

# Stats 401: Analysis

Matt Suba and Matthew Varela

2025-04-30

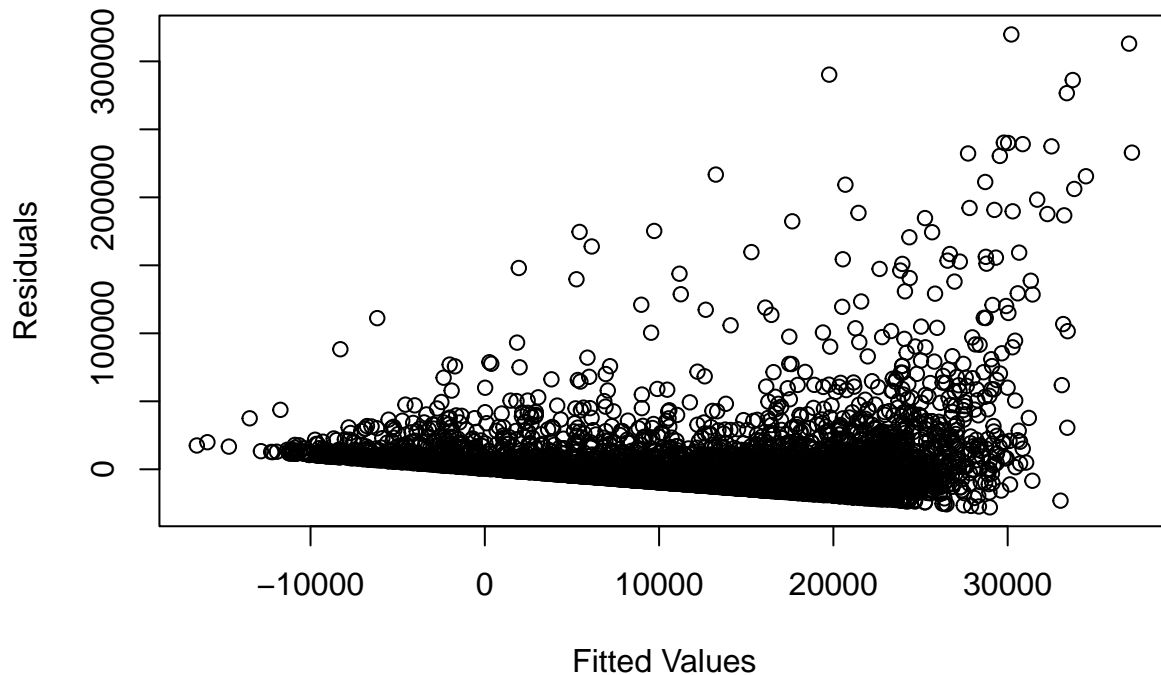
## Contents

Figure 1 . . . . .	2
Figure 2 . . . . .	4
Figure 3 . . . . .	5
Figure 4 . . . . .	7
Figure 5 . . . . .	9
Figure 6 . . . . .	12
Figure 7 . . . . .	13
Figure 8 . . . . .	14

Figure 1

```
##
## Call:
## lm(formula = wage_eur ~ shooting + pace + attacking_crossing +
##     preferred_foot, data = fifa)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -27976 -10873  -5621   3328 319790
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -27567.68    1887.88  -14.602 < 2e-16 ***
## shooting         350.20      23.26   15.055 < 2e-16 ***
## pace          -13.49       26.66   -0.506  0.61299
## attacking_crossing  355.63      26.03   13.661 < 2e-16 ***
## preferred_footRight 2013.25     658.55    3.057  0.00224 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23380 on 7005 degrees of freedom
## Multiple R-squared:  0.1173, Adjusted R-squared:  0.1168
## F-statistic: 232.8 on 4 and 7005 DF,  p-value: < 2.2e-16
```

Figure 1a: Residuals vs. Fitted (Initial Model)



