library(ISLR)

#9a

pairs(Auto)

#9b

> cor(subset(Auto, select=-name))

mpg cylinders displacement horsepower weight acceleration year origin

mpg 1.0000000 -0.7776175 -0.8051269 -0.7784268 -0.8322442 0.4233285 0.5805410 0.5652088

cylinders -0.7776175 1.0000000 0.9508233 0.8429834 0.8975273 -0.5046834 -0.3456474 -0.5689316

displacement -0.8051269 0.9508233 1.0000000 0.8972570 0.9329944 -0.5438005 -0.3698552 -0.6145351

horsepower -0.7784268 0.8429834 0.8972570 1.0000000 0.8645377 -0.6891955 -0.4163615 -0.4551715

weight -0.8322442 0.8975273 0.9329944 0.8645377 1.0000000 -0.4168392 -0.3091199 -0.5850054

acceleration 0.4233285 -0.5046834 -0.5438005 -0.6891955 -0.4168392 1.0000000 0.2903161 0.2127458

year 0.5805410 -0.3456474 -0.3698552 -0.4163615 -0.3091199 0.2903161 1.0000000 0.1815277

origin 0.5652088 -0.5689316 -0.6145351 -0.4551715 -0.5850054 0.2127458 0.1815277 1.0000000

#9c

> lm.fit1 = lm(mpg~.-name, data=Auto)

> summary(lm.fit1)

Call:

lm(formula = mpg ~ . - name, data = Auto)

Residuals:

Min 1Q Median 3Q Max

-9.5903 -2.1565 -0.1169 1.8690 13.0604

Coefficients:

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

(Intercept) -17.218435 4.644294 -3.707 0.00024 \*\*\*

cylinders -0.493376 0.323282 -1.526 0.12780

displacement 0.019896 0.007515 2.647 0.00844 \*\*

horsepower -0.016951 0.013787 -1.230 0.21963

weight -0.006474 0.000652 -9.929 < 2e-16 \*\*\*

acceleration 0.080576 0.098845 0.815 0.41548

year 0.750773 0.050973 14.729 < 2e-16 \*\*\*

origin 1.426141 0.278136 5.127 4.67e-07 \*\*\*

---

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

Residual standard error: 3.328 on 384 degrees of freedom

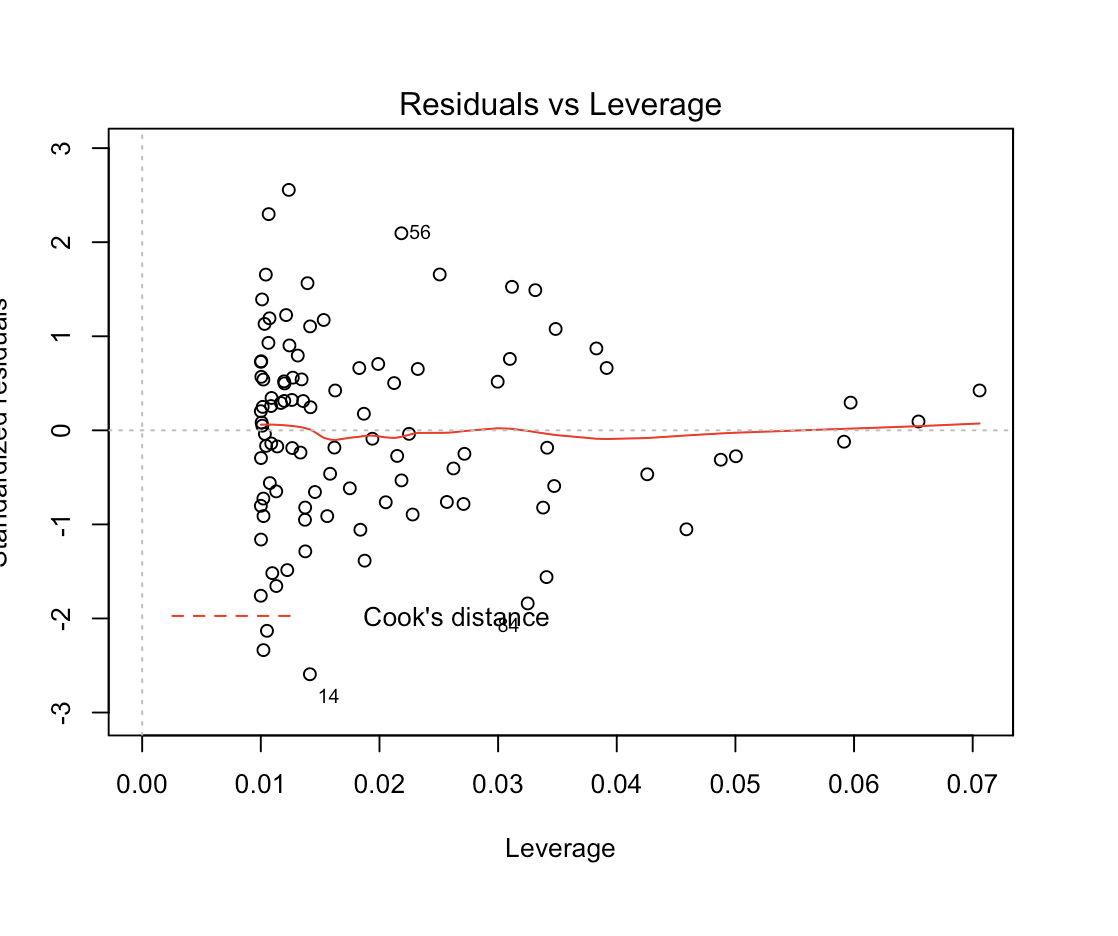
Multiple R-squared: 0.8215, Adjusted R-squared: 0.8182

