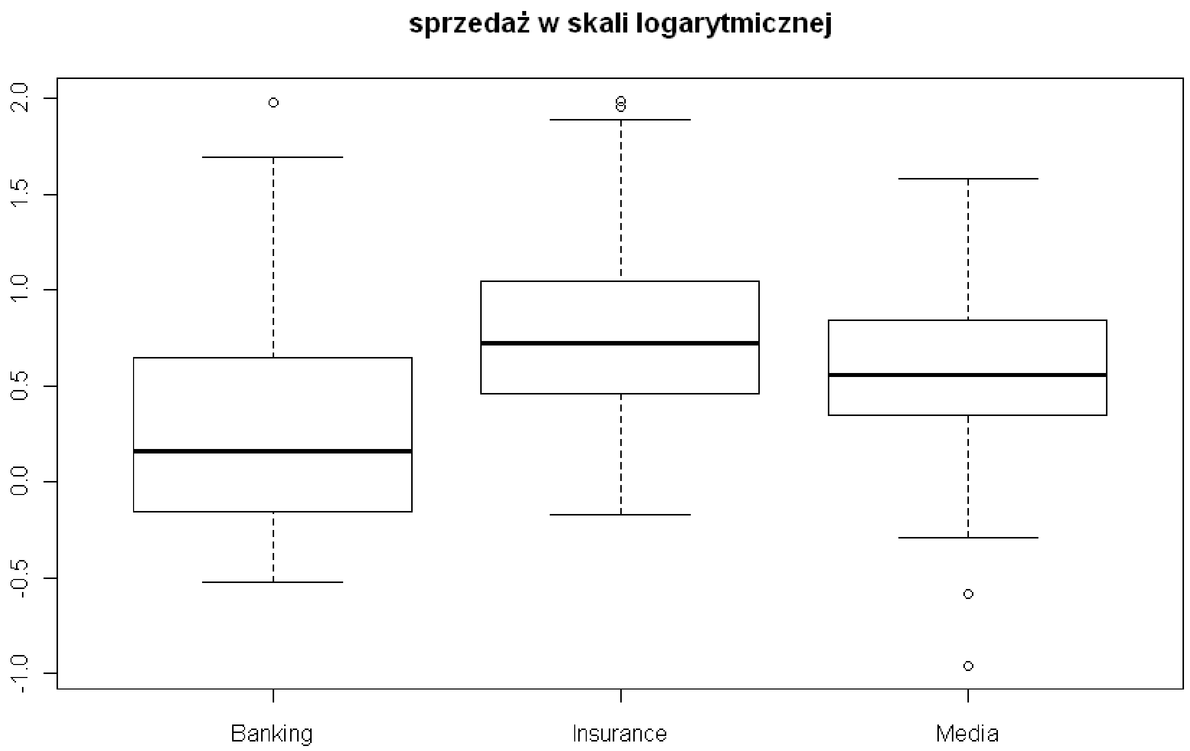
Projekt ***„Nowa oferta edukacyjna Uniwersytetu Wrocławskiego odpowiedzią na współczesne potrzeby rynku pracy i gospodarki opartej na wiedzy"***

ANALIZA PORÓWNAŃ WIELOKROTNYCH GDY WARIANCJE SĄ NIERÓWNE

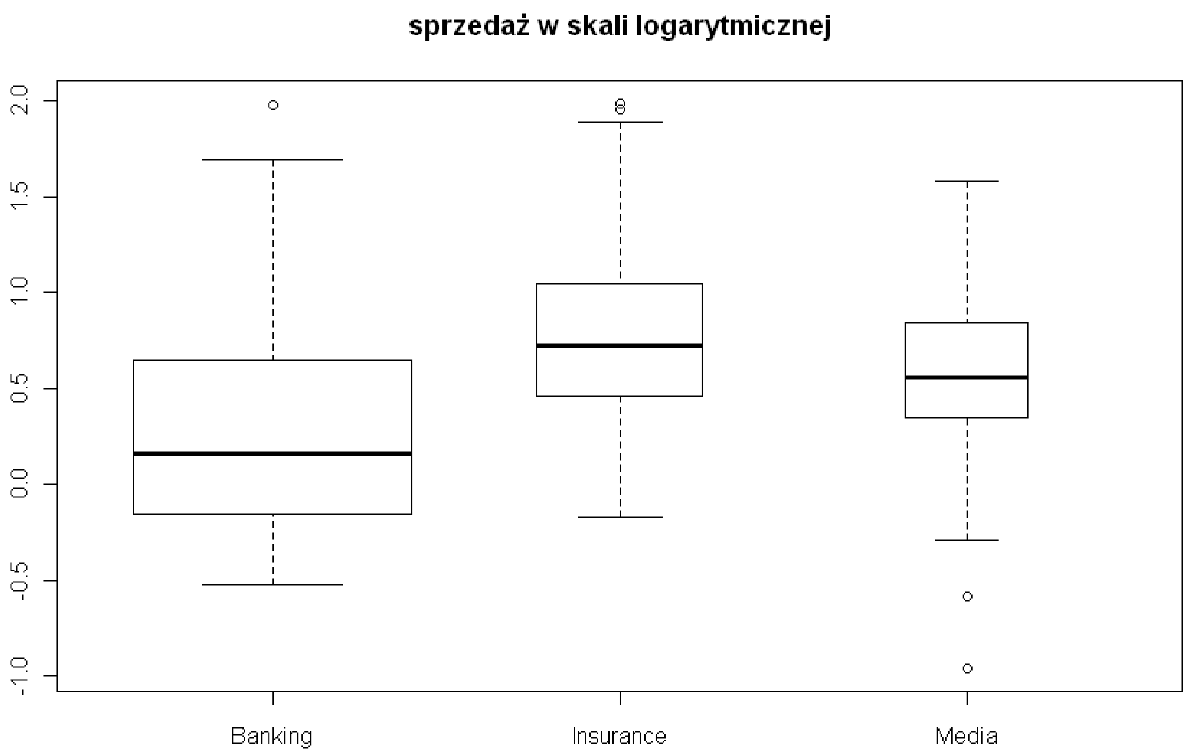
```
lsales.BIM <- log10(sales[category %in% c("Banking","Insurance","Media")])
catBIM <- as.factor(category[category %in% c("Banking","Insurance","Media"),
                             drop=T])#usunięcie zbędnych poziomów
sBIM <- data.frame(lsales.BIM,catBIM)
```

```
levels(catBIM)
[1] "Banking"   "Insurance" "Media"
```

```
boxplot(lsales.BIM~catBIM, main="sprzedaż w skali logarytmicznej")
```



```
boxplot(lsales.BIM~catBIM, main="sprzedaż w skali logarytmicznej",varwidth=T)
```




```
cat("n BIM: ",format(by(lsales.BIM, catBIM, length),digits=2))
cat("średnie BIM: ",format(by(lsales.BIM, catBIM, mean),digits=2))
cat("sd BIM: ",format(by(lsales.BIM, catBIM, sd),digits=2))
```

```
n BIM:  313 112  61
średnie BIM:  0.31 0.78 0.58
sd BIM:  0.56 0.48 0.46
```

```
library("car")
leveneTest(lsales.BIM,catBIM)
```

```
Levene's Test for Homogeneity of Variance (center = median)
       Df F value  Pr(>F)
group   2  3.6119 0.02773 *
      483
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
kruskal.test(lsales.BIM~catBIM)
```

```
        Kruskal-Wallis rank sum test

data:  lsales.BIM by catBIM
Kruskal-Wallis chi-squared = 70.2038, df = 2, p-value = 5.694e-16
```

