

Analiza zależności między wydatkami na reklamę a wartością sprzedaży

Dane

```
sprzedaz = read.csv('sprzedaz.txt')
names(sprzedaz) = c('Firma','reklama', 'przychód')

sprzedaz$reklama = sprzedaz$reklama /1000
head(sprzedaz)
```

```
##   Firma reklama przychód
## 1 2013Q1    0.32    10.2
## 2 2013Q2    0.25     9.3
## 3 2013Q3    0.34     9.6
## 4 2013Q4    0.41    11.1
## 5 2014Q1    0.35     9.9
## 6 2014Q2    0.50    12.9
```

Zależność pomiędzy wartością sprzedaży a wydatkami na reklamę

```
sprzedaz_df = subset(sprzedaz, select = c('reklama', 'przychód'))
names(sprzedaz_df) = c('reklama','przychód')

lm_sprzedaz = lm(przychód ~ reklama, data = sprzedaz_df)
summary(lm_sprzedaz)
```

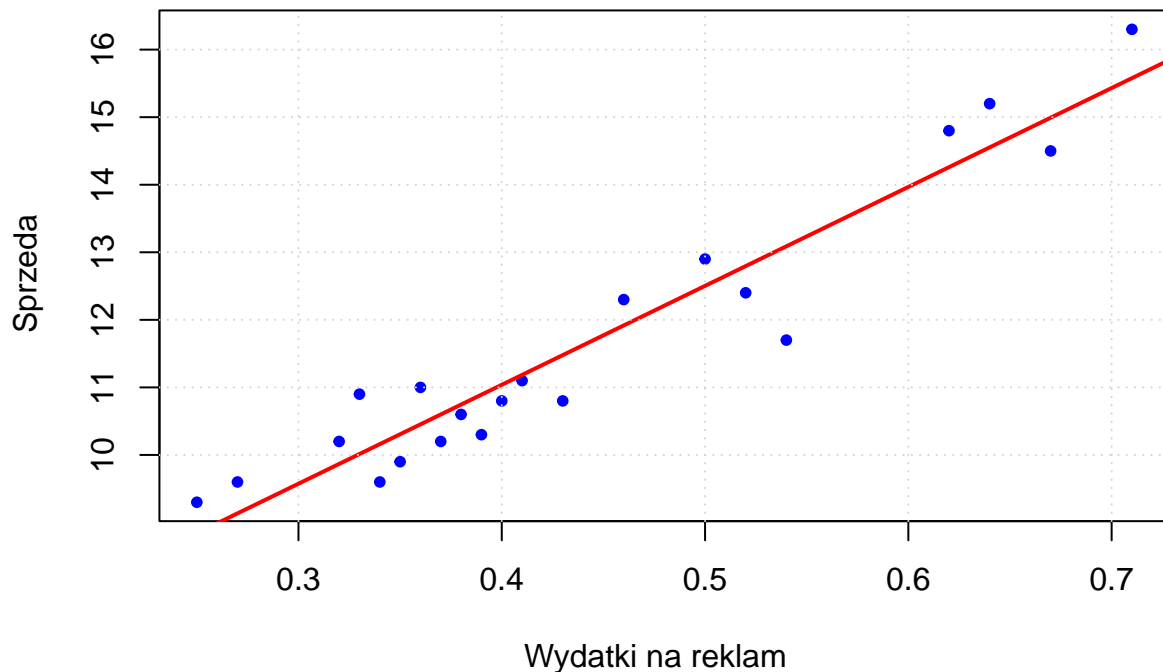
```
##
## Call:
## lm(formula = przychód ~ reklama, data = sprzedaz_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38840 -0.40633 -0.08488  0.46507  0.88652
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.1815     0.4732   10.95 1.20e-09 ***
## reklama      14.6425     1.0300   14.22 1.41e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6078 on 19 degrees of freedom
## Multiple R-squared:  0.9141, Adjusted R-squared:  0.9095
## F-statistic: 202.1 on 1 and 19 DF,  p-value: 1.41e-11
```

```
plot(sprzedaz$reklama, sprzedaz$przychód,
     main = "Zależność sprzedaży od wydatków na reklamę",
     xlab = "Wydatki na reklamę",
     ylab = "Sprzedaż",
     pch = 20,
     col = "blue")

abline(lm_sprzedaz, col = "red", lwd = 2)

grid()
```

Zależność sprzedaży od wydatków na reklamę



Prognostowane wartości sprzedaży przy wydatkach: 300, 500, 700 tys. zł.

```
pred_res = predict(lm_sprzedaz, newdata = data.frame(reklama = c(300, 500, 700)), se.fit = TRUE)
print("Prognostowane wartości sprzedaży:")
```

```
## [1] "Prognostowane wartości sprzedaży:"
```

```
print(pred_res)
```

```
## $fit
##      1      2      3
## 4397.930 7326.429 10254.928
##
## $se.fit
##      1      2      3
## 308.5519 514.5560 720.5600
##
## $df
## [1] 19
##
## $residual.scale
## [1] 0.6077826
```

Odchylenie standardowe błędów z jakim wyznaczono prognozowane wartości sprzedaży dla poszczególnych wartości wydatków na reklamę.

```
print("Odchylenie standardowe błędów predykcji:")
```

```
## [1] "Odchylenie standardowe błędów predykcji:"
```

```
print(pred_res$se.fit)
```

```
##          1          2          3
## 308.5519 514.5560 720.5600
```

Odchylenie standardowe błędu prognozy wynosi 308.55, 514.56, 720.56.

