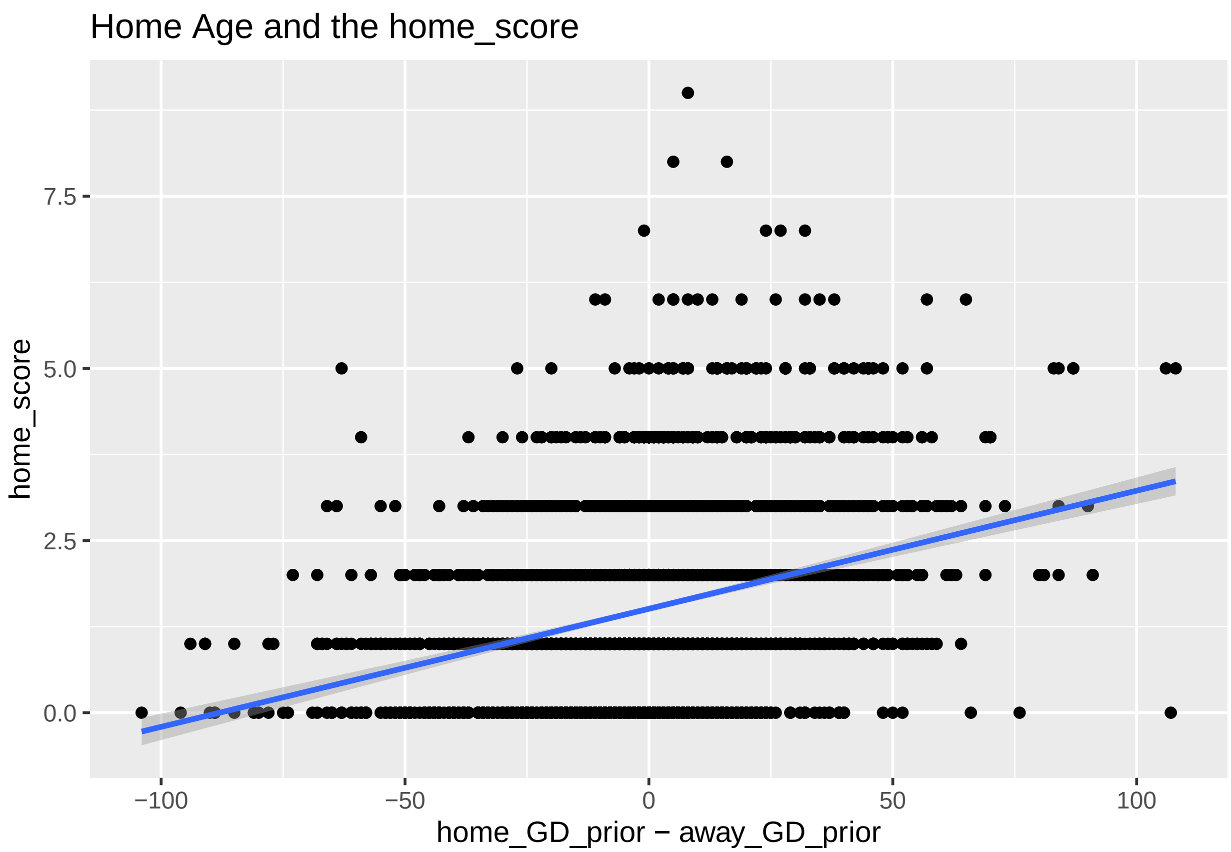
```
ggplot(data = data, aes(x = HomePlayer_Age_mean,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Age and the home_score")
```

