

A puzzling example: introduction to multiple linear regression.

R : Copyright 2001, The R Development Core Team  
Version 1.2.3 (2001-04-26)

R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.  
Type 'contributors()' for more information.

Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for a HTML browser interface to help.  
Type 'q()' to quit R.

```
# Load dataset
```

```
> sales <- read.table("/home/amafaculty/emmanuel/Data/sales.dat", header = T)
```

```
# Data
```

```
> sales[1:5,]  
  OwnPrice OtherPrice    Sales  
1 5.135670  5.204186 144.48788  
2 3.495460  8.059732 637.24524  
3 7.275341 11.675979 620.78693  
4 4.662816  8.364421 549.00714  
5 3.584537  2.150292  20.42542
```

```
> attach(sales)
```

```
# Plot data
```

```
> plot(OwnPrice, Sales)
```

```
# Regress Sales on OwnPrice
```

```
> sales1.lm <- lm(Sales ~ OwnPrice)  
> summary(sales1.lm)
```

```
Call:
```

```
lm(formula = Sales ~ OwnPrice)
```

```
Residuals:
```

```
      Min       1Q   Median       3Q      Max  
-513.912 -157.695  -1.425  155.195  650.196
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	211.16	66.49	3.176	0.00200	**
OwnPrice	63.71	13.04	4.886	4.01e-06	***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 223.4 on 98 degrees of freedom

Multiple R-Squared: 0.1959, Adjusted R-squared: 0.1877

F-statistic: 23.87 on 1 and 98 degrees of freedom, p-value: 4.015e-06

# Other Plots

```
> par(mfrow = c(2,2))
> plot(OwnPrice,OtherPrice)
> plot(OtherPrice,Sales)
> plot(OwnPrice,OtherPrice)
```

# Isolate entries with a roughly fixed value of competitor's price

```
> length(OtherPrice)
[1] 100
> index <- (1:100)[(OtherPrice > 4)]
> index <- index[OtherPrice[index] < 6]
> index
[1] 1 7 12 29 48 51 53 56 67 68 69 70 76 89
```

# Plot data again

```
> postscript("/home/amafaculty/emmanuel/sales3.eps")
> par(mfrow = c(1,1))
> plot(OwnPrice[-index], Sales[-index],
      main = "Sales vs. Own Price: Competitor Level Fixed")
> points(OwnPrice[index], Sales[index], type = "p", pch = 0)
> dev.off()
```

X11

2

# Regress Sales on Own Price and Competitor's Price.

```
> Sales.lm <- lm(Sales ~ OwnPrice + OtherPrice)
> Sales.lm
```

Call:

```
lm(formula = Sales ~ OwnPrice + OtherPrice)
```

Coefficients:

```
(Intercept)    OwnPrice    OtherPrice
      115.72         -97.66         108.80
```

```
> summary(Sales.lm)
```

```
Call:
```

```
lm(formula = Sales ~ OwnPrice + OtherPrice)
```

```
Residuals:
```

```
      Min       1Q   Median       3Q      Max
-66.9161 -15.6634  -0.5095  18.9038  63.3021
```

```
Coefficients:
```

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  115.717      8.548   13.54  <2e-16 ***
OwnPrice     -97.657      2.669  -36.59  <2e-16 ***
OtherPrice   108.800      1.409   77.20  <2e-16 ***
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 28.42 on 97 degrees of freedom
```

```
Multiple R-Squared: 0.9871,    Adjusted R-squared: 0.9869
```

```
F-statistic: 3717 on 2 and 97 degrees of freedom,    p-value: 0
```