**Figure 1b: Normal Q–Q (Initial Model)**

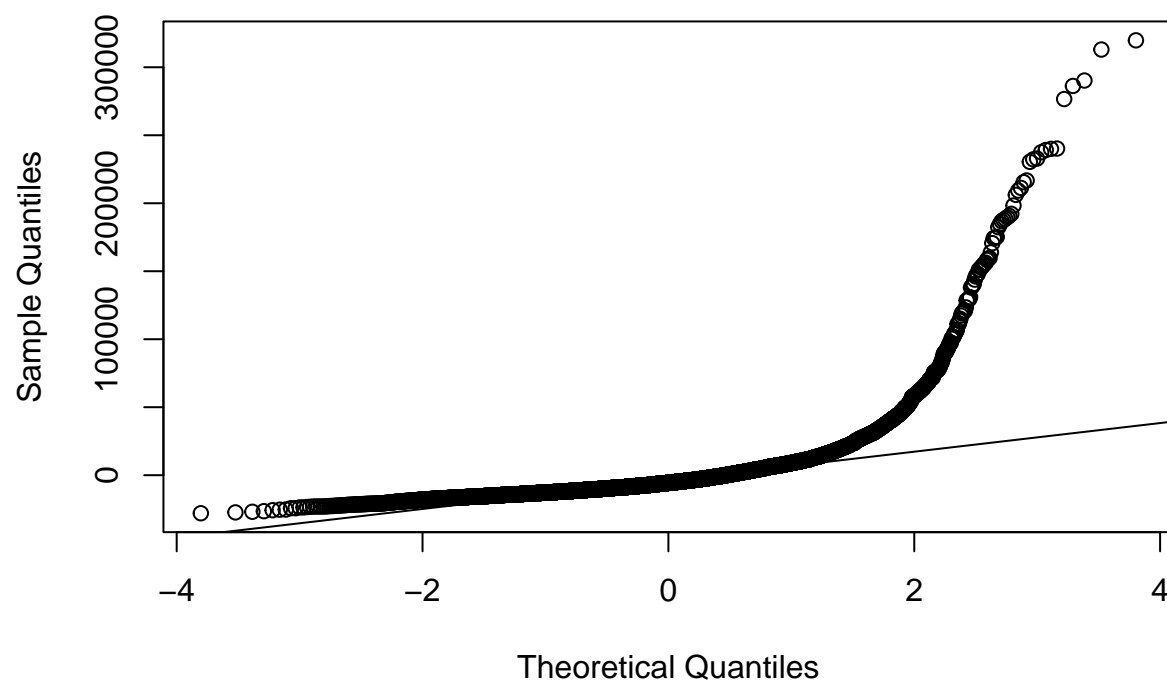


