AFL With Elo

```
## New names:
## * `` -> ...1
```

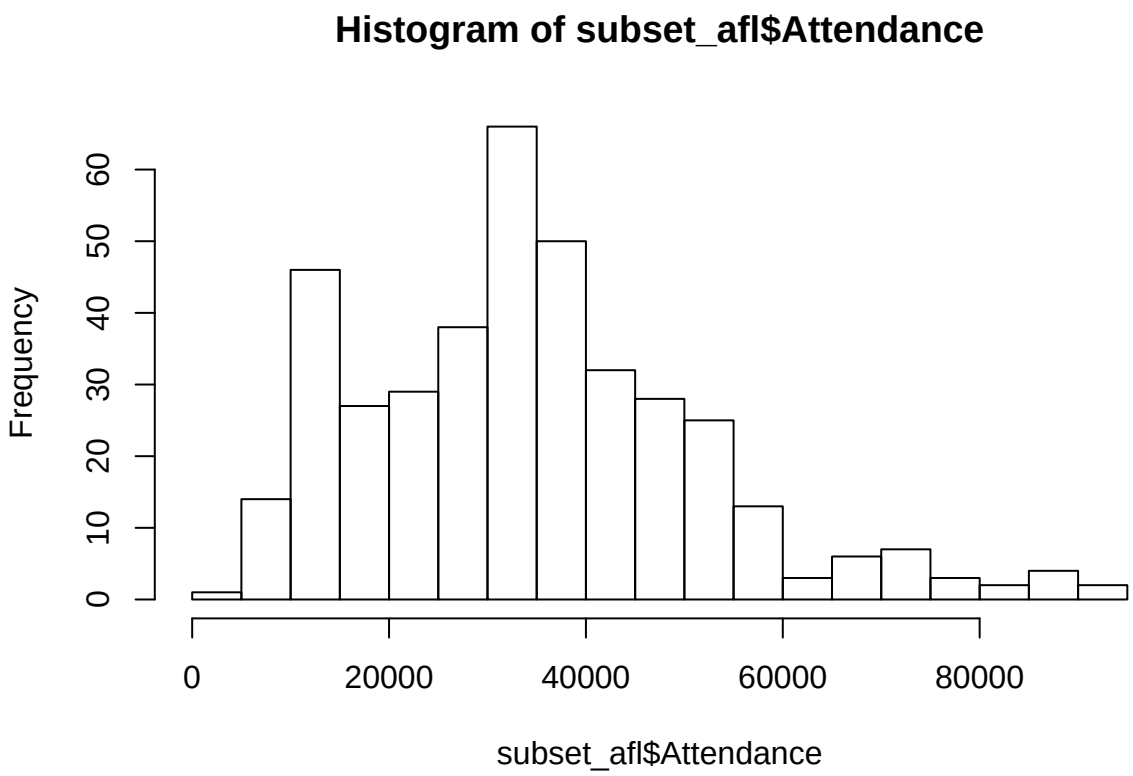
```
subset_afl <- afl[which(!is.na(afl$Home_top8)),]
write.csv(subset_afl, "subset_afl_elo.csv")
```

```
subset_afl <- read.csv("subset_afl_elo.csv")
library(car)
```

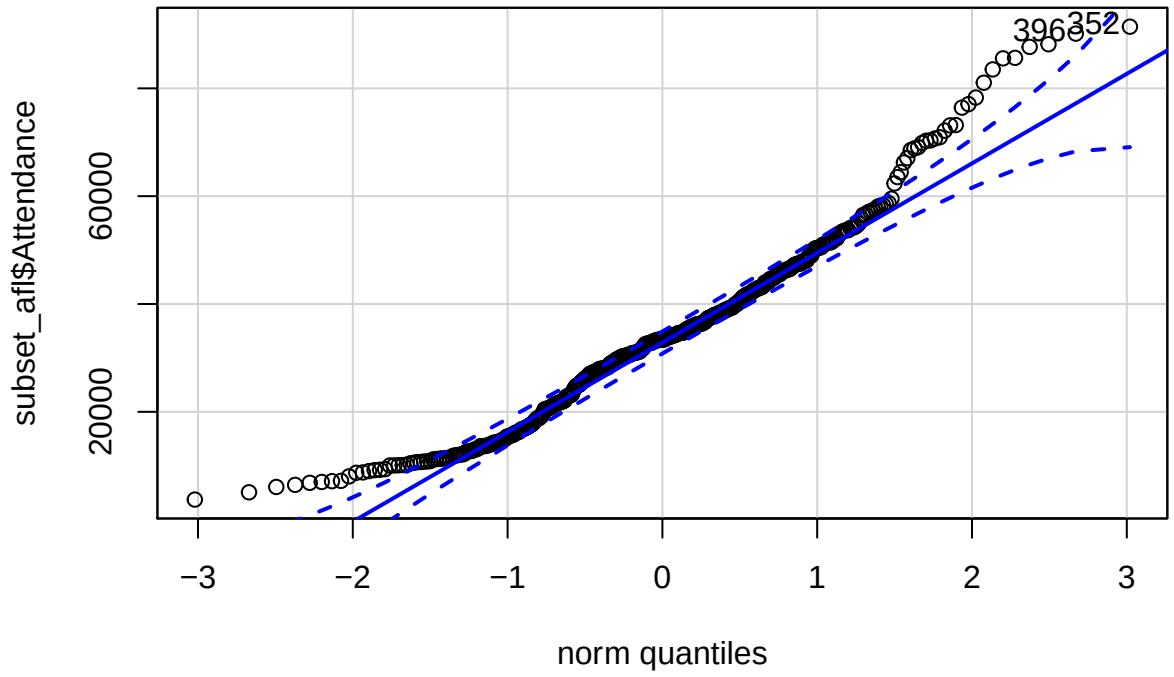
```
## Loading required package: carData
```

```
# Create var for difference in Elo between the teams
subset_afl$Diff_Elo <- abs(subset_afl$Home_elo - subset_afl$Away_elo)
# Create var combining the Elo of the teams
subset_afl$Elo_Combined <- subset_afl$Home_elo+subset_afl$Away_elo
```

```
hist(subset_afl$Attendance, breaks = 20)
```




```
qqPlot(subset_afl$Attendance)
```



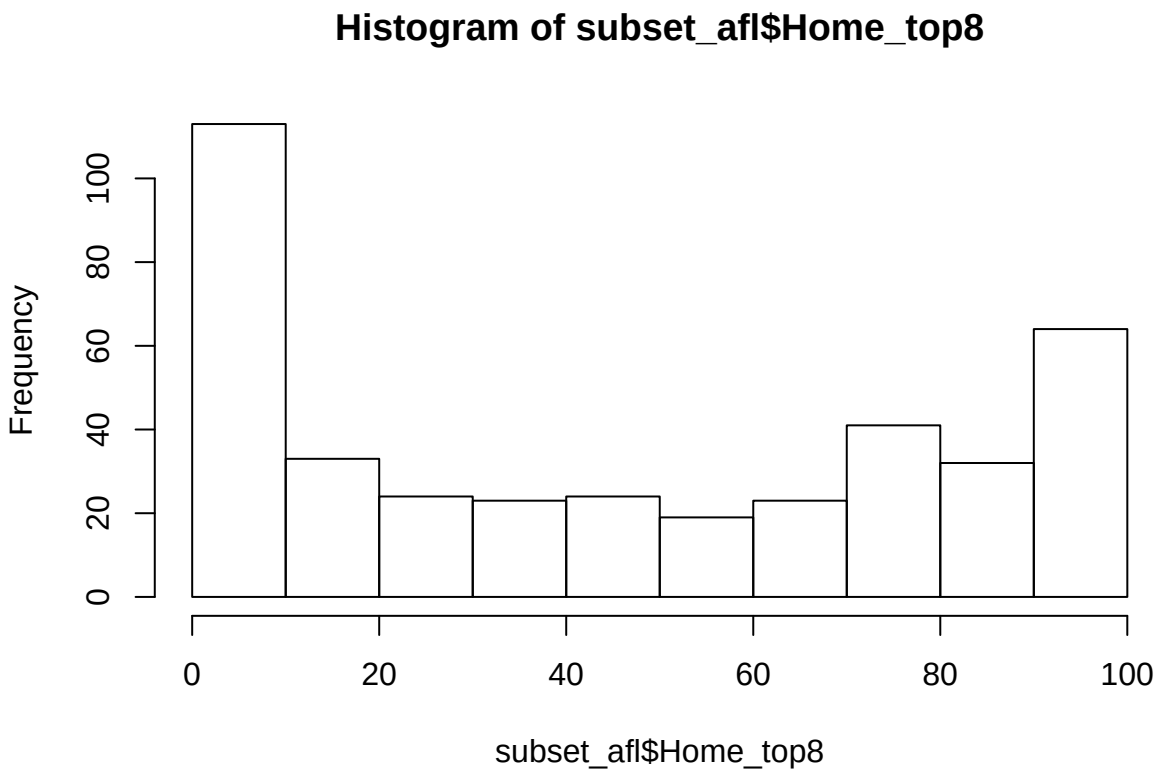

```
## [1] 352 396
```

```
shapiro.test(subset_afl$Attendance)
```

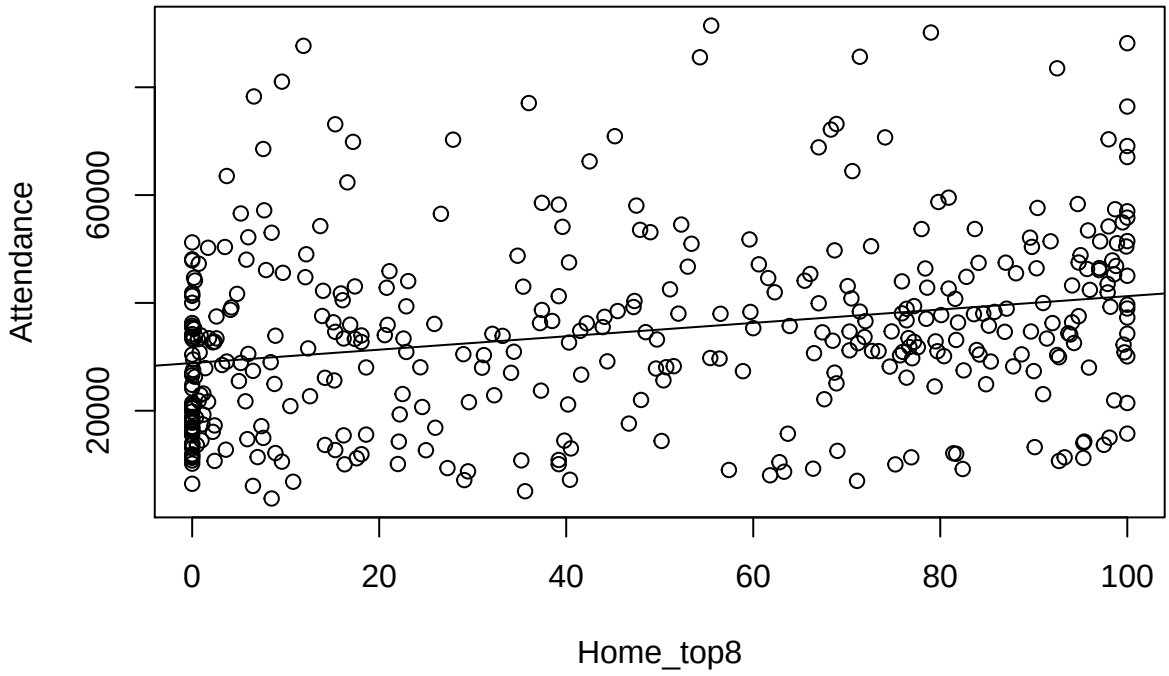
```
##
##  Shapiro-Wilk normality test
##
## data:  subset_afl$Attendance
## W = 0.95986, p-value = 6.251e-09
```

We can see that Attendance is not normally distributed. For simplicity I will continue running linear models with the assumption that Attendance is normally distributed but we will need to address this later on. It helps that the sample size of 396 is somewhat large.

```
hist(subset_afl$Home_top8)
```



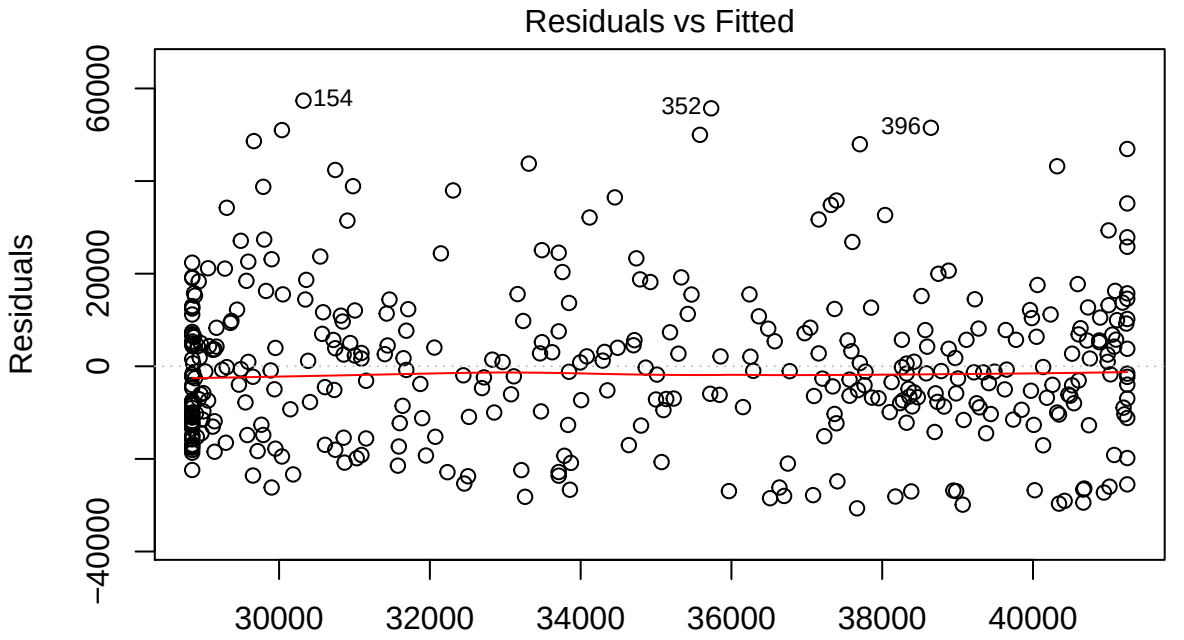

```
plot(Attendance~Home_top8, data = subset_afl)
fit <- lm(Attendance~Home_top8, data=subset_afl)
abline(fit)
```

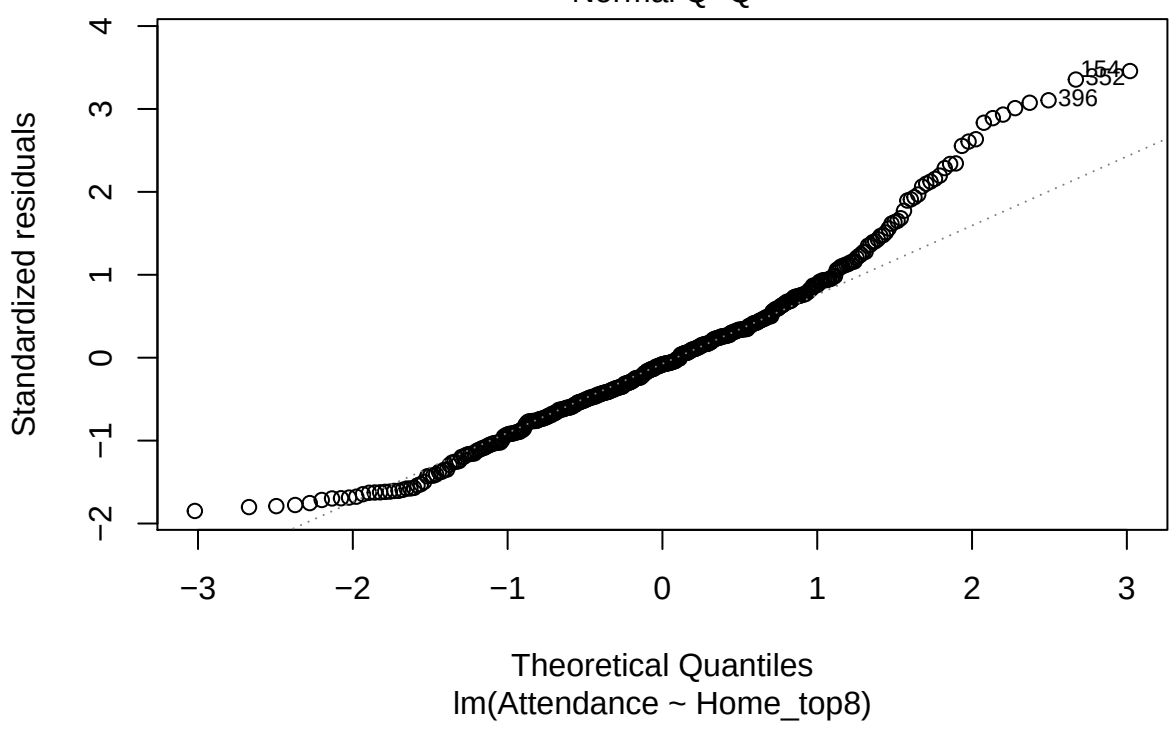


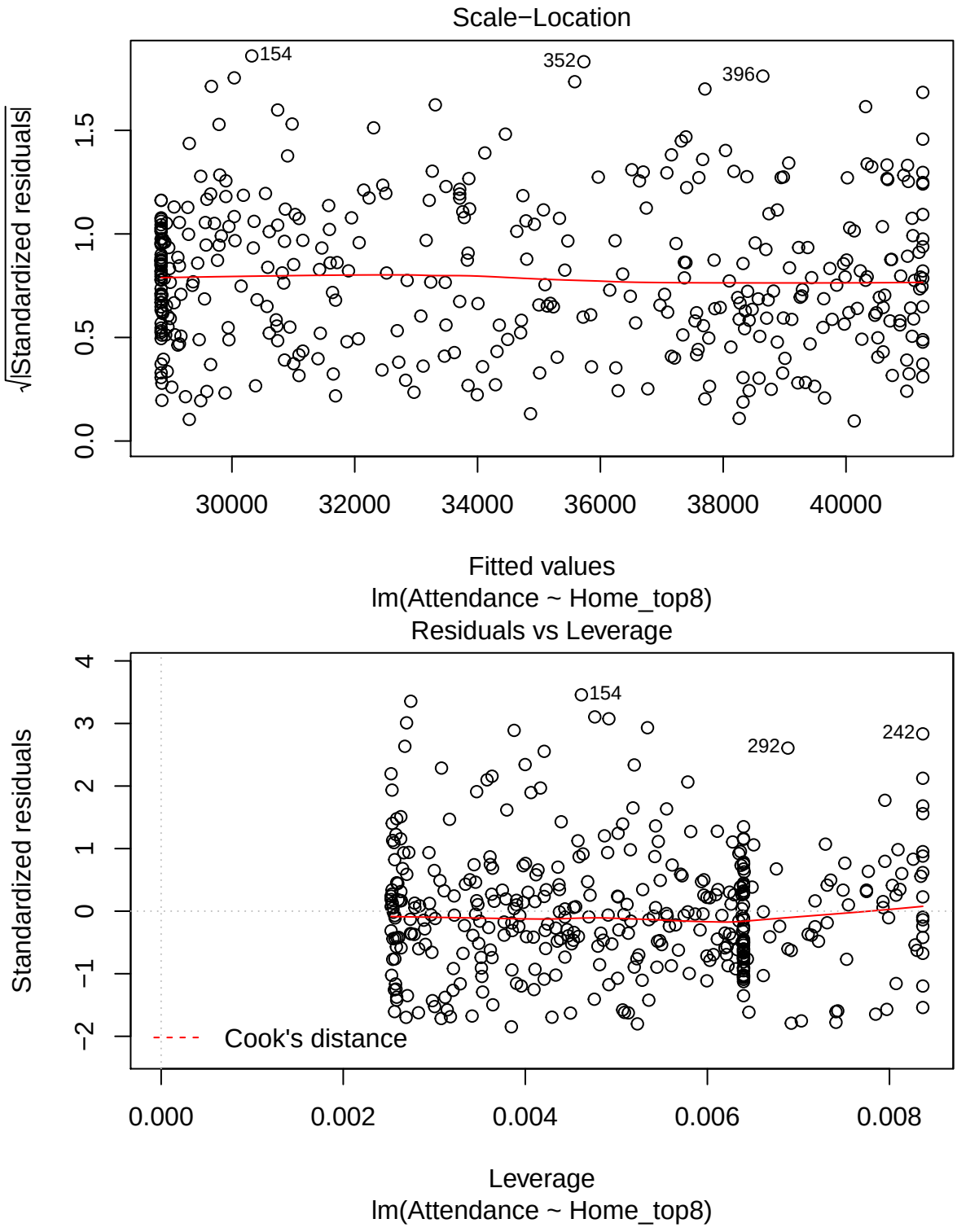

```
cor(subset_afl$Attendance, subset_afl$Home_top8)
```

```
## [1] 0.2615824
```

```r
plot(fit)
```


Residuals vs Fitted
Residuals
60000
20000
0
−40000
154
352
396
30000
32000
34000
36000
38000
40000
Fitted values
lm(Attendance ~ Home_top8)


Normal Q−Q
Standardized residuals
4
3
2
1
0
−1
−2
154
352
396
−3
−2
−1
0
1
2
3
Theoretical Quantiles
lm(Attendance ~ Home_top8)



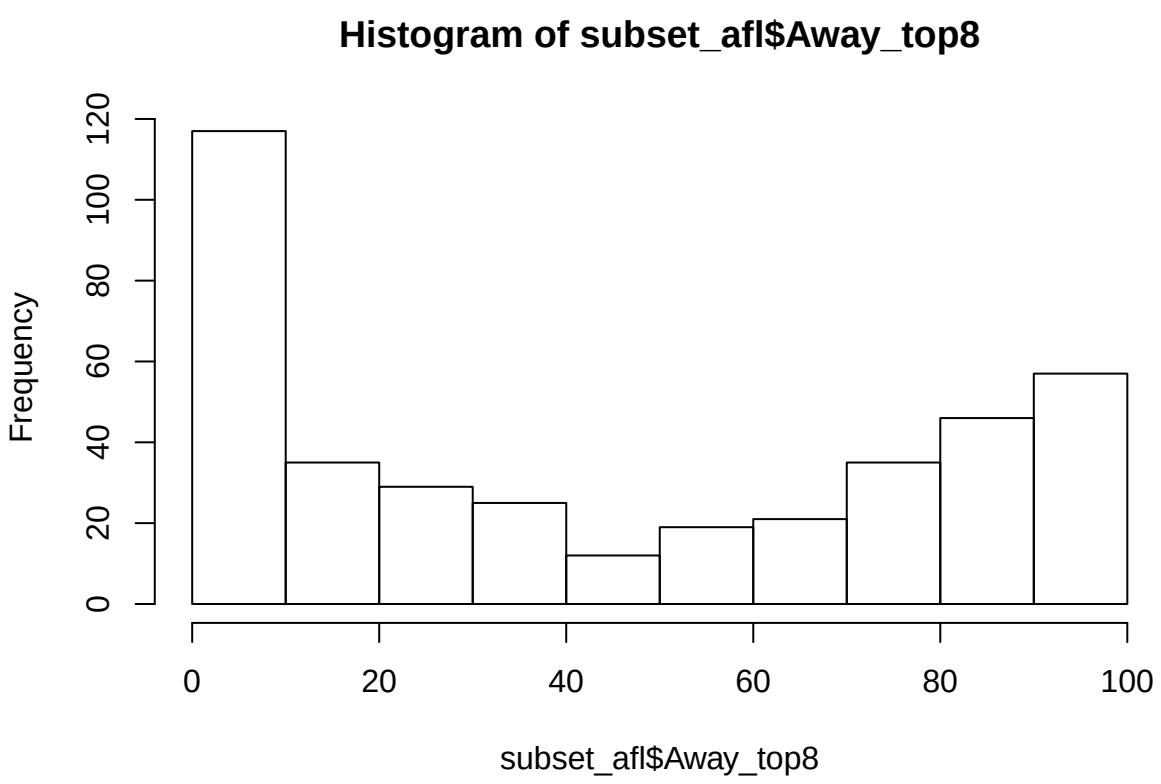

```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8, data = subset_afl)
##
```

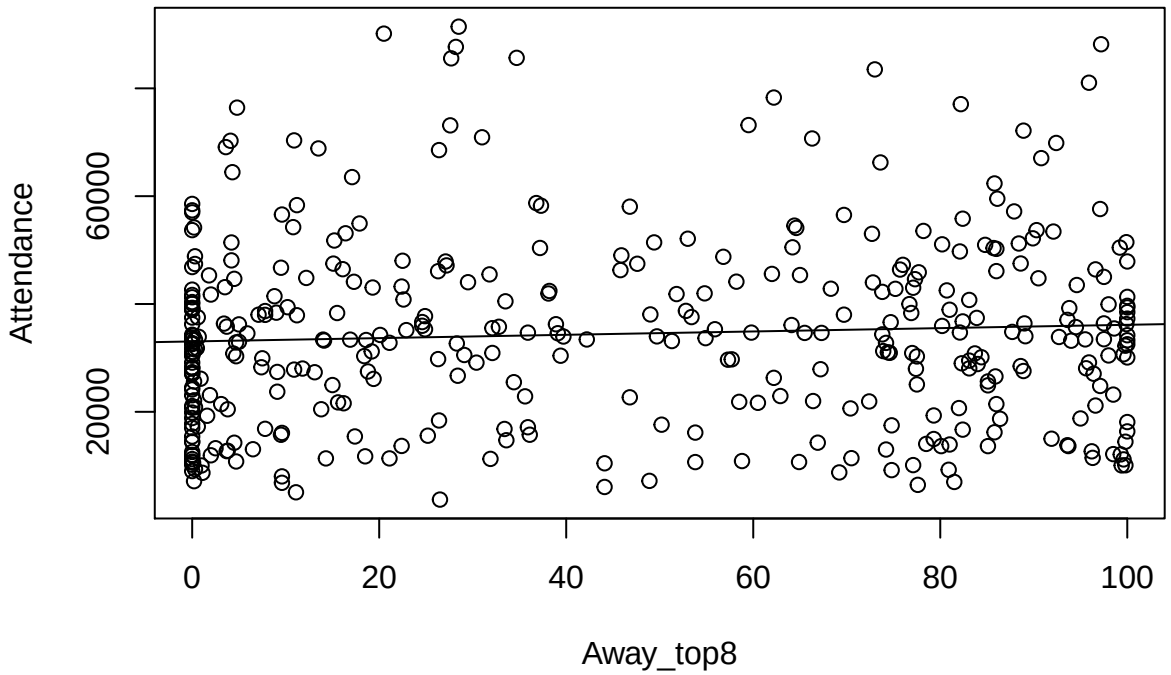
```
## Residuals:
##    Min     1Q Median     3Q    Max
## -30677 -10526  -1315   8122  57362
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 28847.26    1330.04   21.69  < 2e-16 ***
## Home_top8     124.03      23.06    5.38 1.28e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16630 on 394 degrees of freedom
## Multiple R-squared:  0.06843,    Adjusted R-squared:  0.06606
## F-statistic: 28.94 on 1 and 394 DF,  p-value: 1.283e-07
```

The data here is bimodal as the percent chance of a team being in the top 8 is often either 0 or 100. The plot doesn't show a great relationship between the two variables, but the correlation coefficient of 0.261 shows a weak correlation. The residuals vs. fitted shows that there are a few high-attendance games that have a strong pull on the linear model; this is due to specific games being attendance draws for reasons other than playoff probability (i.e. April 25, Round 1). The model shows that on average we can expect an increase in 124 fans for every 1 percent increase in playoff probability.

```
hist(subset_afl$Away_top8)
```



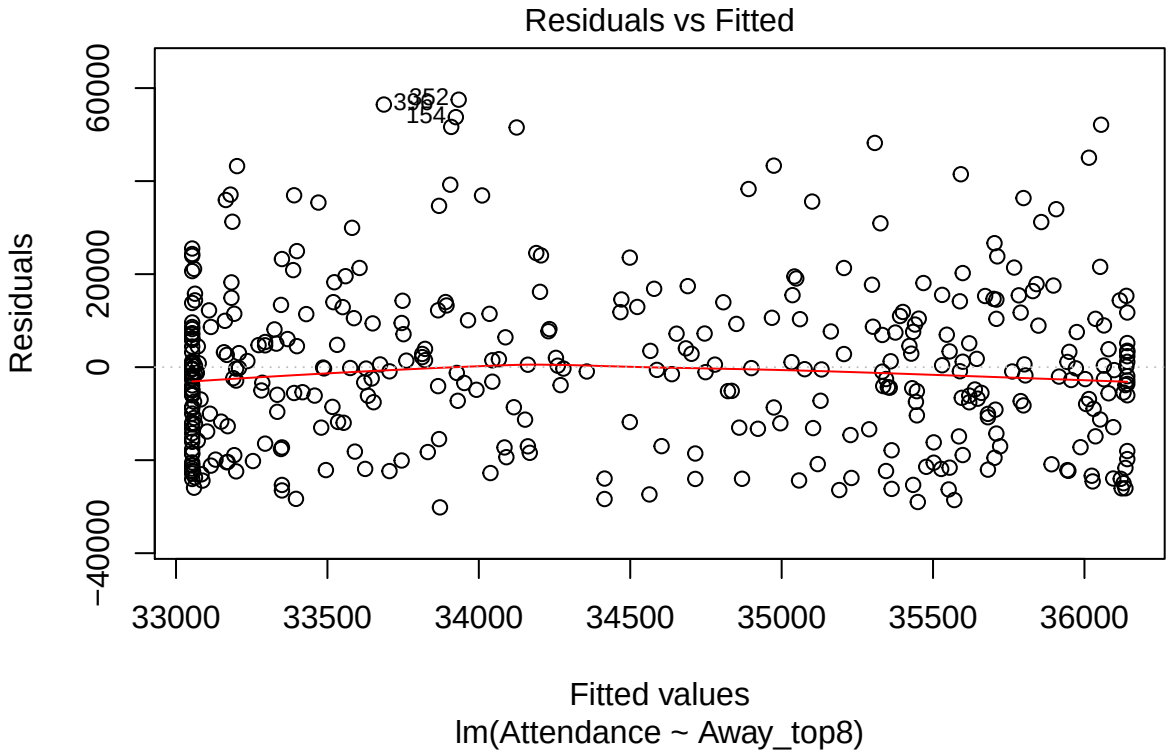

```
plot(Attendance~Away_top8, data = subset_afl)
fit <- lm(Attendance~Away_top8, data=subset_afl)
abline(fit)
```

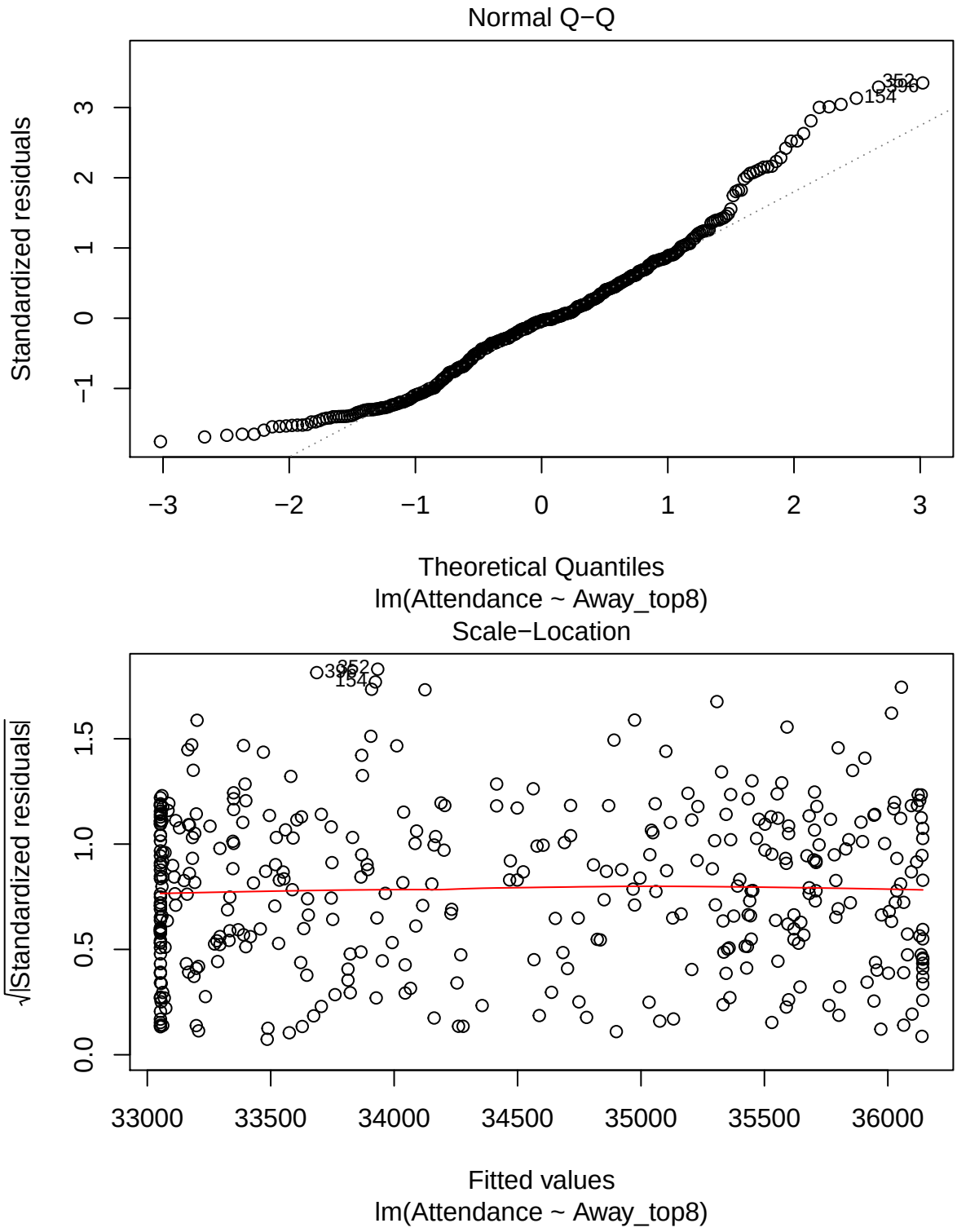


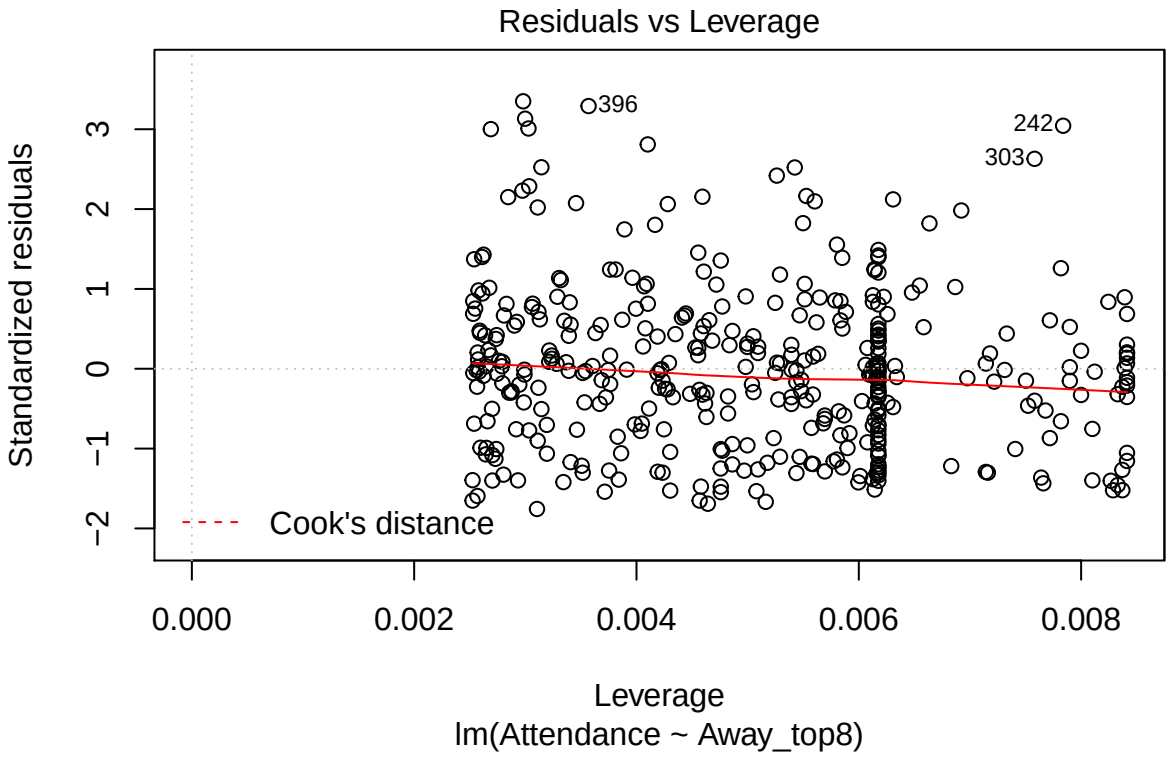

```
cor(subset_afl$Attendance, subset_afl$Away_top8)
```

```
## [1] 0.06583929
```

```
plot(fit)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Away_top8)
Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Away_top8)

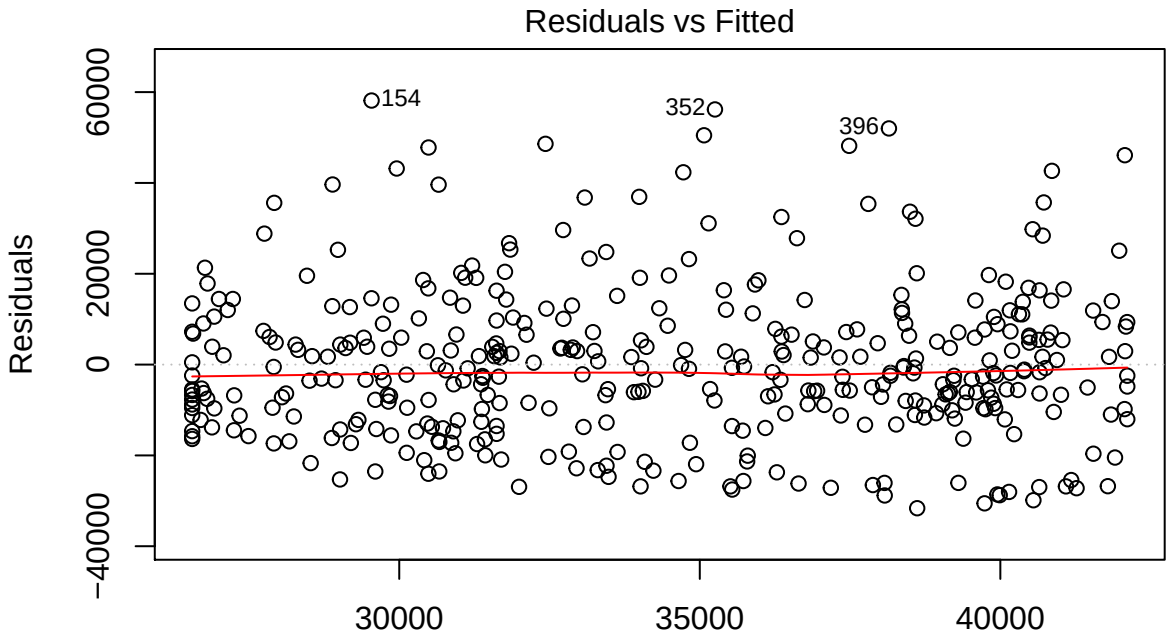


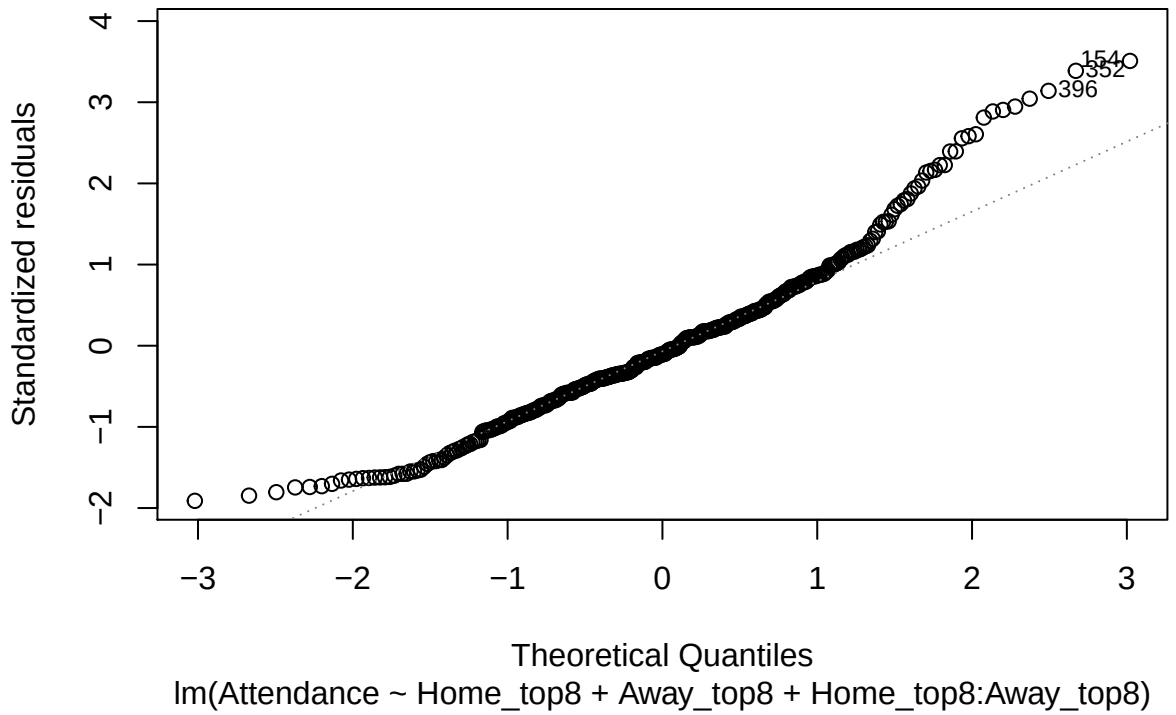

```
summary(fit)
```

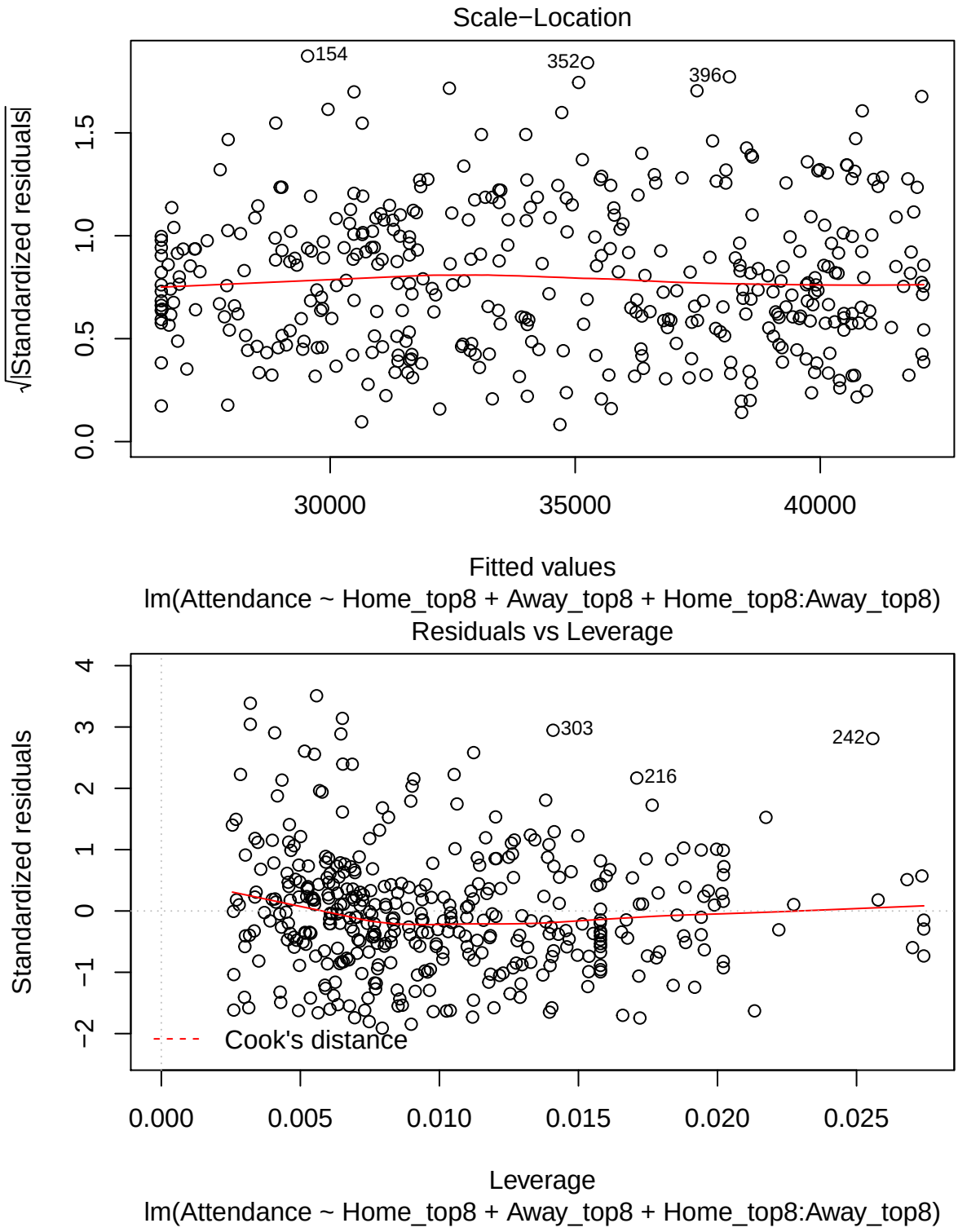
```
##
## Call:
## lm(formula = Attendance ~ Away_top8, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -30150 -12403   -616   9429  57507
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 33053.34    1351.05   24.46   <2e-16 ***
## Away_top8      30.88      23.58    1.31    0.191
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17190 on 394 degrees of freedom
## Multiple R-squared:  0.004335,   Adjusted R-squared:  0.001808
## F-statistic: 1.715 on 1 and 394 DF,  p-value: 0.1911
```

The Away playoff odds data shown here is heavily weighted towards 0, even moreso than the Home playoff odds, likely due to the fact that away teams have a greater chance of losing the upcoming game. The plot of Away playoff odds against Attendance appears to be a random scatter, and the correlation coefficient is .06 which means there is almost no correlation. The residuals vs. fitted shows that there are a few high-attendance games that have a strong pull on the linear model; this is due to specific games being attendance draws for reasons other than playoff probability (i.e. April 25, Round 1). The model shows that on average we can expect an increase in 31 fans for every 1 percent increase in playoff probability, however this is not a significant result (p-value of 0.191), so we fail to reject the null hypothesis that the Away team's chances of making the playoffs changes the number of fans that attend the game.

```
fit <- lm(Attendance~Home_top8+Away_top8+Home_top8:Away_top8, data=subset_afl)
plot(fit)
```








```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Away_top8 + Home_top8:Away_top8,
##     data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -31628 -10728  -1749   8418  58145
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)         26555.1342  2088.1337  12.717  < 2e-16 ***
## Home_top8             140.9749    35.4124   3.981 8.18e-05 ***
## Away_top8              50.6391    35.6336   1.421    0.156
## Home_top8:Away_top8    -0.3605     0.6073  -0.593    0.553
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16620 on 392 degrees of freedom
## Multiple R-squared:  0.07463,    Adjusted R-squared:  0.06755
## F-statistic: 10.54 on 3 and 392 DF,  p-value: 1.114e-06
```

```
fit1 <- step(fit)
```

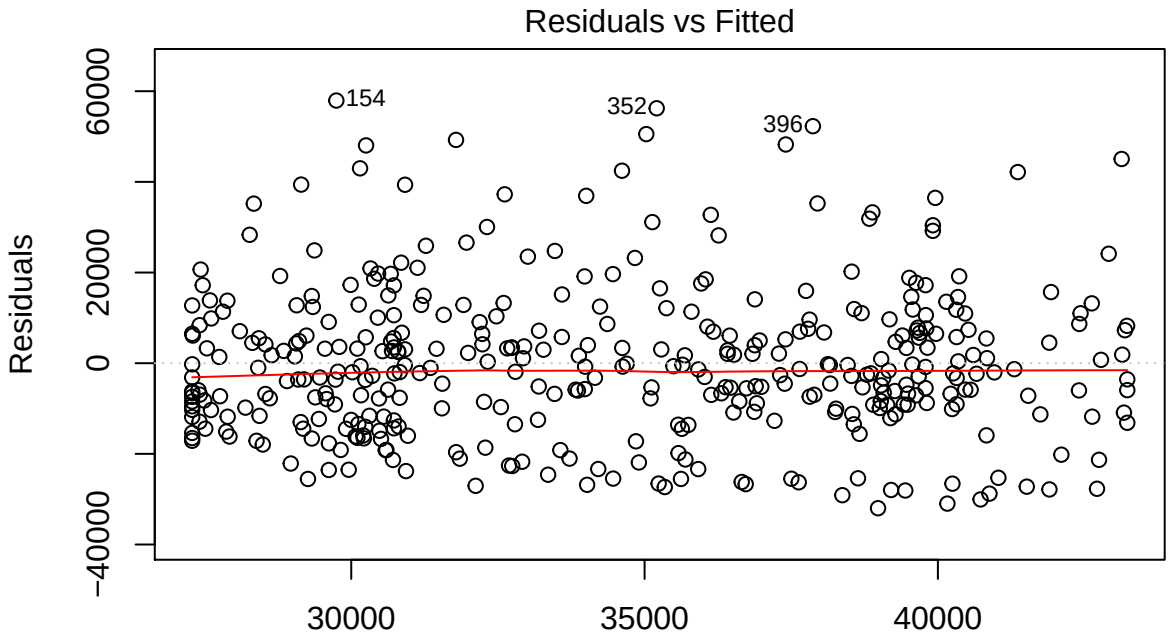
```
## Start:  AIC=7700.72
## Attendance ~ Home_top8 + Away_top8 + Home_top8:Away_top8
##
##                       Df Sum of Sq        RSS    AIC
## - Home_top8:Away_top8  1  97244683 1.0832e+11 7699.1
## <none>                              1.0822e+11 7700.7
##
## Step:  AIC=7699.07
## Attendance ~ Home_top8 + Away_top8
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    1.0832e+11 7699.1
## - Away_top8  1  628048258 1.0895e+11 7699.4
## - Home_top8  1 8123521074 1.1644e+11 7725.7
```

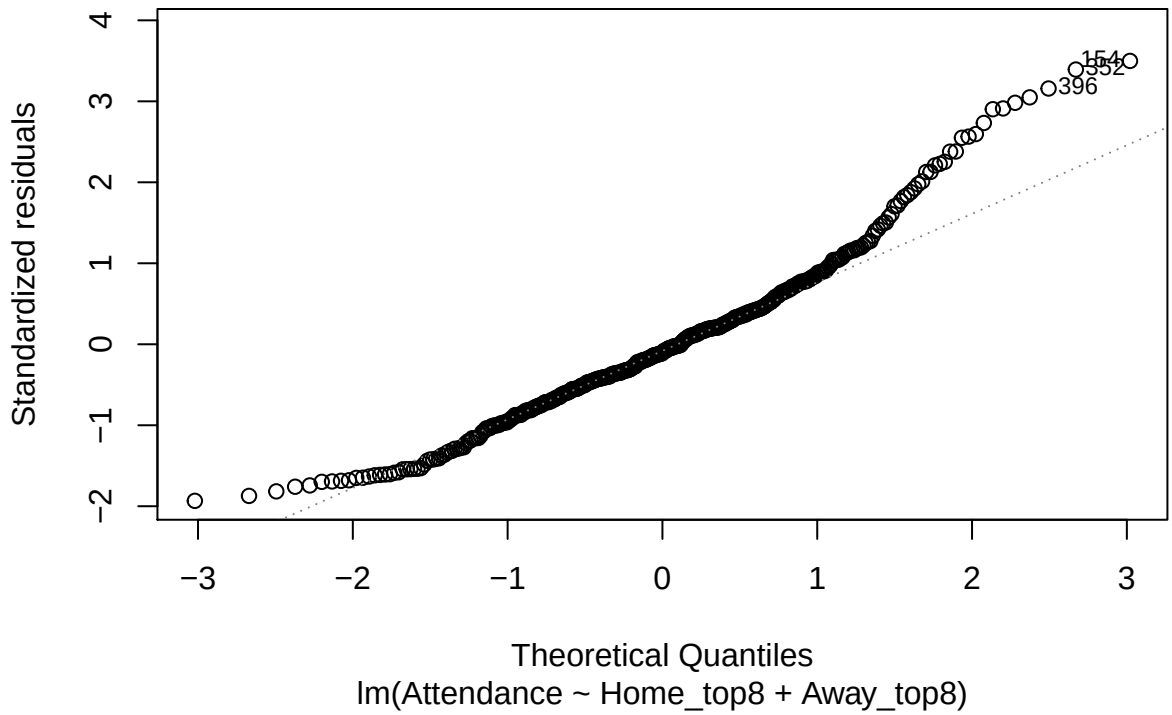
```
summary(fit1)
```

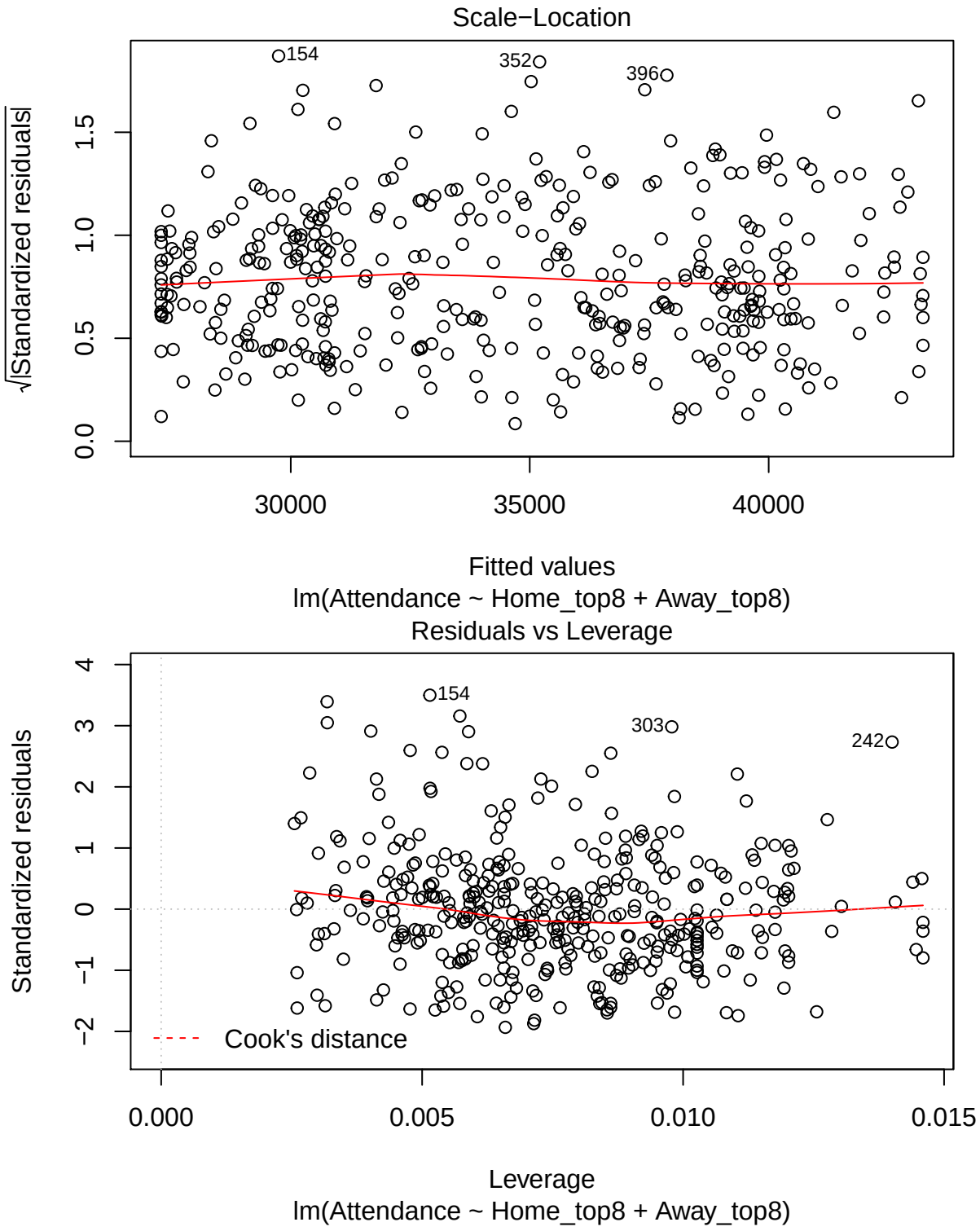
```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Away_top8, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -31990 -10796  -1751   8096  57939
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 27288.19    1682.27  16.221  < 2e-16 ***
## Home_top8     125.02      23.03   5.429 9.95e-08 ***
## Away_top8      34.39      22.78   1.510    0.132
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16600 on 393 degrees of freedom
## Multiple R-squared:  0.0738, Adjusted R-squared:  0.06908
```

```
## F-statistic: 15.66 on 2 and 393 DF,  p-value: 2.87e-07
```

```
plot(fit1)
```





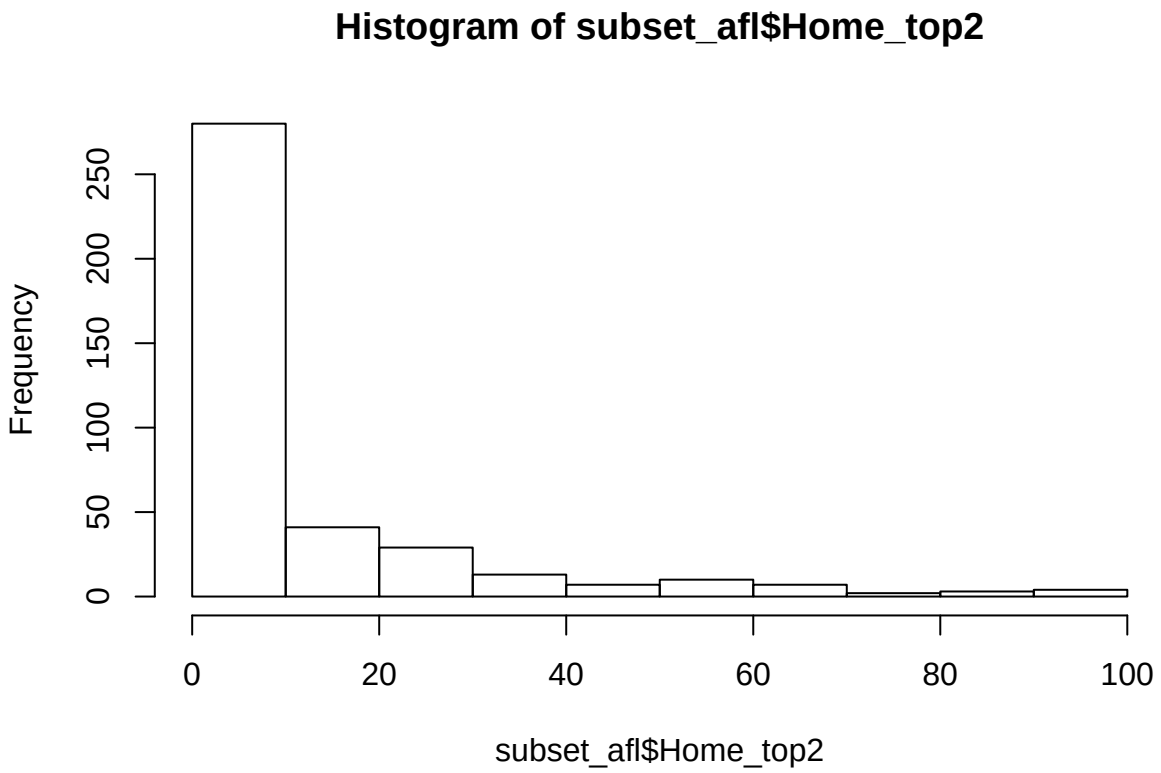


The residuals vs. fitted shows that there are a few high-attendance games that have a strong pull on the linear model; this is due to specific games being attendance draws for reasons other than playoff probability (i.e. April 25, Round 1). This means that we should look to see if there are other factors we can add to the model that might help to explain some of this variability. The model shows that on average we can expect an increase of 141 fans for every 1 percent increase in Home probability of making the playoffs and an increase of 50 fans for every 1 percent increase in Away probability of making the playoffs, although Away probability

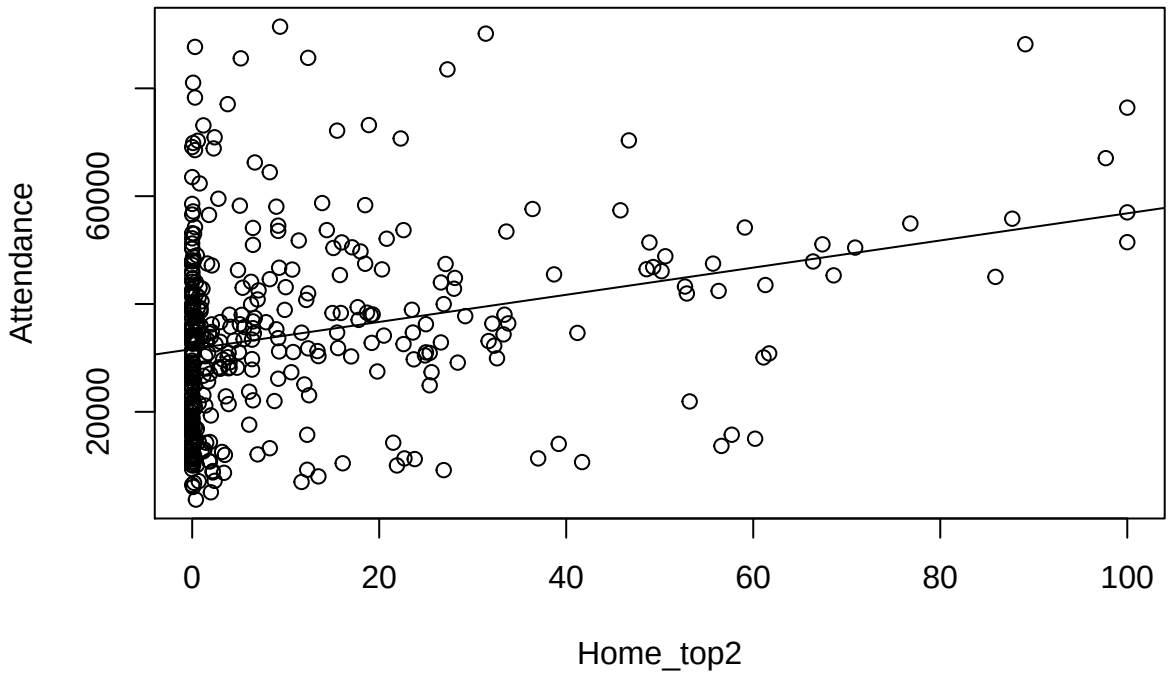
is not a significant result (p-value of 0.156), and the interaction term has a p-value of 0.553, bad enough for us to remove it from the model when using stepwise variable selection. Using the reduced model, the increase due to Home playoff probability drops to 125 fans per 1 percent increase while the increase due to Away playoff probability drops to 34.4 fans per 1 percent increase. The p-values are lower, with Home playoff probability remaining near zero while Away playoff probability has a p-value of 0.132, closer to being statistically significant and important enough to remain in the stepwise model.

Now we will use a similar process for championship appearance probability.

