

This report was automatically generated with the R package **knitr** (version 1.5).

```
library(faraway)
data(gala, package = "faraway")
head(gala[, -2])
```

n=30
number of species found
on the various Galapagos
Islands
nearest - distance to the
nearest island
Scruz - distance to Santa
Cruz Island
Adjacent - area of the
adjacent island

	Species	Area	Elevation	Nearest	Scruz	Adjacent
Bal tra	58	25.09	346	0.6	0.6	1.84
Bartolome	31	1.24	109	0.6	26.3	572.33
Caldwell	3	0.21	114	2.8	58.7	0.78
Champion	25	0.10	46	1.9	47.4	0.18
Coamano	2	0.05	77	1.9	1.9	903.82
Daphne. Major	18	0.34	119	8.0	8.0	1.84

```
lmod <- lm(Species ~ Area + Elevation + Nearest + Scruz + Adjacent, data = gala)
summary(lmod)
```

```
Call:
lm(formula = Species ~ Area + Elevation + Nearest + Scruz + Adjacent,
    data = gala)
```

What does first row of X matrix
look like?

Residuals:

Min	1Q	Median	3Q	Max
-111.68	-34.90	-7.86	33.46	182.58

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.06822	19.15420	0.37	0.7154
Area	-0.02394	0.02242	-1.07	0.2963
Elevation	0.31946	0.05366	5.95	3.8e-06
Nearest	0.00914	1.05414	0.01	0.9932
Scruz	-0.24052	0.21540	-1.12	0.2752
Adjacent	-0.07480	0.01770	-4.23	0.0003

Residual standard error: 61 on 24 degrees of freedom

Multiple R-squared: 0.766, Adjusted R-squared: 0.717

F-statistic: 15.7 on 5 and 24 DF, p-value: 6.84e-07

```
require(faraway)
summary(lmod)
```

reduced information - note the missing 'm' in summary

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.06822	19.15420	0.37	0.7154
Area	-0.02394	0.02242	-1.07	0.2963
Elevation	0.31946	0.05366	5.95	3.8e-06
Nearest	0.00914	1.05414	0.01	0.9932

```
Scruz      - 0.24052    0.21540   - 1.12    0.2752
Adjacent   - 0.07480    0.01770   - 4.23    0.0003
```

n = 30, p = 6, Residual SE = 60.98, R-Squared = 0.77

```
x <- model.matrix(~Area + Elevation + Nearest + Scruz + Adjacent, gala)
y <- gala$Species
xtxi <- solve(t(x) %*% x)
xtxi %*% t(x) %*% y
```

extract X matrix. extract y vector.

jump to odor
data set

```
      [, 1]
(Intercept) 7.068221
Area         -0.023938
Elevation    0.319465
Nearest      0.009144
Scruz        -0.240524
Adjacent     -0.074805
```

```
solve(crossprod(x, x), crossprod(x, y))
```

```
      [, 1]
(Intercept) 7.068221
Area         -0.023938
Elevation    0.319465
Nearest      0.009144
Scruz        -0.240524
Adjacent     -0.074805
```

```
names(lmod)
```

```
[1] "coefficients" "residuals"    "effects"      "rank"
[5] "fitted.values" "assign"        "qr"           "df.residual"
[9] "xlevels"       "call"         "terms"        "model"
```

```
lmodsum <- summary(lmod)
names(lmodsum)
```

```
[1] "call"          "terms"         "residuals"     "coefficients"
[5] "aliases"       "sigma"         "df"            "r.squared"
[9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
sqrt(deviance(lmod)/df.residual(lmod))
```

```
[1] 60.98
```

```
lmodsum$sigma
```

4

```
[1] 60.98
```

```
xtxi <- lmodsum$cov.unscaled  
sqrt(diag(xtxi)) * 60.975
```

(Intercept)	Area	Elevation	Nearest	Scruz	Adjacent
19.15414	0.02242	0.05366	1.05413	0.21540	0.01770

```
lmodsum$coef[, 2]
```

(Intercept)	Area	Elevation	Nearest	Scruz	Adjacent
19.15420	0.02242	0.05366	1.05414	0.21540	0.01770

```
qrx <- qr(x)  
dim(qr.Q(qrx))
```

```
[1] 30 6
```

```
(f <- t(qr.Q(qrx)) %*% y)
```

```
      [,1]  
[1,] -466.842  
[2,]  381.406  
[3,]  256.250  
[4,]    5.408  
[5,] -119.498  
[6,]  257.694
```

```
backsolve(qr.R(qrx), f)
```

```
      [,1]  
[1,]  7.068221  
[2,] -0.023938  
[3,]  0.319465  
[4,]  0.009144  
[5,] -0.240524  
[6,] -0.074805
```

```
gala$Adiff <- gala$Area - gala$Adjacent  
lmod <- lm(Species ~ Area + Elevation + Nearest + Scruz + Adjacent + Adiff,  
          gala)  
summary(lmod)
```

