Alexander Martinez

Hw 5

9.10 b. The scatterplot shows that each of the predictor variables has atleast somewhat of a linear relationship w/ Y, which is good.

I am worried about Multicollinearity between the X variables, there is high correlation between X_3 and X_4 and also X_2 and X_3

c. $\hat{Y} = -124.38182 + .29573X_1 + .04827X_2 + 1.36601X_3 + .51982X_4$

Only $\hat{\beta}_2$ is not significant, maybe there needs to be further investigation on whether it or others should be retained.

9.11 a.

Subset	$R^2_{a,p}$
X_1, X_3, X_4	.956
X_1, X_2, X_3, X_4	.9555
X_1, X_3	.9269
X_1, X_2, X_3	.9247

b. We can use C_p, BIC, AIC b/c they penalize for adding predictors

9.16 a. Using Backwards elimination, the ones that stayed in the final model are X_1, X_3, and X_4

15.16 a. Done in R (w/ graph at the end)

$|t_i| \leq t^*$, no outliers
if else, outliers

Conclude No outliers

b. Done in R

3, 5, 16, 21, 22, 43, 44, 48 — All High leveraged points

C. Graph Done in R, it doesn't appear to extrapolate.

I got .022217 = h_{new}, new, the book got a different answer; mine are inconsistent

d. Conclude that cases 3, 5, 16, 21, 22, 43, 44 and 48 are still consistent as high influence measures w/ cook's and/or Dffits and/or DFBETAS

f. Done in R, 43, 16, 51, 32, 35, 14.
Hard to read.

10.16 a. There is no obvious correlation between the variables. Highest corr is between X_3 and X_2

b. No, the individual VIF is < 10 by a lot and the mean is less than 6 or 7.

10.24 $H = X(X'X)^{-1}X'$

$= XX^{-1}(X')^{-1}X'$

$= I$

$h_{ii} = I$ and $\hat{Y}_i = Y_i$

10.25 $MSE_{(i)} = \dfrac{\left[\dfrac{(n-p)SSE}{n-p} - \dfrac{e_i^2}{1-h_{ii}}\right]}{(n-p-1)}$

$t_i = \dfrac{e_i}{\sqrt{MSE_{(i)}(1-h_{ii})}} = 10.24a$

$t_i = \dfrac{e_i}{\left[\dfrac{(n-p)SSE}{n-p} - \dfrac{e_i^2}{1-h_{ii}}\right](1-h_{ii})}{(n-p-1)}$

→ on Back

$$e_i \left[\frac{n-p-1}{SSE(1-h_{ii}) - e_i^2} \right]^{1/2}$$

11.1 They are not completely invalid if they depart from homogeneity, unless there are no effective remedial measures. This heterogeneity can throw of Se's and hypothesis testing, which is crucial.

11.2 Incorrect, we are trying to minimize the errors caused by things, such as outliers, so we definitely worry.

11.3 Yes, it is true. We use smoothing to analyze "windows" of relationships between variables, and we cannot seem to do that if the data becomes too sparse.

11.7 a. $\hat{Y} = -5.7500 + .1875X_1$

Graph Done in R

The variance increases w/ speed.

11.7 b. p value = .08312 < .10. Thus is significant departure from the constant σ^2

c. Graph in R
It shows that they variance may be growing faster at higher rates ie, exponential growth

d. $\hat{y} = -186.083 + 1.244x$
weights found in R

e. $\hat{Y} = -6.23322 + .18911X$ (in R)