```
hist(subset_afl$Home_top2)
```



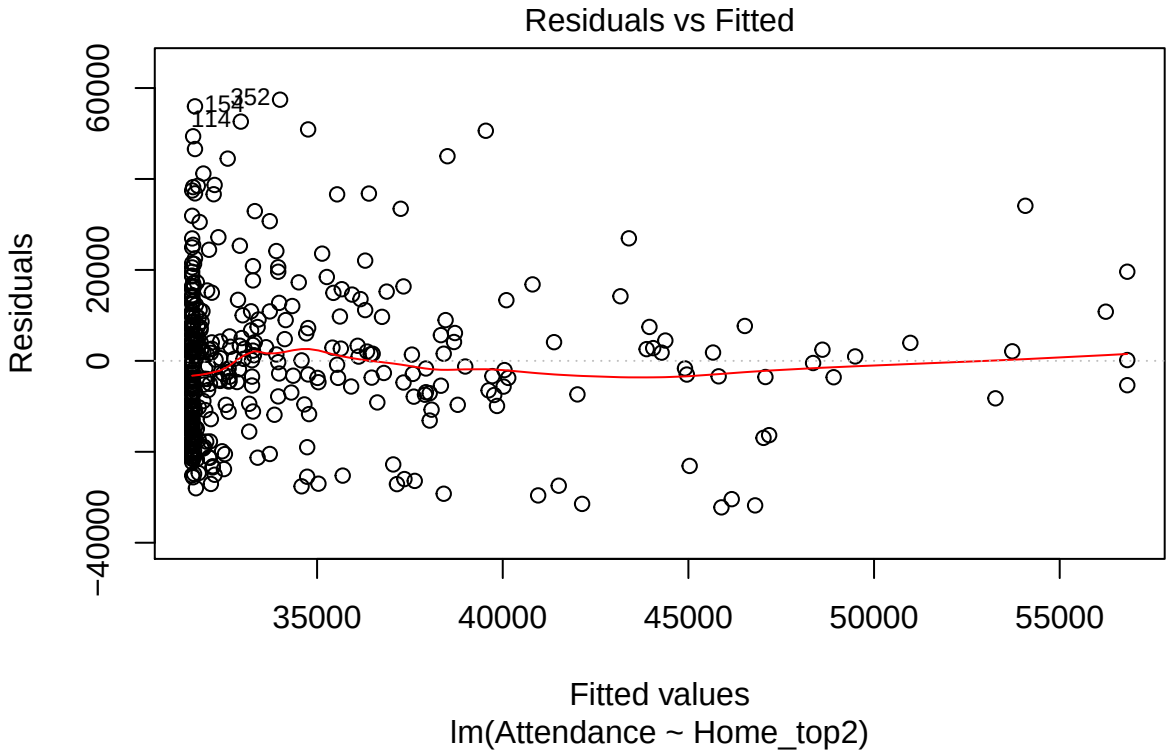

```
plot(Attendance~Home_top2, data = subset_afl)
fit <- lm(Attendance~Home_top2, data=subset_afl)
abline(fit)
```

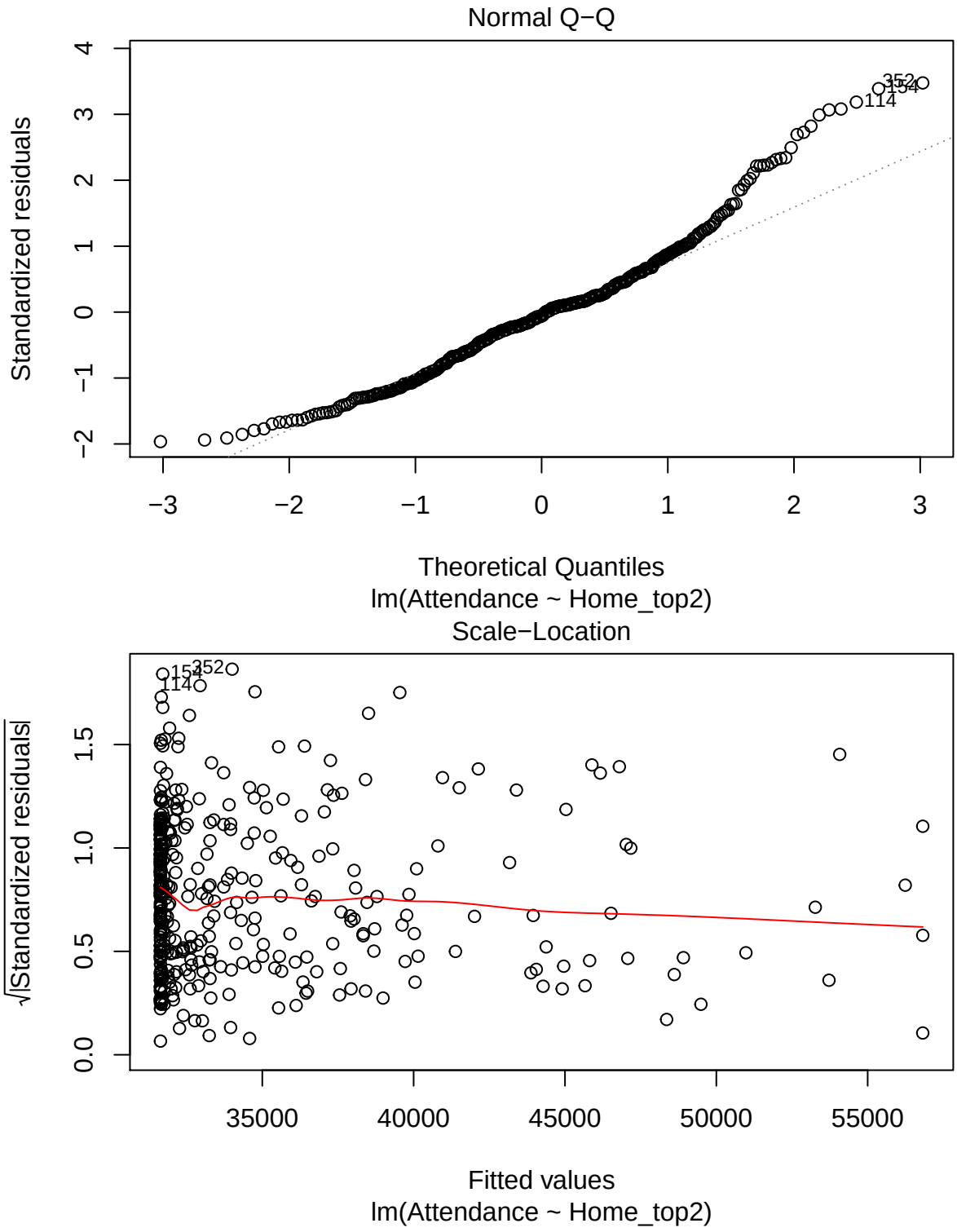


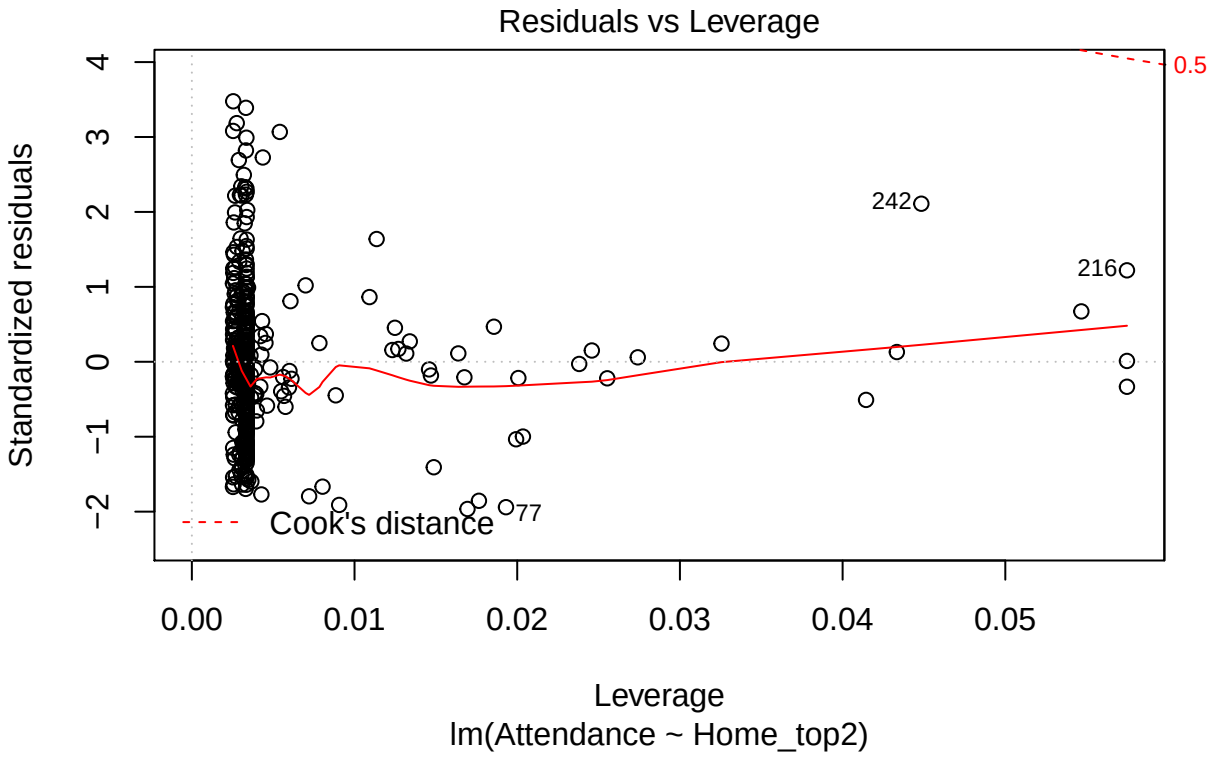

```
cor(subset_afl$Attendance, subset_afl$Home_top2)
```

```
## [1] 0.2794917
```

```
plot(fit)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2)
Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Home_top2)




```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -32219 -11011   -918   7798  57437
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 31636.13     960.27  32.945  < 2e-16 ***
## Home_top2     251.84      43.59   5.778 1.54e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16540 on 394 degrees of freedom
## Multiple R-squared:  0.07812,    Adjusted R-squared:  0.07578
## F-statistic: 33.39 on 1 and 394 DF,  p-value: 1.537e-08
```

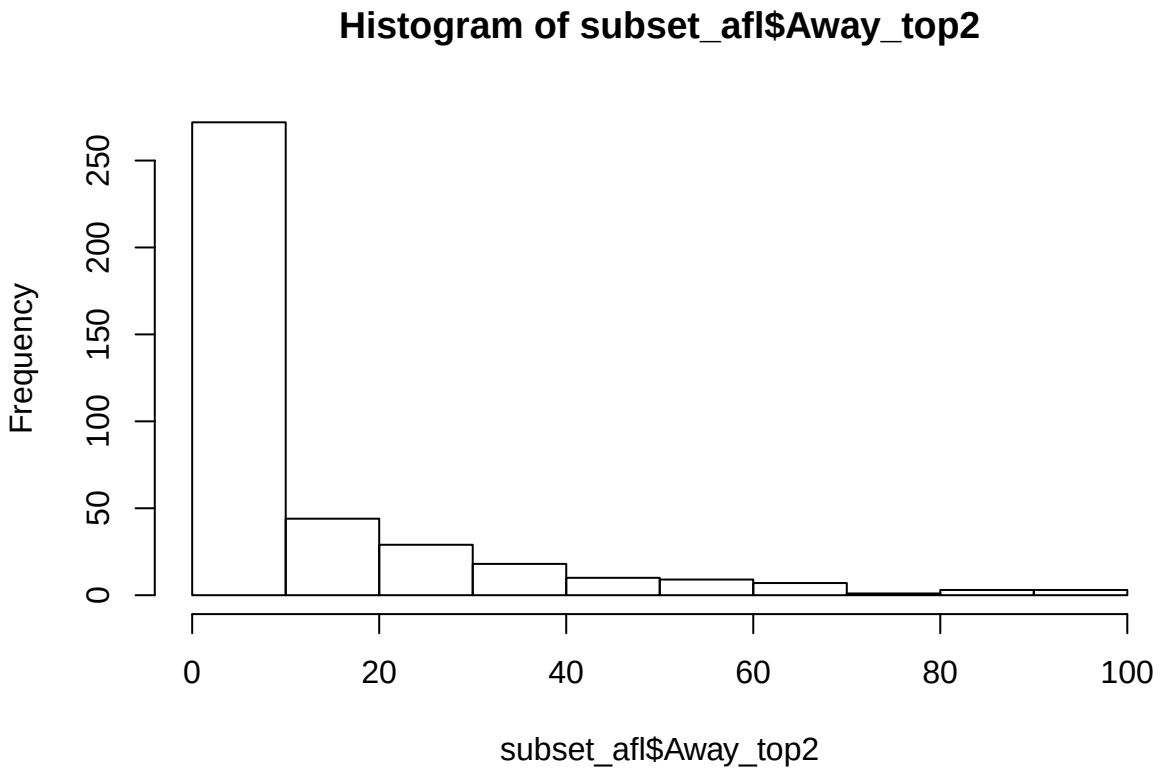
```
nrow(subset_afl[subset_afl$Home_top2==0,])/nrow(subset_afl)
```

```
## [1] 0.2954545
```

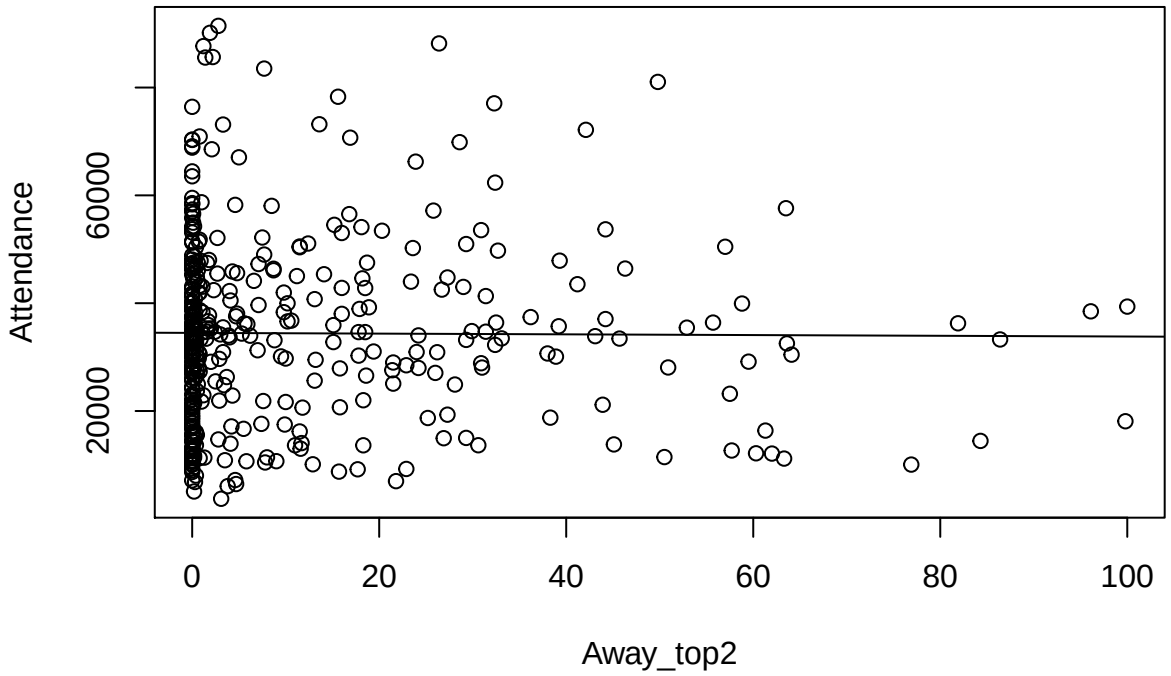
The data here is very right skewed because the chances of making the championship are generally very low. The plot doesn't show a great relationship between Attendance and Home championship probability, but the correlation coefficient of 0.279 shows a weak correlation. The residuals vs. fitted shows that there are a few high-attendance games that have a strong pull on the linear model; this is due to specific games being attendance draws for reasons other than playoff probability (i.e. April 25, Round 1). The model shows that on average we can expect an increase in 252 fans for every 1 percent increase in championship appearance

probability. We considered a log transformation but due to the high number of zeroes (about 29.5 percent of the data points) this is likely futile. We could consider a zero-inflated model for further analysis.

```
hist(subset_afl$Away_top2)
```



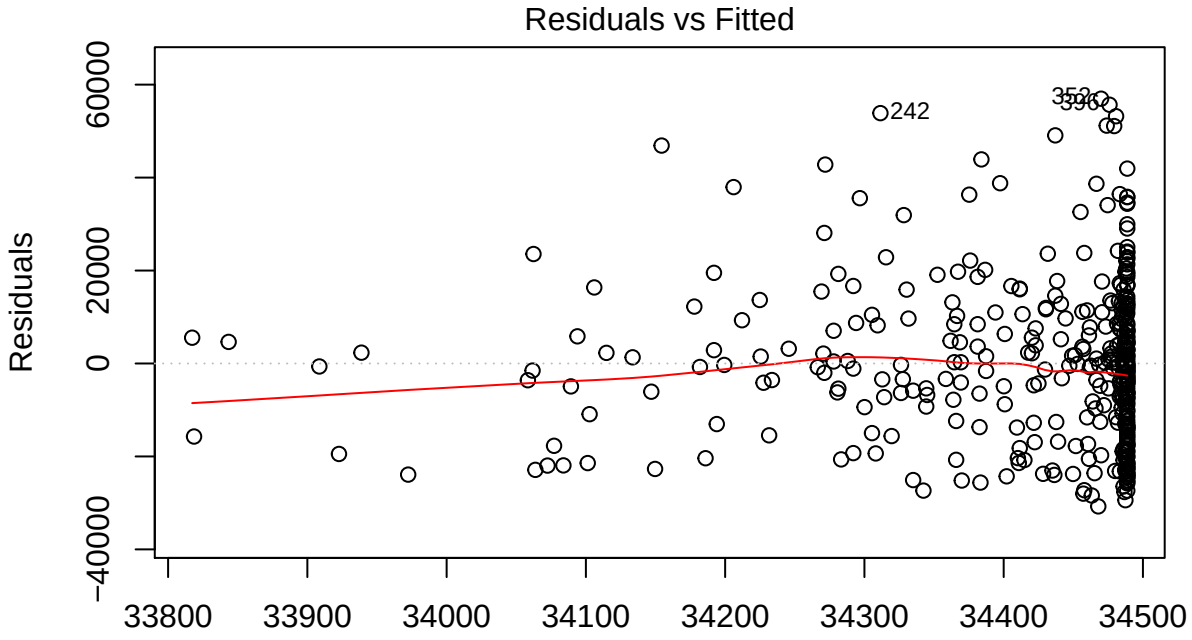

```
plot(Attendance~Away_top2, data = subset_afl)
fit <- lm(Attendance~Away_top2, data=subset_afl)
abline(fit)
```

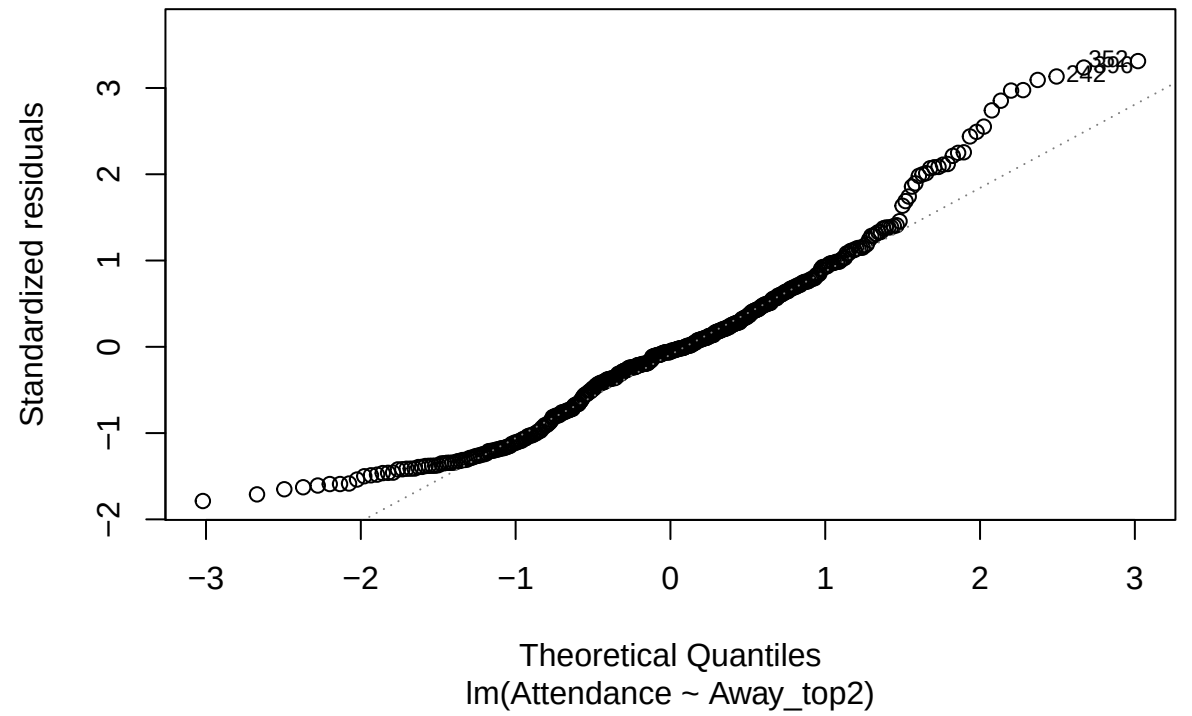


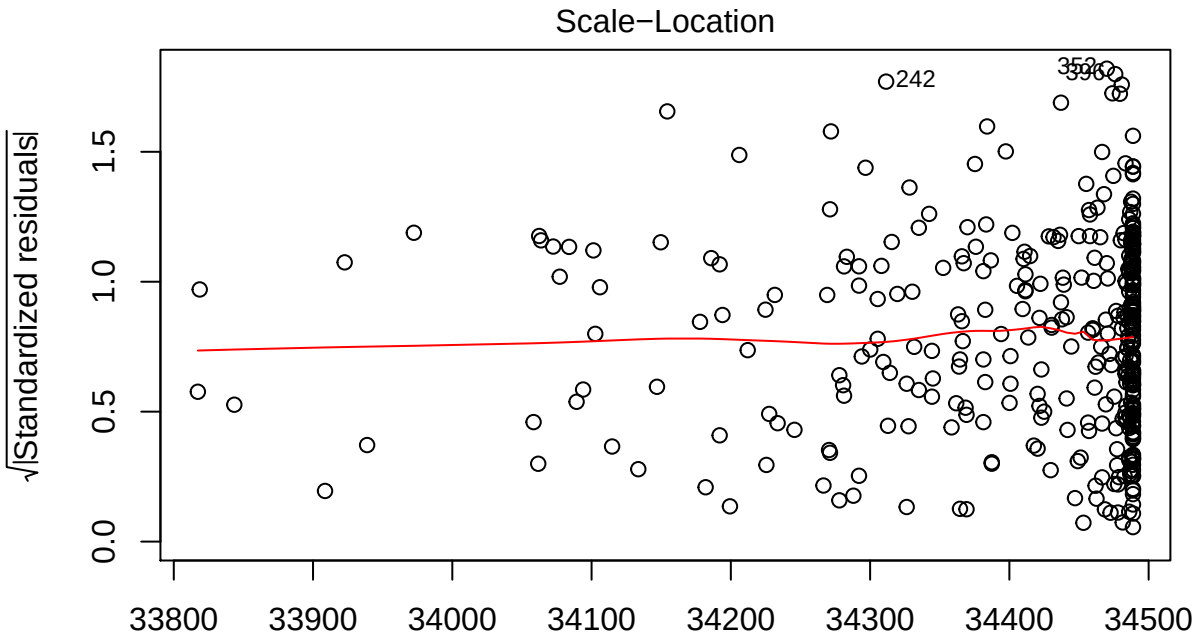
```r
cor(subset_afl$Attendance, subset_afl$Away_top2)
```

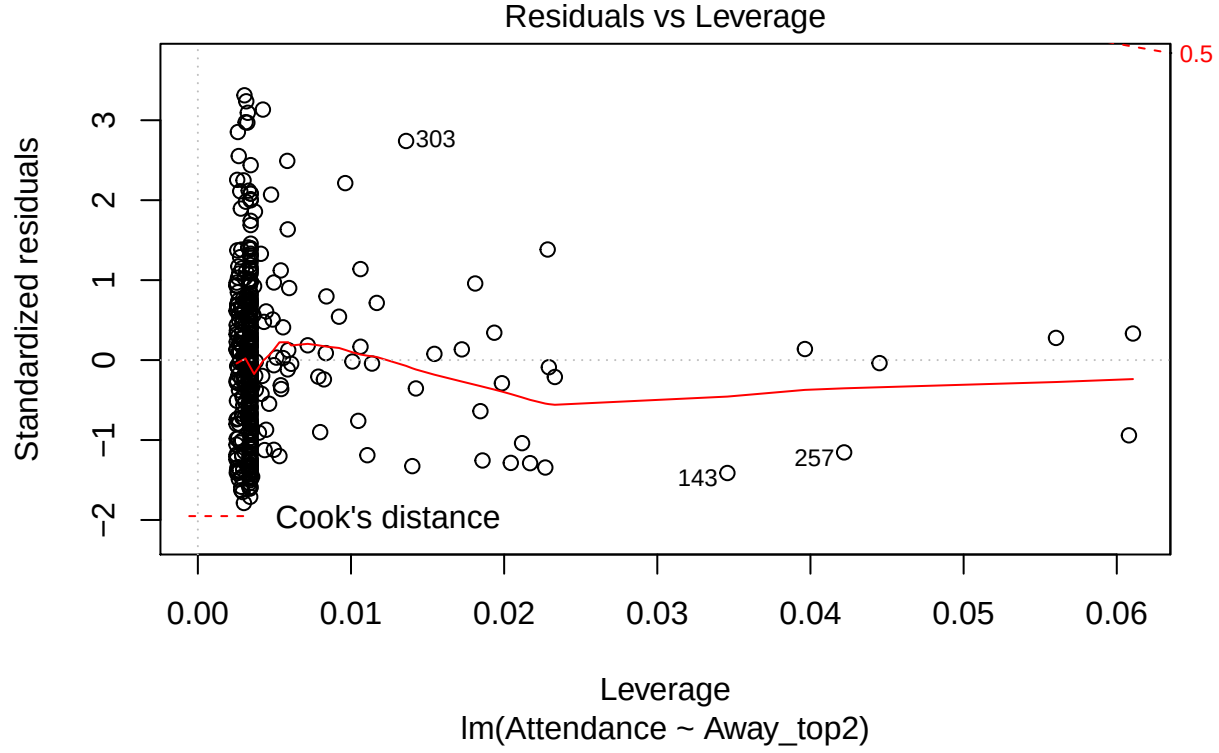
```
## [1] -0.007209847
```

```r
plot(fit)
```










```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Away_top2, data = subset_afl)
##
```

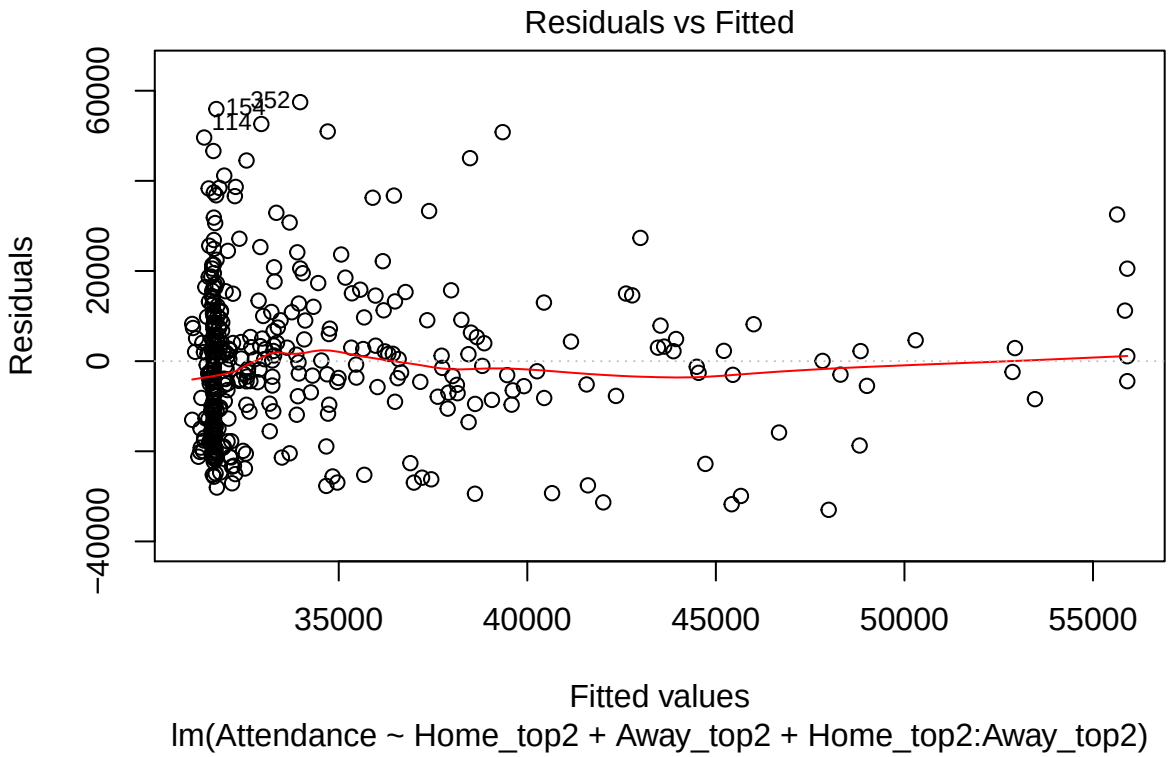
```
## Residuals:
##    Min     1Q Median     3Q    Max
## -30746 -12789   -929   9640  56970
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 34488.758   1011.819  34.086   <2e-16 ***
## Away_top2      -6.715     46.918  -0.143    0.886
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17230 on 394 degrees of freedom
## Multiple R-squared:  5.198e-05,  Adjusted R-squared:  -0.002486
## F-statistic: 0.02048 on 1 and 394 DF,  p-value: 0.8863
```

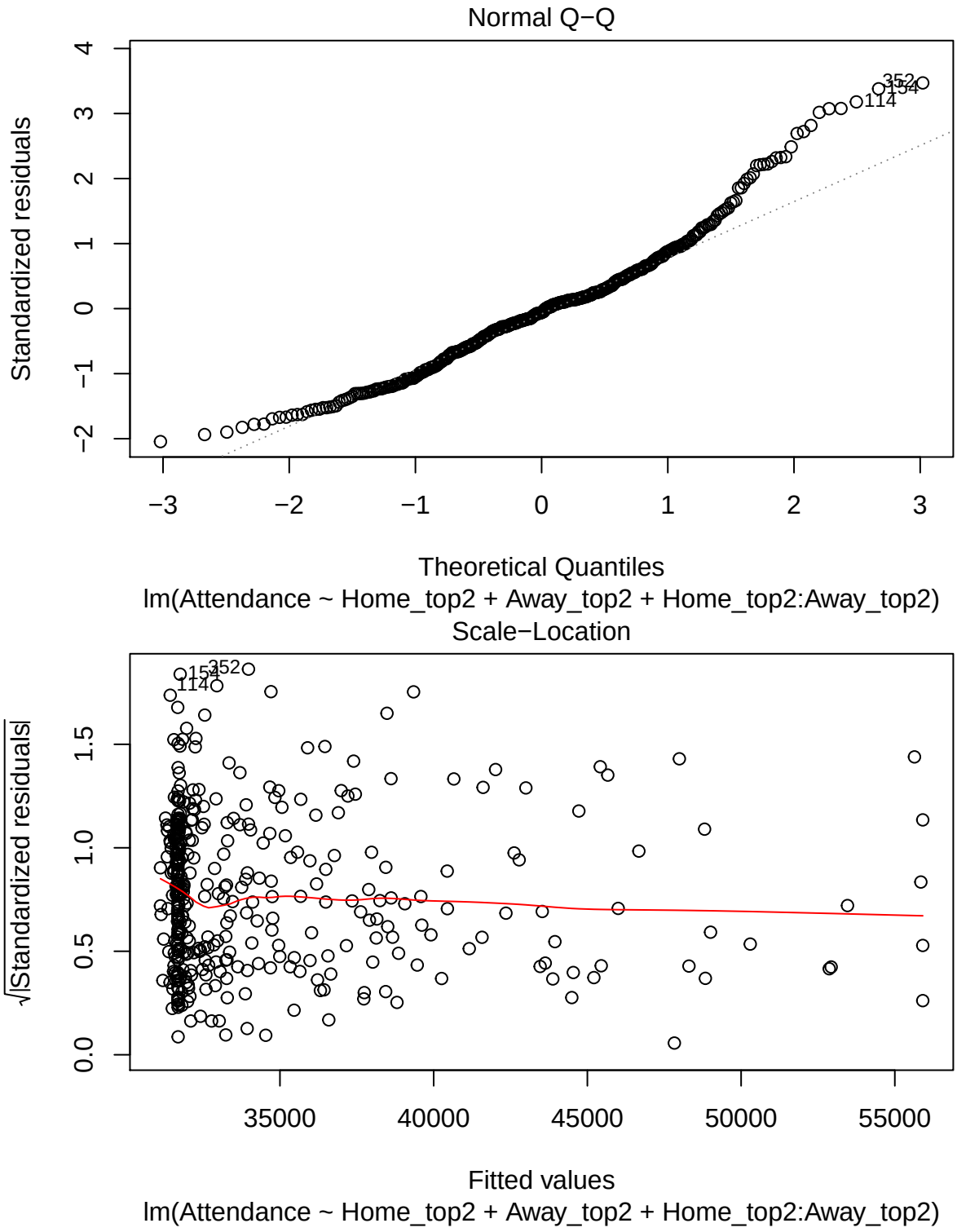
```
nrow(subset_afl[subset_afl$Away_top2==0,])/nrow(subset_afl)
```

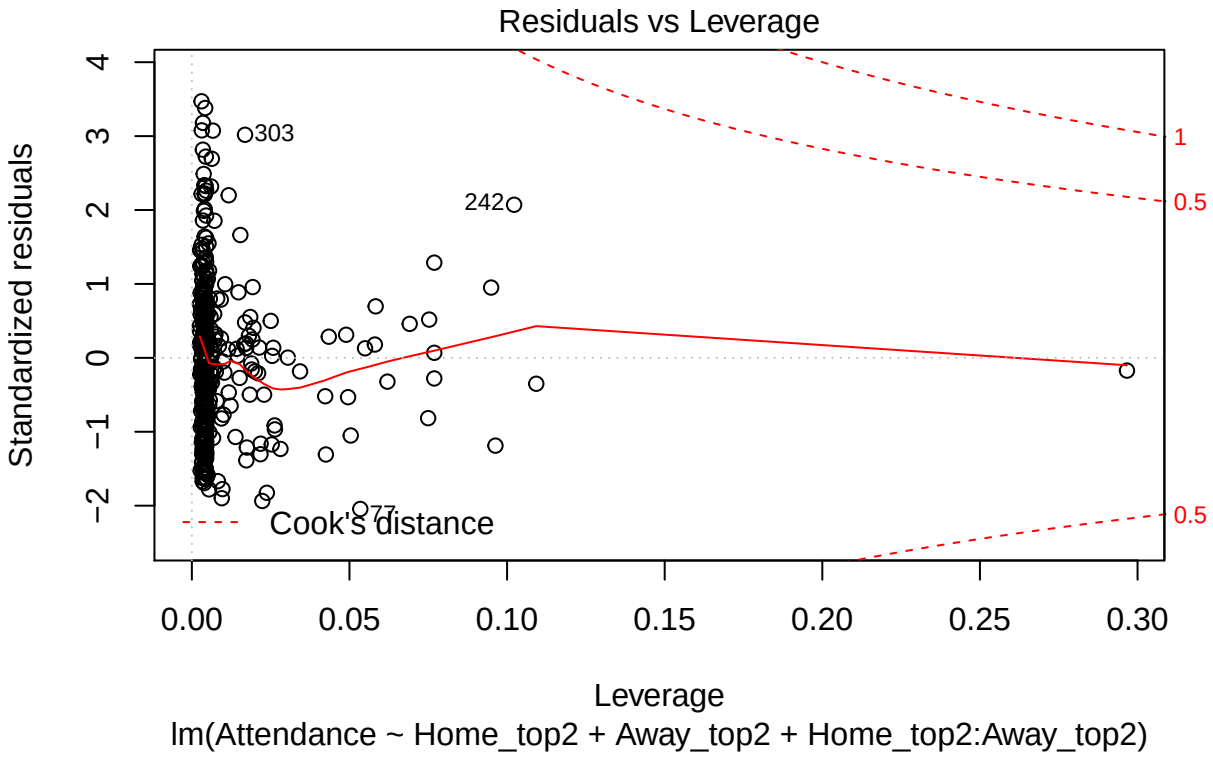
```
## [1] 0.2979798
```

The Away playoff odds data shown here is heavily weighted towards 0 as championship appearance probability is generally very low. The plot of Away playoff odds against Attendance appears to be a random scatter, and the correlation coefficient is -0.007 which means there is no correlation. The residuals vs. fitted still shows larger positive than negative residuals. The model shows that on average we can expect a decrease of 7 fans for every 1 percent increase in Away playoff probability, however this result is meaningless with a p-value of 0.88.

```
fit <- lm(Attendance~Home_top2+Away_top2+Home_top2:Away_top2, data=subset_afl)
plot(fit)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Away_top2 + Home_top2:Away_top2)
Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Home_top2 + Away_top2 + Home_top2:Away_top2)




```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Away_top2 + Home_top2:Away_top2,
##     data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -32985 -10985   -964   8202  57463
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)         31688.129   1115.325  28.412  < 2e-16 ***
## Home_top2             242.208     50.335   4.812 2.14e-06 ***
## Away_top2              -5.746     50.090  -0.115    0.909
## Home_top2:Away_top2     1.072      2.750   0.390    0.697
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16580 on 392 degrees of freedom
## Multiple R-squared:  0.07848,    Adjusted R-squared:  0.07143
## F-statistic: 11.13 on 3 and 392 DF,  p-value: 5.035e-07
```

```
fit1 <- step(fit)
```

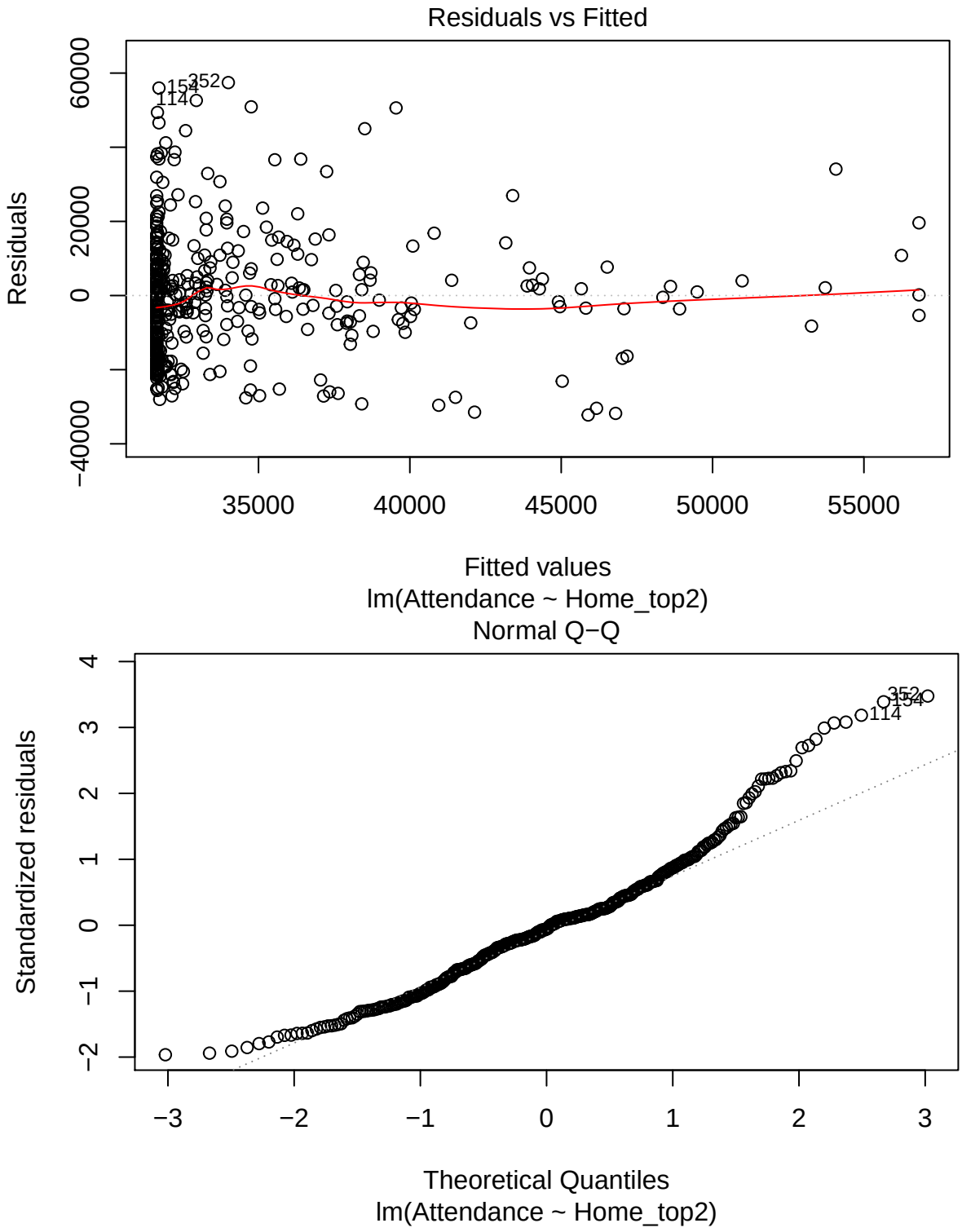
```
## Start:  AIC=7699.06
## Attendance ~ Home_top2 + Away_top2 + Home_top2:Away_top2
##
##                        Df Sum of Sq        RSS    AIC
## - Home_top2:Away_top2  1  41775323 1.0781e+11 7697.2
```

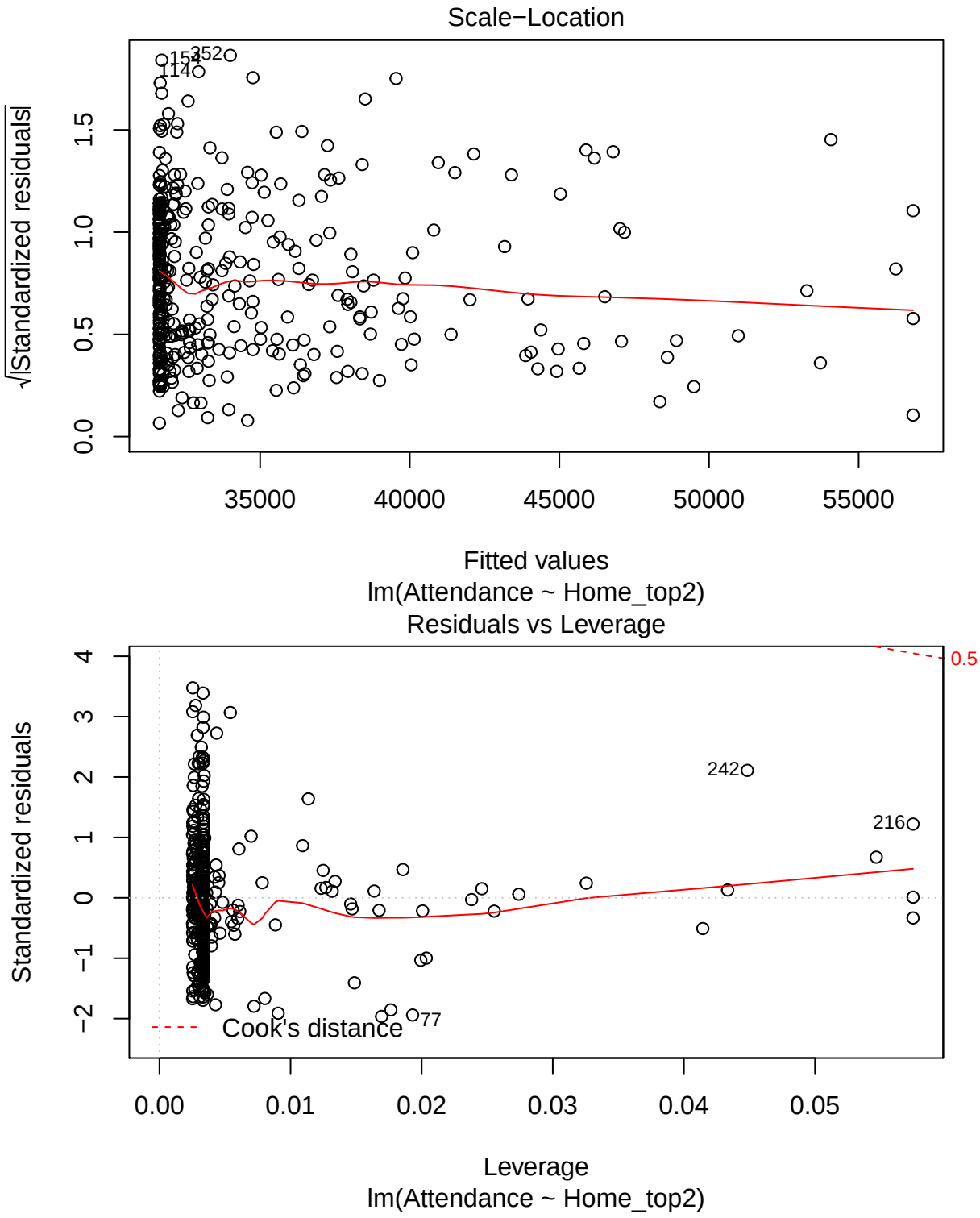
```
## <none>                             1.0777e+11 7699.1
##
## Step:  AIC=7697.22
## Attendance ~ Home_top2 + Away_top2
##
##             Df  Sum of Sq        RSS    AIC
## - Away_top2  1     967634 1.0782e+11 7695.2
## <none>                    1.0781e+11 7697.2
## - Home_top2  1 9130608430 1.1695e+11 7727.4
##
## Step:  AIC=7695.22
## Attendance ~ Home_top2
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    1.0782e+11 7695.2
## - Home_top2  1 9135720145 1.1695e+11 7725.4
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -32219 -11011   -918   7798  57437
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 31636.13     960.27  32.945  < 2e-16 ***
## Home_top2     251.84      43.59   5.778 1.54e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16540 on 394 degrees of freedom
## Multiple R-squared:  0.07812,    Adjusted R-squared:  0.07578
## F-statistic: 33.39 on 1 and 394 DF,  p-value: 1.537e-08
```

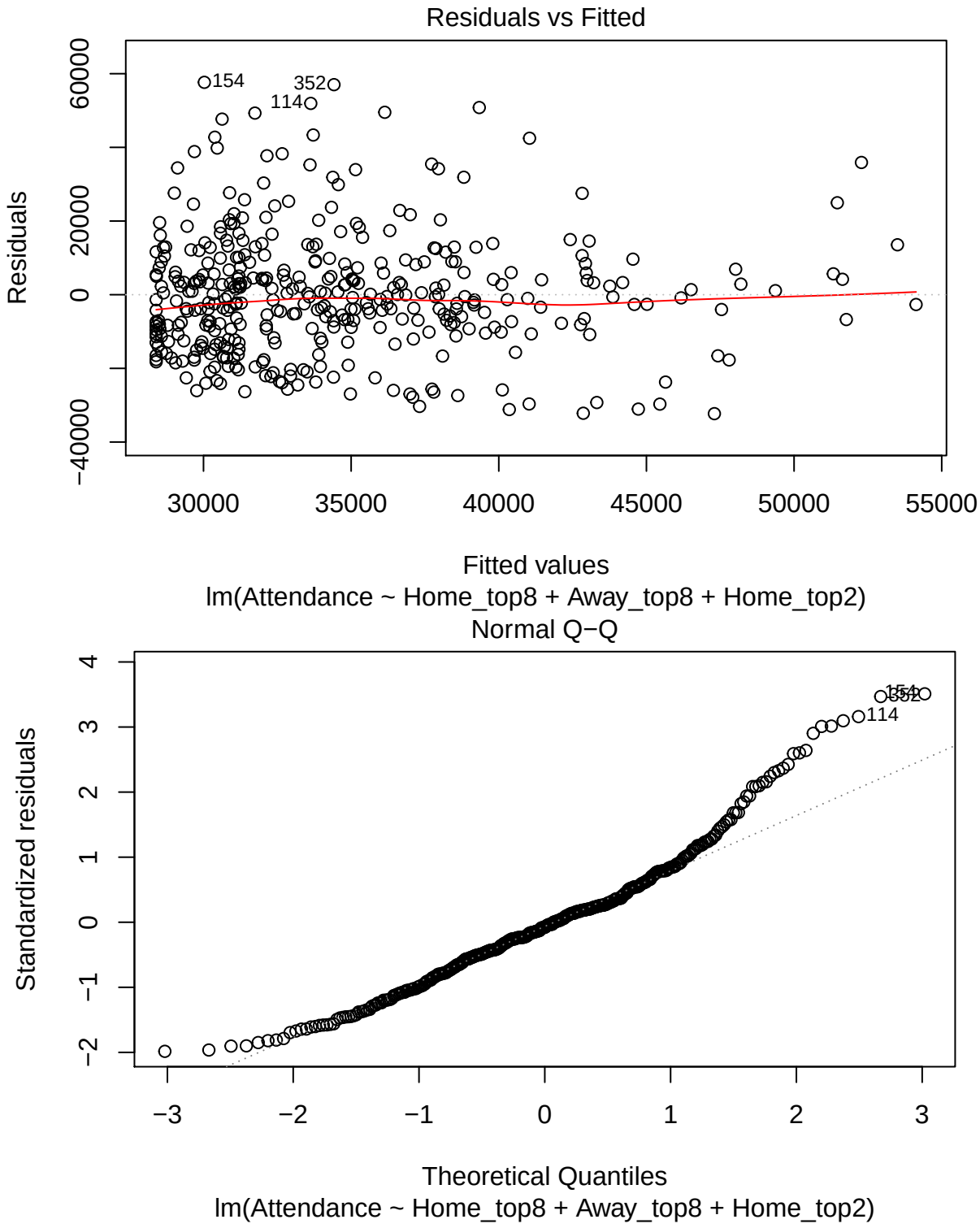
```
plot(fit1)
```

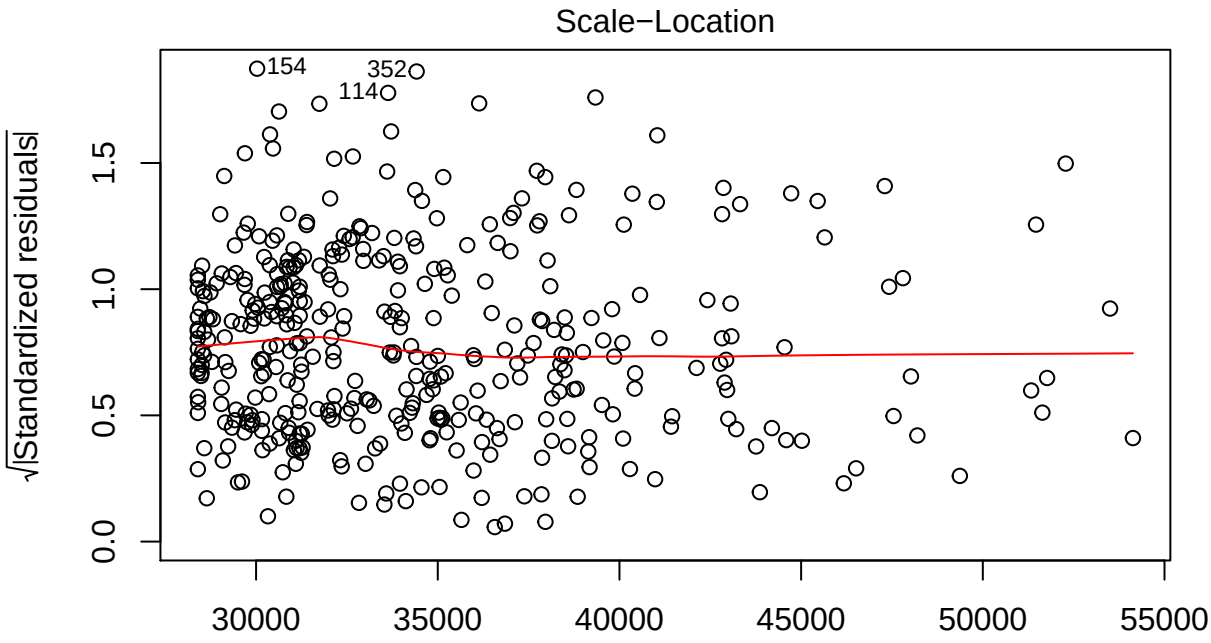

Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_top2)
352
154
114
60000
20000
0
−40000
35000
40000
45000
50000
55000
Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2)
4
3
2
1
0
−1
−2
−3

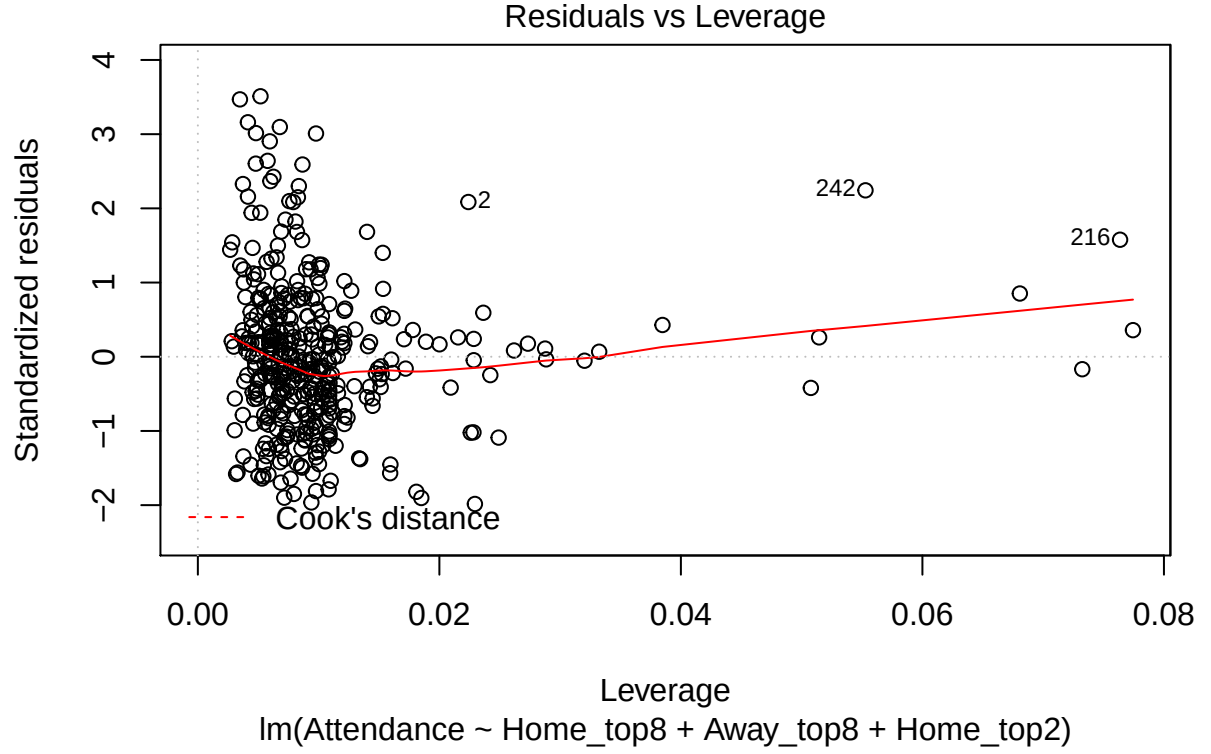


There is no need to run a model with additional interactions. We can see here that only Home championship probability affects the model. We can try to run the model in conjunction with the playoff probability model.

