MBA6693 Assignment 1

Xinkai Zhou, 3326722

Due 11/7/2020

The file indu.csv contains historical monthly returns on the Fama French 49 industry portfolios for the U.S. market over a 72 month period. The file 5factors.csv contains historical monthly returns on the set of factor returns for the U.S. market over the same 72 month period, as well as the risk free rate. The factors in the data set are:

- RF: The risk free rate of return for that month
- Mkt: The market return
- Mkt-RF: The market excess return for that month
- SMB: The size factor return (the return of a portfolio that is long smaller stocks and short larger stocks)
- HML: The value factor return (the return of a portfolio that is long cheaper value stocks and short more expensive stocks)
- RMW: the profitability factor (the return of a portfolio that is long robust operating profitability stock and short weak operating profitability stocks)
- CMA: the investment factor (the return of a portfolio that is long conservative investment stocks and short the aggressive investment stocks).

Before start this assignment, we need to load the required data and packages:

```
setwd("D:/My Documents/GitHub/MBA6693")  #set working diretory
rm(list=ls()) #clear all variables out of the environment

fac <- read.csv("5factors.csv", header=TRUE) #load Fama French data
indu <- read.csv("indu.csv", header=TRUE) #load industry portfolios data

indu = indu - fac[,2] #compute the excess return
```

Model 1

Transform the industry returns into excess returns and then regress each industry portfolio's return on the market excess return Mkt - RF by using single linear regression model:

$$r_I = \alpha + \beta(r_m - r_f) + \epsilon$$

In this case I will only analizing the Business Service portfolio returns. Which is the 35th column.

```
colnames(indu)[35] #show Business Service column is the 35th column
```

```
## [1] "BusSv"
```

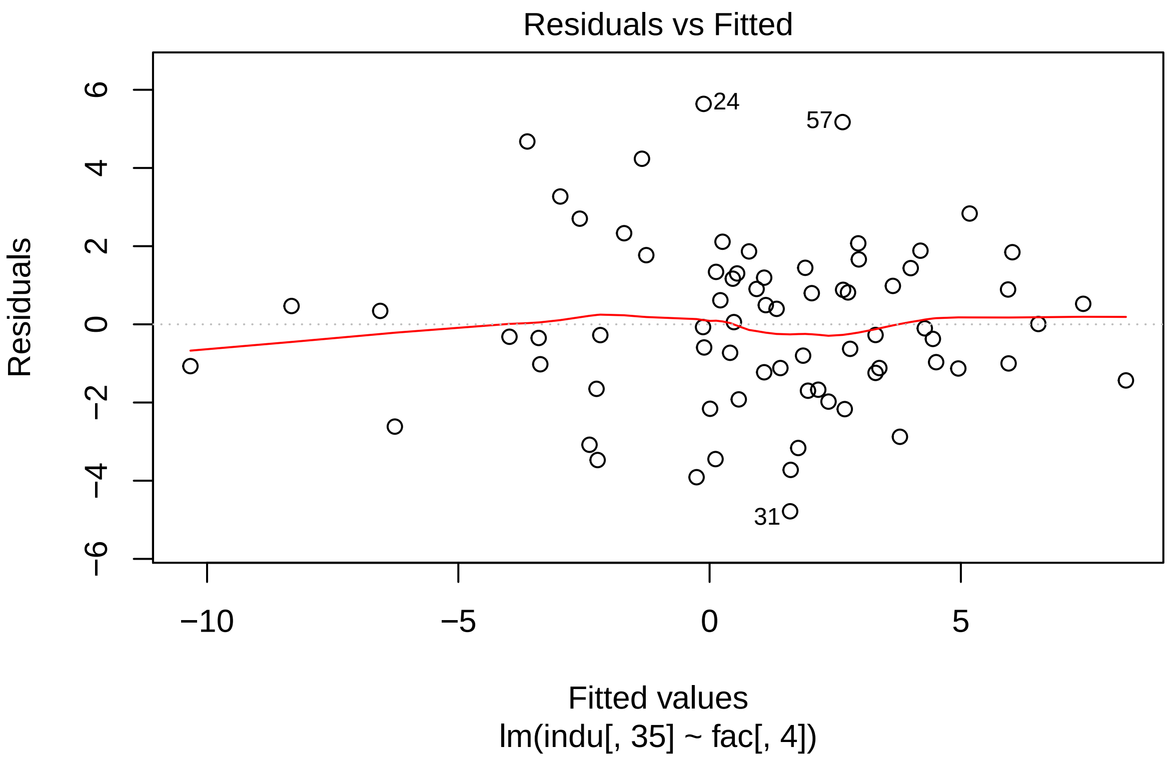
```r
colnames(fac) #show factors column names
```

