Multicollinearity

MP

2015-06-05

```
> setwd("D:/Dropbox/R/2015-NUS/Session-2/(c) Data Modelling - More Selective/Multicollinearity")
```

```
> Dataset <-
+   read.table("D:/Dropbox/R/2015-NUS/Session-2/(c) Data Modelling - More Selective/Multicollinearity/Table 5.1 Sales-and-Assets.cs
+    header=TRUE, sep=",", na.strings="NA", dec=".", strip.white=TRUE)
```

```
> library(grid, pos=14)
> library(lattice, pos=14)
> library(survival, pos=14)
> library(Formula, pos=14)
> library(ggplot2, pos=14)
> library(Hmisc, pos=14)
> rcorr.adjust(Dataset[,c("Assets","Sales")], type="pearson", use="complete")
```

```
 Pearson correlations:
       Assets  Sales
Assets 1.0000 0.9488
Sales  0.9488 1.0000

 Number of observations: 15

 Pairwise two-sided p-values:
       Assets Sales
Assets        <.0001
Sales  <.0001

 Adjusted p-values (Holm's method)
       Assets Sales
Assets        <.0001
Sales  <.0001
```

```
> RegModel.1 <- lm(Profit~Assets+Sales, data=Dataset)
> summary(RegModel.1)
```

```
Call:
lm(formula = Profit ~ Assets + Sales, data = Dataset)

Residuals:
    Min      1Q  Median      3Q     Max
-909.09  -39.15   30.83  117.77  316.85

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -88.156655 127.569747  -0.691    0.503
Assets        0.005871   0.012341   0.476    0.643
Sales         0.020070   0.020180   0.995    0.340

Residual standard error: 319.3 on 12 degrees of freedom
Multiple R-squared:  0.6385,    Adjusted R-squared:  0.5783
F-statistic:  10.6 on 2 and 12 DF,  p-value: 0.002231
```

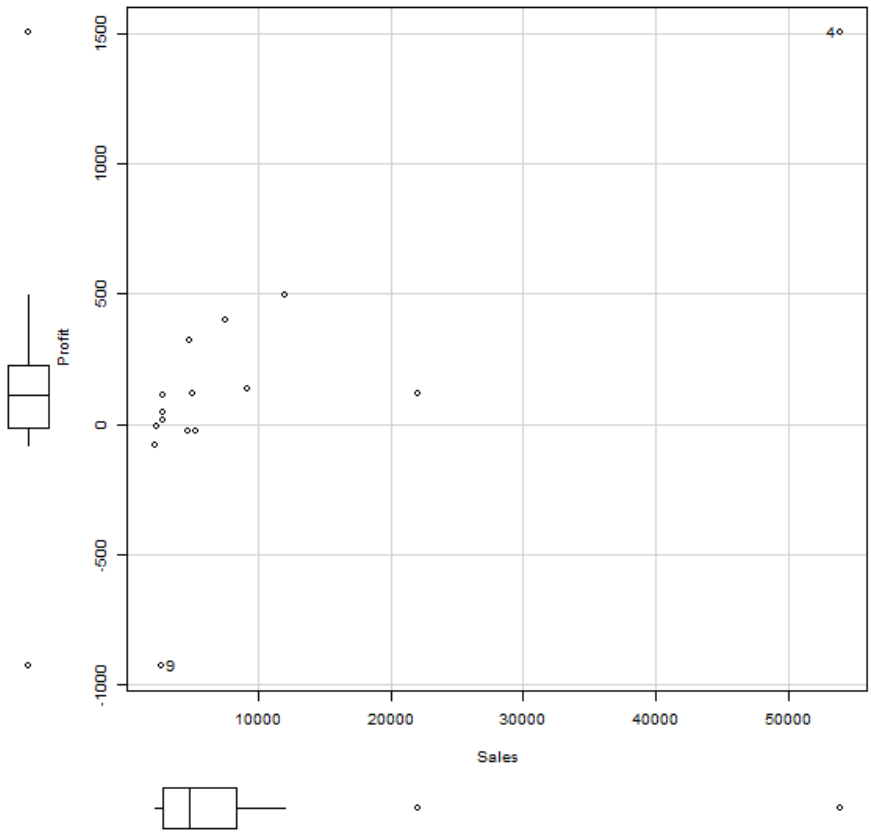
```
> RegModel.2 <- lm(Profit~Sales, data=Dataset)
> summary(RegModel.2)
```

```
Call:
lm(formula = Profit ~ Sales, data = Dataset)

Residuals:
    Min      1Q  Median      3Q     Max
-879.31  -25.10   57.02  125.48  306.25

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.249e+02  9.853e+01  -1.267 0.227313
Sales        2.918e-02  6.179e-03   4.722 0.000399 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 309.7 on 13 degrees of freedom
Multiple R-squared:  0.6317,    Adjusted R-squared:  0.6034
F-statistic:  22.3 on 1 and 13 DF,  p-value: 0.000399
```

```
> RegModel.3 <- lm(Profit~Assets, data=Dataset)
> summary(RegModel.3)
```