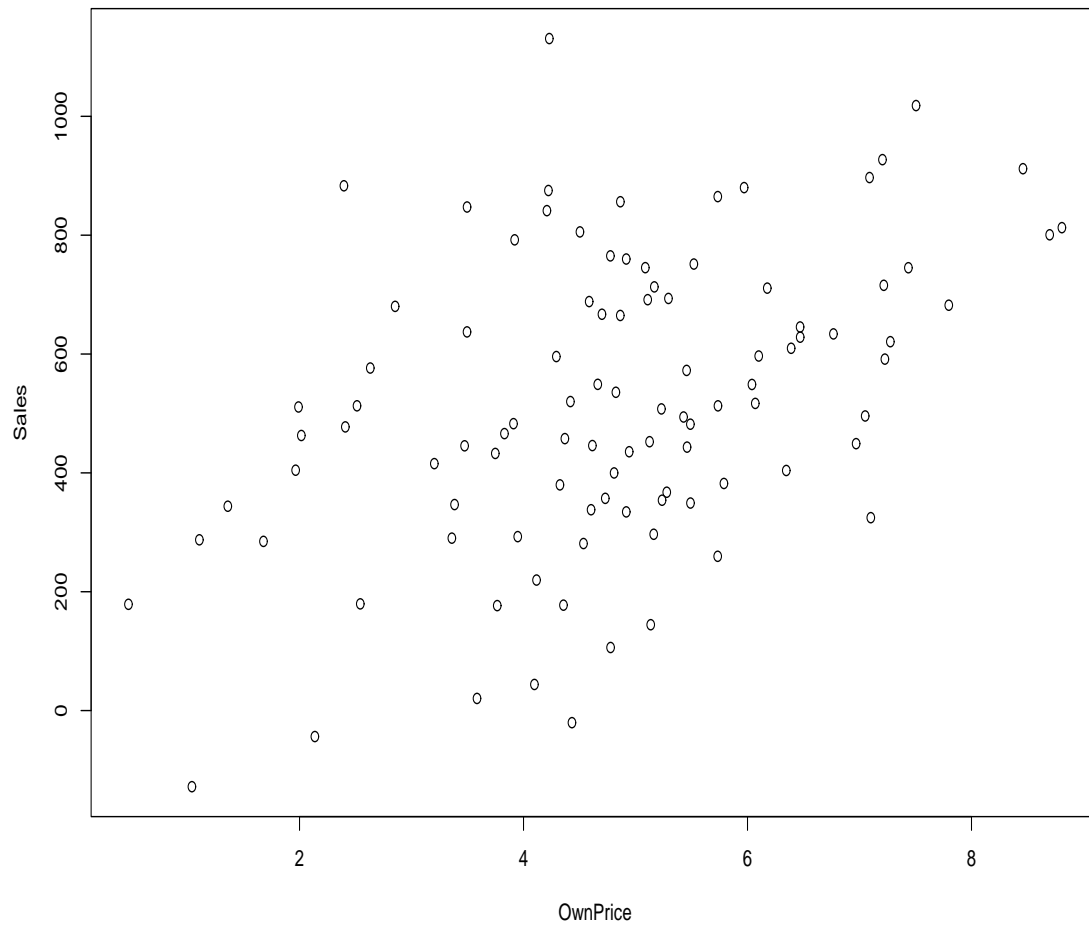


Figure 1: Sales vs. Own Price

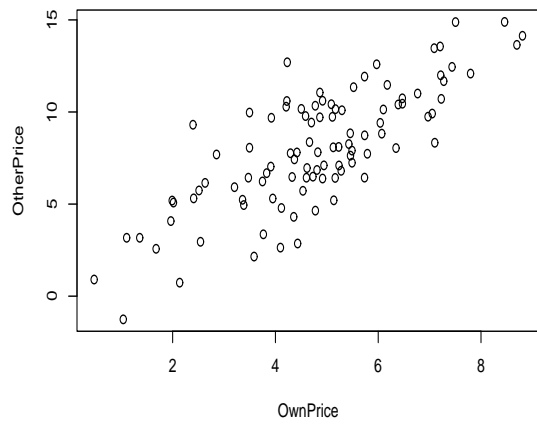
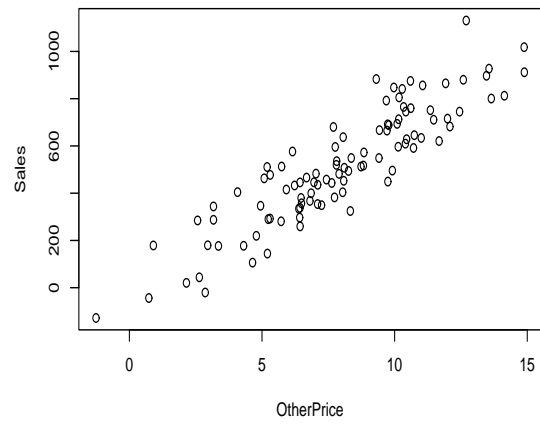
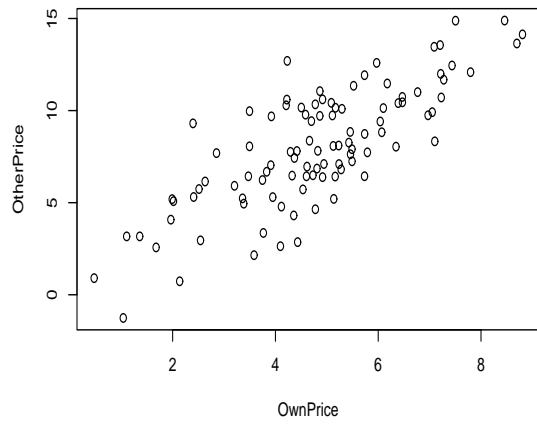