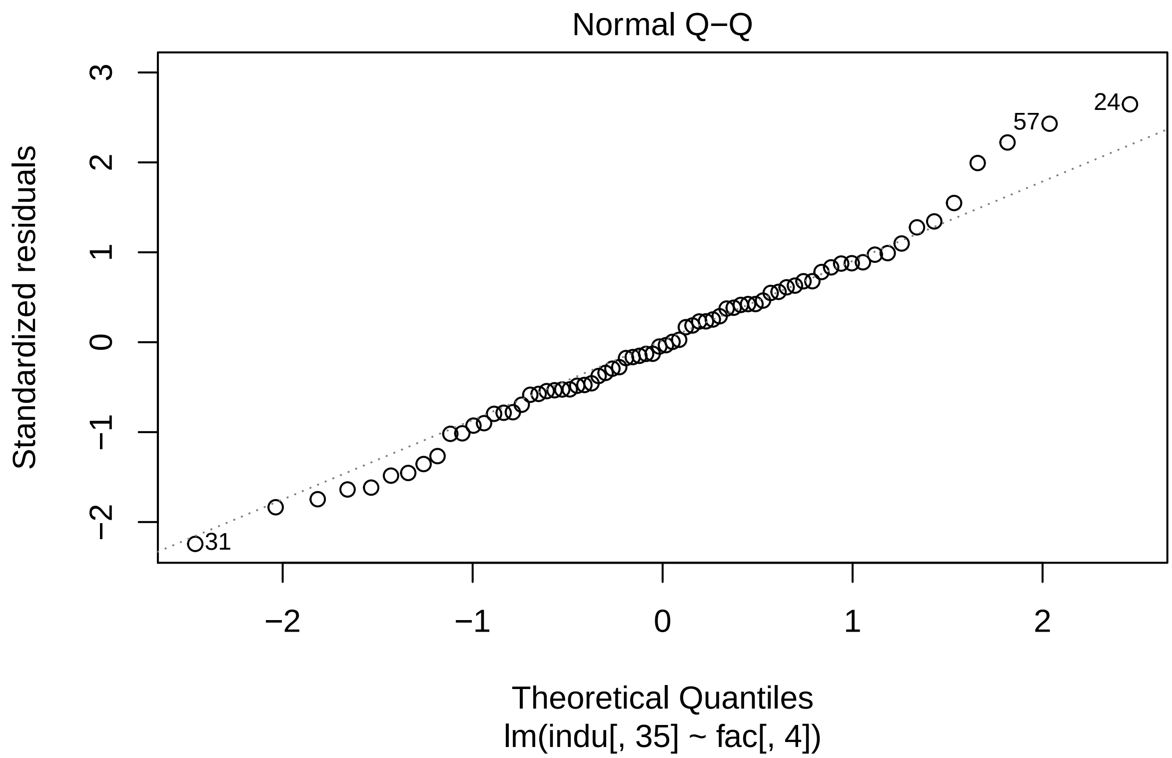
```
## [1] "Month"   "RF"      "Mkt"     "Mkt.RF" "SMB"     "HML"     "RMW"     "CMA"
```

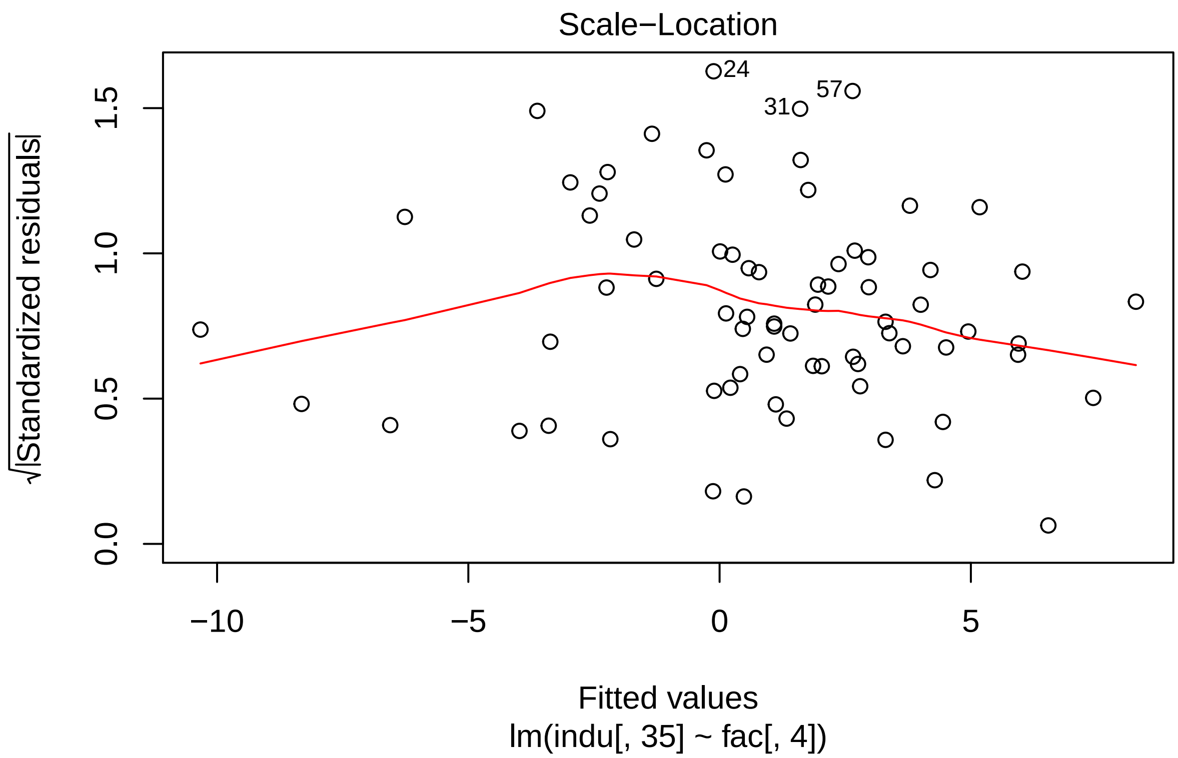
```r
m1 <- lm(indu[,35] ~ fac[,4])
summary(m1)
```