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



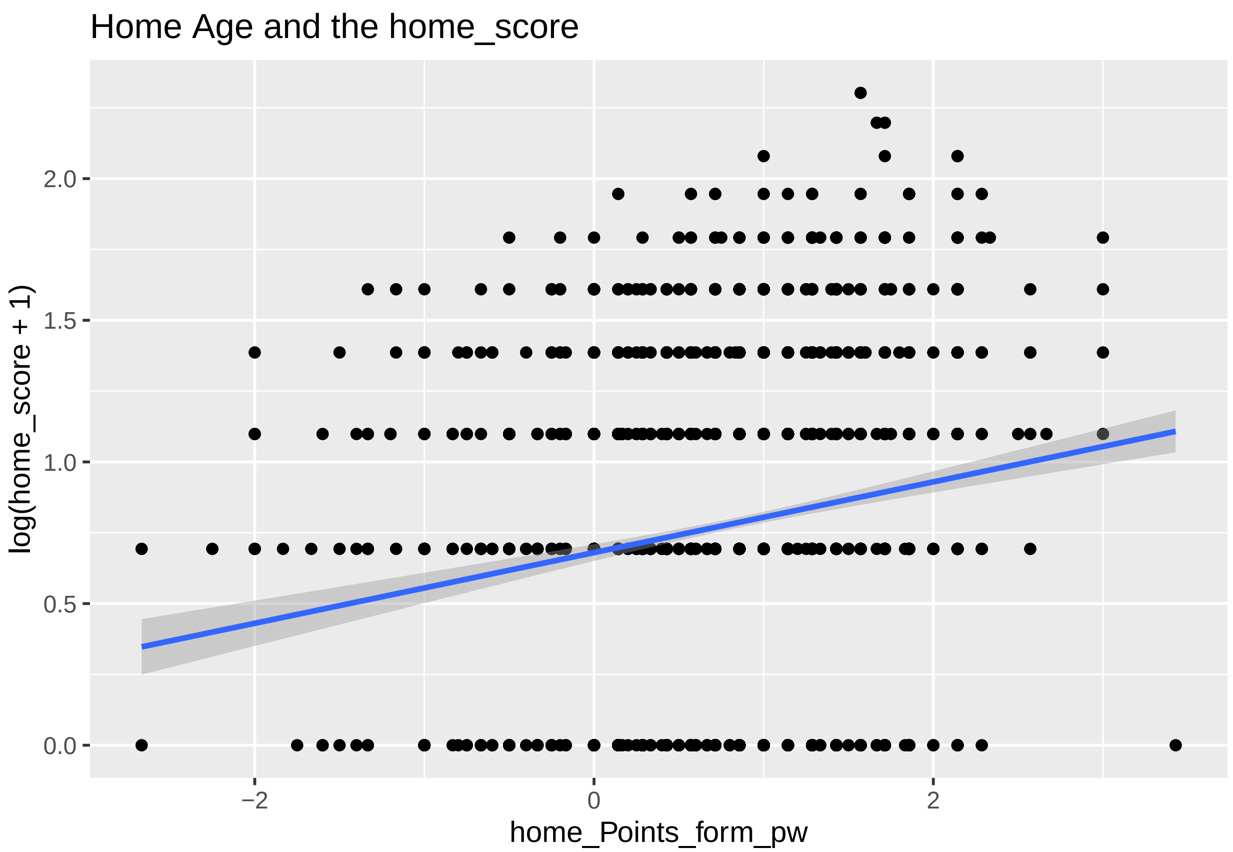
```
ggplot(data = data, aes(x = home_GD_prior- away_GD_prior,
                        y = home_score)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Age and the home_score")
```