**Poprawka Bonferroniego**

```
lsB <- split(sBIM,catBIM)[[1]]$lsales.BIM
lsI <- split(sBIM,catBIM)[[2]]$lsales.BIM
lsM <- split(sBIM,catBIM)[[3]]$lsales.BIM
```

```
t.test(lsB,lsI,conf.level=1-0.05/3)
```

```
        Welch Two Sample t-test

data:  lsB and lsI
t = -8.6395, df = 226.478, p-value = 1.028e-15
alternative hypothesis: true difference in means is not equal to 0
98.33333 percent confidence interval:
 -0.6134633 -0.3456952
sample estimates:
mean of x mean of y
0.3050799 0.7846592
```

```
t.test(lsB,lsM,conf.level=1-0.05/3)
```

```
        Welch Two Sample t-test

data:  lsB and lsM
t = -4.0471, df = 97.966, p-value = 0.0001036
alternative hypothesis: true difference in means is not equal to 0
98.33333 percent confidence interval:
 -0.4357291 -0.1083001
sample estimates:
mean of x mean of y
0.3050799 0.5770945
```

```
t.test(lsM,lsI,conf.level=1-0.05/3)
```

```
        Welch Two Sample t-test

data:  lsM and lsI
t = -2.7789, df = 127.715, p-value = 0.006277
alternative hypothesis: true difference in means is not equal to 0
98.33333 percent confidence interval:
 -0.38876160 -0.02636769
sample estimates:
mean of x mean of y
0.5770945 0.7846592
```

## ANALIZA WIELOCZYNNIKOWA

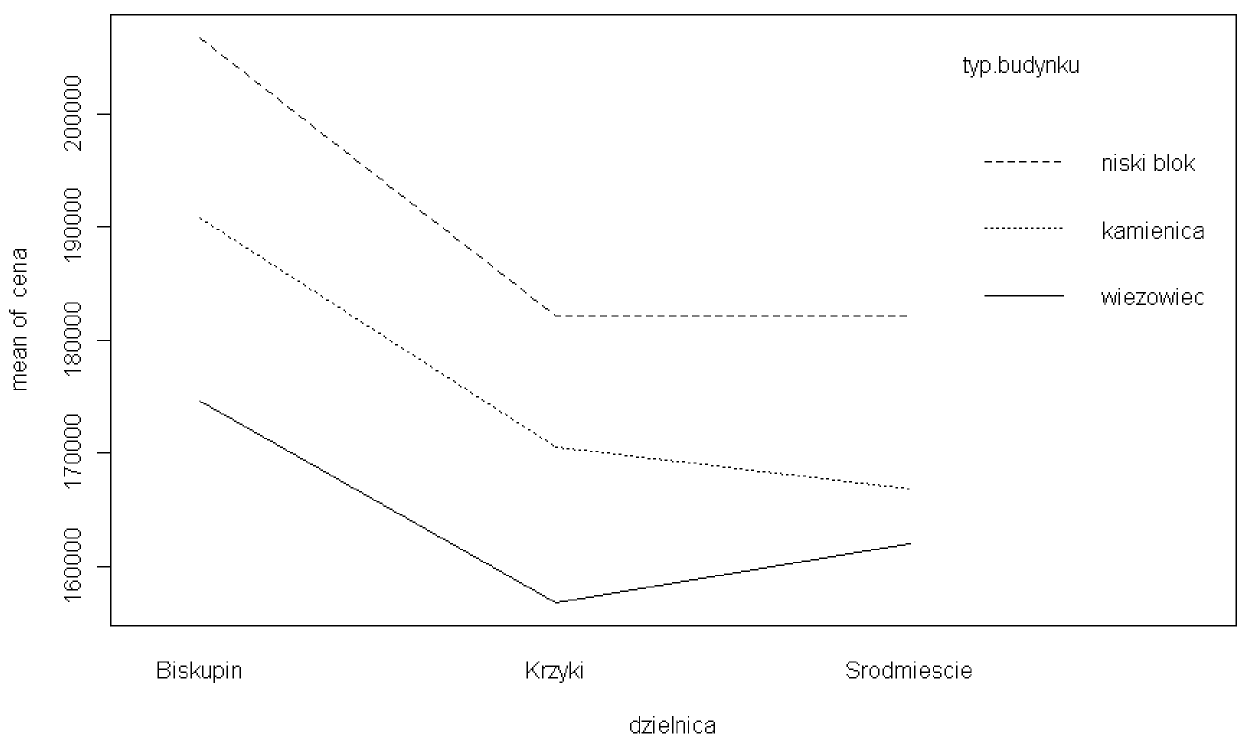
```
levene.test(cena, typ.budynku)
```

```
Levene's Test for Homogeneity of Variance (center = median)
       Df F value Pr(>F)
group   2  0.1698  0.844
      197
```

```
(anova(lm(cena ~ dzielnica * typ.budynku)))>
```

```
Analysis of Variance Table

Response: cena
                       Df     Sum Sq    Mean Sq F value   Pr(>F)
dzielnica               2 1.7995e+10 8.9977e+09  5.2461 0.006054 **
typ.budynku             2 2.2719e+10 1.1359e+10  6.6231 0.001656 **
dzielnica:typ.budynku   4 9.9528e+08 2.4882e+08  0.1451 0.964995
Residuals             191 3.2759e+11 1.7151e+09
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
interaction.plot(dzielnica,typ.budynku,cena)
```




*

```
DB.aov <- aov(lm(cena ~ dzielnica * typ.budynku))
TukeyHSD(DB.aov,"typ.budynku")
```

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = lm(cena ~ dzielnica * typ.budynku))

$typ.budynku
                          diff        lwr       upr     p adj
niski blok-kamienica  13451.96  -4120.284 31024.202 0.1695988
wiezowiec-kamienica  -12138.74 -28955.415  4677.934 0.2059238
wiezowiec-niski blok -25590.70 -42258.638 -8922.761 0.0010688
```

```
(tapply(cena/1000,typ.budynku,mean))
```

```
kamienica niski blok  wiezowiec
  178.3177   188.7909   163.3632
```

| | | | |
|---|---|---|---|
| **W** | 163,4 | * | |
| **K** | 178,3 | * | * |
| **N** | 188,8 | - | * |

{W,k}{N} lub {w} {K,N}
Dla dzielnic
{K,S} {B}

```
library("agricolae")
```

```
(DB.aovag <- aov(cena ~ dzielnica * typ.budynku))
HSD.test(DB.aovag,"typ.budynku")
```

```
Study:

HSD Test for cena

Mean Square Error:  1715123954

typ.budynku,  means

                cena  std.err replication
kamienica  178317.7 5464.928          61
niski blok 188790.9 5144.235          63
wiezowiec  163363.2 4845.965          76

alpha: 0.05 ; Df Error: 191
Critical Value of Studentized Range: 3.340556

Harmonic Mean of Cell Sizes  66.04387
Honestly Significant Difference: 17023.56

Means with the same letter are not significantly different.

Groups, Treatments and means
a        niski blok      188790.9
ab       kamienica       178317.7
 b       wiezowiec       163363.2
```

```
HSD.test(DB.aovag,"dzielnica")
```

```
Study:

HSD Test for cena

Mean Square Error:  1715123954

dzielnica,  means

                cena  std.err replication
Biskupin    189494.0 5041.724          65
Krzyki      168173.0 5078.924          79
Srodmiescie 171143.5 5302.788          56

alpha: 0.05 ; Df Error: 191
Critical Value of Studentized Range: 3.340556

Harmonic Mean of Cell Sizes  65.3595
Honestly Significant Difference: 17112.45

Means with the same letter are not significantly different.