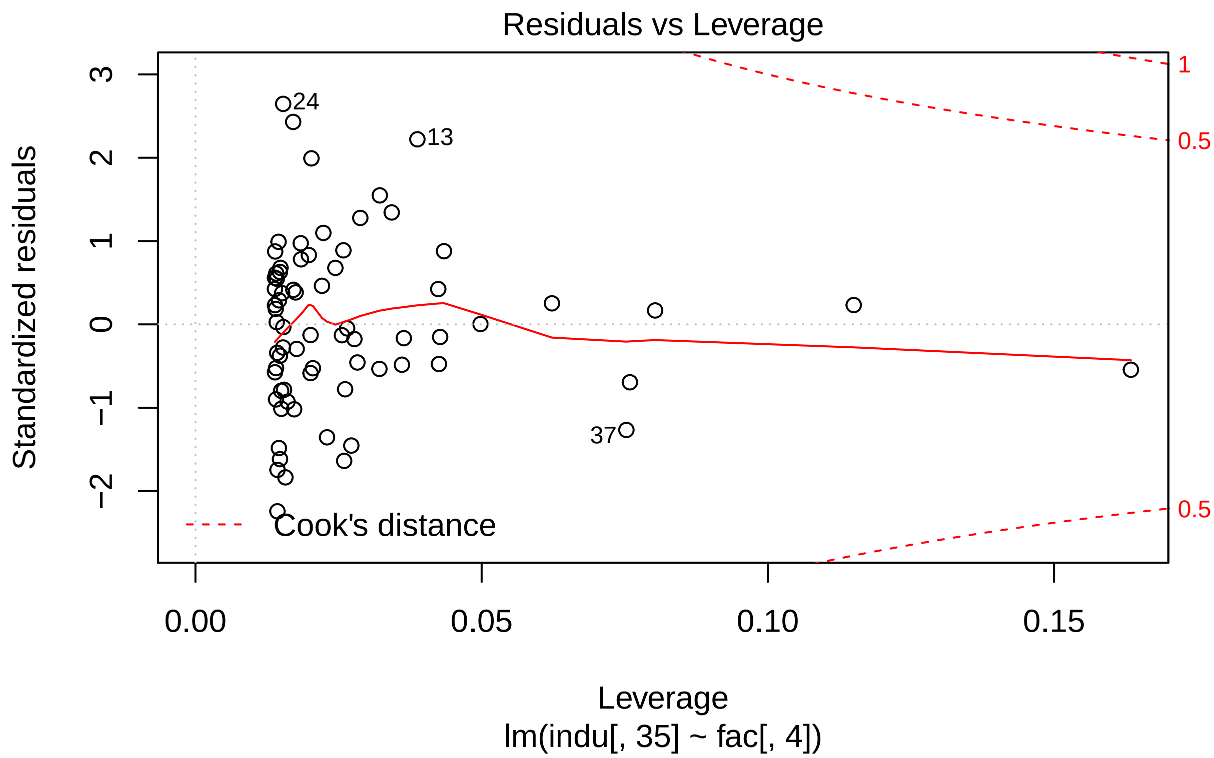
```
##
## Call:
## lm(formula = indu[, 35] ~ fac[, 4])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.7824 -1.2301 -0.0859  1.3121  5.6393
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.05473    0.26449  -0.207    0.837
## fac[, 4]     1.07607    0.07892  13.635   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.148 on 70 degrees of freedom
## Multiple R-squared:  0.7265, Adjusted R-squared:  0.7226
## F-statistic: 185.9 on 1 and 70 DF,  p-value: < 2.2e-16
```

```r
plot(m1)
```


Residuals vs Fitted
Residuals
6
4
2
0
−2
−4
−6
24
57
31
−10
−5
0
5
Fitted values
lm(indu[, 35] ~ fac[, 4])


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(indu[, 35] ~ fac[, 4])
−2
−1
0
1
2
3
24
57
31


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
24
57
31
−10
−5
0
5
Fitted values
lm(indu[, 35] ~ fac[, 4])



By using the market excess return, , the portfolio return is produced by:

$$r_I = -0.05473 + 1.07607(r_m - r_f) + \epsilon$$

Which means the portfolio returns increase by 1.08 if market excess return increase by 1.

The market excess return explains 72.26% of the variability in the portfolio returns.