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



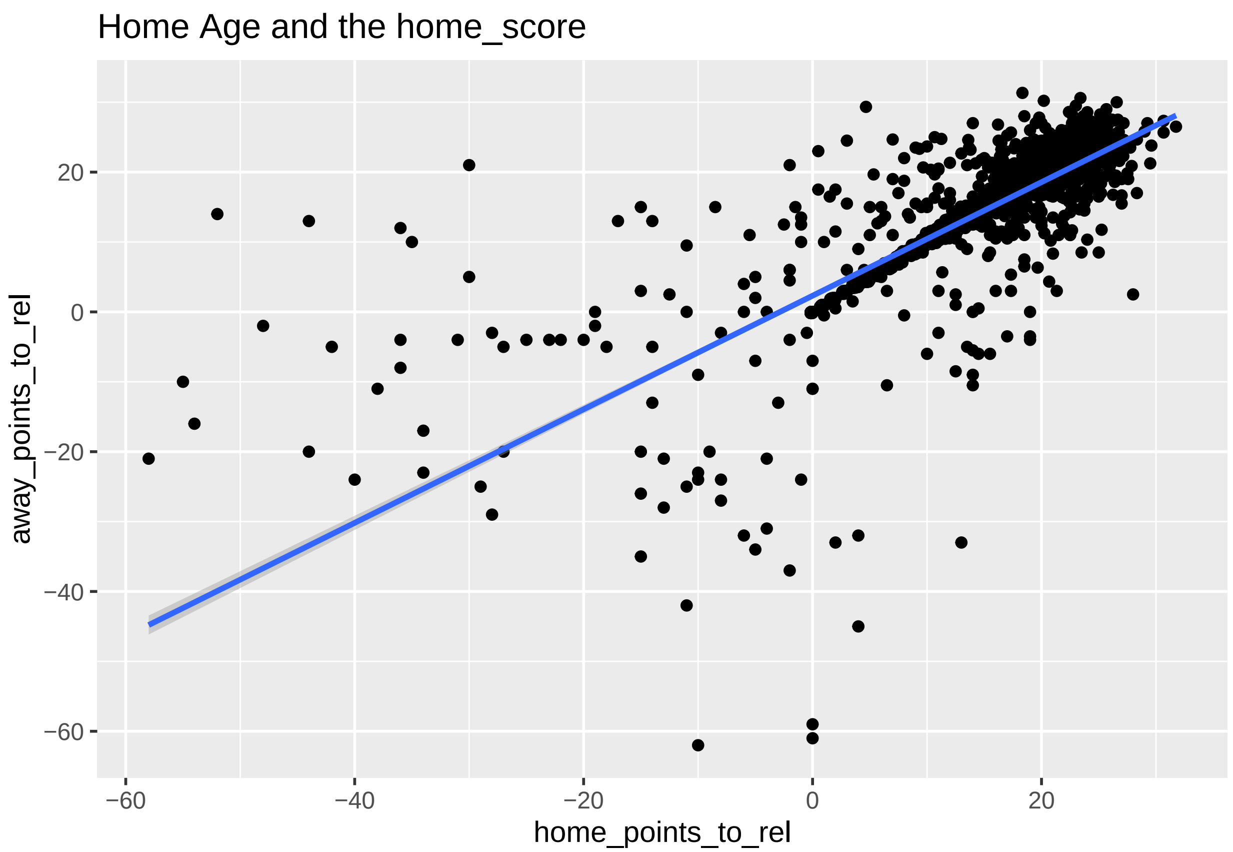
```
ggplot(data = data, aes(x = home_Points_form_pw,
                       y = log(home_score+1))) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home Age and the home_score")
```

```
## ‘geom_smooth()‘ using formula = ’y ~ x’
```



```
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’
```



Now we will see different models to check the following:

1. Does the standard deviation have a positive effect on the teams performance?
2. How is the motivation to not get relegated affects the team different than the motivation to reach top 4 / win the league
3. Which is better in estimating form? Goal differences before the game or points gained

First we will see which is better at estimating form: points form or GD form over the past 5-6 (per week)

```
data <- data%>%
  mutate(GD = home_score-away_score)
GD_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 | 0 | home_team_name)

Points_model <- felm(data=data, formula = GD ~
                HomePlayer_Overall_mean+
              HomePlayer_Overall_sd +
                AwayPlayer_Overall_mean+
                AwayPlayer_Overall_sd+
                home_Points_form_pw +
                away_Points_form_pw
               | home_team_name + away_team_name | 0 | home_team_name)

summary(GD_model)
```

```
##
## Call:
##    felm(formula = GD ~ HomePlayer_Overall_mean + HomePlayer_Overall_sd +      AwayPlayer_Overall_mean
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.7618 -1.0386 -0.0212  1.0493  7.9060
##
## Coefficients:
##                         Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.04022      0.01610   2.499  0.01798 *
## HomePlayer_Overall_sd    0.07088      0.02649   2.676  0.01180 *
## AwayPlayer_Overall_mean -0.04364      0.02202  -1.982  0.05637 .
## AwayPlayer_Overall_sd   -0.02199      0.02078  -1.058  0.29807
## home_GD_form_pw          0.08652      0.04851   1.784  0.08429 .
## away_GD_form_pw         -0.21837      0.05348  -4.083  0.00029 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 1.646 on 2971 degrees of freedom
## Multiple R-squared(full model): 0.2448   Adjusted R-squared: 0.2275
## Multiple R-squared(proj model): 0.01331   Adjusted R-squared: -0.009269
## F-statistic(full model, *iid*):14.16 on 68 and 2971 DF, p-value: < 2.2e-16
## F-statistic(proj model):  6.23 on 6 and 31 DF, p-value: 0.0002253
```

```
summary(Points_model)
```

```
##
## Call:
##    felm(formula = GD ~ HomePlayer_Overall_mean + HomePlayer_Overall_sd +      AwayPlayer_Overall_mean
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.8721 -1.0598 -0.0232  1.0390  7.8976
##
## Coefficients:
##                         Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.04221      0.01635   2.581   0.0148 *
## HomePlayer_Overall_sd    0.06922      0.02723   2.542   0.0162 *
## AwayPlayer_Overall_mean -0.04593      0.02202  -2.085   0.0454 *
## AwayPlayer_Overall_sd   -0.02688      0.02009  -1.338   0.1906
## home_Points_form_pw      0.10381      0.04979   2.085   0.0454 *
## away_Points_form_pw     -0.13316      0.08593  -1.550   0.1314
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 1.65 on 2971 degrees of freedom
## Multiple R-squared(full model): 0.2411   Adjusted R-squared: 0.2237
## Multiple R-squared(proj model): 0.008456   Adjusted R-squared: -0.01424
## F-statistic(full model, *iid*):13.88 on 68 and 2971 DF, p-value: < 2.2e-16
## F-statistic(proj model): 4.341 on 6 and 31 DF, p-value: 0.002726
```