Groups, Treatments and means
a        Biskupin        189494
 b       Srodmiescie     171143.5
 b       Krzyki          168173.0
```

```
(tapply(cena/1000,dzielnica:typ.budynku,mean))
```

```
    Biskupin:kamienica    Biskupin:niski blok     Biskupin:wiezowiec
              190.8044               206.6991               174.6505
      Krzyki:kamienica      Krzyki:niski blok       Krzyki:wiezowiec
              170.5303               182.1891               156.8233
 Srodmiescie:kamienica Srodmiescie:niski blok  Srodmiescie:wiezowiec
              166.8094               182.1547               162.0649
```

```
(cDB=rbind(c(191,207,175),c(171,182,157),c(167,182,162)))
```

```
     [,1] [,2] [,3]
[1,]  191  207  175
[2,]  171  182  157
[3,]  167  182  162
167 182 162
```

```
colnames(cDB) <- c("kam","nisbl","wiez")
rownames(cDB) <- c("Bis","Krz","Srd")
```

```
    kam nisbl wiez
Bis 191   207  175
Krz 171   182  157
Srd 167   182  162
```

```
medpolish(cDB,eps=0.001,maxiter=20)
```

```
Median Polish Results (Dataset: "cDB")

Overall: 171

Row Effects:
Bis Krz Srd
 20   0  -4

Column Effects:
  kam nisbl  wiez
    0    15   -14

Residuals:
    kam nisbl wiez
Bis   0     1   -2
Krz   0    -4    0
Srd   0     0    9
```

```
medianPolish.AD<-function(d,logg=F){
 print(d)
 if (logg==F) ld<-d else ld<-log(d)
 med.ld <- medpolish(ld)
 if (logg==F) {
 print("OGÓLNA",quote=FALSE)
 print(med.ld$overall)
 print("WIERSZE",quote=FALSE)
 print(med.ld$row,digits=3)
 print("KOLUMNY",quote=FALSE)
 print(med.ld$col,digits=3)
 print("INTERAKCJE",quote=FALSE)
 print(med.ld$residuals,digits=3)
 }
 else {
 print("OGÓLNA",quote=FALSE)
 print(exp(med.ld$overall))
 print("WIERSZE",quote=FALSE)
 print((exp(med.ld$row)-1)*100,digits=3)
 print("KOLUMNY",quote=FALSE)
 print((exp(med.ld$col)-1)*100,digits=3)
 print("INTERAKCJE",quote=FALSE)
 print((exp(med.ld$residuals)-1)*100,digits=3)
}
}
```

```
medianPolish.AD(cDB,logg=T)
```

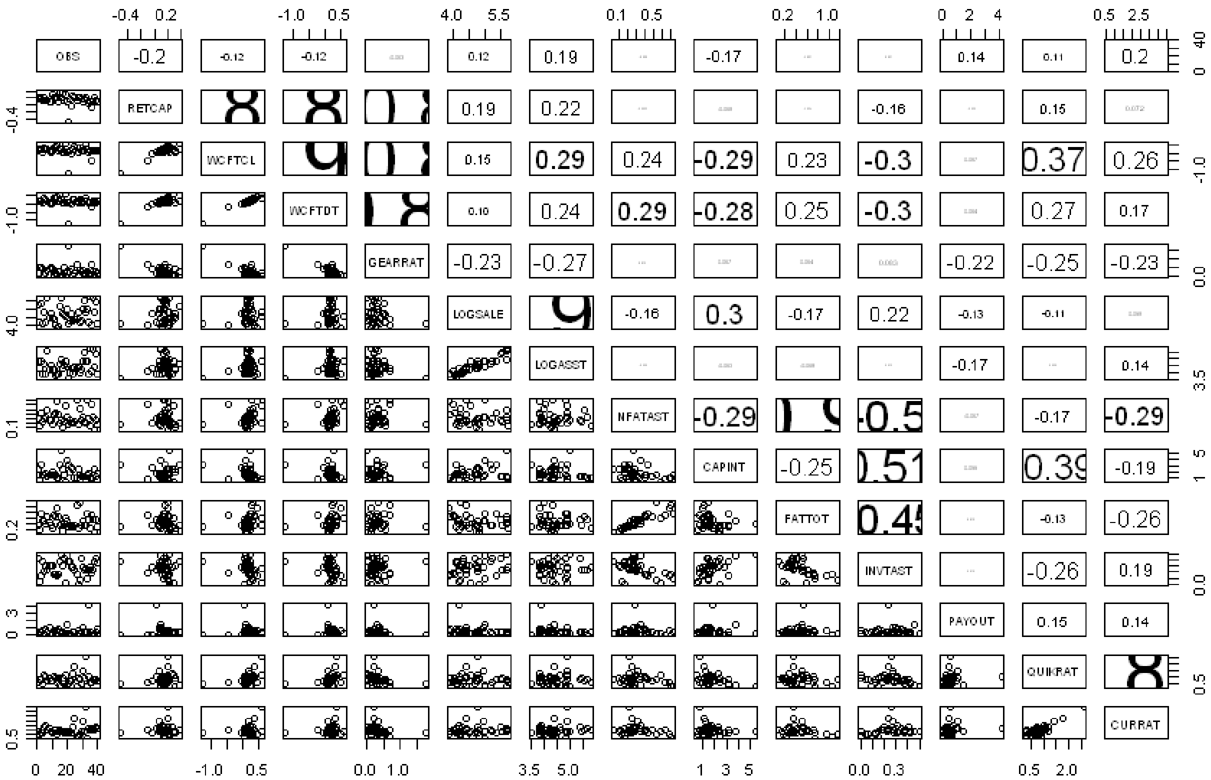
```
    kam nisbl wiez
Bis 191   207  175
Krz 171   182  157
Srd 167   182  162

