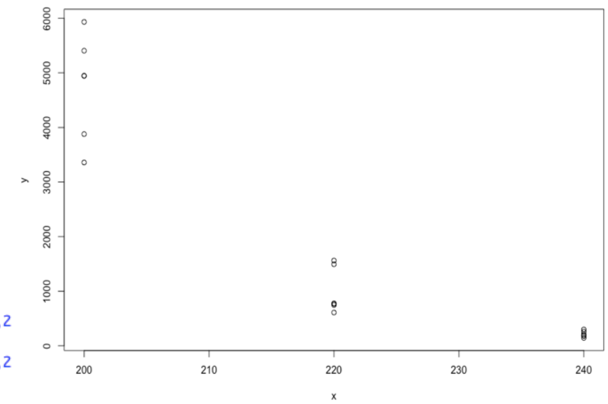
13.2.19 Thermal endurance tests were performed to study the relationship between temperature and lifetime of polyester enameled wire **("Thermal Endurance of Polyester Enameled Wires Using Twisted Wire Specimens," *IEEE Trans. Insulation*, 1965: 38–44)**, resulting in the following data.

Temp.	200	200	200	200	200	200
Lifetime	5933	5404	4947	4963	3358	3878
Temp.	220	220	220	220	220	220
Lifetime	1561	1494	747	768	609	777
Temp.	240	240	240	240	240	240
Lifetime	258	299	209	144	180	184

(a) Does a scatterplot of the data suggest a linear probabilistic relationship between lifetime and temperature?

```
> x<-c(200,200,200,200,200,200,220,220,220,220,220,220,240,240,240,240,2
40,240)
> y<-c(5933,5405,4947,4943,3358,3878,1561,1494,747,768,609,777,258,299,2
09,144,180,184)
> plot(y~x)
```

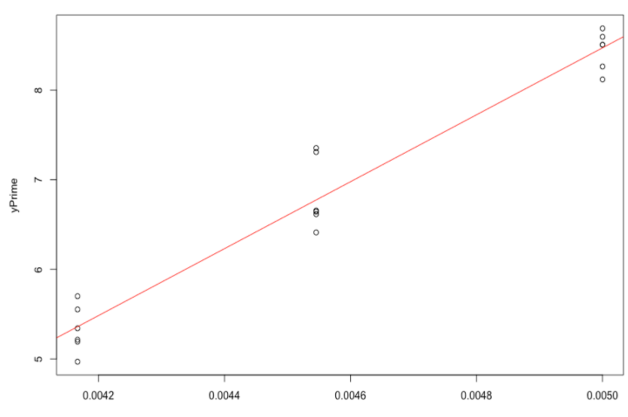


The scatterplot does not suggest linear relationship between variables, no curvature either.

Ans: No

(b) What model is implied by a linear relationship between expected ln(lifetime) and 1/temperature? Does a scatterplot of the transformed data appear consistent with this relationship?

```
> xPrime<-1/x
> yPrime<-log(y)
> plot(yPrime~xPrime)
> fit<-lm(yPrime~xPrime)
> abline(fit,col="red")
```



Ans: $Y' = \beta_0 + \beta_1 \cdot (1/t) + \epsilon'$**, where** $Y' = \ln(Y)$**, so** $Y = \alpha e^{\beta/t} \cdot \epsilon$

Ans: Yes, the scatterplot of the transformed data appear consistent with this relationship.

(c) Estimate the parameters of the model suggested in part (b). What lifetime would you predict for a temperature of 220?

```
> summary(fit)

Call:
lm(formula = yPrime ~ xPrime)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.39016 -0.16490 -0.06841  0.17546  0.57847 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -10.2011     0.9366  -10.89 8.26e-09 ***
xPrime      3734.6594   204.3376   18.28 3.82e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2953 on 16 degrees of freedom
Multiple R-squared:  0.9543,	Adjusted R-squared:  0.9514 
F-statistic:   334 on 1 and 16 DF,  p-value: 3.82e-12
```

```
> exp(-10.2011+3734.6594/220)
[1] 875.3506
```

Ans: $\boldsymbol{\hat{\beta} = \hat{\beta}_1 = 3734.6594, \hat{\beta}_0 = -10.2011, \hat{y} = 875.3506}$

(d) Because there are multiple observations at each *x* value, the method in Exercise 14 can be used to test the null hypothesis that states that the model suggested in part (b) is correct. Carry out the test at level .01.