Now we see looking at the adjusted R squared for the full model, it's better off to use to GD which makes sense because there will be instances where 2 teams are in the same form point wise but one team scores more goals.

Now we will add motivation to the mix:

```
staying_league_model <- felm(data=data, formula = GD ~
                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 | 0 |home_team_name)


tr_model <- felm(data=data, formula = GD ~
                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 | 0 | home_team_name)

gi_model <- felm(data=data, formula = GD ~
                HomePlayer_Overall_mean + AwayPlayer_Overall_mean+
              HomePlayer_Overall_sd+
                AwayPlayer_Overall_sd+
                 +home_GD_form_pw + away_GD_form_pw+
                home_match_importance + away_match_importance
               | home_team_name + away_team_name | 0 | home_team_name)

summary(gi_model)
```

```
##
## Call:
##    felm(formula = GD ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +      HomePlayer_Overall_s
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.7565 -1.0336 -0.0194  1.0516  7.9023
##
## Coefficients:
##                         Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.03957      0.01602   2.470 0.019240 *
## AwayPlayer_Overall_mean -0.04359      0.02227  -1.958 0.059342 .
## HomePlayer_Overall_sd    0.07062      0.02670   2.645 0.012719 *
## AwayPlayer_Overall_sd   -0.02167      0.02065  -1.049 0.302074
## home_GD_form_pw          0.08607      0.04878   1.765 0.087491 .
## away_GD_form_pw         -0.21617      0.05322  -4.062 0.000307 ***
## home_match_importance   -0.10055      0.08780  -1.145 0.260895
## away_match_importance    0.11685      0.09993   1.169 0.251177
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 1.646 on 2969 degrees of freedom
## Multiple R-squared(full model): 0.2452   Adjusted R-squared: 0.2274
## Multiple R-squared(proj model): 0.01375   Adjusted R-squared: -0.009507
## F-statistic(full model, *iid*):13.78 on 70 and 2969 DF, p-value: < 2.2e-16
## F-statistic(proj model): 4.938 on 8 and 31 DF, p-value: 0.0005327
```

```r
summary(tr_model)
```

```
##
## Call:
##    felm(formula = GD ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +      HomePlayer_Overall_s
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.7419 -1.0283 -0.0165  1.0365  7.9212
##
## Coefficients:
##                              Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean      0.041293     0.016030   2.576 0.014988 *
## AwayPlayer_Overall_mean     -0.042186     0.021998  -1.918 0.064394 .
## HomePlayer_Overall_sd        0.066671     0.026333   2.532 0.016630 *
## AwayPlayer_Overall_sd       -0.025572     0.021417  -1.194 0.241548
## home_GD_form_pw              0.082225     0.050088   1.642 0.110781
## away_GD_form_pw             -0.212584     0.052842  -4.023 0.000343 ***
## home_points_to_championship -0.007794     0.007826  -0.996 0.327019
## away_points_to_championship  0.009505     0.007893   1.204 0.237620
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 1.646 on 2969 degrees of freedom
## Multiple R-squared(full model): 0.2457   Adjusted R-squared: 0.2279
## Multiple R-squared(proj model): 0.01449   Adjusted R-squared: -0.008745
## F-statistic(full model, *iid*):13.82 on 70 and 2969 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.593 on 8 and 31 DF, p-value: 0.0002019
```

```r
summary(staying_league_model)
```

```
##
## Call:
##    felm(formula = GD ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +      HomePlayer_Overall_s
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.7670 -1.0324 -0.0101  1.0400  7.9111
##
## Coefficients:
##                          Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.040650     0.016234   2.504 0.017755 *
## AwayPlayer_Overall_mean -0.043090     0.021934  -1.965 0.058487 .
## HomePlayer_Overall_sd    0.070378     0.026463   2.660 0.012273 *
## AwayPlayer_Overall_sd   -0.021887     0.021783  -1.005 0.322778
## home_GD_form_pw          0.080000     0.049799   1.606 0.118313
## away_GD_form_pw         -0.211521     0.053062  -3.986 0.000379 ***
## home_points_to_rel      -0.008169     0.007610  -1.073 0.291340
## away_points_to_rel       0.009348     0.008864   1.055 0.299730
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 1.646 on 2969 degrees of freedom
## Multiple R-squared(full model): 0.2454   Adjusted R-squared: 0.2276
## Multiple R-squared(proj model): 0.01403   Adjusted R-squared: -0.009216
## F-statistic(full model, *iid*):13.79 on 70 and 2969 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.669 on 8 and 31 DF, p-value: 0.0001811
```

```
linearHypothesis(gi_model, c("home_match_importance= 0", "away_match_importance=0"))
```

```
## Linear hypothesis test
##
## Hypothesis:
## home_match_importance = 0
## away_match_importance = 0
##
## Model 1: restricted model
## Model 2: GD ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean + HomePlayer_Overall_sd +
##     AwayPlayer_Overall_sd + +home_GD_form_pw + away_GD_form_pw +
##     home_match_importance + away_match_importance | home_team_name +
##     away_team_name | 0 | home_team_name
##
##   Res.Df Df Chisq Pr(>Chisq)
## 1     33
## 2     31  2 1.415     0.4929
```

```
linearHypothesis(tr_model, c("home_points_to_championship= 0", "away_points_to_championship=0"))
```