F-statistic: 252.4 on 7 and 384 DF, p-value: < 2.2e-16

#9d

> par(mfrow=c(2,2))

> plot(lm.fit1)



#9e

lm.fit2 = lm(mpg~cylinders\*displacement+displacement\*weight)

summary(lm.fit2)

#9f

lm.fit3 = lm(mpg~log(weight)+sqrt(horsepower)+acceleration+I(acceleration^2))

summary(lm.fit3)

par(mfrow=c(2,2))

plot(lm.fit3)

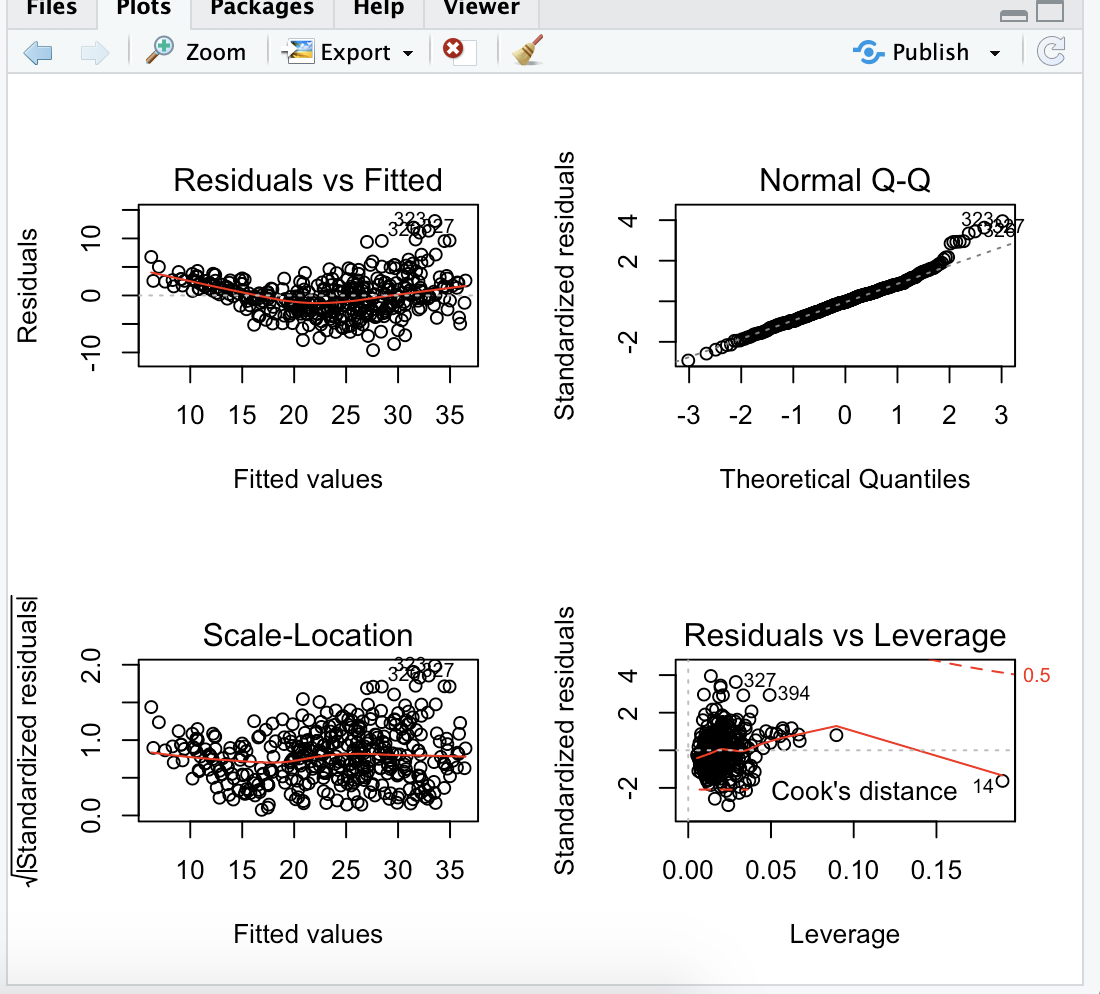
lm.fit2<-lm(log(mpg)~cylinders+displacement+horsepower+weight+acceleration+year+origin,data=Auto)

summary(lm.fit2)

par(mfrow=c(2,2))

plot(lm.fit2)

plot(predict(lm.fit2),rstudent(lm.fit2))



#13a

> set.seed(1)

> x = rnorm(100)

#13b

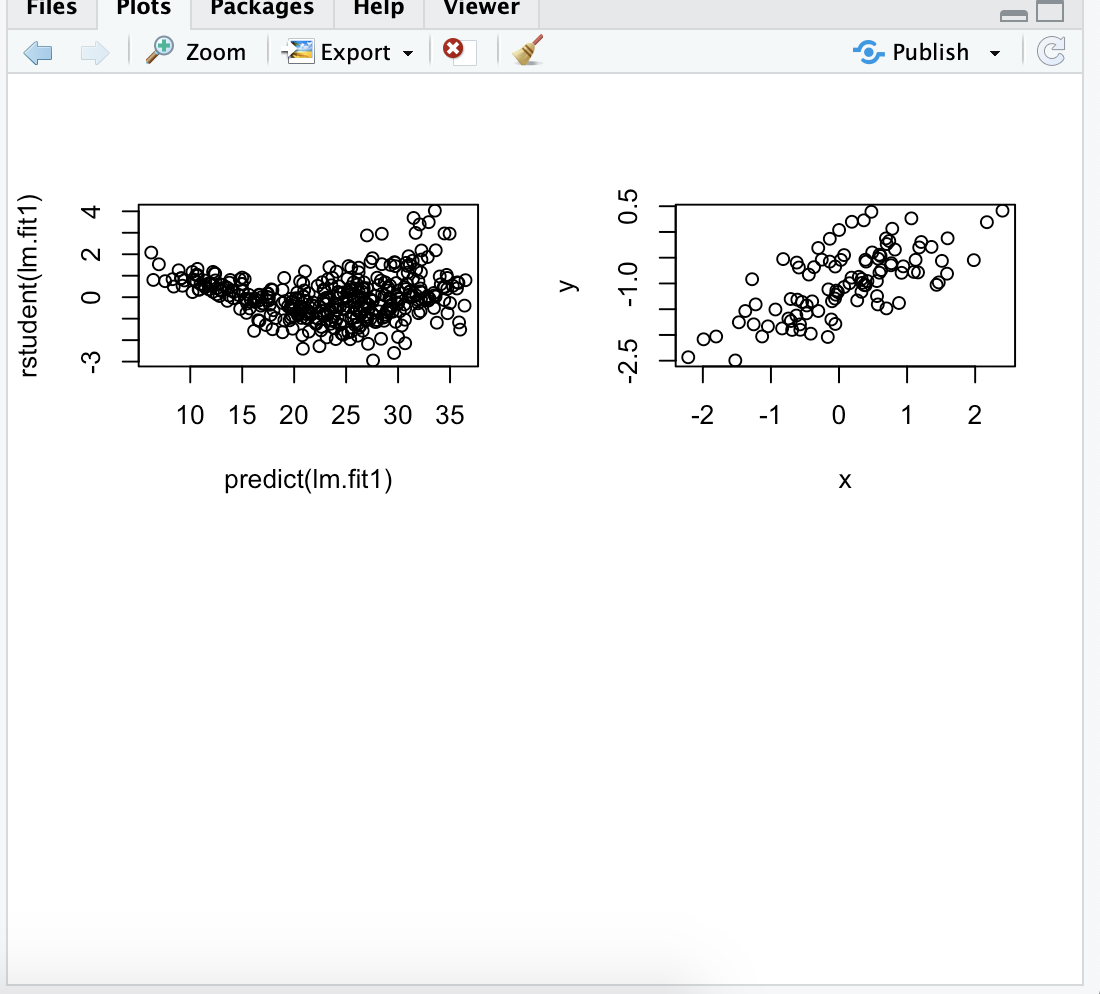
> eps = rnorm(100, 0, sqrt(0.25))