Figure 2: Other Plots

Sales vs. Own Price: Competitor Level Fixed

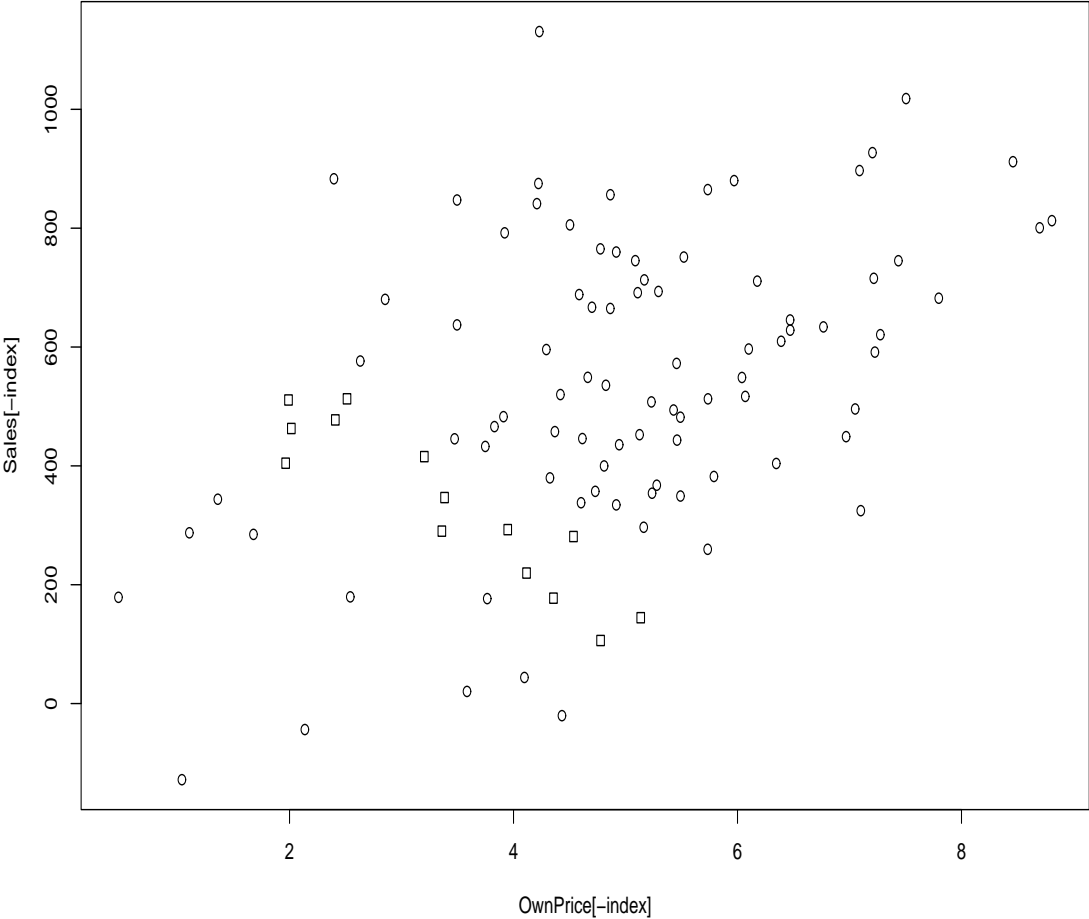


Figure 3: Sales vs. Own Price: Competitor Level Fixed

```
> plot(OwnPrice,Sales)
> plot(OtherPrice,Sales)
> plot(OwnPrice,OtherPrice)
> plot(Prices.lm$resid,sales1.lm$resid,xlab="Residuals from
OtherPrice on OwnPrice",ylab="Residuals from Sales on OwnPrice")
```

Multiple linear regression of Sales on OwnPrice and OtherPrice:

```
> sales.lm <- lm(Sales ~ OwnPrice + OtherPrice)
> sales.lm
Call:
lm(formula = Sales ~ OwnPrice + OtherPrice)
```

```
Coefficients:
(Intercept)  OwnPrice  OtherPrice
  115.7172  -97.65737   108.7999
```

```
Degrees of freedom: 100 total; 97 residual
Residual standard error: 28.41801
```

Simple linear regression of Sales on OwnPrice:

```
> sales1.lm <- lm(Sales ~ OwnPrice)
> sales1.lm
Call:
lm(formula = Sales ~ OwnPrice)
```

```
Coefficients:
(Intercept)  OwnPrice
  211.1646  63.71296
```

```
Degrees of freedom: 100 total; 98 residual
Residual standard error: 223.4006
```

