```
achiev <- read.table("achiev.txt", header = T)

achiev

    Ind ac teach text

1      1 70      1      1

2      2 67      1      1

3      3 65      1      1

attach(achiev)

boxplot(ac ~ teach, data=achiev, main="Achievement as a Function of
Teacher",xlab = "Teacher", ylab="Achievement")
```
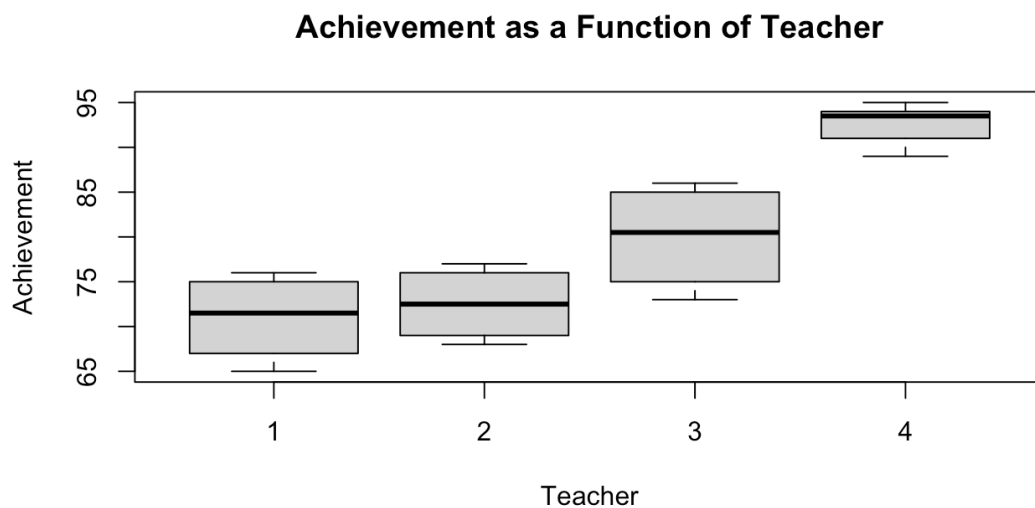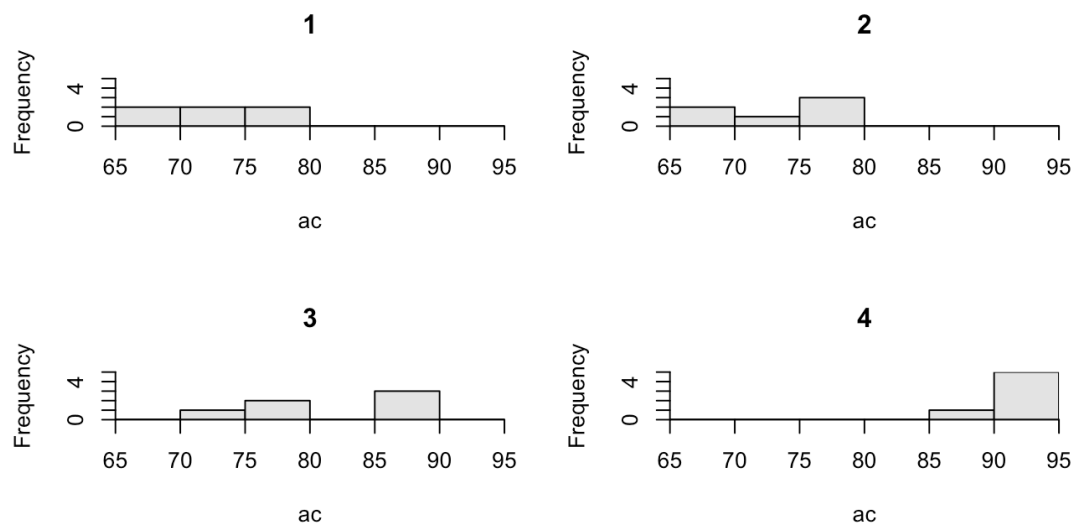
Achievement as a Function of Teacher


```
install.packages('FSA')

library(FSA)

f.teach <- factor(teach)

hist(ac~f.teach, data = achiev)
```

shapiro.test(ac)