Model 2

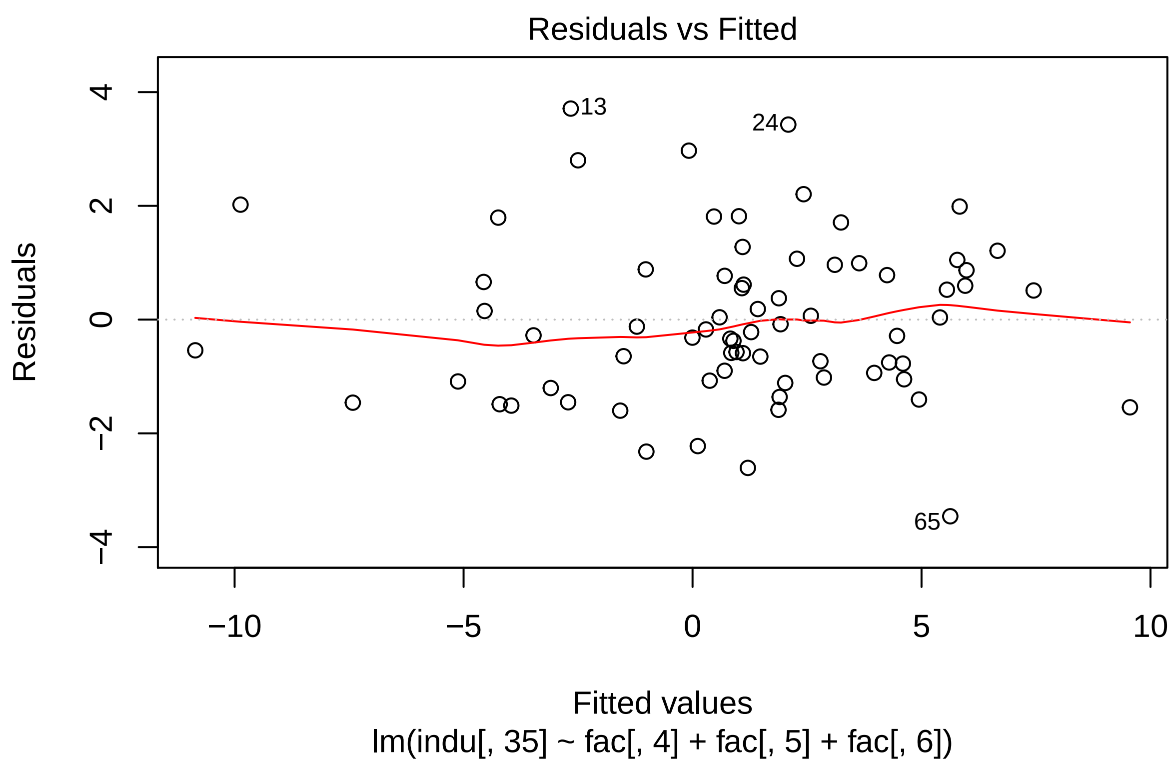
Next perform the same analysis as in before, but instead regress the industry portfolio excess returns on the three Fama French factors (Mkt-RF, SMB and HML) in a multivariate model.

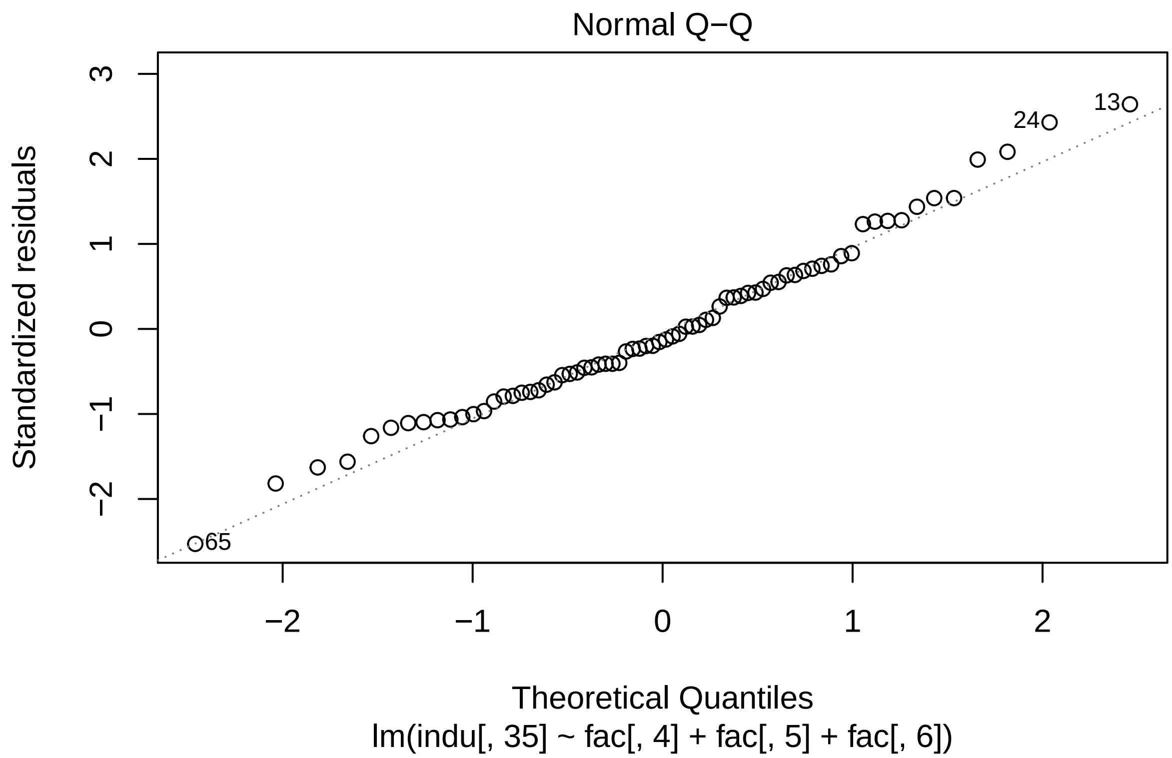
```
m2 <- lm(indu[,35] ~ fac[,4]+fac[,5]+fac[,6])
summary(m2)
```

```
##
## Call:
## lm(formula = indu[, 35] ~ fac[, 4] + fac[, 5] + fac[, 6])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4580 -1.0268 -0.1963  0.8718  3.7100
```