```
fit <- lm(Attendance~Home_top8+Away_top8+Home_top2, data=subset_afl)
plot(fit)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_top8 + Away_top8 + Home_top2)
Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top8 + Away_top8 + Home_top2)
154
352
114






```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Away_top8 + Home_top2,
##     data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -32294 -10479  -1268   8369  57657
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 28397.12    1717.14  16.537  < 2e-16 ***
## Home_top8      66.52      31.29   2.126  0.03415 *
## Away_top8      28.05      22.71   1.235  0.21763
## Home_top2     162.81      59.54   2.735  0.00653 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16470 on 392 degrees of freedom
## Multiple R-squared:  0.09113,    Adjusted R-squared:  0.08418
## F-statistic:  13.1 on 3 and 392 DF,  p-value: 3.596e-08
```

```
fit1 <- step(fit)
```

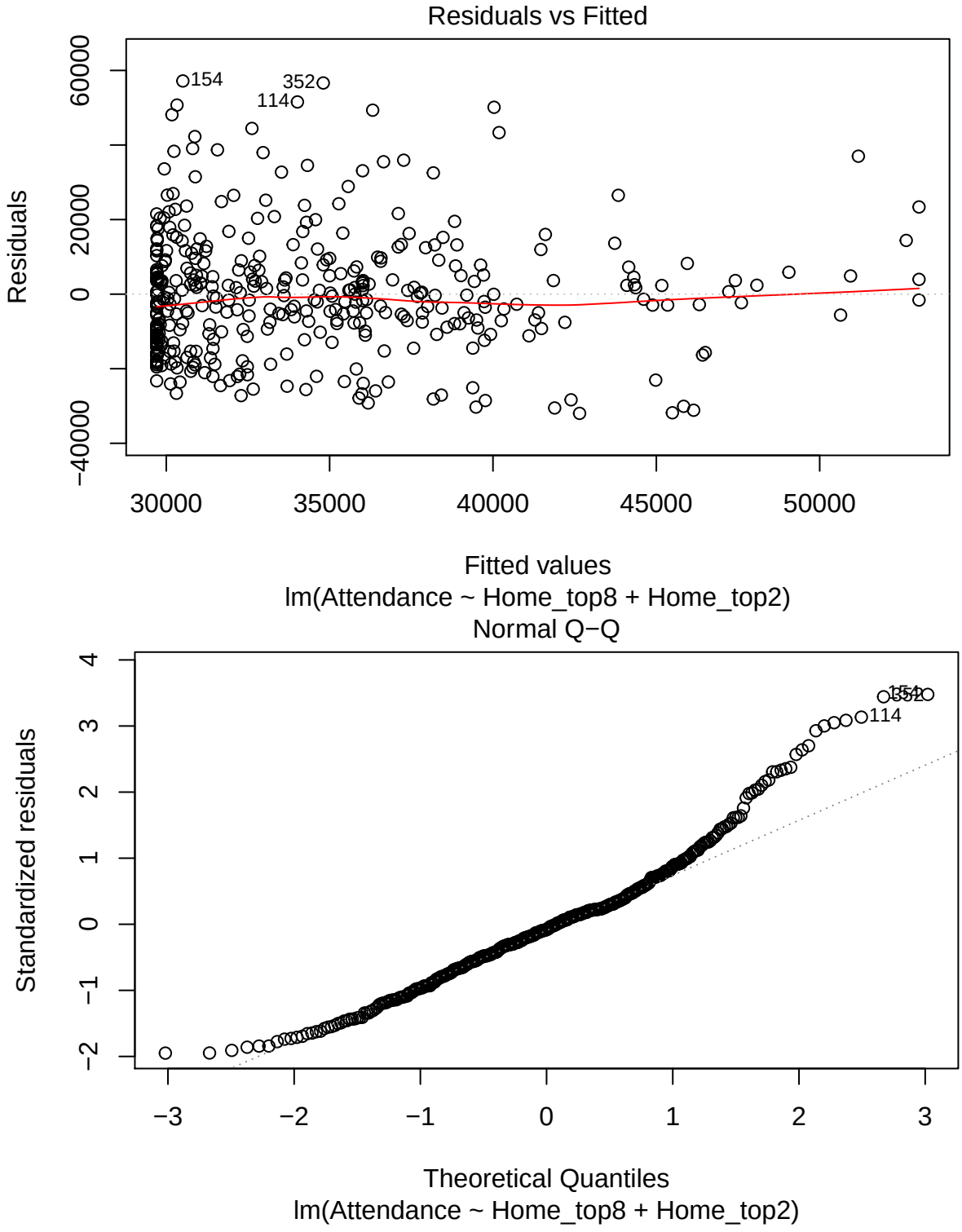
```
## Start:  AIC=7693.59
## Attendance ~ Home_top8 + Away_top8 + Home_top2
##
##             Df  Sum of Sq        RSS    AIC
## - Away_top8  1  413462857 1.0671e+11 7693.1
## <none>                    1.0629e+11 7693.6
## - Home_top8  1 1225241733 1.0752e+11 7696.1
## - Home_top2  1 2027760750 1.0832e+11 7699.1
##
## Step:  AIC=7693.13
## Attendance ~ Home_top8 + Home_top2
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    1.0671e+11 7693.1
## - Home_top8  1 1109060582 1.0782e+11 7695.2
## - Home_top2  1 2242346151 1.0895e+11 7699.4
```

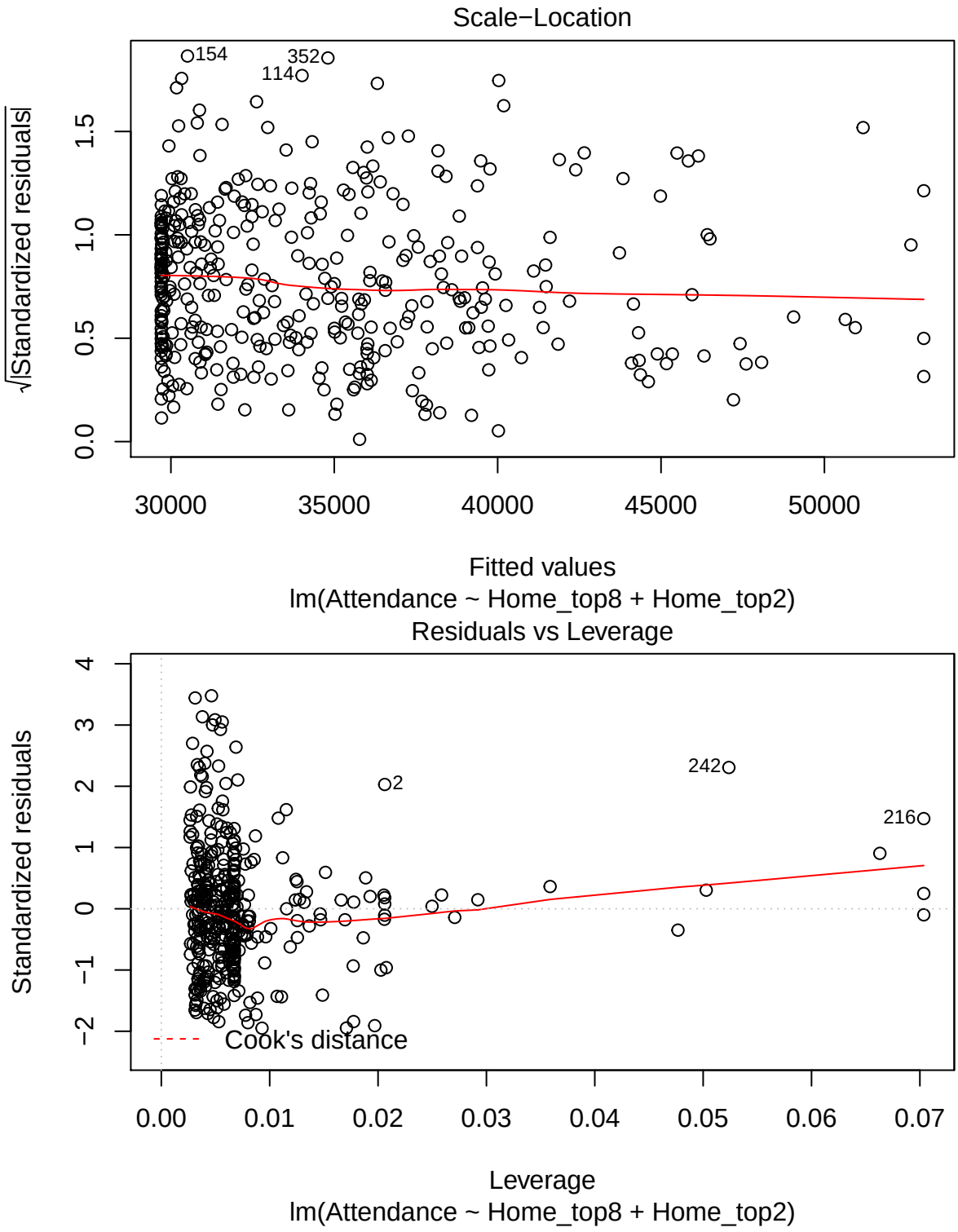
```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Home_top2, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -31974 -10799  -1315   7690  57177
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 29706.62    1351.45  21.981  < 2e-16 ***
## Home_top8      63.03      31.19   2.021  0.04395 *
## Home_top2     170.31      59.26   2.874  0.00428 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 16480 on 393 degrees of freedom
## Multiple R-squared:  0.0876, Adjusted R-squared:  0.08296
## F-statistic: 18.87 on 2 and 393 DF,  p-value: 1.502e-08
```

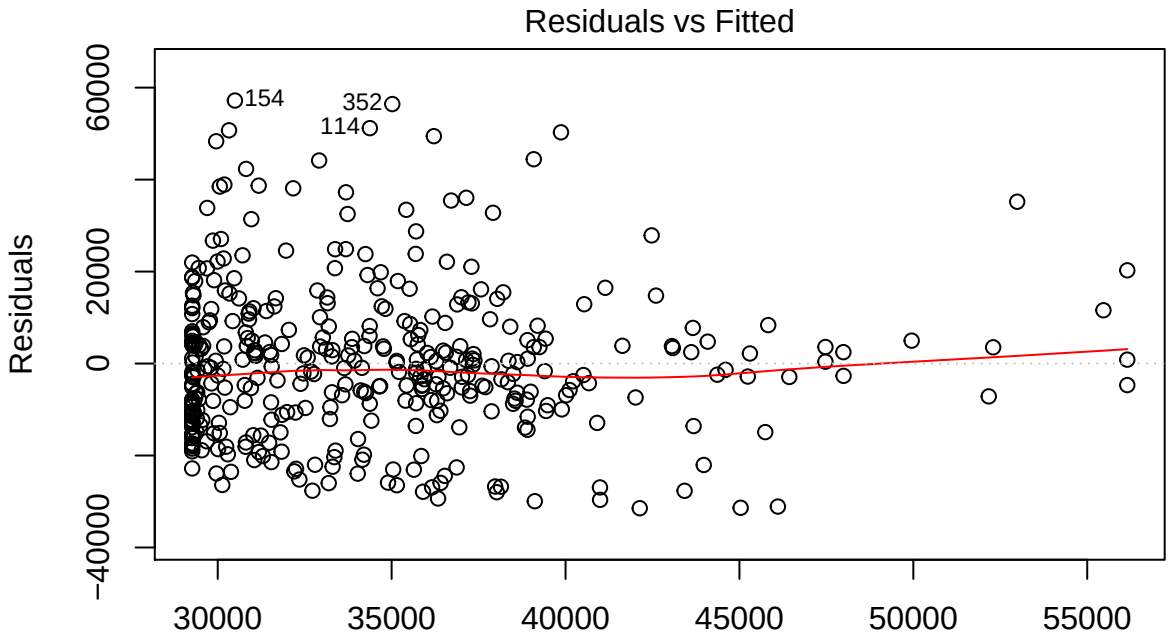
```
plot(fit1)
```

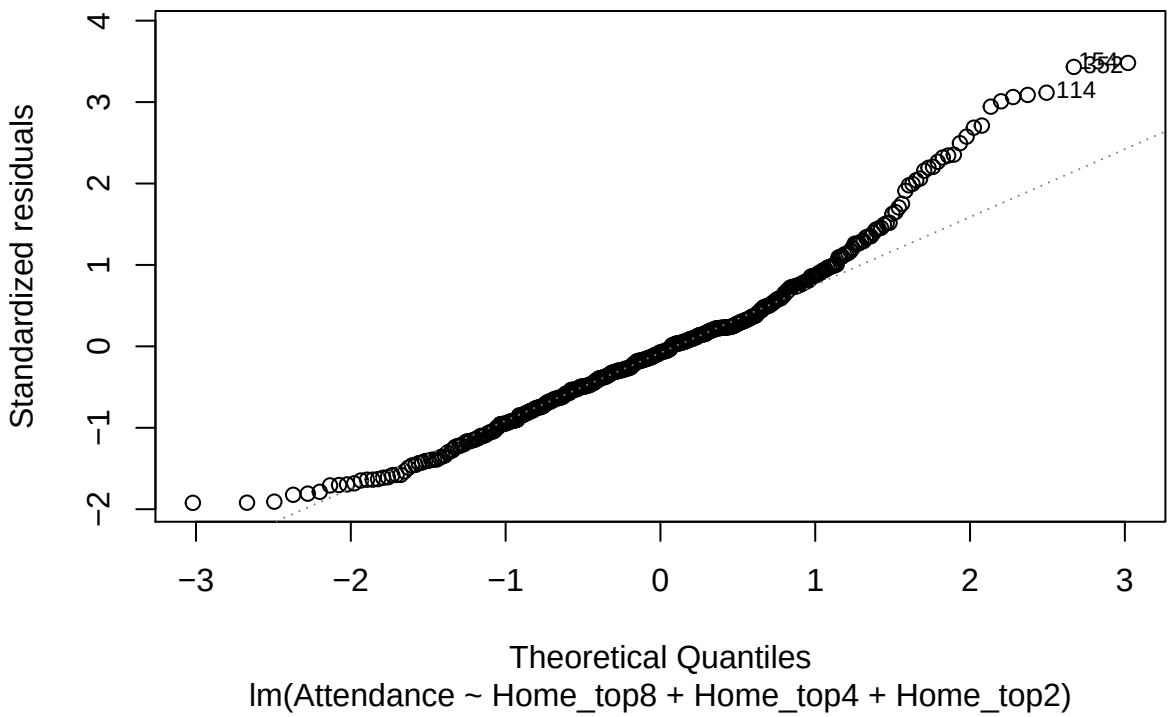


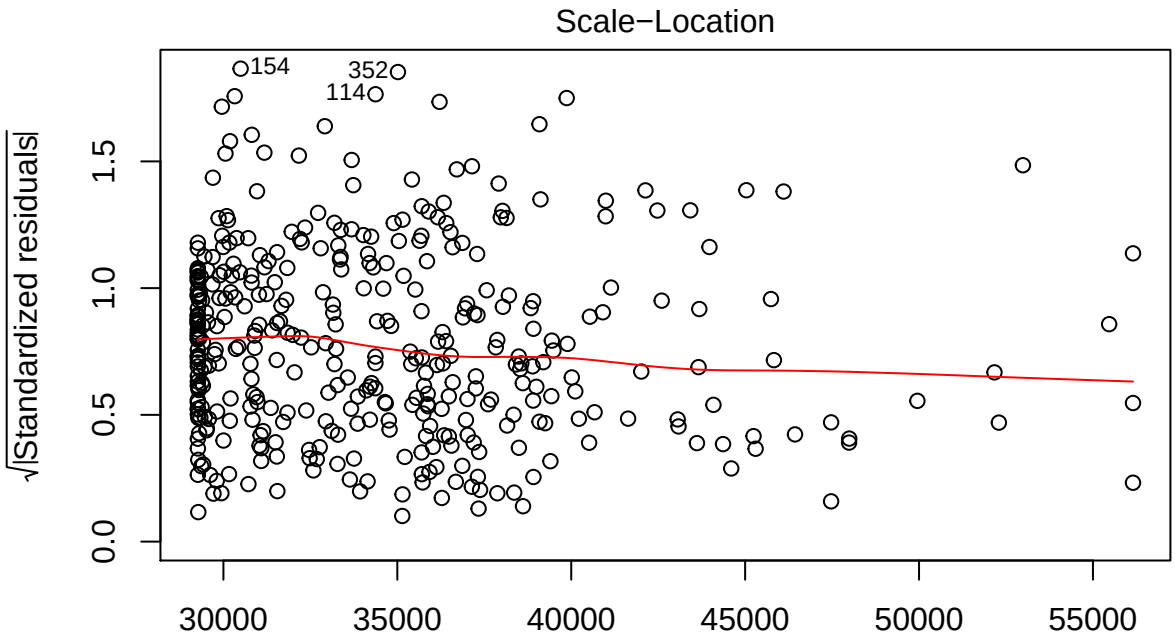


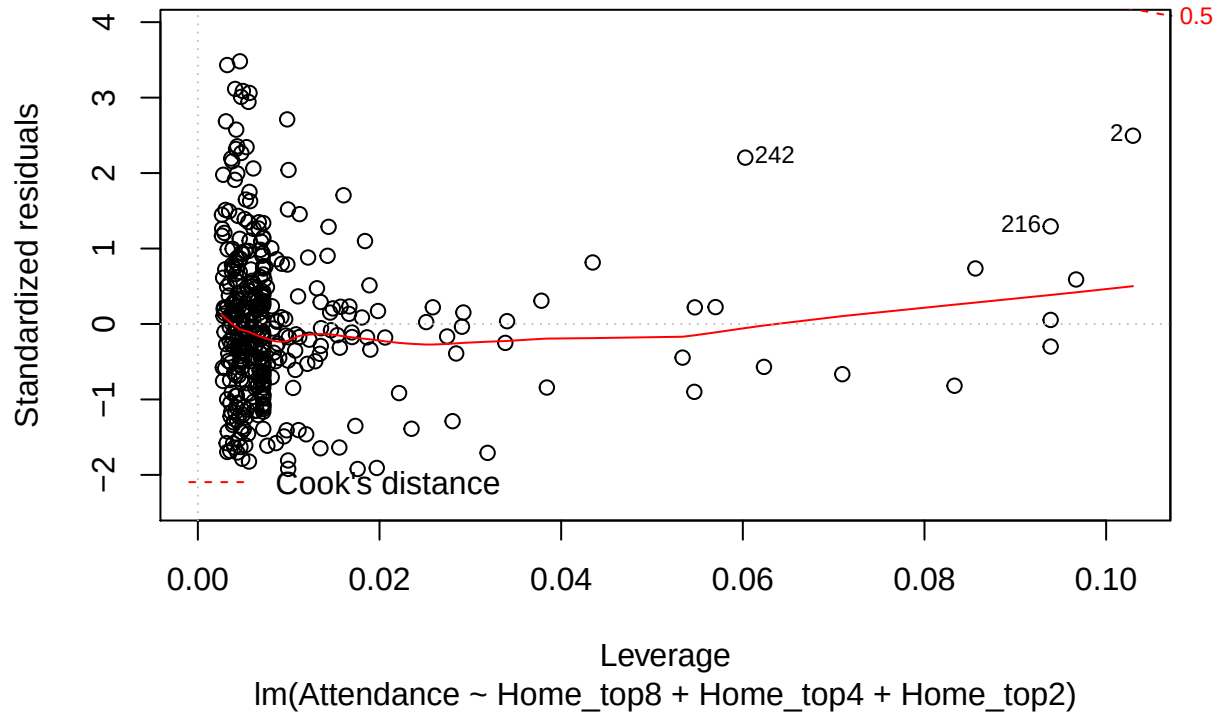
Adding Home championship probability to the model with Home and Away playoff probability makes the model better. A stepwise regression shows that we should remove Away playoff probability from the model. Our new model shows that as the Home playoff probability increases by 1 percent, we can expect an increase in the number of fans of 63, while Home championship probability increases attendance by 170 fans for every one percent increase.

```r
fit <- lm(Attendance~Home_top8+Home_top4+Home_top2, data=subset_afl)
plot(fit)
```










```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Home_top4 + Home_top2,
##     data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -31457 -10512  -1202   7930  57191
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 29264.94    1397.36  20.943   <2e-16 ***
## Home_top8     121.42      56.73   2.140   0.0330 *
## Home_top4    -173.10     140.54  -1.232   0.2188
## Home_top2     320.56     135.60   2.364   0.0186 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16470 on 392 degrees of freedom
## Multiple R-squared:  0.09112,    Adjusted R-squared:  0.08416
## F-statistic:  13.1 on 3 and 392 DF,  p-value: 3.61e-08
```

```
fit1 <- step(fit)
```

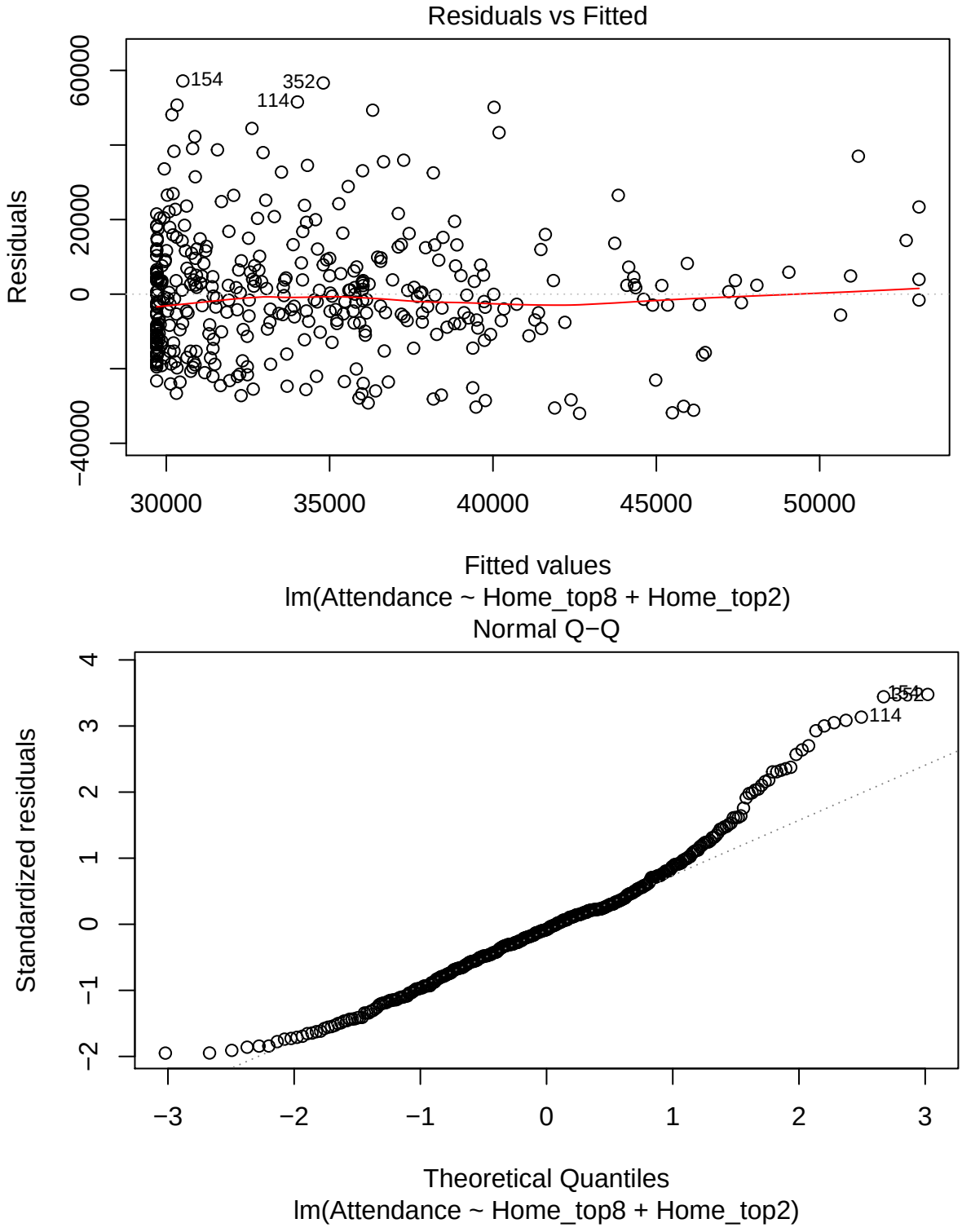
```
## Start:  AIC=7693.6
## Attendance ~ Home_top8 + Home_top4 + Home_top2
##
##             Df  Sum of Sq        RSS    AIC
## - Home_top4  1  411377685 1.0671e+11 7693.1
## <none>                    1.0630e+11 7693.6
## - Home_top8  1 1242049505 1.0754e+11 7696.2
## - Home_top2  1 1515380766 1.0781e+11 7697.2
##
## Step:  AIC=7693.13
## Attendance ~ Home_top8 + Home_top2
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    1.0671e+11 7693.1
## - Home_top8  1 1109060582 1.0782e+11 7695.2
## - Home_top2  1 2242346151 1.0895e+11 7699.4
```

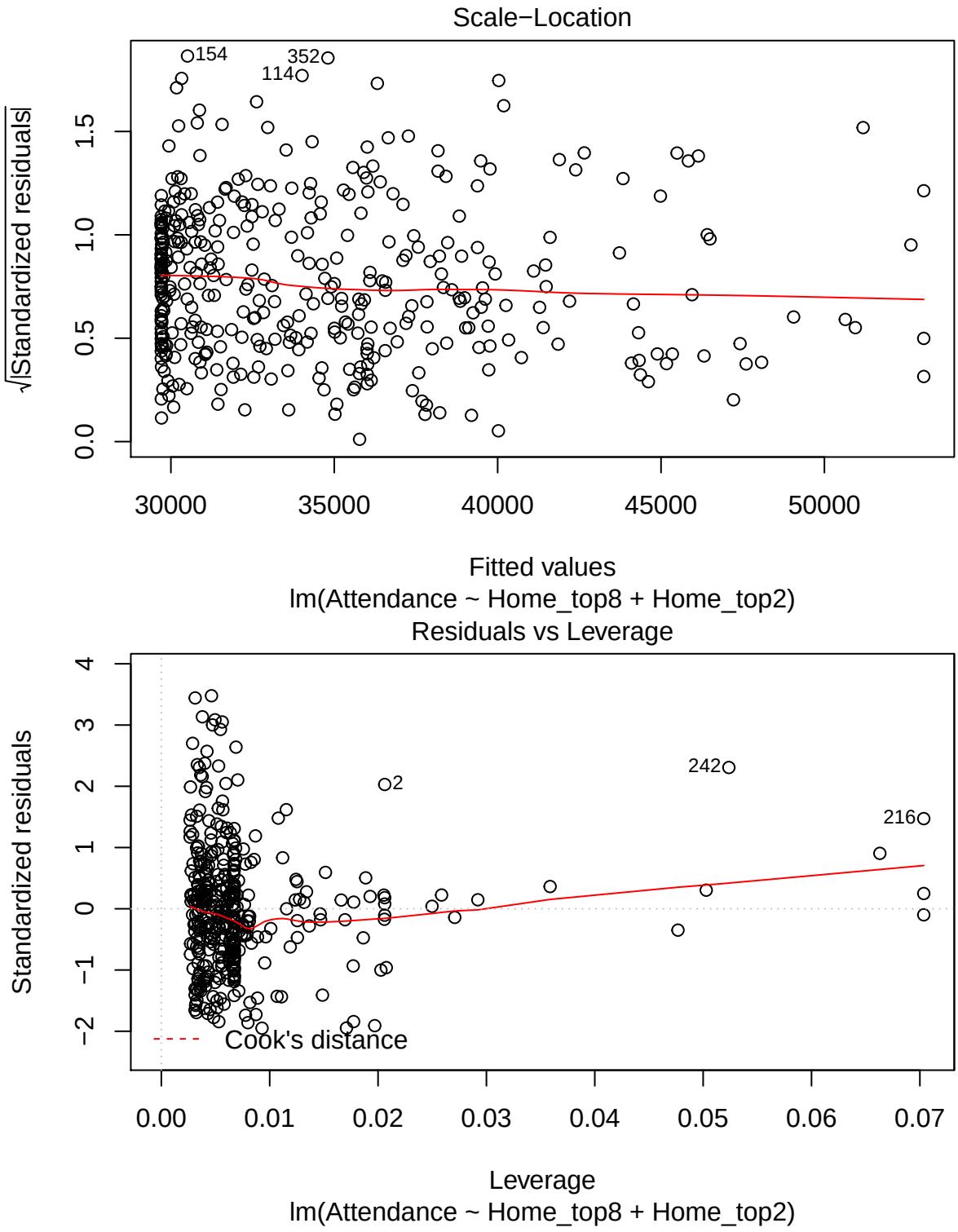
```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Home_top2, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -31974 -10799  -1315   7690  57177
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 29706.62    1351.45  21.981  < 2e-16 ***
## Home_top8      63.03      31.19   2.021  0.04395 *
## Home_top2     170.31      59.26   2.874  0.00428 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 16480 on 393 degrees of freedom
## Multiple R-squared:  0.0876, Adjusted R-squared:  0.08296
## F-statistic: 18.87 on 2 and 393 DF,  p-value: 1.502e-08
```

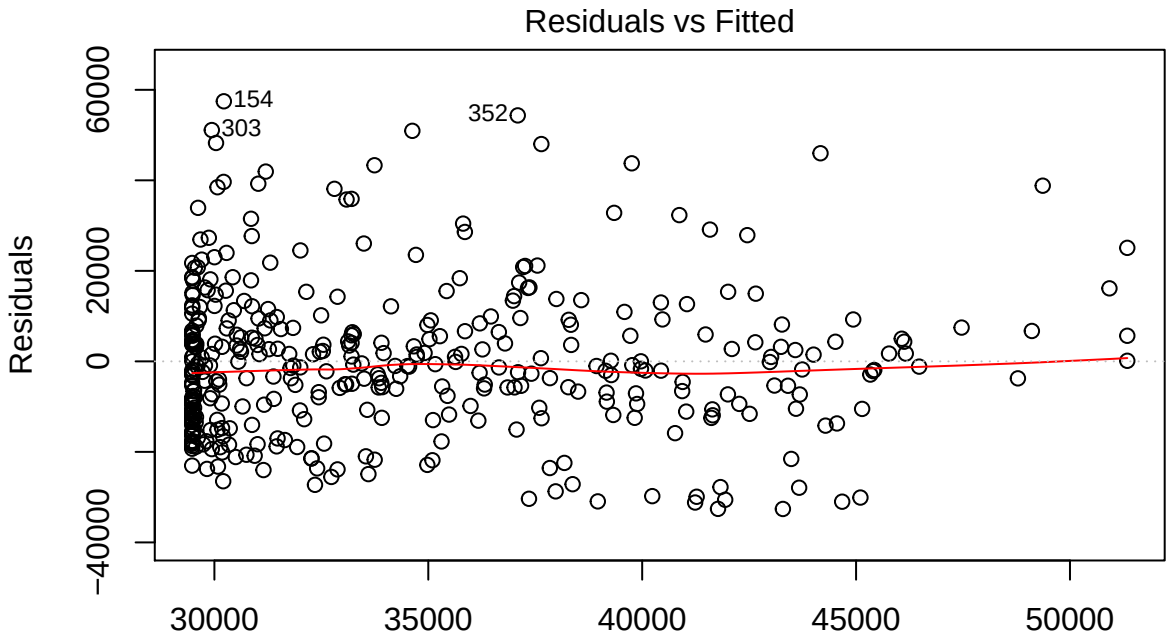
```
plot(fit1)
```

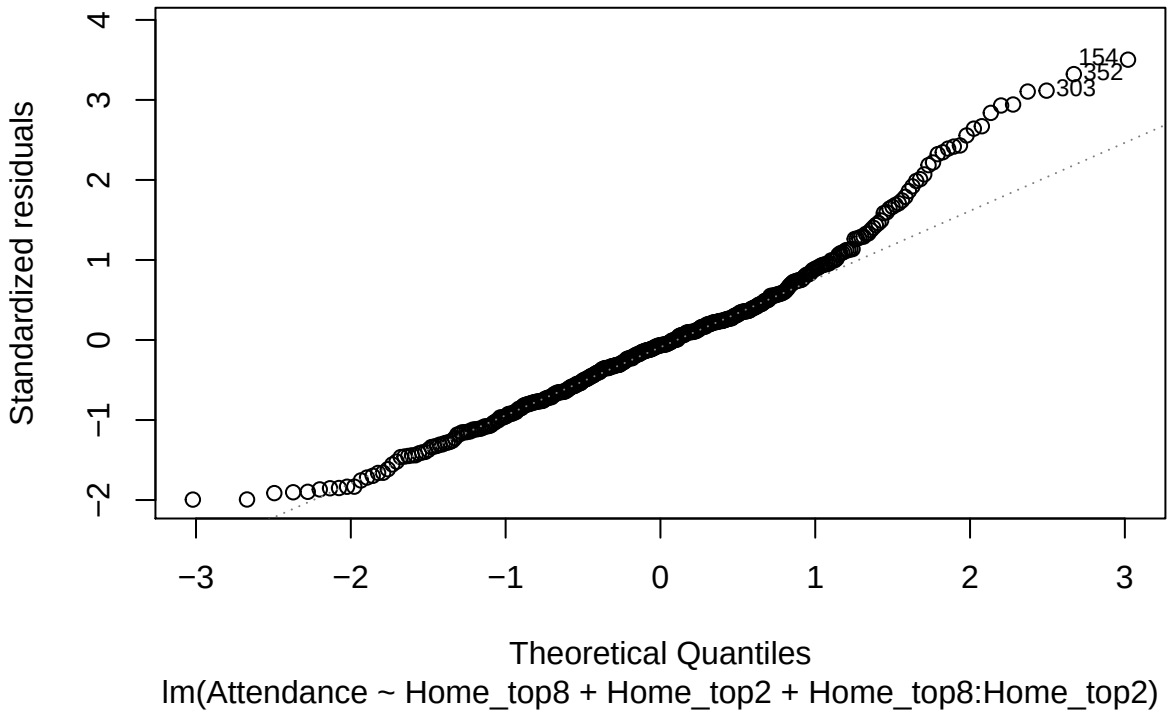


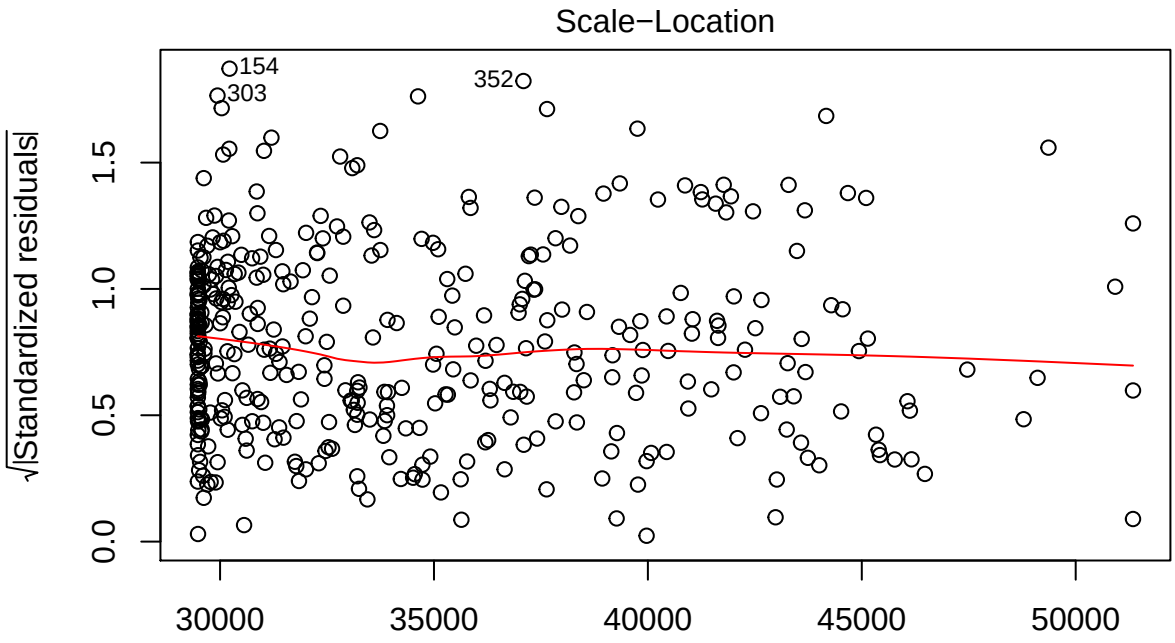


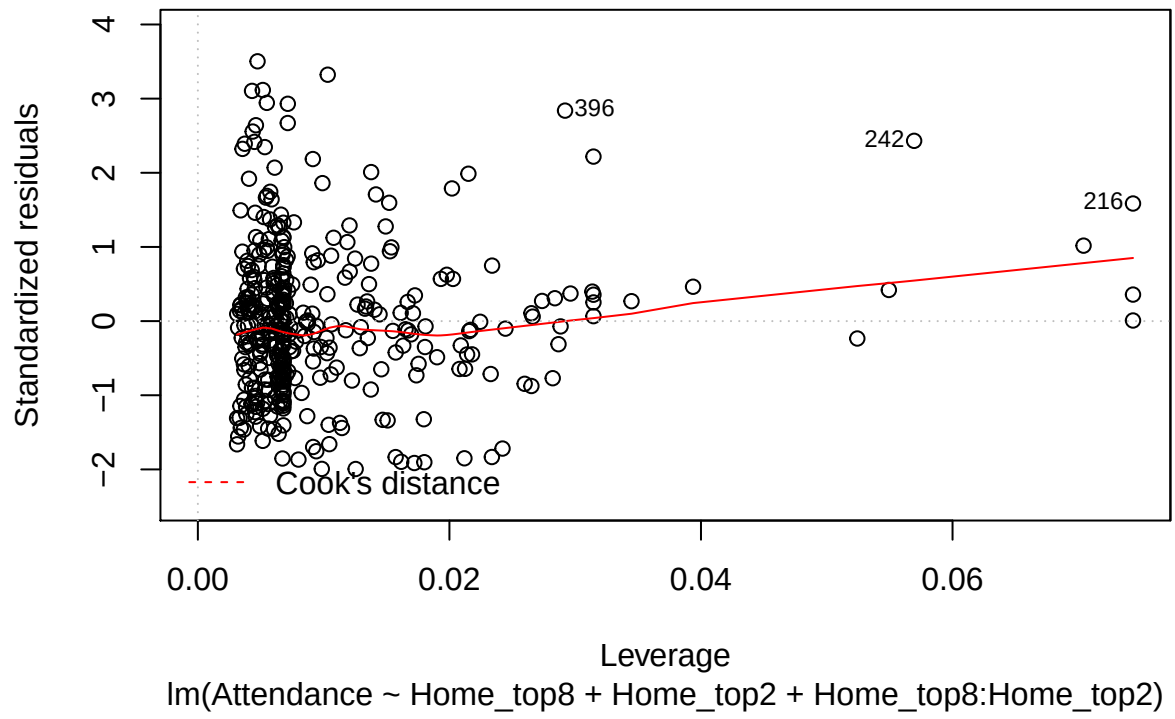
Adding in Home chances of being in the top 4 does not improve the model.

```
fit <- lm(Attendance~Home_top8+Home_top2+Home_top8:Home_top2, data=subset_afl)
plot(fit)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_top8 + Home_top2 + Home_top8:Home_top2)
154
303
352
60000
20000
0
−40000
30000
35000
40000
45000
50000


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top8 + Home_top2 + Home_top8:Home_top2)
154
352
303
4
3
2
1
0
−1
−2
−3






```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Home_top2 + Home_top8:Home_top2,
##     data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -32612 -10902  -1080   7938  57464
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         29480.484   1355.619  21.747   <2e-16 ***
## Home_top8              37.257     34.872   1.068   0.2860
## Home_top2            1098.906    570.053   1.928   0.0546 .
## Home_top8:Home_top2    -9.176      5.602  -1.638   0.1023
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16440 on 392 degrees of freedom
## Multiple R-squared:  0.0938, Adjusted R-squared:  0.08686
## F-statistic: 13.53 on 3 and 392 DF,  p-value: 2.05e-08
```

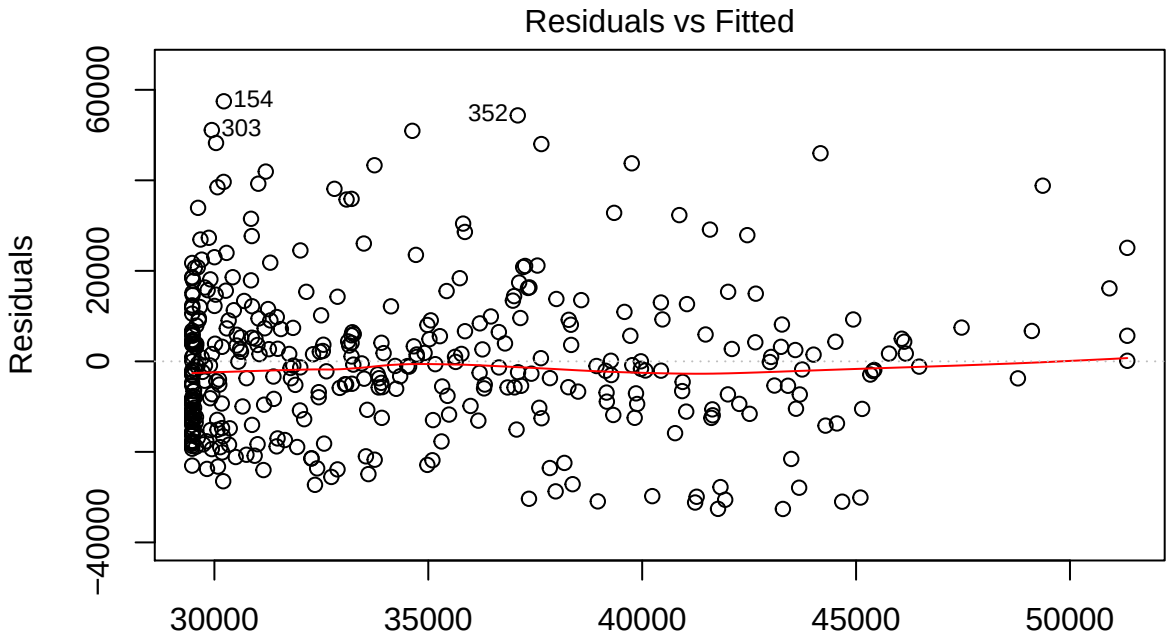
```
fit1 <- step(fit)
```

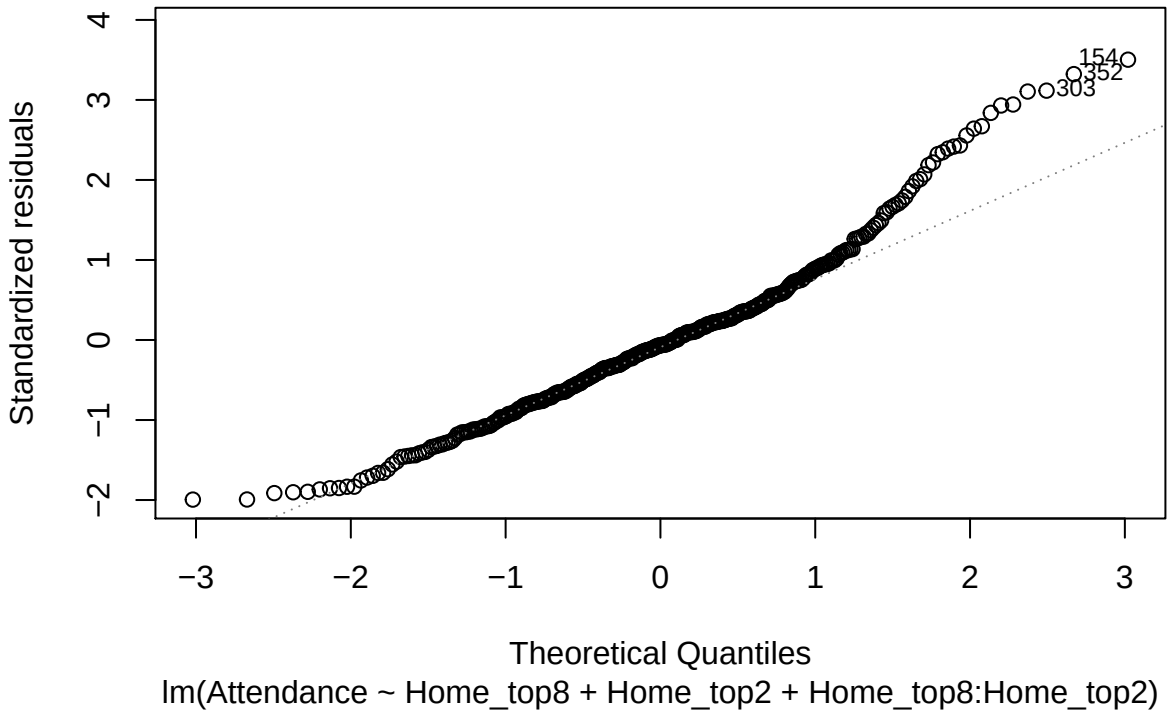
```
## Start:  AIC=7692.43
## Attendance ~ Home_top8 + Home_top2 + Home_top8:Home_top2
##
##                       Df Sum of Sq        RSS    AIC
## <none>                               1.0598e+11 7692.4
## - Home_top8:Home_top2  1 725211557 1.0671e+11 7693.1
```

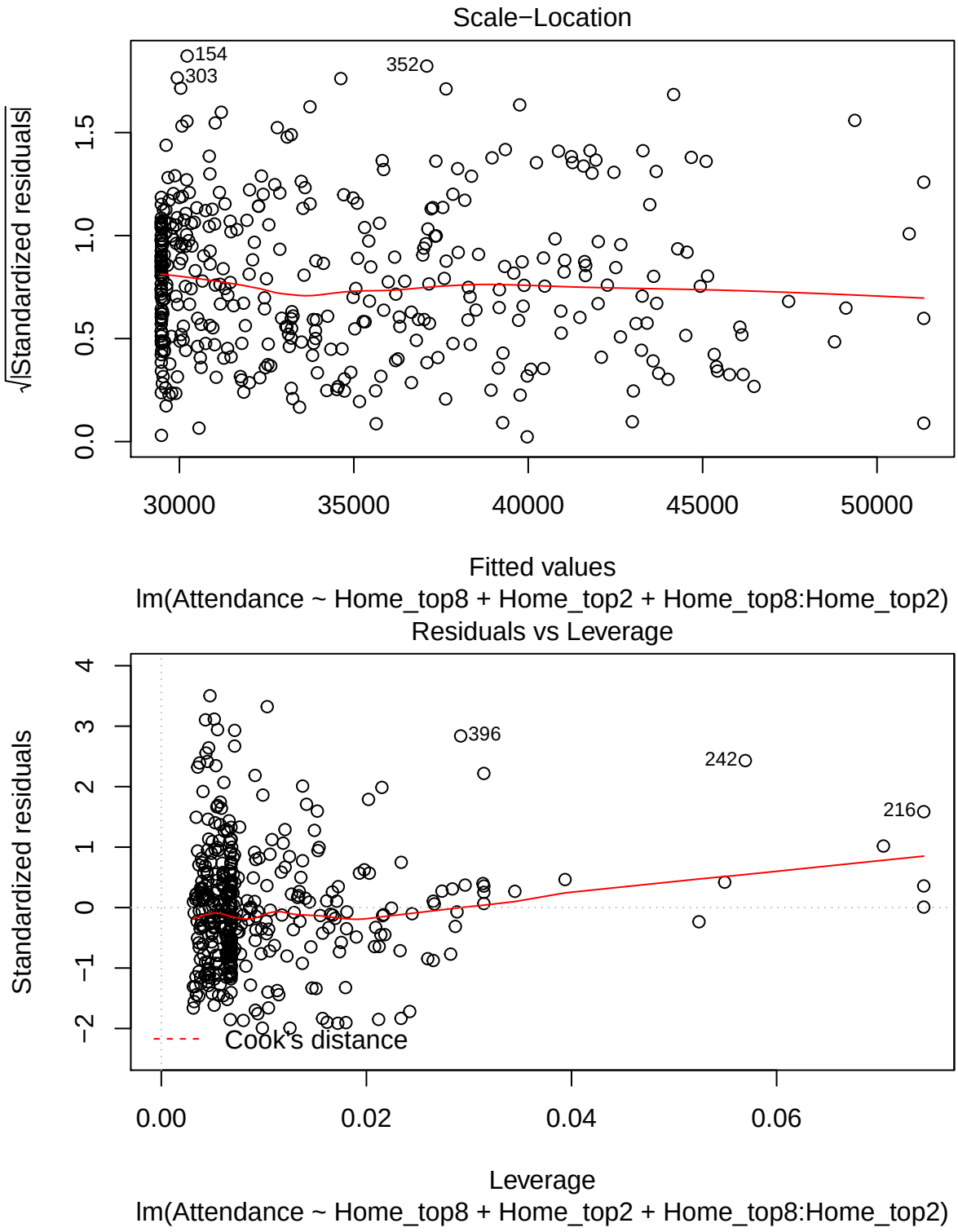
```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Home_top2 + Home_top8:Home_top2,
##     data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -32612 -10902  -1080   7938  57464
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         29480.484   1355.619  21.747   <2e-16 ***
## Home_top8              37.257     34.872   1.068   0.2860
## Home_top2            1098.906    570.053   1.928   0.0546 .
## Home_top8:Home_top2    -9.176      5.602  -1.638   0.1023
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16440 on 392 degrees of freedom
## Multiple R-squared:  0.0938, Adjusted R-squared:  0.08686
## F-statistic: 13.53 on 3 and 392 DF,  p-value: 2.05e-08
```

```
plot(fit1)
```

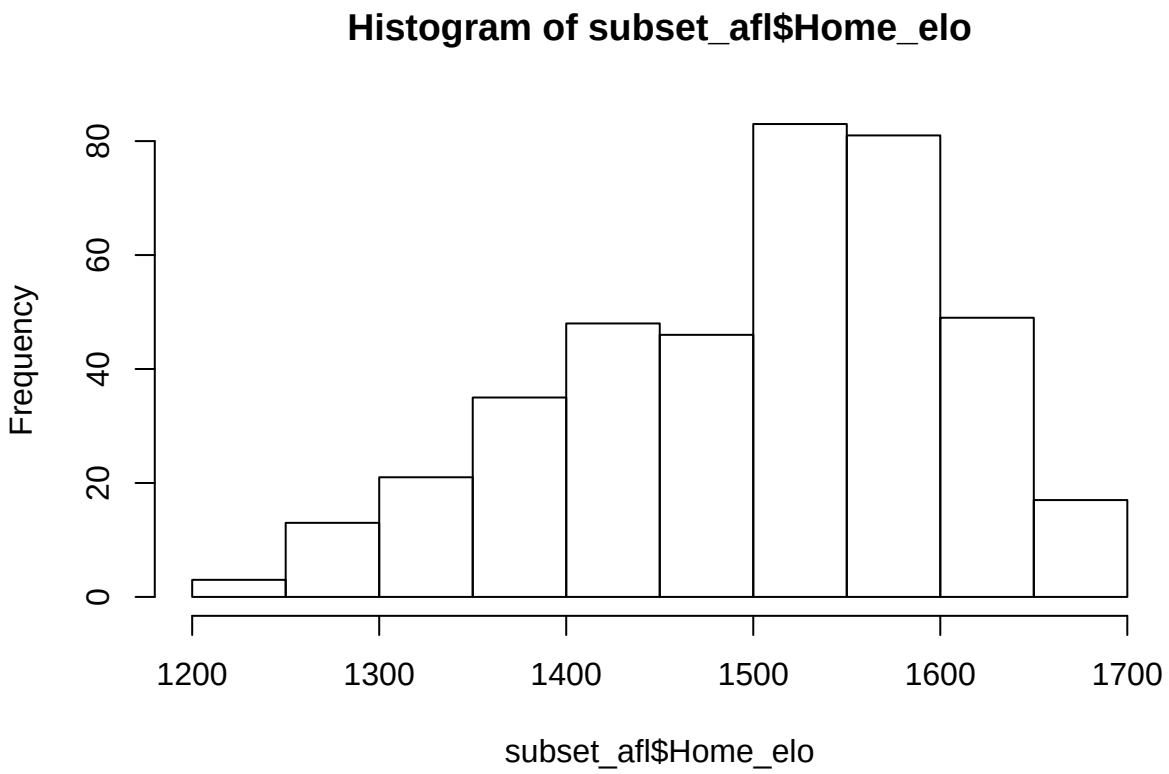

Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_top8 + Home_top2 + Home_top8:Home_top2)
154
303
352
60000
20000
0
−40000
30000
35000
40000
45000
50000


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top8 + Home_top2 + Home_top8:Home_top2)
154
352
303
4
3
2
1
0
−1
−2
−3



Adding an interaction term between home playoff probability and home championship appearance probability does make the model better, although the p-values are now higher.