        Shapiro-Wilk normality test

data: ac

W = 0.90565, p-value = 0.02842

fligner.test(ac~f.teach, data = achiev)

        Fligner-Killeen test of homogeneity of variances

data: ac by f.teach

Fligner-Killeen:med chi-squared = 10.813, df = 3, p-value = 0.01278

aggregate(ac ~ f.teach, FUN = var)

   f.teach   ac

1   1 19.600000

2   2 15.500000

3   3 35.200000

4   4   5.066667

aggregate(ac ~ f.teach, FUN = mean)

   f.teach   ac

1   1 71.00000

2   2 72.50000

3   3 80.00000

4   4 92.66667

```
anova.fit <- aov(ac ~ f.teach, data = achiev)

summary(anova.fit)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
f.teach	3	1764.1	588.0	31.21	9.68e-08	***
Residuals	20	376.8	18.8			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
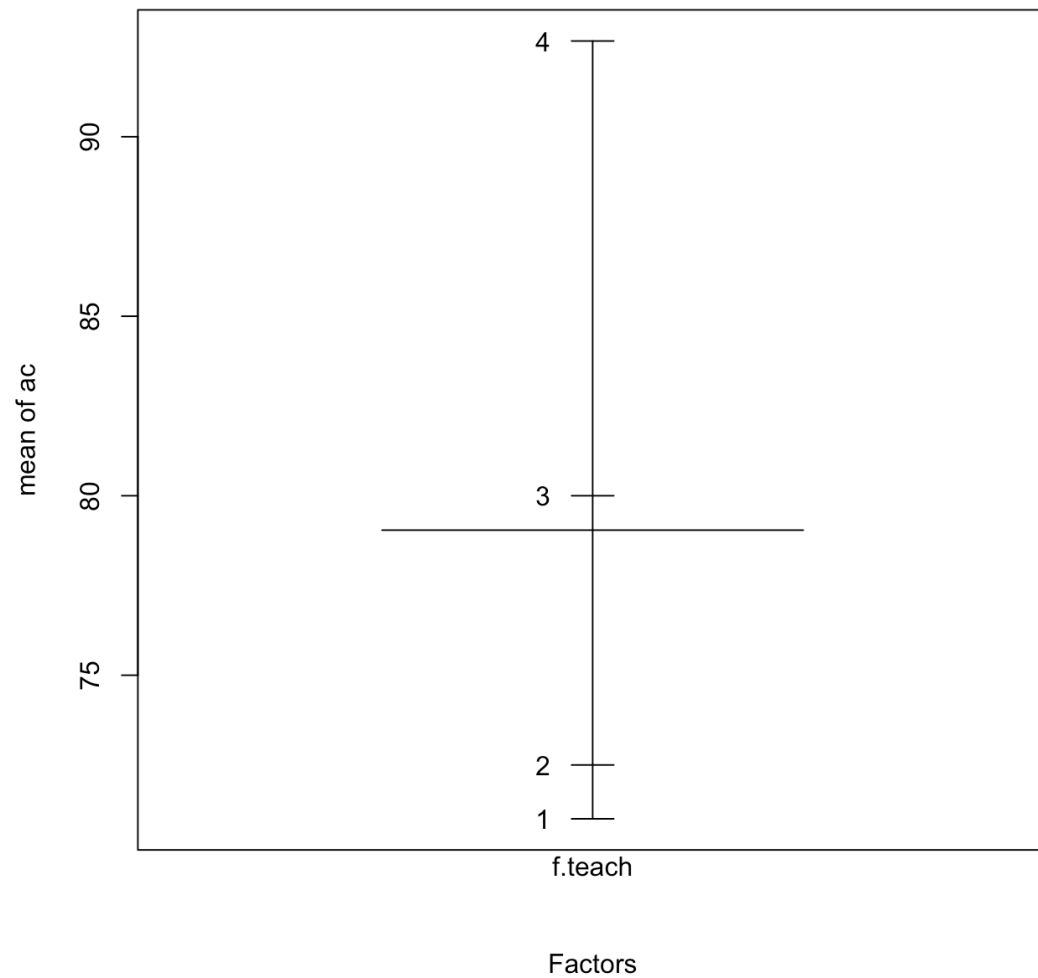
```
model.tables(anova.fit)
```

Tables of effects

 f.teach

f.teach

1	2	3	4
-8.042	-6.542	0.958	13.625