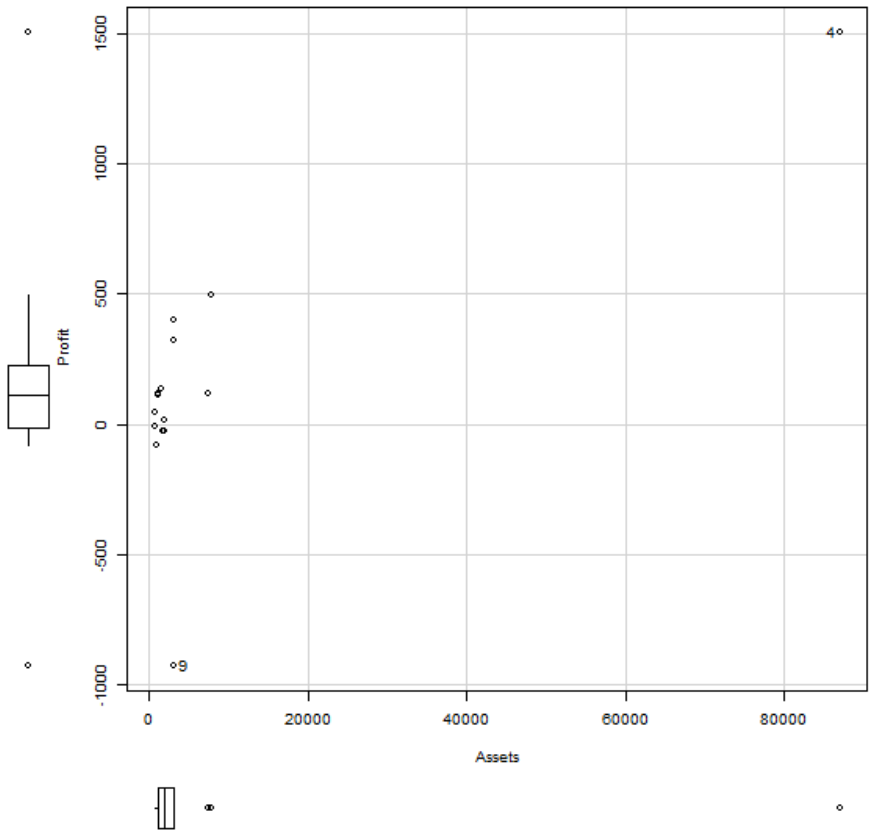
```
Call:
lm(formula = Profit ~ Assets, data = Dataset)

Residuals:
    Min      1Q  Median      3Q     Max
-981.39  -41.00  -14.78   98.88  358.04

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.291274  88.390498   0.037   0.9709
Assets       0.017517   0.003895   4.497   0.0006 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 319.2 on 13 degrees of freedom
Multiple R-squared:  0.6087,    Adjusted R-squared:  0.5786
F-statistic: 20.22 on 1 and 13 DF,  p-value: 0.0006004
```

```
> scatterplot(Profit~Sales, reg.line=FALSE, smooth=FALSE, spread=TRUE,
+   id.method='mahal', id.n = 2, boxplots='xy', span=0.5, data=Dataset)
```




```
4 9
4 9
```

```
> scatterplot(Profit~Assets, reg.line=FALSE, smooth=FALSE, spread=TRUE,
+   id.method='mahal', id.n = 2, boxplots='xy', span=0.5, data=Dataset)
```


Profit
1500
1000
500
0
-500
-1000
4
9
0
20000
40000
60000
80000
Assets

```
4 9
4 9
```