Figure 2

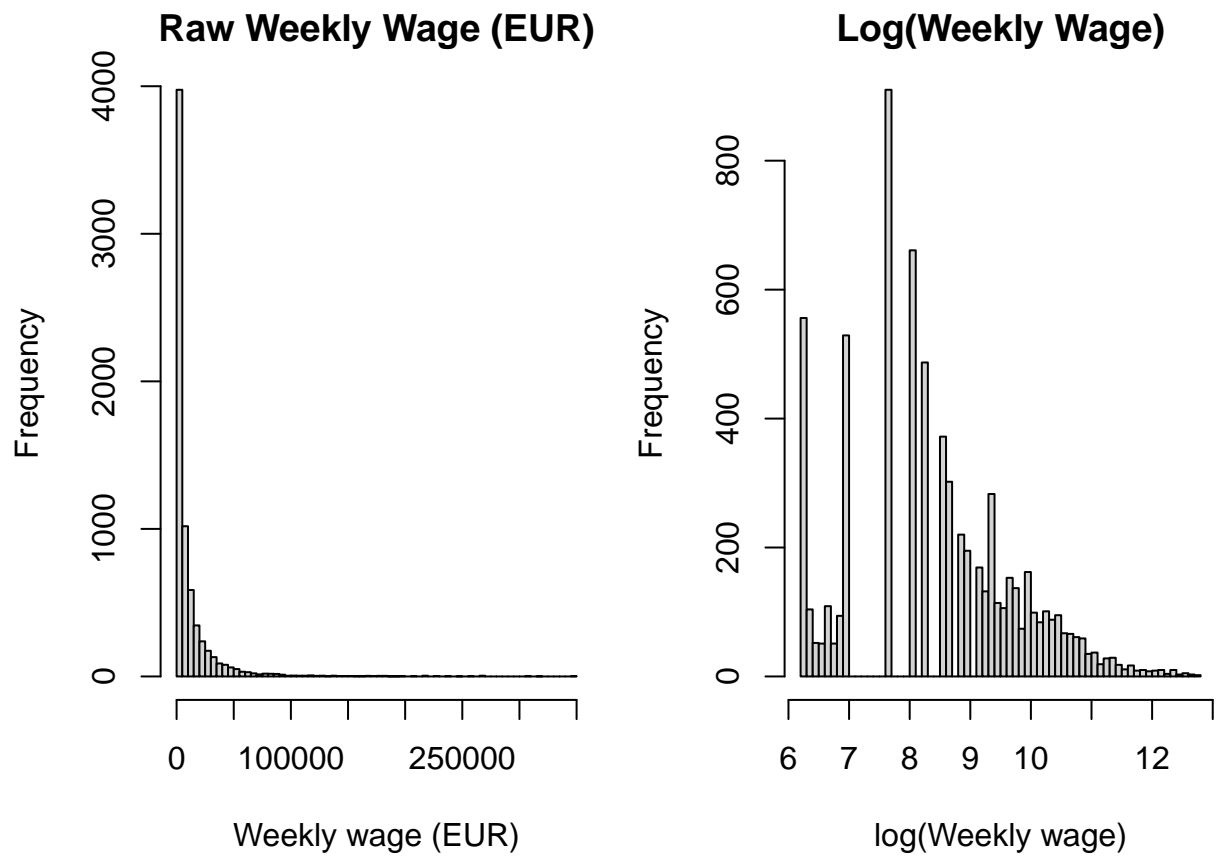
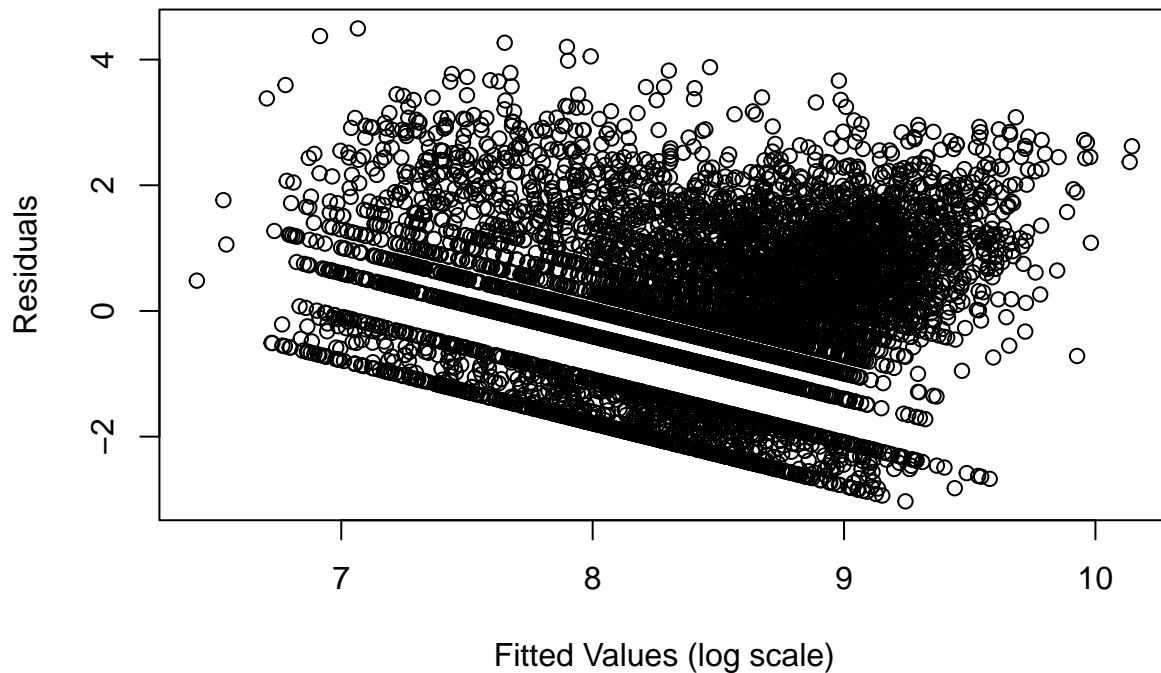