```
plot.design(ac~f.teach)
```

TukeyHSD(anova.fit)

   Tukey multiple comparisons of means

    95% family-wise confidence level

Fit: aov(formula = ac ~ f.teach, data = achiev)

$f.teach

	diff	lwr	upr	p adj
2-1	1.50000	-5.5144241	8.514424	0.9313130

3-1   9.00000   1.9855759 16.014424 0.0090868

4-1 21.66667 14.6522425 28.681091 0.0000002

3-2   7.50000   0.4855759 14.514424 0.0334428

```
delivery <- read.csv('delivery.csv', header = T)

delivery <- read.table('delivery.csv', sep=',', header = T)

attach(delivery)

library(car)

scatterplot(time ~ distance)
```
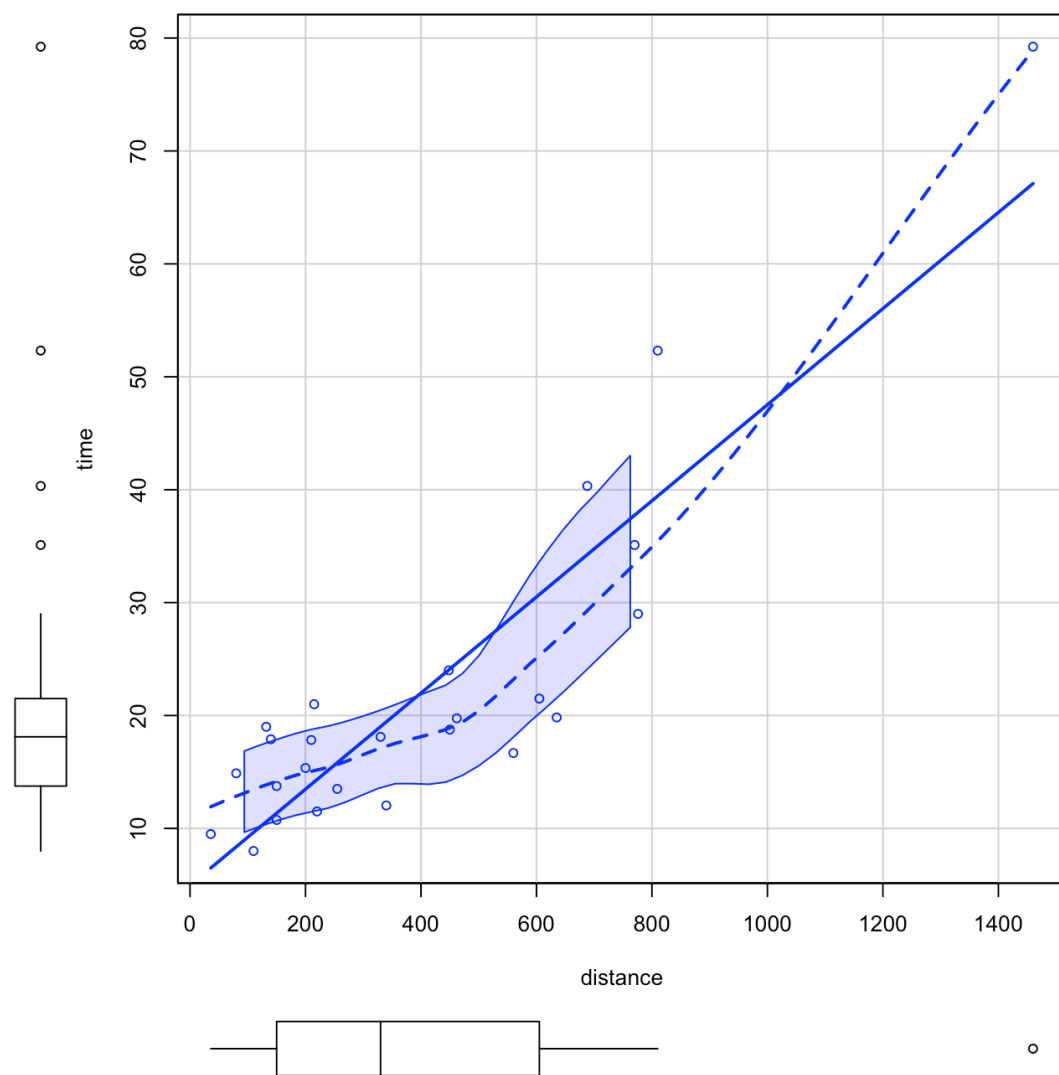


```
scatterplot(time ~ distance)
```

```
fit <- lm(time ~ distance)