[1] OGÓLNA
[1] 171

[1] WIERSZE
  Bis   Krz   Srd
11.70  0.00 -1.79
[1] KOLUMNY

 kam nisbl  wiez
 0.00  8.38 -8.19

[1] INTERAKCJE
       kam nisbl   wiez
Bis  0.000  0.00 -0.207
Krz  0.000 -1.79  0.000
Srd -0.555  0.00  5.070
```

## Regresja wielokrotna

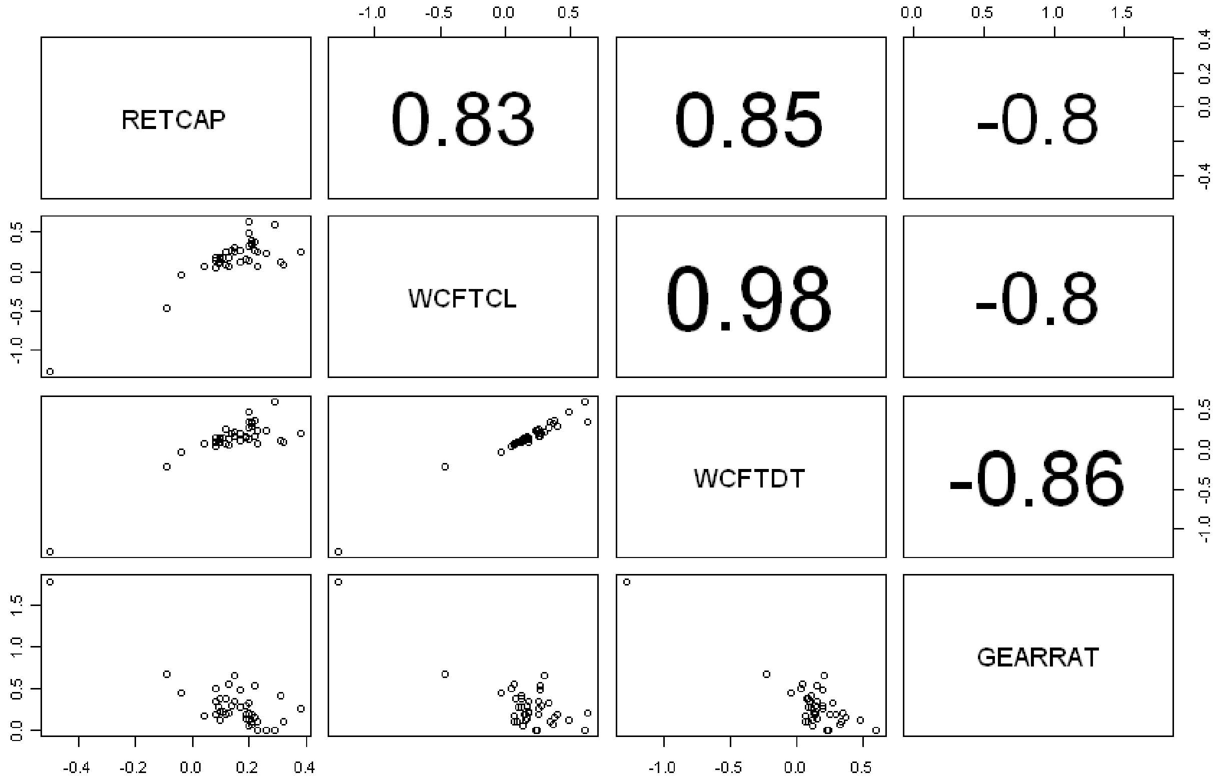
Przykład 40 przedsiębiorstw brytyjskich 1983

```
RETCAP – return of capital employed zysk z kapitału
       = zysk po opodatkowaniu kapitału

WCFTCL – ratio of working capital flow to total liabilities
= stosunek kapitału aktywnego do bieżących płatności

WCFTDT  -  ratio of working capital flow to total debt= stosunek kapitału
aktywnego do bieżącego zadłużenia

GEARRAT – gearing ratio (debt – equity ratio)= zadłużenia do kapitału w akcjach
LOGSALE logarytm sprzedaży
LOGASST logarytm aktywów (majątku)
NFATAST  ratio of net fixed assets to total assets stosunek dóbr trwałych
nieobciązonych do całkowitego majątku
CAPINT  capital intensity = stosunek sprzedaży całkowitej do całkowitego majątku
FATTOT stosunek majątku trwałego o dużej wartości do całkowitego majątku
INVTAST  stosunek zapasów do całkowitego majątki
PAYOUT  stosunek dywidend do zysków
QUIKRAT   stosunek aktywów bieżących- zapasy do bieżących płatności
CURRAT stosunek aktywów bieżących (obrotowych) do pasywów bieżących
```

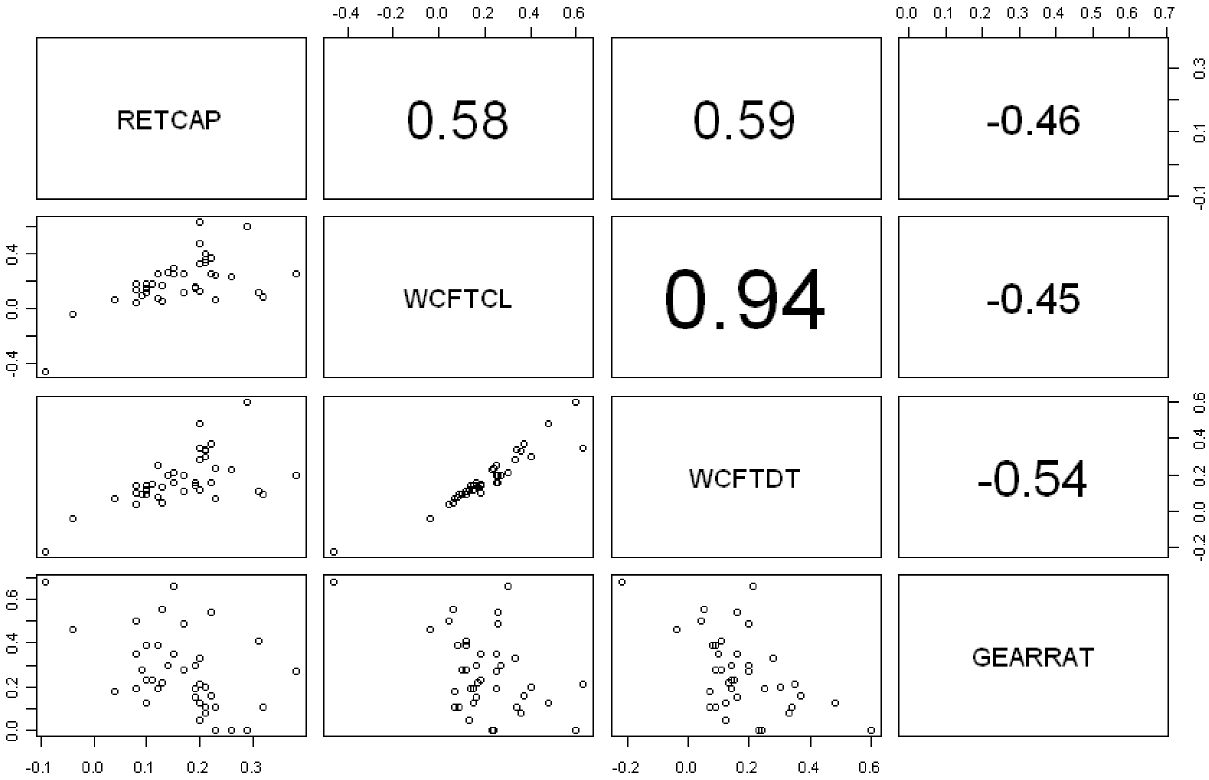
```
panel.cor <- function(x, y, ...)
{
  par(usr = c(0, 1, 0, 1))
  txt <- as.character(format(cor(x, y), digits=2))
  text(0.5, 0.5, txt,  cex = 6* abs(cor(x, y)))
}
pairs(fad, upper.panel=panel.cor)
```



```
ret <- subset(fad,select=RETCAP:GEARRAT)
pairs(ret, upper.panel=panel.cor)
```




```
ret <- subset(fad,select=RETCAP:GEARRAT,RETCAP>-0.4)
pairs(ret, upper.panel=panel.cor)
```




```
model1 <- lm(RETCAP~.,data=ret)
summary(model1)
```

```
Call:
lm(formula = RETCAP ~ ., data = ret)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.11926 -0.05137 -0.01167  0.02711  0.20516 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.14610    0.03763   3.882 0.000438 ***
WCFTCL       0.18867    0.20019   0.942 0.352416    
WCFTDT       0.07324    0.27194   0.269 0.789273    
GEARRAT     -0.12254    0.08351  -1.467 0.151195    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.07405 on 35 degrees of freedom
Multiple R-squared: 0.3937,     Adjusted R-squared: 0.3417 
F-statistic: 7.576 on 3 and 35 DF,  p-value: 0.0004952
```

**Wybór zmiennych**
Cele:

1. Prognoza zmiennej zależnej. Zalecana AIC
2. Wybór dla opisu zależności, niekoniecznie dla prognozy Zalecane BIC.
3. Chcemy dobrej estymacji parametrów VIF.

**Kryteria GIC**

$$GIC(M) = -2\log L(M|y, X) + h|M|$$

gdzie

- $M$ - model regresji
- $L$ - funkcja wiarygodności
- $h$ - współczynnik - im mniejszy tych bogatszy model preferowany. Dla AIC $h = 2$, dla BIC $h = \log(n)$
- $|M|$ - liczba parametrów modelu

Przy $n \to \infty$ w tempie $O(\log(n))$ model $M$ jest zgodny.

MODELE DLA PROGNOZY -> AIC