Yes, except the intercept is a little more negative

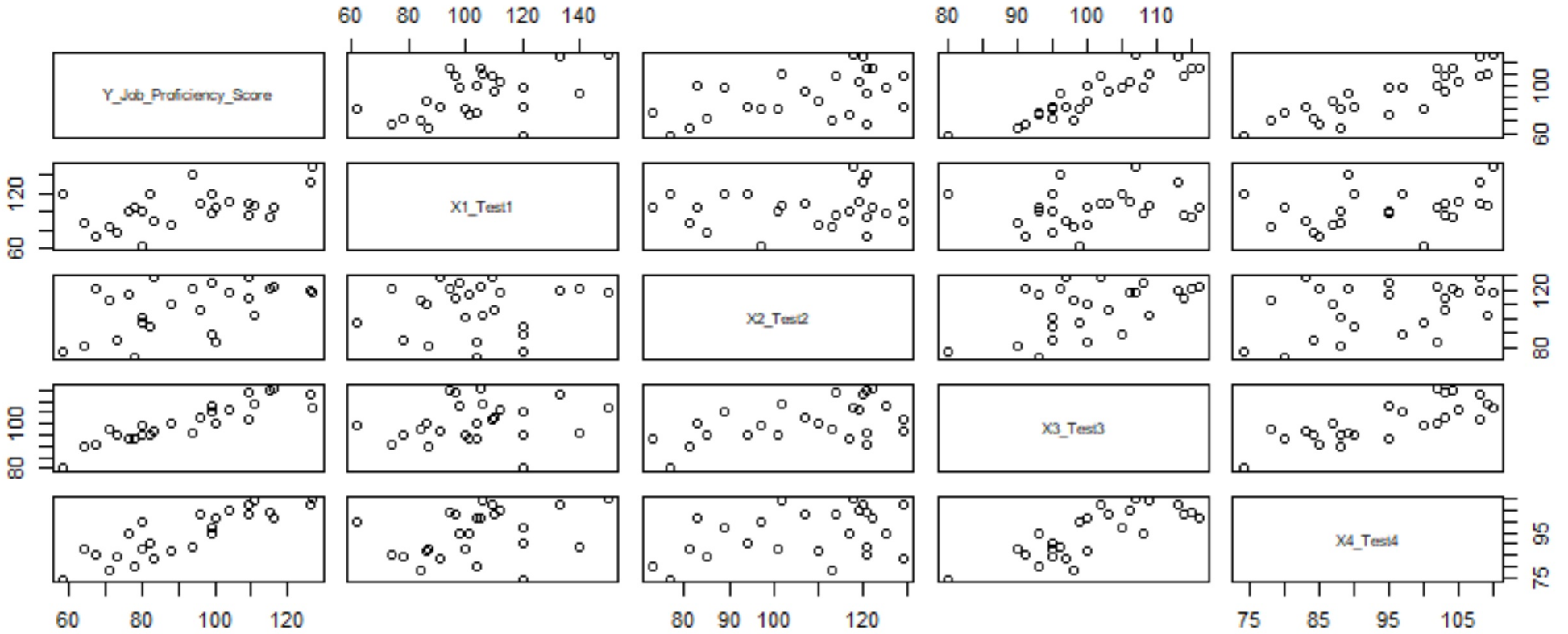
f. Done in R

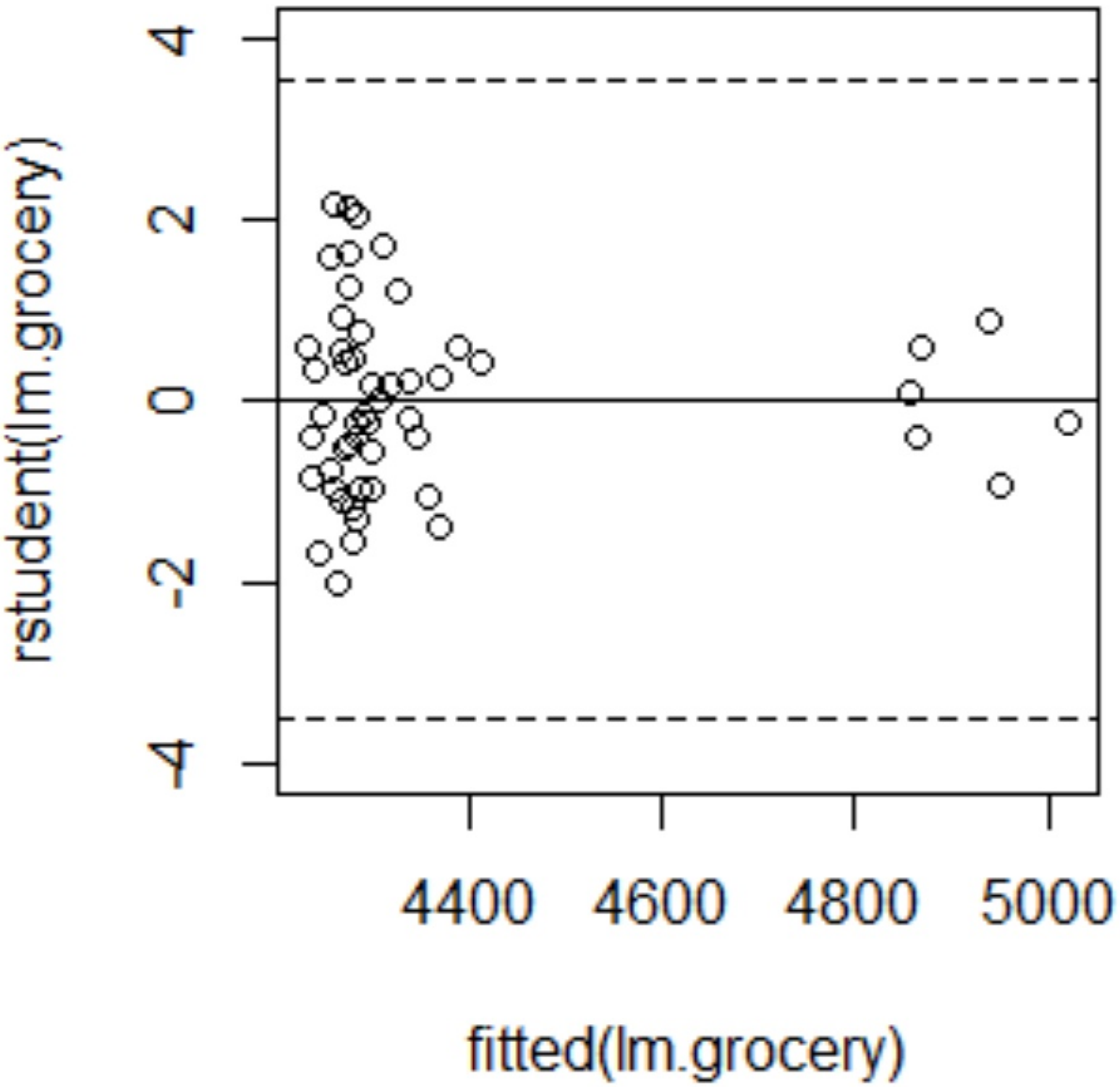
$Se_0 = 16.73052$
$Se_1 = .05381$