Figure 3

```
##
## Call:
## lm(formula = log(wage_eur) ~ shooting + pace + attacking_crossing +
##     preferred_foot, data = fifa)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.0296 -0.8388 -0.0283  0.8316  4.4957
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.861581   0.101031  58.017 < 2e-16 ***
## shooting        0.025907   0.001245  20.811 < 2e-16 ***
## pace          -0.004185   0.001427  -2.933 0.00336 **
## attacking_crossing 0.024307   0.001393  17.447 < 2e-16 ***
## preferred_footRight 0.088879   0.035243   2.522 0.01169 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.251 on 7005 degrees of freedom
## Multiple R-squared:  0.1835, Adjusted R-squared:  0.1831
## F-statistic: 393.7 on 4 and 7005 DF,  p-value: < 2.2e-16
```

Figure 3a: Residuals vs. Fitted (Log-Response Model)



**Figure 3b: Normal Q-Q (Log-Response Model)**

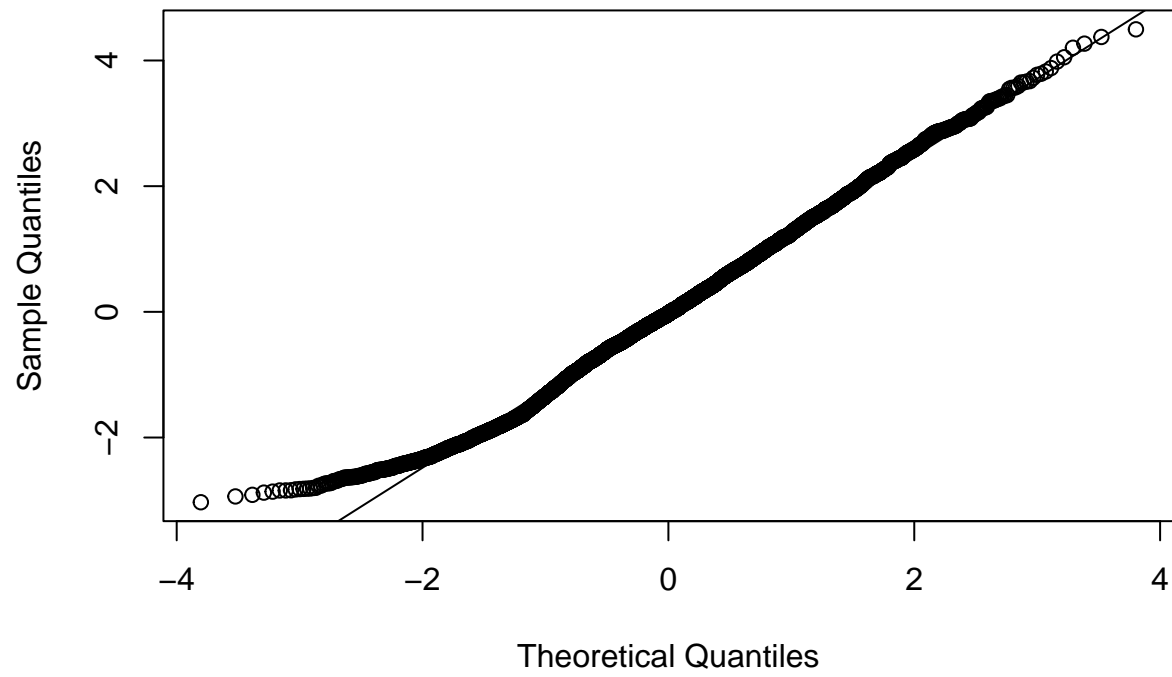


Figure 4

```
##
## Call:
## lm(formula = log(wage_eur) ~ shooting + pace + attacking_crossing *
##     preferred_foot, data = fifa)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.0032 -0.8396 -0.0262  0.8323  4.4450
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   5.431571   0.164592  33.000 < 2e-16
## shooting                      0.025915   0.001244  20.832 < 2e-16
## pace                        -0.004120   0.001426  -2.889 0.003874
## attacking_crossing            0.031247   0.002518  12.410 < 2e-16
## preferred_footRight           0.623420   0.165384   3.770 0.000165
## attacking_crossing:preferred_footRight -0.008919  0.002696  -3.308 0.000944
##
## (Intercept)                  ***
## shooting                     ***
## pace                         **
## attacking_crossing            ***
## preferred_footRight           ***
## attacking_crossing:preferred_footRight ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.25 on 7004 degrees of freedom
## Multiple R-squared:  0.1848, Adjusted R-squared:  0.1842
## F-statistic: 317.6 on 5 and 7004 DF,  p-value: < 2.2e-16
```