```
hist(subset_afl$Home_elo, breaks = 10)
```



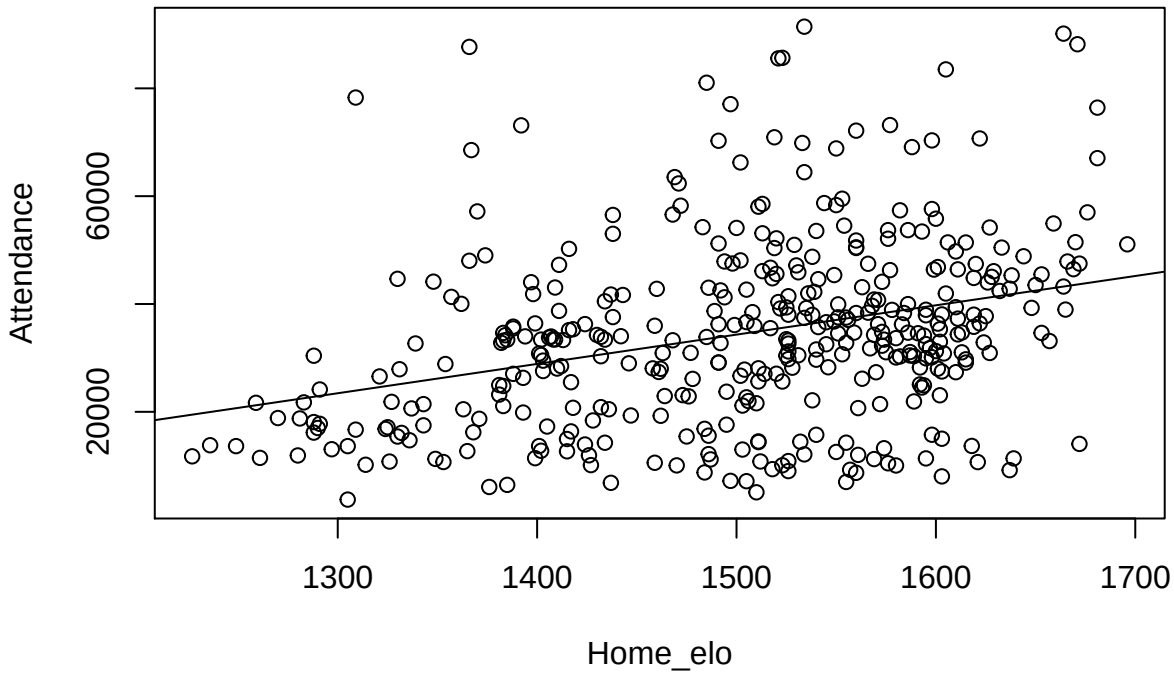

```
mean(subset_afl$Home_elo)
```

```
## [1] 1501.707
```

```
sd(subset_afl$Home_elo)
```

```
## [1] 101.0113
```

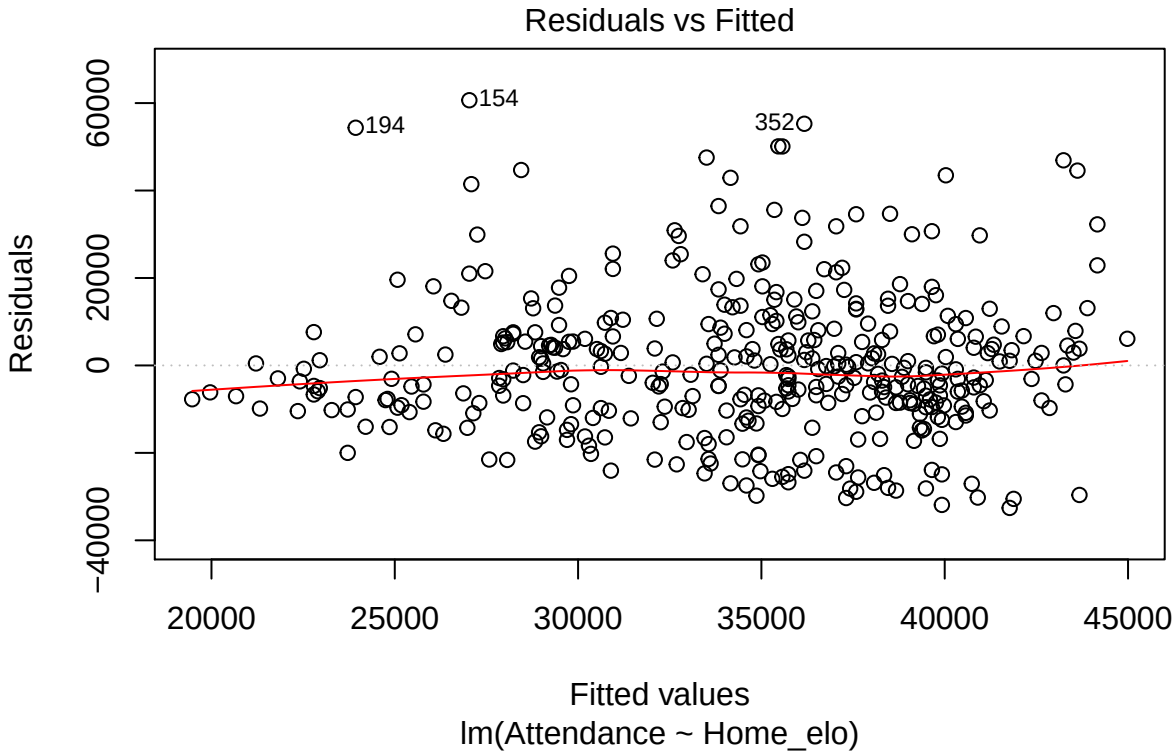
```
plot(Attendance~Home_elo, data = subset_afl)
fit <- lm(Attendance~Home_elo, data=subset_afl)
abline(fit)
```

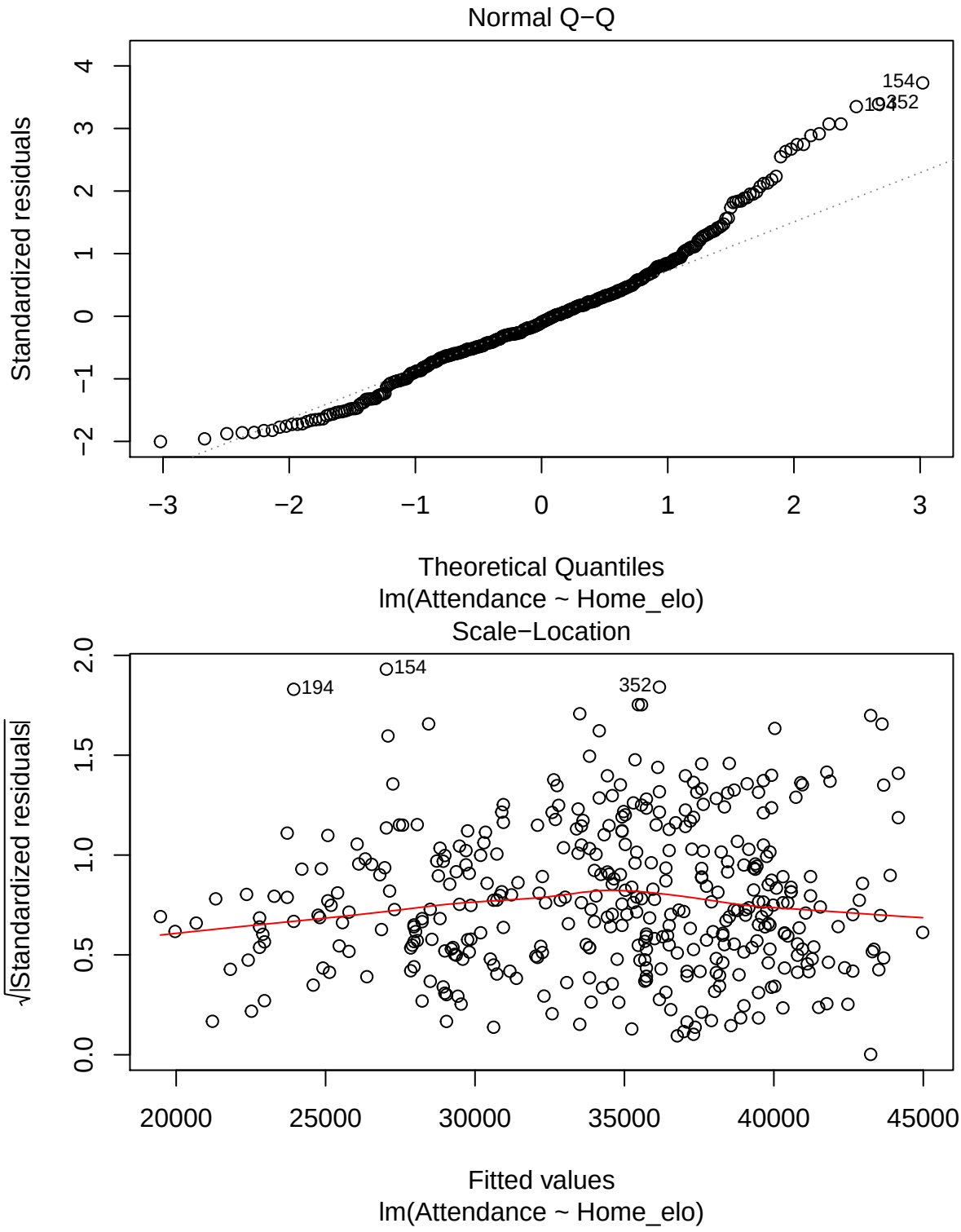


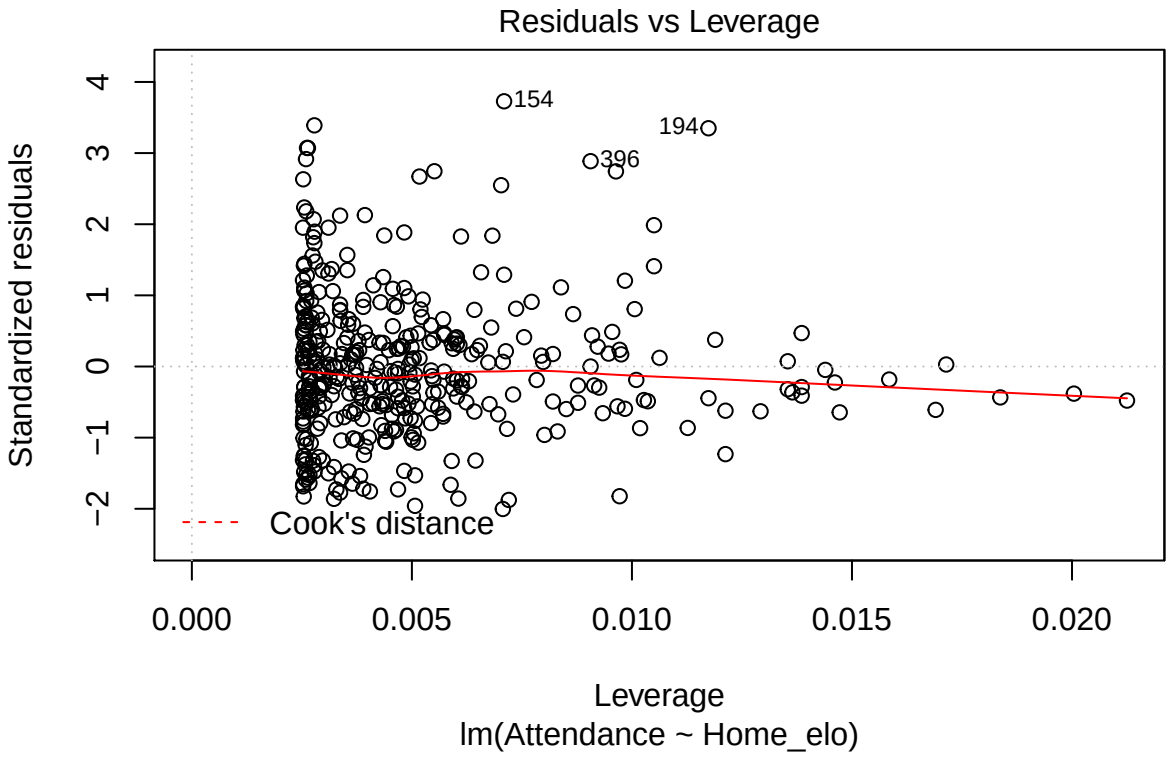

```
cor(subset_afl$Attendance, subset_afl$Home_elo)
```

```
## [1] 0.3192542
```

```
plot(fit)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_elo)
Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Home_elo)
154
194
352

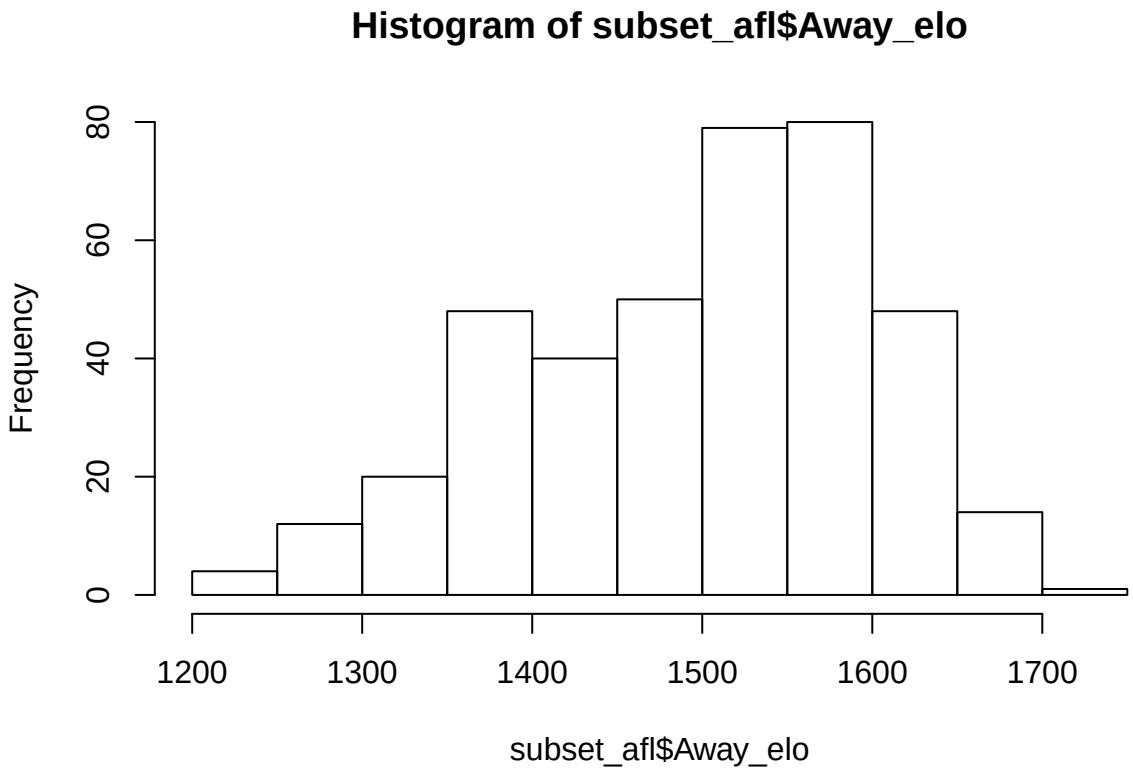



```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_elo, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -32587  -9677  -1527   7570  60651
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -47254.922  12240.680  -3.860 0.000132 ***
## Home_elo        54.384      8.133   6.687 7.82e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16330 on 394 degrees of freedom
## Multiple R-squared:  0.1019, Adjusted R-squared:  0.09964
## F-statistic: 44.72 on 1 and 394 DF,  p-value: 7.823e-11
```

As we would expect, the mean for home team Elo is very close to 1500 (1501.707). However, the histogram shows that it is not normally distributed: the really bad teams are worse than the really good teams are good. The plot of Home_elo and Attendance appears to show a positive relationship; the correlation coefficient is 0.319, indicating a moderate positive correlation. Similar to above, high residuals are more common than low ones. The model shows that for every one-point increase in Elo, attendance rises by 54.4 fans.

```
hist(subset_afl$Away_elo, breaks = 10)
```



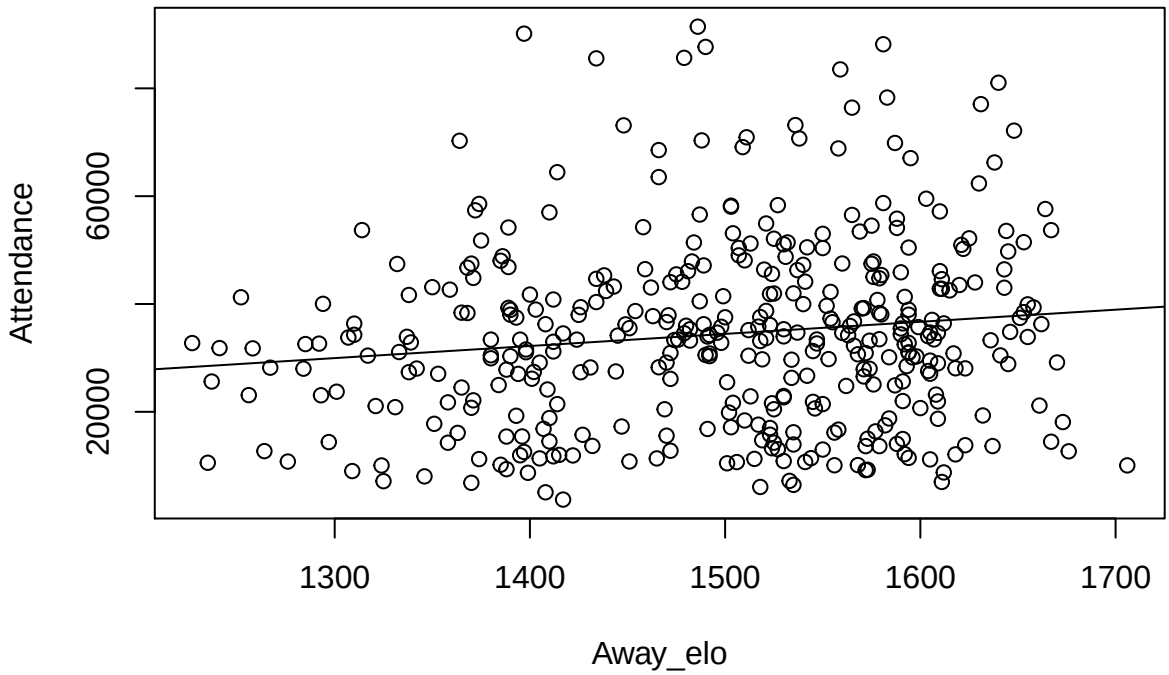

```
mean(subset_afl$Away_elo)
```

```
## [1] 1498.308
```

```
sd(subset_afl$Away_elo)
```

```
## [1] 102.593
```

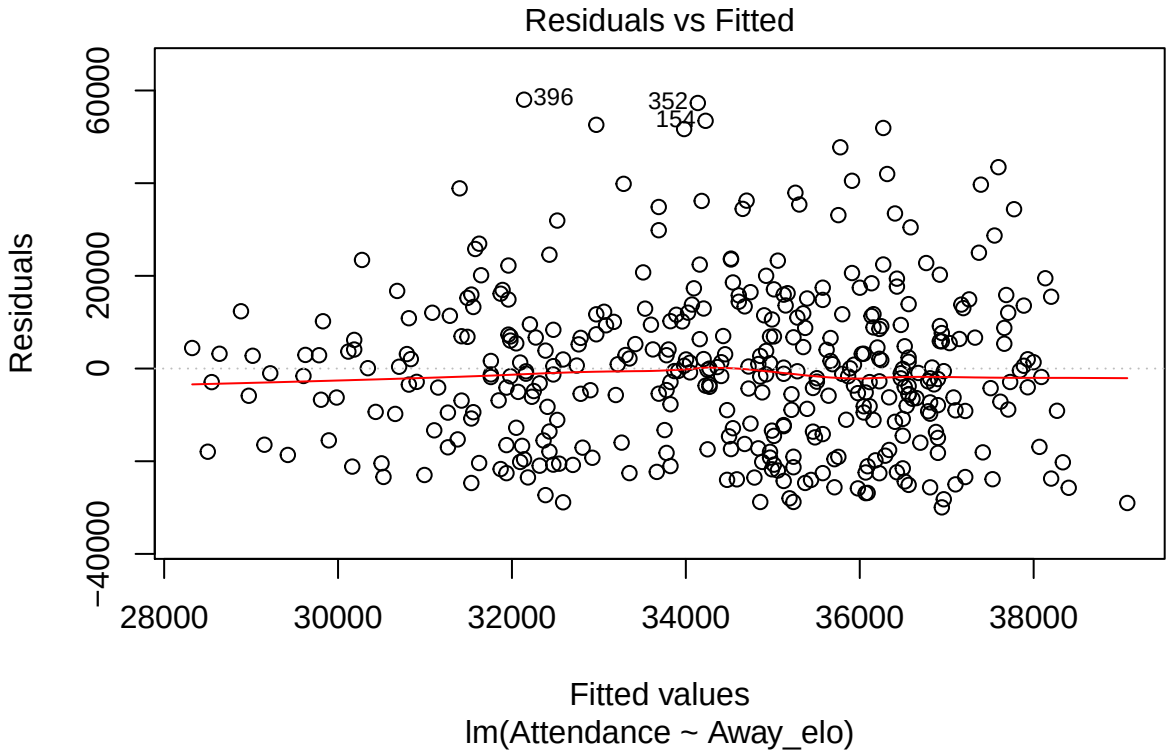
```
plot(Attendance~Away_elo, data = subset_afl)
fit <- lm(Attendance~Away_elo, data=subset_afl)
abline(fit)
```

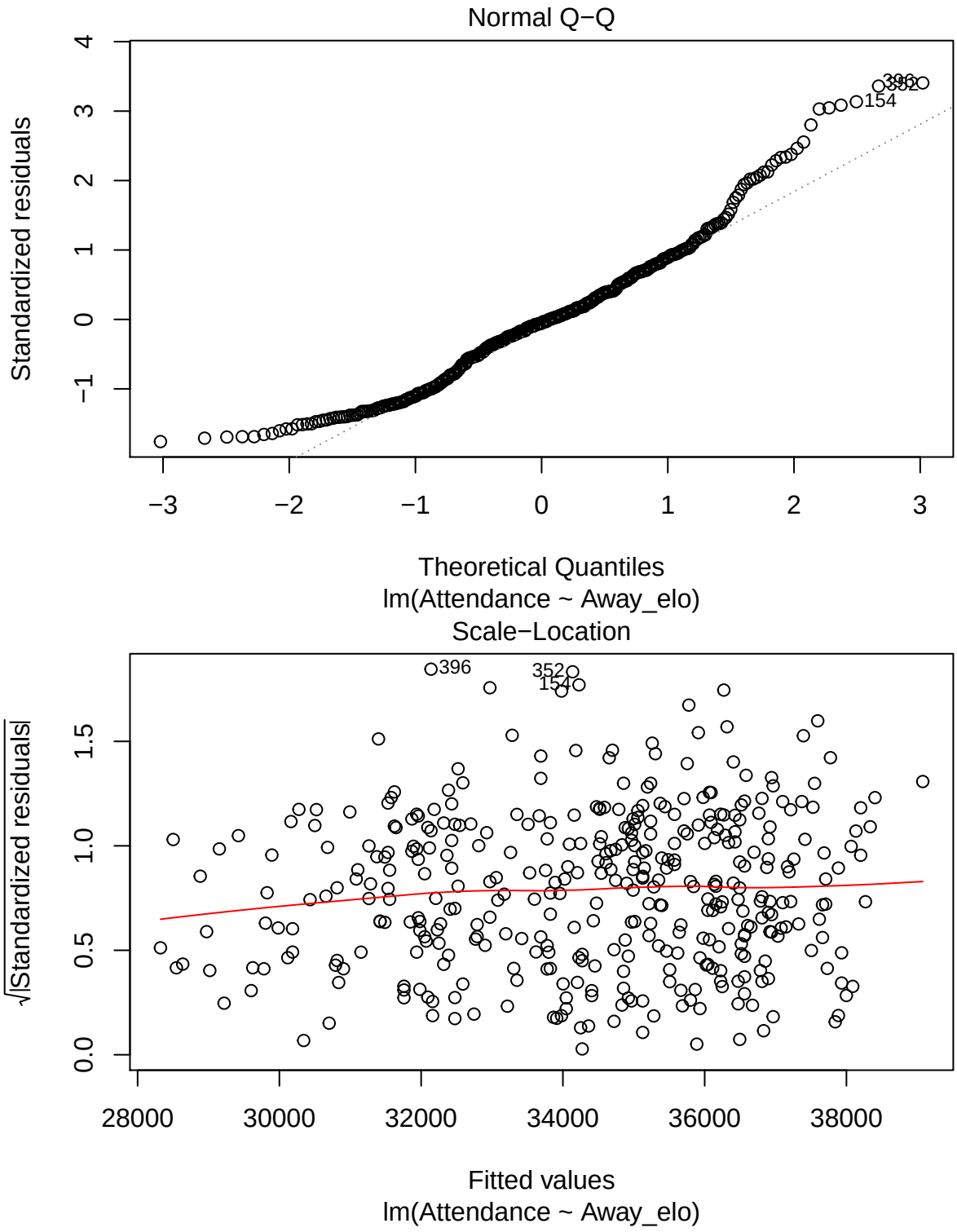


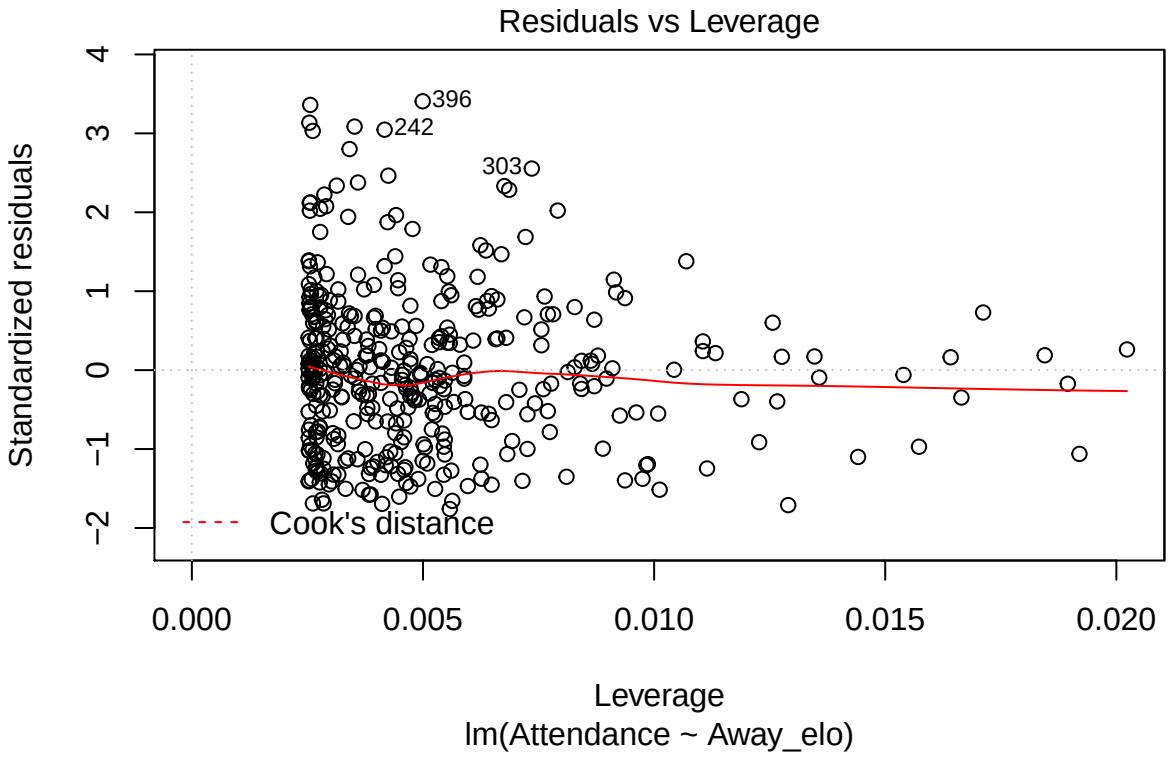

```
cor(subset_afl$Attendance, subset_afl$Away_elo)
```

```
## [1] 0.1338678
```

```
plot(fit)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Away_elo)
Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Away_elo)

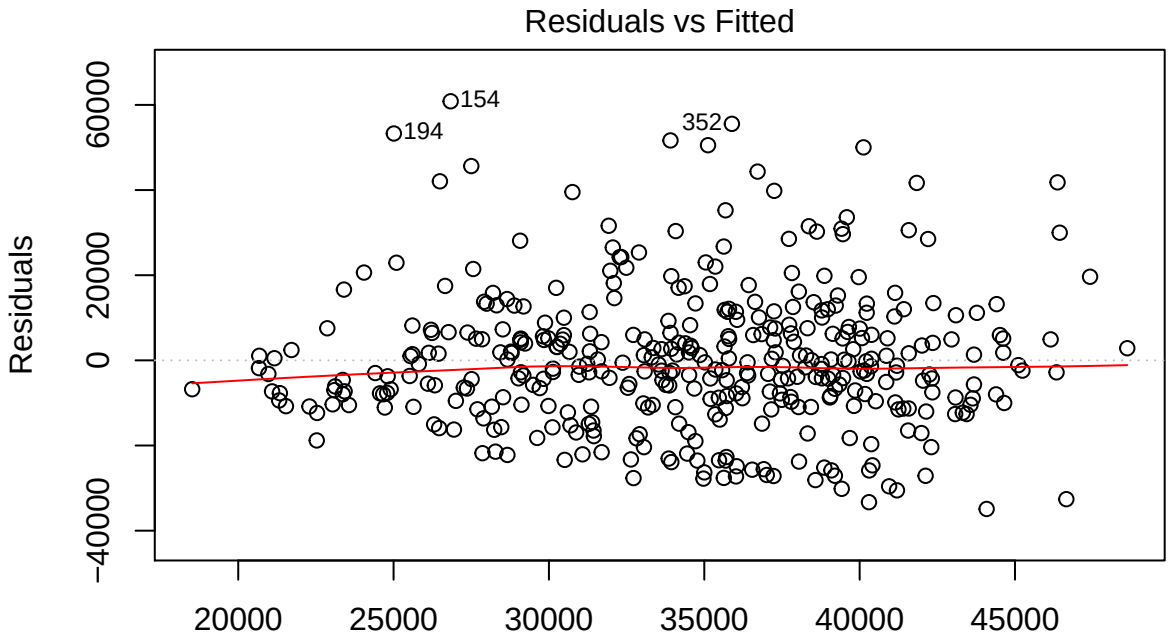


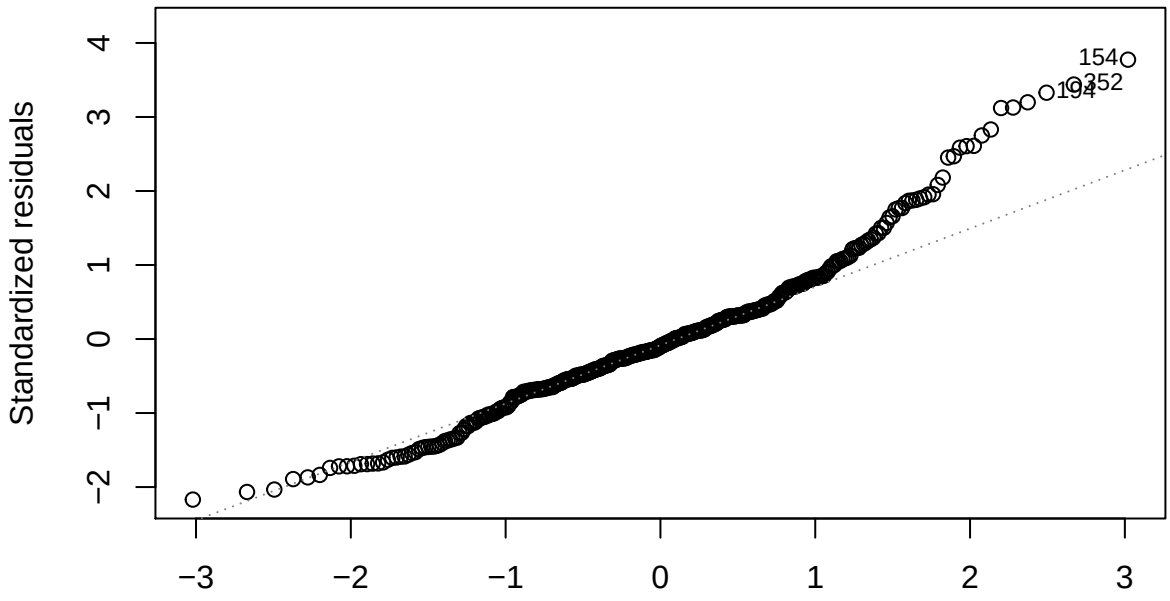

```
summary(fit)
```

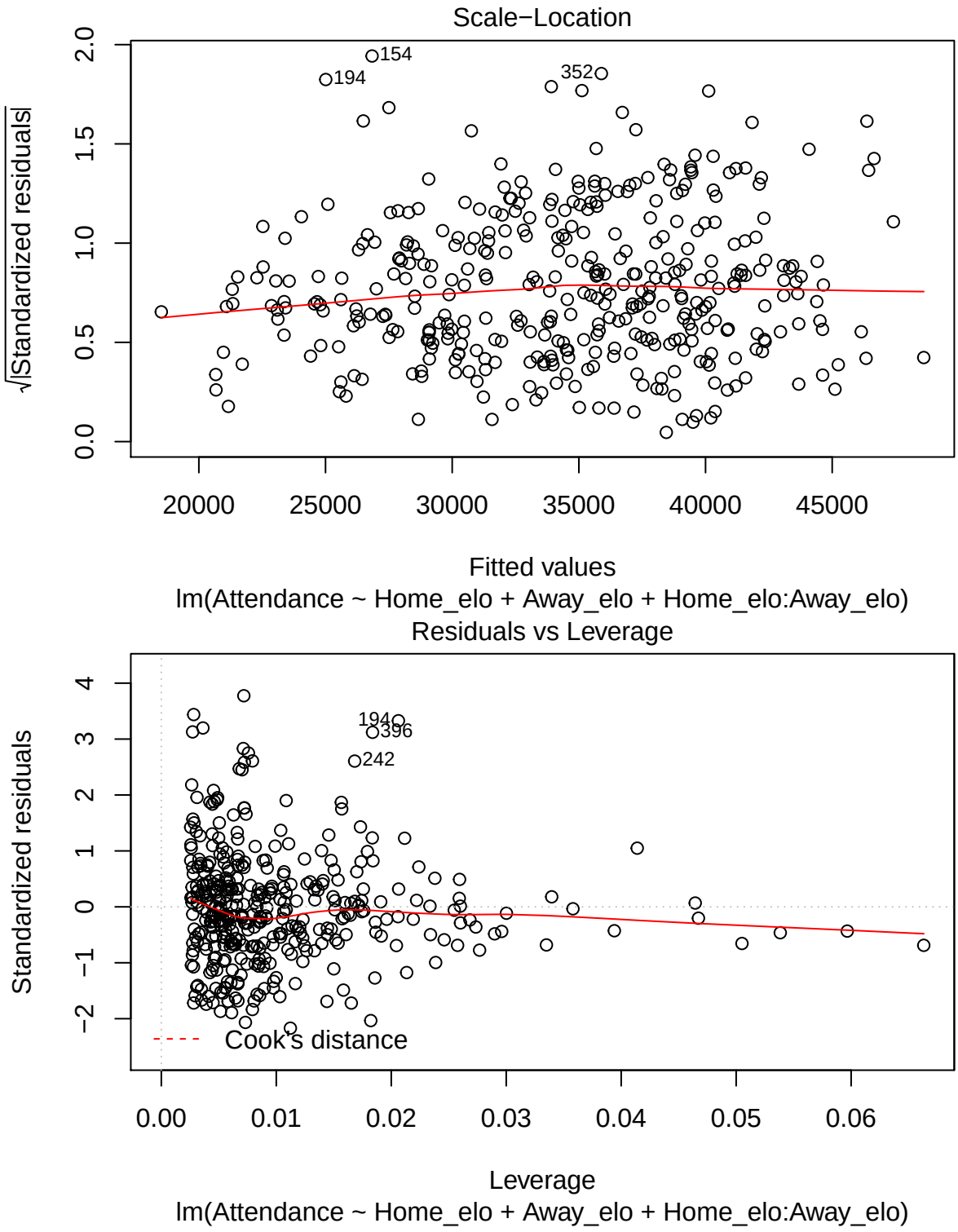
```
##
## Call:
## lm(formula = Attendance ~ Away_elo, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -29955 -12805   -893   9479  58012
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   773.222  12575.513   0.061  0.95100
## Away_elo       22.452      8.374   2.681  0.00764 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17070 on 394 degrees of freedom
## Multiple R-squared:  0.01792,    Adjusted R-squared:  0.01543
## F-statistic:  7.19 on 1 and 394 DF,  p-value: 0.007641
```

The mean for away team Elo is, as we expect, also close to 1500. It is slightly below 1500 (1498.308) by random chance, but making up for the fact that the home teams have an Elo slightly above 1500. We see the same trend in the histogram: left skewness. The plot of Away Elo against Attendance appears to have a slight positive relationship, and the correlation coefficient is 0.134 which means there is a minor positive correlation. The residuals vs. fitted still shows larger positive than negative residuals. The model shows that on average we can expect an increase of 22.5 fans per one point increase in Away Elo, and this is significant with a p-value of 0.008.

```
fit <- lm(Attendance~Home_elo+Away_elo+Home_elo:Away_elo, data=subset_afl)
plot(fit)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_elo + Away_elo + Home_elo:Away_elo)
154
194
352
60000
20000
0
−40000
20000
25000
30000
35000
40000
45000


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_elo + Away_elo + Home_elo:Away_elo)
154
352
194
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3




```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_elo + Away_elo + Home_elo:Away_elo,
##     data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34902  -9918  -1557   7210  60848
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.068e+04  1.922e+05   0.160    0.873
## Home_elo          -2.094e+01  1.275e+02  -0.164    0.870
## Away_elo          -5.252e+01  1.280e+02  -0.410    0.682
## Home_elo:Away_elo  5.061e-02  8.490e-02   0.596    0.551
##
## Residual standard error: 16180 on 392 degrees of freedom
## Multiple R-squared:  0.1225, Adjusted R-squared:  0.1158
## F-statistic: 18.25 on 3 and 392 DF,  p-value: 4.206e-11
```

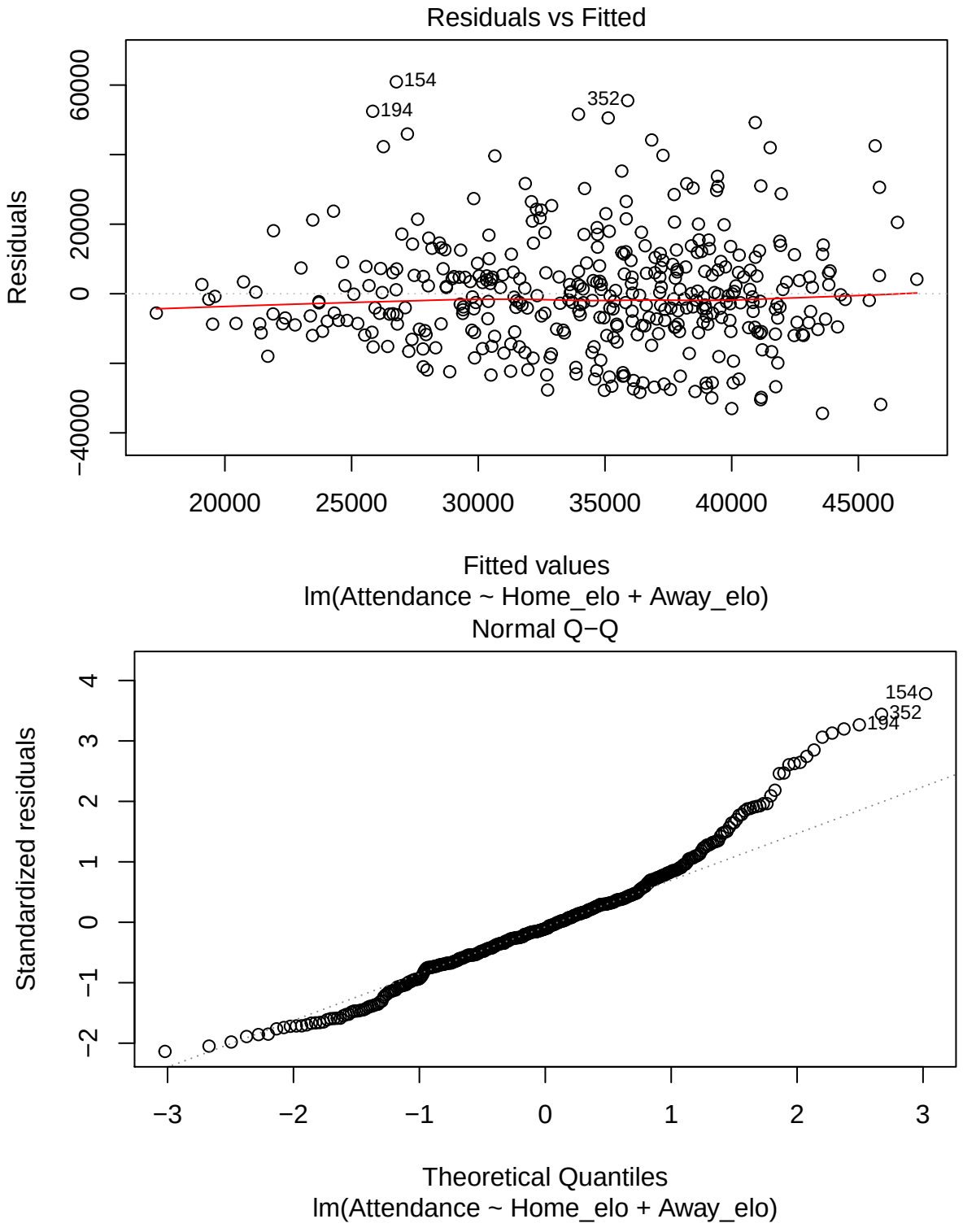
```
fit1 <- step(fit)
```

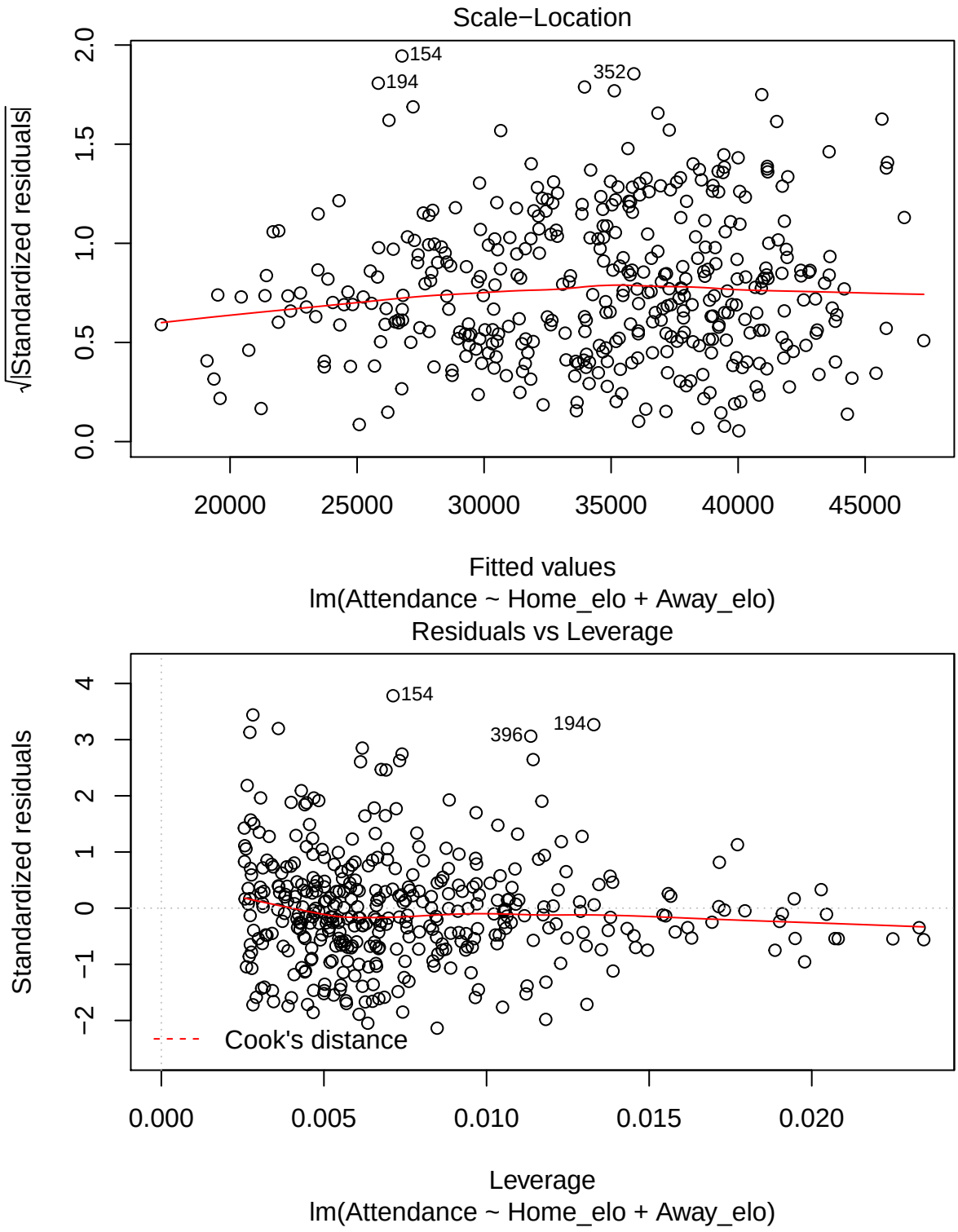
```
## Start:  AIC=7679.66
## Attendance ~ Home_elo + Away_elo + Home_elo:Away_elo
##
##                     Df Sum of Sq        RSS    AIC
## - Home_elo:Away_elo  1  93033971 1.0271e+11 7678.0
## <none>                           1.0262e+11 7679.7
##
## Step:  AIC=7678.02
## Attendance ~ Home_elo + Away_elo
##
##            Df  Sum of Sq        RSS    AIC
## <none>                   1.0271e+11 7678.0
## - Away_elo  1 2.3180e+09 1.0503e+11 7684.9
## - Home_elo  1 1.2142e+10 1.1486e+11 7720.3
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_elo + Away_elo, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34397  -9654  -1696   7150  60918
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -83418.177  17156.808  -4.862 1.68e-06 ***
## Home_elo        54.901      8.055   6.816 3.53e-11 ***
## Away_elo        23.618      7.931   2.978  0.00308 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16170 on 393 degrees of freedom
## Multiple R-squared:  0.1217, Adjusted R-squared:  0.1173
## F-statistic: 27.24 on 2 and 393 DF,  p-value: 8.349e-12
```

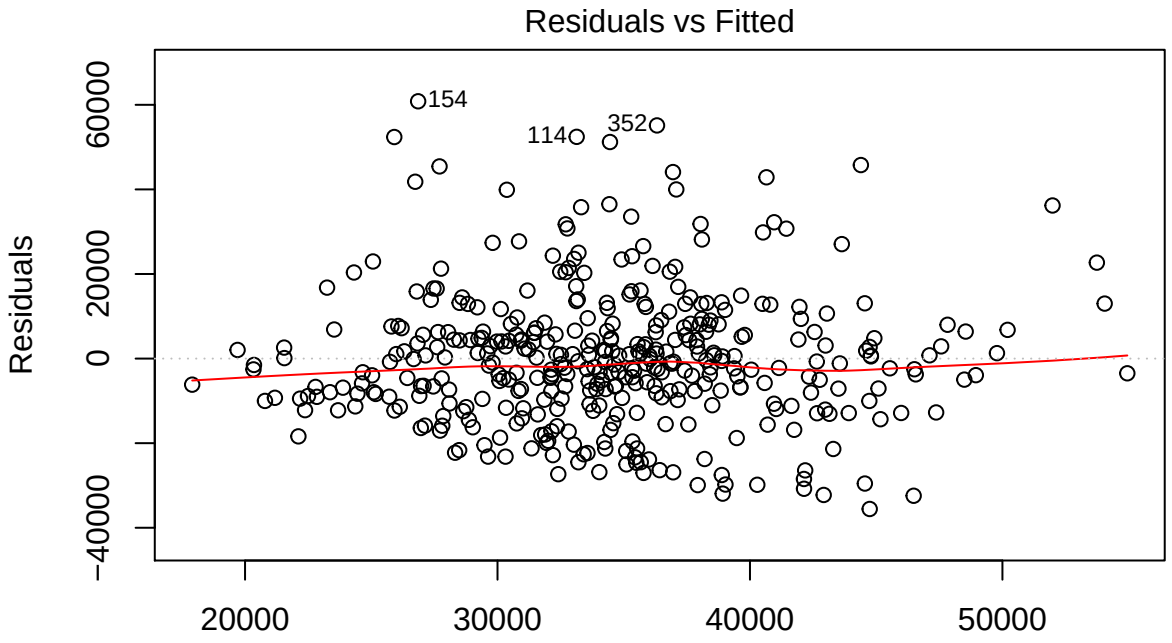
```
plot(fit1)
```

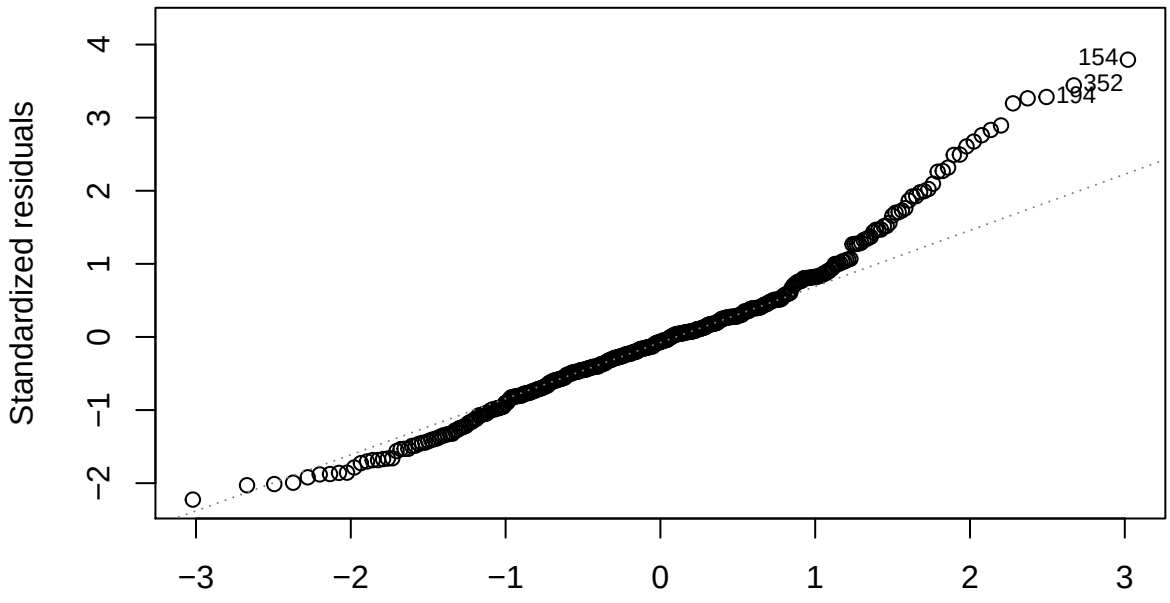

Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_elo + Away_elo)
154
194
352
Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_elo + Away_elo)

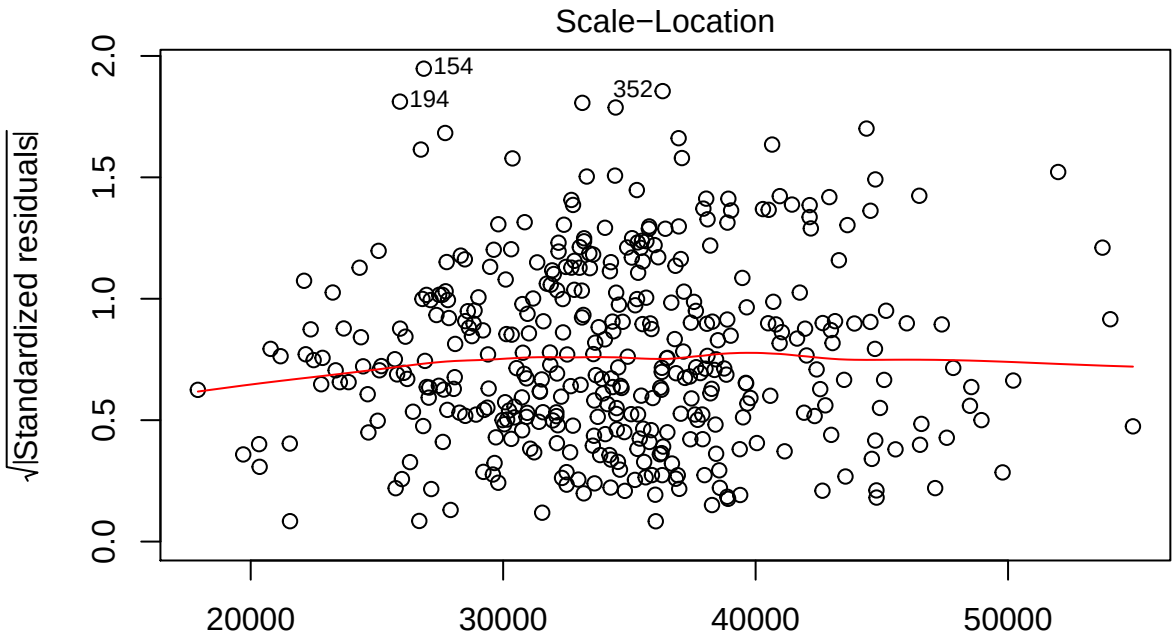


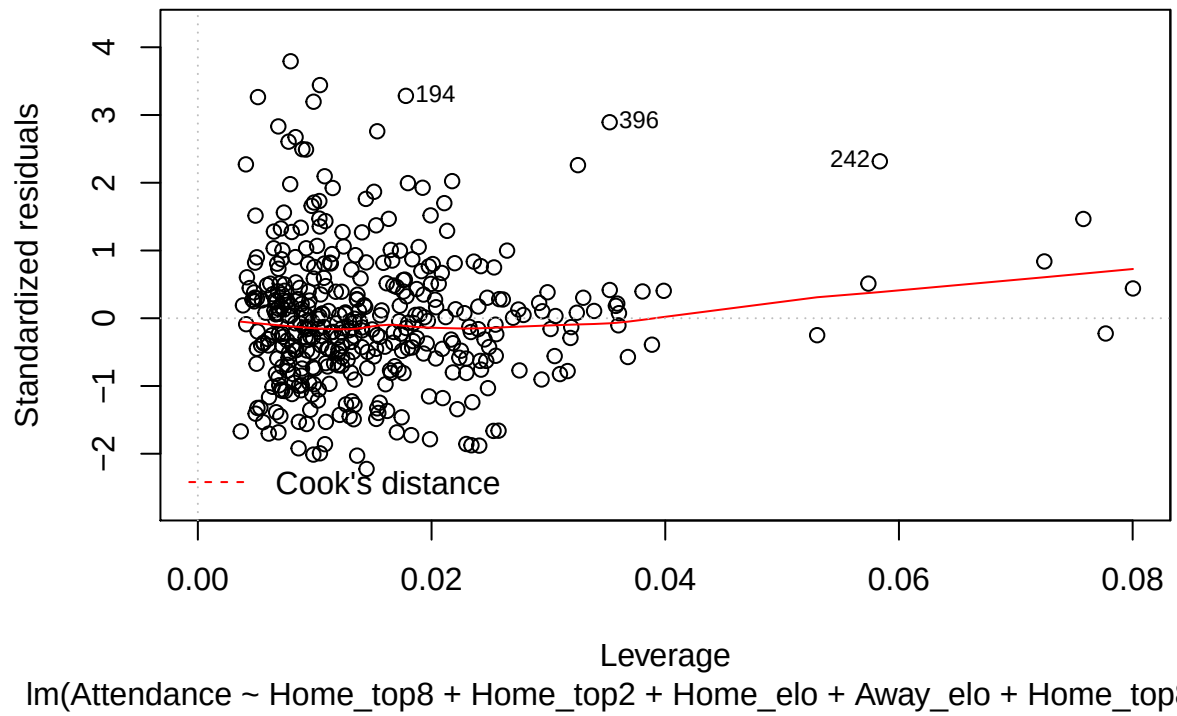
There is no need to run a model with the interaction term. The model with Home Elo and Away Elo shows that there is an increase of 54.9 fans per one point of Home Elo and an increase of 23.6 fans per one point of Away Elo.

```
fit <- lm(Attendance~Home_top8+Home_top2+Home_elo+Away_elo+Home_top8:Home_top2, data=subset_afl)
plot(fit)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_top8 + Home_top2 + Home_elo + Away_elo + Home_top8:Hor
154
114
352
60000
20000
0
−40000
20000
30000
40000
50000


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top8 + Home_top2 + Home_elo + Away_elo + Home_top8:Hor
154
352
194
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3






```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top8 + Home_top2 + Home_elo +
##     Away_elo + Home_top8:Home_top2, data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -35558  -9466  -1107   6947  60829
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)         -82697.011  26328.197  -3.141 0.001812 **
## Home_top8              -75.243     49.629  -1.516 0.130301
## Home_top2              607.684    571.148   1.064 0.288000
## Home_elo                57.889     16.785   3.449 0.000624 ***
## Away_elo                20.944      7.972   2.627 0.008948 **
## Home_top8:Home_top2     -4.689      5.600  -0.837 0.402885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16100 on 390 degrees of freedom
## Multiple R-squared:  0.1354, Adjusted R-squared:  0.1243
## F-statistic: 12.22 on 5 and 390 DF,  p-value: 5.141e-11
```

```
fit1 <- step(fit)
```

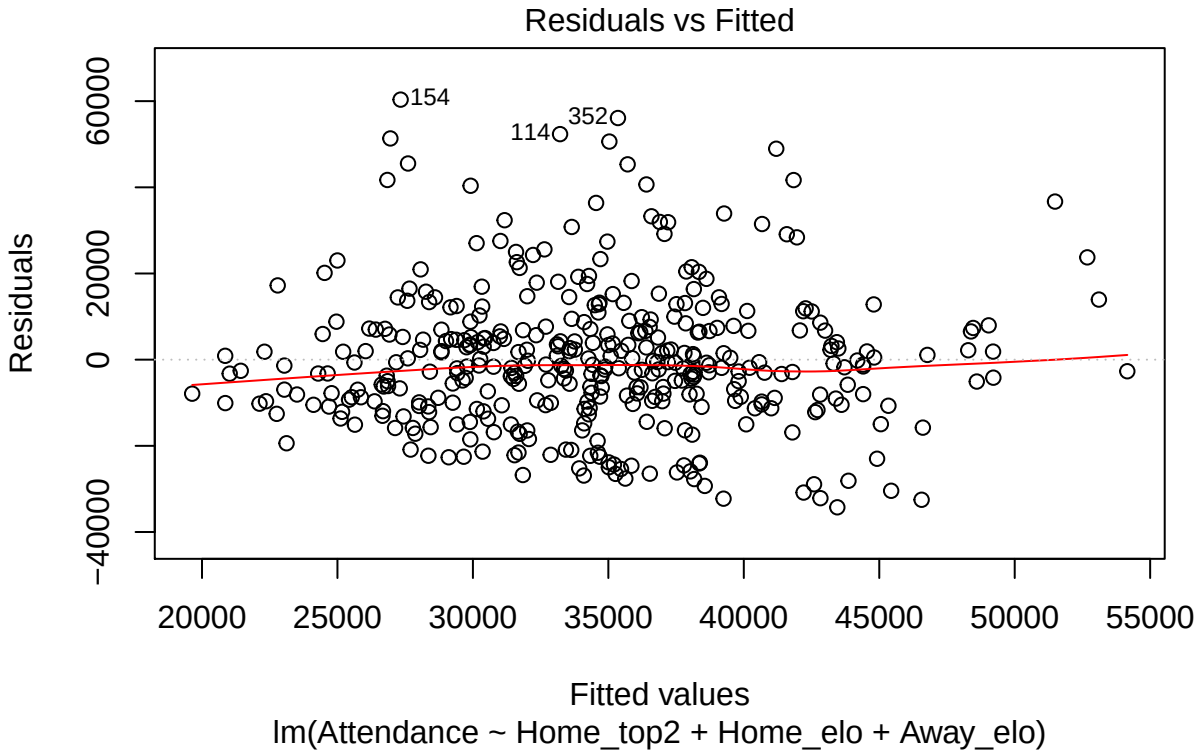
```
## Start:  AIC=7677.81
## Attendance ~ Home_top8 + Home_top2 + Home_elo + Away_elo + Home_top8:Home_top2
##
##                        Df  Sum of Sq        RSS    AIC
## - Home_top8:Home_top2   1  181804057 1.0130e+11 7676.5
## <none>                               1.0111e+11 7677.8
## - Away_elo              1 1789515726 1.0290e+11 7682.8
## - Home_elo              1 3083987294 1.0420e+11 7687.7
##
## Step:  AIC=7676.52
## Attendance ~ Home_top8 + Home_top2 + Home_elo + Away_elo
##
##             Df  Sum of Sq        RSS    AIC
## - Home_top8  1  504560693 1.0180e+11 7676.5
## <none>                    1.0130e+11 7676.5
## - Home_top2  1 1315121052 1.0261e+11 7679.6
## - Away_elo   1 1852012217 1.0315e+11 7681.7
## - Home_elo   1 3514269210 1.0481e+11 7688.0
##
## Step:  AIC=7676.49
## Attendance ~ Home_top2 + Home_elo + Away_elo
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    1.0180e+11 7676.5
## - Home_top2  1  913310675 1.0271e+11 7678.0
## - Away_elo   1 2012010767 1.0381e+11 7682.2
## - Home_elo   1 4443742079 1.0624e+11 7691.4
```

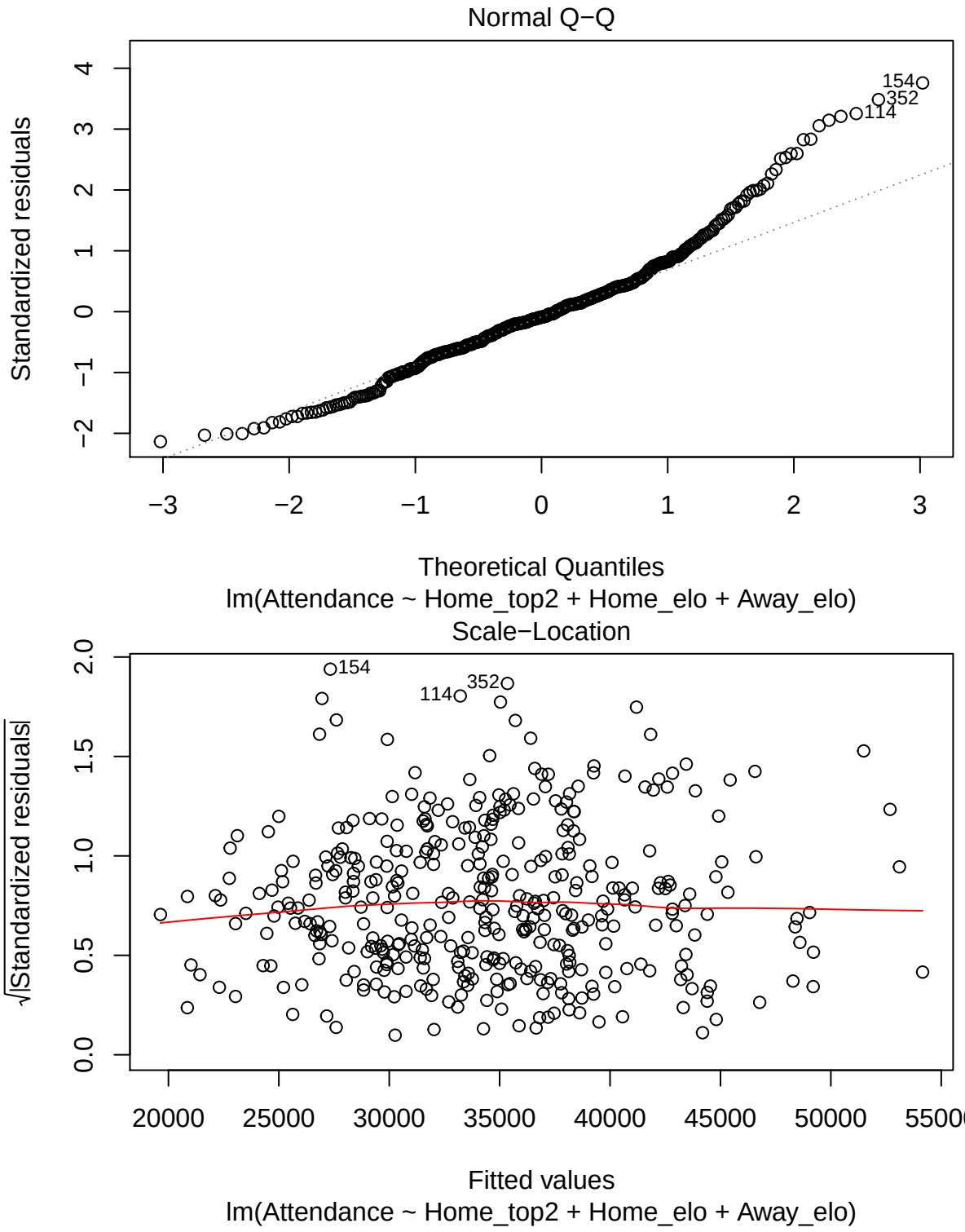
```
summary(fit1)
```

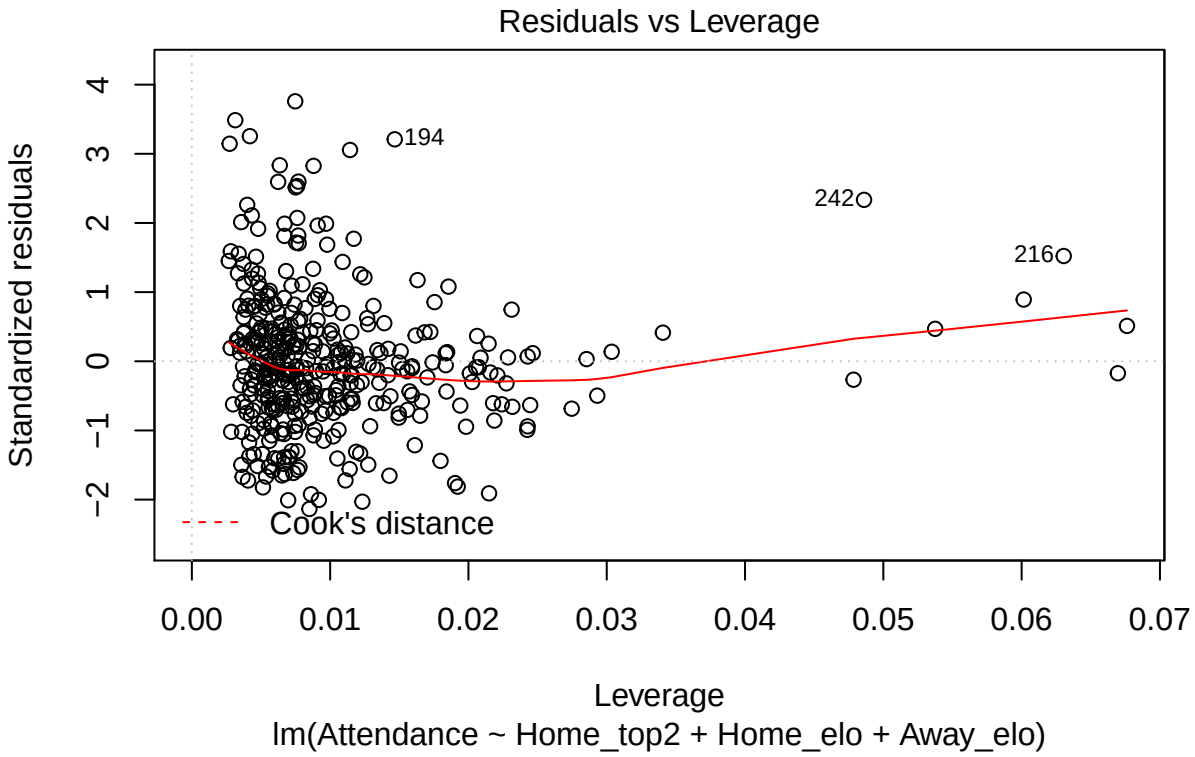
```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo, data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34267  -9785  -1433   7010  60354
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -64008.711  19990.068  -3.202  0.00148 **
## Home_top2      102.638     54.731   1.875  0.06149 .
## Home_elo        42.720     10.327   4.137 4.32e-05 ***
## Away_elo        22.116      7.946   2.783  0.00564 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16120 on 392 degrees of freedom
## Multiple R-squared:  0.1296, Adjusted R-squared:  0.1229
## F-statistic: 19.45 on 3 and 392 DF,  p-value: 8.962e-12
```