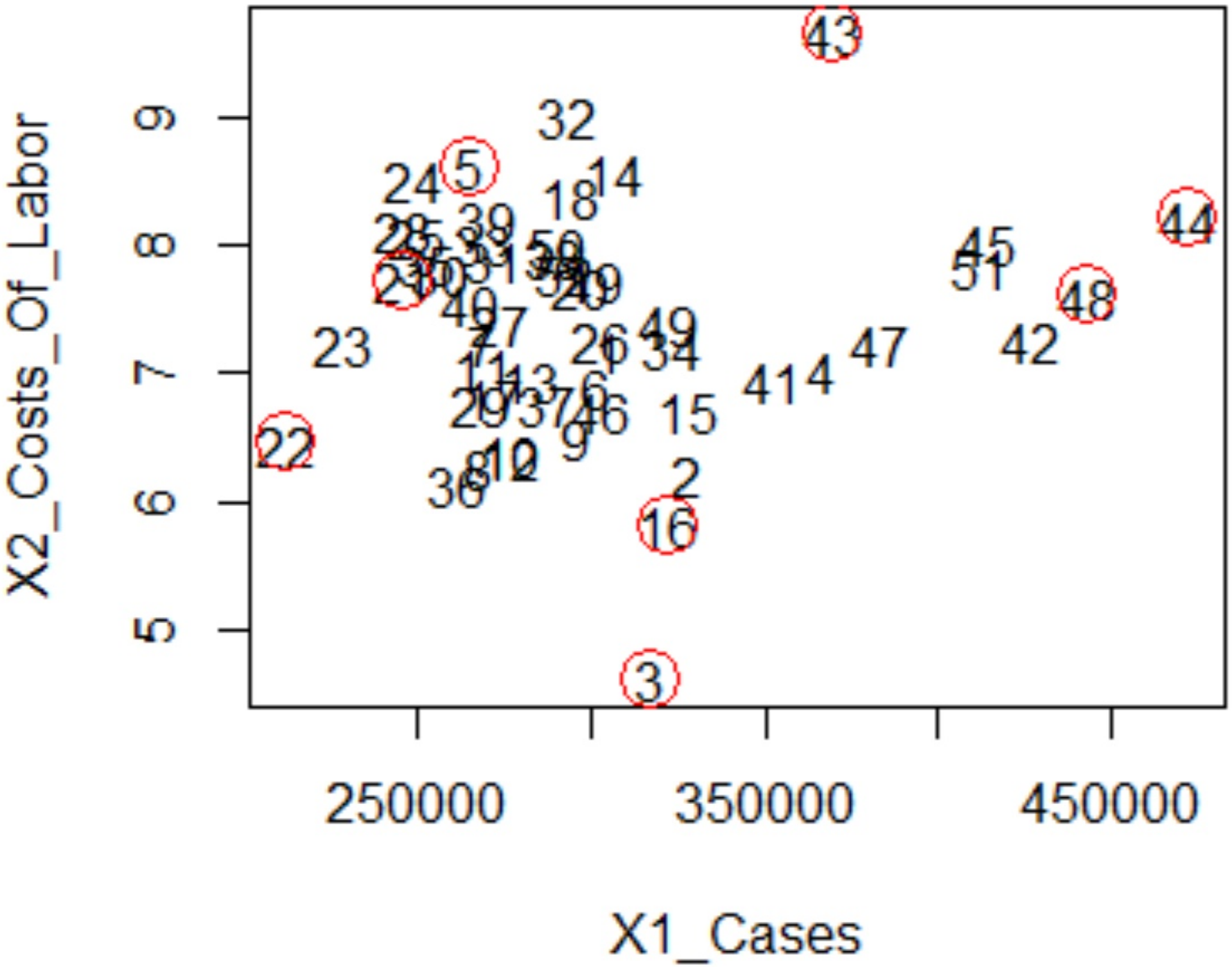
$Se_{0w} = 13.16843$
$Se_{1w} = .05056$

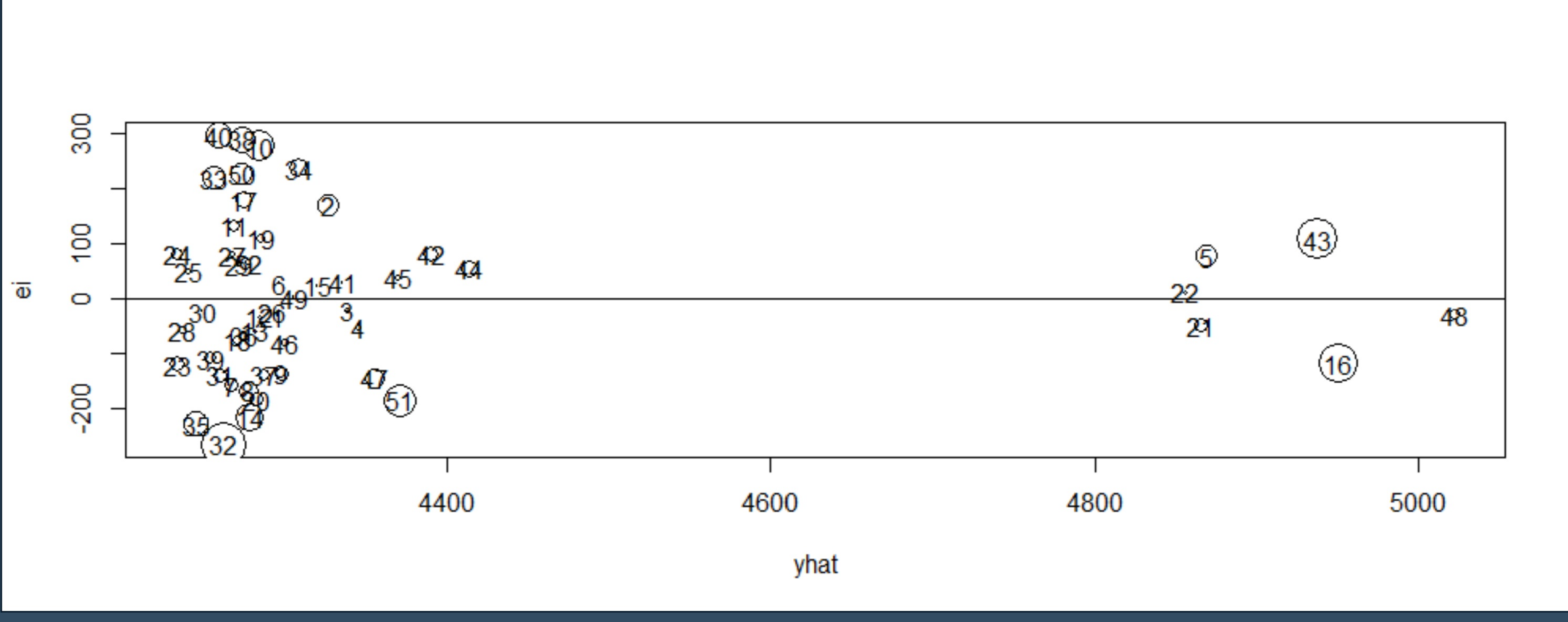
11.25 a. $\hat{Y} = 40.10114 + 3.09514 X_1 + -0.53 X_2 + -.29286 X_2^2$