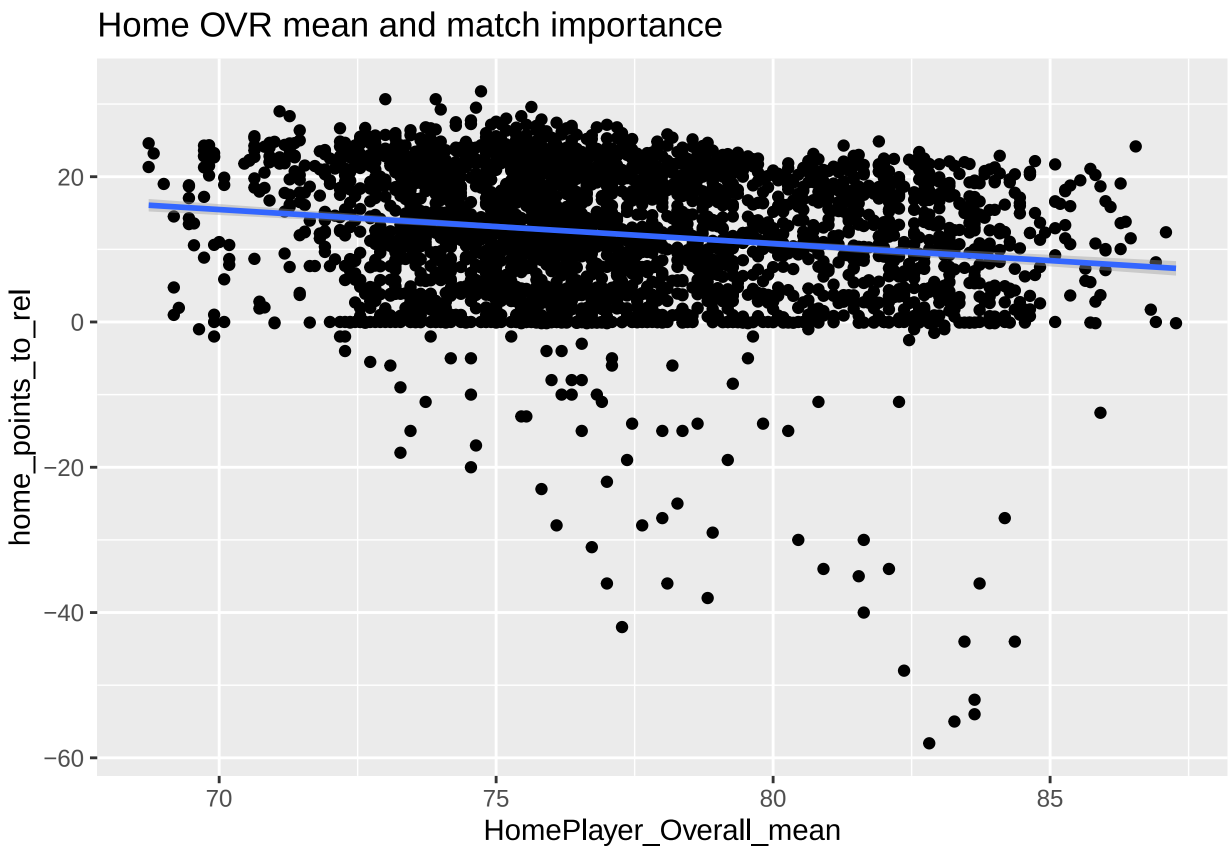
```
## Linear hypothesis test
##
## Hypothesis:
## home_points_to_championship = 0
## away_points_to_championship = 0
##
## Model 1: restricted model
## Model 2: GD ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean + HomePlayer_Overall_sd +
##     AwayPlayer_Overall_sd + +home_GD_form_pw + away_GD_form_pw +
##     home_points_to_championship + away_points_to_championship |
##     home_team_name + away_team_name | 0 | home_team_name
##
##   Res.Df Df  Chisq Pr(>Chisq)
## 1     33
## 2     31  2 2.8826     0.2366
```

```r
linearHypothesis(staying_league_model, c("home_points_to_rel= 0", "away_points_to_rel=0"))
```

```
## Linear hypothesis test
##
## Hypothesis:
## home_points_to_rel = 0
## away_points_to_rel = 0
##
## Model 1: restricted model
## Model 2: GD ~ 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 | 0 | home_team_name
##
##   Res.Df Df Chisq Pr(>Chisq)
## 1     33
## 2     31  2 1.207     0.5469
```

```r
ggplot(data = data, aes(x = HomePlayer_Overall_mean,
                        y = home_points_to_rel)) +
  geom_point() + stat_smooth(method = "lm") +
  ggtitle("Home OVR mean and match importance")
```