From Exercise 14:

$H_0: \mu_{Y \cdot x} = \beta_0 + \beta_1 x$ for some values β_0, β_1 (the true regression function is linear

$H_a: H_0$ is not true (the true regression function is not linear)

$$\text{SSE} = \text{SSPE} + \text{SSLF}, \text{SSPE} = \sum\sum Y_{ij}^2 - \sum\sum n_i \bar{Y}_{i\cdot}^2$$

```
> y1HatPrime<-mean(yPrime[1:6])
> y2HatPrime<-mean(yPrime[7:12])
> y3HatPrime<-mean(yPrime[13:18])
> sspe<-sum(yPrimeSQ[1:18])-(6*y1HatPrime^2+6*y2HatPrime^2+6*y3HatPrime^
2)
> sspe
[1] 1.366356
> xPrimeyPrime<-xPrime*yPrime
> sse<-sum(yPrimeSQ[1:18])-(-10.2011)*sum(yPrime[1:18])-3734.6594*sum(xP
rimeyPrime)
> sse
[1] 1.394542
> sslf<-sse-sspe
> sslf
[1] 0.02818591
> f<-(sslf/1)/(sspe/15)
> f
[1] 0.3094278
```

Conclusion: Do not reject H_0 (f = 0.30942778 < 8.53 (Table A.9)

Ans: A linear model is appropriate

13.3.30 The accompanying data was extracted from the article **"Effects of Cold and Warm Temperatures on Springback of Aluminum-Magnesium Alloy 5083-H111" (*J. of Engr. Manuf.*, 2009: 427–431)**. The response variable is yield strength (MPa), and the predictor is temperature (°C).

x	–50	25	100	200	300
y	91.0	120.5	136.0	133.1	120.8

Here is Minitab output from fitting the quadratic regression model (a graph in the cited paper suggests that the authors did this):

```
Predictor        Coef    SE Coef      T      P
Constant      111.277      2.100  52.98  0.000
temp          0.32845    0.03303   9.94  0.010
tempsqd    −0.0010050  0.0001213  −8.29  0.014

S = 3.44398   R-Sq = 98.1%     R-Sq(adj) = 96.3%

Analysis of Variance

Source          DF       SS      MS      F      P
Regression       2  1245.39  622.69  52.50  0.019
Residual Error   2    23.72   11.86
Total            4  1269.11
```

(a) What proportion of observed variation in strength can be attributed to the model relationship?

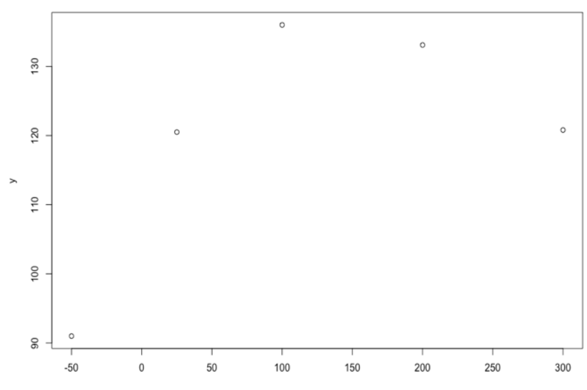
```
> x<-c(-50,25,100,200,300)
> y<-c(91,120.5,136,133.1,120.8)
> xSQ<-x^2
> fit<-lm(y~x+xSQ)
> plot(y~x)
> summary(fit)

Call:
lm(formula = y ~ x + xSQ)

Residuals:
     1      2      3      4      5
-1.342  1.639  1.928 -3.666  1.441

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.113e+02  2.100e+00  52.979 0.000356 ***
x            3.285e-01  3.303e-02   9.944 0.009962 **
xSQ         -1.005e-03  1.213e-04  -8.286 0.014255 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.444 on 2 degrees of freedom
Multiple R-squared:  0.9813,    Adjusted R-squared:  0.9626
F-statistic:  52.5 on 2 and 2 DF,  p-value: 0.01869
```



Ans: 98.13% of observed variation in density can be attributed to the model relationship

(b) Carry out a test of hypotheses at significance level .05 to decide if the quadratic predictor provides useful information over and above that provided by the linear predictor.

$H_0: \beta_2 = 0; H_a: \beta_2 \neq 0$. From the screenshot in (a), t = –8.286, P-value = 0.014 < 0.05 (reject H_0)

Ans: The quadratic predictor provides useful information

(c) For a strength value of 100, $\hat{y} = 134.07, s_{\hat{y}} = 2.38$. Estimate true average strength when temperature is 100, in a way that conveys information about precision and reliability.

```
> predict(fit,newdata=data.frame(x=100,xSQ=10000),interval='confidence')
       fit      lwr      upr
1 134.0721 123.8475 144.2967
```

Ans: 95% CI = (123.8475, 144.2967)