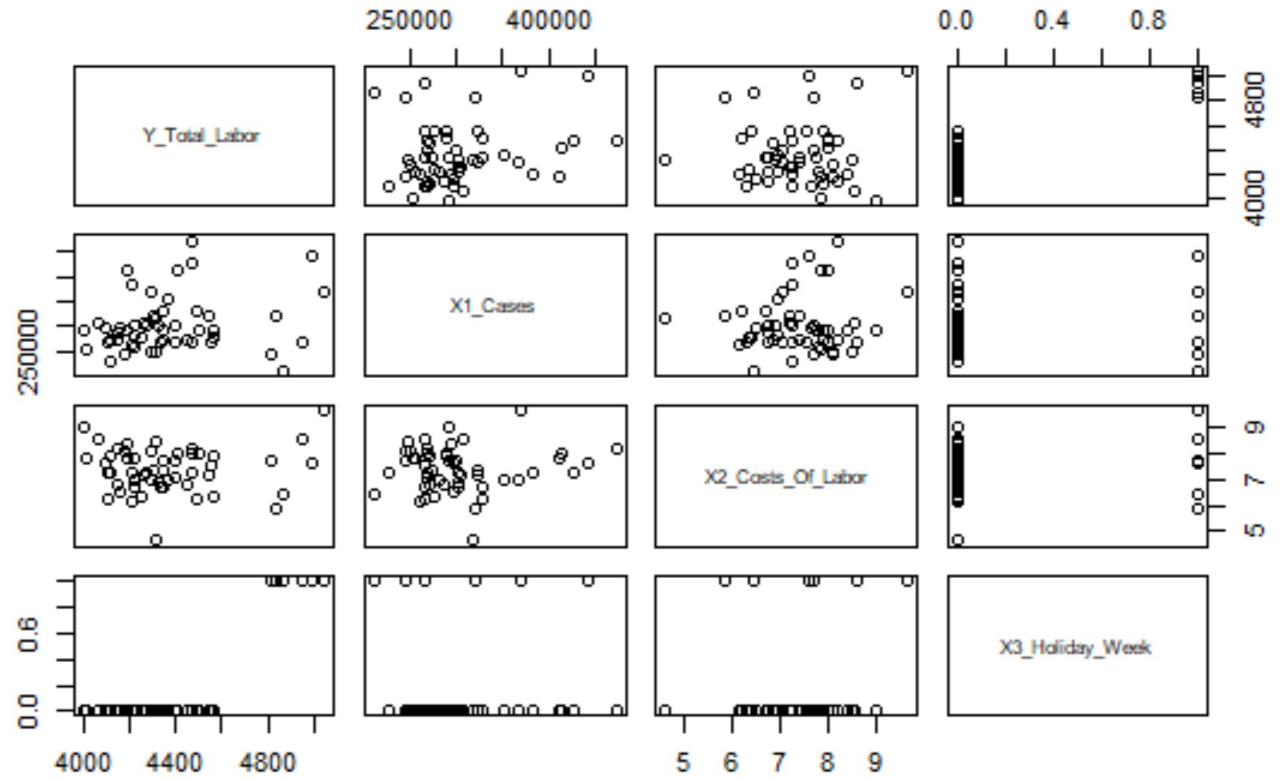
b.

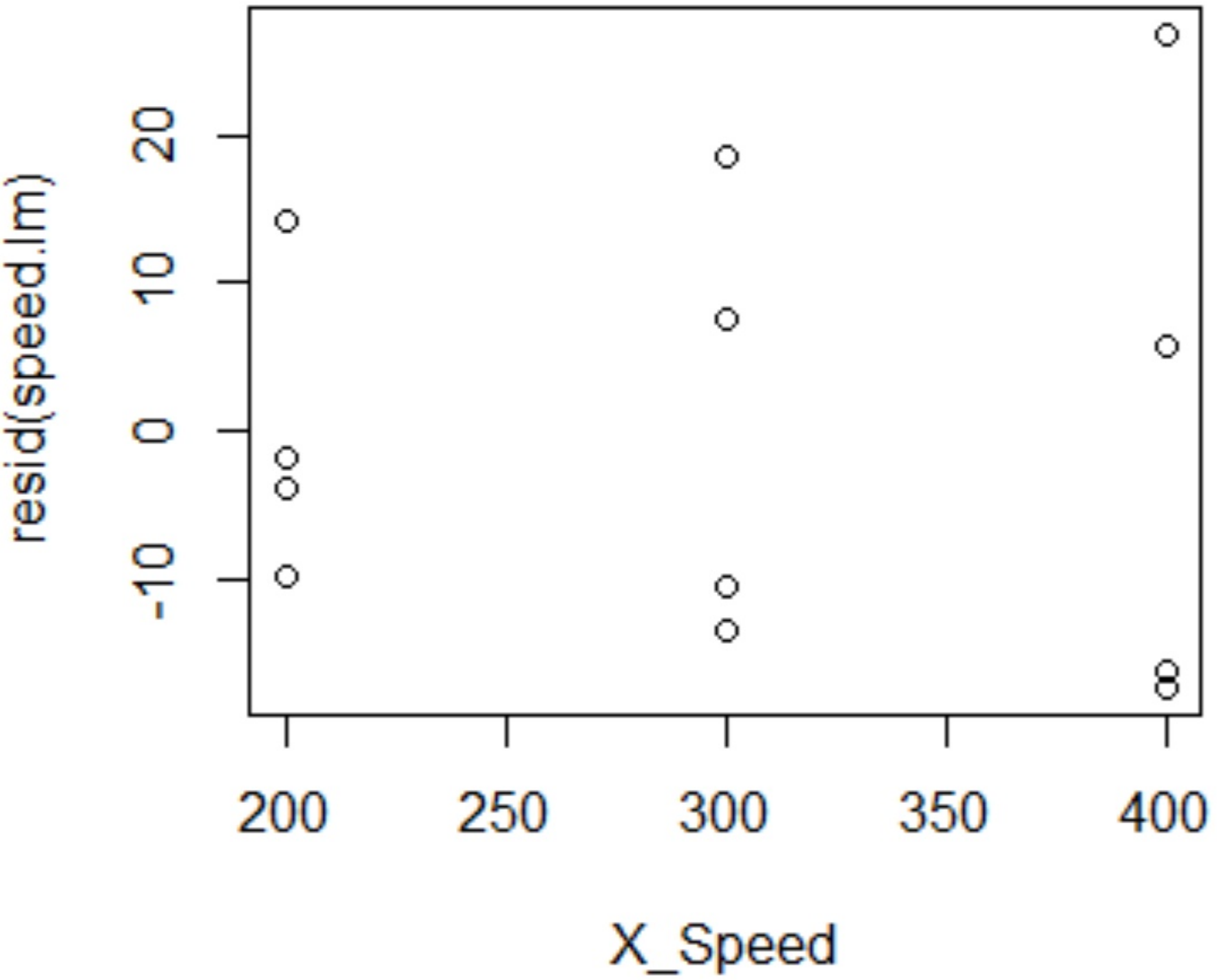

Y_Job_Proficiency_Score
X1_Test1
X2_Test2
X3_Test3
X4_Test4
60 80 100 120 140
80 90 100 110
60 80 100 120
80 90 100 120
75 85 95 105
60 120
80 100
60 100
80 120
75 95