#13c

> y = -1 + 0.5\*x + eps

#13d

> plot(x, y)



#13e

> lm.fit = lm(y~x)

> summary(lm.fit)

Call:

lm(formula = y ~ x)

Residuals:

Min 1Q Median 3Q Max

-0.93842 -0.30688 -0.06975 0.26970 1.17309

Coefficients:

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

(Intercept) -1.01885 0.04849 -21.010 < 2e-16 \*\*\*

x 0.49947 0.05386 9.273 4.58e-15 \*\*\*

---

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

Residual standard error: 0.4814 on 98 degrees of freedom

Multiple R-squared: 0.4674, Adjusted R-squared: 0.4619

F-statistic: 85.99 on 1 and 98 DF, p-value: 4.583e-15

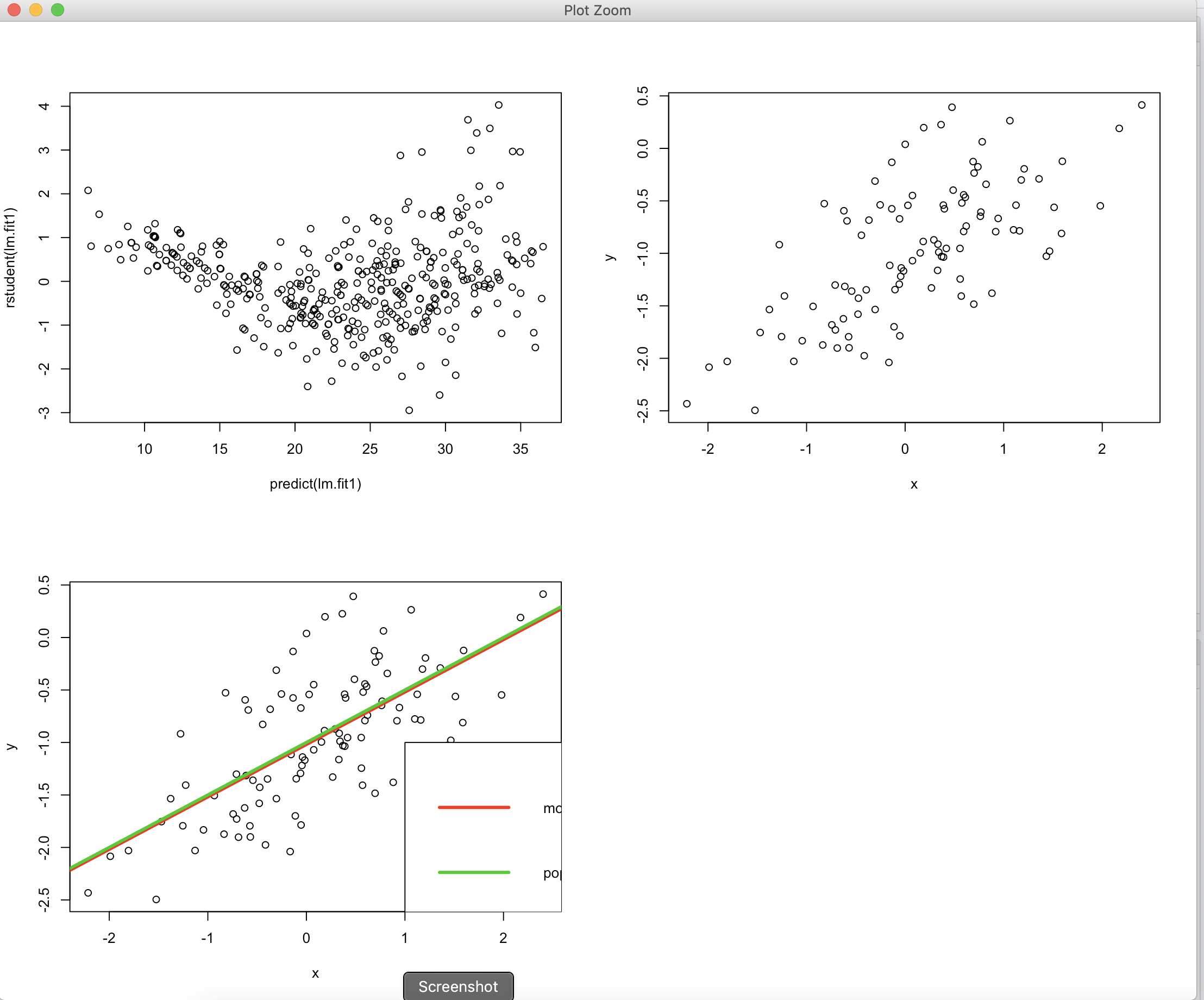
#13f

> plot(x, y)

> abline(lm.fit, lwd=3, col=2)

> abline(-1, 0.5, lwd=3, col=3)

> legend(-1, legend = c("model fit", "pop. regression"), col=2:3, lwd=3)



#13g

> lm.fit\_sq = lm(y~x+I(x^2))

> summary(lm.fit\_sq)

Call:

lm(formula = y ~ x + I(x^2))

Residuals:

Min 1Q Median 3Q Max

-0.98252 -0.31270 -0.06441 0.29014 1.13500

Coefficients:

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

(Intercept) -0.97164 0.05883 -16.517 < 2e-16 \*\*\*

x 0.50858 0.05399 9.420 2.4e-15 \*\*\*

I(x^2) -0.05946 0.04238 -1.403 0.164

---

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

Residual standard error: 0.479 on 97 degrees of freedom

Multiple R-squared: 0.4779, Adjusted R-squared: 0.4672

F-statistic: 44.4 on 2 and 97 DF, p-value: 2.038e-14

#13h

> set.seed(1)