(d) Use the information in (c) to predict strength for a single observation to be made when temperature is 100, and do so in a way that conveys information about precision and reliability. Then compare this prediction to the estimate obtained in (c).

```
> predict(fit,newdata=data.frame(x=100,xSQ=10000),interval='prediction')
       fit      lwr      upr
1 134.0721 116.0687 152.0755
```

Ans: 95% PI = (116.0687, 152.0755)

13.4.43 An experiment carried out to study the effect of the mole contents of cobalt (x_1) and the calcination temperature (x_2) on the surface area of an iron-cobalt hydroxide catalyst (y) resulted in the accompanying data **("Structural Changes and Surface Properties of $Co_xFe_{3-x}O_4$ Spinels," *J. of Chemical Tech. and Biotech.*, 1994: 161–170)**. A request to the SAS package to fit $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$, where $x_3 = x_1 x_2$ (an interaction predictor) yields the output below.

x_1	.6	.6	.6	.6	.6	1.0	1.0
x_2	200	250	400	500	600	200	250
y	90.6	82.7	58.7	43.2	25.0	127.1	112.3

x_1	1.0	1.0	1.0	2.6	2.6	2.6	2.6
x_2	400	500	600	200	250	400	500
y	19.6	17.8	9.1	53.1	52.0	43.4	42.4

x_1	2.6	2.8	2.8	2.8	2.8	2.8
x_2	600	200	250	400	500	600
y	31.6	40.9	37.9	27.5	27.3	19.0

(a) Predict the value of surface area when cobalt content is 2.6 and temperature is 250, and calculate the value of the corresponding residual.

```
> x1<-c(.6,.6,.6,.6,.6,1,1,1,1,1,2.6,2.6,2.6,2.6,2.6,2.8,2.8,2.8,2.8,2.8
)
> x2<-c(200,250,400,500,600,200,250,400,500,600,200,250,400,500,600,200,
250,400,500,600)
> y<-c(90.6,82.7,58.7,43.2,25,127.1,112.3,19.6,17.8,9.1,53.1,52,43.4,42.
4,31.6,40.9,37.9,27.5,27.3,19)
> x1x2<-x1*x2
> fit<-lm(y~x1+x2+x1x2)
> fit

Call:
lm(formula = y ~ x1 + x2 + x1x2)

Coefficients:
(Intercept)           x1           x2         x1x2
   185.4857     -45.9695      -0.3015       0.0888

> predict(fit,newdata=data.frame(x1=2.6,x2=250,x1x2=2.6*250))
       1
48.31031
> 52.0 - predict(fit,newdata=data.frame(x1=2.6,x2=250,x1x2=2.6*250))
      1
3.68969
```

Ans: $\hat{y} = 48.31031$, $\epsilon = 3.68969$

(b) Since $\hat{\beta}_1 = -46.0$, is it legitimate to conclude that if cobalt content increases by 1 unit while the values of the other predictors remain fixed, surface area can be expected to decrease by roughly 46 units? Explain your reasoning.

No, because when x_1 changes, x_3 changes too. Therefore you cannot increase the cobalt content while keeping x_3 constant.

Ans: No. If x_1 increases, either x_3 or x_2 must change.

(c) Does there appear to be a useful linear relationship between y and the predictors?

```
> summary(fit)

Call:
lm(formula = y ~ x1 + x2 + x1x2)

Residuals:
    Min      1Q  Median      3Q     Max 
-34.836  -6.574  -1.355   8.357  30.124 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 185.48574   21.19748   8.750 1.70e-07 ***
x1          -45.96947   10.61201  -4.332 0.000515 ***
x2           -0.30150    0.05074  -5.942 2.07e-05 ***
x1x2          0.08880    0.02540   3.496 0.002991 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16.38 on 16 degrees of freedom
Multiple R-squared:  0.7801,	Adjusted R-squared:  0.7389 
F-statistic: 18.92 on 3 and 16 DF,  p-value: 1.635e-05
```

Ans: Yes, since f = 18.92, P-value = 1.635e-05

(d) Given that mole contents and calcination temperature remains in the model, does the interaction predictor x_3 provide useful information about y? State and test the appropriate hypotheses using a significance level of .01.

Ans: Yes, t = 3.496 and P-value = .002991 < 0.01 (from screenshot of (c))

(e) The estimated standard deviation of $\hat{Y}$ when mole contents is 2.0 and calcination temperature is 500 is $s_{\hat{y}} = 4.69$. Calculate a 95% confidence interval for the mean value of surface area under these circumstances.

```
> predict(fit,newdata=data.frame(x1=2.0,x2=500,x1x2=2.0*500),interval='c
onfidence')
       fit    lwr      upr
1 31.59676 21.644 41.54952
```

Ans: 95% CI = (21.644, 41.54952)