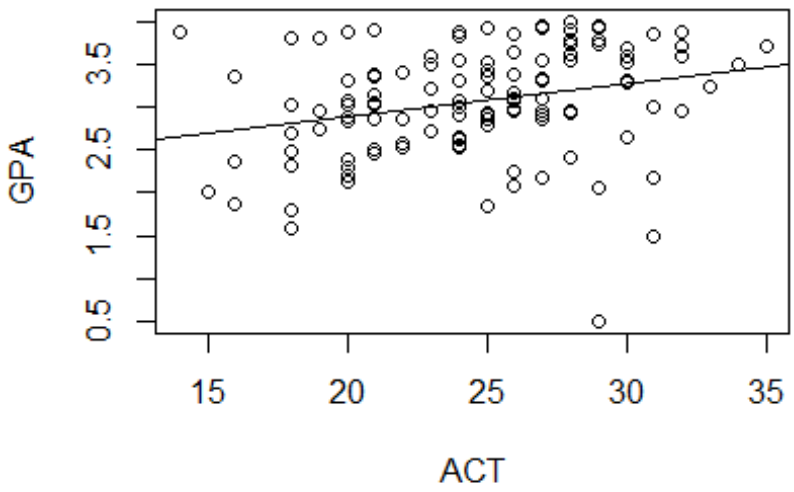
Chapter 1 Question 19 Michael Streyle

```
##
## Call:
## lm(formula = GPA ~ ACT)
##
## Coefficients:
## (Intercept)         ACT
##     2.11405      0.03883

## Analysis of Variance Table
##
## Response: GPA
##           Df Sum Sq Mean Sq F value   Pr(>F)
## ACT        1  3.588  3.5878  9.2402 0.002917 **
## Residuals 118 45.818  0.3883
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
## (Intercept)
##    3.278863

## (Intercept)
##     3.31769
```

The intercept is 2.11 GPA units and the slope is 0.0388 GPA unit per ACT score increment.

Answers Written Out

19)

a) The least squares estimate of β_0 and β_1 are 2.114049 and 0.03882713 respectively, so the estimated regression function is $\hat{Y} = 2.114049 + 0.03882713(X_i)$

b) The best fit line fits the data pretty well, considering the distribution of the data points.

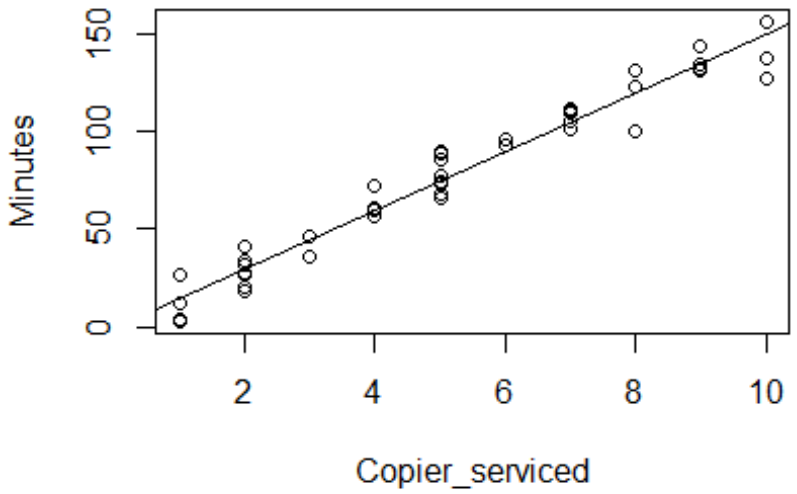
c) For a student with an ACT score of 30, the point estimate of the mean freshman GPA is 3.278863.

d) When the entry ACT score increases to 31, the point estimate of the mean freshman GPA is 3.31769. This means that the mean freshman GPA changes by 0.038827.