Interaction: Attacking Crossing × Preferred Foot

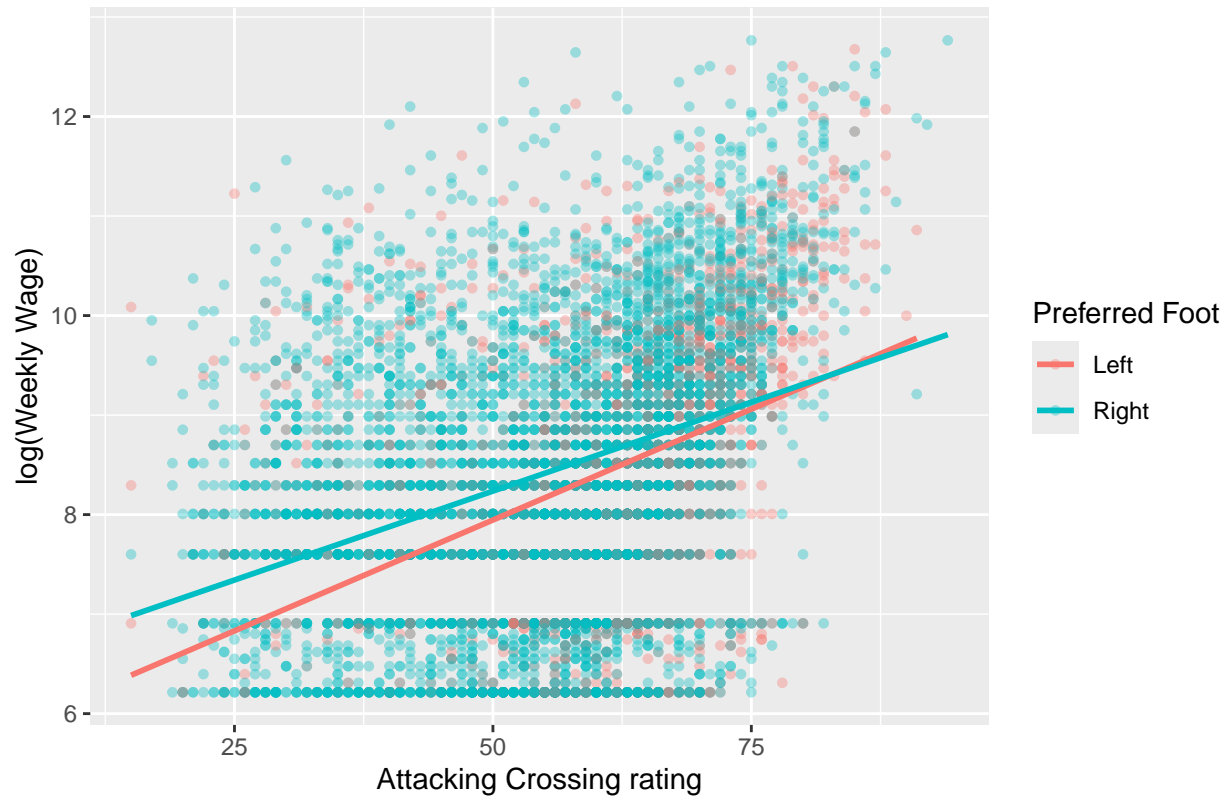
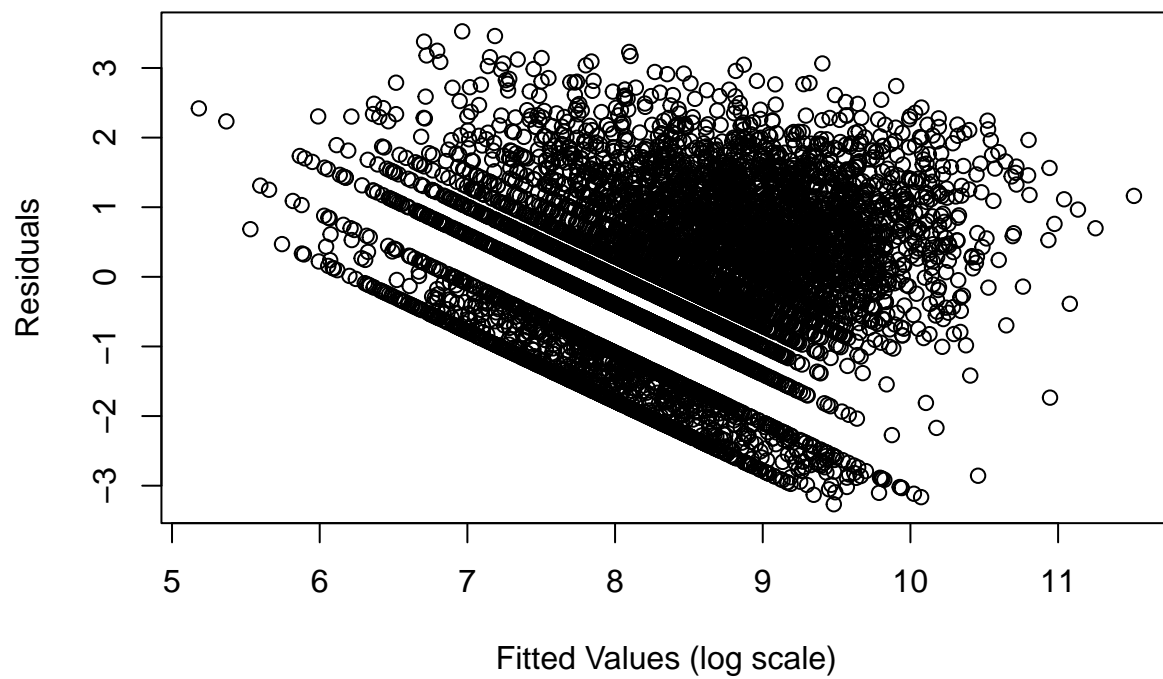