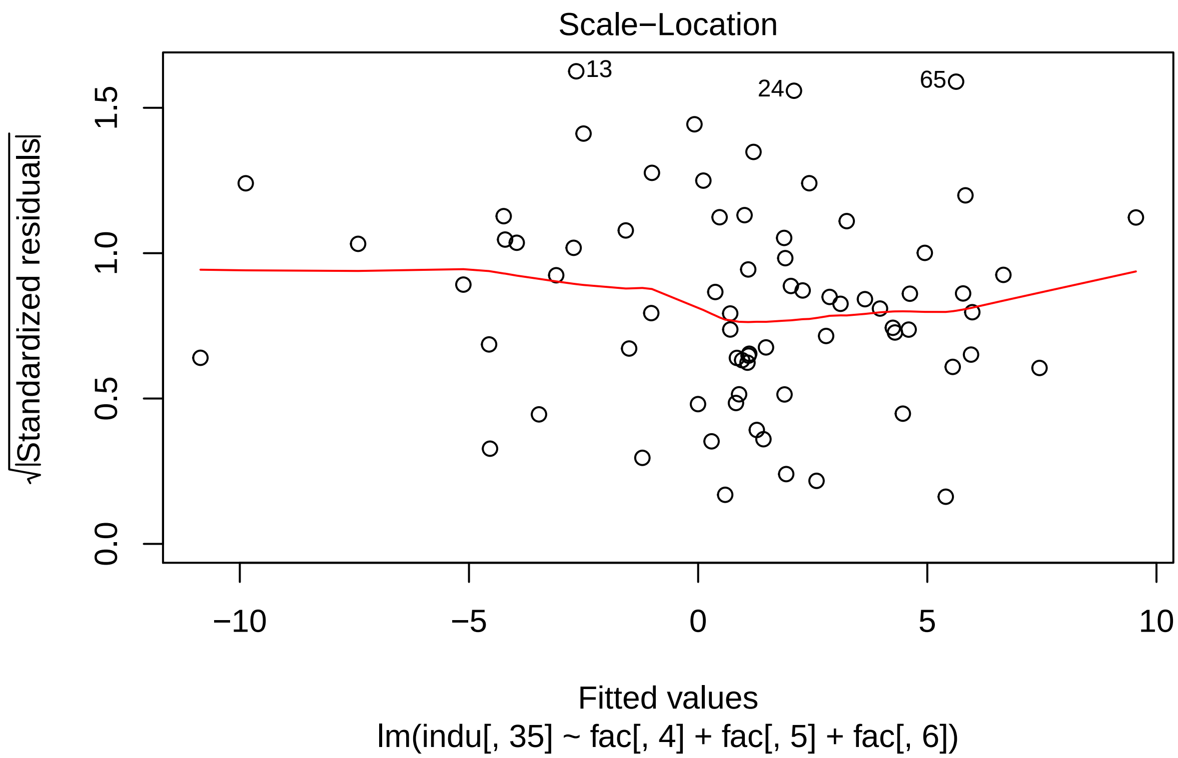
```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.15764    0.17989   0.876    0.384
## fac[, 4]     0.93614    0.05546  16.881  < 2e-16 ***
## fac[, 5]     0.66108    0.07329   9.020 3.11e-13 ***
## fac[, 6]     0.04137    0.07730   0.535    0.594
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.447 on 68 degrees of freedom
## Multiple R-squared:  0.8793, Adjusted R-squared:  0.874
## F-statistic: 165.1 on 3 and 68 DF,  p-value: < 2.2e-16
```

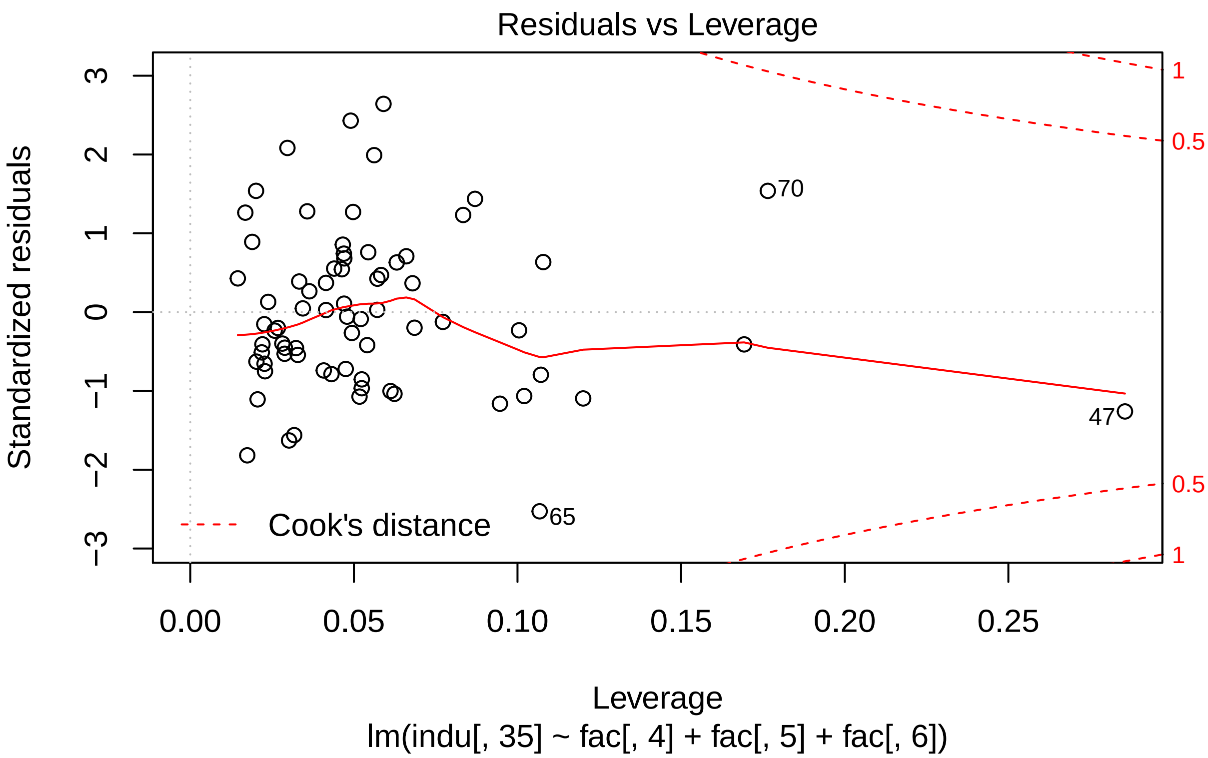
```r
plot(m2)
```





Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(indu[, 35] ~ fac[, 4] + fac[, 5] + fac[, 6])
13
24
65
−2
−1
0
1
2
3