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.06822	19.15420	0.37	0.7154
Area	-0.02394	0.02242	-1.07	0.2963
Elevation	0.31946	0.05366	5.95	3.8e-06
Nearest	0.00914	1.05414	0.01	0.9932
Scruz	-0.24052	0.21540	-1.12	0.2752
Adjacent	-0.07480	0.01770	-4.23	0.0003

n = 30, p = 6, Residual SE = 60.98, R-Squared = 0.77

```
set.seed(123)
```

```
Adiffe <- gala$Adifff + 0.001 * (runif(30) - 0.5)
```

```
lmod <- lm(Species ~ Area + Elevation + Nearest + Scruz + Adjacent + Adiffe,
  gala)
```

```
summary(lmod)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.30e+00	1.94e+01	0.17	0.87
Area	-4.51e+04	4.26e+04	-1.06	0.30
Elevation	3.13e-01	5.39e-02	5.81	6.4e-06
Nearest	3.83e-01	1.11e+00	0.35	0.73
Scruz	-2.62e-01	2.16e-01	-1.21	0.24
Adjacent	4.51e+04	4.26e+04	1.06	0.30
Adiffe	4.51e+04	4.26e+04	1.06	0.30

n = 30, p = 7, Residual SE = 60.82, R-Squared = 0.78

```
data(odor, package = "faraway")
```

```
odor
```

	odor	temp	gas	pack
1	66	-1	-1	0
2	39	1	-1	0
3	43	-1	1	0
4	49	1	1	0
5	58	-1	0	-1
6	17	1	0	-1
7	-5	-1	0	1
8	-40	1	0	1
9	65	0	-1	-1
10	7	0	1	-1
11	43	0	-1	1
12	-22	0	1	1
13	-31	0	0	0
14	-35	0	0	0

designed experiment to reduce the unpleasant odor of a chemical product sold for household use.

Factors are column temperature, gas/liquid ratio, and packing height. Factors are coded to -1, 0, 1.

Note: experimenter has control of the levels of the factors.

```
cov(odor[, -1])
```

```
      temp    gas    pack
temp 0.5714 0.0000 0.0000
gas   0.0000 0.5714 0.0000
pack  0.0000 0.0000 0.5714
```

```
lmod <- lm(odor ~ temp + gas + pack, odor)
summary(lmod, cor = T)
```

Call:

```
lm(formula = odor ~ temp + gas + pack, data = odor)
```

Residuals:

Min	1Q	Median	3Q	Max
-50.20	-17.14	1.18	20.30	62.93

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	15.2	9.3	1.63	0.13
temp	-12.1	12.7	-0.95	0.36
gas	-17.0	12.7	-1.34	0.21
pack	-21.4	12.7	-1.68	0.12

Residual standard error: 36 on 11 degrees of freedom

Multiple R-squared: 0.334, Adjusted R-squared: 0.152

F-statistic: 1.84 on 3 and 11 DF, p-value: 0.199

Correlation of Coefficients:

	(Intercept)	temp	gas
temp	0.00		
gas	0.00	0.00	
pack	0.00	0.00	0.00

```
lmod <- lm(odor ~ gas + pack, odor)
summary(lmod)
```

Call:

```
lm(formula = odor ~ gas + pack, data = odor)
```

Residuals:

Min	1Q	Median	3Q	Max
-50.20	-26.70	1.17	26.80	50.80

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	15.20	9.26	1.64	0.13
gas	-17.00	12.68	-1.34	0.20
pack	-21.37	12.68	-1.69	0.12

Residual standard error: 35.9 on 12 degrees of freedom

Multiple R-squared: 0.279, Adjusted R-squared: 0.159

F-statistic: 2.32 on 2 and 12 DF, p-value: 0.141

```
x <- 1:20
y <- x + rnorm(20)
```

The R session information (including the OS info, R version and all packages used):

```
sessionInfo()
```

R version 3.1.0 (2014-04-10)

Platform: x86_64-apple-darwin13.1.0 (64-bit)

locale:

[1] en_GB.UTF-8/en_GB.UTF-8/en_GB.UTF-8/C/en_GB.UTF-8/en_GB.UTF-8

attached base packages:

[1] graphics grDevices utils datasets methods stats base

other attached packages:

[1] faraway_1.0.6 knitr_1.5 ggplot2_0.9.3.1

loaded via a namespace (and not attached):

[1] colorspace_1.2-4 dichromat_2.0-0 digest_0.6.4
 [4] evaluate_0.5.3 formatR_0.10 grid_3.1.0
 [7] gtable_0.1.2 labeling_0.2 MASS_7.3-31
 [10] munsell_0.4.2 plyr_1.8.1 proto_0.3-10
 [13] RColorBrewer_1.0-5 Rcpp_0.11.1 reshape2_1.2.2
 [16] scales_0.2.3 stringr_0.6.2 tools_3.1.0

```
Sys.time()
```

[1] "2014-06-16 14:01:04 BST"