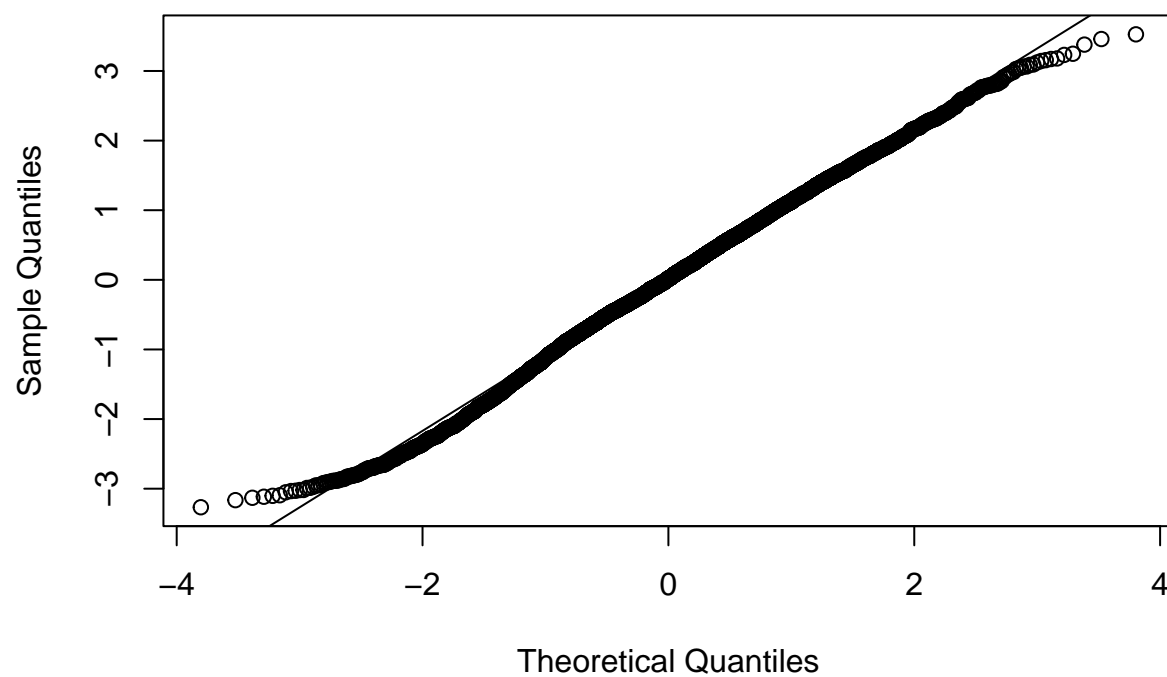
Figure 5

```
##
## Call:
## lm(formula = log(wage_eur) ~ shooting + pace + attacking_crossing *
##     preferred_foot + age + passing + dribbling, data = fifa)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.2666 -0.7133  0.0150  0.7688  3.5263
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.545599   0.200103   7.724 1.28e-14
## shooting        -0.014824   0.001538  -9.641 < 2e-16
## pace            -0.001109   0.001640  -0.676  0.4991
## attacking_crossing -0.027659   0.002754 -10.044 < 2e-16
## preferred_footRight  0.251901   0.148574   1.695  0.0900
## age              0.026724   0.003561   7.505 6.90e-14
## passing           0.075389   0.003648  20.668 < 2e-16
## dribbling         0.063243   0.003720  17.001 < 2e-16
## attacking_crossing:preferred_footRight -0.004429   0.002420  -1.830  0.0673
##
## (Intercept)          ***
## shooting              ***
## pace
## attacking_crossing    ***
## preferred_footRight   .
## age                  ***
## passing               ***
## dribbling             ***
## attacking_crossing:preferred_footRight .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.121 on 7001 degrees of freedom
## Multiple R-squared:  0.3452, Adjusted R-squared:  0.3444
## F-statistic: 461.3 on 8 and 7001 DF, p-value: < 2.2e-16
```

**Figure 5a: Residuals vs. Fitted (Extended Model)**



**Figure 5b: Normal Q-Q (Extended Model)**



**Figure 6**

##	Model	RMSE_EUR	R2_adj
## 1	Initial (raw wage)	23368.7	0.117
## 2	Log-response	24572.1	0.183
## 3	Extended + dribbling and passing	22385.7	0.344

**Figure 7**

##	shooting	pace	attacking_crossing	age
##	2.707593	2.053224	3.743577	1.259438
##	passing	dribbling		
##	6.799330	6.769001		

## Figure 8

```
## Analysis of Variance Table
##
## Model 1: log(wage_eur) ~ shooting + pace + attacking_crossing + preferred_foot
## Model 2: log(wage_eur) ~ shooting + pace + attacking_crossing * preferred_foot
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     7005 10964
## 2     7004 10946   1    17.102 10.943 0.0009444 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```