Porównanie wartości współczynnika determinacji R^2 dla dwóch modeli: regresji liniowej oraz zależności kwadratowej.

```
lm_sprzedaz = lm(przychód ~ reklama, data = sprzedaz_df)
summary(lm_sprzedaz)
```

```
##
## Call:
## lm(formula = przychód ~ reklama, data = sprzedaz_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38840 -0.40633 -0.08488  0.46507  0.88652
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.1815     0.4732   10.95 1.20e-09 ***
## reklama      14.6425     1.0300   14.22 1.41e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6078 on 19 degrees of freedom
## Multiple R-squared:  0.9141, Adjusted R-squared:  0.9095
## F-statistic: 202.1 on 1 and 19 DF,  p-value: 1.41e-11
```

```
lm_sprzedaz2 = lm(przychód ~ I(reklama^2) - 1, sprzedaz_df)
summary(lm_sprzedaz2)
```

```
##
## Call:
## lm(formula = przychód ~ I(reklama^2) - 1, data = sprzedaz_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.2626  0.2974  3.5761  4.4259  6.5026
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## I(reklama^2)  44.758      3.946   11.34 3.65e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.437 on 20 degrees of freedom
## Multiple R-squared:  0.8655, Adjusted R-squared:  0.8587
## F-statistic: 128.7 on 1 and 20 DF,  p-value: 3.655e-10
```

```
R2 = summary(lm_sprzedaz)$r.squared
R2_2 = summary(lm_sprzedaz2)$r.squared
```

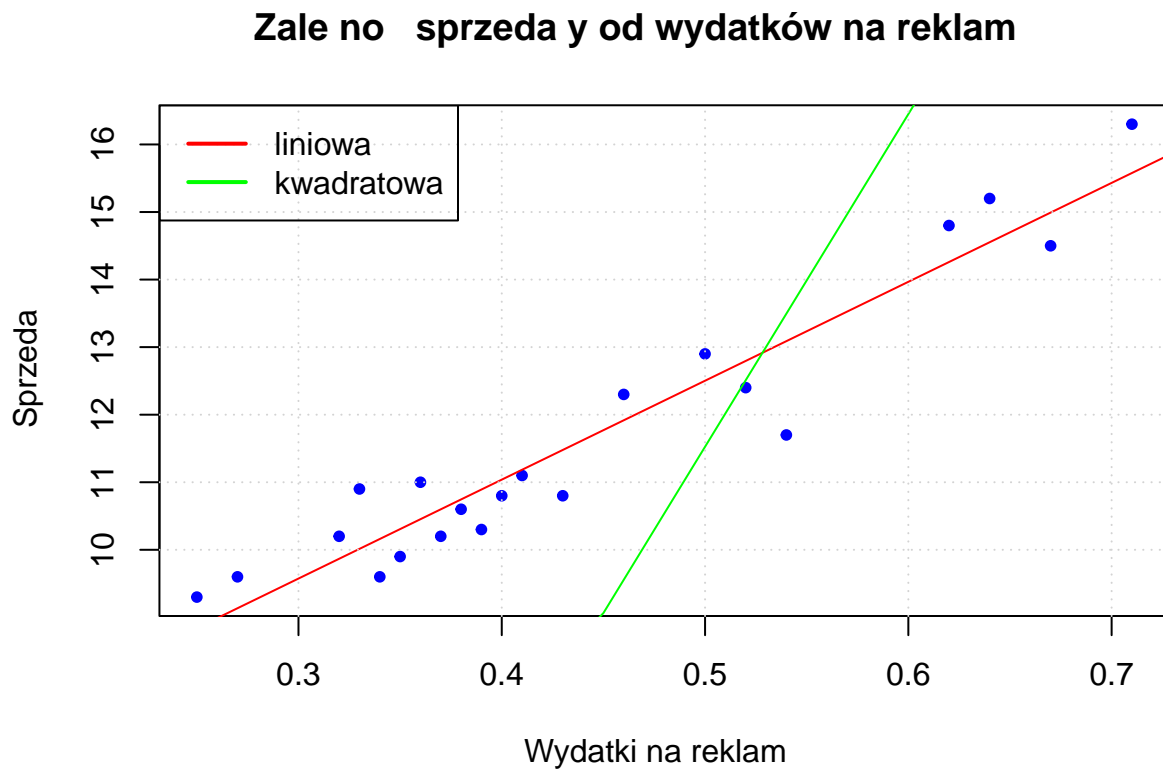
```

plot(przychód ~ reklama, sprzedaz_df,
     main = "Zależność sprzedaży od wydatków na reklamę",
     xlab = "Wydatki na reklamę",
     ylab = "Sprzedaż",
     pch = 20,
     col = "blue")
abline(lm_sprzedaz, col = 'red')

reklama = seq(min(sprzedaz_df$reklama), max(sprzedaz_df$reklama), by = 0.2)
DF_przewidywania = data.frame(reklama = reklama)
DF_przewidywania$przychód = predict(lm_sprzedaz2, DF_przewidywania)
lines(przychód ~ reklama, DF_przewidywania, col = 'green')

grid()
legend('topleft', c('liniowa', 'kwadratowa'), col = c('red', 'green'), lwd = 2)

```



Zależność kwadratowa nie byłaby lepszym modelem zależności między wartością wydatków na reklamę a wartością sprzedaży.