```
plot(fit1)
```

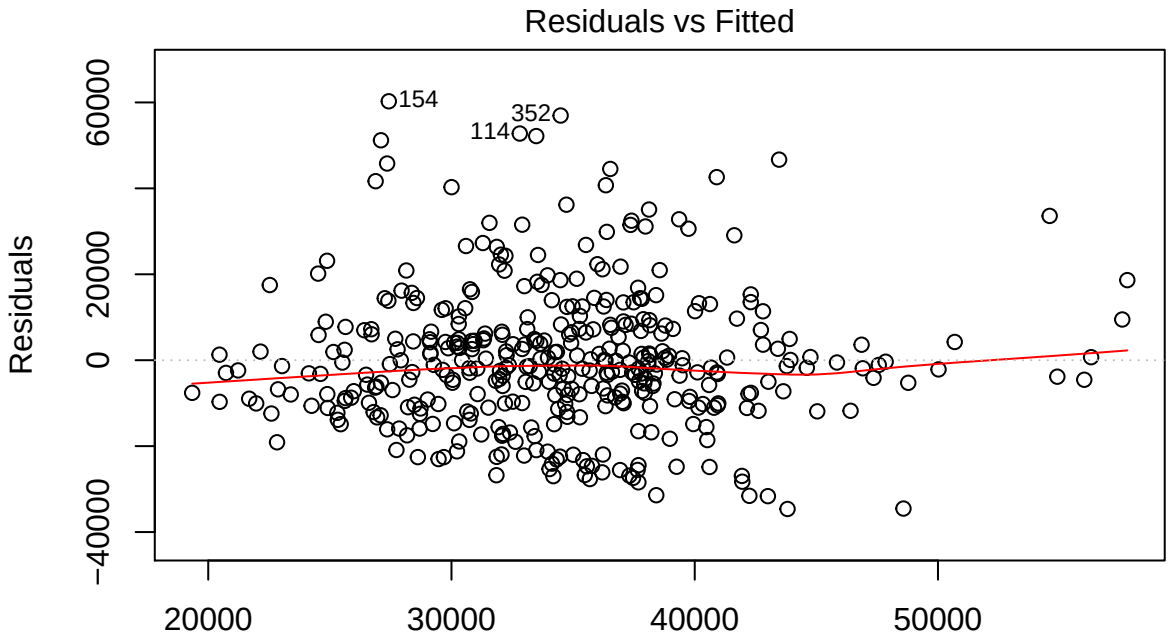


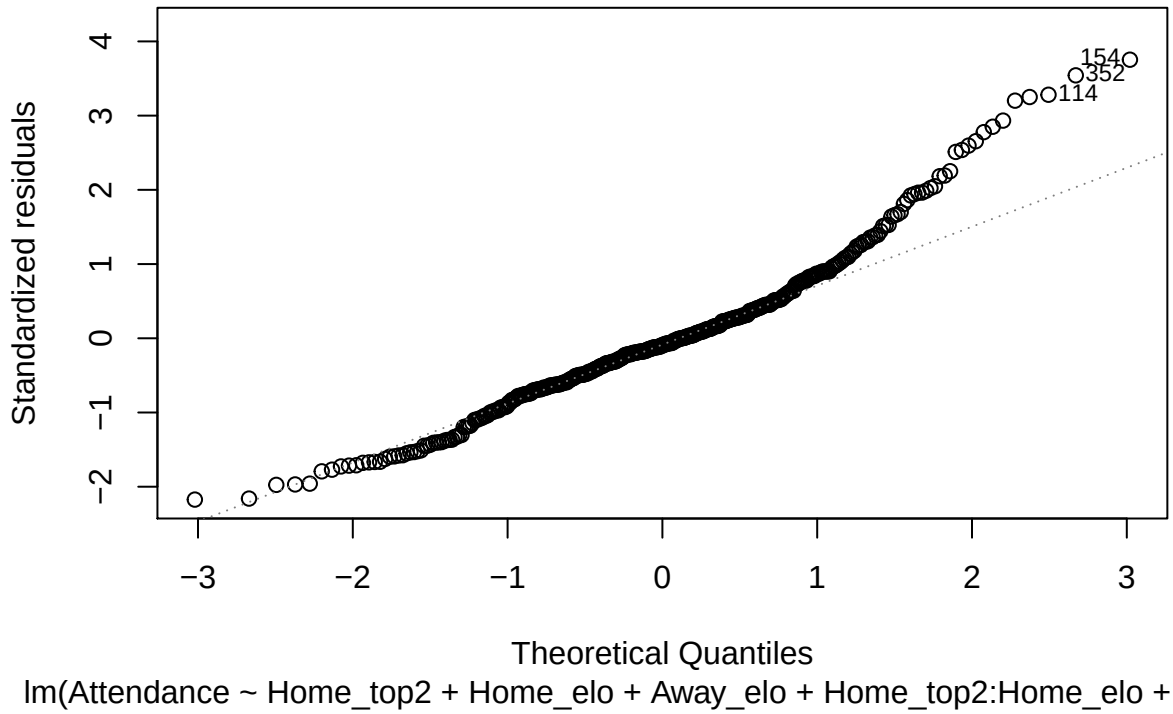

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo)
154
352
114
Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo)

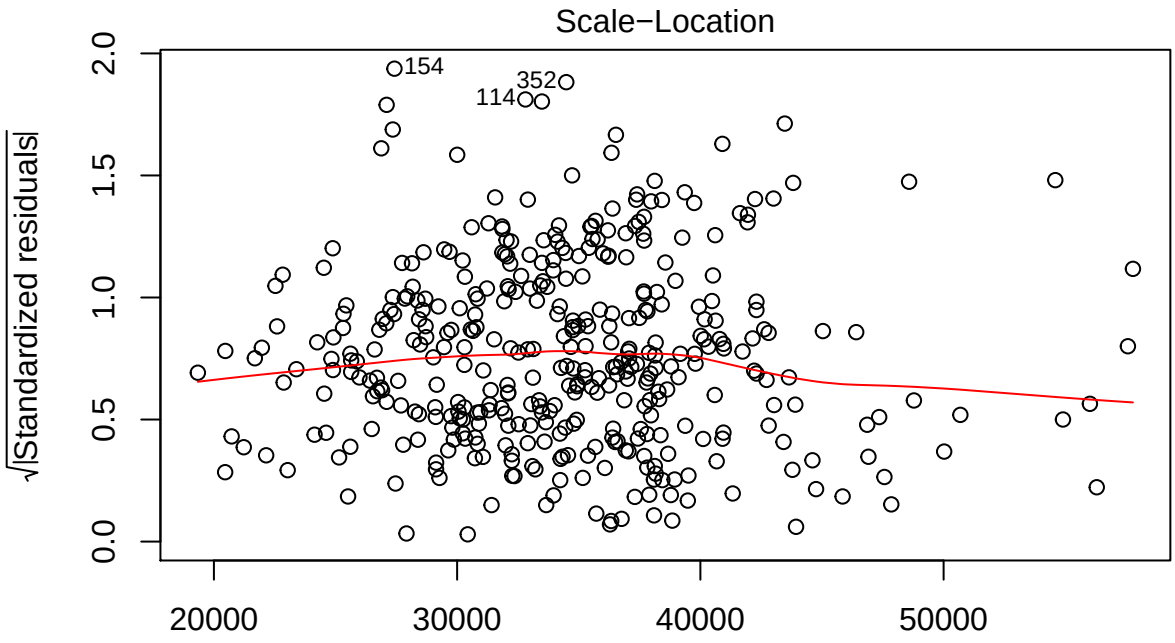


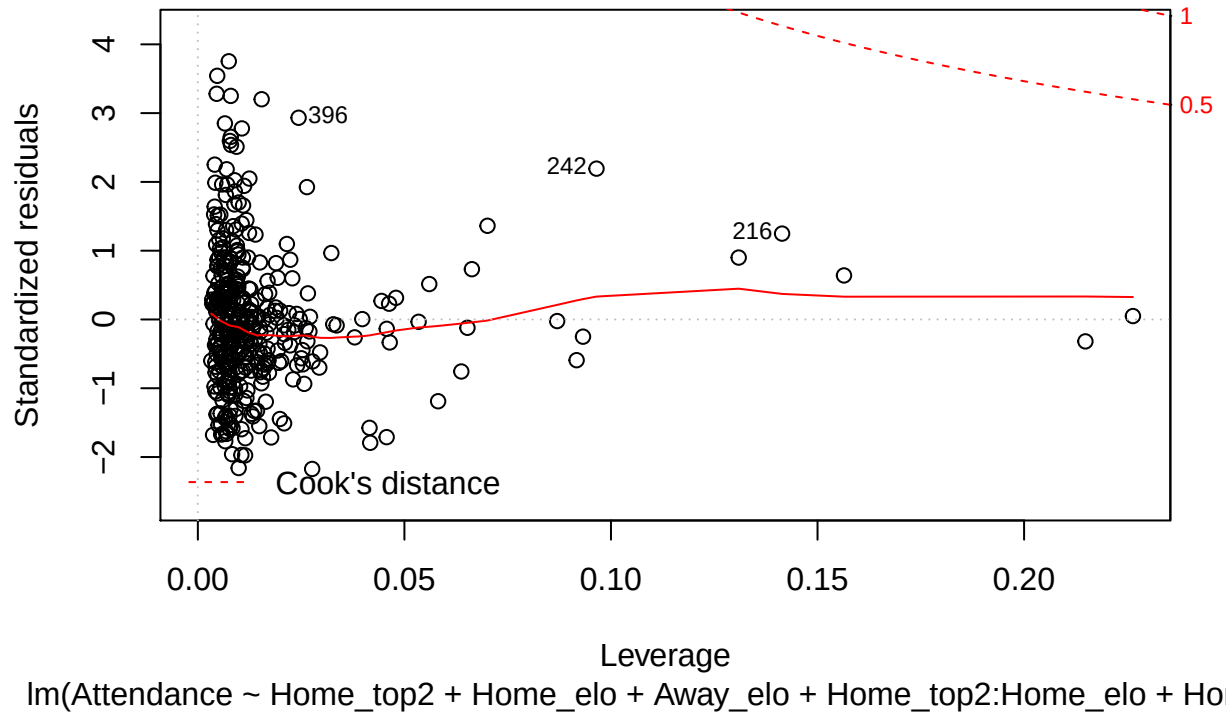
Adding Home and Away Elo to the model with Home playoff probability and Home championship probability makes the model better. A stepwise regression shows that we should remove Home playoff probability from the model. Our new model shows that as the Home championship probability increases by 1 percent, we can expect an increase in the number of fans of 103, while Home Elo increases the attendance by 42.7 fans per point and Away Elo increases the number of fans by 22 per point.

```
fit <- lm(Attendance~Home_top2+Home_elo+Away_elo+Home_top2:Home_elo+Home_top2:Away_elo, data=subset_afl
plot(fit)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Home_top2:Home_elo + Hom
154
352
114
60000
20000
0
−40000
20000
30000
40000
50000


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Home_top2:Home_elo + Hom
154
352
114
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3






```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Home_top2:Home_elo +
##     Home_top2:Away_elo, data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34630  -9985  -1463   7220  60263
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -7.050e+04  2.093e+04  -3.368 0.000832 ***
## Home_top2          -1.947e+03  1.861e+03  -1.046 0.296077
## Home_elo            4.504e+01  1.056e+01   4.265 2.52e-05 ***
## Away_elo            2.449e+01  9.117e+00   2.686 0.007543 **
## Home_top2:Home_elo  1.439e+00  1.053e+00   1.366 0.172782
## Home_top2:Away_elo -2.098e-01  4.812e-01  -0.436 0.663044
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16110 on 390 degrees of freedom
## Multiple R-squared:  0.134,  Adjusted R-squared:  0.1229
## F-statistic: 12.07 on 5 and 390 DF,  p-value: 6.938e-11
```

```
fit1 <- step(fit)
```

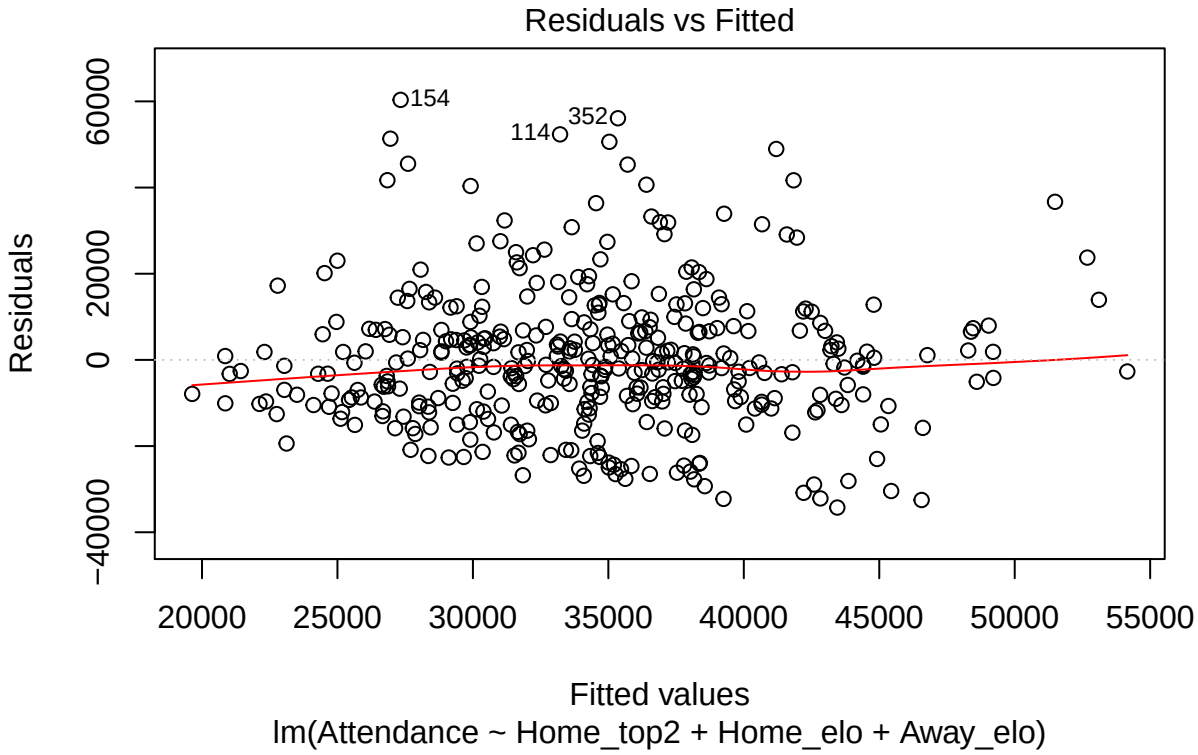
```
## Start:  AIC=7678.45
## Attendance ~ Home_top2 + Home_elo + Away_elo + Home_top2:Home_elo +
##     Home_top2:Away_elo
##
##                       Df Sum of Sq        RSS    AIC
## - Home_top2:Away_elo  1  49375481 1.0133e+11 7676.6
## - Home_top2:Home_elo  1 484434043 1.0176e+11 7678.3
## <none>                             1.0128e+11 7678.4
##
## Step:  AIC=7676.64
## Attendance ~ Home_top2 + Home_elo + Away_elo + Home_top2:Home_elo
##
##                       Df  Sum of Sq        RSS    AIC
## - Home_top2:Home_elo  1  473658414 1.0180e+11 7676.5
## <none>                              1.0133e+11 7676.6
## - Away_elo            1 2086911773 1.0341e+11 7682.7
##
## Step:  AIC=7676.49
## Attendance ~ Home_top2 + Home_elo + Away_elo
##
##             Df  Sum of Sq        RSS    AIC
## <none>                     1.0180e+11 7676.5
## - Home_top2  1  913310675 1.0271e+11 7678.0
## - Away_elo   1 2012010767 1.0381e+11 7682.2
## - Home_elo   1 4443742079 1.0624e+11 7691.4
```

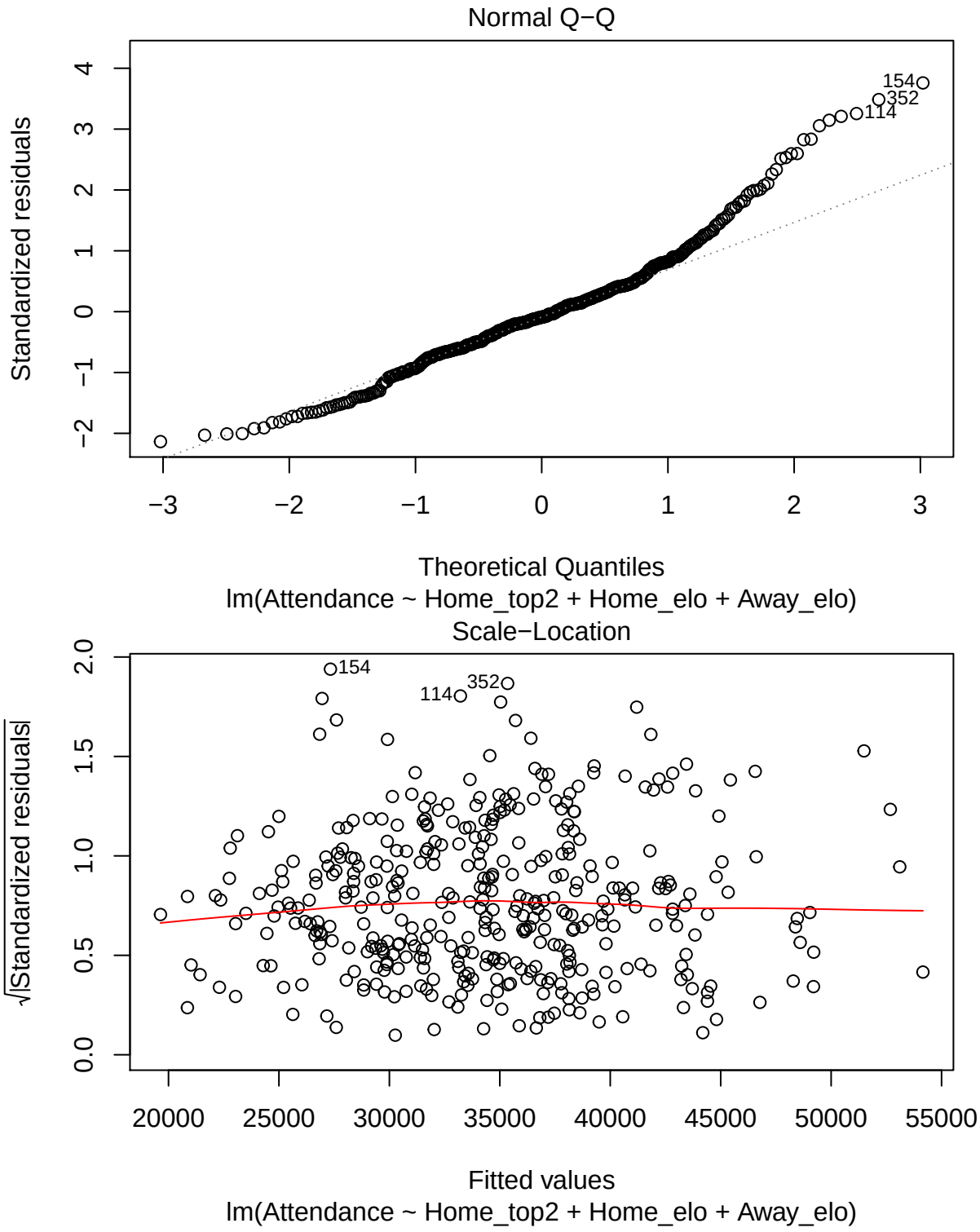
```
summary(fit1)
```

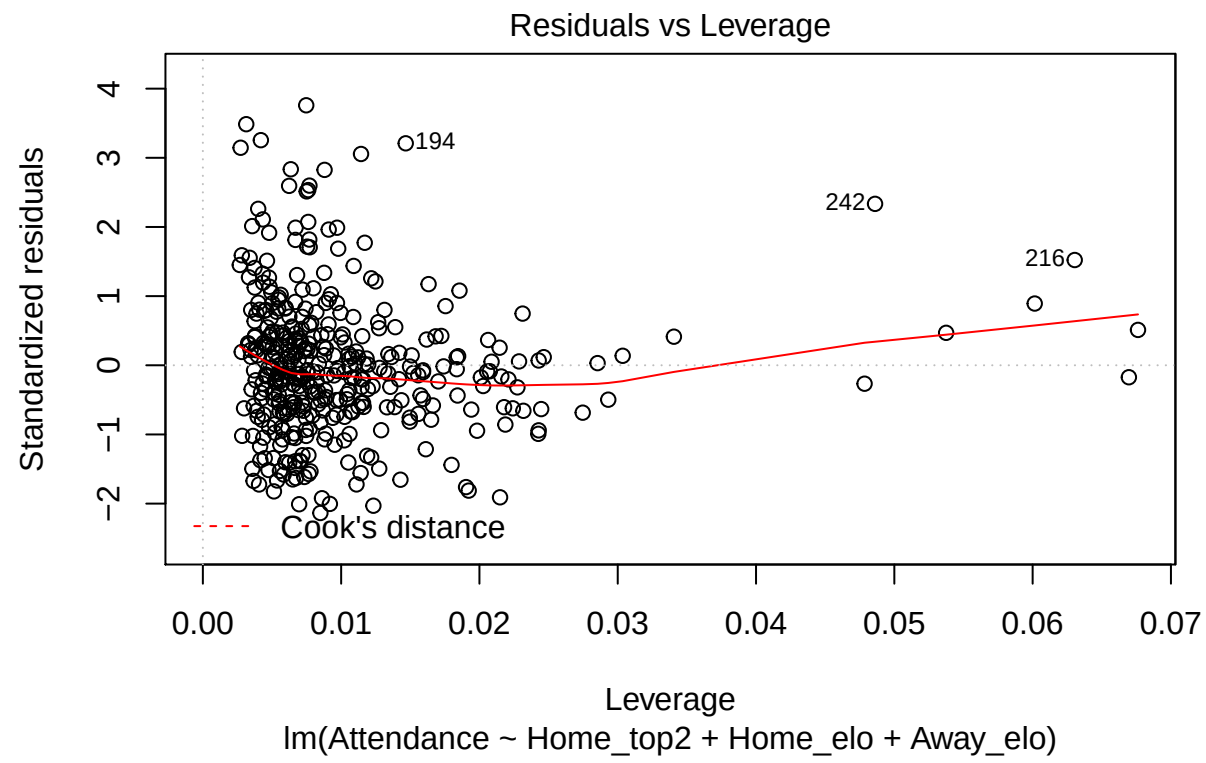
```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo, data = subset_afl)
##
## Residuals:
```

```
##    Min     1Q Median     3Q    Max
## -34267  -9785  -1433   7010  60354
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -64008.711  19990.068  -3.202  0.00148 **
## Home_top2      102.638     54.731   1.875  0.06149 .
## Home_elo        42.720     10.327   4.137 4.32e-05 ***
## Away_elo        22.116      7.946   2.783  0.00564 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16120 on 392 degrees of freedom
## Multiple R-squared:  0.1296, Adjusted R-squared:  0.1229
## F-statistic: 19.45 on 3 and 392 DF,  p-value: 8.962e-12
```

```
plot(fit1)
```

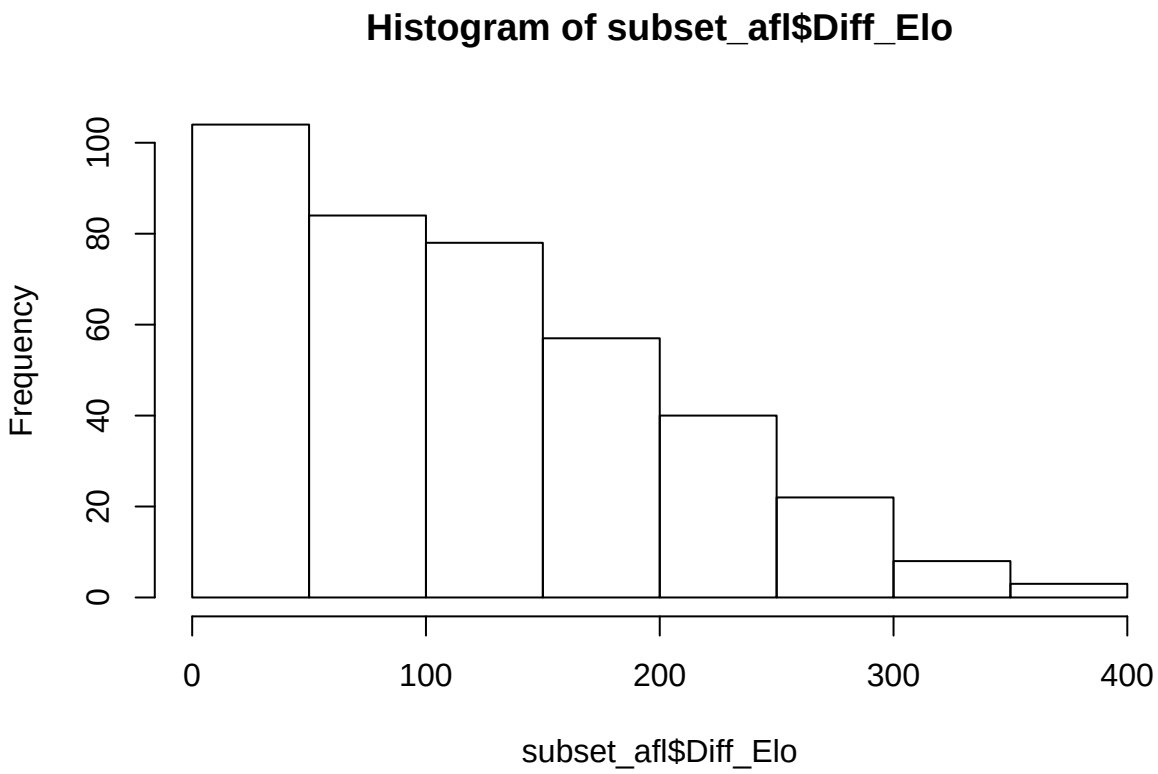



Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo)
154
352
114
Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo)



Adding in interaction terms does not improve the model.

```
hist(subset_afl$Diff_Elo)
```

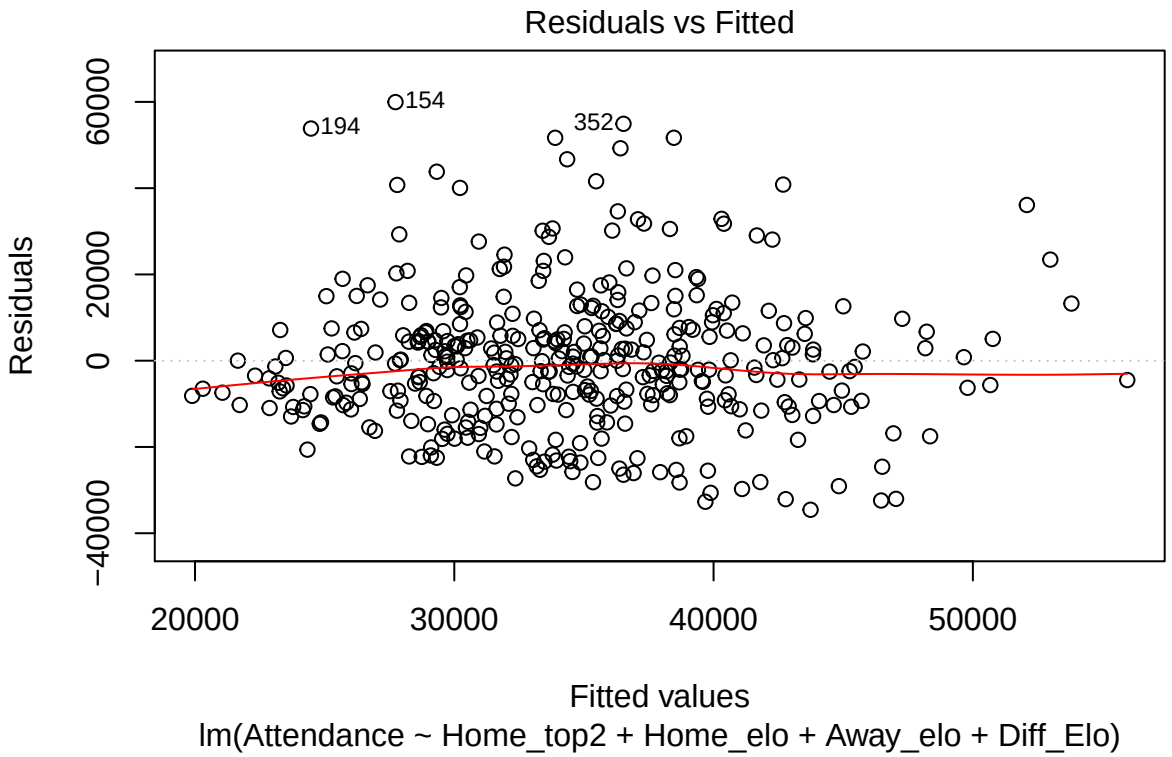


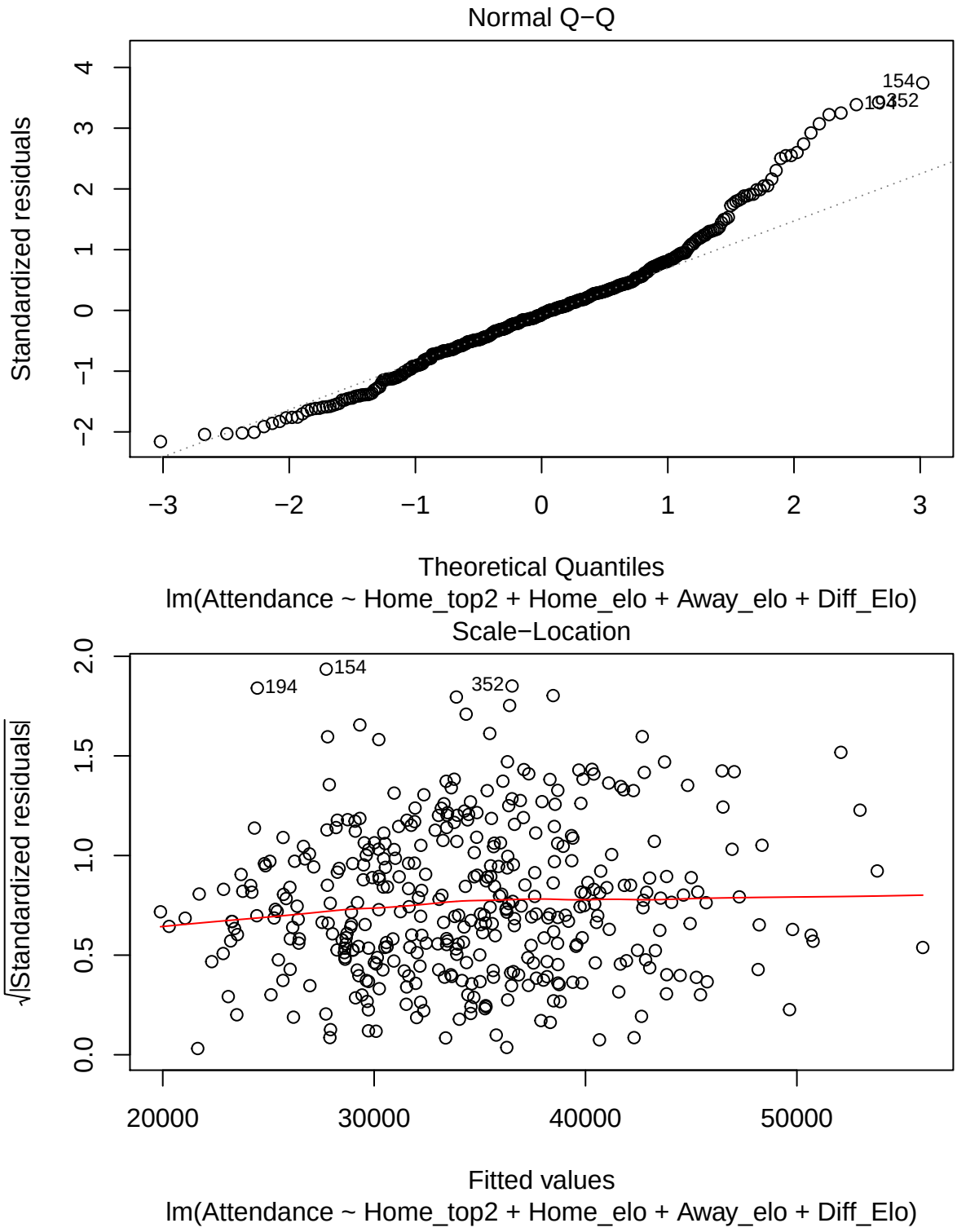

```
fit <- lm(Attendance~Diff_Elo, data = subset_afl)
summary(fit)
```

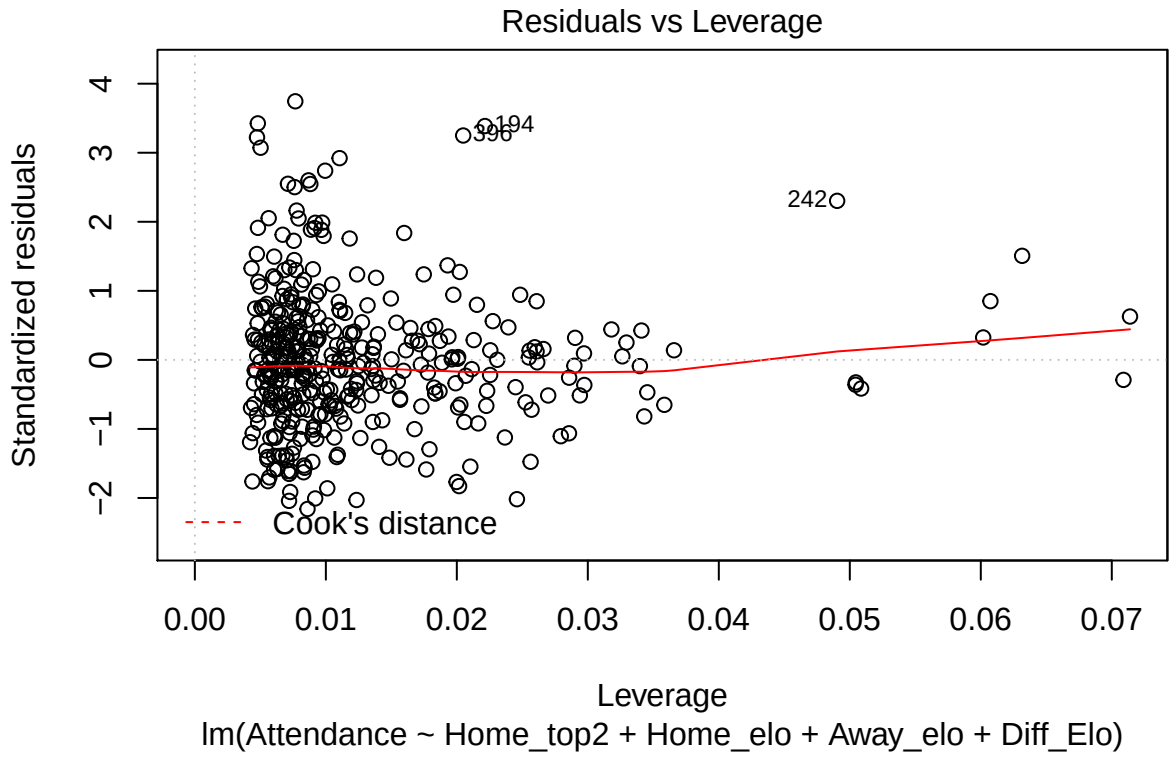
```
##
## Call:
## lm(formula = Attendance ~ Diff_Elo, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -30908 -13031   -821   9342  60584
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 38288.03    1478.04  25.905  < 2e-16 ***
## Diff_Elo      -32.66      10.17  -3.213  0.00142 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17010 on 394 degrees of freedom
## Multiple R-squared:  0.02553,    Adjusted R-squared:  0.02306
## F-statistic: 10.32 on 1 and 394 DF,  p-value: 0.001423
```

Difference in Elo is monotonically decreasing as games are more likely to be between evenly matched teams. The linear model shows it to be a significant predictor.

```
fit <- lm(Attendance~Home_top2+Home_elo+Away_elo+Diff_Elo, data = subset_afl)
plot(fit)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo)
Scale−Location
√|Standardized residuals|
154
194
352
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo)




```
summary(fit)
```

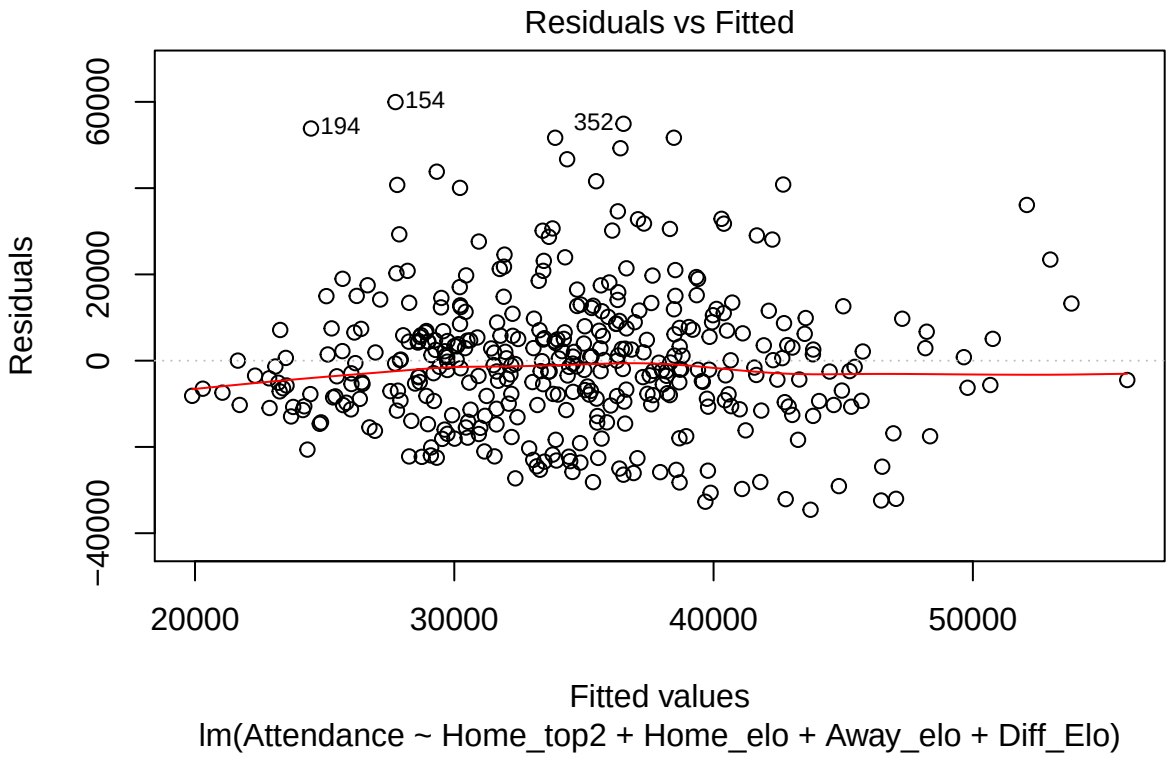
```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo,
##     data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34555  -9650  -1068   7083  59953
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -48678.620  21705.837  -2.243 0.025479 *
## Home_top2      118.080     55.260   2.137 0.033233 *
## Home_elo        38.082     10.621   3.585 0.000379 ***
## Away_elo        17.871      8.273   2.160 0.031359 *
## Diff_Elo       -18.333     10.270  -1.785 0.075033 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16070 on 391 degrees of freedom
## Multiple R-squared:  0.1366, Adjusted R-squared:  0.1278
## F-statistic: 15.46 on 4 and 391 DF,  p-value: 9.4e-12
```

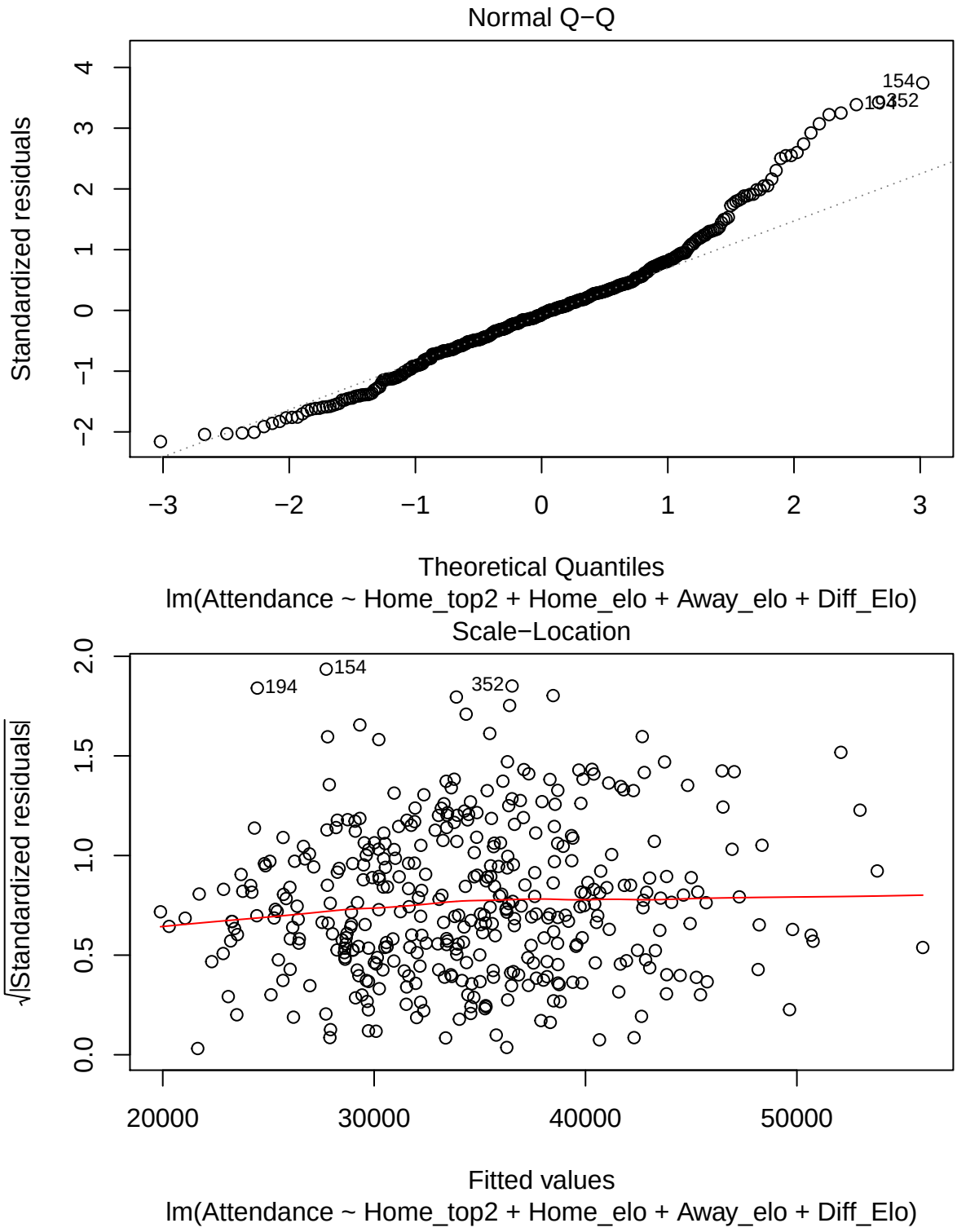
```
fit1 <- step(fit)
```

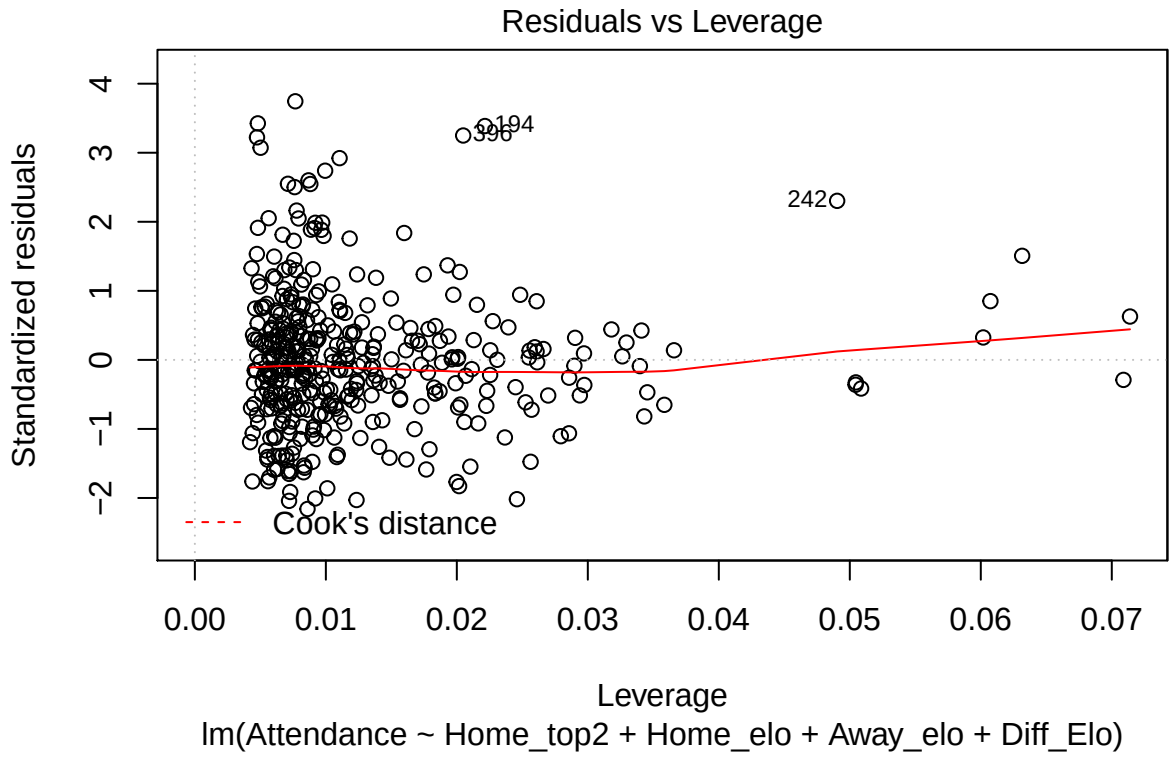
```
## Start:  AIC=7675.27
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo
##
##             Df  Sum of Sq        RSS    AIC
```

```
## <none>                    1.0098e+11 7675.3
## - Diff_Elo    1  822872878 1.0180e+11 7676.5
## - Home_top2   1 1179174887 1.0216e+11 7677.9
## - Away_elo    1 1205206581 1.0218e+11 7678.0
## - Home_elo    1 3319775118 1.0430e+11 7686.1
```

```
plot(fit1)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo)
Scale−Location
√|Standardized residuals|
154
194
352
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo)

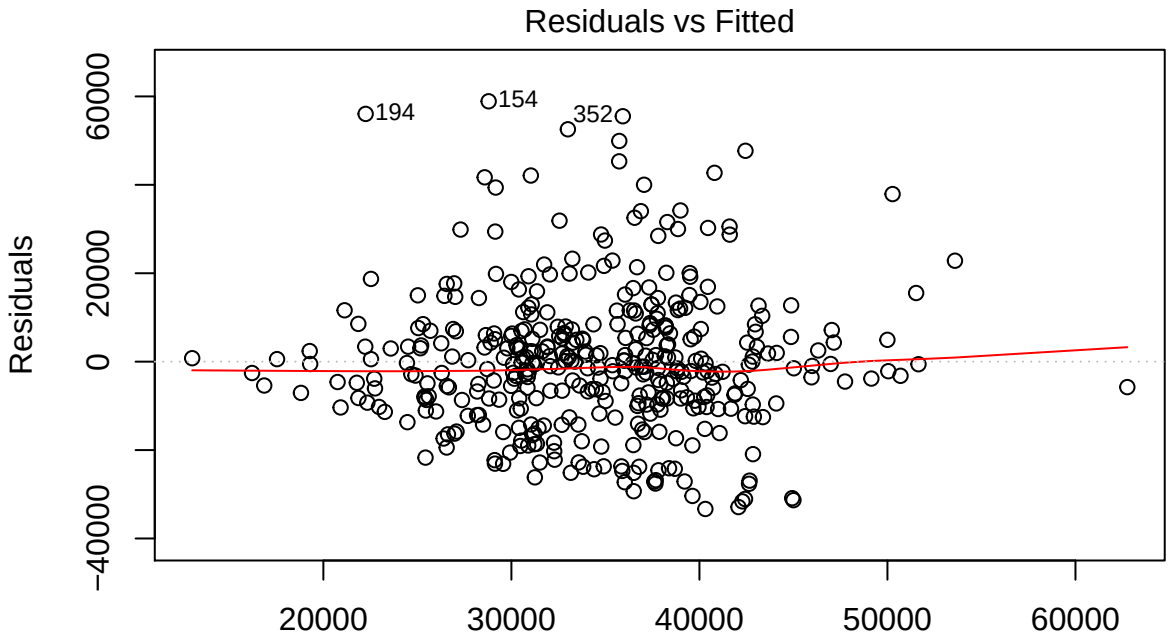


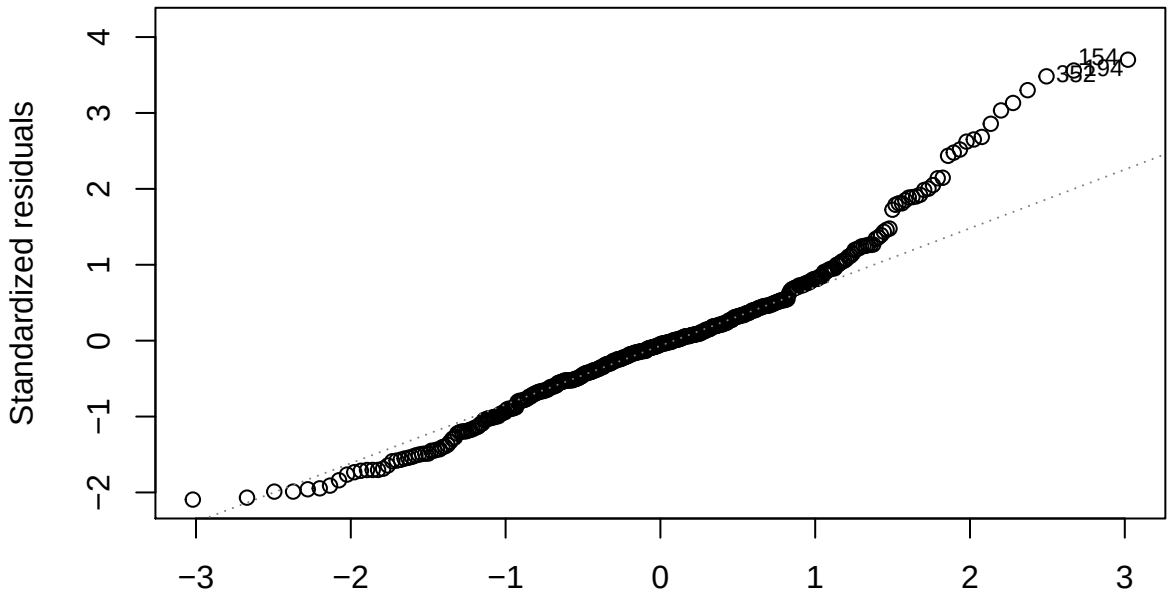

```
summary(fit1)
```

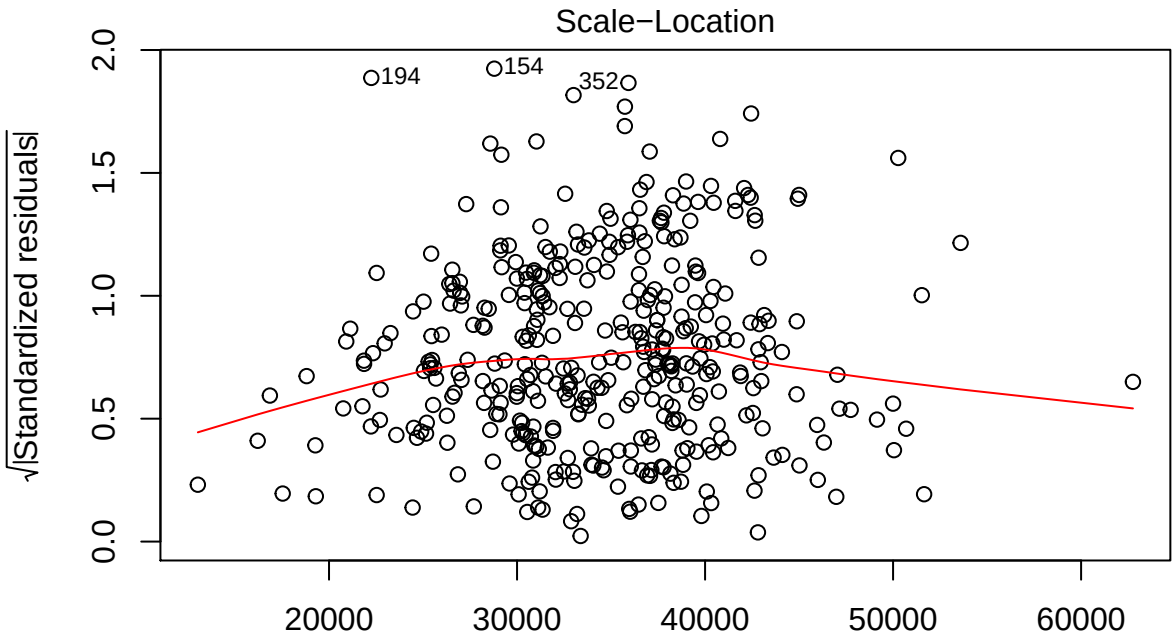
```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo,
##     data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34555  -9650  -1068   7083  59953
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -48678.620  21705.837  -2.243 0.025479 *
## Home_top2      118.080     55.260   2.137 0.033233 *
## Home_elo        38.082     10.621   3.585 0.000379 ***
## Away_elo        17.871      8.273   2.160 0.031359 *
## Diff_Elo       -18.333     10.270  -1.785 0.075033 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16070 on 391 degrees of freedom
## Multiple R-squared:  0.1366, Adjusted R-squared:  0.1278
## F-statistic: 15.46 on 4 and 391 DF,  p-value: 9.4e-12
```