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



Naively we can say there is a problem when motivation, but that can be due to the high correlations between both the teams quality, and their position in the league, and we can gain information mainly from non linear models that can use such information to predict better.

```
lm_model <- felm(data=data, formula = GD ~
                 HomePlayer_Overall_mean +
                 HomePlayer_Overall_min*
                 HomePlayer_Overall_max +

                 AwayPlayer_Overall_mean +
                 AwayPlayer_Overall_min*
                 AwayPlayer_Overall_max +
 home_GD_form_pw +away_GD_form_pw
               | home_team_name + away_team_name | 0 | home_team_name)
summary(lm_model)
```

```
##
## Call:
##    felm(formula = GD ~ HomePlayer_Overall_mean + HomePlayer_Overall_min *      HomePlayer_Overall_ma
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.8520 -1.0313 -0.0198  1.0593  7.8955
##
## Coefficients:
##                                                 Estimate Cluster s.e. t value
## HomePlayer_Overall_mean                        0.0589418    0.0296532   1.988
## HomePlayer_Overall_min                        -0.1803598    0.1462135  -1.234
## HomePlayer_Overall_max                        -0.1222168    0.1284245  -0.952
## AwayPlayer_Overall_mean                       -0.0407594    0.0304499  -1.339
## AwayPlayer_Overall_min                        -0.0526484    0.1026241  -0.513
## AwayPlayer_Overall_max                        -0.0671599    0.0879776  -0.763
## home_GD_form_pw                                0.0781577    0.0503293   1.553
## away_GD_form_pw                               -0.2172026    0.0533514  -4.071
## HomePlayer_Overall_min:HomePlayer_Overall_max  0.0018678    0.0017544   1.065
## AwayPlayer_Overall_min:AwayPlayer_Overall_max  0.0007321    0.0012437   0.589
##                                               Pr(>|t|)
## HomePlayer_Overall_mean                         0.0557 .
## HomePlayer_Overall_min                          0.2266
## HomePlayer_Overall_max                          0.3486
## AwayPlayer_Overall_mean                         0.1904
## AwayPlayer_Overall_min                          0.6116
## AwayPlayer_Overall_max                          0.4510
## home_GD_form_pw                                 0.1306
## away_GD_form_pw                                 0.0003 ***
## HomePlayer_Overall_min:HomePlayer_Overall_max   0.2953
## AwayPlayer_Overall_min:AwayPlayer_Overall_max   0.5603
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 1.645 on 2967 degrees of freedom
## Multiple R-squared(full model): 0.2464   Adjusted R-squared: 0.2281
## Multiple R-squared(proj model): 0.01541   Adjusted R-squared: -0.008478
## F-statistic(full model, *iid*):13.48 on 72 and 2967 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.002 on 10 and 31 DF, p-value: 0.0002511
```

```
res_model <- felm(data=data, formula = GD ~
                 HomePlayer_Overall_mean +
                 AwayPlayer_Overall_mean +
 home_GD_form_pw +away_GD_form_pw
               | home_team_name + away_team_name | 0 | home_team_name)
summary(res_model)
```

```
##
## Call:
##    felm(formula = GD ~ HomePlayer_Overall_mean + AwayPlayer_Overall_mean +      home_GD_form_pw + awa
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.6627 -1.0475 -0.0163  1.0446  7.7842