Chapter 1 Question 20 Michael Streyle

```
##
## Call:
## lm(formula = Minutes ~ Copier_serviced)
##
## Coefficients:
##     (Intercept)  Copier_serviced
##         -0.5802          15.0352

## Analysis of Variance Table
##
## Response: Minutes
##                Df Sum Sq Mean Sq F value    Pr(>F)
## Copier_serviced  1  76960   76960  968.66 < 2.2e-16 ***
## Residuals       43   3416      79
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
## (Intercept)
##    74.59608
```

The intercept is -0.58 minutes and the slope is 15.0352 minutes unit per copier serviced.

Answers Written Out #20

a) The estimated regression function is $\hat{Y} = -0.5802 + 15.0352(X)$.

b) The estimated regression function appears to fit the data very well.

c) The β_0 value is -0.5802 which means for 0 copier machines, it takes -0.5802 minutes of service. This doesn't provide any relevant information here, because if there are no machines to be serviced, it would take zero minutes. Also, Y is in minutes and negative minutes are not very plausible.

d) When X = 5, the point estimate of the mean service time is 74.59608 minutes.

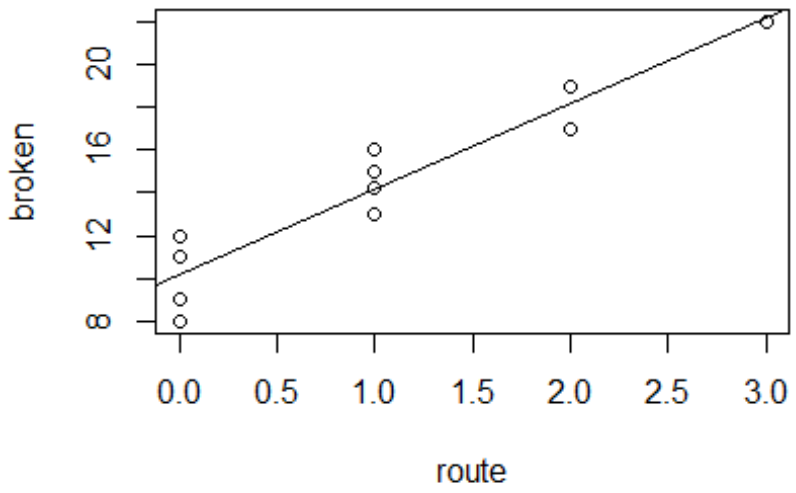
Chapter 1 Problem 21 Michael Streyle

```
##
## Call:
## lm(formula = broken ~ route)
##
## Coefficients:
## (Intercept)       route
##        10.2         4.0

## (Intercept)
##        14.2

## (Intercept)
##        18.2

## (Intercept)
##          4
```



The intercept is 10.2 broken ampules and the slope is 4 broken ampules per route.

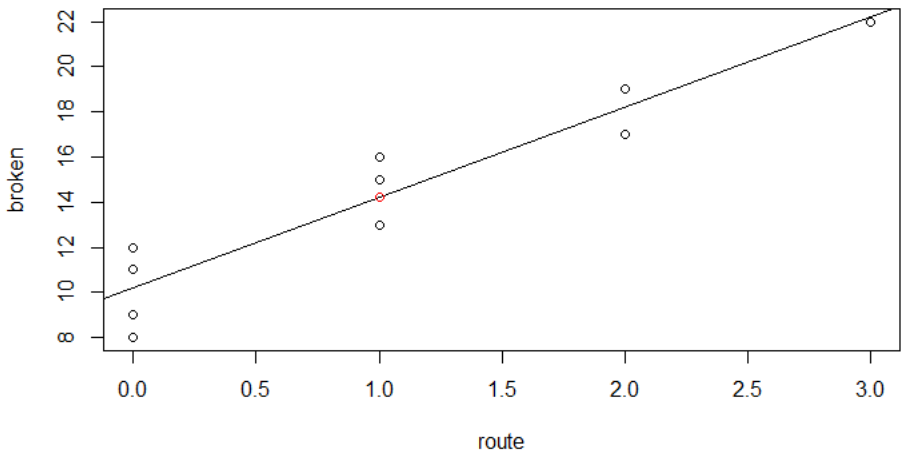
Written Answers to Question 21

a) The estimated regression function is $\hat{Y}$ = 10.2 + 4(X). The linear regression line does seem to give a good fit here – the line follows the data very well.

b) The point estimate of the expected number of broken ampules when X = 1 is 14.2.