```
m1.step<-step(model1,trace=1)
```

```
Start:  AIC=-199.26
RETCAP ~ WCFTCL + WCFTDT + GEARRAT

          Df Sum of Sq     RSS     AIC
- WCFTDT   1 0.0003977 0.19231 -201.18
- WCFTCL   1 0.0048703 0.19678 -200.28
<none>                 0.19191 -199.26
- GEARRAT  1 0.0118066 0.20371 -198.93

Step:  AIC=-201.18
RETCAP ~ WCFTCL + GEARRAT

          Df Sum of Sq     RSS     AIC
<none>                 0.19231 -201.18
- GEARRAT  1  0.015917 0.20822 -200.08
- WCFTCL   1  0.055857 0.24816 -193.23
```

```
summary(m1.step)
```

```
Call:
lm(formula = RETCAP ~ WCFTCL + GEARRAT, data = ret)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.121252 -0.048481 -0.007882  0.025652  0.204596 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.15116    0.03219   4.697 3.78e-05 ***
WCFTCL       0.23868    0.07381   3.234  0.00262 ** 
GEARRAT     -0.13123    0.07602  -1.726  0.09289 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07309 on 36 degrees of freedom
Multiple R-squared: 0.3924,	Adjusted R-squared: 0.3587 
F-statistic: 11.63 on 2 and 36 DF,  p-value: 0.0001272
```

```
ret0 <- subset(fad,RETCAP>-0.4)
model2 <- lm(RETCAP~.,data=ret0)
```

```
m2.step<-step(model2,trace=0)
summary(m2.step)
```

```
Call:
lm(formula = RETCAP ~ WCFTDT + LOGSALE + LOGASST + NFATAST +
    INVTAST + PAYOUT + QUIKRAT, data = ret0)

Residuals:
      Min        1Q    Median        3Q       Max
-0.168485 -0.027825  0.002484  0.024422  0.135629

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.25148    0.10666   2.358  0.02488 *
WCFTDT       0.52047    0.08339   6.241 6.19e-07 ***
LOGSALE      0.10036    0.06086   1.649  0.10924
LOGASST     -0.10351    0.06174  -1.676  0.10370
NFATAST     -0.28317    0.07932  -3.570  0.00119 **
INVTAST     -0.16258    0.10506  -1.548  0.13188
PAYOUT      -0.02418    0.01547  -1.563  0.12812
QUIKRAT     -0.04725    0.02998  -1.576  0.12517
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06122 on 31 degrees of freedom
Multiple R-squared: 0.633,     Adjusted R-squared: 0.5501
F-statistic: 7.637 on 7 and 31 DF,  p-value: 2.315e-05
```

```
model3 <- lm(RETCAP~WCFTDT+NFATAST ,data=ret0)
summary(model3)
```

```
Call:
lm(formula = RETCAP ~ WCFTDT + NFATAST, data = ret0)

Residuals:
      Min        1Q    Median        3Q       Max
-0.131761 -0.035365 -0.008204  0.026340  0.180692

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.15216    0.02401   6.338 2.47e-07 ***
WCFTDT       0.46642    0.08086   5.768 1.42e-06 ***
NFATAST     -0.21967    0.06762  -3.249  0.00251 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06654 on 36 degrees of freedom
Multiple R-squared: 0.4965,    Adjusted R-squared: 0.4685
F-statistic: 17.75 on 2 and 36 DF,  p-value: 4.328e-06
```

```
anova(model2,model3)
```

```
Analysis of Variance Table

Model 1: RETCAP ~ OBS + WCFTCL + WCFTDT + GEARRAT + LOGSALE + LOGASST +
    NFATAST + CAPINT + FATTOT + INVTAST + PAYOUT + QUIKRAT +
    CURRAT
Model 2: RETCAP ~ WCFTDT + NFATAST
  Res.Df     RSS  Df Sum of Sq     F Pr(>F)
1     25 0.10957
2     36 0.15937 -11 -0.049804 1.033 0.4484
```

MODELE DLA OPISU ZALEŻNOŚCI -> BIC

```
m2.stepBIC<-step(model2,trace=0,k=log(length(RETCAP)))
summary(m2.stepBIC)
```

```
Call:
lm(formula = RETCAP ~ WCFTDT + NFATAST + QUIKRAT, data = ret0)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.127125 -0.033206 -0.006826  0.029472  0.142593 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.20573    0.03206   6.417 2.19e-07 ***
WCFTDT       0.55093    0.08419   6.544 1.49e-07 ***
NFATAST     -0.27338    0.06766  -4.041 0.000277 ***
QUIKRAT     -0.06198    0.02628  -2.358 0.024096 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.06268 on 35 degrees of freedom
Multiple R-squared: 0.5655,    Adjusted R-squared: 0.5283 
F-statistic: 15.18 on 3 and 35 DF,  p-value: 1.709e-06
```

```
m2.stepBIC<-step(model2,trace=1,k=log(length(RETCAP)))
```

```
Start:  AIC=-177.47
RETCAP ~ OBS + WCFTCL + WCFTDT + GEARRAT + LOGSALE + LOGASST + 
    NFATAST + CAPINT + FATTOT + INVTAST + PAYOUT + QUIKRAT + 
    CURRAT

          Df Sum of Sq     RSS     AIC
- GEARRAT  1 0.0000352 0.10961 -181.15
- OBS      1 0.0002007 0.10977 -181.09
- CURRAT   1 0.0003030 0.10987 -181.05

Step:  AIC=-181.15
RETCAP ~ OBS + WCFTCL + WCFTDT + LOGSALE + LOGASST + NFATAST + 
    CAPINT + FATTOT + INVTAST + PAYOUT + QUIKRAT + CURRAT

          Df Sum of Sq     RSS     AIC
- CURRAT   1 0.0002947 0.10990 -184.73
- OBS      1 0.0004080 0.11001 -184.69
- CAPINT   1 0.0006341 0.11024 -184.61

….
Step:  AIC=-203.65
RETCAP ~ WCFTDT + NFATAST + PAYOUT + QUIKRAT

          Df Sum of Sq     RSS     AIC
- PAYOUT   1  0.006354 0.13753 -205.50
<none>                 0.13117 -203.65
- QUIKRAT  1  0.017494 0.14867 -202.46
```

## INFLACJA WARIANCJI

$$V(b_j^*) = \frac{\frac{\sigma^2}{n-1}}{(1-R_j^2)s_{x_j}^2} = \frac{\frac{\sigma^2}{n-1}}{s_{x_j}^2}V_j$$

$$d^2(b^*,\beta^*) = \sigma^2 \sum_{j=1}^{p} V_j s_{x_j}^2$$

10 - Wartość progowa dla VIF

```
library("car")
vif(lm(RETCAP~.,data=ret0))
```

```
      OBS    WCFTCL    WCFTDT   GEARRAT   LOGSALE   LOGASST   NFATAST    CAPINT
 1.782154 22.745876 36.946340  5.772543 99.363725 84.294839  7.808239 11.222472
   FATTOT   INVTAST    PAYOUT   QUIKRAT    CURRAT
 6.462455  7.475291  1.446606 22.655241 20.090718
```