fit

Call:

lm(formula = time ~ distance)

Coefficients:

(Intercept)        distance

    4.96116         0.04257

coef(fit)

(Intercept)       distance

  4.96115886    0.04256949

cor(fitted(fit),time)^2

[1] 0.7950755

summary(fit)

Call:

lm(formula = time ~ distance)

Residuals:

      Min         1Q     Median         3Q         Max

-12.1628    -4.8783    -0.5966     6.0810    12.8776

Coefficients:

              Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.961159     2.337360     2.123      0.0448 *

distance     0.042569     0.004506     9.447 2.21e-09 ***

---

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

Residual standard error: 7.179 on 23 degrees of freedom

Multiple R-squared: 0.7951,   Adjusted R-squared: 0.7862

F-statistic: 89.24 on 1 and 23 DF,   p-value: 2.214e-09

anova(fit)

Analysis of Variance Table

Response: time

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
distance	1	4599.1	4599.1	89.237	2.214e-09 ***
Residuals	23	1185.4	51.5		

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

new <- data.frame(distance=c(100, 500, 1000))

predict(fit, new)

1	2	3
9.218108	26.245904	47.530650

fit.mr <- lm(time~distance+case)

summary(fit.mr)

Call:

lm(formula = time ~ distance + case)

Residuals:

Min	1Q	Median	3Q	Max
-5.7880	-0.6629	0.4364	1.1566	7.4197

anova(fit.mr)

Analysis of Variance Table

Response: time

```
           Df Sum Sq Mean Sq F value      Pr(>F)

distance    1 4599.1   4599.1 432.895 5.825e-16 ***

case        1   951.7    951.7  89.575 3.255e-09 ***

Residuals 22   233.7     10.6
```
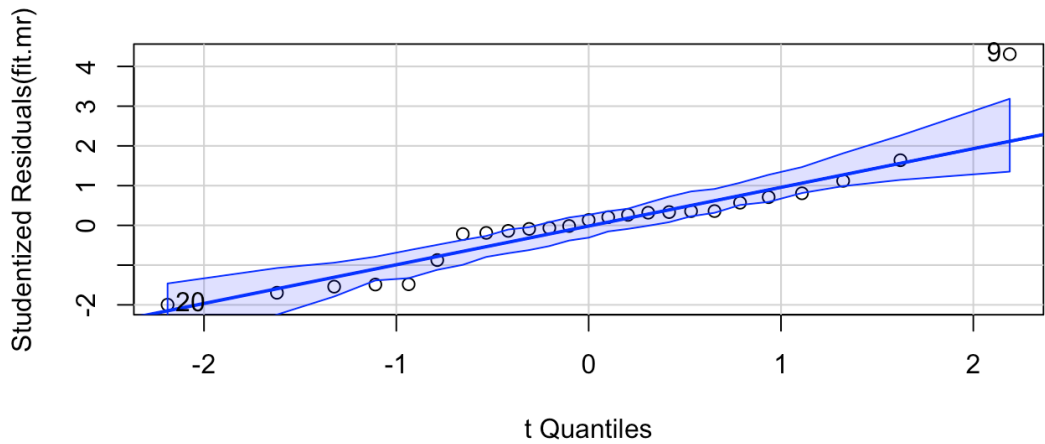

