Regression-Analysis.R

Finley

2025-04-09

```r
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse
2.0.0 ──
## ✓ dplyr     1.1.4     ✓ readr     2.1.5
## ✓ forcats   1.0.0     ✓ stringr   1.5.1
## ✓ ggplot2   3.5.1     ✓ tibble    3.2.1
## ✓ lubridate 1.9.4     ✓ tidyr     1.3.1
## ✓ purrr     1.0.4
## ── Conflicts ────────────────────────────────────────
tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
```

```r
library(ggplot2)
library(dplyr)

RD <- read.csv("C:\\Users\\finle\\OneDrive\\University\\Econometrics and
Statistics\\Computer stuff\\RDCHEM.csv", header = TRUE)
var.labels <- c(rd = "R&D Spending, £M", sales = "firm Sales, £M", profits =
"Profits, £M")
attr(RD, "variable.labels") <- var.labels

glimpse(RD)
```

```
## Rows: 32
## Columns: 3
## $ rd      <dbl> 430.6, 59.0, 23.5, 3.5, 1.7, 8.4, 2.5, 39.9, 1136.0,
1428.0, 4…
## $ sales   <dbl> 4570.2, 2830.0, 596.8, 133.6, 42.0, 390.0, 93.9, 907.9,
19773.…
## $ profits <dbl> 186.9, 467.0, 107.4, -4.3, 8.0, 47.3, 0.9, 77.4, 2563.0,
4154.…
```

```r
RD <- RD %>%
  mutate(
    profmarg = profits/sales * 100,
    rdintens = rd / sales * 100,
    salessq = sales ** 2,
    lnsales = log(sales),
    lnrd = log(rd))
```

```r
glimpse(RD)
```

```
## Rows: 32
## Columns: 8
## $ rd       <dbl> 430.6, 59.0, 23.5, 3.5, 1.7, 8.4, 2.5, 39.9, 1136.0,
1428.0, …
## $ sales    <dbl> 4570.2, 2830.0, 596.8, 133.6, 42.0, 390.0, 93.9, 907.9,
19773…
## $ profits  <dbl> 186.9, 467.0, 107.4, -4.3, 8.0, 47.3, 0.9, 77.4, 2563.0,
4154…
## $ profmarg <dbl> 4.0895366, 16.5017668, 17.9959786, -3.2185629,
19.0476190, 12…
## $ rdintens <dbl> 9.421907, 2.084806, 3.937668, 2.619760, 4.047619,
2.153846, 2…
## $ salessq  <dbl> 2.088673e+07, 8.008900e+06, 3.561702e+05, 1.784896e+04,
1.764…
## $ lnsales  <dbl> 8.427312, 7.948032, 6.391582, 4.894850, 3.737670,
5.966147, 4…
## $ lnrd     <dbl> 6.0651796, 4.0775374, 3.1570004, 1.2527630, 0.5306283,
2.1282…
```

```r
var.labels <- c(var.labels, profmarg = "Profit Margins (%)",
                rdintens = "R&D as a percentage of sales",
                salessq = "Sales squared",
                lnsales = "natural log of sales",
                lnrd = "natural log of R&D")
rd.lm <- lm(data =RD, rdintens~ lnsales + profmarg)
summary(rd.lm)
```

```
##
## Call:
## lm(formula = rdintens ~ lnsales + profmarg, data = RD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3016 -1.2707 -0.6895  0.8785  6.0369
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.47225    1.67606   0.282    0.780
## lnsales      0.32135    0.21557   1.491    0.147
## profmarg     0.05004    0.04578   1.093    0.283
##
## Residual standard error: 1.839 on 29 degrees of freedom
## Multiple R-squared:  0.09847,    Adjusted R-squared:  0.0363
## F-statistic: 1.584 on 2 and 29 DF,  p-value: 0.2224
```