```
vif(lm(RETCAP~.-LOGSALE,data=ret0))
```

```
      OBS    WCFTCL    WCFTDT   GEARRAT   LOGASST   NFATAST    CAPINT    FATTOT
 1.684443 19.598095 28.870444  4.723694  1.556263  7.697887  1.964764  5.777506
  INVTAST    PAYOUT   QUIKRAT    CURRAT
 7.161620  1.413401 22.358857 19.428026
```

```
vif(lm(RETCAP~.- OBS-LOGSALE-WCFTDT,data=ret0))
```

```
   WCFTCL   GEARRAT   LOGASST   NFATAST    CAPINT    FATTOT   INVTAST    PAYOUT
 2.159916  1.822016  1.254053  6.692956  1.695955  5.641044  7.044536  1.228259
  QUIKRAT    CURRAT
21.756643 18.560793
```

```
vif(lm(RETCAP~.- OBS-LOGSALE-WCFTDT-QUIKRAT,data=ret0))
```

```
  WCFTCL  GEARRAT  LOGASST  NFATAST   CAPINT   FATTOT  INVTAST   PAYOUT   CURRAT
2.026018 1.821984 1.093693 5.439908 1.689753 5.557206 1.932474 1.225436 1.628328
```

```
model4 <- lm(RETCAP~.- OBS-LOGSALE-WCFTDT-QUIKRAT,data=ret0)
summary(model4)
```

```
Call:
lm(formula = RETCAP ~ . - OBS - LOGSALE - WCFTDT - QUIKRAT, data = ret0)

Residuals:
      Min        1Q    Median        3Q       Max
-0.179248 -0.023350 -0.003953  0.017935  0.154733

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.32836    0.10987   2.989 0.005656 **
WCFTCL       0.36385    0.08134   4.473 0.000109 ***
GEARRAT     -0.06406    0.07944  -0.806 0.426593
LOGASST     -0.01872    0.02012  -0.931 0.359700
NFATAST     -0.10745    0.14261  -0.753 0.457251
CAPINT       0.01671    0.01406   1.189 0.244122
FATTOT      -0.08201    0.09940  -0.825 0.416030
INVTAST     -0.04971    0.11410  -0.436 0.666308
PAYOUT      -0.02807    0.01670  -1.680 0.103617
CURRAT      -0.04536    0.02252  -2.015 0.053312 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06346 on 29 degrees of freedom
Multiple R-squared: 0.6311,    Adjusted R-squared: 0.5166
F-statistic: 5.512 on 9 and 29 DF,  p-value: 0.0001975
```

```
m4.step<-step(model4,trace=0)
summary(m4.step)
```

```
Call:
lm(formula = RETCAP ~ WCFTCL + CAPINT + FATTOT + CURRAT, data = ret0)

Residuals:
      Min        1Q    Median        3Q       Max
-0.161659 -0.028978 -0.007494  0.021926  0.148380

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.20465    0.05053   4.050 0.000280 ***
WCFTCL       0.41225    0.06250   6.596 1.47e-07 ***
CAPINT       0.01551    0.01128   1.376 0.177925
FATTOT      -0.15888    0.04745  -3.349 0.001996 **
CURRAT      -0.05310    0.01999  -2.656 0.011954 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0622 on 34 degrees of freedom
Multiple R-squared: 0.5844,    Adjusted R-squared: 0.5355
F-statistic: 11.95 on 4 and 34 DF,  p-value: 3.602e-06
```

```
m4.stepBIC<-step(model4,trace=0,k=length(RETCAP))
summary(m4.stepBIC)
```

```
Call:
lm(formula = RETCAP ~ 1, data = ret0)

Residuals:
      Min        1Q    Median        3Q       Max
-0.253846 -0.058846  0.006154  0.046154  0.216154

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.16385    0.01461   11.21 1.30e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.09127 on 38 degrees of freedom
```

## WSPÓŁCZYNNIKI STANDARYZOWANE (bety)

W przypadku jednej zmiennej beta = korelacja między y a x

```
m4.step$coeff["WCFTCL"]
m4.step$coeff["FATTOT"]
```

```
 0.412251
-0.1588808
```

```
with(ret0,
     m4.step$coeff["WCFTCL"] * sd(WCFTCL) / sd(RETCAP)
)
```

```
0.8136877
```

```
with(ret0,
     m4.step$coeff["FATTOT"] * sd(FATTOT) / sd(RETCAP)
)
```

```
-0.4250227
```

```
library("QuantPsyc")
lm.beta(m4.step)
```

```
    WCFTCL     CAPINT     FATTOT     CURRAT
 0.8136877  0.1617887 -0.4250227 -0.3393619
```

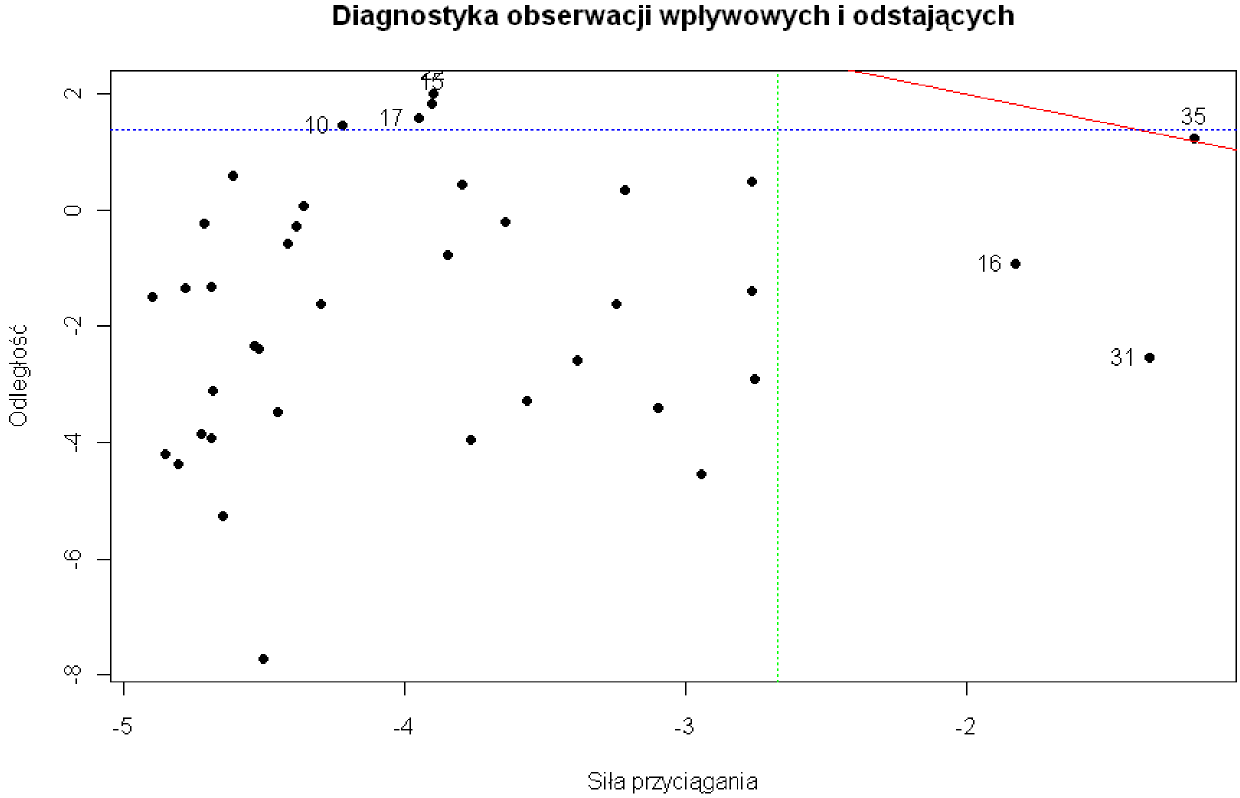
```
lm.beta(model3)
```