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
13
24
65
−10
−5
0
5
10
Fitted values
lm(indu[, 35] ~ fac[, 4] + fac[, 5] + fac[, 6])



By using the market excess return, the size factor return and the value factor return, the portfolio return is produced by:

$$r_I = 0.15764 + 0.93614(r_m - r_f) + 0.66108SMB + 0.04137HML + \epsilon$$

All three factors show positive correlation with portfolio returns, the portfolio rtn increases by 0.94, 0.66, 0.04 if market excess return, size factor and value factor increases by 1.

```
summary(m2)$adj.r.squared
```

```
## [1] 0.8739841
```

The market excess return,the size factor return and the value factor return explains 87.398% of the variability in the portfolio returns.

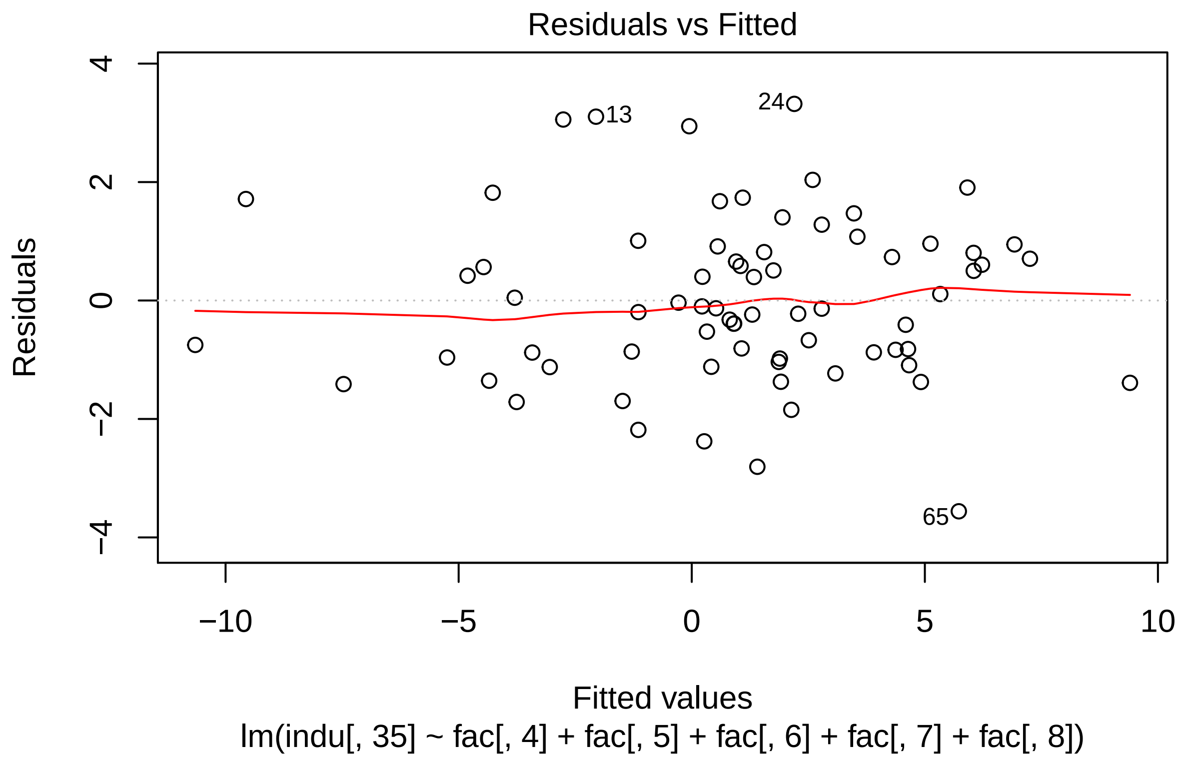
Model 3

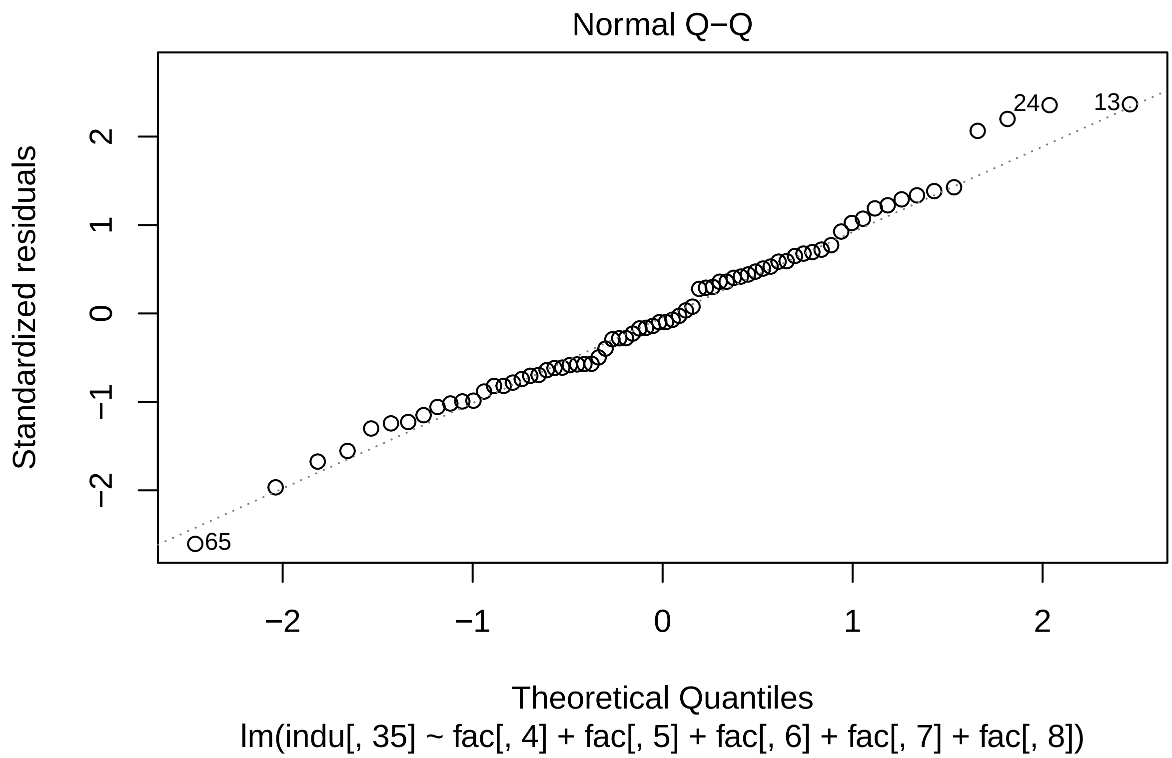
Perform the same analysis as in before, but instead regress the industry portfolio excess returns on the five Fama French factors (Mkt.RF, SMB, HML, RMW and CMA) in a multivariate model.

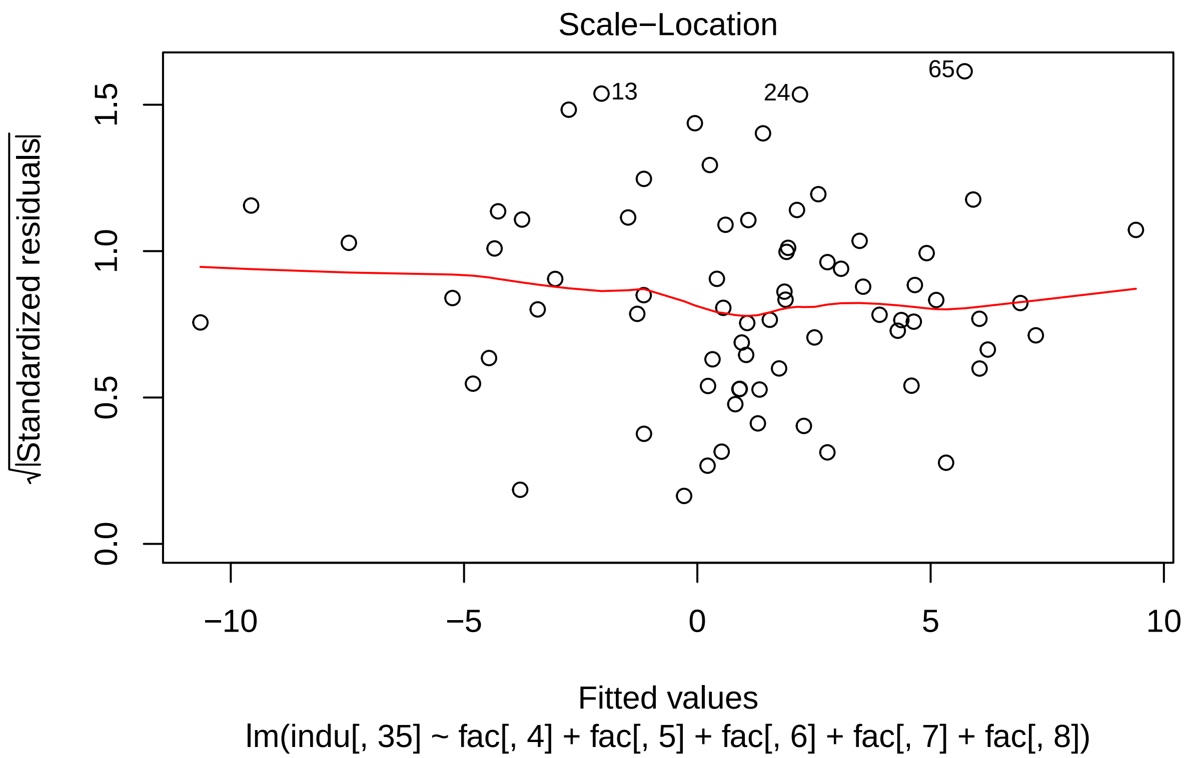
```r
m3 <- lm(indu[,35] ~ fac[,4]+fac[,5]+fac[,6]+fac[,7]+fac[,8])
summary(m3)
```