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

cor(fitted(fit),time)^2

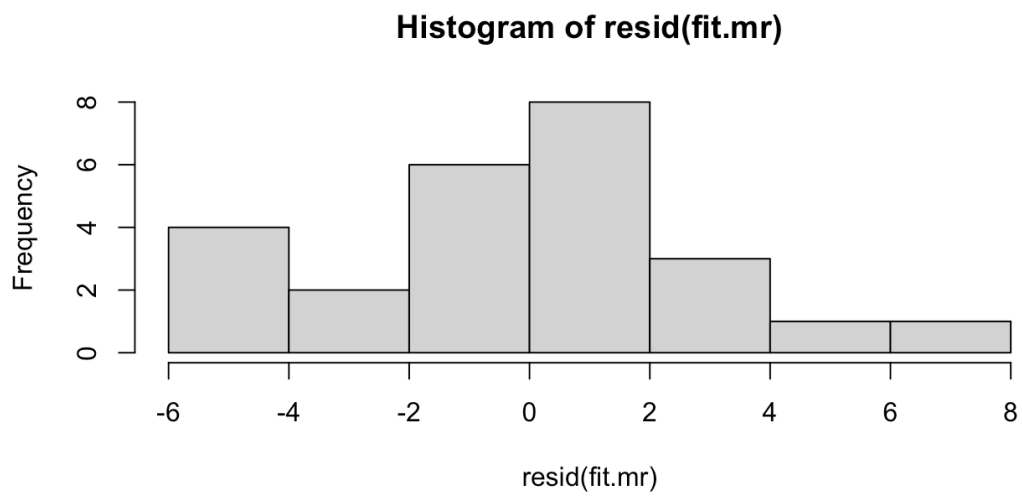[1] 0.7950755

qqPlot(fit.mr)

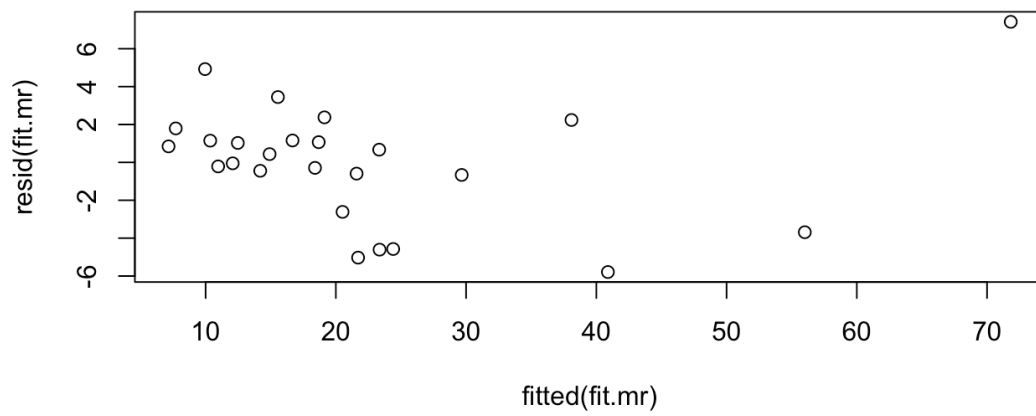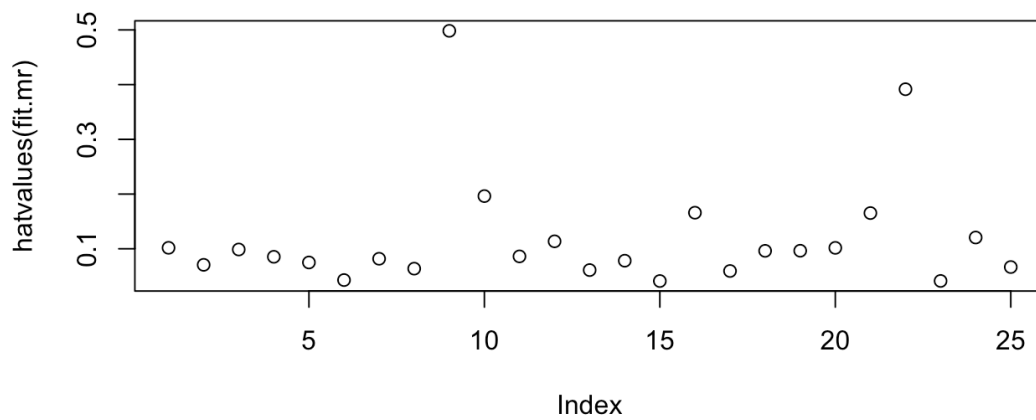[1]   9 20

hist(resid(fit.mr))