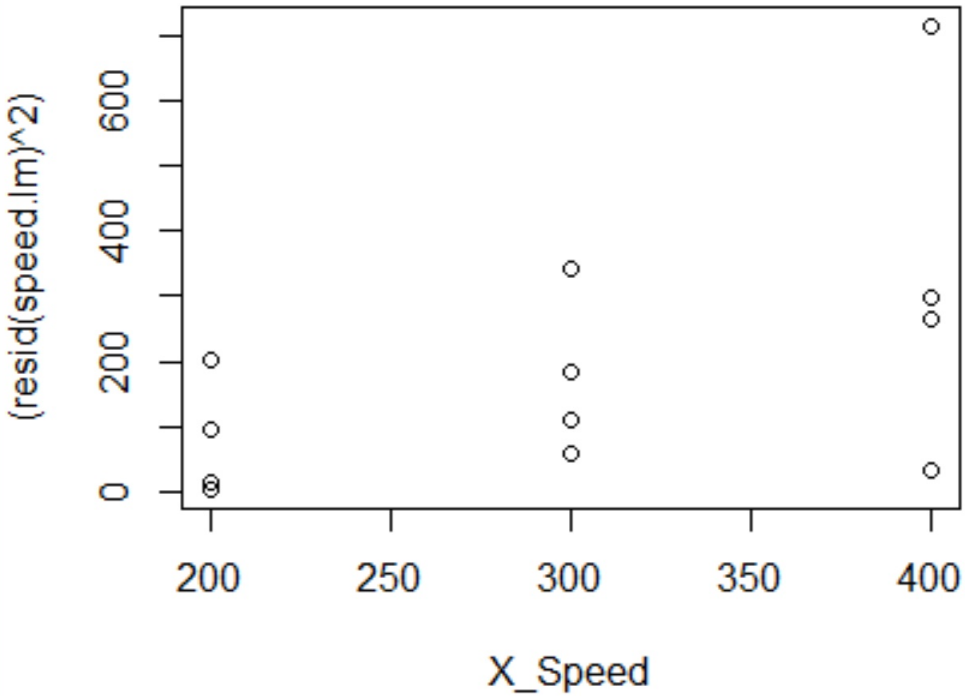

rstudent(lm.grocery)
4
2
0
-2
-4
4400
4600
4800
5000
fitted(lm.grocery)


X2_Costs_Of_Labor
X1_Cases
250000
350000
450000
5
6
7
8
9


ei
yhat
300
100
0
-200
4400
4600
4800
5000


Y_Total_Labor
X1_Cases
X2_Costs_Of_Labor
X3_Holiday_Week
250000
400000
0.0
0.4
0.8
4000
4800
250000
5
7
9
0.0
0.6
4000
4400
4800
5
6
7
8
9