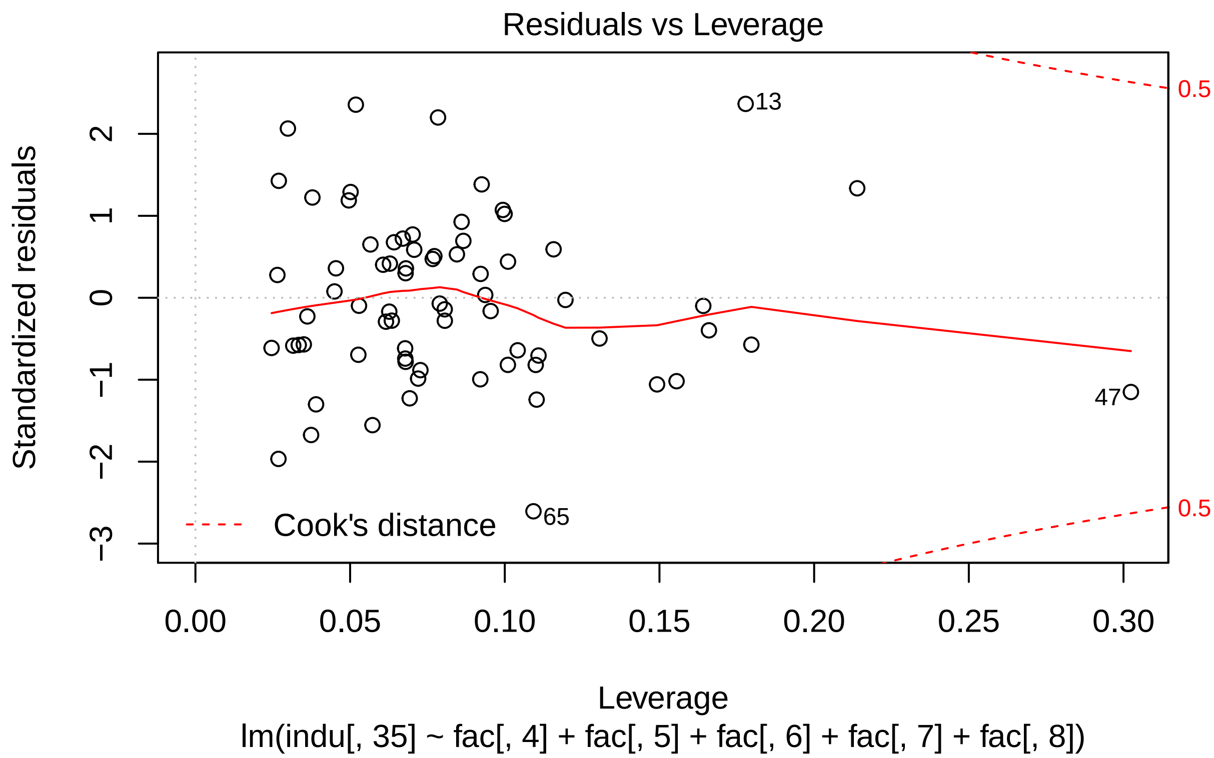
```
##
## Call:
## lm(formula = indu[, 35] ~ fac[, 4] + fac[, 5] + fac[, 6] + fac[,
##     7] + fac[, 8])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5590 -0.9668 -0.1344  0.8414  3.3200
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.172854   0.180197   0.959    0.341
## fac[, 4]     0.941237   0.056014  16.804  < 2e-16 ***
## fac[, 5]     0.612116   0.082886   7.385 3.33e-10 ***
## fac[, 6]     0.005922   0.102519   0.058    0.954
## fac[, 7]    -0.180305   0.134808  -1.337    0.186
## fac[, 8]     0.112817   0.161616   0.698    0.488
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.447 on 66 degrees of freedom
## Multiple R-squared:  0.8829, Adjusted R-squared:  0.874
## F-statistic: 99.51 on 5 and 66 DF,  p-value: < 2.2e-16
```

```r
plot(m3)
```


Residuals vs Fitted
Residuals
Fitted values
lm(indu[, 35] ~ fac[, 4] + fac[, 5] + fac[, 6] + fac[, 7] + fac[, 8])
13
24
65
4
2
0
−2
−4
−10
−5
0
5
10


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(indu[, 35] ~ fac[, 4] + fac[, 5] + fac[, 6] + fac[, 7] + fac[, 8])
24
13
65
−2
−1
0
1
2


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
−10
−5
0
5
10
13
24
65
Fitted values
lm(indu[, 35] ~ fac[, 4] + fac[, 5] + fac[, 6] + fac[, 7] + fac[, 8])



By using all Fama French factors, the portfolio return is produced by:

$$r_I = 0.172854 + 0.941237(r_m - r_f) + 0.612116SMB + 0.005922HML - 0.180305RMW + 0.112817CMA + \epsilon$$

Only the profitability factor shows negative correlation with the portfolio return, the portfolio returns decreases by 0.18 if the profitability factor increases by 1. the rest factors are positively correlated with the portfolio return,it increases by 0.94, 0.61, 0.01 and 0.11 if market excess return, size factor, value factor and investment factor increases by 1.

```
summary(m3)$adj.r.squared
```

```
## [1] 0.8740128
```

All 5 factors explains 87.401% of the variability in the portfolio returns.

Conclusion

By adding more factors, the R-Squared slightly increases, we clearly can say 5-factor model is better than 3- and 1-factor.

Compare all three models the third's residuals vs fitted plot is better, it better fitted the data. The normal QQ plot also shows the third are better fit as normal. The size of the standardized residuals is more consistent in the third model.

I would choose 5-factor models to explain the cross-section of excess returns of industry portfolios, the R-sq is slightly better than 3-factor model.