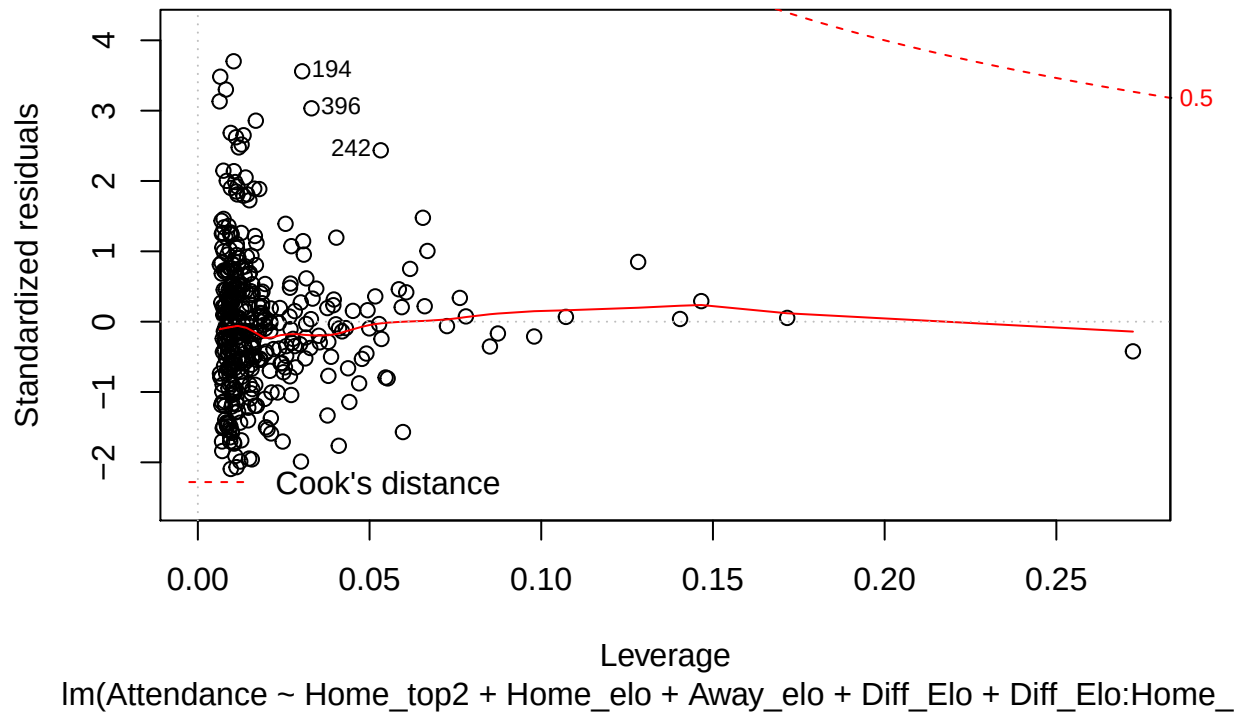
Difference in Elo is a good predictor to add to the model. None of the other predictors should be removed from the model.

```
fit <- lm(Attendance~Home_top2+Home_elo+Away_elo+Diff_Elo+Diff_Elo:Home_top2+Diff_Elo:Home_elo+Diff_Elo
plot(fit)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Diff_Elo:Home_ .
194
154
352
60000
20000
0
−40000
20000
30000
40000
50000
60000


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Diff_Elo:Home_ .
154
352
194
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3






```
summary(fit)
```

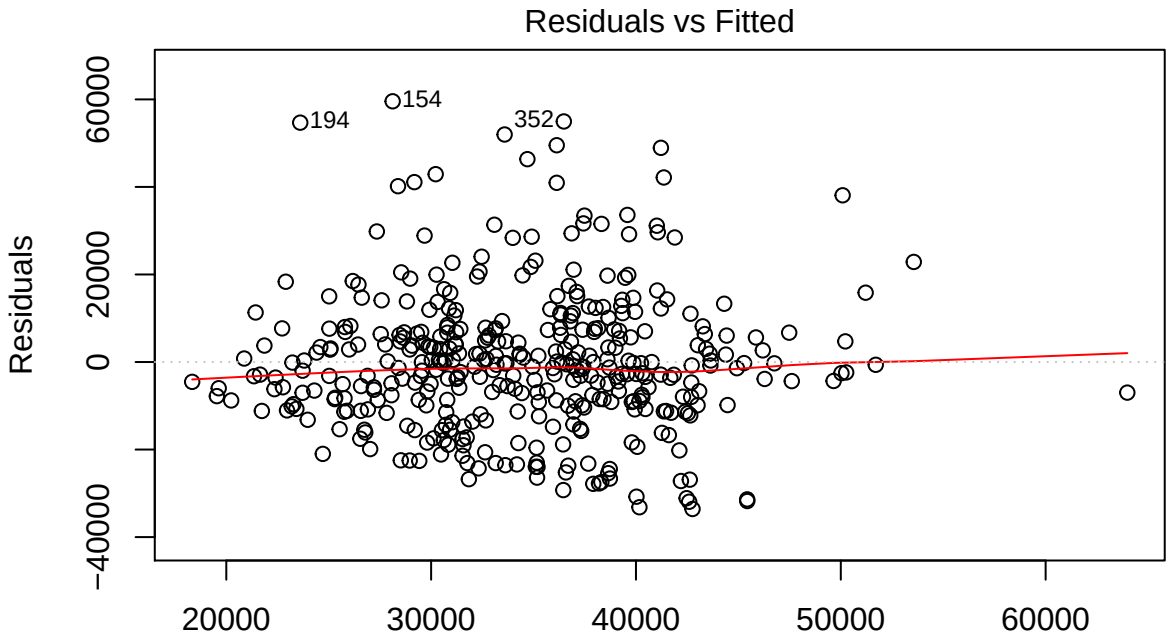
```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo +
##     Diff_Elo:Home_top2 + Diff_Elo:Home_elo + Diff_Elo:Away_elo,
```

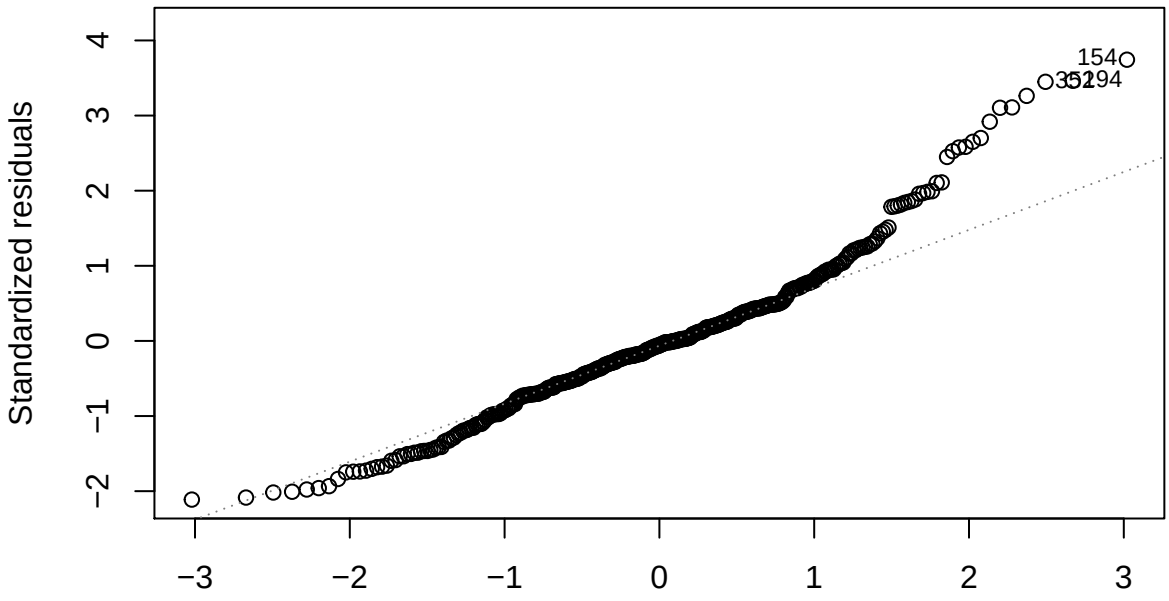
```
##     data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33327  -9311   -737   7204  58894
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -3.332e+04  3.470e+04  -0.960    0.338
## Home_top2           1.877e+01  9.083e+01   0.207    0.836
## Home_elo            1.448e+01  2.232e+01   0.649    0.517
## Away_elo            3.200e+01  1.957e+01   1.635    0.103
## Diff_Elo           -1.799e+02  2.662e+02  -0.676    0.500
## Home_top2:Diff_Elo  9.939e-01  6.411e-01   1.550    0.122
## Home_elo:Diff_Elo   1.166e-01  1.268e-01   0.920    0.358
## Away_elo:Diff_Elo  -1.538e-02  1.100e-01  -0.140    0.889
##
## Residual standard error: 15990 on 388 degrees of freedom
## Multiple R-squared:  0.1515, Adjusted R-squared:  0.1362
## F-statistic: 9.894 on 7 and 388 DF,  p-value: 2.25e-11
```

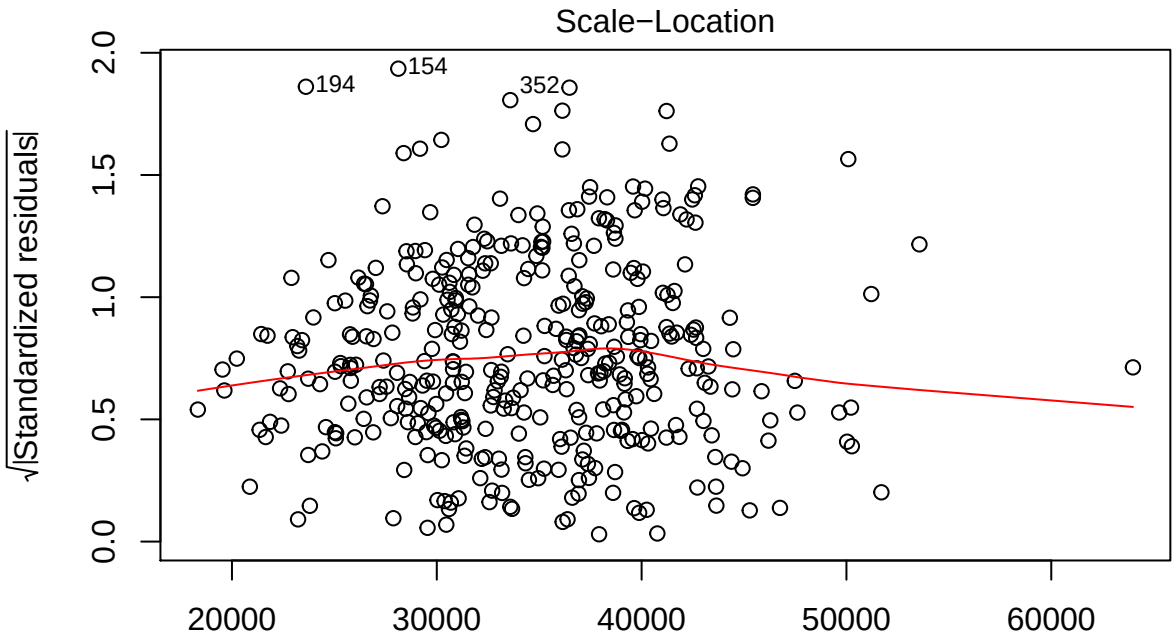
```
fit1 <- step(fit)
```

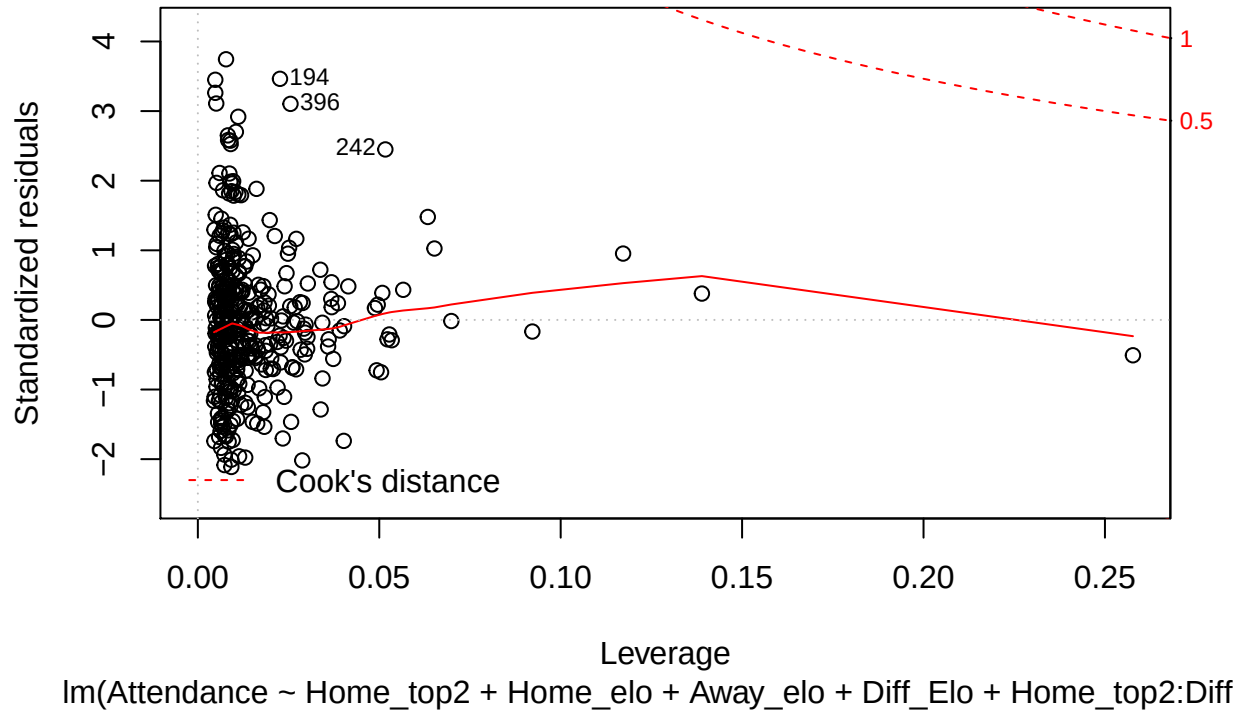
```
## Start:  AIC=7674.39
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Diff_Elo:Home_top2 +
##     Diff_Elo:Home_elo + Diff_Elo:Away_elo
##
##                       Df Sum of Sq        RSS    AIC
## - Away_elo:Diff_Elo    1   4997985 9.9243e+10 7672.4
## - Home_elo:Diff_Elo    1 216482274 9.9454e+10 7673.3
## <none>                             9.9238e+10 7674.4
## - Home_top2:Diff_Elo   1 614629551 9.9852e+10 7674.8
##
## Step:  AIC=7672.41
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Home_elo:Diff_Elo
##
##                       Df  Sum of Sq        RSS    AIC
## - Home_elo:Diff_Elo    1  235461739 9.9478e+10 7671.3
## <none>                             9.9243e+10 7672.4
## - Home_top2:Diff_Elo   1  623460890 9.9866e+10 7672.9
## - Away_elo             1 2119070871 1.0136e+11 7678.8
##
## Step:  AIC=7671.35
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                               9.9478e+10 7671.3
## - Home_top2:Diff_Elo   1 1498888462 1.0098e+11 7675.3
## - Away_elo             1 2017461720 1.0150e+11 7677.3
## - Home_elo             1 2471896704 1.0195e+11 7679.1
```

```
plot(fit1)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff .
194
154
352
60000
20000
0
−40000
20000
30000
40000
50000
60000


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff .
154
352
194
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3






```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo +
##     Home_top2:Diff_Elo, data = subset_afl)
```

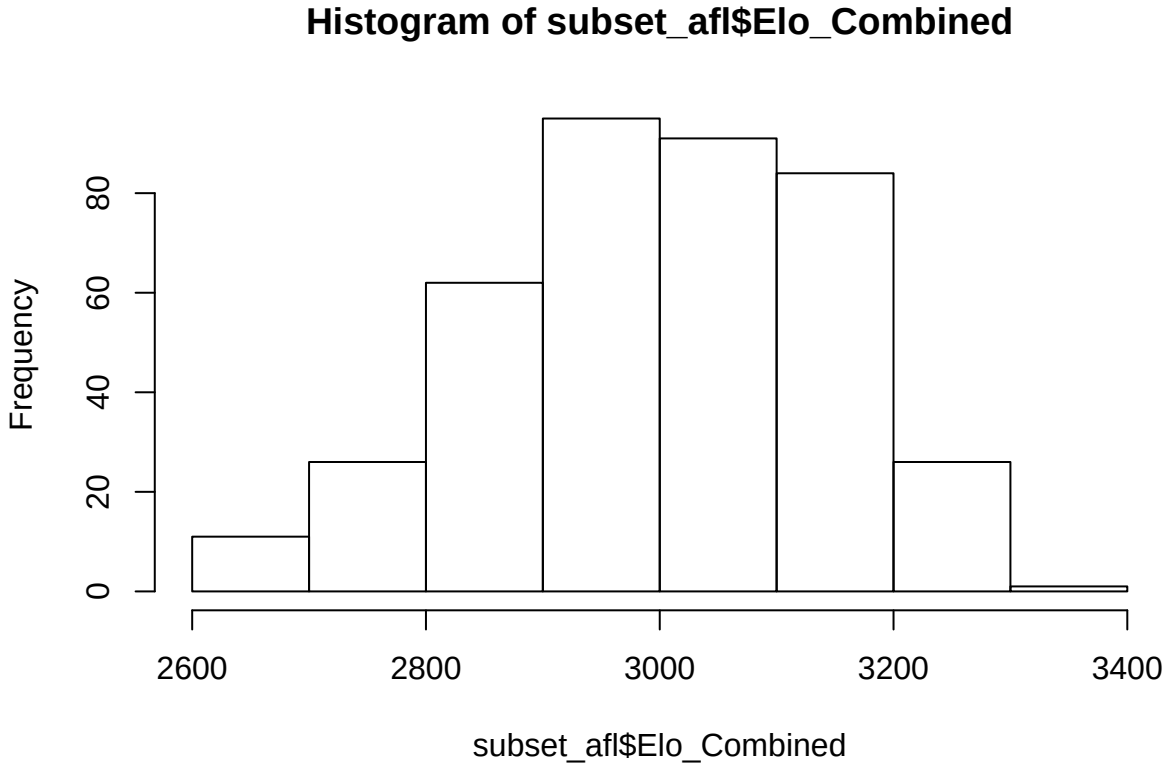
```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33577  -9273   -904   7257  59563
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.968e+04  2.158e+04  -2.302  0.02184 *
## Home_top2          -2.035e+01  7.923e+01  -0.257  0.79743
## Home_elo            3.341e+01  1.073e+01   3.113  0.00199 **
## Away_elo            2.429e+01  8.638e+00   2.812  0.00517 **
## Diff_Elo           -3.283e+01  1.183e+01  -2.775  0.00578 **
## Home_top2:Diff_Elo  1.285e+00  5.303e-01   2.424  0.01580 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15970 on 390 degrees of freedom
## Multiple R-squared:  0.1494, Adjusted R-squared:  0.1385
## F-statistic:  13.7 on 5 and 390 DF,  p-value: 2.465e-12
```

```
vif(fit1)
```

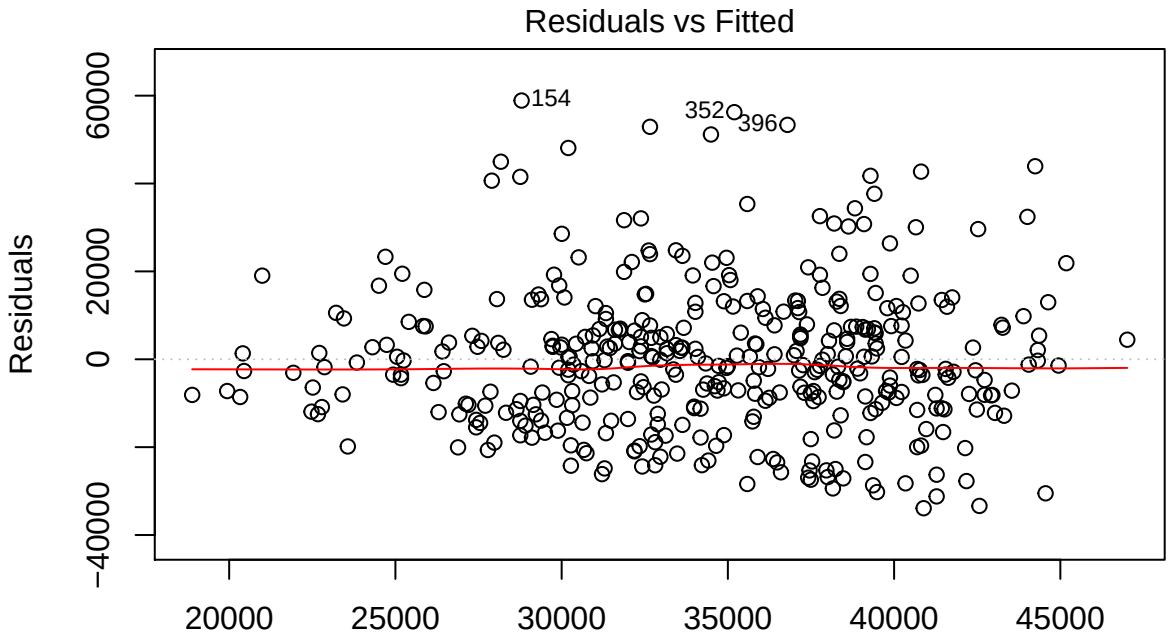
```
##          Home_top2           Home_elo           Away_elo
##           3.544770           1.819406           1.216143
##           Diff_Elo Home_top2:Diff_Elo
##           1.535136           3.790851
```

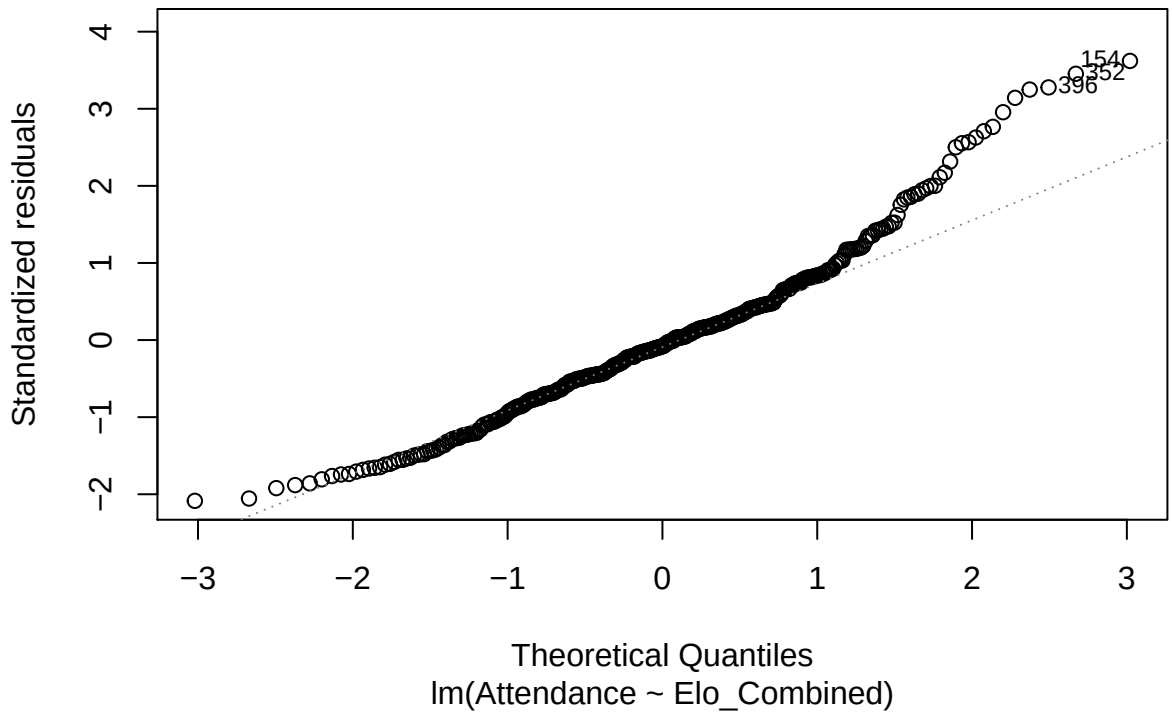
The interaction between Home championship probability and Difference in Elo does appear to be a significant predictor, although it makes Home championship probability on its own have a higher p-value.

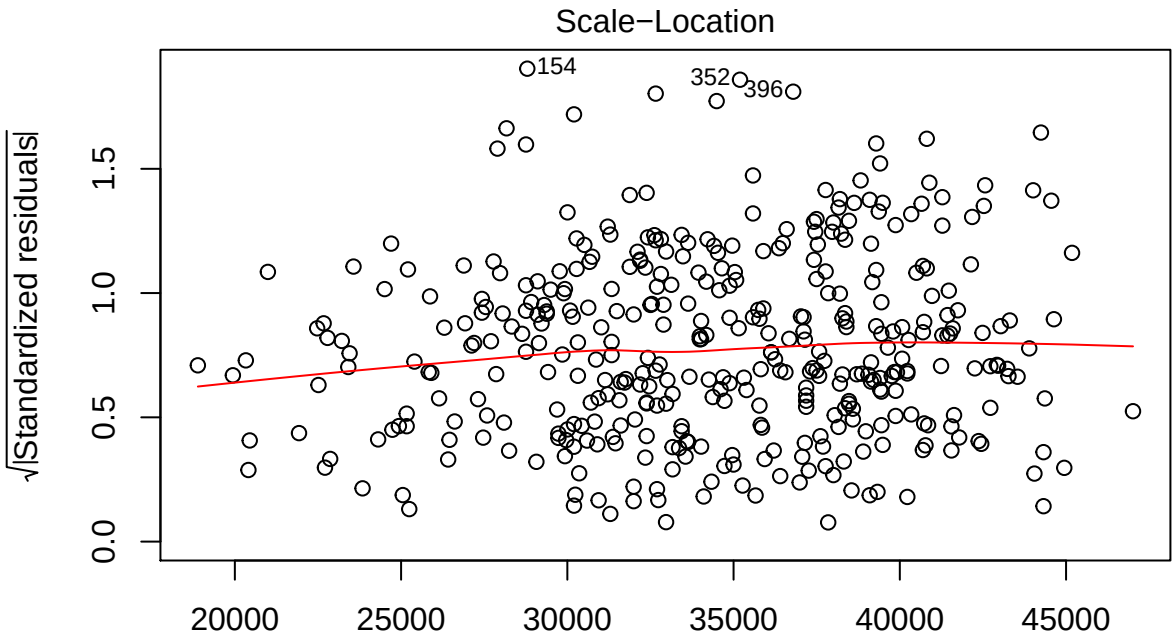
```
hist(subset_afl$Elo_Combined)
```

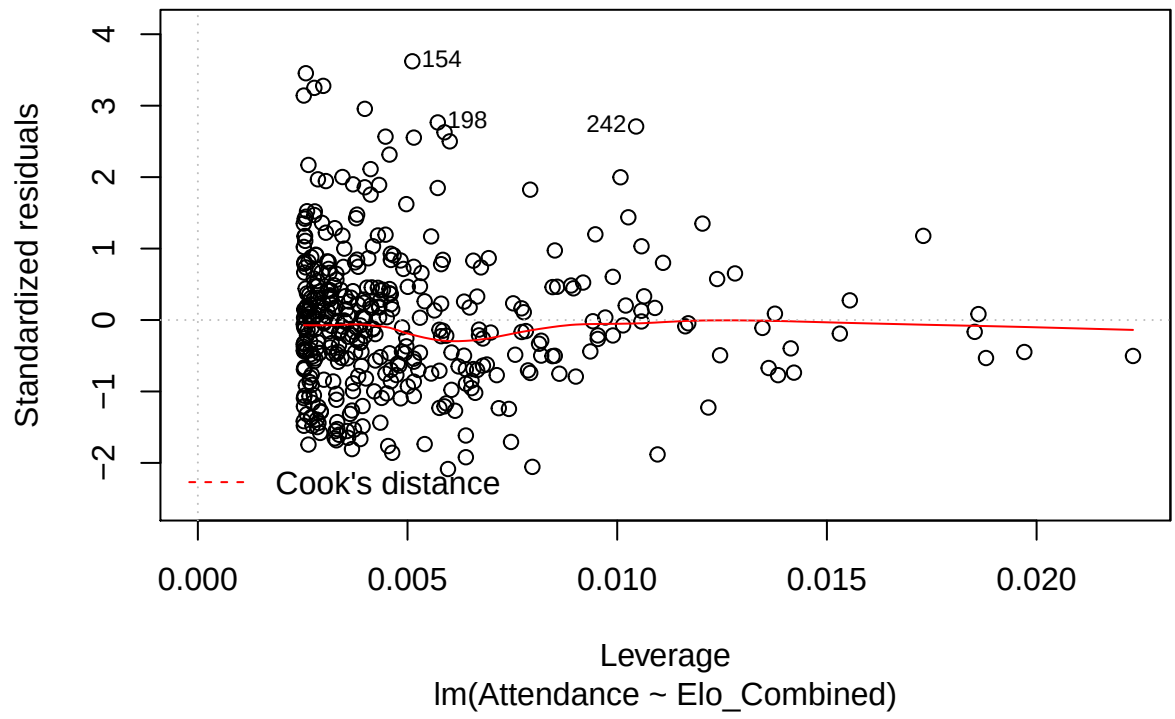


```r
fit <- lm(Attendance~Elo_Combined, data = subset_afl)
plot(fit)
```







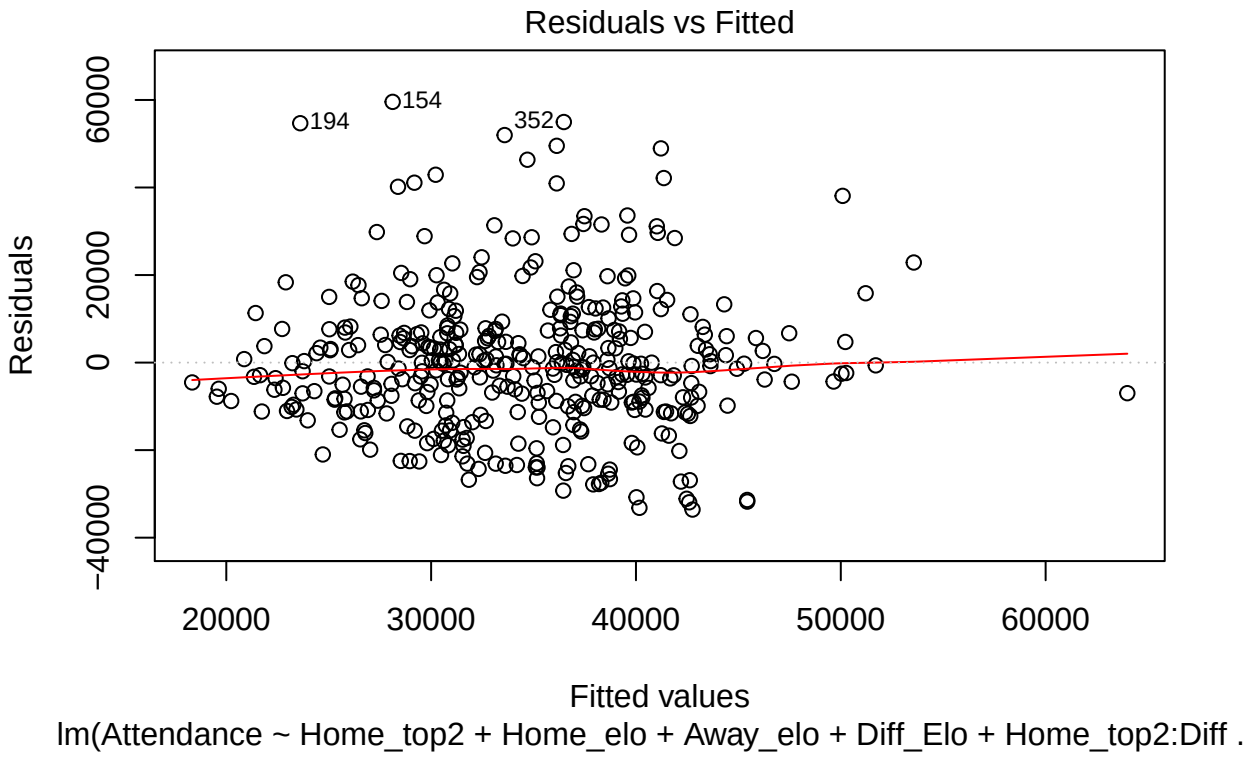


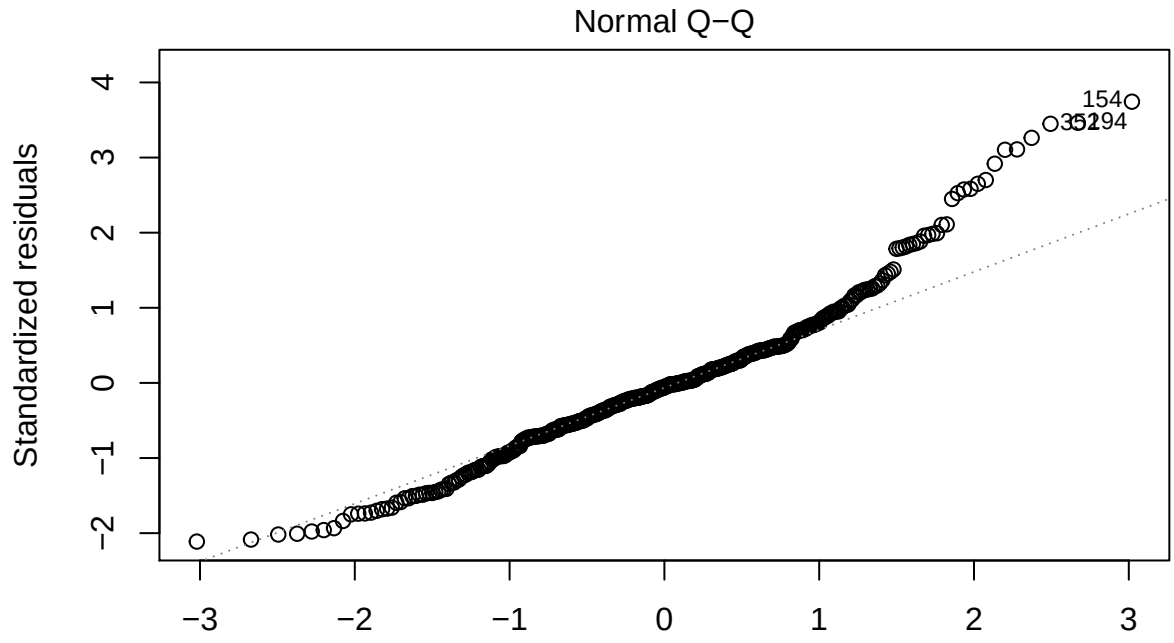

```
summary(fit)
```

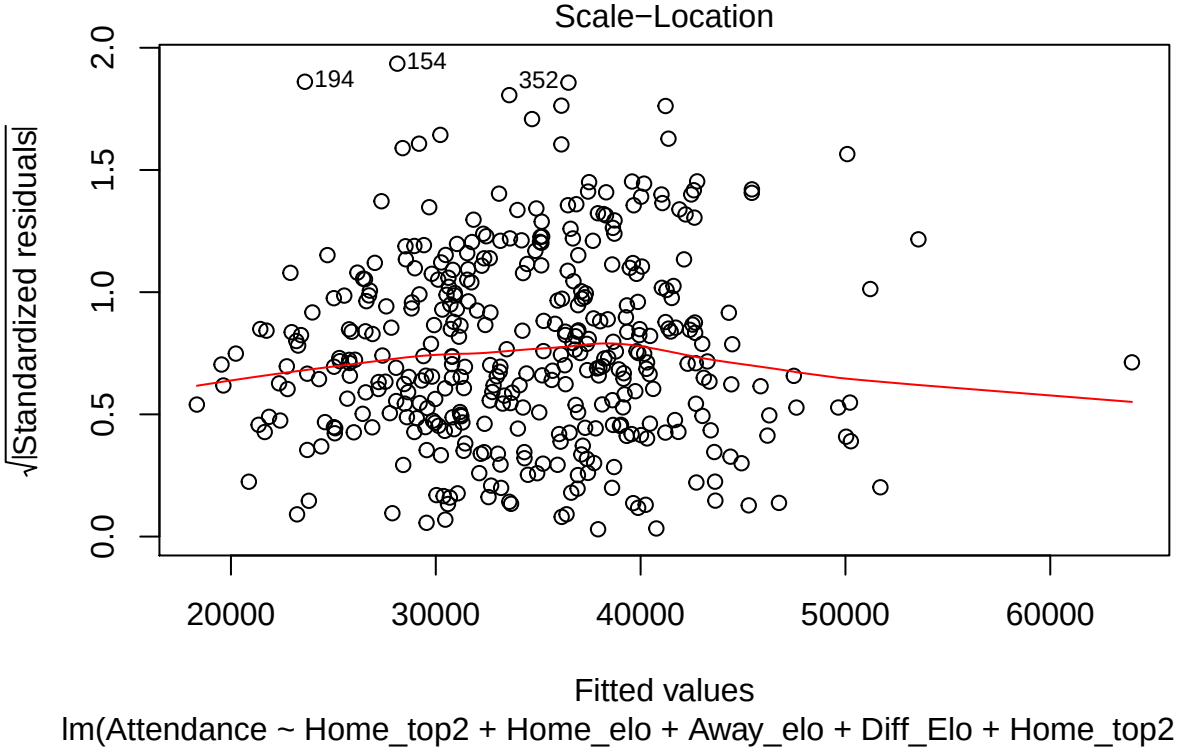
```
##
## Call:
## lm(formula = Attendance ~ Elo_Combined, data = subset_afl)
##
```

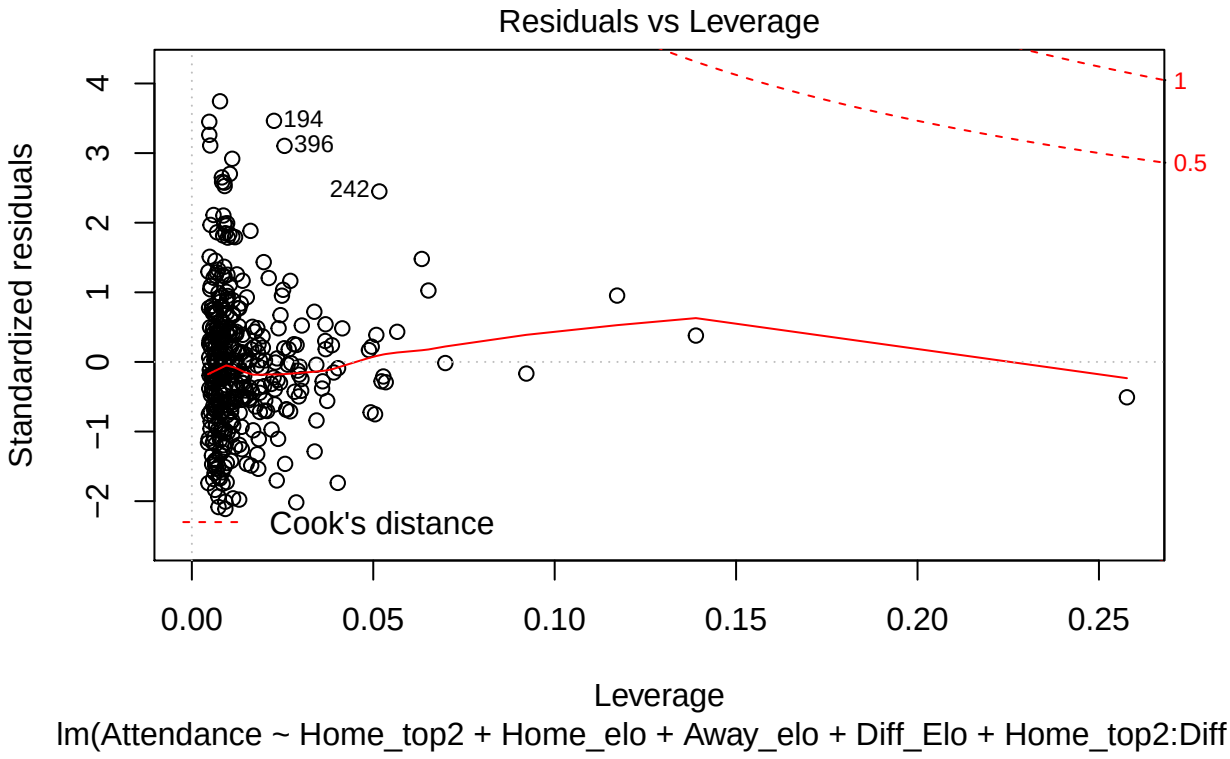
```
## Residuals:
##    Min     1Q Median     3Q    Max
## -33900 -10479  -1279   7552  58889
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -82619.926  17302.442  -4.775 2.54e-06 ***
## Elo_Combined     39.011      5.761   6.772 4.64e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16310 on 394 degrees of freedom
## Multiple R-squared:  0.1042, Adjusted R-squared:  0.102
## F-statistic: 45.85 on 1 and 394 DF,  p-value: 4.643e-11
```

```
fit <- lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo + Elo_Combined, d
plot(fit)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff .
154
352
194


Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff .
194
154
352




```
summary(fit)
```

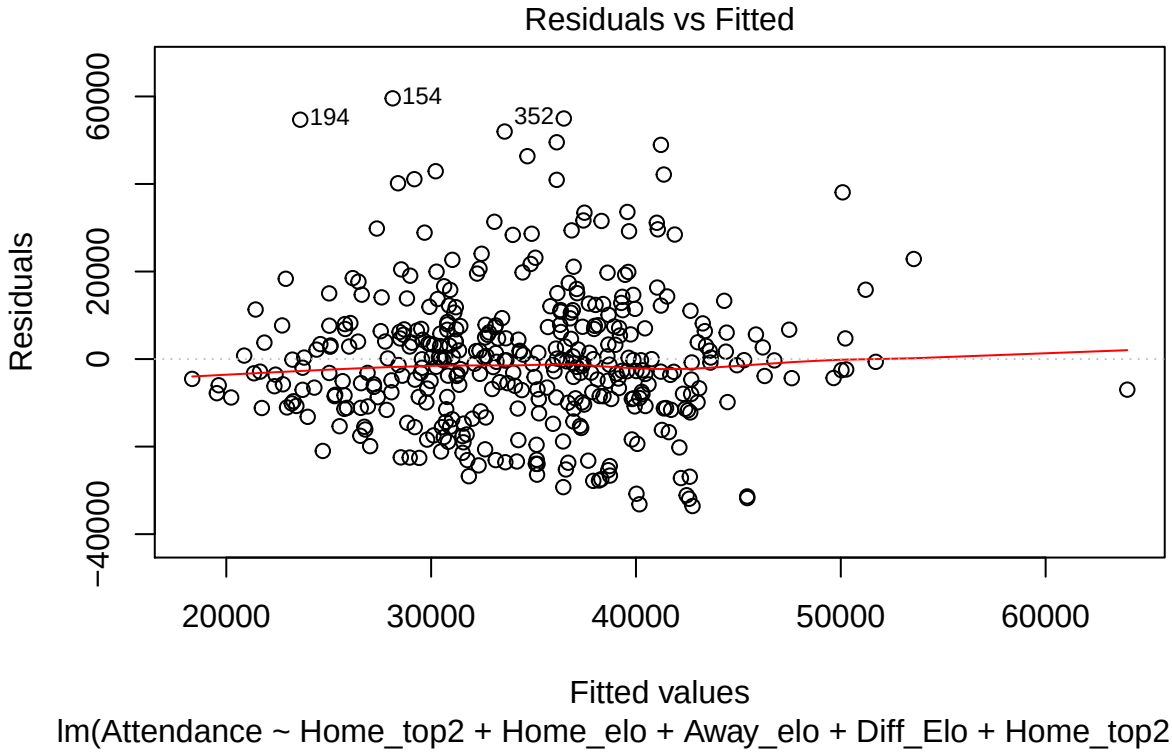
```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo +
##     Home_top2:Diff_Elo + Elo_Combined, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33577  -9273   -904   7257  59563
##
## Coefficients: (1 not defined because of singularities)
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.968e+04  2.158e+04  -2.302  0.02184 *
## Home_top2          -2.035e+01  7.923e+01  -0.257  0.79743
## Home_elo            3.341e+01  1.073e+01   3.113  0.00199 **
## Away_elo            2.429e+01  8.638e+00   2.812  0.00517 **
## Diff_Elo           -3.283e+01  1.183e+01  -2.775  0.00578 **
## Elo_Combined               NA         NA      NA       NA
## Home_top2:Diff_Elo  1.285e+00  5.303e-01   2.424  0.01580 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15970 on 390 degrees of freedom
## Multiple R-squared:  0.1494, Adjusted R-squared:  0.1385
## F-statistic:  13.7 on 5 and 390 DF,  p-value: 2.465e-12
```

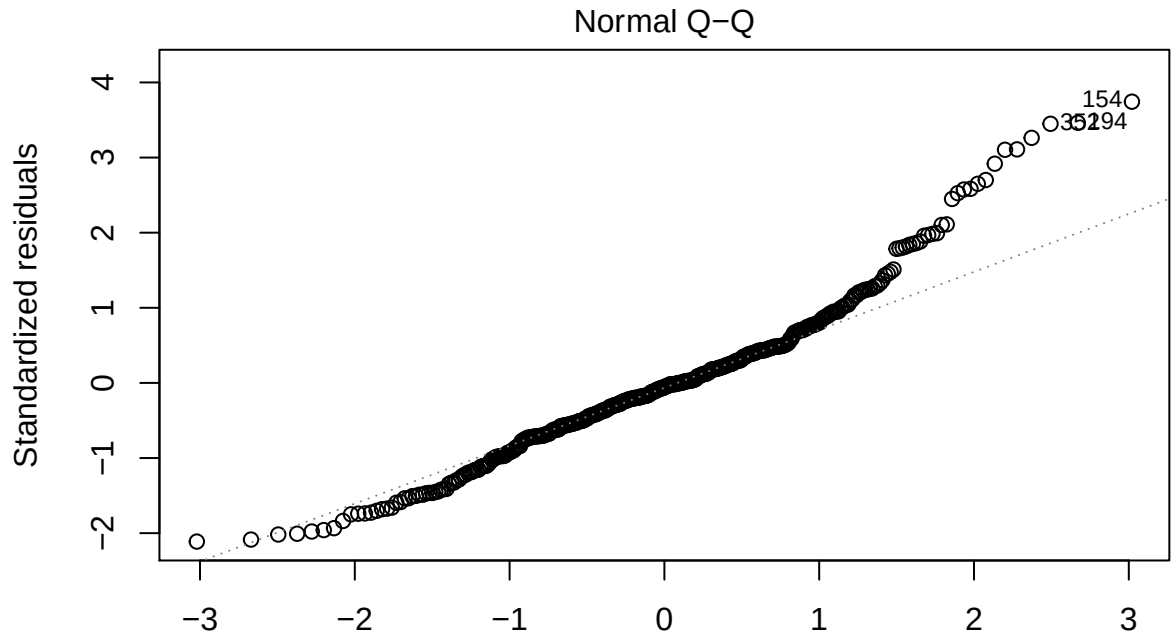
```
fit1 <- step(fit)
```

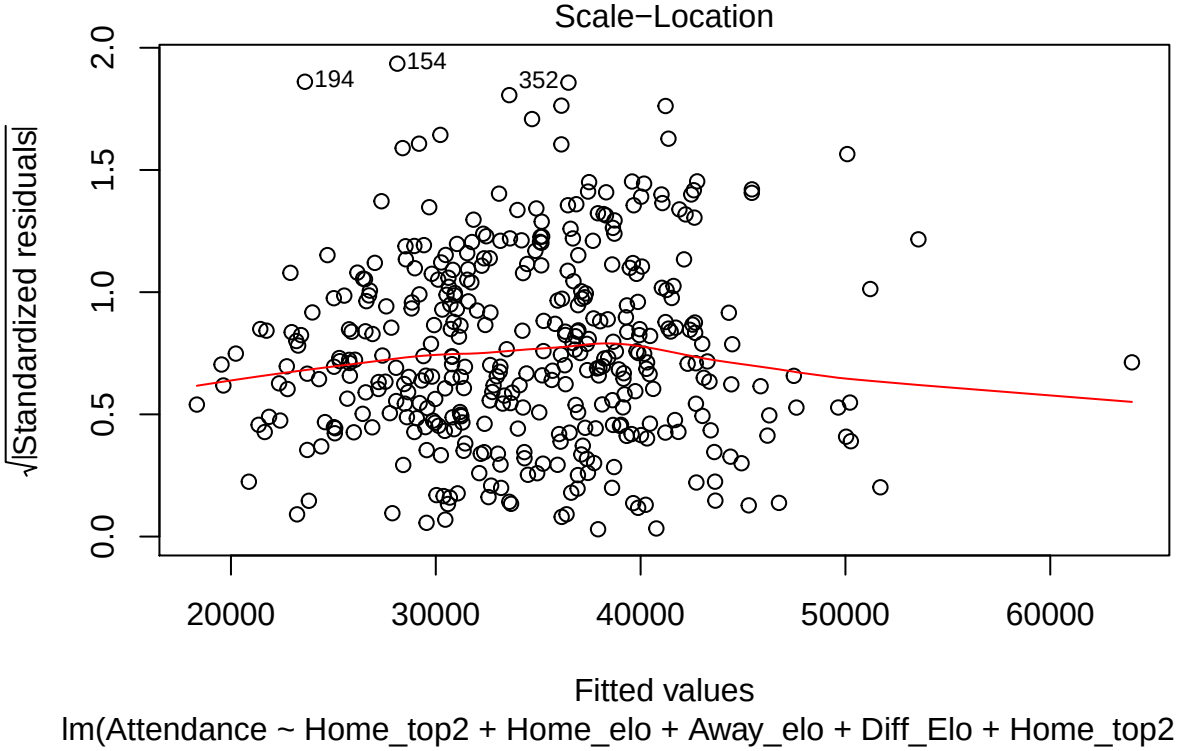
```
## Start:  AIC=7671.35
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo +
```

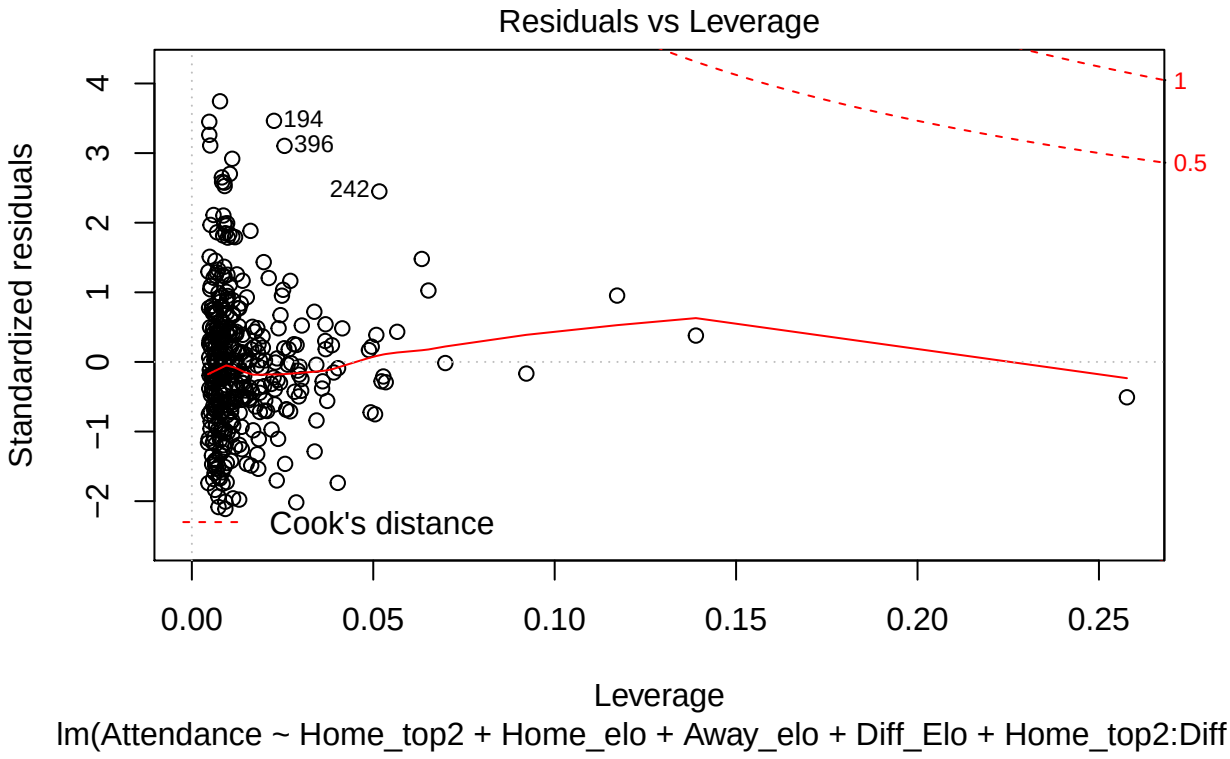
```
##     Elo_Combined
##
##
## Step:  AIC=7671.35
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                                9.9478e+10 7671.3
## - Home_top2:Diff_Elo   1 1498888462 1.0098e+11 7675.3
## - Away_elo             1 2017461720 1.0150e+11 7677.3
## - Home_elo             1 2471896704 1.0195e+11 7679.1
```

```
plot(fit1)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff .
154
352
194


Scale−Location
√|Standardized residuals|
Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff .
194
154
352




```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo +
##     Home_top2:Diff_Elo, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33577  -9273   -904   7257  59563
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.968e+04  2.158e+04  -2.302  0.02184 *
## Home_top2          -2.035e+01  7.923e+01  -0.257  0.79743
## Home_elo            3.341e+01  1.073e+01   3.113  0.00199 **
## Away_elo            2.429e+01  8.638e+00   2.812  0.00517 **
## Diff_Elo           -3.283e+01  1.183e+01  -2.775  0.00578 **
## Home_top2:Diff_Elo  1.285e+00  5.303e-01   2.424  0.01580 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15970 on 390 degrees of freedom
## Multiple R-squared:  0.1494, Adjusted R-squared:  0.1385
## F-statistic:  13.7 on 5 and 390 DF,  p-value: 2.465e-12
```

We see here that combined Elo is not worth including in the model, likely due to the fact that we have already accounted for Home Elo and Away Elo.

```
fit <- lm(Attendance ~ TimeBucket, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ TimeBucket, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -31636 -12599   -842   9668  57513
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    32496       1703  19.082   <2e-16 ***
## TimeBucket       477        365   1.307    0.192
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17190 on 394 degrees of freedom
## Multiple R-squared:  0.004314,   Adjusted R-squared:  0.001787
## F-statistic: 1.707 on 1 and 394 DF,  p-value: 0.1921
```

```
fit <- lm(Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo +
    Home_top2:Diff_Elo + TimeBucket, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo +
##     Home_top2:Diff_Elo + TimeBucket, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33529  -9529   -884   7230  59644
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.970e+04  2.160e+04  -2.301  0.02192 *
## Home_top2          -2.218e+01  7.960e+01  -0.279  0.78066
## Home_elo            3.344e+01  1.074e+01   3.112  0.00199 **
## Away_elo            2.403e+01  8.702e+00   2.761  0.00603 **
## Diff_Elo           -3.276e+01  1.185e+01  -2.766  0.00595 **
## TimeBucket          9.461e+01  3.461e+02   0.273  0.78472
## Home_top2:Diff_Elo  1.287e+00  5.309e-01   2.425  0.01577 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15990 on 389 degrees of freedom
## Multiple R-squared:  0.1496, Adjusted R-squared:  0.1365
## F-statistic:  11.4 on 6 and 389 DF,  p-value: 9.403e-12
```

```
fit1 <- step(fit)
```

```
## Start:  AIC=7673.27
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo +
##     TimeBucket
##
##                       Df  Sum of Sq        RSS    AIC
```

```
## - TimeBucket          1   19106424 9.9478e+10 7671.3
## <none>                              9.9459e+10 7673.3
## - Home_top2:Diff_Elo  1 1503239574 1.0096e+11 7677.2
## - Away_elo            1 1949684950 1.0141e+11 7679.0
## - Home_elo            1 2476613738 1.0194e+11 7681.0
##
## Step:  AIC=7671.35
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                              9.9478e+10 7671.3
## - Home_top2:Diff_Elo  1 1498888462 1.0098e+11 7675.3
## - Away_elo            1 2017461720 1.0150e+11 7677.3
## - Home_elo            1 2471896704 1.0195e+11 7679.1
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo +
##     Home_top2:Diff_Elo, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33577  -9273   -904   7257  59563
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.968e+04  2.158e+04  -2.302  0.02184 *
## Home_top2          -2.035e+01  7.923e+01  -0.257  0.79743
## Home_elo            3.341e+01  1.073e+01   3.113  0.00199 **
## Away_elo            2.429e+01  8.638e+00   2.812  0.00517 **
## Diff_Elo           -3.283e+01  1.183e+01  -2.775  0.00578 **
## Home_top2:Diff_Elo  1.285e+00  5.303e-01   2.424  0.01580 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15970 on 390 degrees of freedom
## Multiple R-squared:  0.1494, Adjusted R-squared:  0.1385
## F-statistic:  13.7 on 5 and 390 DF,  p-value: 2.465e-12
```

```
vif(fit1)
```

```
##          Home_top2           Home_elo           Away_elo
##           3.544770           1.819406           1.216143
##           Diff_Elo Home_top2:Diff_Elo
##           1.535136           3.790851
```

TimeBucket is not necessary in the model.

```
fit <- lm(Attendance ~ Differential, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Differential, data = subset_afl)
##
```

```
## Residuals:
##    Min     1Q Median     3Q    Max
## -33079 -13637   -839   9512  56884
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   36874.0     1215.5  30.336  < 2e-16 ***
## Differential  -1042.5      365.3  -2.854  0.00455 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17050 on 394 degrees of freedom
## Multiple R-squared:  0.02025,    Adjusted R-squared:  0.01777
## F-statistic: 8.145 on 1 and 394 DF,  p-value: 0.004545
```

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo + Diff
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo +
##     Home_top2:Diff_Elo + Differential, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34216 -10175  -1177   7397  58439
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.887e+04  2.147e+04  -2.277 0.023351 *
## Home_top2          -5.457e+00  7.909e+01  -0.069 0.945024
## Home_elo            4.397e+01  1.165e+01   3.774 0.000185 ***
## Away_elo            1.300e+01  9.937e+00   1.308 0.191717
## Diff_Elo            1.399e-01  1.872e+01   0.007 0.994044
## Differential       -1.507e+03  6.657e+02  -2.264 0.024146 *
## Home_top2:Diff_Elo  1.075e+00  5.356e-01   2.008 0.045383 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15890 on 389 degrees of freedom
## Multiple R-squared:  0.1605, Adjusted R-squared:  0.1475
## F-statistic: 12.39 on 6 and 389 DF,  p-value: 8.773e-13
```

```
fit1 <- step(fit)
```

```
## Start:  AIC=7668.17
## Attendance ~ Home_top2 + Home_elo + Away_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Differential
##
##                       Df  Sum of Sq        RSS    AIC
## - Away_elo             1  431689194 9.8617e+10 7667.9
## <none>                              9.8185e+10 7668.2
## - Home_top2:Diff_Elo   1 1017245959 9.9202e+10 7670.2
## - Differential         1 1293335400 9.9478e+10 7671.3
## - Home_elo             1 3595541774 1.0178e+11 7680.4
```

```
##
## Step:  AIC=7667.9
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Differential +
##     Home_top2:Diff_Elo
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                               9.8617e+10 7667.9
## - Home_top2:Diff_Elo  1  691974667 9.9309e+10 7668.7
## - Differential        1 2879107926 1.0150e+11 7677.3
## - Home_elo            1 4022862078 1.0264e+11 7681.7
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Differential +
##     Home_top2:Diff_Elo, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34969 -10231   -900   7965  57979
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -3.257e+04  1.749e+04  -1.862 0.063366 .
## Home_top2           2.682e+01  7.521e+01   0.357 0.721616
## Home_elo            4.607e+01  1.155e+01   3.989 7.94e-05 ***
## Diff_Elo            8.627e+00  1.758e+01   0.491 0.623857
## Differential       -1.944e+03  5.761e+02  -3.374 0.000814 ***
## Home_top2:Diff_Elo  8.313e-01  5.025e-01   1.654 0.098880 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15900 on 390 degrees of freedom
## Multiple R-squared:  0.1568, Adjusted R-squared:  0.146
## F-statistic:  14.5 on 5 and 390 DF,  p-value: 4.847e-13
```

```
vif(fit1)
```

```
##          Home_top2           Home_elo           Diff_Elo
##           3.221943           2.126381           3.419713
##       Differential Home_top2:Diff_Elo
##           2.861168           3.434350
```

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Differential +
    Home_top2:Diff_Elo, data = subset_afl)
```

Difference in odds is significant enough for inclusion in the model but replaces Away Elo.

```
fit <- lm(Attendance ~ Underdog, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Underdog, data = subset_afl)
##
## Residuals:
```

```
##    Min     1Q Median     3Q    Max
## -32706 -13545   -866   9700  56701
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  38540.2     1729.1  22.289  < 2e-16 ***
## Underdog     -1094.6      398.3  -2.748  0.00627 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17070 on 394 degrees of freedom
## Multiple R-squared:  0.01881,    Adjusted R-squared:  0.01632
## F-statistic: 7.553 on 1 and 394 DF,  p-value: 0.006266
```

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Differential +
    Home_top2:Diff_Elo + Underdog, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Differential +
##     Home_top2:Diff_Elo + Underdog, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34162 -10492   -896   7540  57464
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.828e+04  1.990e+04  -2.426 0.015723 *
## Home_top2           3.889e+01  7.540e+01   0.516 0.606342
## Home_elo            4.483e+01  1.155e+01   3.881 0.000122 ***
## Diff_Elo            1.454e+01  1.791e+01   0.812 0.417183
## Differential       -1.186e+04  6.064e+03  -1.956 0.051141 .
## Underdog            1.066e+04  6.485e+03   1.643 0.101164
## Home_top2:Diff_Elo  8.159e-01  5.015e-01   1.627 0.104603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15870 on 389 degrees of freedom
## Multiple R-squared:  0.1626, Adjusted R-squared:  0.1497
## F-statistic: 12.59 on 6 and 389 DF,  p-value: 5.502e-13
```

```
fit1 <- step(fit)
```

```
## Start:  AIC=7667.16
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Differential +
##     Home_top2:Diff_Elo + Underdog
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                               9.7937e+10 7667.2
## - Home_top2:Diff_Elo  1  666241221 9.8603e+10 7667.8
## - Underdog            1  679734014 9.8617e+10 7667.9
## - Differential        1  963566267 9.8900e+10 7669.0
## - Home_elo            1 3792309621 1.0173e+11 7680.2
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Differential +
##     Home_top2:Diff_Elo + Underdog, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34162 -10492   -896   7540  57464
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.828e+04  1.990e+04  -2.426 0.015723 *
## Home_top2           3.889e+01  7.540e+01   0.516 0.606342
## Home_elo            4.483e+01  1.155e+01   3.881 0.000122 ***
## Diff_Elo            1.454e+01  1.791e+01   0.812 0.417183
## Differential       -1.186e+04  6.064e+03  -1.956 0.051141 .
## Underdog            1.066e+04  6.485e+03   1.643 0.101164
## Home_top2:Diff_Elo  8.159e-01  5.015e-01   1.627 0.104603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15870 on 389 degrees of freedom
## Multiple R-squared:  0.1626, Adjusted R-squared:  0.1497
## F-statistic: 12.59 on 6 and 389 DF,  p-value: 5.502e-13
```

```
vif(fit1)
```

```
##          Home_top2           Home_elo           Diff_Elo
##           3.252820           2.135547           3.563789
##       Differential           Underdog Home_top2:Diff_Elo
##         318.318884         306.683522           3.435559
```

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo +
    Home_top2:Diff_Elo + Underdog, data = subset_afl)
vif(fit)
```

```
##          Home_top2           Home_elo           Diff_Elo
##           3.227318           2.133801           3.305356
##           Underdog Home_top2:Diff_Elo
##           2.756585           3.434384
```

Underdog odds are added but Differential is removed due to multicollinearity.

```
fit <- lm(Attendance ~ Favorite, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Favorite, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -36115 -13785   -749   9255  58387
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    17426       5204   3.349 0.000891 ***
## Favorite       12049       3641   3.309 0.001022 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16990 on 394 degrees of freedom
## Multiple R-squared:  0.02704,    Adjusted R-squared:  0.02457
## F-statistic: 10.95 on 1 and 394 DF,  p-value: 0.001022
```

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo +
    Home_top2:Diff_Elo + Underdog + Favorite, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34162 -10492   -896   7540  57464
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.828e+04  1.990e+04  -2.426 0.015723 *
## Home_top2           3.889e+01  7.540e+01   0.516 0.606342
## Home_elo            4.483e+01  1.155e+01   3.881 0.000122 ***
## Diff_Elo            1.454e+01  1.791e+01   0.812 0.417183
## Underdog           -1.207e+03  7.291e+02  -1.656 0.098532 .
## Favorite            1.186e+04  6.064e+03   1.956 0.051141 .
## Home_top2:Diff_Elo  8.159e-01  5.015e-01   1.627 0.104603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15870 on 389 degrees of freedom
## Multiple R-squared:  0.1626, Adjusted R-squared:  0.1497
## F-statistic: 12.59 on 6 and 389 DF,  p-value: 5.502e-13
```

```
fit1 <- step(fit)
```

```
## Start:  AIC=7667.16
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                                9.7937e+10 7667.2
## - Home_top2:Diff_Elo  1  666241221 9.8603e+10 7667.8
## - Underdog            1  690409373 9.8627e+10 7667.9
## - Favorite            1  963566267 9.8900e+10 7669.0
## - Home_elo            1 3792309621 1.0173e+11 7680.2
```

```
summary(fit1)
```

```
##
## Call:
```

```
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34162 -10492   -896   7540  57464
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.828e+04  1.990e+04  -2.426 0.015723 *
## Home_top2           3.889e+01  7.540e+01   0.516 0.606342
## Home_elo            4.483e+01  1.155e+01   3.881 0.000122 ***
## Diff_Elo            1.454e+01  1.791e+01   0.812 0.417183
## Underdog           -1.207e+03  7.291e+02  -1.656 0.098532 .
## Favorite            1.186e+04  6.064e+03   1.956 0.051141 .
## Home_top2:Diff_Elo  8.159e-01  5.015e-01   1.627 0.104603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15870 on 389 degrees of freedom
## Multiple R-squared:  0.1626, Adjusted R-squared:  0.1497
## F-statistic: 12.59 on 6 and 389 DF,  p-value: 5.502e-13
```

```
vif(fit1)
```

```
##          Home_top2           Home_elo           Diff_Elo
##           3.252820           2.135547           3.563789
##           Underdog           Favorite Home_top2:Diff_Elo
##           3.876425           3.181450           3.435559
```

Favorite odds is significant enough for inclusion in the model.