resid(speed.lm)
-10
0
10
20
200
250
300
350
400
X_Speed


(resid(speed.lm)^2)
0
200
400
600
200
250
300
350
400
X_Speed

```
> #b.
> aptitude <- read.delim(file = "https://www.math.arizona.edu/~piegorsch/571A/Data/Chapter09/CH09PR10.txt",
+                        header = F, sep = "")
> Y_Job_Proficiency_Score <- aptitude$V1
> X1_Test1 <- aptitude$V2
> X2_Test2 <- aptitude$V3
> X3_Test3 <- aptitude$V4
> X4_Test4 <- aptitude$V5
> CH09PR10.df <- data.frame(Y_Job_Proficiency_Score, X1_Test1,X2_Test2,
+                           X3_Test3,X4_Test4)
> pairs(CH09PR10.df)
> cor(CH09PR10.df)
                        Y_Job_Proficiency_Score  X1_Test1  X2_Test2  X3_Test3  X4_Test4
Y_Job_Proficiency_Score               1.0000000 0.5144107 0.4970057 0.8970645 0.8693865
X1_Test1                              0.5144107 1.0000000 0.1022689 0.1807692 0.3266632
X2_Test2                              0.4970057 0.1022689 1.0000000 0.5190448 0.3967101
X3_Test3                              0.8970645 0.1807692 0.5190448 1.0000000 0.7820385
X4_Test4                              0.8693865 0.3266632 0.3967101 0.7820385 1.0000000
> #c.
> CH09PR10.lm <- lm(Y_Job_Proficiency_Score ~ X1_Test1 + X2_Test2 + X3_Test3 + X4_Test4)
> CH09PR10.lm

Call:
lm(formula = Y_Job_Proficiency_Score ~ X1_Test1 + X2_Test2 +
    X3_Test3 + X4_Test4)

Coefficients:
(Intercept)     X1_Test1     X2_Test2     X3_Test3     X4_Test4
 -124.38182      0.29573      0.04829      1.30601      0.51982

> summary(CH09PR10.lm)

Call:
lm(formula = Y_Job_Proficiency_Score ~ X1_Test1 + X2_Test2 +
    X3_Test3 + X4_Test4)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-5.9779 -3.4506  0.0941  2.4749  5.9959

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -124.38182    9.94106 -12.512 6.48e-11 ***
X1_Test1       0.29573    0.04397   6.725 1.52e-06 ***
X2_Test2       0.04829    0.05662   0.853  0.40383
X3_Test3       1.30601    0.16409   7.959 1.26e-07 ***
X4_Test4       0.51982    0.13194   3.940  0.00081 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.099 on 20 degrees of freedom
Multiple R-squared:  0.9629,    Adjusted R-squared:  0.9555
F-statistic: 129.7 on 4 and 20 DF,  p-value: 5.262e-14

> #a.
> library(leaps)
> best <- function(CH09PR10.lm, ...)
+ {
+   subsets <- regsubsets(formula(CH09PR10.lm), model.frame(CH09PR10.lm), ...)
+   subsets <- with(summary(subsets),
+                   cbind(p = as.numeric(rownames(which)), which, adjr2))
+
+   return(subsets)
+ }
> round(best(CH09PR10.lm, nbest = 6), 4)
  p (Intercept) X1_Test1 X2_Test2 X3_Test3 X4_Test4  adjr2
1 1           1        0        0        1        0 0.7962
1 1           1        0        0        0        1 0.7452
1 1           1        1        0        0        0 0.2326
1 1           1        0        1        0        0 0.2143
2 2           1        1        0        1        0 0.9269
2 2           1        0        0        1        1 0.8661
2 2           1        1        0        0        1 0.7985
2 2           1        0        1        1        0 0.7884
2 2           1        0        1        0        1 0.7636
2 2           1        1        1        0        0 0.4155
```

```
2 2           1       1       1       0       0 0.4155
3 3           1       1       0       1       1 0.9560
3 3           1       1       1       1       0 0.9247
3 3           1       0       1       1       1 0.8617
3 3           1       1       1       0       1 0.8233
4 4           1       1       1       1       1 0.9555
> #a.
> step(CH09PR10.lm, direction = "backward", k = 2)
Start:  AIC=74.95
Y_Job_Proficiency_Score ~ X1_Test1 + X2_Test2 + X3_Test3 + X4_Test4

           Df Sum of Sq     RSS     AIC
- X2_Test2  1     12.22  348.20  73.847
<none>                   335.98  74.954
- X4_Test4  1    260.74  596.72  87.314
- X1_Test1  1    759.83 1095.81 102.509
- X3_Test3  1   1064.15 1400.13 108.636

Step:  AIC=73.85
Y_Job_Proficiency_Score ~ X1_Test1 + X3_Test3 + X4_Test4

           Df Sum of Sq     RSS     AIC
<none>                  348.20  73.847
- X4_Test4  1    258.46  606.66  85.727
- X1_Test1  1    763.12 1111.31 100.861
- X3_Test3  1   1324.39 1672.59 111.081

Call:
lm(formula = Y_Job_Proficiency_Score ~ X1_Test1 + X3_Test3 +
    X4_Test4)

Coefficients:
(Intercept)     X1_Test1     X3_Test3     X4_Test4
  -124.2000       0.2963       1.3570       0.5174

> library(rms)
> CH09PR10.ols <- ols(CH09PR10.lm)
> fastbw(fit=CH09PR10.ols, rule = "p", type = "individual", sls = .05)
```

```
 Deleted  Chi-Sq d.f. P      Residual d.f. P      AIC   R2
 X2_Test2 0.73   1    0.3937 0.73     1    0.3937 -1.27 0.962

Approximate Estimates after Deleting Factors

              Coef    S.E.  Wald Z         P
Intercept -124.2000 9.93878 -12.497 0.000e+00
X1_Test1     0.2963 0.04397   6.740 1.585e-11
X3_Test3     1.3570 0.15283   8.879 0.000e+00
X4_Test4     0.5174 0.13191   3.922 8.765e-05

Factors in Final Model

[1] X1_Test1 X3_Test3 X4_Test4
> #a.
> n = length(Y_Total_Labor)
> plot(rstudent(lm.grocery)~ fitted(lm.grocery),ylim = c(-4,4))
> abline(h=0)
> tcrit = qt(1- .5*(.05/n), n-4-1)
> abline(h= tcrit, lty = 2)
> abline(h = -tcrit, lty = 2)
> tcrit
[1] 3.52308
Warning messages:
1: In doTryCatch(return(expr), name, parentenv, handler) :
  display list redraw incomplete
2: In doTryCatch(return(expr), name, parentenv, handler) :
  invalid graphics state
3: In doTryCatch(return(expr), name, parentenv, handler) :
  invalid graphics state
> g.res <- rstudent(lm.grocery)
> g.res
          1           2           3           4           5           6           7           8           9          10          11          12
-0.22408724  1.22549009 -0.17058921 -0.38465346  0.59079243  0.19612403 -1.11220903 -1.20529304 -0.97317140  2.03651764  0.93459516 -0.23775605
         13          14          15          16          17          18          19          20          21          22          23          24
-0.41516269 -1.57563574  0.16177701 -0.94585538  1.27571169 -0.53946536  0.76695374 -1.30688333 -0.37866873  0.09348669 -0.86199416  0.58204380
         25          26          27          28          29          30          31          32          33          34          35          36
 0.34737819 -0.16427705  0.55395260 -0.41694583  0.43222641 -0.16381817 -0.98793522 -1.99766654  1.58403006  1.70041654 -1.66686277 -0.48548061
```

```
-0.98726030  2.11878596 -0.77401415  2.17827186  0.20535372  0.59660183  0.88556630  0.43246959  0.27680521 -0.56690329 -1.04682238 -0.23443689
         49          50          51          52
 0.01909697  1.63020460 -1.37470514  0.45278959
> #b.
> hat <- as.numeric(hatvalues(lm.grocery))
> p = 4
> 2*p/n
[1] 0.1538462
> ifelse(hat > .1538462, "High Leverage", "Low Leverage")
 [1] "Low Leverage"  "Low Leverage"  "High Leverage" "Low Leverage"  "High Leverage" "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"

[10] "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "High Leverage" "Low Leverage"  "Low Leverage"

[19] "Low Leverage"  "Low Leverage"  "High Leverage" "High Leverage" "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"

[28] "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"

[37] "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"  "High Leverage" "High Leverage" "Low Leverage"

[46] "Low Leverage"  "Low Leverage"  "High Leverage" "Low Leverage"  "Low Leverage"  "Low Leverage"  "Low Leverage"
> #c.
> plot(X2_Costs_Of_Labor~ X1_Cases, pch='')
Error in plot.new() : figure margins too large
> text(X1_Cases, X2_Costs_Of_Labor, label=as.character(1:52))
Error in text.default(X1_Cases, X2_Costs_Of_Labor, label = as.character(1:52)) :