Histogram of resid(fit.mr)

plot(fitted(fit.mr), resid(fit.mr))

plot(hatvalues(fit.mr))

```
which.max(hatvalues(fit.mr))
```

9

9

```
cooks.distance(fit.mr)
```

| 1 | 2 | 3 | 4 | 5 |
| 6 | 7 | | | |

1.000921e-01 3.375704e-03 9.455785e-06 7.764718e-02 5.432217e-04
1.231067e-04 2.171604e-03

| 8 | 9 | 10 | 11 | 12 |
| 13 | 14 | | | |

3.051135e-03 3.419318e+00 5.384516e-02 1.619975e-02 1.596392e-03
2.294737e-03 3.292786e-03

| 15 | 16 | 17 | 18 | 19 |
| 20 | 21 | | | |

6.319880e-04 3.289086e-03 4.013419e-04 4.397807e-02 1.191868e-02
1.324449e-01 5.086063e-02

22	23	24	25

4.510455e-01 2.989892e-02 1.023224e-01 1.084694e-04

```
p1 <- c(2, 4, 6, 8, 10)
```

```
p2 <- c(4, 8, 12, 16, 20)
```

```
y <- c(12, 15, 8, 4, 1)
```

```
model <- lm(y ~ p1 + p2)
```

```
summary(model)
```

Call:

```
lm(formula = y ~ p1 + p2)
```

Residuals:

1	2	3	4	5
-2.600e+00	3.700e+00	6.661e-16	-7.000e-01	-4.000e-01

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	17.9000	2.7815	6.435	0.00761	**
p1	-1.6500	0.4193	-3.935	0.02924	*
p2	NA	NA	NA	NA	

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

Residual standard error: 2.652 on 3 degrees of freedom

Multiple R-squared: 0.8377,   Adjusted R-squared: 0.7836

F-statistic: 15.48 on 1 and 3 DF,   p-value: 0.02924

plot(p1,p2)

model.1 <- lm(y ~ p1)

summary(model.1)

Call:

lm(formula = y ~ p1)

Residuals:

1	2	3	4	5

-2.600e+00   3.700e+00   6.661e-16 -7.000e-01 -4.000e-01

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept)   17.9000   2.7815   6.435   0.00761 **

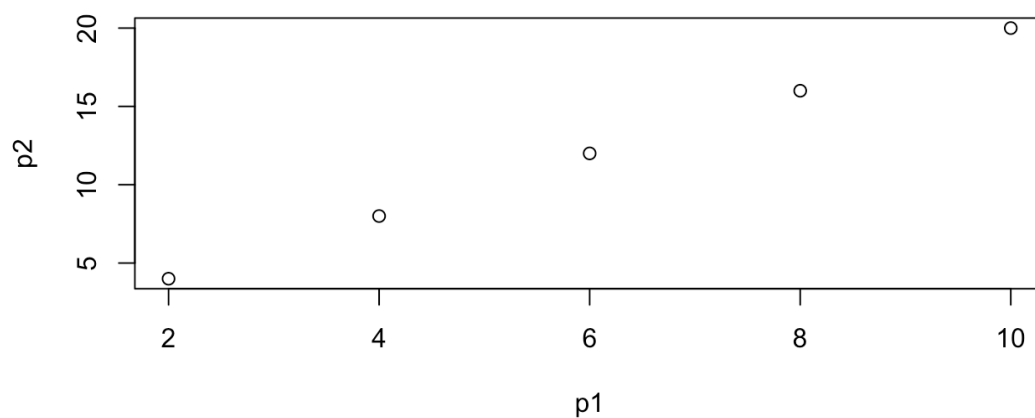p1   -1.6500   0.4193   -3.935   0.02924 *

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

Residual standard error: 2.652 on 3 degrees of freedom

Multiple R-squared: 0.8377,   Adjusted R-squared: 0.7836

F-statistic: 15.48 on 1 and 3 DF,   p-value: 0.02924

vif(model)

Error in vif.default(model) : there are aliased coefficients in the model

p2 <- c(4, 8, 12, 16, 21)

model <- lm(y ~ p1 + p2)

summary(model)