##
## Coefficients:
##                         Estimate Cluster s.e. t value Pr(>|t|)
## HomePlayer_Overall_mean  0.01626      0.01874   0.868 0.392332
## AwayPlayer_Overall_mean -0.03109      0.02026  -1.534 0.135096
## home_GD_form_pw          0.09740      0.04788   2.034 0.050547 .
## away_GD_form_pw         -0.21773      0.05426  -4.013 0.000353 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.649 on 2973 degrees of freedom
## Multiple R-squared(full model): 0.2417   Adjusted R-squared: 0.2249
## Multiple R-squared(proj model): 0.009216   Adjusted R-squared: -0.01278
## F-statistic(full model, *iid*):14.36 on 66 and 2973 DF, p-value: < 2.2e-16
## F-statistic(proj model): 6.352 on 4 and 31 DF, p-value: 0.0007453
```

```
summary(lm_model)
```

```
##
## Call:
##    felm(formula = GD ~ HomePlayer_Overall_mean + HomePlayer_Overall_min *      HomePlayer_Overall_ma
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.8520 -1.0313 -0.0198  1.0593  7.8955
##
## Coefficients:
##                                                 Estimate Cluster s.e. t value
## HomePlayer_Overall_mean                        0.0589418    0.0296532   1.988
## HomePlayer_Overall_min                        -0.1803598    0.1462135  -1.234
## HomePlayer_Overall_max                        -0.1222168    0.1284245  -0.952
## AwayPlayer_Overall_mean                       -0.0407594    0.0304499  -1.339
## AwayPlayer_Overall_min                        -0.0526484    0.1026241  -0.513
## AwayPlayer_Overall_max                        -0.0671599    0.0879776  -0.763
## home_GD_form_pw                                0.0781577    0.0503293   1.553
## away_GD_form_pw                               -0.2172026    0.0533514  -4.071
## HomePlayer_Overall_min:HomePlayer_Overall_max  0.0018678    0.0017544   1.065
## AwayPlayer_Overall_min:AwayPlayer_Overall_max  0.0007321    0.0012437   0.589
##                                                Pr(>|t|)
## HomePlayer_Overall_mean                          0.0557 .
## HomePlayer_Overall_min                           0.2266
## HomePlayer_Overall_max                           0.3486
## AwayPlayer_Overall_mean                          0.1904
## AwayPlayer_Overall_min                           0.6116
## AwayPlayer_Overall_max                           0.4510
## home_GD_form_pw                                  0.1306
## away_GD_form_pw                                  0.0003 ***
## HomePlayer_Overall_min:HomePlayer_Overall_max   0.2953
## AwayPlayer_Overall_min:AwayPlayer_Overall_max   0.5603
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 1.645 on 2967 degrees of freedom
## Multiple R-squared(full model): 0.2464   Adjusted R-squared: 0.2281
## Multiple R-squared(proj model): 0.01541   Adjusted R-squared: -0.008478
## F-statistic(full model, *iid*):13.48 on 72 and 2967 DF, p-value: < 2.2e-16
## F-statistic(proj model): 5.002 on 10 and 31 DF, p-value: 0.0002511
```

Now we can do an F test to see weather or not min-max variables are statistically significant together, as for the restricted model, we will get: ((0.2464-0.2417)/5)/((1-0.2464)/(3040-5)) = 3.78569532909 > 2.0986 meaning we can say that these values are statistically significant together

now if we derive by the max and min for the home team (assuming its pretty much the same effect on the away team) we will get: -0.1803598+0.0018678max by min and -0.1222168 +0.0018678min by max

lets look when the marginal effect for increasing our best player is greater than the marginal effect of increasing our worst: -0.1222168 +0.0018678min > -0.1803598+0.0018678max and we will get it holds for max-min < 31.1291 (https://www.wolframalpha.com/input?i2d=true&i=-0.1222168+%2B0.0018678x+%3E+-0.1803598%2B0.0018678y when x is the min and y is the max)