```
fit <- lm(Attendance ~ Line, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Line, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33889 -12622   -834   8877  57422
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  56924.7    10927.3   5.209 3.06e-07 ***
## Line          -128.3       62.1  -2.066   0.0394 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17140 on 394 degrees of freedom
## Multiple R-squared:  0.01072,    Adjusted R-squared:  0.008211
## F-statistic:  4.27 on 1 and 394 DF,  p-value: 0.03944
```

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo +
    Home_top2:Diff_Elo + Underdog + Favorite + Line, data = subset_afl)
```

```r
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite + Line, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33056 -10344   -727   7158  56564
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -3.257e+04  2.216e+04  -1.470 0.142361
## Home_top2           3.257e+01  7.536e+01   0.432 0.665819
## Home_elo            4.530e+01  1.153e+01   3.928 0.000101 ***
## Diff_Elo            2.060e+01  1.827e+01   1.128 0.260076
## Underdog           -1.362e+03  7.340e+02  -1.855 0.064335 .
## Favorite            1.199e+04  6.052e+03   1.981 0.048266 *
## Line               -9.505e+01  5.946e+01  -1.599 0.110716
## Home_top2:Diff_Elo  8.269e-01  5.006e-01   1.652 0.099371 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15840 on 388 degrees of freedom
## Multiple R-squared:  0.1681, Adjusted R-squared:  0.1531
## F-statistic:  11.2 on 7 and 388 DF,  p-value: 6.259e-13
```

```r
fit1 <- step(fit)
```

```
## Start:  AIC=7666.57
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite + Line
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                               9.7296e+10 7666.6
## - Line                 1  640856979 9.7937e+10 7667.2
## - Home_top2:Diff_Elo   1  684246545 9.7980e+10 7667.3
## - Underdog             1  863007845 9.8159e+10 7668.1
## - Favorite             1  984352876 9.8280e+10 7668.6
## - Home_elo             1 3869798958 1.0117e+11 7680.0
```

```r
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite + Line, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -33056 -10344   -727   7158  56564
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)        -3.257e+04  2.216e+04  -1.470 0.142361
## Home_top2           3.257e+01  7.536e+01   0.432 0.665819
## Home_elo            4.530e+01  1.153e+01   3.928 0.000101 ***
## Diff_Elo            2.060e+01  1.827e+01   1.128 0.260076
## Underdog           -1.362e+03  7.340e+02  -1.855 0.064335 .
## Favorite            1.199e+04  6.052e+03   1.981 0.048266 *
## Line               -9.505e+01  5.946e+01  -1.599 0.110716
## Home_top2:Diff_Elo  8.269e-01  5.006e-01   1.652 0.099371 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15840 on 388 degrees of freedom
## Multiple R-squared:  0.1681, Adjusted R-squared:  0.1531
## F-statistic:  11.2 on 7 and 388 DF,  p-value: 6.259e-13
```

```
vif(fit1)
```

```
##          Home_top2           Home_elo           Diff_Elo
##           3.261786           2.136939           3.724194
##           Underdog           Favorite               Line
##           3.944665           3.182009           1.073421
## Home_top2:Diff_Elo
##           3.436212
```

Line should also be included in the model. ##NOTE: Line means Over/Under##

```
fit <- lm(Attendance ~ Home.Odds, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home.Odds, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -31069 -12085   -587   9193  56243
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  38703.8     1632.1  23.715  < 2e-16 ***
## Home.Odds    -2027.3      656.8  -3.087  0.00217 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17020 on 394 degrees of freedom
## Multiple R-squared:  0.02361,    Adjusted R-squared:  0.02113
## F-statistic: 9.527 on 1 and 394 DF,  p-value: 0.002168
```

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
    Underdog + Favorite + Line + Home.Odds, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite + Line + Home.Odds, data = subset_afl)
##
```

```
## Residuals:
##    Min     1Q Median     3Q    Max
## -33217  -9841   -659   7303  57972
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -58555.131  25526.489  -2.294   0.0223 *
## Home_top2             -11.278     78.115  -0.144   0.8853
## Home_elo               59.566     13.473   4.421 1.28e-05 ***
## Diff_Elo               10.441     18.875   0.553   0.5805
## Underdog            -1328.534    731.248  -1.817   0.0700 .
## Favorite            11833.041   6028.572   1.963   0.0504 .
## Line                  -81.990     59.569  -1.376   0.1695
## Home.Odds            1752.307    865.158   2.025   0.0435 *
## Home_top2:Diff_Elo      1.096      0.516   2.124   0.0343 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15770 on 387 degrees of freedom
## Multiple R-squared:  0.1768, Adjusted R-squared:  0.1598
## F-statistic: 10.39 on 8 and 387 DF,  p-value: 3.315e-13
```

```
fit1 <- step(fit)
```

```
## Start:  AIC=7664.39
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite + Line + Home.Odds
##
##                       Df  Sum of Sq        RSS    AIC
## - Line                 1  471293808 9.6747e+10 7664.3
## <none>                              9.6275e+10 7664.4
## - Underdog             1  821144793 9.7097e+10 7665.8
## - Favorite             1  958447240 9.7234e+10 7666.3
## - Home.Odds            1 1020548874 9.7296e+10 7666.6
## - Home_top2:Diff_Elo   1 1122391617 9.7398e+10 7667.0
## - Home_elo             1 4862301979 1.0114e+11 7681.9
##
## Step:  AIC=7664.32
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog + Favorite +
##     Home.Odds + Home_top2:Diff_Elo
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                              9.6747e+10 7664.3
## - Underdog             1  675827458 9.7423e+10 7665.1
## - Favorite             1  939144382 9.7686e+10 7666.1
## - Home_top2:Diff_Elo   1 1144013592 9.7891e+10 7667.0
## - Home.Odds            1 1190112044 9.7937e+10 7667.2
## - Home_elo             1 4974802236 1.0172e+11 7682.2
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog +
##     Favorite + Home.Odds + Home_top2:Diff_Elo, data = subset_afl)
```

```
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34172  -9519   -792   7560  58844
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -7.385e+04  2.301e+04  -3.210  0.00144 **
## Home_top2          -9.119e+00  7.819e+01  -0.117  0.90722
## Home_elo            6.021e+01  1.348e+01   4.467 1.04e-05 ***
## Diff_Elo            4.526e+00  1.840e+01   0.246  0.80585
## Underdog           -1.195e+03  7.256e+02  -1.646  0.10051
## Favorite            1.171e+04  6.035e+03   1.941  0.05302 .
## Home.Odds           1.881e+03  8.611e+02   2.185  0.02951 *
## Home_top2:Diff_Elo  1.106e+00  5.165e-01   2.142  0.03282 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15790 on 388 degrees of freedom
## Multiple R-squared:  0.1728, Adjusted R-squared:  0.1578
## F-statistic: 11.58 on 7 and 388 DF,  p-value: 2.231e-13
```

```
vif(fit1)
```

```
##          Home_top2           Home_elo           Diff_Elo
##           3.531737           2.937414           3.799746
##           Underdog           Favorite          Home.Odds
##           3.876677           3.181864           1.997717
## Home_top2:Diff_Elo
##           3.679447
```

Including Home Odds bumps Line from the model.

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
    Underdog + Favorite + Home.Odds + Away.Odds, data = subset_afl)
fit1 <- step(fit)
```

```
## Start:  AIC=7664.32
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Home_top2:Diff_Elo +
##     Underdog + Favorite + Home.Odds + Away.Odds
##
##
## Step:  AIC=7664.32
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog + Favorite +
##     Home.Odds + Home_top2:Diff_Elo
##
##                       Df  Sum of Sq        RSS    AIC
## <none>                               9.6747e+10 7664.3
## - Underdog             1  675827458 9.7423e+10 7665.1
## - Favorite             1  939144382 9.7686e+10 7666.1
## - Home_top2:Diff_Elo   1 1144013592 9.7891e+10 7667.0
## - Home.Odds            1 1190112044 9.7937e+10 7667.2
## - Home_elo             1 4974802236 1.0172e+11 7682.2
```

```
summary(fit1)
```

```
##
```

```
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog +
##     Favorite + Home.Odds + Home_top2:Diff_Elo, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -34172  -9519   -792   7560  58844
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -7.385e+04  2.301e+04  -3.210  0.00144 **
## Home_top2          -9.119e+00  7.819e+01  -0.117  0.90722
## Home_elo            6.021e+01  1.348e+01   4.467 1.04e-05 ***
## Diff_Elo            4.526e+00  1.840e+01   0.246  0.80585
## Underdog           -1.195e+03  7.256e+02  -1.646  0.10051
## Favorite            1.171e+04  6.035e+03   1.941  0.05302 .
## Home.Odds           1.881e+03  8.611e+02   2.185  0.02951 *
## Home_top2:Diff_Elo  1.106e+00  5.165e-01   2.142  0.03282 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15790 on 388 degrees of freedom
## Multiple R-squared:  0.1728, Adjusted R-squared:  0.1578
## F-statistic: 11.58 on 7 and 388 DF,  p-value: 2.231e-13
```

```
vif(fit1)
```

```
##          Home_top2           Home_elo           Diff_Elo
##           3.531737           2.937414           3.799746
##           Underdog           Favorite          Home.Odds
##           3.876677           3.181864           1.997717
## Home_top2:Diff_Elo
##           3.679447
```

Away odds is not good enough for inclusion in the model.

```
fit <- lm(Attendance ~ Venue, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Venue, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -29275  -4613    -99   4181  39093
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)       42108       1561  26.979  < 2e-16 ***
## VenueBellerive   -31208       4506  -6.927 1.86e-11 ***
## VenueCarrara     -28265       2956  -9.560  < 2e-16 ***
## VenueCazaly's    -35564       7485  -4.751 2.87e-06 ***
## VenueDocklands   -11266       1901  -5.926 6.98e-09 ***
## VenueEureka      -34318       6178  -5.555 5.24e-08 ***
## VenueGabba       -25172       2664  -9.449  < 2e-16 ***
```

```
## VenueJiangwan    -31704       7485  -4.236 2.87e-05 ***
## VenueKardinia    -13018       3022  -4.307 2.11e-05 ***
## VenueM.C.G.       10240       1901   5.387 1.27e-07 ***
## VenueManuka      -30013       4506  -6.661 9.59e-11 ***
## VenueMarrara     -31711       7485  -4.237 2.85e-05 ***
## VenuePerth         4786       2664   1.796  0.07322 .
## VenueS.C.G.       -8429       2703  -3.118  0.00196 **
## VenueSubiaco      -7545       2703  -2.791  0.00552 **
## VenueSydney      -29380       3022  -9.721  < 2e-16 ***
## VenueTraeger     -36077       7485  -4.820 2.08e-06 ***
## VenueYork        -29220       3979  -7.343 1.29e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10350 on 378 degrees of freedom
## Multiple R-squared:  0.6536, Adjusted R-squared:  0.638
## F-statistic: 41.95 on 17 and 378 DF,  p-value: < 2.2e-16
```

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog +
    Favorite + Home.Odds + Home_top2:Diff_Elo + Venue, data = subset_afl)
vif(fit)
```

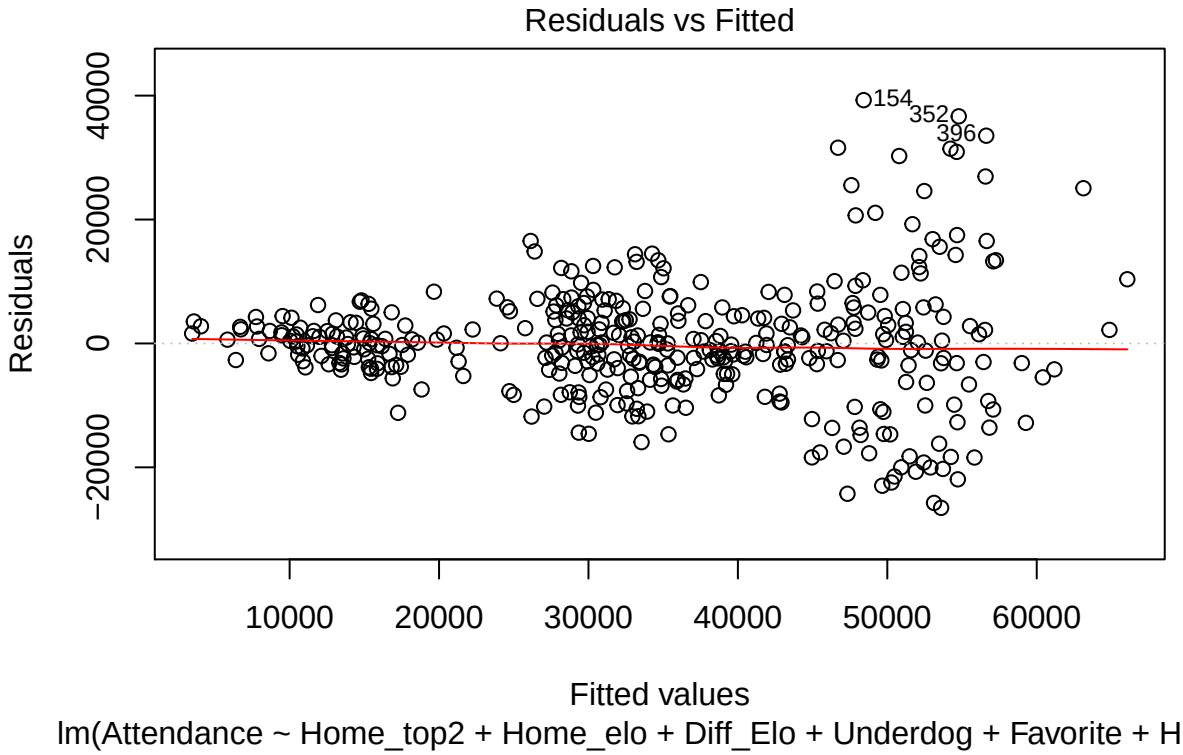
```
##                       GVIF Df GVIF^(1/(2*Df))
## Home_top2          3.930989  1        1.982672
## Home_elo           4.893140  1        2.212044
## Diff_Elo           4.039176  1        2.009770
## Underdog           4.123634  1        2.030673
## Favorite           3.314534  1        1.820586
## Home.Odds          2.246036  1        1.498678
## Venue              3.359557 17        1.036284
## Home_top2:Diff_Elo 3.927300  1        1.981742
```

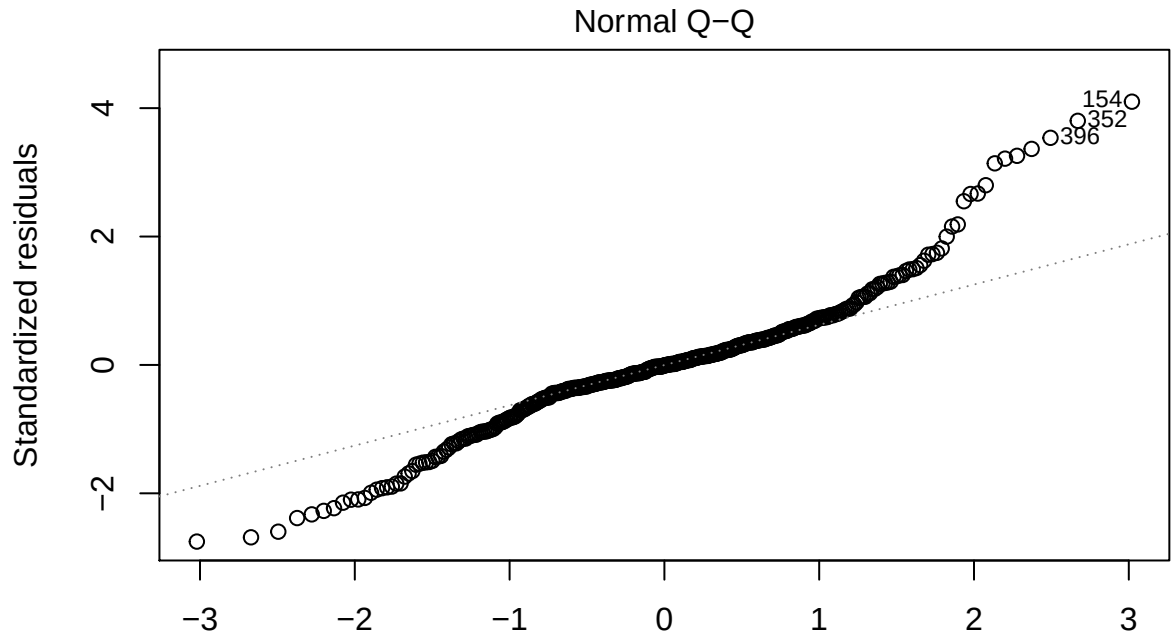
```
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog +
##     Favorite + Home.Odds + Home_top2:Diff_Elo + Venue, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -26535  -3853    -69   4035  39268
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -1.813e+04  1.756e+04  -1.033 0.302492
## Home_top2          9.950e+01  5.078e+01   1.959 0.050810 .
## Home_elo           3.005e+01  1.071e+01   2.806 0.005287 **
## Diff_Elo           9.896e+00  1.168e+01   0.847 0.397336
## Underdog          -6.249e+02  4.606e+02  -1.357 0.175749
## Favorite           7.989e+03  3.791e+03   2.107 0.035776 *
## Home.Odds          6.727e+02  5.620e+02   1.197 0.232080
## VenueBellerive    -2.732e+04  4.346e+03  -6.287 9.09e-10 ***
## VenueCarrara      -2.123e+04  3.254e+03  -6.523 2.26e-10 ***
## VenueCazaly's     -3.079e+04  7.202e+03  -4.275 2.44e-05 ***
```

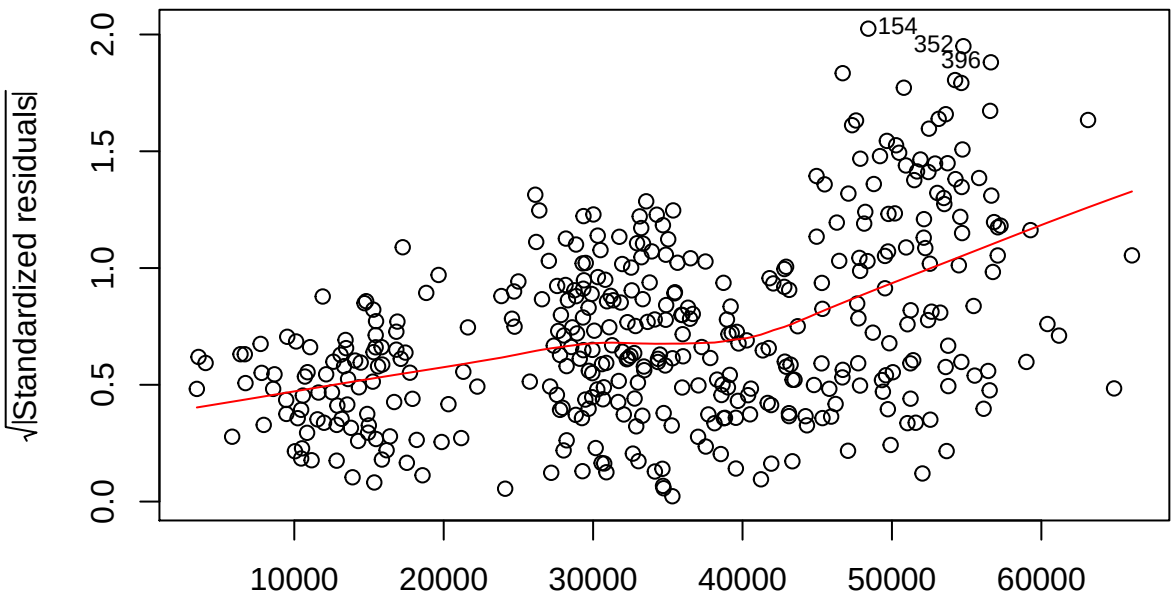
```
## VenueDocklands     -7.588e+03  1.984e+03  -3.825 0.000153 ***
## VenueEureka        -2.961e+04  5.897e+03  -5.021 8.01e-07 ***
## VenueGabba         -1.778e+04  3.235e+03  -5.497 7.20e-08 ***
## VenueJiangwan      -2.485e+04  7.258e+03  -3.424 0.000687 ***
## VenueKardinia      -1.296e+04  2.911e+03  -4.453 1.12e-05 ***
## VenueM.C.G.         1.174e+04  1.881e+03   6.241 1.19e-09 ***
## VenueManuka        -3.085e+04  4.266e+03  -7.233 2.73e-12 ***
## VenueMarrara       -2.743e+04  7.082e+03  -3.873 0.000127 ***
## VenuePerth          6.828e+03  2.616e+03   2.611 0.009403 **
## VenueS.C.G.        -7.613e+03  2.606e+03  -2.922 0.003695 **
## VenueSubiaco       -4.449e+03  2.674e+03  -1.664 0.097044 .
## VenueSydney        -2.869e+04  2.847e+03 -10.078  < 2e-16 ***
## VenueTraeger       -3.490e+04  7.064e+03  -4.941 1.18e-06 ***
## VenueYork          -2.535e+04  3.813e+03  -6.648 1.06e-10 ***
## Home_top2:Diff_Elo -1.922e-02  3.285e-01  -0.059 0.953373
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9720 on 371 degrees of freedom
## Multiple R-squared:  0.7003, Adjusted R-squared:  0.6809
## F-statistic: 36.12 on 24 and 371 DF,  p-value: < 2.2e-16
```

```
plot(fit)
```

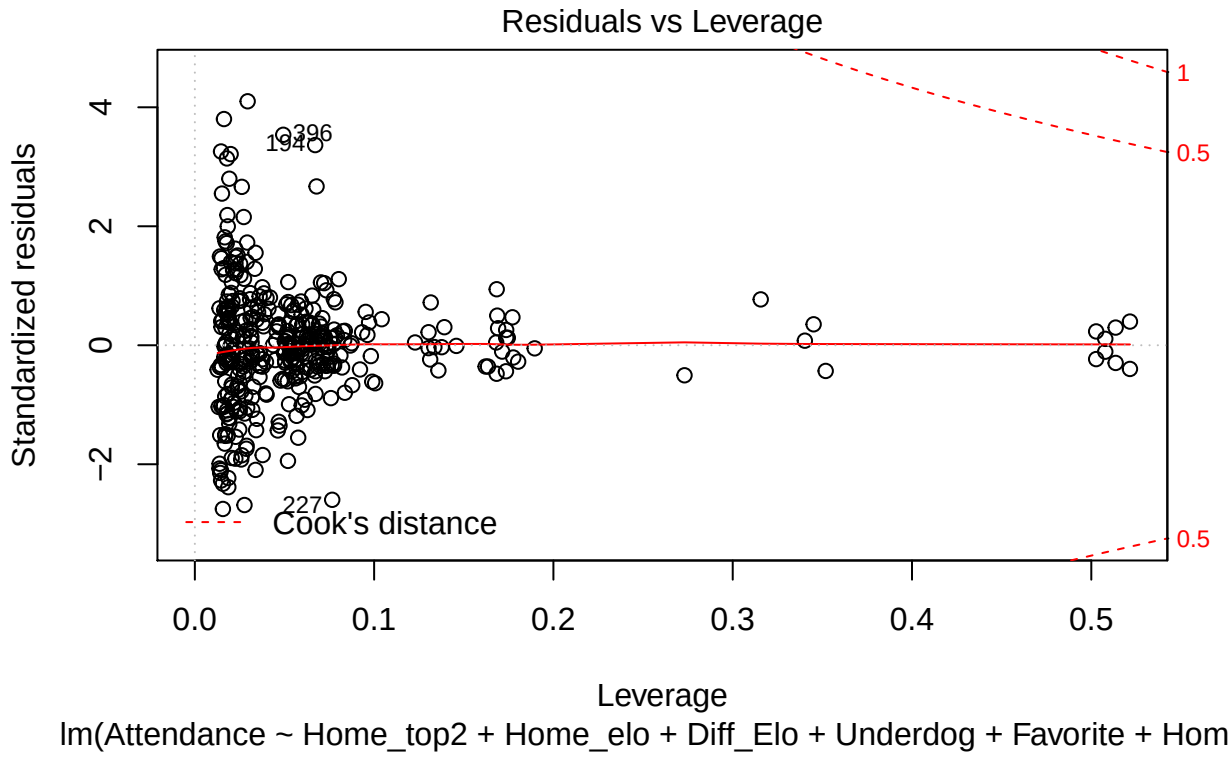




Theoretical Quantiles
lm(Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog + Favorite + Hom .



Fitted values
lm(Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog + Favorite + Hom .




```
fit1 <- step(fit)
```

```
## Start:  AIC=7296.29
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog + Favorite +
##     Home.Odds + Home_top2:Diff_Elo + Venue
##
##                       Df  Sum of Sq        RSS    AIC
## - Home_top2:Diff_Elo   1 3.2347e+05 3.5053e+10 7294.3
## - Home.Odds            1 1.3537e+08 3.5188e+10 7295.8
## - Underdog             1 1.7387e+08 3.5227e+10 7296.2
## <none>                               3.5053e+10 7296.3
## - Favorite             1 4.1950e+08 3.5472e+10 7299.0
## - Home_elo             1 7.4369e+08 3.5796e+10 7302.6
## - Venue               17 6.1694e+10 9.6747e+10 7664.3
##
## Step:  AIC=7294.29
## Attendance ~ Home_top2 + Home_elo + Diff_Elo + Underdog + Favorite +
##     Home.Odds + Venue
##
##             Df  Sum of Sq        RSS    AIC
## - Diff_Elo   1 7.6242e+07 3.5129e+10 7293.2
## - Home.Odds  1 1.4914e+08 3.5202e+10 7294.0
## - Underdog   1 1.7494e+08 3.5228e+10 7294.3
## <none>                   3.5053e+10 7294.3
## - Favorite   1 4.1919e+08 3.5472e+10 7297.0
## - Home_top2  1 6.6254e+08 3.5716e+10 7299.7
## - Home_elo   1 7.4371e+08 3.5797e+10 7300.6
## - Venue     17 6.2838e+10 9.7891e+10 7667.0
##
## Step:  AIC=7293.15
```

```
## Attendance ~ Home_top2 + Home_elo + Underdog + Favorite + Home.Odds +
##     Venue
##
##             Df  Sum of Sq        RSS    AIC
## - Underdog   1 1.0517e+08 3.5234e+10 7292.3
## <none>                    3.5129e+10 7293.2
## - Home.Odds  1 2.0559e+08 3.5335e+10 7293.5
## - Favorite   1 3.5178e+08 3.5481e+10 7295.1
## - Home_top2  1 6.9203e+08 3.5821e+10 7298.9
## - Home_elo   1 7.1119e+08 3.5840e+10 7299.1
## - Venue     17 6.3138e+10 9.8267e+10 7666.5
##
## Step:  AIC=7292.33
## Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds + Venue
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    3.5234e+10 7292.3
## - Home.Odds  1 1.7959e+08 3.5414e+10 7292.3
## - Home_elo   1 6.3986e+08 3.5874e+10 7297.5
## - Home_top2  1 7.8585e+08 3.6020e+10 7299.1
## - Favorite   1 2.0341e+09 3.7269e+10 7312.6
## - Venue     17 6.3377e+10 9.8611e+10 7665.9
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
##     Venue, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -26525  -3851    -90   3920  39438
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -18582.72   17011.84  -1.092 0.275387
## Home_top2         104.88      36.32   2.888 0.004100 **
## Home_elo           27.36      10.50   2.606 0.009523 **
## Favorite        10071.96    2167.56   4.647 4.68e-06 ***
## Home.Odds         727.88     527.19   1.381 0.168200
## VenueBellerive -26972.41    4328.49  -6.231 1.25e-09 ***
## VenueCarrara   -21000.24    3238.40  -6.485 2.82e-10 ***
## VenueCazaly's  -30439.39    7172.03  -4.244 2.77e-05 ***
## VenueDocklands  -7348.03    1972.42  -3.725 0.000225 ***
## VenueEureka    -29613.43    5883.78  -5.033 7.52e-07 ***
## VenueGabba     -17260.37    3177.66  -5.432 1.01e-07 ***
## VenueJiangwan  -25160.25    7231.59  -3.479 0.000562 ***
## VenueKardinia  -13406.47    2847.99  -4.707 3.54e-06 ***
## VenueM.C.G.     12024.13    1865.40   6.446 3.55e-10 ***
## VenueManuka    -30432.45    4231.64  -7.192 3.51e-12 ***
## VenueMarrara   -27617.33    7057.39  -3.913 0.000108 ***
## VenuePerth       6894.96    2609.90   2.642 0.008592 **
## VenueS.C.G.     -7546.30    2589.22  -2.915 0.003777 **
## VenueSubiaco    -4283.26    2667.15  -1.606 0.109133
```

```
## VenueSydney    -28618.43    2836.54 -10.089  < 2e-16 ***
## VenueTraeger   -34502.32    7047.93  -4.895 1.46e-06 ***
## VenueYork      -25397.06    3801.88  -6.680 8.67e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9706 on 374 degrees of freedom
## Multiple R-squared:  0.6987, Adjusted R-squared:  0.6818
## F-statistic:  41.3 on 21 and 374 DF,  p-value: < 2.2e-16
```

```
vif(fit1)
```

```
##               GVIF Df GVIF^(1/(2*Df))
## Home_top2 2.016377  1        1.419992
## Home_elo  4.714634  1        2.171321
## Favorite  1.086414  1        1.042312
## Home.Odds 1.981963  1        1.407822
## Venue     2.786421 17        1.030599
```

Venue is a very good predictor. It effectively removes Diff Elo, Underdog, and Home_top2:Diff Elo from the model.

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
    Venue + Away, data = subset_afl)
vif(fit)
```

```
##               GVIF Df GVIF^(1/(2*Df))
## Home_top2 2.105621  1        1.451076
## Home_elo  5.349093  1        2.312811
## Favorite  1.789143  1        1.337589
## Home.Odds 3.398363  1        1.843465
## Venue     5.509693 17        1.051472
## Away      4.839327 17        1.047468
```

```
fit1 <- step(fit)
```

```
## Start:  AIC=7214.96
## Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds + Venue +
##     Away
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    2.6596e+10 7215.0
## - Home.Odds  1 1.4671e+08 2.6743e+10 7215.1
## - Home_top2  1 4.3544e+08 2.7032e+10 7219.4
## - Home_elo   1 1.0108e+09 2.7607e+10 7227.7
## - Favorite   1 1.0156e+09 2.7612e+10 7227.8
## - Away      17 8.6381e+09 3.5234e+10 7292.3
## - Venue     17 5.3754e+10 8.0350e+10 7618.8
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
##     Venue + Away, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
```

```
## -21559  -4998   -113   3449  33342
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -35417.753  17327.904  -2.044 0.041690 *
## Home_top2                     79.783     33.000   2.418 0.016122 *
## Home_elo                      36.628      9.944   3.684 0.000266 ***
## Favorite                    9133.055   2473.568   3.692 0.000257 ***
## Home.Odds                    861.457    613.874   1.403 0.161392
## VenueBellerive             -24184.905   3920.957  -6.168 1.87e-09 ***
## VenueCarrara               -22850.124   2928.213  -7.803 6.73e-14 ***
## VenueCazaly's              -25660.420   6514.980  -3.939 9.85e-05 ***
## VenueDocklands              -6015.862   1789.398  -3.362 0.000858 ***
## VenueEureka                -22895.956   5430.422  -4.216 3.15e-05 ***
## VenueGabba                 -15563.363   2858.603  -5.444 9.69e-08 ***
## VenueJiangwan              -17772.132   6719.393  -2.645 0.008532 **
## VenueKardinia              -11374.140   2567.340  -4.430 1.25e-05 ***
## VenueM.C.G.                 10869.036   1679.841   6.470 3.23e-10 ***
## VenueManuka                -27674.592   3844.245  -7.199 3.61e-12 ***
## VenueMarrara               -24460.066   6399.269  -3.822 0.000156 ***
## VenuePerth                   7821.555   2340.594   3.342 0.000921 ***
## VenueS.C.G.                 -7657.190   2330.101  -3.286 0.001116 **
## VenueSubiaco                -3997.704   2389.106  -1.673 0.095143 .
## VenueSydney                -29648.762   2559.627 -11.583  < 2e-16 ***
## VenueTraeger               -30127.160   6407.972  -4.702 3.69e-06 ***
## VenueYork                  -21832.435   3466.490  -6.298 8.86e-10 ***
## AwayBrisbane Lions            731.562   3043.191   0.240 0.810163
## AwayCarlton                 10080.751   2956.182   3.410 0.000724 ***
## AwayCollingwood             12558.662   2692.592   4.664 4.39e-06 ***
## AwayEssendon                12025.062   2719.489   4.422 1.30e-05 ***
## AwayFremantle                 617.645   2860.976   0.216 0.829200
## AwayGeelong                  6661.858   2673.631   2.492 0.013167 *
## AwayGold Coast               -689.277   2941.620  -0.234 0.814872
## AwayGreater Western Sydney  -2763.415   2660.201  -1.039 0.299601
## AwayHawthorn                 8986.127   2710.302   3.316 0.001009 **
## AwayMelbourne                4307.433   2657.176   1.621 0.105888
## AwayNorth Melbourne          1156.662   2783.924   0.415 0.678041
## AwayPort Adelaide           -2997.106   2772.437  -1.081 0.280411
## AwayRichmond                10545.861   2662.979   3.960 9.04e-05 ***
## AwaySt Kilda                 3810.070   2799.481   1.361 0.174375
## AwaySydney                   1360.879   2696.518   0.505 0.614095
## AwayWest Coast               -416.617   2700.056  -0.154 0.877461
## AwayWestern Bulldogs         1767.807   2715.766   0.651 0.515503
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8631 on 357 degrees of freedom
## Multiple R-squared:  0.7726, Adjusted R-squared:  0.7484
## F-statistic: 31.92 on 38 and 357 DF,  p-value: < 2.2e-16
```

```
vif(fit1)
```

```
##               GVIF Df GVIF^(1/(2*Df))
## Home_top2 2.105621  1        1.451076
## Home_elo  5.349093  1        2.312811
```

```
## Favorite  1.789143  1        1.337589
## Home.Odds 3.398363  1        1.843465
## Venue     5.509693 17        1.051472
## Away      4.839327 17        1.047468
```

Away is a good predictor.

```
fit <- fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
    Venue + Away + Day, data = subset_afl)
vif(fit)
```

```
##               GVIF Df GVIF^(1/(2*Df))
## Home_top2 2.149049  1        1.465964
## Home_elo  5.465958  1        2.337939
## Favorite  1.824748  1        1.350832
## Home.Odds 3.434431  1        1.853222
## Venue     7.640446 17        1.061632
## Away      7.007250 17        1.058934
## Day       2.281825  6        1.071166
```

```
fit1 <- step(fit)
```

```
## Start:  AIC=7150.84
## Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds + Venue +
##     Away + Day
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    2.1945e+10 7150.8
## - Home.Odds  1 1.5788e+08 2.2103e+10 7151.7
## - Home_top2  1 4.5384e+08 2.2399e+10 7156.9
## - Favorite   1 7.8991e+08 2.2735e+10 7162.8
## - Home_elo   1 1.0010e+09 2.2946e+10 7166.5
## - Day        6 4.6511e+09 2.6596e+10 7215.0
## - Away      17 6.4751e+09 2.8420e+10 7219.2
## - Venue     17 4.6303e+10 6.8248e+10 7566.1
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
##     Venue + Away + Day, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -18961  -4926    -95   3282  32009
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -33306.791  16163.530  -2.061 0.040076 *
## Home_top2                     82.286     30.542   2.694 0.007394 **
## Home_elo                      36.845      9.208   4.001 7.69e-05 ***
## Favorite                    8134.233   2288.457   3.554 0.000431 ***
## Home.Odds                    898.375    565.342   1.589 0.112942
## VenueBellerive             -24362.207   3606.472  -6.755 5.94e-11 ***
## VenueCarrara               -22482.745   2711.322  -8.292 2.40e-15 ***
## VenueCazaly's              -25868.477   5975.429  -4.329 1.96e-05 ***
```

```
## VenueDocklands              -5588.355   1660.193  -3.366 0.000847 ***
## VenueEureka                -22711.529   4981.656  -4.559 7.11e-06 ***
## VenueGabba                 -15020.967   2638.933  -5.692 2.65e-08 ***
## VenueJiangwan              -17822.878   6172.862  -2.887 0.004126 **
## VenueKardinia              -11691.474   2354.526  -4.966 1.07e-06 ***
## VenueM.C.G.                  9582.318   1588.500   6.032 4.10e-09 ***
## VenueManuka                -27788.688   3536.989  -7.857 4.87e-14 ***
## VenueMarrara               -25082.244   5871.790  -4.272 2.50e-05 ***
## VenuePerth                   8881.279   2202.964   4.032 6.80e-05 ***
## VenueS.C.G.                 -7879.628   2139.906  -3.682 0.000268 ***
## VenueSubiaco                -3080.267   2218.115  -1.389 0.165808
## VenueSydney                -29643.254   2406.613 -12.317  < 2e-16 ***
## VenueTraeger               -29274.938   5889.548  -4.971 1.05e-06 ***
## VenueYork                  -22024.314   3188.998  -6.906 2.34e-11 ***
## AwayBrisbane Lions            942.036   2800.104   0.336 0.736749
## AwayCarlton                 10176.599   2714.733   3.749 0.000208 ***
## AwayCollingwood              9991.454   2503.938   3.990 8.04e-05 ***
## AwayEssendon                10780.240   2525.337   4.269 2.53e-05 ***
## AwayFremantle                 931.960   2635.166   0.354 0.723804
## AwayGeelong                  6085.366   2474.080   2.460 0.014389 *
## AwayGold Coast               -819.358   2703.954  -0.303 0.762052
## AwayGreater Western Sydney  -2371.004   2449.440  -0.968 0.333722
## AwayHawthorn                 8725.965   2508.114   3.479 0.000566 ***
## AwayMelbourne                3901.828   2452.894   1.591 0.112576
## AwayNorth Melbourne          1619.173   2566.909   0.631 0.528590
## AwayPort Adelaide           -2560.459   2559.427  -1.000 0.317805
## AwayRichmond                 9506.587   2472.930   3.844 0.000144 ***
## AwaySt Kilda                 3817.873   2565.573   1.488 0.137618
## AwaySydney                   1366.845   2481.971   0.551 0.582183
## AwayWest Coast                370.859   2499.115   0.148 0.882115
## AwayWestern Bulldogs         1473.214   2520.456   0.585 0.559257
## DayMon                      18879.304   3955.869   4.772 2.67e-06 ***
## DaySat                       -872.713   1493.713  -0.584 0.559422
## DaySun                      -2691.089   1573.299  -1.710 0.088062 .
## DayThu                       -961.362   2427.572  -0.396 0.692332
## DayTue                      28749.404   5968.365   4.817 2.17e-06 ***
## DayWed                      31463.642   8225.807   3.825 0.000155 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7907 on 351 degrees of freedom
## Multiple R-squared:  0.8124, Adjusted R-squared:  0.7888
## F-statistic: 34.54 on 44 and 351 DF,  p-value: < 2.2e-16
```

```
vif(fit1)
```

```
##                GVIF Df GVIF^(1/(2*Df))
## Home_top2 2.149049  1        1.465964
## Home_elo  5.465958  1        2.337939
## Favorite  1.824748  1        1.350832
## Home.Odds 3.434431  1        1.853222
## Venue     7.640446 17        1.061632
## Away      7.007250 17        1.058934
## Day       2.281825  6        1.071166
```

Day is a good predictor.

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
    Venue + Away + Day + Round, data = subset_afl)
fit1 <- step(fit)
```

```
## Start:  AIC=7152.82
## Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds + Venue +
##     Away + Day + Round
##
##             Df  Sum of Sq        RSS    AIC
## - Round      1 1.1877e+06 2.1945e+10 7150.8
## <none>                    2.1944e+10 7152.8
## - Home.Odds  1 1.5555e+08 2.2100e+10 7153.6
## - Home_top2  1 4.5426e+08 2.2398e+10 7158.9
## - Favorite   1 7.8948e+08 2.2733e+10 7164.8
## - Home_elo   1 9.9152e+08 2.2936e+10 7168.3
## - Day        6 4.6085e+09 2.6553e+10 7216.3
## - Away      17 6.4720e+09 2.8416e+10 7221.2
## - Venue     17 4.6220e+10 6.8164e+10 7567.7
##
## Step:  AIC=7150.84
## Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds + Venue +
##     Away + Day
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    2.1945e+10 7150.8
## - Home.Odds  1 1.5788e+08 2.2103e+10 7151.7
## - Home_top2  1 4.5384e+08 2.2399e+10 7156.9
## - Favorite   1 7.8991e+08 2.2735e+10 7162.8
## - Home_elo   1 1.0010e+09 2.2946e+10 7166.5
## - Day        6 4.6511e+09 2.6596e+10 7215.0
## - Away      17 6.4751e+09 2.8420e+10 7219.2
## - Venue     17 4.6303e+10 6.8248e+10 7566.1
```

```
vif(fit1)
```

```
##               GVIF Df GVIF^(1/(2*Df))
## Home_top2 2.149049  1        1.465964
## Home_elo  5.465958  1        2.337939
## Favorite  1.824748  1        1.350832
## Home.Odds 3.434431  1        1.853222
## Venue     7.640446 17        1.061632
## Away      7.007250 17        1.058934
## Day       2.281825  6        1.071166
```

Round is not a good predictor.

```
subset_afl$Round <- as.factor(subset_afl$Round)
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
    Venue + Away + Day + Round, data = subset_afl)
vif(fit)
```

```
##               GVIF Df GVIF^(1/(2*Df))
## Home_top2  2.291235  1        1.513683
## Home_elo   5.714823  1        2.390570
## Favorite   1.923974  1        1.387074
```

```
## Home.Odds  3.605758  1        1.898883
## Venue     17.523422 17        1.087870
## Away      13.304966 17        1.079094
## Day        3.378396  6        1.106775
## Round      6.929046 22        1.044976
```

```
fit1 <- step(fit)
```

```
## Start:  AIC=7153.28
## Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds + Venue +
##     Away + Day + Round
##
##             Df  Sum of Sq        RSS    AIC
## - Round     22 2.1865e+09 2.1945e+10 7150.8
## <none>                    1.9759e+10 7153.3
## - Home.Odds  1 1.5884e+08 1.9918e+10 7154.4
## - Home_top2  1 4.1311e+08 2.0172e+10 7159.5
## - Favorite   1 7.7913e+08 2.0538e+10 7166.6
## - Home_elo   1 1.0056e+09 2.0764e+10 7170.9
## - Day        6 4.3192e+09 2.4078e+10 7219.6
## - Away      17 6.3267e+09 2.6085e+10 7229.3
## - Venue     17 4.4859e+10 6.4618e+10 7588.5
##
## Step:  AIC=7150.84
## Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds + Venue +
##     Away + Day
##
##             Df  Sum of Sq        RSS    AIC
## <none>                    2.1945e+10 7150.8
## - Home.Odds  1 1.5788e+08 2.2103e+10 7151.7
## - Home_top2  1 4.5384e+08 2.2399e+10 7156.9
## - Favorite   1 7.8991e+08 2.2735e+10 7162.8
## - Home_elo   1 1.0010e+09 2.2946e+10 7166.5
## - Day        6 4.6511e+09 2.6596e+10 7215.0
## - Away      17 6.4751e+09 2.8420e+10 7219.2
## - Venue     17 4.6303e+10 6.8248e+10 7566.1
```

```
summary(fit1)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
##     Venue + Away + Day, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -18961  -4926    -95   3282  32009
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -33306.791  16163.530  -2.061 0.040076 *
## Home_top2                     82.286     30.542   2.694 0.007394 **
## Home_elo                      36.845      9.208   4.001 7.69e-05 ***
## Favorite                    8134.233   2288.457   3.554 0.000431 ***
## Home.Odds                    898.375    565.342   1.589 0.112942
```

```
## VenueBellerive             -24362.207   3606.472  -6.755 5.94e-11 ***
## VenueCarrara               -22482.745   2711.322  -8.292 2.40e-15 ***
## VenueCazaly's              -25868.477   5975.429  -4.329 1.96e-05 ***
## VenueDocklands              -5588.355   1660.193  -3.366 0.000847 ***
## VenueEureka                -22711.529   4981.656  -4.559 7.11e-06 ***
## VenueGabba                 -15020.967   2638.933  -5.692 2.65e-08 ***
## VenueJiangwan              -17822.878   6172.862  -2.887 0.004126 **
## VenueKardinia              -11691.474   2354.526  -4.966 1.07e-06 ***
## VenueM.C.G.                  9582.318   1588.500   6.032 4.10e-09 ***
## VenueManuka                -27788.688   3536.989  -7.857 4.87e-14 ***
## VenueMarrara               -25082.244   5871.790  -4.272 2.50e-05 ***
## VenuePerth                   8881.279   2202.964   4.032 6.80e-05 ***
## VenueS.C.G.                 -7879.628   2139.906  -3.682 0.000268 ***
## VenueSubiaco                -3080.267   2218.115  -1.389 0.165808
## VenueSydney                -29643.254   2406.613 -12.317  < 2e-16 ***
## VenueTraeger               -29274.938   5889.548  -4.971 1.05e-06 ***
## VenueYork                  -22024.314   3188.998  -6.906 2.34e-11 ***
## AwayBrisbane Lions            942.036   2800.104   0.336 0.736749
## AwayCarlton                 10176.599   2714.733   3.749 0.000208 ***
## AwayCollingwood              9991.454   2503.938   3.990 8.04e-05 ***
## AwayEssendon                10780.240   2525.337   4.269 2.53e-05 ***
## AwayFremantle                 931.960   2635.166   0.354 0.723804
## AwayGeelong                  6085.366   2474.080   2.460 0.014389 *
## AwayGold Coast               -819.358   2703.954  -0.303 0.762052
## AwayGreater Western Sydney  -2371.004   2449.440  -0.968 0.333722
## AwayHawthorn                 8725.965   2508.114   3.479 0.000566 ***
## AwayMelbourne                3901.828   2452.894   1.591 0.112576
## AwayNorth Melbourne          1619.173   2566.909   0.631 0.528590
## AwayPort Adelaide           -2560.459   2559.427  -1.000 0.317805
## AwayRichmond                 9506.587   2472.930   3.844 0.000144 ***
## AwaySt Kilda                 3817.873   2565.573   1.488 0.137618
## AwaySydney                   1366.845   2481.971   0.551 0.582183
## AwayWest Coast                370.859   2499.115   0.148 0.882115
## AwayWestern Bulldogs         1473.214   2520.456   0.585 0.559257
## DayMon                      18879.304   3955.869   4.772 2.67e-06 ***
## DaySat                       -872.713   1493.713  -0.584 0.559422
## DaySun                      -2691.089   1573.299  -1.710 0.088062 .
## DayThu                       -961.362   2427.572  -0.396 0.692332
## DayTue                      28749.404   5968.365   4.817 2.17e-06 ***
## DayWed                      31463.642   8225.807   3.825 0.000155 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7907 on 351 degrees of freedom
## Multiple R-squared:  0.8124, Adjusted R-squared:  0.7888
## F-statistic: 34.54 on 44 and 351 DF,  p-value: < 2.2e-16
```

```
vif(fit1)
```

```
##               GVIF Df GVIF^(1/(2*Df))
## Home_top2 2.149049  1        1.465964
## Home_elo  5.465958  1        2.337939
## Favorite  1.824748  1        1.350832
## Home.Odds 3.434431  1        1.853222
## Venue     7.640446 17        1.061632
```

```
## Away      7.007250 17        1.058934
## Day       2.281825  6        1.071166
```

Round is not a good predictor even as a factor.

Our final model is:

lm(Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds + Venue + Away + Day, data = subset_afl)

The summary of this model is as follows:

```
fit <- lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
    Venue + Away + Day, data = subset_afl)
summary(fit)
```

```
##
## Call:
## lm(formula = Attendance ~ Home_top2 + Home_elo + Favorite + Home.Odds +
##     Venue + Away + Day, data = subset_afl)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -18961  -4926    -95   3282  32009
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -33306.791  16163.530  -2.061 0.040076 *
## Home_top2                     82.286     30.542   2.694 0.007394 **
## Home_elo                      36.845      9.208   4.001 7.69e-05 ***
## Favorite                    8134.233   2288.457   3.554 0.000431 ***
## Home.Odds                    898.375    565.342   1.589 0.112942
## VenueBellerive             -24362.207   3606.472  -6.755 5.94e-11 ***
## VenueCarrara               -22482.745   2711.322  -8.292 2.40e-15 ***
## VenueCazaly's              -25868.477   5975.429  -4.329 1.96e-05 ***
## VenueDocklands              -5588.355   1660.193  -3.366 0.000847 ***
## VenueEureka                -22711.529   4981.656  -4.559 7.11e-06 ***
## VenueGabba                 -15020.967   2638.933  -5.692 2.65e-08 ***
## VenueJiangwan              -17822.878   6172.862  -2.887 0.004126 **
## VenueKardinia              -11691.474   2354.526  -4.966 1.07e-06 ***
## VenueM.C.G.                  9582.318   1588.500   6.032 4.10e-09 ***
## VenueManuka                -27788.688   3536.989  -7.857 4.87e-14 ***
## VenueMarrara               -25082.244   5871.790  -4.272 2.50e-05 ***
## VenuePerth                   8881.279   2202.964   4.032 6.80e-05 ***
## VenueS.C.G.                 -7879.628   2139.906  -3.682 0.000268 ***
## VenueSubiaco                -3080.267   2218.115  -1.389 0.165808
## VenueSydney                -29643.254   2406.613 -12.317  < 2e-16 ***
## VenueTraeger               -29274.938   5889.548  -4.971 1.05e-06 ***
## VenueYork                  -22024.314   3188.998  -6.906 2.34e-11 ***
## AwayBrisbane Lions            942.036   2800.104   0.336 0.736749
## AwayCarlton                 10176.599   2714.733   3.749 0.000208 ***
## AwayCollingwood              9991.454   2503.938   3.990 8.04e-05 ***
## AwayEssendon                10780.240   2525.337   4.269 2.53e-05 ***
## AwayFremantle                 931.960   2635.166   0.354 0.723804
## AwayGeelong                  6085.366   2474.080   2.460 0.014389 *
## AwayGold Coast               -819.358   2703.954  -0.303 0.762052
## AwayGreater Western Sydney  -2371.004   2449.440  -0.968 0.333722
```

```
## AwayHawthorn                 8725.965   2508.114   3.479 0.000566 ***
## AwayMelbourne                3901.828   2452.894   1.591 0.112576
## AwayNorth Melbourne          1619.173   2566.909   0.631 0.528590
## AwayPort Adelaide           -2560.459   2559.427  -1.000 0.317805
## AwayRichmond                 9506.587   2472.930   3.844 0.000144 ***
## AwaySt Kilda                 3817.873   2565.573   1.488 0.137618
## AwaySydney                   1366.845   2481.971   0.551 0.582183
## AwayWest Coast                370.859   2499.115   0.148 0.882115
## AwayWestern Bulldogs         1473.214   2520.456   0.585 0.559257
## DayMon                      18879.304   3955.869   4.772 2.67e-06 ***
## DaySat                       -872.713   1493.713  -0.584 0.559422
## DaySun                      -2691.089   1573.299  -1.710 0.088062 .
## DayThu                       -961.362   2427.572  -0.396 0.692332
## DayTue                      28749.404   5968.365   4.817 2.17e-06 ***
## DayWed                      31463.642   8225.807   3.825 0.000155 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7907 on 351 degrees of freedom
## Multiple R-squared:  0.8124, Adjusted R-squared:  0.7888
## F-statistic: 34.54 on 44 and 351 DF,  p-value: < 2.2e-16
```

Here are the stepwise models:

```
library(MASS)
fit <- step(lm(Attendance~., data = subset_afl), direction="forward")
```

```
## Start:  AIC=-18762.06
## Attendance ~ X + ...1 + Round + Day + Date + Year + Time + Home +
##     Away + Venue + Home.Odds + Away.Odds + Line + Favorite +
##     Underdog + Differential + venueavg + Attendance_Differential +
##     HomeVenueAvg + HomeDifferential + HomeDifferentialPct + TimeBucket +
##     Home_elo + Home_top8 + Home_top4 + Home_top2 + Away_elo +
##     Away_top8 + Away_top4 + Away_top2 + Diff_Elo + Elo_Combined
```

```
min.model <- lm(Attendance ~ 1, data=subset_afl)
max.model <- formula(lm(Attendance ~ X + ...1 + Round + Day + Year + Time +
    Away + Venue + Home.Odds + Away.Odds + Line + Favorite +
    Underdog + Differential + venueavg +
    HomeVenueAvg + TimeBucket +
    Home_elo + Home_top8 + Home_top4 + Home_top2 + Away_elo +
    Away_top8 + Away_top4 + Away_top2, data = subset_afl))
max.model2 <- lm(Attendance ~ Round + Day + Year + Time +
    Away + Venue + Home.Odds + Away.Odds + Line + Favorite +
    Underdog + Differential + venueavg +
    HomeVenueAvg + TimeBucket +
    Home_elo + Home_top8 + Home_top4 + Home_top2 + Away_elo +
    Away_top8 + Away_top4 + Away_top2, data = subset_afl)
fwd.model <- step(min.model, direction='forward', scope=max.model)
```