  Call:

lm(formula = y ~ p1 + p2)

Residuals:

1	2	3	4	5

-2.400e+00   3.700e+00 -2.000e-01 -1.100e+00   1.965e-14

Coefficients:

```
                    Estimate Std. Error t value Pr(>|t|)

(Intercept)    17.500          3.940     4.441     0.0471 *

p1                 0.450       10.694     0.042     0.9703

p2                -1.000        5.087    -0.197     0.8623

---

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

Residual standard error: 3.217 on 2 degrees of freedom

Multiple R-squared:    0.8408,    Adjusted R-squared:    0.6815

F-statistic:    5.28 on 2 and 2 DF,    p-value: 0.1592

gaf.data <- read.table("gaf.txt", header = T)

gaf.data

ID gaf age pretherapy n_therapy

1    1    25   21          52           6

2    2    25   19          61          17

3    3    0    18          50           5

attach(gaf.data)

gaf.fit <- lm(gaf ~ age)

summary(gaf.fit)

Call:

lm(formula = gaf ~ age)

Residuals:

     Min        1Q    Median        3Q        Max

-13.657   -9.196     4.359     7.617     9.784

Coefficients:
```

```
                Estimate Std. Error t value Pr(>|t|)

(Intercept) -15.6806      12.1427   -1.291    0.23264

age                1.6299        0.4369     3.731    0.00578 **

---

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

Residual standard error: 10.19 on 8 degrees of freedom

Multiple R-squared:   0.635,      Adjusted R-squared:   0.5894

F-statistic: 13.92 on 1 and 8 DF,    p-value: 0.00578

gaf.fit <- lm(gaf ~ age + pretherapy)

summary(gaf.fit)

Call:

lm(formula = gaf ~ age + pretherapy)

Residuals:

    Min        1Q    Median        3Q        Max

-8.8399 -7.2939 -0.0415    6.6845    9.2966

Coefficients:

                Estimate Std. Error t value Pr(>|t|)

(Intercept) -102.7837      39.6255   -2.594    0.0357 *

age                1.2671        0.3888    3.259    0.0139 *

pretherapy     1.7669        0.7786    2.269    0.0575 .

---

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

Residual standard error: 8.265 on 7 degrees of freedom

Multiple R-squared:   0.7897,    Adjusted R-squared:   0.7297
```

F-statistic: 13.15 on 2 and 7 DF,   p-value: 0.004263

mod.1 <- lm(gaf ~ age)

summary(mod.1)

Call:

lm(formula = gaf ~ age)

Residuals:

     Min   1Q   Median   3Q   Max

-13.657   -9.196   4.359   7.617   9.784

Coefficients:

                Estimate Std. Error t value Pr(>|t|)

(Intercept) -15.6806   12.1427   -1.291   0.23264

age   1.6299   0.4369   3.731   0.00578 **

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

Residual standard error: 10.19 on 8 degrees of freedom

Multiple R-squared: 0.635,   Adjusted R-squared: 0.5894

F-statistic: 13.92 on 1 and 8 DF,   p-value: 0.00578

mod.2 <- lm(gaf ~ age + pretherapy)

summary(mod.2)

Call:

lm(formula = gaf ~ age + pretherapy)

Residuals:

     Min   1Q   Median   3Q   Max

-8.8399 -7.2939 -0.0415   6.6845   9.2966

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -102.7837   39.6255   -2.594   0.0357 *

age   1.2671   0.3888   3.259   0.0139 *

pretherapy   1.7669   0.7786   2.269   0.0575 .

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

Residual standard error: 8.265 on 7 degrees of freedom

Multiple R-squared: 0.7897,   Adjusted R-squared: 0.7297

F-statistic: 13.15 on 2 and 7 DF,   p-value: 0.004263

library(MASS)

gaf.fit <- lm(time ~ distance + case, data = delivery)

step <- stepAIC(gaf.fit, direction = "both")

Start: AIC=61.88

time ~ distance + case

	Df	Sum of Sq	RSS	AIC
<none>			233.73	61.882
- distance	1	168.40	402.13	73.448
- case	1	951.66	1185.39	100.474