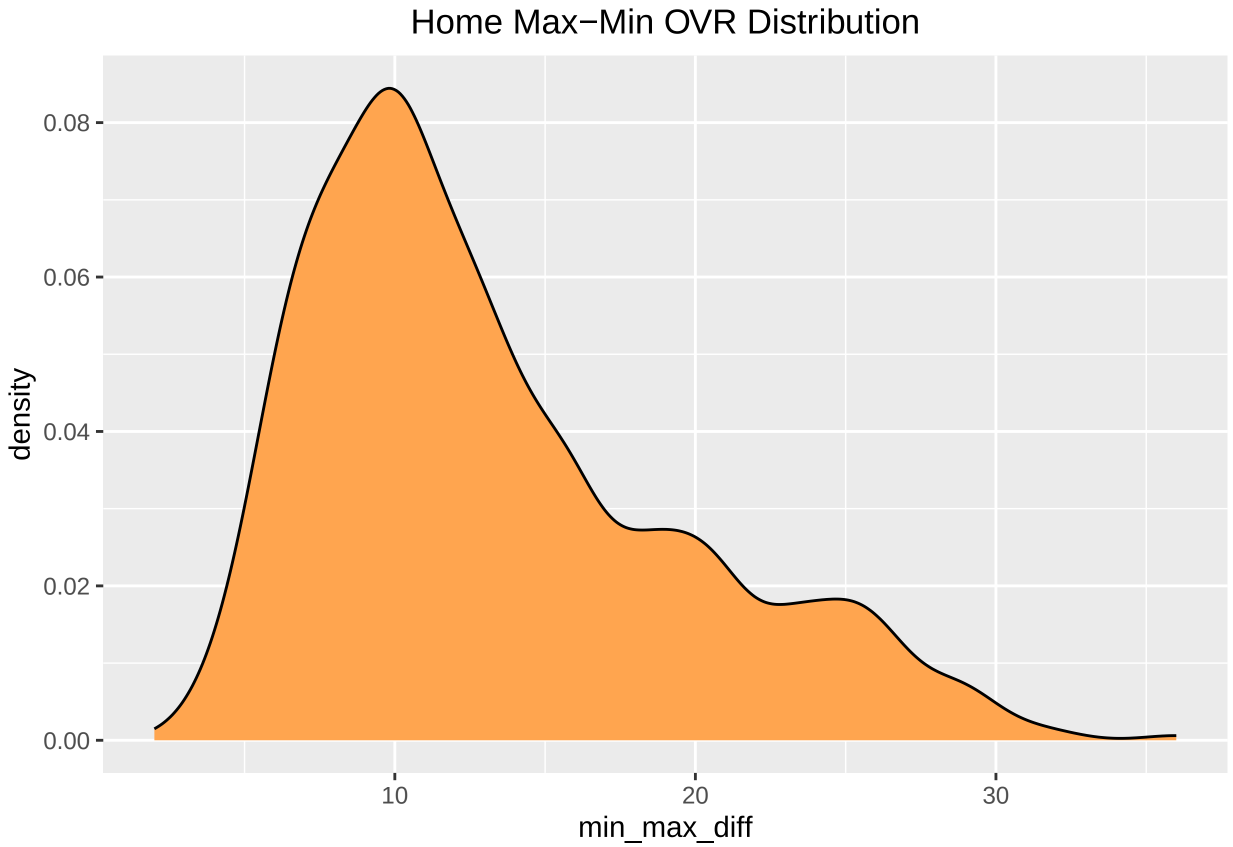
now the funny thing is, it holds for almost all premier league teams, lets show below:

```r
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
```

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



and what we find is the te gap rarely goes above 31, and even then, it could either stem from actually having different marginal effects, but also due to players being new and automatically assigned the lowest value of the premier league when we set up the data, meaning if we would've removed matches containing those players, we would've removed these matches.

- Conclusion *

We see that we can't assume that there are diminishing returns on the talent of a player for the the teams. If we look at football matches as experiments, and there are no other variables missing to our data, we can assume that there are increasing returns, but even if we don't, we can still raise questions about how the talent of players are affecting the team, and if EA ranks the players in a way which means that for example a 91 rated player in real life is significantly better than a 90 player (in the same league).