```
    WCFTDT    NFATAST
 0.7195529 -0.4052404
```

WYKRES DIAGNOSTYCZNY

```
# Wykres pokazuje położenie punktów obserwacyjnych względem zmiennych
# diagnostycznych: siła przyciągania oraz odległości standaryzowanej
#(obie w skali logarytmicznej).
# Linie diagnostyczne:
#   pionowa, zielona wskazuje na punkty wpływowe regresji (na prawo od linii)
#   pozioma, niebieska wskazuje na punkty dalekie od regresji (w górę od linii)
#   ukośna, czerwona, wskazuje na punkty o dużej odl. Cooka (w górę od linii)
# Punkty wybiera się klikając myszką. Po zakończeniu (ESC) na wykresie
#pojawią się etykiety punktów a na konsoli ich opis.
# Autor: Andrzej Dąbrowski, Agata Zawadzka
```

```
wykresDiagnostyczny(m4.step)
```




{diagnostyka1 [5]}

```
[1] Numery wyróżnionych punktów:
[1] 10 15 16 17 22 31 35
[1] Siła przyciągania:
       10        15        16 
-4.217198 -3.900918 -1.825000 
       17        23        32 
-3.949107 -3.897516 -1.348618 
       36 
-1.190532 
[1] Odległość:
        10         15         16 
 1.4644666  1.8350964 -0.9275395 
        17         23         32 
 1.5833804  2.0068318 -2.5286044 
        36 
 1.2400804 
```

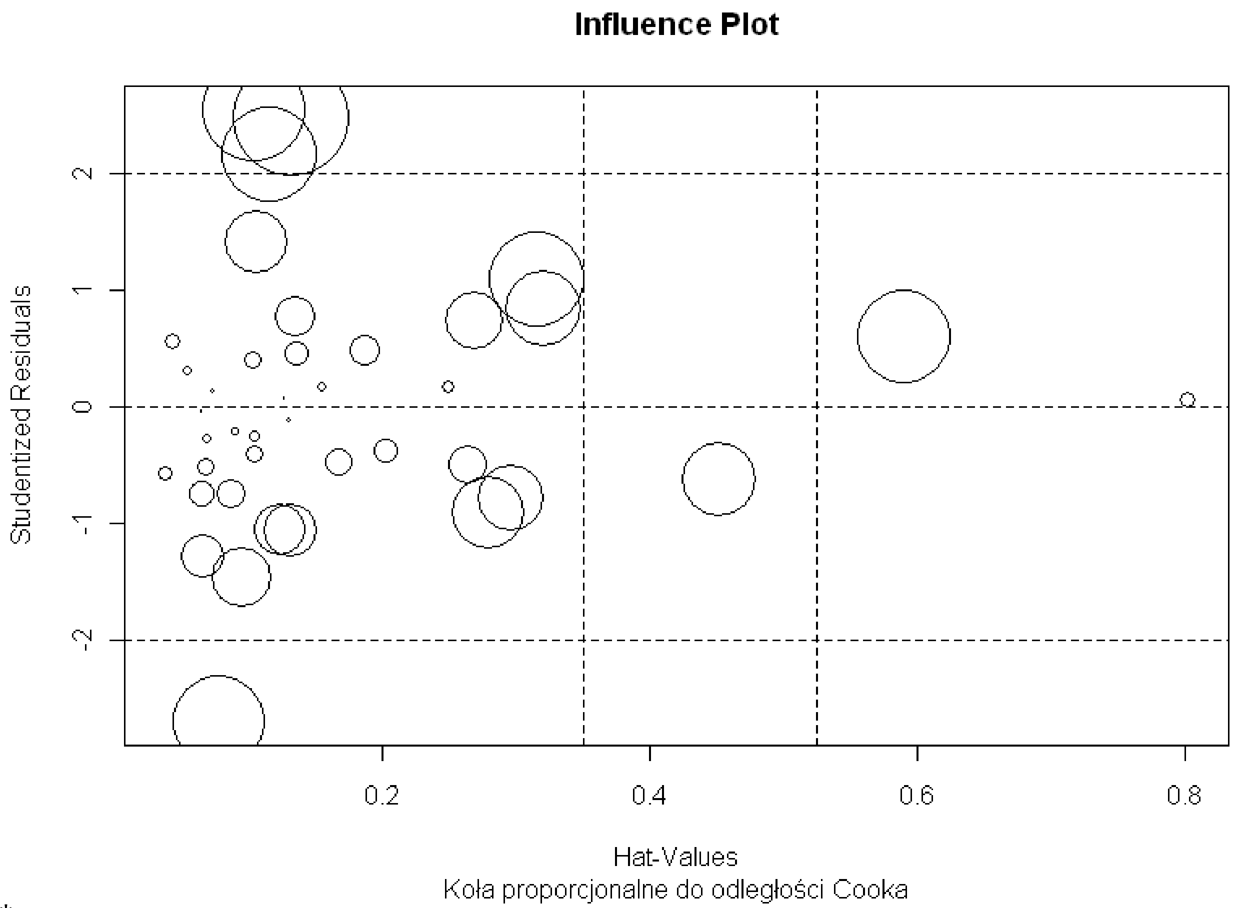
|  | WCFTCL | CAPINT | FATTOT | CURRAT |
|---|---|---|---|---|
| 36 | -0.46 | 0.77 | 0.71 | 1.45 |
| 32 | 0.63 | 0.81 | 0.34 | 3.98 |
| 16 | 0.24 | 5.44 | 0.38 | 1.29 |
| 23 | -0.04 | 2.79 | 0.32 | 1.11 |

```
ret04.1<-subset(ret0,select=c(RETCAP,WCFTCL,CAPINT,FATTOT,CURRAT),
                ! rownames(ret0) %in% c("36" ))
m41<- lm(RETCAP~.,data=ret04.1)
summary(m41)
lm.beta(m41)
```

```
Call:
lm(formula = RETCAP ~ ., data = ret04.1)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.14262 -0.02681 -0.00062  0.02273  0.13133 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.21088    0.04872   4.328 0.000131 ***
WCFTCL       0.54694    0.09206   5.941 1.15e-06 ***
CAPINT       0.02068    0.01117   1.851 0.073182 .  
FATTOT      -0.20747    0.05212  -3.981 0.000355 ***
CURRAT      -0.06999    0.02113  -3.313 0.002248 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.05984 on 33 degrees of freedom
Multiple R-squared: 0.528,	Adjusted R-squared: 0.4708 
F-statistic: 9.229 on 4 and 33 DF,  p-value: 4.047e-05