```
## Start:  AIC=7725.43
## Attendance ~ 1
##
##                Df  Sum of Sq        RSS    AIC
## + HomeVenueAvg  1 7.7715e+10 3.9237e+10 7294.9
## + venueavg      1 7.5068e+10 4.1883e+10 7320.8
```

```
## + Venue        17 7.6437e+10 4.0515e+10 7339.6
## + Day           6 2.4228e+10 9.2723e+10 7645.5
## + Time         27 2.4876e+10 9.2075e+10 7684.7
## + Home_elo      1 1.1920e+10 1.0503e+11 7684.9
## + Away         17 1.9818e+10 9.7134e+10 7685.9
## + Home_top2     1 9.1357e+09 1.0782e+11 7695.2
## + Home_top4     1 8.9079e+09 1.0804e+11 7696.1
## + Home_top8     1 8.0024e+09 1.0895e+11 7699.4
## + Favorite      1 3.1625e+09 1.1379e+11 7716.6
## + Home.Odds     1 2.7612e+09 1.1419e+11 7718.0
## + Differential  1 2.3688e+09 1.1458e+11 7719.3
## + Underdog      1 2.1998e+09 1.1475e+11 7719.9
## + Away_elo      1 2.0958e+09 1.1486e+11 7720.3
## + Line          1 1.2540e+09 1.1570e+11 7723.2
## <none>                       1.1695e+11 7725.4
## + Away_top8     1 5.0696e+08 1.1644e+11 7725.7
## + TimeBucket    1 5.0459e+08 1.1645e+11 7725.7
## + X             1 3.3128e+08 1.1662e+11 7726.3
## + ...1          1 3.3128e+08 1.1662e+11 7726.3
## + Away.Odds     1 7.0321e+07 1.1688e+11 7727.2
## + Year          1 6.4605e+07 1.1689e+11 7727.2
## + Away_top4     1 3.6961e+07 1.1691e+11 7727.3
## + Away_top2     1 6.0793e+06 1.1695e+11 7727.4
## + Round        22 4.6336e+09 1.1232e+11 7753.4
##
## Step:  AIC=7294.94
## Attendance ~ HomeVenueAvg
##
##                Df  Sum of Sq        RSS    AIC
## + Day           6 9471152276 2.9765e+10 7197.5
## + Away         17 8403193856 3.0833e+10 7233.5
## + Home_elo      1 2481354964 3.6755e+10 7271.1
## + Home_top8     1 2248023449 3.6989e+10 7273.6
## + Home_top4     1 1985522530 3.7251e+10 7276.4
## + Home_top2     1 1686609858 3.7550e+10 7279.5
## + Favorite      1 1457279300 3.7779e+10 7282.0
## + TimeBucket    1 1407451944 3.7829e+10 7282.5
## + venueavg      1 1137131203 3.8099e+10 7285.3
## + Differential  1 1132750214 3.8104e+10 7285.3
## + Underdog      1 1056942559 3.8180e+10 7286.1
## + Home.Odds     1  772678142 3.8464e+10 7289.1
## + Away_elo      1  534143508 3.8702e+10 7291.5
## <none>                       3.9237e+10 7294.9
## + Away_top8     1  167026407 3.9070e+10 7295.2
## + Away.Odds     1  116447983 3.9120e+10 7295.8
## + Year          1   12609850 3.9224e+10 7296.8
## + Away_top2     1    4603869 3.9232e+10 7296.9
## + Away_top4     1    4431984 3.9232e+10 7296.9
## + Line          1    3221182 3.9233e+10 7296.9
## + X             1       2947 3.9237e+10 7296.9
## + ...1          1       2947 3.9237e+10 7296.9
## + Venue        17 2510325276 3.6726e+10 7302.8
## + Time         27 3810630656 3.5426e+10 7308.5
## + Round        22 2411307590 3.6825e+10 7313.8
```

```
##
## Step:  AIC=7197.54
## Attendance ~ HomeVenueAvg + Day
##
##                 Df  Sum of Sq        RSS    AIC
## + Away          17 5369650157 2.4396e+10 7152.8
## + Home_elo       1 1871538300 2.7894e+10 7173.8
## + Home_top8      1 1747319444 2.8018e+10 7175.6
## + Home_top4      1 1543552726 2.8222e+10 7178.5
## + Home_top2      1 1320882149 2.8444e+10 7181.6
## + Favorite       1  701419170 2.9064e+10 7190.1
## + Home.Odds      1  602959957 2.9162e+10 7191.4
## + venueavg       1  572035629 2.9193e+10 7191.9
## + Differential   1  525450515 2.9240e+10 7192.5
## + Underdog       1  487553280 2.9278e+10 7193.0
## + TimeBucket     1  334016126 2.9431e+10 7195.1
## + Away_elo       1  159776801 2.9606e+10 7197.4
## <none>                       2.9765e+10 7197.5
## + Away_top8      1   45466808 2.9720e+10 7198.9
## + X              1   30172458 2.9735e+10 7199.1
## + ...1           1   30172458 2.9735e+10 7199.1
## + Away_top2      1   23871855 2.9742e+10 7199.2
## + Away.Odds      1   15588847 2.9750e+10 7199.3
## + Year           1    8478494 2.9757e+10 7199.4
## + Away_top4      1    2452444 2.9763e+10 7199.5
## + Line           1    2022817 2.9763e+10 7199.5
## + Round         22 2384073363 2.7381e+10 7208.5
## + Venue         17 1439494156 2.8326e+10 7211.9
## + Time          27 1937247037 2.7828e+10 7224.9
##
## Step:  AIC=7152.76
## Attendance ~ HomeVenueAvg + Day + Away
##
##                 Df  Sum of Sq        RSS    AIC
## + Home_elo       1 2733583231 2.1662e+10 7107.7
## + Home_top8      1 2448332486 2.1947e+10 7112.9
## + Home_top4      1 1651693604 2.2744e+10 7127.0
## + Home_top2      1 1265742853 2.3130e+10 7133.7
## + Home.Odds      1 1243394076 2.3152e+10 7134.0
## + venueavg       1  569576554 2.3826e+10 7145.4
## + Favorite       1  500489857 2.3895e+10 7146.6
## + Away_elo       1  306010091 2.4090e+10 7149.8
## + Away_top8      1  276682957 2.4119e+10 7150.2
## + Differential   1  233818786 2.4162e+10 7150.9
## + Underdog       1  199129849 2.4197e+10 7151.5
## <none>                       2.4396e+10 7152.8
## + Away.Odds      1  119622643 2.4276e+10 7152.8
## + Round         22 2536377691 2.1859e+10 7153.3
## + Venue         17 1969687410 2.2426e+10 7153.4
## + Line           1   63417021 2.4332e+10 7153.7
## + Away_top4      1   23798960 2.4372e+10 7154.4
## + X              1   18377200 2.4377e+10 7154.5
## + ...1           1   18377200 2.4377e+10 7154.5
## + TimeBucket     1    8224709 2.4388e+10 7154.6
```

```
## + Year          1    6295171 2.4389e+10 7154.7
## + Away_top2     1     323601 2.4395e+10 7154.8
## + Time         27 1521980137 2.2874e+10 7181.2
##
## Step:  AIC=7107.7
## Attendance ~ HomeVenueAvg + Day + Away + Home_elo
##
##                Df  Sum of Sq        RSS    AIC
## + Favorite      1  637732482 2.1024e+10 7097.9
## + Away.Odds     1  585738834 2.1076e+10 7098.8
## + Differential  1  503030911 2.1159e+10 7100.4
## + venueavg      1  501145362 2.1161e+10 7100.4
## + Underdog      1  464260612 2.1198e+10 7101.1
## + Away_elo      1  342245975 2.1320e+10 7103.4
## + Away_top8     1  285641210 2.1377e+10 7104.4
## + Round        22 2407907312 1.9254e+10 7105.0
## + Line          1  138109485 2.1524e+10 7107.2
## <none>                       2.1662e+10 7107.7
## + Home_top8     1   76305294 2.1586e+10 7108.3
## + Away_top4     1   46902746 2.1615e+10 7108.8
## + Home_top2     1   32037936 2.1630e+10 7109.1
## + TimeBucket    1   13738920 2.1648e+10 7109.4
## + Home_top4     1   12170476 2.1650e+10 7109.5
## + Venue        17 1687943749 1.9974e+10 7109.6
## + X             1    6545323 2.1656e+10 7109.6
## + ...1          1    6545323 2.1656e+10 7109.6
## + Away_top2     1    4662465 2.1657e+10 7109.6
## + Year          1    3300634 2.1659e+10 7109.6
## + Home.Odds     1      91378 2.1662e+10 7109.7
## + Time         27 1730661889 1.9931e+10 7128.7
##
## Step:  AIC=7097.86
## Attendance ~ HomeVenueAvg + Day + Away + Home_elo + Favorite
##
##                Df  Sum of Sq        RSS    AIC
## + venueavg      1  489577756 2.0535e+10 7090.5
## + Round        22 2378365662 1.8646e+10 7094.3
## + Away_top8     1  209675445 2.0815e+10 7095.9
## + Home.Odds     1  205403125 2.0819e+10 7096.0
## + Away_elo      1  200341124 2.0824e+10 7096.1
## + Line          1  157162647 2.0867e+10 7096.9
## + Away.Odds     1  156899719 2.0868e+10 7096.9
## + Home_top8     1  139865580 2.0885e+10 7097.2
## + Home_top2     1  116910772 2.0908e+10 7097.7
## <none>                       2.1024e+10 7097.9
## + Home_top4     1   65237639 2.0959e+10 7098.6
## + Away_top4     1   39765819 2.0985e+10 7099.1
## + TimeBucket    1   15553100 2.1009e+10 7099.6
## + Underdog      1   13503467 2.1011e+10 7099.6
## + Differential  1   13503467 2.1011e+10 7099.6
## + Away_top2     1    6012881 2.1018e+10 7099.8
## + X             1    1149536 2.1023e+10 7099.8
## + ...1          1    1149536 2.1023e+10 7099.8
## + Year          1      75400 2.1024e+10 7099.9
```

```
## + Venue        17 1528278024 1.9496e+10 7102.0
## + Time         27 1562290299 1.9462e+10 7121.3
##
## Step:  AIC=7090.53
## Attendance ~ HomeVenueAvg + Day + Away + Home_elo + Favorite +
##     venueavg
##
##                Df  Sum of Sq        RSS    AIC
## + Away_top8     1  241554367 2.0293e+10 7087.8
## + Away_elo      1  234636258 2.0300e+10 7088.0
## + Home.Odds     1  223323411 2.0312e+10 7088.2
## + Round        22 2266963595 1.8268e+10 7088.2
## + Home_top2     1  179528292 2.0355e+10 7089.1
## + Line          1  169739818 2.0365e+10 7089.2
## + Away.Odds     1  168989194 2.0366e+10 7089.3
## <none>                      2.0535e+10 7090.5
## + Home_top4     1   99084282 2.0436e+10 7090.6
## + Home_top8     1   90574976 2.0444e+10 7090.8
## + Away_top4     1   67856945 2.0467e+10 7091.2
## + Away_top2     1   22783975 2.0512e+10 7092.1
## + TimeBucket    1   17151478 2.0518e+10 7092.2
## + Underdog      1   14218132 2.0521e+10 7092.3
## + Differential  1   14218132 2.0521e+10 7092.3
## + X             1    3385253 2.0531e+10 7092.5
## + ...1          1    3385253 2.0531e+10 7092.5
## + Year          1     367105 2.0534e+10 7092.5
## + Venue        16 1038700268 1.9496e+10 7102.0
## + Time         27 1566669949 1.8968e+10 7113.1
##
## Step:  AIC=7087.85
## Attendance ~ HomeVenueAvg + Day + Away + Home_elo + Favorite +
##     venueavg + Away_top8
##
##                Df  Sum of Sq        RSS    AIC
## + Round        22 2228494905 1.8065e+10 7085.8
## + Line          1  183023623 2.0110e+10 7086.3
## + Home_top2     1  180697021 2.0113e+10 7086.3
## + Home_top4     1  106050935 2.0187e+10 7087.8
## <none>                      2.0293e+10 7087.8
## + Home_top8     1   97303584 2.0196e+10 7087.9
## + Home.Odds     1   86910375 2.0206e+10 7088.1
## + Away.Odds     1   79131629 2.0214e+10 7088.3
## + Away_top4     1   48940448 2.0244e+10 7088.9
## + Away_elo      1   42365493 2.0251e+10 7089.0
## + Away_top2     1   31100075 2.0262e+10 7089.2
## + TimeBucket    1   21309183 2.0272e+10 7089.4
## + Underdog      1   10803289 2.0282e+10 7089.6
## + Differential  1   10803289 2.0282e+10 7089.6
## + X             1    5202316 2.0288e+10 7089.7
## + ...1          1    5202316 2.0288e+10 7089.7
## + Year          1     873802 2.0292e+10 7089.8
## + Venue        16  985309758 1.9308e+10 7100.1
## + Time         27 1464032984 1.8829e+10 7112.2
##
```

```
## Step:  AIC=7085.78
## Attendance ~ HomeVenueAvg + Day + Away + Home_elo + Favorite +
##     venueavg + Away_top8 + Round
##
##                Df  Sum of Sq        RSS    AIC
## + Home_top2     1  187405558 1.7877e+10 7083.7
## + Line          1  158073278 1.7907e+10 7084.3
## + Home_top8     1  127508309 1.7937e+10 7085.0
## + Home_top4     1  119690950 1.7945e+10 7085.2
## <none>                       1.8065e+10 7085.8
## + Away.Odds     1   74841741 1.7990e+10 7086.1
## + Home.Odds     1   69966177 1.7995e+10 7086.2
## + Away_elo      1   48795765 1.8016e+10 7086.7
## + Away_top4     1   44563193 1.8020e+10 7086.8
## + Away_top2     1   27512481 1.8037e+10 7087.2
## + Underdog      1   12818602 1.8052e+10 7087.5
## + Differential  1   12818602 1.8052e+10 7087.5
## + TimeBucket    1    6691659 1.8058e+10 7087.6
## + X             1     625942 1.8064e+10 7087.8
## + ...1          1     625942 1.8064e+10 7087.8
## + Year          1     595315 1.8064e+10 7087.8
## + Venue        16 1048343590 1.7016e+10 7094.1
## + Time         27 1526368035 1.6538e+10 7104.8
##
## Step:  AIC=7083.65
## Attendance ~ HomeVenueAvg + Day + Away + Home_elo + Favorite +
##     venueavg + Away_top8 + Round + Home_top2
##
##                Df  Sum of Sq        RSS    AIC
## + Line          1  152339938 1.7725e+10 7082.3
## <none>                       1.7877e+10 7083.7
## + Home.Odds     1   67858443 1.7810e+10 7084.1
## + Away.Odds     1   62306926 1.7815e+10 7084.3
## + Away_elo      1   61865168 1.7816e+10 7084.3
## + Away_top4     1   43191426 1.7834e+10 7084.7
## + Home_top8     1   41028438 1.7836e+10 7084.7
## + Away_top2     1   21112626 1.7856e+10 7085.2
## + Underdog      1    8498583 1.7869e+10 7085.5
## + Differential  1    8498583 1.7869e+10 7085.5
## + TimeBucket    1    4297743 1.7873e+10 7085.6
## + Home_top4     1    3282789 1.7874e+10 7085.6
## + X             1     579833 1.7877e+10 7085.6
## + ...1          1     579833 1.7877e+10 7085.6
## + Year          1     568147 1.7877e+10 7085.6
## + Venue        16 1065123720 1.6812e+10 7091.3
## + Time         27 1616287198 1.6261e+10 7100.1
##
## Step:  AIC=7082.26
## Attendance ~ HomeVenueAvg + Day + Away + Home_elo + Favorite +
##     venueavg + Away_top8 + Round + Home_top2 + Line
##
##                Df  Sum of Sq        RSS    AIC
## <none>                       1.7725e+10 7082.3
## + Home.Odds     1   80908174 1.7644e+10 7082.5
```

```
## + Away.Odds     1   70817976 1.7654e+10 7082.7
## + Away_top4     1   61192852 1.7664e+10 7082.9
## + Away_elo      1   58687043 1.7666e+10 7083.0
## + Home_top8     1   58176611 1.7667e+10 7083.0
## + Year          1   43859568 1.7681e+10 7083.3
## + X             1   43404840 1.7682e+10 7083.3
## + ...1          1   43404840 1.7682e+10 7083.3
## + Away_top2     1   37223184 1.7688e+10 7083.4
## + Underdog      1    8937415 1.7716e+10 7084.1
## + Differential  1    8937415 1.7716e+10 7084.1
## + TimeBucket    1    8790559 1.7716e+10 7084.1
## + Home_top4     1       1700 1.7725e+10 7084.3
## + Venue        16 1044952123 1.6680e+10 7090.2
## + Time         27 1707683069 1.6017e+10 7096.1
```

```
summary(fwd.model)
```

```
##
## Call:
## lm(formula = Attendance ~ HomeVenueAvg + Day + Away + Home_elo +
##     Favorite + venueavg + Away_top8 + Round + Home_top2 + Line,
##     data = subset_afl)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18655.9  -4263.7   -263.2   3618.3  27070.7
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   -4.738e+04  1.046e+04  -4.529 8.18e-06 ***
## HomeVenueAvg                   6.219e-01  8.001e-02   7.772 9.05e-14 ***
## DayMon                         2.354e+04  3.688e+03   6.383 5.65e-10 ***
## DaySat                        -2.072e+03  1.305e+03  -1.587 0.113350
## DaySun                        -2.803e+03  1.380e+03  -2.032 0.042959 *
## DayThu                        -7.049e+02  2.247e+03  -0.314 0.753949
## DayTue                         2.860e+04  5.696e+03   5.021 8.27e-07 ***
## DayWed                         2.639e+04  7.724e+03   3.417 0.000708 ***
## AwayBrisbane Lions             2.141e+03  2.562e+03   0.836 0.403953
## AwayCarlton                    1.207e+04  2.554e+03   4.727 3.34e-06 ***
## AwayCollingwood                1.042e+04  2.309e+03   4.511 8.86e-06 ***
## AwayEssendon                   1.255e+04  2.448e+03   5.124 5.00e-07 ***
## AwayFremantle                  3.599e+03  2.498e+03   1.441 0.150537
## AwayGeelong                    4.350e+03  2.257e+03   1.928 0.054736 .
## AwayGold Coast                 1.646e+03  2.538e+03   0.649 0.517048
## AwayGreater Western Sydney    -2.516e+03  2.217e+03  -1.135 0.257205
## AwayHawthorn                   8.055e+03  2.295e+03   3.510 0.000509 ***
## AwayMelbourne                  4.145e+03  2.242e+03   1.849 0.065381 .
## AwayNorth Melbourne            3.361e+03  2.372e+03   1.417 0.157522
## AwayPort Adelaide             -5.899e+02  2.268e+03  -0.260 0.794980
## AwayRichmond                   9.541e+03  2.280e+03   4.185 3.63e-05 ***
## AwaySt Kilda                   7.522e+03  2.456e+03   3.063 0.002362 **
## AwaySydney                     4.214e+03  2.294e+03   1.837 0.067022 .
## AwayWest Coast                -3.003e+02  2.231e+03  -0.135 0.893014
## AwayWestern Bulldogs           2.678e+03  2.359e+03   1.135 0.257044
## Home_elo                       2.305e+01  5.007e+00   4.604 5.86e-06 ***
```

```
## Favorite                    6.471e+03  1.841e+03   3.515 0.000499 ***
## venueavg                    2.628e-01  8.611e-02   3.052 0.002454 **
## Away_top8                   3.635e+01  1.777e+01   2.046 0.041513 *
## Round2                     -9.547e+03  2.468e+03  -3.868 0.000132 ***
## Round3                     -5.637e+03  2.441e+03  -2.309 0.021517 *
## Round4                     -9.751e+03  2.453e+03  -3.975 8.57e-05 ***
## Round5                     -6.542e+03  2.562e+03  -2.554 0.011095 *
## Round6                     -7.297e+03  2.451e+03  -2.977 0.003120 **
## Round7                     -6.848e+03  2.449e+03  -2.796 0.005466 **
## Round8                     -1.043e+04  2.432e+03  -4.289 2.33e-05 ***
## Round9                     -7.915e+03  2.474e+03  -3.200 0.001504 **
## Round10                    -6.191e+03  2.493e+03  -2.484 0.013485 *
## Round11                    -7.775e+03  2.556e+03  -3.041 0.002537 **
## Round12                    -9.324e+03  2.683e+03  -3.475 0.000575 ***
## Round13                    -1.203e+04  2.751e+03  -4.374 1.62e-05 ***
## Round14                    -9.188e+03  2.597e+03  -3.538 0.000458 ***
## Round15                    -6.397e+03  2.469e+03  -2.591 0.009987 **
## Round16                    -5.599e+03  2.447e+03  -2.289 0.022711 *
## Round17                    -7.019e+03  2.441e+03  -2.875 0.004290 **
## Round18                    -7.294e+03  2.455e+03  -2.971 0.003176 **
## Round19                    -5.309e+03  2.494e+03  -2.129 0.033959 *
## Round20                    -5.813e+03  2.461e+03  -2.362 0.018729 *
## Round21                    -4.014e+03  2.467e+03  -1.627 0.104658
## Round22                    -5.986e+03  2.466e+03  -2.428 0.015715 *
## Round23                    -4.029e+03  2.448e+03  -1.646 0.100749
## Home_top2                   4.968e+01  2.649e+01   1.875 0.061644 .
## Line                        5.385e+01  3.136e+01   1.717 0.086889 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7189 on 343 degrees of freedom
## Multiple R-squared:  0.8484, Adjusted R-squared:  0.8255
## F-statistic: 36.93 on 52 and 343 DF,  p-value: < 2.2e-16
```

```
both.model <- step(max.model2, direction='both')
```

```
## Start:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + Underdog + Differential + venueavg +
##     HomeVenueAvg + TimeBucket + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_elo + Away_top8 + Away_top4 + Away_top2
##
##
## Step:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + Underdog + Differential + venueavg +
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2 +
##     Away_elo + Away_top8 + Away_top4 + Away_top2
##
##
## Step:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + Underdog + Differential + HomeVenueAvg +
##     Home_elo + Home_top8 + Home_top4 + Home_top2 + Away_elo +
##     Away_top8 + Away_top4 + Away_top2
```

```
##
##
## Step:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + Underdog + HomeVenueAvg + Home_elo +
##     Home_top8 + Home_top4 + Home_top2 + Away_elo + Away_top8 +
##     Away_top4 + Away_top2
##
##
## Step:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 +
##     Home_top4 + Home_top2 + Away_elo + Away_top8 + Away_top4 +
##     Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Time         27 1220358880 1.6371e+10 7098.8
## - Venue        17  800553764 1.5951e+10 7108.5
## - Away_top2     1    3998937 1.5154e+10 7120.2
## - Away_elo      1    8845338 1.5159e+10 7120.3
## - Away_top4     1   11765079 1.5162e+10 7120.4
## - Away_top8     1   13878702 1.5164e+10 7120.5
## - Away.Odds     1   15059933 1.5165e+10 7120.5
## - Home.Odds     1   15162751 1.5166e+10 7120.5
## - Year          1   16150021 1.5167e+10 7120.5
## - Home_top4     1   37178902 1.5188e+10 7121.1
## <none>                       1.5150e+10 7122.1
## - Home_top2     1  115605534 1.5266e+10 7123.1
## - Home_top8     1  129268554 1.5280e+10 7123.5
## - Home_elo      1  146562061 1.5297e+10 7123.9
## - Line          1  184838203 1.5335e+10 7124.9
## - Favorite      1  272549273 1.5423e+10 7127.2
## - Round        22 2425971131 1.7576e+10 7136.9
## - HomeVenueAvg  1 3095155609 1.8246e+10 7193.7
## - Away         17 5266349331 2.0417e+10 7206.3
## - Day           6 4198481838 1.9349e+10 7207.0
##
## Step:  AIC=7098.79
## Attendance ~ Round + Day + Year + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 +
##     Home_top4 + Home_top2 + Away_elo + Away_top8 + Away_top4 +
##     Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Venue        17 1226992719 1.7598e+10 7093.4
## - Away_top2     1      59551 1.6371e+10 7096.8
## - Home.Odds     1    7896049 1.6379e+10 7097.0
## - Away_top4     1   10314007 1.6381e+10 7097.0
## - Away_elo      1   18088561 1.6389e+10 7097.2
## - Year          1   19836896 1.6391e+10 7097.3
## - Away.Odds     1   20123412 1.6391e+10 7097.3
## - Away_top8     1   38787050 1.6410e+10 7097.7
## - Home_top4     1   38868590 1.6410e+10 7097.7
## <none>                       1.6371e+10 7098.8
```

```
## + TimeBucket    1   66903563 1.6304e+10 7099.2
## - Home_top2     1  102942499 1.6474e+10 7099.3
## - Home_top8     1  132097734 1.6503e+10 7100.0
## - Home_elo      1  158399991 1.6529e+10 7100.6
## - Line          1  169585591 1.6540e+10 7100.9
## - Favorite      1  236511266 1.6607e+10 7102.5
## - Round        22 2256426972 1.8627e+10 7105.9
## + Time         27 1220358880 1.5150e+10 7122.1
## - HomeVenueAvg  1 3080778894 1.9452e+10 7165.1
## - Day           6 4435713117 2.0806e+10 7181.7
## - Away         17 6136061670 2.2507e+10 7190.8
##
## Step:  AIC=7093.41
## Attendance ~ Round + Day + Year + Away + Home.Odds + Away.Odds +
##     Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_elo + Away_top8 + Away_top4 + Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## + venueavg      1 3.0262e+08 1.7295e+10 7088.5
## - Away_elo      1 3.3622e+05 1.7598e+10 7091.4
## - Away_top4     1 1.8128e+06 1.7600e+10 7091.5
## - Away_top2     1 8.9731e+06 1.7607e+10 7091.6
## - Away_top8     1 4.2757e+07 1.7641e+10 7092.4
## - Home.Odds     1 5.0882e+07 1.7649e+10 7092.6
## - Year          1 6.8563e+07 1.7666e+10 7093.0
## - Away.Odds     1 7.8575e+07 1.7676e+10 7093.2
## <none>                       1.7598e+10 7093.4
## + TimeBucket    1 1.4742e+07 1.7583e+10 7095.1
## - Home_top4     1 1.8813e+08 1.7786e+10 7095.6
## - Home_top2     1 2.1238e+08 1.7810e+10 7096.2
## - Home_elo      1 2.3375e+08 1.7832e+10 7096.6
## - Line          1 2.3819e+08 1.7836e+10 7096.7
## - Favorite      1 2.8964e+08 1.7887e+10 7097.9
## - Home_top8     1 3.2660e+08 1.7924e+10 7098.7
## + Venue        17 1.2270e+09 1.6371e+10 7098.8
## - Round        22 2.3579e+09 1.9956e+10 7099.2
## + Time         27 1.6468e+09 1.5951e+10 7108.5
## - Day           6 4.7685e+09 2.2366e+10 7176.4
## - Away         17 6.0553e+09 2.3653e+10 7176.5
## - HomeVenueAvg  1 4.5913e+10 6.3511e+10 7599.7
##
## Step:  AIC=7088.54
## Attendance ~ Round + Day + Year + Away + Home.Odds + Away.Odds +
##     Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_elo + Away_top8 + Away_top4 + Away_top2 +
##     venueavg
##
##                Df  Sum of Sq        RSS    AIC
## - Away_elo      1    1921258 1.7297e+10 7086.6
## - Away_top2     1    3074510 1.7298e+10 7086.6
## - Away_top4     1    3620856 1.7299e+10 7086.6
## - Away_top8     1   45081359 1.7340e+10 7087.6
## - Home.Odds     1   46064942 1.7341e+10 7087.6
## - Year          1   55158375 1.7350e+10 7087.8
```

```
## - Away.Odds     1   60221424 1.7355e+10 7087.9
## <none>                       1.7295e+10 7088.5
## - Home_top4     1  114484715 1.7410e+10 7089.2
## + TimeBucket    1   14983556 1.7280e+10 7090.2
## - Home_top2     1  176747078 1.7472e+10 7090.6
## - Home_top8     1  187726519 1.7483e+10 7090.8
## - Line          1  232730089 1.7528e+10 7091.8
## - Home_elo      1  257069349 1.7552e+10 7092.4
## - Round        22 2252558666 1.9548e+10 7093.0
## - Favorite      1  296001860 1.7591e+10 7093.3
## - venueavg      1  302615370 1.7598e+10 7093.4
## + Venue        16  924377349 1.6371e+10 7098.8
## + Time         27 1630365080 1.5665e+10 7103.3
## - HomeVenueAvg  1 3316482950 2.0612e+10 7156.0
## - Day           6 4567785300 2.1863e+10 7169.4
## - Away         17 5996932971 2.3292e+10 7172.4
##
## Step:  AIC=7086.59
## Attendance ~ Round + Day + Year + Away + Home.Odds + Away.Odds +
##     Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_top8 + Away_top4 + Away_top2 + venueavg
##
##                Df  Sum of Sq        RSS    AIC
## - Away_top2     1    2552000 1.7300e+10 7084.6
## - Away_top4     1    3756862 1.7301e+10 7084.7
## - Away_top8     1   55614132 1.7353e+10 7085.9
## - Year          1   56065810 1.7353e+10 7085.9
## - Home.Odds     1   57297879 1.7354e+10 7085.9
## - Away.Odds     1   85760167 1.7383e+10 7086.5
## <none>                       1.7297e+10 7086.6
## - Home_top4     1  119028291 1.7416e+10 7087.3
## + TimeBucket    1   15627184 1.7281e+10 7088.2
## + Away_elo      1    1921258 1.7295e+10 7088.5
## - Home_top2     1  178461198 1.7476e+10 7088.7
## - Home_top8     1  197181652 1.7494e+10 7089.1
## - Line          1  235234985 1.7532e+10 7089.9
## - Round        22 2251486512 1.9549e+10 7091.0
## - Home_elo      1  296147556 1.7593e+10 7091.3
## - venueavg      1  301030334 1.7598e+10 7091.4
## - Favorite      1  301199111 1.7598e+10 7091.4
## + Venue        16  908210046 1.6389e+10 7097.2
## + Time         27 1630180801 1.5667e+10 7101.4
## - HomeVenueAvg  1 3334069516 2.0631e+10 7154.4
## - Day           6 4568597432 2.1866e+10 7167.4
## - Away         17 6006888519 2.3304e+10 7170.6
##
## Step:  AIC=7084.64
## Attendance ~ Round + Day + Year + Away + Home.Odds + Away.Odds +
##     Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_top8 + Away_top4 + venueavg
##
##                Df  Sum of Sq        RSS    AIC
## - Home.Odds     1   54749708 1.7354e+10 7083.9
## - Year          1   54877360 1.7354e+10 7083.9
```

```
## - Away_top4     1   56969527 1.7357e+10 7083.9
## - Away.Odds     1   84896288 1.7385e+10 7084.6
## <none>                       1.7300e+10 7084.6
## - Away_top8     1   96134183 1.7396e+10 7084.8
## - Home_top4     1  116936136 1.7417e+10 7085.3
## + TimeBucket    1   16040837 1.7284e+10 7086.3
## + Away_top2     1    2552000 1.7297e+10 7086.6
## + Away_elo      1    1398748 1.7298e+10 7086.6
## - Home_top2     1  177184394 1.7477e+10 7086.7
## - Home_top8     1  195170445 1.7495e+10 7087.1
## - Line          1  232702363 1.7532e+10 7087.9
## - Round        22 2250285369 1.9550e+10 7089.1
## - Home_elo      1  294257382 1.7594e+10 7089.3
## - Favorite      1  298715125 1.7598e+10 7089.4
## - venueavg      1  307158436 1.7607e+10 7089.6
## + Venue        16  910714495 1.6389e+10 7095.2
## + Time         27 1632525056 1.5667e+10 7099.4
## - HomeVenueAvg  1 3336477139 2.0636e+10 7152.5
## - Day           6 4566348628 2.1866e+10 7165.4
## - Away         17 6004341201 2.3304e+10 7168.6
##
## Step:  AIC=7083.9
## Attendance ~ Round + Day + Year + Away + Away.Odds + Line + Favorite +
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2 +
##     Away_top8 + Away_top4 + venueavg
##
##                 Df  Sum of Sq        RSS    AIC
## - Away_top4      1   50315111 1.7405e+10 7083.0
## - Year           1   61500207 1.7416e+10 7083.3
## <none>                        1.7354e+10 7083.9
## - Home_top4      1  112236712 1.7467e+10 7084.4
## + Home.Odds      1   54749708 1.7300e+10 7084.6
## + Underdog       1   54749708 1.7300e+10 7084.6
## + Differential   1   54749708 1.7300e+10 7084.6
## - Away.Odds      1  124718078 1.7479e+10 7084.7
## - Away_top8      1  132409166 1.7487e+10 7084.9
## + Away_elo       1   12789519 1.7342e+10 7085.6
## + TimeBucket     1    8908449 1.7345e+10 7085.7
## - Home_top2      1  170671357 1.7525e+10 7085.8
## + Away_top2      1       3829 1.7354e+10 7085.9
## - Home_top8      1  194975281 1.7549e+10 7086.3
## - Line           1  229567084 1.7584e+10 7087.1
## - Home_elo       1  239675331 1.7594e+10 7087.3
## - Favorite       1  252329406 1.7607e+10 7087.6
## - Round         22 2260756189 1.9615e+10 7088.4
## - venueavg       1  304285405 1.7659e+10 7088.8
## + Venue         16  943921903 1.6410e+10 7093.7
## + Time          27 1650753801 1.5704e+10 7098.3
## - HomeVenueAvg   1 3368975554 2.0723e+10 7152.2
## - Day            6 4593478747 2.1948e+10 7164.9
## - Away          17 6041300055 2.3396e+10 7168.2
##
## Step:  AIC=7083.04
## Attendance ~ Round + Day + Year + Away + Away.Odds + Line + Favorite +
```

```
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2 +
##     Away_top8 + venueavg
##
##                 Df  Sum of Sq        RSS    AIC
## - Year           1   60971231 1.7466e+10 7082.4
## <none>                        1.7405e+10 7083.0
## - Away_top8      1   93649275 1.7498e+10 7083.2
## - Home_top4      1  106899204 1.7512e+10 7083.5
## + Away_top4      1   50315111 1.7354e+10 7083.9
## + Home.Odds      1   48095292 1.7357e+10 7083.9
## + Underdog       1   48095292 1.7357e+10 7083.9
## + Differential   1   48095292 1.7357e+10 7083.9
## - Away.Odds      1  136352652 1.7541e+10 7084.1
## + Away_top2      1   39693069 1.7365e+10 7084.1
## - Home_top2      1  162980574 1.7568e+10 7084.7
## + TimeBucket     1    9109049 1.7396e+10 7084.8
## + Away_elo       1    4552245 1.7400e+10 7084.9
## - Home_top8      1  198694090 1.7603e+10 7085.5
## - Line           1  215098308 1.7620e+10 7085.9
## - Home_elo       1  242526349 1.7647e+10 7086.5
## - Favorite       1  263086355 1.7668e+10 7087.0
## - Round         22 2262321251 1.9667e+10 7087.4
## - venueavg       1  326966803 1.7732e+10 7088.4
## + Venue         16  949934023 1.6455e+10 7092.8
## + Time          27 1687391068 1.5717e+10 7096.7
## - HomeVenueAvg   1 3323224382 2.0728e+10 7150.2
## - Day            6 4601896164 2.2007e+10 7163.9
## - Away          17 6017178318 2.3422e+10 7166.6
##
## Step:  AIC=7082.43
## Attendance ~ Round + Day + Away + Away.Odds + Line + Favorite +
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2 +
##     Away_top8 + venueavg
##
##                 Df  Sum of Sq        RSS    AIC
## <none>                        1.7466e+10 7082.4
## - Away_top8      1   97939960 1.7564e+10 7082.6
## - Home_top4      1  108809656 1.7574e+10 7082.9
## + Year           1   60971231 1.7405e+10 7083.0
## + Home.Odds      1   54317730 1.7411e+10 7083.2
## + Underdog       1   54317730 1.7411e+10 7083.2
## + Differential   1   54317730 1.7411e+10 7083.2
## - Away.Odds      1  125070105 1.7591e+10 7083.3
## + Away_top4      1   49786134 1.7416e+10 7083.3
## + Away_top2      1   36240264 1.7429e+10 7083.6
## - Line           1  154142531 1.7620e+10 7083.9
## - Home_top2      1  169525329 1.7635e+10 7084.3
## + TimeBucket     1    7440932 1.7458e+10 7084.3
## + Away_elo       1    6849352 1.7459e+10 7084.3
## - Home_top8      1  188568838 1.7654e+10 7084.7
## - Home_elo       1  241649088 1.7707e+10 7085.9
## - Favorite       1  247435009 1.7713e+10 7086.0
## - Round         22 2245130032 1.9711e+10 7086.3
## - venueavg       1  339429158 1.7805e+10 7088.0
```

```
## + Venue        16  985707985 1.6480e+10 7091.4
## + Time         27 1698086048 1.5768e+10 7095.9
## - HomeVenueAvg  1 3286176862 2.0752e+10 7148.7
## - Day           6 4604563230 2.2070e+10 7163.1
## - Away         17 5957971714 2.3424e+10 7164.7
```

```
summary(both.model)
```

```
##
## Call:
## lm(formula = Attendance ~ Round + Day + Away + Away.Odds + Line +
##     Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_top8 + venueavg, data = subset_afl)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18748.0  -4012.9   -206.7   3434.1  26843.1
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)               -3.821e+04  1.333e+04  -2.866 0.004420 **
## Round2                    -9.511e+03  2.462e+03  -3.864 0.000134 ***
## Round3                    -5.496e+03  2.435e+03  -2.257 0.024621 *
## Round4                    -9.758e+03  2.450e+03  -3.983 8.33e-05 ***
## Round5                    -6.217e+03  2.562e+03  -2.427 0.015764 *
## Round6                    -7.322e+03  2.445e+03  -2.994 0.002954 **
## Round7                    -6.481e+03  2.449e+03  -2.646 0.008523 **
## Round8                    -1.050e+04  2.426e+03  -4.331 1.96e-05 ***
## Round9                    -7.946e+03  2.467e+03  -3.221 0.001400 **
## Round10                   -5.959e+03  2.492e+03  -2.392 0.017322 *
## Round11                   -7.363e+03  2.564e+03  -2.871 0.004346 **
## Round12                   -8.832e+03  2.685e+03  -3.289 0.001112 **
## Round13                   -1.198e+04  2.752e+03  -4.353 1.78e-05 ***
## Round14                   -8.893e+03  2.604e+03  -3.415 0.000716 ***
## Round15                   -6.036e+03  2.472e+03  -2.442 0.015109 *
## Round16                   -5.240e+03  2.446e+03  -2.142 0.032875 *
## Round17                   -7.079e+03  2.435e+03  -2.907 0.003888 **
## Round18                   -7.196e+03  2.453e+03  -2.934 0.003576 **
## Round19                   -4.678e+03  2.510e+03  -1.863 0.063266 .
## Round20                   -5.550e+03  2.457e+03  -2.258 0.024560 *
## Round21                   -4.308e+03  2.464e+03  -1.748 0.081288 .
## Round22                   -6.023e+03  2.464e+03  -2.444 0.015026 *
## Round23                   -3.658e+03  2.464e+03  -1.484 0.138654
## DayMon                     2.329e+04  3.691e+03   6.309 8.73e-10 ***
## DaySat                    -2.145e+03  1.318e+03  -1.627 0.104673
## DaySun                    -2.903e+03  1.388e+03  -2.092 0.037178 *
## DayThu                    -6.815e+02  2.245e+03  -0.304 0.761593
## DayTue                     2.830e+04  5.683e+03   4.981 1.01e-06 ***
## DayWed                     2.591e+04  7.713e+03   3.360 0.000868 ***
## AwayBrisbane Lions         2.842e+03  2.626e+03   1.082 0.279873
## AwayCarlton                1.275e+04  2.606e+03   4.892 1.54e-06 ***
## AwayCollingwood            1.054e+04  2.303e+03   4.576 6.66e-06 ***
## AwayEssendon               1.239e+04  2.445e+03   5.068 6.62e-07 ***
## AwayFremantle              3.904e+03  2.530e+03   1.543 0.123770
## AwayGeelong                4.703e+03  2.259e+03   2.082 0.038076 *
```

```
## AwayGold Coast              2.673e+03  2.669e+03   1.001 0.317313
## AwayGreater Western Sydney -2.674e+03  2.212e+03  -1.209 0.227569
## AwayHawthorn                8.382e+03  2.295e+03   3.652 0.000301 ***
## AwayMelbourne               4.306e+03  2.238e+03   1.924 0.055148 .
## AwayNorth Melbourne         3.870e+03  2.379e+03   1.627 0.104707
## AwayPort Adelaide          -3.962e+02  2.267e+03  -0.175 0.861385
## AwayRichmond                9.860e+03  2.279e+03   4.326 2.00e-05 ***
## AwaySt Kilda                7.966e+03  2.460e+03   3.238 0.001324 **
## AwaySydney                  4.482e+03  2.291e+03   1.956 0.051241 .
## AwayWest Coast              2.680e+02  2.243e+03   0.119 0.904959
## AwayWestern Bulldogs        2.921e+03  2.357e+03   1.239 0.216128
## Away.Odds                  -5.329e+02  3.415e+02  -1.560 0.119606
## Line                        5.478e+01  3.162e+01   1.732 0.084138 .
## Favorite                    4.815e+03  2.194e+03   2.195 0.028860 *
## HomeVenueAvg                6.558e-01  8.199e-02   7.998 2.00e-14 ***
## Home_elo                    1.841e+01  8.488e+00   2.169 0.030783 *
## Home_top8                   7.022e+01  3.665e+01   1.916 0.056211 .
## Home_top4                  -9.728e+01  6.684e+01  -1.455 0.146483
## Home_top2                   1.163e+02  6.403e+01   1.817 0.070156 .
## Away_top8                   2.580e+01  1.869e+01   1.381 0.168251
## venueavg                    2.262e-01  8.798e-02   2.571 0.010580 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7167 on 340 degrees of freedom
## Multiple R-squared:  0.8507, Adjusted R-squared:  0.8265
## F-statistic: 35.21 on 55 and 340 DF,  p-value: < 2.2e-16
```

```
back.model <- step(max.model2)
```

```
## Start:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + Underdog + Differential + venueavg +
##     HomeVenueAvg + TimeBucket + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_elo + Away_top8 + Away_top4 + Away_top2
##
##
## Step:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + Underdog + Differential + venueavg +
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2 +
##     Away_elo + Away_top8 + Away_top4 + Away_top2
##
##
## Step:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + Underdog + Differential + HomeVenueAvg +
##     Home_elo + Home_top8 + Home_top4 + Home_top2 + Away_elo +
##     Away_top8 + Away_top4 + Away_top2
##
##
## Step:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + Underdog + HomeVenueAvg + Home_elo +
##     Home_top8 + Home_top4 + Home_top2 + Away_elo + Away_top8 +
```

```
##     Away_top4 + Away_top2
##
##
## Step:  AIC=7122.11
## Attendance ~ Round + Day + Year + Time + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 +
##     Home_top4 + Home_top2 + Away_elo + Away_top8 + Away_top4 +
##     Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Time         27 1220358880 1.6371e+10 7098.8
## - Venue        17  800553764 1.5951e+10 7108.5
## - Away_top2     1    3998937 1.5154e+10 7120.2
## - Away_elo      1    8845338 1.5159e+10 7120.3
## - Away_top4     1   11765079 1.5162e+10 7120.4
## - Away_top8     1   13878702 1.5164e+10 7120.5
## - Away.Odds     1   15059933 1.5165e+10 7120.5
## - Home.Odds     1   15162751 1.5166e+10 7120.5
## - Year          1   16150021 1.5167e+10 7120.5
## - Home_top4     1   37178902 1.5188e+10 7121.1
## <none>                       1.5150e+10 7122.1
## - Home_top2     1  115605534 1.5266e+10 7123.1
## - Home_top8     1  129268554 1.5280e+10 7123.5
## - Home_elo      1  146562061 1.5297e+10 7123.9
## - Line          1  184838203 1.5335e+10 7124.9
## - Favorite      1  272549273 1.5423e+10 7127.2
## - Round        22 2425971131 1.7576e+10 7136.9
## - HomeVenueAvg  1 3095155609 1.8246e+10 7193.7
## - Away         17 5266349331 2.0417e+10 7206.3
## - Day           6 4198481838 1.9349e+10 7207.0
##
## Step:  AIC=7098.79
## Attendance ~ Round + Day + Year + Away + Venue + Home.Odds +
##     Away.Odds + Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 +
##     Home_top4 + Home_top2 + Away_elo + Away_top8 + Away_top4 +
##     Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Venue        17 1226992719 1.7598e+10 7093.4
## - Away_top2     1      59551 1.6371e+10 7096.8
## - Home.Odds     1    7896049 1.6379e+10 7097.0
## - Away_top4     1   10314007 1.6381e+10 7097.0
## - Away_elo      1   18088561 1.6389e+10 7097.2
## - Year          1   19836896 1.6391e+10 7097.3
## - Away.Odds     1   20123412 1.6391e+10 7097.3
## - Away_top8     1   38787050 1.6410e+10 7097.7
## - Home_top4     1   38868590 1.6410e+10 7097.7
## <none>                       1.6371e+10 7098.8
## - Home_top2     1  102942499 1.6474e+10 7099.3
## - Home_top8     1  132097734 1.6503e+10 7100.0
## - Home_elo      1  158399991 1.6529e+10 7100.6
## - Line          1  169585591 1.6540e+10 7100.9
## - Favorite      1  236511266 1.6607e+10 7102.5
## - Round        22 2256426972 1.8627e+10 7105.9
```

```
## - HomeVenueAvg  1 3080778894 1.9452e+10 7165.1
## - Day           6 4435713117 2.0806e+10 7181.7
## - Away         17 6136061670 2.2507e+10 7190.8
##
## Step:  AIC=7093.41
## Attendance ~ Round + Day + Year + Away + Home.Odds + Away.Odds +
##     Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_elo + Away_top8 + Away_top4 + Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Away_elo      1 3.3622e+05 1.7598e+10 7091.4
## - Away_top4     1 1.8128e+06 1.7600e+10 7091.5
## - Away_top2     1 8.9731e+06 1.7607e+10 7091.6
## - Away_top8     1 4.2757e+07 1.7641e+10 7092.4
## - Home.Odds     1 5.0882e+07 1.7649e+10 7092.6
## - Year          1 6.8563e+07 1.7666e+10 7093.0
## - Away.Odds     1 7.8575e+07 1.7676e+10 7093.2
## <none>                       1.7598e+10 7093.4
## - Home_top4     1 1.8813e+08 1.7786e+10 7095.6
## - Home_top2     1 2.1238e+08 1.7810e+10 7096.2
## - Home_elo      1 2.3375e+08 1.7832e+10 7096.6
## - Line          1 2.3819e+08 1.7836e+10 7096.7
## - Favorite      1 2.8964e+08 1.7887e+10 7097.9
## - Home_top8     1 3.2660e+08 1.7924e+10 7098.7
## - Round        22 2.3579e+09 1.9956e+10 7099.2
## - Day           6 4.7685e+09 2.2366e+10 7176.4
## - Away         17 6.0553e+09 2.3653e+10 7176.5
## - HomeVenueAvg  1 4.5913e+10 6.3511e+10 7599.7
##
## Step:  AIC=7091.42
## Attendance ~ Round + Day + Year + Away + Home.Odds + Away.Odds +
##     Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_top8 + Away_top4 + Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Away_top4     1 1.8560e+06 1.7600e+10 7089.5
## - Away_top2     1 8.6801e+06 1.7607e+10 7089.6
## - Away_top8     1 4.9374e+07 1.7647e+10 7090.5
## - Home.Odds     1 5.8878e+07 1.7657e+10 7090.7
## - Year          1 6.9020e+07 1.7667e+10 7091.0
## <none>                       1.7598e+10 7091.4
## - Away.Odds     1 1.0180e+08 1.7700e+10 7091.7
## - Home_top4     1 1.9167e+08 1.7790e+10 7093.7
## - Home_top2     1 2.1327e+08 1.7811e+10 7094.2
## - Line          1 2.3954e+08 1.7838e+10 7094.8
## - Home_elo      1 2.6231e+08 1.7860e+10 7095.3
## - Favorite      1 2.9275e+08 1.7891e+10 7096.0
## - Home_top8     1 3.3563e+08 1.7934e+10 7096.9
## - Round        22 2.3576e+09 1.9956e+10 7097.2
## - Day           6 4.7684e+09 2.2366e+10 7174.4
## - Away         17 6.0669e+09 2.3665e+10 7174.7
## - HomeVenueAvg  1 4.5950e+10 6.3549e+10 7597.9
##
## Step:  AIC=7089.46
```

```
## Attendance ~ Round + Day + Year + Away + Home.Odds + Away.Odds +
##     Line + Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2 + Away_top8 + Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Home.Odds     1 6.4271e+07 1.7664e+10 7088.9
## - Year          1 7.0321e+07 1.7670e+10 7089.0
## - Away_top8     1 8.6252e+07 1.7686e+10 7089.4
## - Away_top2     1 8.7651e+07 1.7688e+10 7089.4
## <none>                       1.7600e+10 7089.5
## - Away.Odds     1 1.0396e+08 1.7704e+10 7089.8
## - Home_top4     1 1.9427e+08 1.7794e+10 7091.8
## - Home_top2     1 2.1363e+08 1.7814e+10 7092.2
## - Line          1 2.4237e+08 1.7842e+10 7092.9
## - Home_elo      1 2.7150e+08 1.7871e+10 7093.5
## - Favorite      1 2.9748e+08 1.7897e+10 7094.1
## - Home_top8     1 3.3961e+08 1.7940e+10 7095.0
## - Round        22 2.3587e+09 1.9959e+10 7095.3
## - Day           6 4.7801e+09 2.2380e+10 7172.6
## - Away         17 6.0679e+09 2.3668e+10 7172.8
## - HomeVenueAvg  1 4.5967e+10 6.3567e+10 7596.0
##
## Step:  AIC=7088.9
## Attendance ~ Round + Day + Year + Away + Away.Odds + Line + Favorite +
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2 +
##     Away_top8 + Away_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Away_top2     1 6.7426e+07 1.7732e+10 7088.4
## - Year          1 7.7750e+07 1.7742e+10 7088.6
## <none>                       1.7664e+10 7088.9
## - Away_top8     1 1.3321e+08 1.7797e+10 7089.9
## - Away.Odds     1 1.5236e+08 1.7817e+10 7090.3
## - Home_top4     1 1.8589e+08 1.7850e+10 7091.0
## - Home_top2     1 2.0468e+08 1.7869e+10 7091.5
## - Home_elo      1 2.0749e+08 1.7872e+10 7091.5
## - Line          1 2.3603e+08 1.7900e+10 7092.2
## - Favorite      1 2.3762e+08 1.7902e+10 7092.2
## - Home_top8     1 3.3849e+08 1.8003e+10 7094.4
## - Round        22 2.3684e+09 2.0033e+10 7094.7
## - Day           6 4.8064e+09 2.2471e+10 7172.2
## - Away         17 6.1034e+09 2.3768e+10 7172.4
## - HomeVenueAvg  1 4.6389e+10 6.4054e+10 7597.0
##
## Step:  AIC=7088.41
## Attendance ~ Round + Day + Year + Away + Away.Odds + Line + Favorite +
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2 +
##     Away_top8
##
##                Df  Sum of Sq        RSS    AIC
## - Away_top8     1 6.8473e+07 1.7800e+10 7087.9
## - Year          1 7.3434e+07 1.7805e+10 7088.0
## <none>                       1.7732e+10 7088.4
## - Away.Odds     1 1.5830e+08 1.7890e+10 7089.9
```

```
## - Home_top4     1 1.7506e+08 1.7907e+10 7090.3
## - Home_top2     1 1.9486e+08 1.7927e+10 7090.7
## - Home_elo      1 2.0376e+08 1.7935e+10 7090.9
## - Line          1 2.1234e+08 1.7944e+10 7091.1
## - Favorite      1 2.5539e+08 1.7987e+10 7092.1
## - Home_top8     1 3.4258e+08 1.8074e+10 7094.0
## - Round        22 2.3722e+09 2.0104e+10 7094.1
## - Away         17 6.0686e+09 2.3800e+10 7171.0
## - Day           6 4.8024e+09 2.2534e+10 7171.3
## - HomeVenueAvg  1 4.6334e+10 6.4066e+10 7595.1
##
## Step:  AIC=7087.94
## Attendance ~ Round + Day + Year + Away + Away.Odds + Line + Favorite +
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2
##
##                Df  Sum of Sq        RSS    AIC
## - Year          1 7.6980e+07 1.7877e+10 7087.6
## <none>                       1.7800e+10 7087.9
## - Home_top4     1 1.9598e+08 1.7996e+10 7090.3
## - Line          1 2.1006e+08 1.8010e+10 7090.6
## - Home_top2     1 2.1088e+08 1.8011e+10 7090.6
## - Home_elo      1 2.2608e+08 1.8026e+10 7090.9
## - Favorite      1 2.3697e+08 1.8037e+10 7091.2
## - Away.Odds     1 2.5524e+08 1.8055e+10 7091.6
## - Home_top8     1 3.7723e+08 1.8177e+10 7094.2
## - Round        22 2.4177e+09 2.0218e+10 7094.4
## - Day           6 4.7828e+09 2.2583e+10 7170.2
## - Away         17 6.2183e+09 2.4018e+10 7172.6
## - HomeVenueAvg  1 4.6283e+10 6.4083e+10 7593.2
##
## Step:  AIC=7087.65
## Attendance ~ Round + Day + Away + Away.Odds + Line + Favorite +
##     HomeVenueAvg + Home_elo + Home_top8 + Home_top4 + Home_top2
##
##                Df  Sum of Sq        RSS    AIC
## <none>                       1.7877e+10 7087.6
## - Line          1 1.3490e+08 1.8012e+10 7088.6
## - Home_top4     1 2.0111e+08 1.8078e+10 7090.1
## - Favorite      1 2.1951e+08 1.8097e+10 7090.5
## - Home_top2     1 2.2047e+08 1.8098e+10 7090.5
## - Home_elo      1 2.2485e+08 1.8102e+10 7090.6
## - Away.Odds     1 2.3960e+08 1.8117e+10 7090.9
## - Round        22 2.3993e+09 2.0276e+10 7093.5
## - Home_top8     1 3.6481e+08 1.8242e+10 7093.6
## - Day           6 4.7870e+09 2.2664e+10 7169.6
## - Away         17 6.1431e+09 2.4020e+10 7170.6
## - HomeVenueAvg  1 4.6208e+10 6.4085e+10 7591.2
```

```
summary(back.model)
```

```
##
## Call:
## lm(formula = Attendance ~ Round + Day + Away + Away.Odds + Line +
##     Favorite + HomeVenueAvg + Home_elo + Home_top8 + Home_top4 +
##     Home_top2, data = subset_afl)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20807.9  -4217.0      5.8   3718.0  27612.0
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  -3.344e+04  1.334e+04  -2.506 0.012665 *
## Round2                       -9.431e+03  2.482e+03  -3.800 0.000171 ***
## Round3                       -4.971e+03  2.449e+03  -2.030 0.043111 *
## Round4                       -9.839e+03  2.470e+03  -3.983 8.32e-05 ***
## Round5                       -5.800e+03  2.580e+03  -2.248 0.025181 *
## Round6                       -7.772e+03  2.462e+03  -3.157 0.001734 **
## Round7                       -6.356e+03  2.468e+03  -2.575 0.010435 *
## Round8                       -1.065e+04  2.446e+03  -4.353 1.78e-05 ***
## Round9                       -8.227e+03  2.486e+03  -3.309 0.001035 **
## Round10                      -6.117e+03  2.505e+03  -2.442 0.015123 *
## Round11                      -7.624e+03  2.585e+03  -2.949 0.003405 **
## Round12                      -8.198e+03  2.692e+03  -3.046 0.002501 **
## Round13                      -1.225e+04  2.774e+03  -4.416 1.35e-05 ***
## Round14                      -8.954e+03  2.623e+03  -3.414 0.000717 ***
## Round15                      -5.589e+03  2.487e+03  -2.248 0.025244 *
## Round16                      -5.314e+03  2.466e+03  -2.155 0.031879 *
## Round17                      -7.289e+03  2.455e+03  -2.969 0.003201 **
## Round18                      -7.419e+03  2.473e+03  -3.000 0.002897 **
## Round19                      -4.475e+03  2.515e+03  -1.779 0.076070 .
## Round20                      -5.468e+03  2.479e+03  -2.206 0.028056 *
## Round21                      -4.366e+03  2.483e+03  -1.759 0.079533 .
## Round22                      -6.180e+03  2.485e+03  -2.487 0.013356 *
## Round23                      -3.414e+03  2.482e+03  -1.376 0.169876
## DayMon                        2.403e+04  3.701e+03   6.492 2.99e-10 ***
## DaySat                       -2.433e+03  1.326e+03  -1.835 0.067382 .
## DaySun                       -2.970e+03  1.395e+03  -2.129 0.033963 *
## DayThu                       -9.683e+02  2.255e+03  -0.429 0.667912
## DayTue                        2.834e+04  5.724e+03   4.950 1.17e-06 ***
## DayWed                        2.455e+04  7.760e+03   3.164 0.001696 **
## AwayBrisbane Lions            1.781e+03  2.488e+03   0.716 0.474666
## AwayCarlton                   1.166e+04  2.458e+03   4.743 3.10e-06 ***
## AwayCollingwood               1.009e+04  2.278e+03   4.429 1.28e-05 ***
## AwayEssendon                  1.137e+04  2.294e+03   4.955 1.14e-06 ***
## AwayFremantle                 2.954e+03  2.408e+03   1.227 0.220735
## AwayGeelong                   5.151e+03  2.255e+03   2.284 0.022989 *
## AwayGold Coast                1.723e+03  2.567e+03   0.671 0.502694
## AwayGreater Western Sydney -2.544e+03  2.228e+03  -1.142 0.254187
## AwayHawthorn                  7.845e+03  2.277e+03   3.446 0.000640 ***
## AwayMelbourne                 4.056e+03  2.252e+03   1.801 0.072561 .
## AwayNorth Melbourne           3.218e+03  2.277e+03   1.413 0.158511
## AwayPort Adelaide            -1.689e+02  2.281e+03  -0.074 0.941020
## AwayRichmond                  1.037e+04  2.276e+03   4.558 7.21e-06 ***
## AwaySt Kilda                  7.311e+03  2.319e+03   3.152 0.001764 **
## AwaySydney                    4.880e+03  2.304e+03   2.118 0.034873 *
## AwayWest Coast                5.910e+02  2.255e+03   0.262 0.793417
## AwayWestern Bulldogs          2.361e+03  2.311e+03   1.021 0.307810
## Away.Odds                    -6.999e+02  3.269e+02  -2.141 0.032984 *
```

```
## Line                         5.120e+01  3.187e+01   1.606 0.109096
## Favorite                     4.521e+03  2.206e+03   2.049 0.041201 *
## HomeVenueAvg                 8.528e-01  2.868e-02  29.732  < 2e-16 ***
## Home_elo                     1.765e+01  8.508e+00   2.074 0.038826 *
## Home_top8                    9.487e+01  3.591e+01   2.642 0.008626 **
## Home_top4                   -1.302e+02  6.639e+01  -1.961 0.050637 .
## Home_top2                    1.321e+02  6.432e+01   2.054 0.040765 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7230 on 342 degrees of freedom
## Multiple R-squared:  0.8471, Adjusted R-squared:  0.8235
## F-statistic: 35.76 on 53 and 342 DF,  p-value: < 2.2e-16
```