  invalid graphics state
> points(X1_Cases[hat>2*p/n], X2_Costs_Of_Labor[hat>2*p/n],
+        cex=2.5, col = 'red')
Error in plot.xy(xy.coords(x, y), type = type, ...) :
  invalid graphics state
> XM <- matrix(c(rep(1,52),X1_Cases,X2_Costs_Of_Labor,X3_Holiday_Week),ncol = 4)
> tXMnew <- t(matrix(c(1, 300000,7.2, 0)))
> XMnew <- matrix(c(1, 300000,7.2, 0))
> tXM <- t(XM)
> invXM <- solve(tXM%*%XM)
> hatnew <- tXMnew  %*% invXM %*% XMnew
> hatnew
          [,1]
```

```
> hatnew
           [,1]
[1,] 0.02221728
> influence.measures(lm.grocery)
Influence measures of
         lm(formula = Y_Total_Labor ~ X1_Cases + X2_Costs_Of_Labor + X3_Holiday_Week) :

      dfb.1_  dfb.X1_C  dfb.X2_C  dfb.X3_H    dffit cov.r   cook.d    hat inf
1  -0.007459 -0.002714  6.21e-03  0.010689 -0.03406 1.108 2.96e-04 0.0226
2   0.167778  0.104486 -2.38e-01 -0.040585  0.31452 1.022 2.45e-02 0.0618
3  -0.068934 -0.014358  8.54e-02  0.000549 -0.09030 1.389 2.08e-03 0.2189   *
4   0.009435 -0.066932  2.53e-02  0.019826 -0.09097 1.134 2.11e-03 0.0530
5  -0.048128 -0.082695  1.09e-01  0.244168  0.30123 1.331 2.30e-02 0.2063   *
6   0.015085  0.001275 -1.46e-02 -0.008288  0.03275 1.115 2.74e-04 0.0271
7  -0.079590  0.091650  1.16e-02  0.050624 -0.19091 1.009 9.07e-03 0.0286
8  -0.247737  0.090727  2.04e-01  0.034349 -0.29454 1.021 2.15e-02 0.0564
9  -0.133642  0.002884  1.34e-01  0.034015 -0.19909 1.046 9.92e-03 0.0402
10  0.364075 -0.104403 -3.14e-01 -0.063346  0.45863 0.815 4.94e-02 0.0483
11  0.089702 -0.075212 -3.73e-02 -0.039532  0.16469 1.042 6.80e-03 0.0301
12 -0.043534  0.012159  3.79e-02  0.007270 -0.05441 1.139 7.55e-04 0.0498
13 -0.038686  0.018090  2.52e-02  0.017251 -0.06996 1.102 1.25e-03 0.0276
14  0.247429  0.005724 -3.19e-01  0.116867 -0.39974 0.942 3.88e-02 0.0605
15  0.010756  0.012500 -1.75e-02 -0.006780  0.03196 1.128 2.61e-04 0.0376
16 -0.247689 -0.059782  3.25e-01 -0.452100 -0.55399 1.355 7.69e-02 0.2554   *
17  0.157025 -0.088636 -9.99e-02 -0.049133  0.23659 0.982 1.38e-02 0.0332
18  0.062243  0.018636 -9.42e-02  0.037392 -0.12512 1.118 3.97e-03 0.0510
19 -0.026298 -0.005698  4.84e-02 -0.043874  0.12436 1.062 3.90e-03 0.0256
20  0.029409  0.022300 -7.28e-02  0.073071 -0.20892 0.967 1.08e-02 0.0249
21 -0.029541  0.069015 -1.07e-02 -0.162688 -0.18554 1.333 8.76e-03 0.1936   *
22  0.030423 -0.025287 -1.87e-02  0.044646  0.05509 1.464 7.75e-04 0.2577   *
23 -0.109792  0.165314  3.06e-03  0.036262 -0.21148 1.083 1.12e-02 0.0568
24 -0.043157 -0.091669  1.21e-01 -0.039696  0.17116 1.148 7.43e-03 0.0796
25 -0.006270 -0.050658  4.67e-02 -0.020785  0.08471 1.141 1.83e-03 0.0561
26 -0.004113 -0.000529  2.04e-03  0.008053 -0.02458 1.110 1.54e-04 0.0219
27  0.024541 -0.040323  8.18e-03 -0.027020  0.09223 1.089 2.16e-03 0.0270
28  0.007582  0.065726 -5.94e-02  0.025151 -0.10625 1.142 2.87e-03 0.0610
29  0.059946 -0.034894 -3.84e-02 -0.015942  0.08454 1.112 1.82e-03 0.0368
30 -0.002098  0.020266 -1.38e-02  0.008970 -0.03419 1.133 2.98e-04 0.0417
```

```
31  0.020076  0.083893 -9.62e-02  0.057166 -0.19265 1.040 9.28e-03 0.0366
32  0.409541  0.091342 -5.71e-01  0.165206 -0.65108 0.869 9.98e-02 0.0960
33 -0.052422 -0.148650  1.87e-01 -0.094945  0.33139 0.922 2.66e-02 0.0419
34  0.010502  0.094896 -4.21e-02 -0.085179  0.27328 0.879 1.80e-02 0.0252
35 -0.017011  0.226607 -1.59e-01  0.092622 -0.36690 0.906 3.25e-02 0.0462
36 -0.112901  0.043900  9.26e-02  0.012459 -0.12929 1.142 4.25e-03 0.0662
37 -0.110065  0.028631  9.00e-02  0.038338 -0.17683 1.034 7.82e-03 0.0311
38 -0.099615 -0.082738  2.08e-01 -0.127009  0.38552 0.780 3.46e-02 0.0320
39  0.045537  0.072878 -1.15e-01  0.049168 -0.17576 1.087 7.79e-03 0.0490
40  0.073799 -0.212059  9.33e-02 -0.111047  0.39672 0.765 3.65e-02 0.0321
41  0.000971  0.028148 -1.55e-02 -0.009958  0.04391 1.134 4.92e-04 0.0437
42 -0.082711  0.203604 -2.50e-02 -0.038122  0.22442 1.205 1.28e-02 0.1240
43 -0.357797  0.133842  3.26e-01  0.356628  0.56165 1.428 7.92e-02 0.2869   *
44 -0.153072  0.207164  5.12e-02 -0.039600  0.22969 1.373 1.34e-02 0.2200   *
45 -0.063267  0.081261  2.57e-02 -0.021211  0.09757 1.215 2.43e-03 0.1105
46 -0.055528 -0.005636  5.72e-02  0.022404 -0.10240 1.093 2.66e-03 0.0316
47  0.077563 -0.224277  3.62e-02  0.059928 -0.27588 1.061 1.90e-02 0.0649
48  0.044986 -0.093842  9.01e-03 -0.102215 -0.14684 1.508 5.50e-03 0.2818   *
49 -0.000330  0.000995  9.03e-05 -0.001017  0.00302 1.115 2.34e-06 0.0245
50 -0.093733 -0.061904  1.79e-01 -0.100010  0.30678 0.904 2.27e-02 0.0342
51  0.272088 -0.396715 -8.32e-02  0.099684 -0.46528 1.036 5.31e-02 0.1028
52 -0.013521 -0.013921  3.31e-02 -0.025961  0.07620 1.099 1.48e-03 0.0275
> #f.
> ei = resid(lm.grocery)
> yhat = fitted(lm.grocery)
> radius = sqrt( cooks.distance(lm.grocery)/pi )
> plot( ei ~ yhat, pch=''); abline( h=0 )
Error in plot.new() : figure margins too large
> symbols( yhat, ei, circles=radius, inches=.15,
+          bg='white', fg='black', add=T )
Error in symbols(yhat, ei, circles = radius, inches = 0.15, bg = "white",  :
  invalid graphics state
> text(yhat, ei, label=as.character(1:52))
Error in text.default(yhat, ei, label = as.character(1:52)) :
  invalid graphics state
> #a.
> cor(grocery.df)
                Y_Total_Labor   X1_Cases X2_Costs_Of_Labor X3_Holiday_Week
```

```
Y_Total_Labor         1.0000000 0.20766494        0.06002960      0.81057940
X1_Cases              0.2076649 1.00000000        0.08489639      0.04565698
X2_Costs_Of_Labor     0.0600296 0.08489639        1.00000000      0.11337076
X3_Holiday_Week       0.8105794 0.04565698        0.11337076      1.00000000
> pairs(grocery.df)
Error in plot.new() : figure margins too large
> #b.
> library(car)
> vif(lm.grocery)
         X1_Cases X2_Costs_Of_Labor   X3_Holiday_Week
         1.008596          1.019598          1.014364
> mean(vif(lm.grocery))
[1] 1.014186
> #a.
> machine <- read.delim(file = "https://www.math.arizona.edu/~piegorsch/571A/Data/Chapter11/CH11PR07.txt",
+                       header = F, sep = "")
> Y_Defects <- machine$V1
> X_Speed <-  machine$V2
> speed.lm

Call:
lm(formula = Y_Defects ~ X_Speed)

Coefficients:
(Intercept)      X_Speed
    -5.7500       0.1875

> plot(resid(speed.lm) ~ X_Speed)
Error in plot.new() : figure margins too large
> #b.
> library("lawstat")
> ei <- resid(speed.lm)
> BF.htest <- levene.test(ei[order(X_Speed)], group = c(rep(1,6),rep(2,6)), location = "median")
> BF.htest

        Modified robust Brown-Forsythe Levene-type test based on the absolute deviations from the median

data:  ei[order(X_Speed)]
```

```
Test Statistic = 3.7056, p-value = 0.08312

> sqrt(BF.htest$statistic)
Test Statistic
      1.924992
> #c.
> plot((resid(speed.lm)^2) ~ X_Speed)
Error in plot.new() : figure margins too large
> #d.
> variancef <- lm((speed.lm)^2 ~ X_Speed)
Error in (speed.lm)^2 : non-numeric argument to binary operator
> variancef

Call:
lm(formula = resid(speed.lm)^2 ~ X_Speed)

Coefficients:
(Intercept)      X_Speed
   -180.083        1.244

> v <- fitted(variancef) ; w <- 1/v
> w
          1           2           3           4           5           6           7           8           9          10          11          12
0.014563107 0.003150433 0.005180229 0.003150433 0.014563107 0.005180229 0.005180229 0.003150433 0.014563107 0.003150433 0.014563107 0.005180229
> #e.
> summary(lm(Y_Defects ~ X_Speed, weights= w))

Call:
lm(formula = Y_Defects ~ X_Speed, weights = w)

Weighted Residuals:
    Min      1Q  Median      3Q     Max
-1.1572 -0.9338 -0.3124  0.7377  1.7391

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -6.23322   13.16843  -0.473  0.64613
X_Speed      0.18911    0.05056   3.740  0.00385 **
```