    WCFTCL     CAPINT     FATTOT     CURRAT 
 0.9692862  0.2388338 -0.6158348 -0.5029347
```

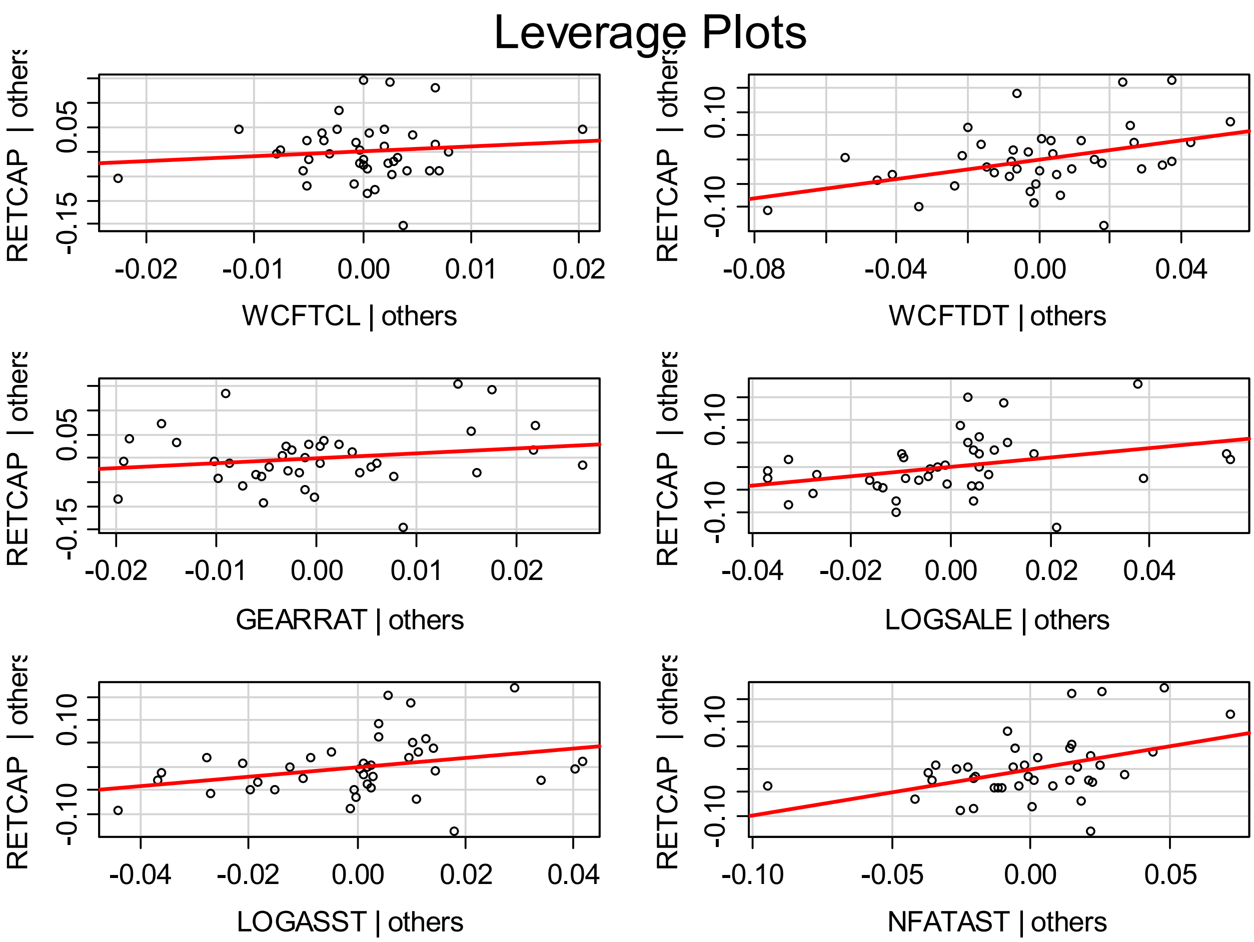
```
ret04.2<-subset(ret04.1,! rownames(ret04.1) %in% c("32" ,"16"))
m42<- lm(RETCAP~.,data=ret04.2)
summary(m42)
lm.beta(m42)
vif(m42)
```

```
Call:
lm(formula = RETCAP ~ ., data = ret04.2)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.144431 -0.028162 -0.001453  0.027169  0.125297 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.16634    0.05920   2.810 0.008506 ** 
WCFTCL       0.60519    0.10050   6.022 1.15e-06 ***
CAPINT       0.03516    0.01542   2.281 0.029597 *  
FATTOT      -0.21175    0.05208  -4.066 0.000304 ***
CURRAT      -0.06129    0.02633  -2.328 0.026620 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0597 on 31 degrees of freedom
Multiple R-squared: 0.5504,	Adjusted R-squared: 0.4924 
F-statistic: 9.487 on 4 and 31 DF,  p-value: 3.963e-05

    WCFTCL     CAPINT     FATTOT     CURRAT 
 0.9535220  0.3087906 -0.6310053 -0.3088687

  WCFTCL   CAPINT   FATTOT   CURRAT 
1.728890 1.263951 1.660818 1.213845
```



Influence Plot
Studentized Residuals
-2
-1
0
1
2
0.2
0.4
0.6
0.8
Hat-Values
Koła proporcjonalne do odległości Cooka


*



Leverage Plots
RETCAP | others
WCFTCL | others
WCFTDT | others
GEARRAT | others
LOGSALE | others
LOGASST | others
NFATAST | others

PRZEDZIAŁY UFNOŚCI DLA PARAMETRÓW

```
confint(m42, level=0.95)
```

```
                   2.5 %       97.5 %
(Intercept)  0.045610489  0.287073492
WCFTCL       0.400209579  0.810163473
CAPINT       0.003717529  0.066597045
FATTOT      -0.317965921 -0.105526037
CURRAT      -0.114991047 -0.007591935
```

$C_p$ Mallowsa I regresja minimalizująca błąd
{diagnostyka 3 [3]}

```
library("wle")
ret.cp <- mle.cp(model2)
summary(ret.cp)
```

```
Mallows Cp:
      (Intercept) OBS WCFTCL WCFTDT GEARRAT LOGSALE LOGASST NFATAST CAPINT FATTOT
 [1,]           1   0      0      1       0       0       0       1      0      0
 [2,]           1   0      0      1       0       0       0       1      0      0
 [3,]           1   0      0      1       0       0       0       1      0      0
 [4,]           1   0      1      0       0       0       0       0      0      1
```

```
      INVTAST PAYOUT QUIKRAT CURRAT      cp
 [1,]       0      0       0      1 -0.1564
 [2,]       0      1       0      1  0.2429
 [3,]       0      0       1      0  0.3788
 [4,]       0      0       0      1  0.6856
```

```
m5 <-lm(RETCAP~WCFTDT+NFATAST+CURRAT, data=ret0)
summary(m5)
lm.beta(m5)
vif(m5)
```

```
Call:
lm(formula = RETCAP ~ WCFTDT + NFATAST + CURRAT, data = ret0)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.142806 -0.033024 -0.004314  0.027566  0.170321 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.23080    0.03861   5.978 8.24e-07 ***
WCFTDT       0.52455    0.07902   6.638 1.12e-07 ***
NFATAST     -0.28553    0.06842  -4.173 0.000189 ***
CURRAT      -0.04753    0.01899  -2.503 0.017145 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 0.06215 on 35 degrees of freedom
Multiple R-squared: 0.5729,    Adjusted R-squared: 0.5363 
F-statistic: 15.65 on 3 and 35 DF,  p-value: 1.273e-06
```

```
lm.beta(m5)
    WCFTDT    NFATAST     CURRAT 
 0.8092370 -0.5267225 -0.3038093
```

```
vif(m5)
  WCFTDT  NFATAST   CURRAT 
1.217791 1.305646 1.207647
```

## WYKRES RESZT CZĄSTKOWYCH

```
crPlots(m5)
```