Understanding the effect of adding one predictor: Simple linear regression of OtherPrice on OwnPrice.

```
> Prices.lm<-lm(OtherPrice ~ OwnPrice)
> Prices.lm
Call:
lm(formula = OtherPrice ~ OwnPrice)
```

```
Coefficients:
```

```
(Intercept) OwnPrice
0.8772747 1.483184
```

```
Degrees of freedom: 100 total; 98 residual
Residual standard error: 2.036807
```

Just checking that the instructor is correct! Regress the residuals.

```
> lm(sales1.lm$resid ~ Prices.lm$resid)
Call:
lm(formula = sales1.lm$resid ~ Prices.lm$resid)
```

```
Coefficients:
(Intercept) Prices.lm$resid
-2.01794e-14      108.7999
```

```
Degrees of freedom: 100 total; 98 residual
Residual standard error: 28.27265
```

Then the slope is  $\beta_2 = 108.8$ , the coefficient associated with OtherPrice in the multiple regression.

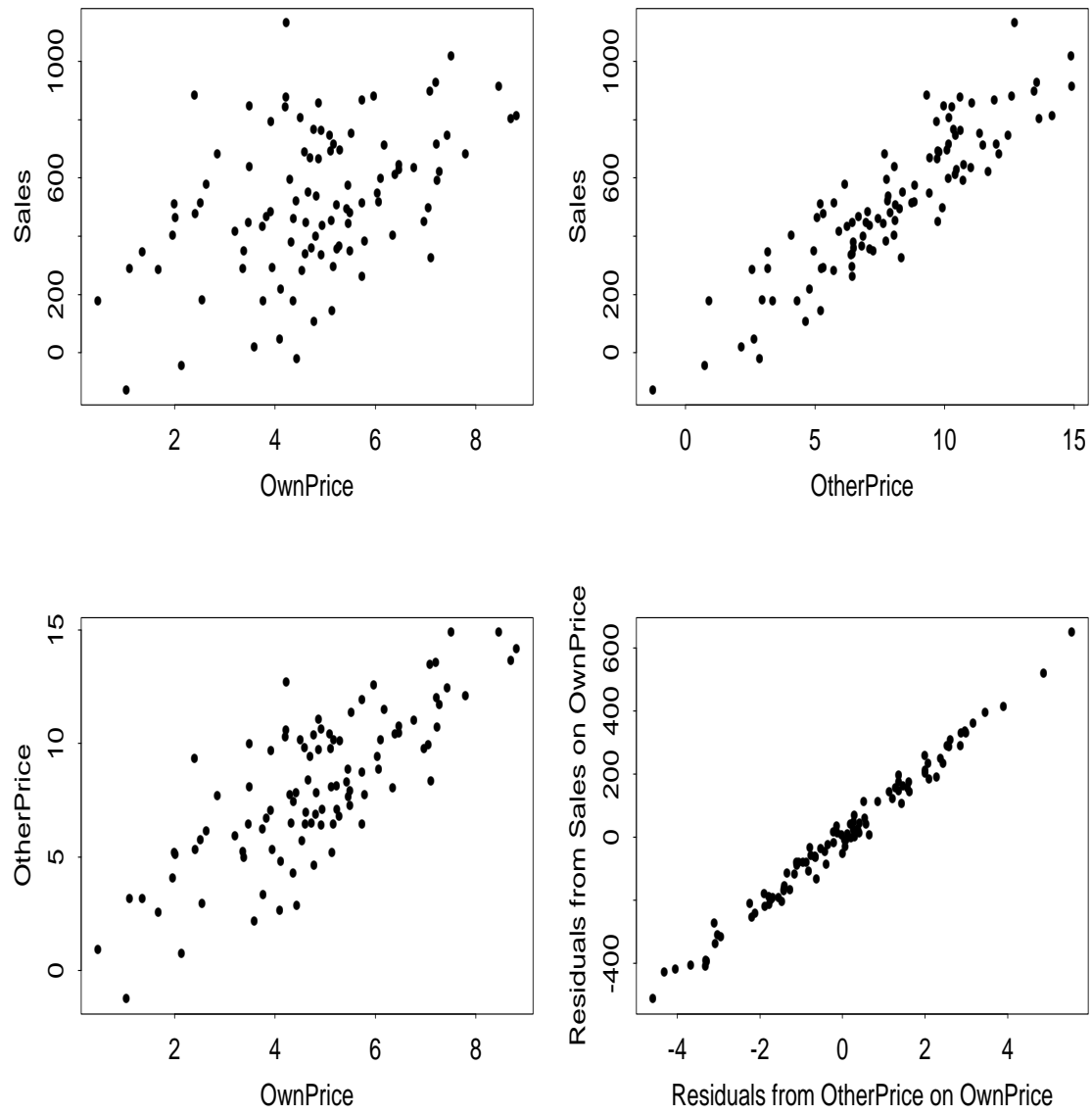


Figure 1: Explaining multiple regression