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

Residual standard error: 1.109 on 10 degrees of freedom
Multiple R-squared:  0.5831,    Adjusted R-squared:  0.5414
F-statistic: 13.99 on 1 and 10 DF,  p-value: 0.003846

> #f
> summary(lm(Y_Defects ~ X_Speed))

Call:
lm(formula = Y_Defects ~ X_Speed)

Residuals:
    Min      1Q  Median      3Q     Max
-17.250 -11.250  -2.750   9.188  26.750

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -5.75000   16.73052  -0.344  0.73820
X_Speed      0.18750    0.05381   3.484  0.00588 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15.22 on 10 degrees of freedom
Multiple R-squared:  0.5484,    Adjusted R-squared:  0.5032
F-statistic: 12.14 on 1 and 10 DF,  p-value: 0.005878

> summary(lm(Y_Defects ~ X_Speed, weights= w))

Call:
lm(formula = Y_Defects ~ X_Speed, weights = w)

Weighted Residuals:
    Min      1Q  Median      3Q     Max
-1.1572 -0.9338 -0.3124  0.7377  1.7391

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

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -6.23322   13.16843  -0.473  0.64613
X_Speed      0.18911    0.05056   3.740  0.00385 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.109 on 10 degrees of freedom
Multiple R-squared:  0.5831,    Adjusted R-squared:  0.5414
F-statistic: 13.99 on 1 and 10 DF,  p-value: 0.003846

> #a.
> cropyield <- read.delim( file = "https://www.math.arizona.edu/~piegorsch/571A/Data/Chapter11/CH11PR25.txt",
+                          header = F,
+                          sep = "")
> X1_Moisture <- cropyield$V2
> X2_Temperature <- cropyield$V3
> Y_tomato <- cropyield$V1
> x1sq <- X1_Moisture^2
> secondorder.lm <- lm(Y_tomato ~ X1_Moisture + X2_Temperature + x1sq )
> summary(secondorder.lm)

Call:
lm(formula = Y_tomato ~ X1_Moisture + X2_Temperature + x1sq)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9994 -0.4691 -0.2331  0.1931  2.0569

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    40.10114    3.58316  11.192 2.61e-10 ***
X1_Moisture     5.09514    0.51142   9.963 2.07e-09 ***
X2_Temperature -0.53000    0.12019  -4.410 0.000244 ***
x1sq           -0.29286    0.02539 -11.533 1.51e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```