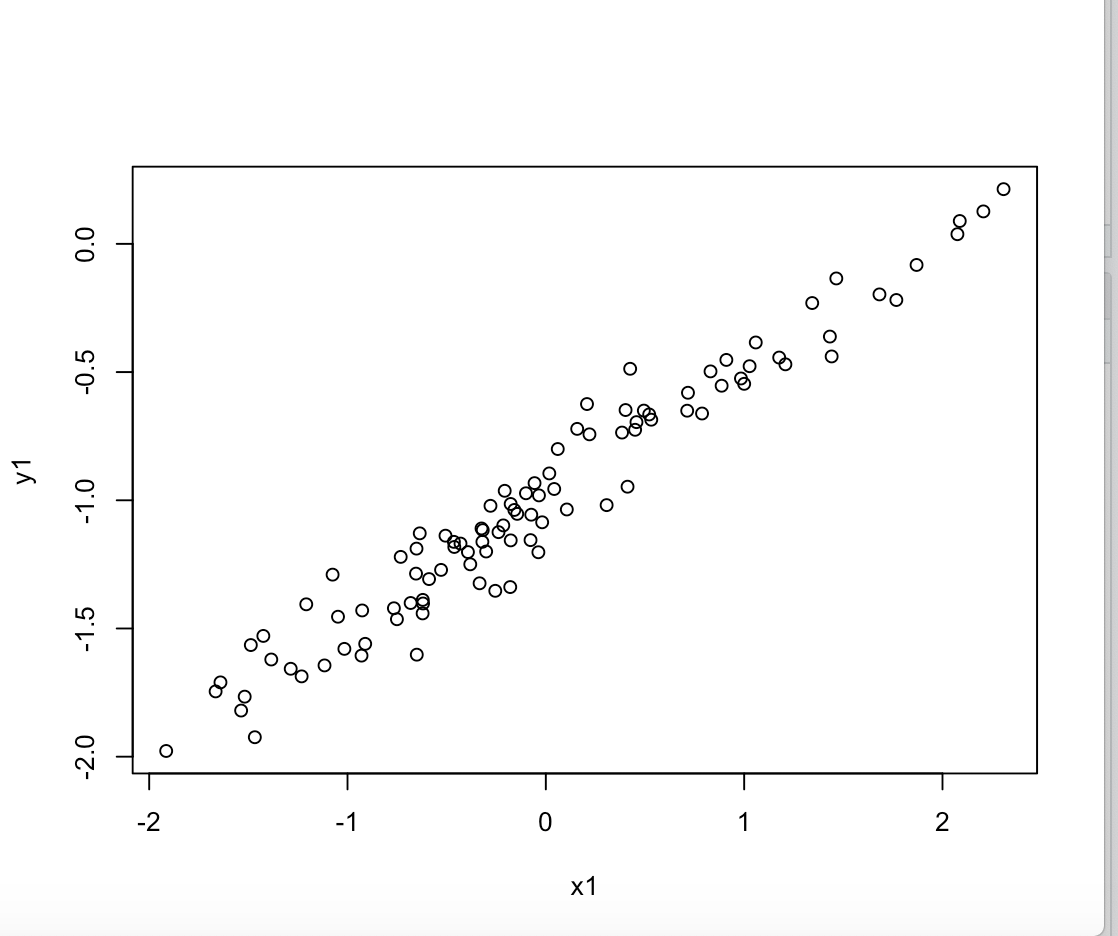
> eps1 = rnorm(100, 0, 0.125)

> x1 = rnorm(100)

> y1 = -1 + 0.5\*x1 + eps1

> plot(x1, y1)

> lm.fit1 = lm(y1~x1)



> summary(lm.fit1)

Call:

lm(formula = y1 ~ x1)

Residuals:

Min 1Q Median 3Q Max

-0.29052 -0.07545 0.00067 0.07288 0.28664

Coefficients:

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

(Intercept) -0.98639 0.01129 -87.34 <2e-16 \*\*\*

x1 0.49988 0.01184 42.22 <2e-16 \*\*\*

---

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

Residual standard error: 0.1128 on 98 degrees of freedom

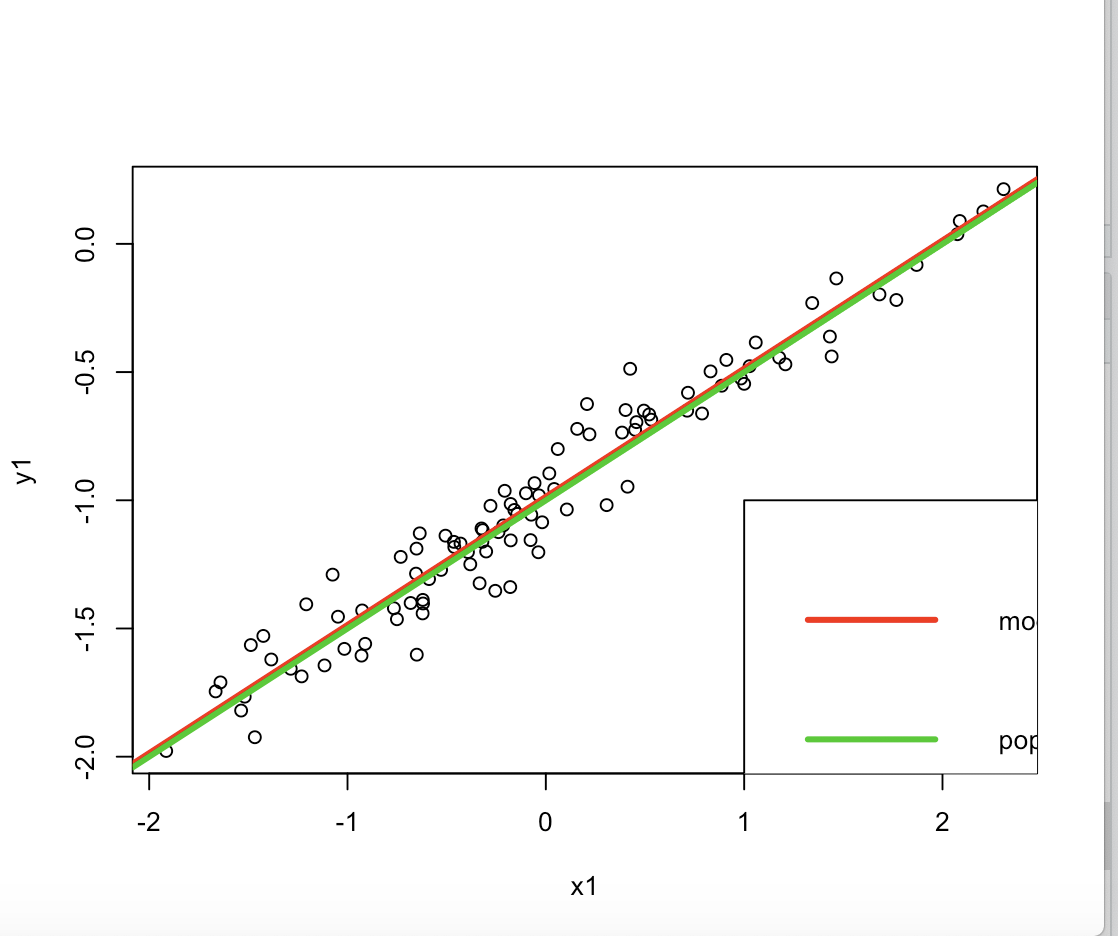
Multiple R-squared: 0.9479, Adjusted R-squared: 0.9474

F-statistic: 1782 on 1 and 98 DF, p-value: < 2.2e-16

> abline(lm.fit1, lwd=3, col=2)

> abline(-1, 0.5, lwd=3, col=3)

> legend(-1, legend = c("model fit", "pop. regression"), col=2:3, lwd=3)



#13i

> set.seed(1)

> eps2 = rnorm(100, 0, 0.5)

> x2 = rnorm(100)

> y2 = -1 + 0.5\*x2 + eps2

> plot(x2, y2)

> lm.fit2 = lm(y2~x2)

> summary(lm.fit2)

Call:

lm(formula = y2 ~ x2)

Residuals:

Min 1Q Median 3Q Max

-1.16208 -0.30181 0.00268 0.29152 1.14658

Coefficients:

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

(Intercept) -0.94557 0.04517 -20.93 <2e-16 \*\*\*

x2 0.49953 0.04736 10.55 <2e-16 \*\*\*

---

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

Residual standard error: 0.4514 on 98 degrees of freedom

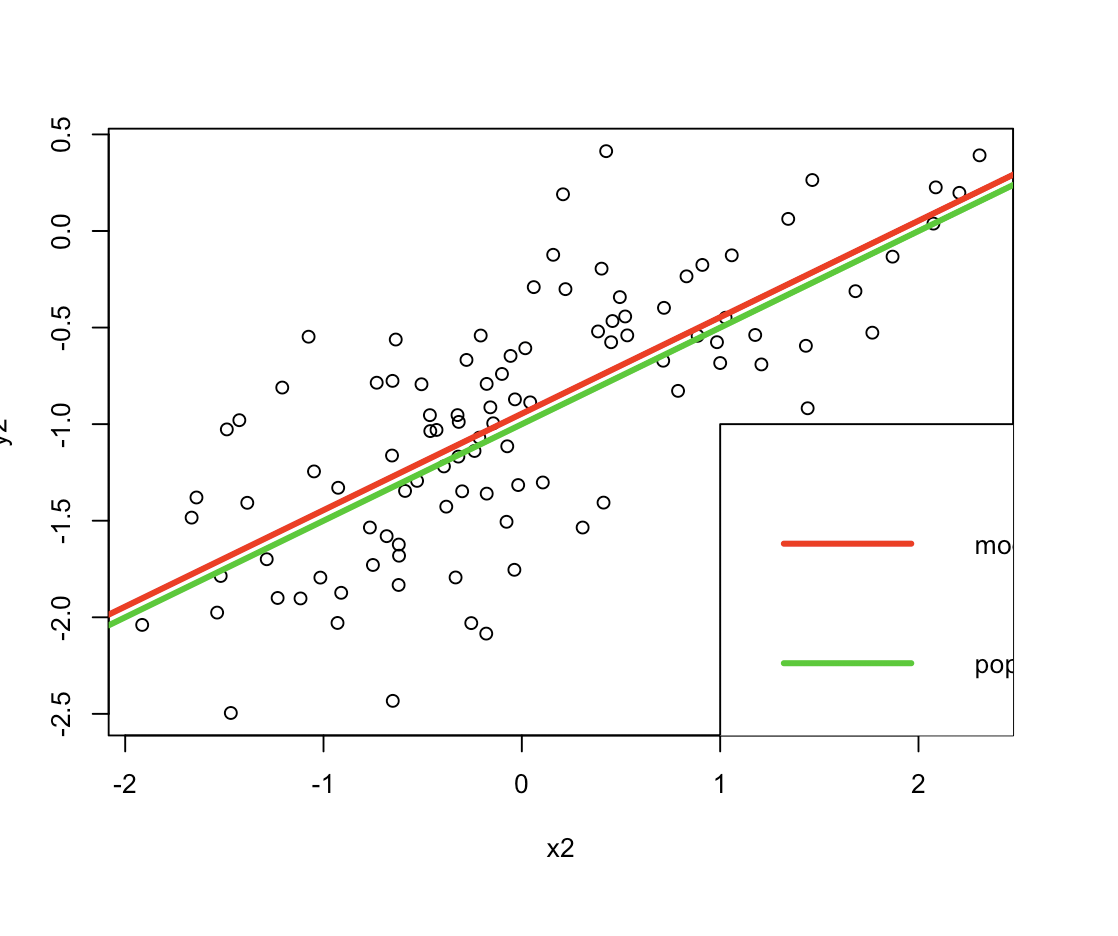
Multiple R-squared: 0.5317, Adjusted R-squared: 0.5269

F-statistic: 111.2 on 1 and 98 DF, p-value: < 2.2e-16

> abline(lm.fit2, lwd=3, col=2)

> abline(-1, 0.5, lwd=3, col=3)

> legend(-1, legend = c("model fit", "pop. regression"), col=2:3, lwd=3)



#13j

> confint(lm.fit)

2.5 % 97.5 %

(Intercept) -1.1150804 -0.9226122

x 0.3925794 0.6063602

> confint(lm.fit1)

2.5 % 97.5 %

(Intercept) -1.008805 -0.9639819

x1 0.476387 0.5233799

> confint(lm.fit2)

2.5 % 97.5 %

(Intercept) -1.0352203 -0.8559276

x2 0.4055479 0.5935197