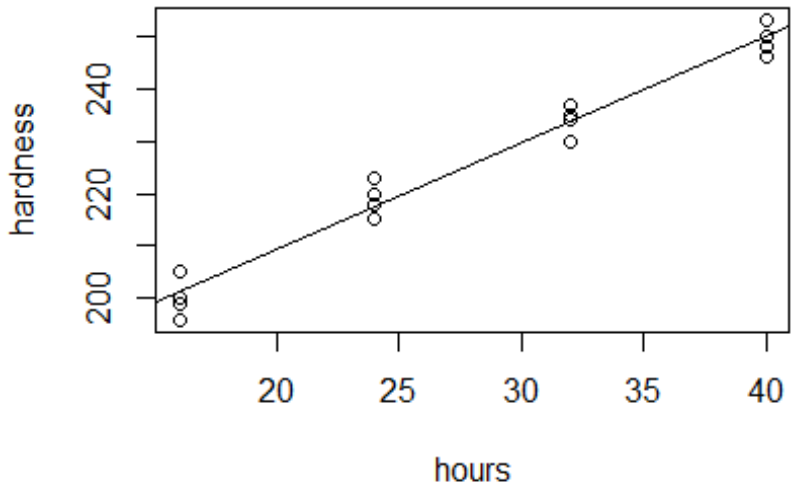
c) When there are two transfers made, the expected number of broken ampules is 18.2 so the difference between one transfer and two transfers is 4.0.

d) By plotting the point ($\acute{X}, \acute{Y}$) on the plot of the data and best fit line, it shows the point is on the fitted regression line. See below in red.



Chapter 1 Problem 22 Michael Streyle

```
##
## Call:
## lm(formula = hardness ~ hours)
##
## Coefficients:
## (Intercept)       hours
##     168.600       2.034
```




```
## (Intercept)
##     249.975

## (Intercept)
##    252.0094

## (Intercept)
##    2.034375
```

The intercept is 168.6 Hardness units and the slope is 2.0344 hardness units per hour.

Written Out Answers to Question 22

a) The estimated regression function is $\hat{Y} = 168.6 + 2.0344(X)$. The regression line does seem to have a good fit for the data – it follows the data points well.

b) A point estimate of the mean hardness when X = 40 hours is 249.975.

c) The point estimate of the mean hardness when X = 41 hours is 252.0094. The difference between when X = 40 and when X = 41 is 2.034375.

Answers Written Out for 23

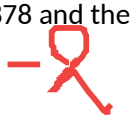
a) Residuals:

fit$residuals

```
         1          2          3          4          5          6          7          8 
 0.96758105  1.22737094  0.57679116 -0.42824608  0.09858105  0.54730978 -0.39451735  0.79861829 
         9         10         11         12         13         14         15         16 
-2.74003597  0.05444541  0.26409967  0.25913691  0.03709967 -0.03290033 -0.15034448 -0.19938171 
        17         18         19         20         21         22         23         24 
 0.43727254 -0.30469022 -0.13772746 -0.77259183 -0.48290033  0.42758105  0.52979116  0.76261829 
        25         26         27         28         29         30         31         32 
 0.35479116 -0.02255459 -0.78120884 -0.38924608  0.74744541  0.13058105  0.84227254 -0.36028332 
        33         34         35         36         37         38         39         40 
-0.27220884  0.25144541 -0.11124608  0.02609967  0.45158105  0.01113691  0.38661829  0.52244541 
        41         42         43         44         45         46         47         48 
-0.14555459 -0.62486309 -0.50590033 -0.87355459 -1.17103597 -0.42890033 -1.13469022 -0.69645619 
        49         50         51         52         53         54         55         56 
 0.10023530  0.99306243 -0.29138171  0.61671668  0.14261829 -0.17155459  0.50109967  0.41213691 
        57         58         59         60         61         62         63         64 
 0.23058105 -0.69659183  0.04413691  0.69596403 -0.16272746 -0.29107321  0.28527254  0.59892679 
        65         66         67         68         69         70         71         72 
-0.63686309 -0.47741895 -0.39090033  0.35748265 -1.00693757  0.50892679  0.14840817 -0.04107321 
        73         74         75         76         77         78         79         80 
-0.33093757 -0.11293757  0.67996403 -0.05659183  0.21492679 -0.03955459  0.79879116  0.07682840 
        81         82         83         84         85         86         87         88 
 0.43240817  0.18140817 -1.04455459  0.51848265  0.12327254 -0.24238171  0.18261829  0.71596403 
        89         90         91         92         93         94         95         96 
 0.95623530 -0.42341895  0.84009967 -0.97938171  0.34427254  0.21106243  0.50996403  0.78709967 
        97         98         99        100        101        102        103        104 
-0.04938171 -0.05441895 -0.10476470 -0.50193757 -1.24372746 -1.22993757 -0.01159183  0.23448265 
       105        106        107        108        109        110        111        112 
-0.13190033  0.24300127 -0.28472746  0.41979116  0.59079116 -0.21772746  0.45075392  0.32113691 
       113        114        115        116        117        118        119        120 
-0.49659183 -0.60459183 -1.83169022  0.99440817  0.55996403  0.71279116 -0.87528332 -0.25320884
```

Sum of Residuals: -2.942091e-15 which is close to 0. Probably not equal to zero because of rounding as it talks about in the textbook.

b) The estimated variance is 3.5878 and the estimated standard deviation is 1.894149. The units of the standard deviation is GPA.



Written Answers for Chapter 1 Problem 24

a) The residuals are as follows:

1	2	3	4	5	6	7	8
-9.4903394	0.4391645	1.4744125	11.5096606	-2.4550914	-12.7723238	-6.5960836	14.4039164
9	10	11	12	13	14	15	16
-10.4550914	2.5096606	9.2629243	6.2276762	3.3686684	-8.5255875	12.4391645	-19.7018277
17	18	19	20	21	22	23	24
0.3334204	11.2981723	-22.7723238	-2.5608355	-8.5960836	-3.6665796	4.3334204	-0.5960836
25	26	27	28	29	30	31	32
-0.7370757	7.3334204	-11.4903394	-1.5960836	6.3334204	6.3686684	3.2981723	15.4039164
33	34	35	36	37	38	39	40
-9.4903394	-1.4903394	-11.4550914	-2.5608355	11.4039164	-2.7370757	7.3334204	12.5449086
41	42	43	44	45			
-3.7370757	4.5096606	-2.4903394	1.4391645	2.4039164			

The sum of the squared residuals is 3416.377. This value is the SSE and it represents the minimum value of Q in (1.8).

b) The variance is the SSE divided by 2-n. The MSE (variance) is 3416.377/43 which is equal to 79.45. The $\sqrt{MSE}$ is equal to 8.91. The MSE is expressed in minutes.

Written Answers to Chapter 1 Problem 25

a) The residual for case 1 is 1.800. Found by using fit$residuals. It is the error value showing how far the data point for case one is from the estimated regression line
b) sum((fit$residuals)^2) gives the value of 17.600, which is the sum of the squared residuals (SSE). The MSE is the SSE/(n-2), which is 2.200. The MSE is an estimate of the variance (σ^2).