STA 108: Homework 2
Hangshi Jin
913142686
Christiana Drake
Section A01

Problem 5.6

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 0 \\ 1 & 3 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 0 \end{bmatrix}, Y = \begin{bmatrix} 16 \\ 9 \\ 17 \\ 12 \\ 22 \\ 13 \\ 8 \\ 15 \\ 19 \\ 11 \end{bmatrix} \Rightarrow Y'Y = 2194, X'X = \begin{bmatrix} 10 & 10 \\ 10 & 20 \end{bmatrix}, X'Y = \begin{bmatrix} 142 \\ 182 \end{bmatrix}$$

Problem 5.9

(a)Yes, since k_1 can be any number while $k_2 = k_3 = 0$.

(b)when r scalars $k_1, \cdots, k_r$, not all zero, can be found such that:

$$k_1 R_1 + k_2 R_2 + \cdots + k_r R_r = 0$$

where 0 denotes the zero row vector, the r row vectors are linearly dependent. If the only set of scalars for which the equality holds is $k_1 = \cdots = k_r = 0$, the set of r row vectors is linearly independent.

The row vectors of A are linearly dependent.

(c)$Rank(A) = 2$

(d)

$$det(A) = 0$$

Problem 5.25

(a)

$$(X'X)^{-1} = \begin{bmatrix} 10 & 10 \\ 10 & 20 \end{bmatrix}^{-1} = \begin{bmatrix} 0.2 & -0.1 \\ -0.1 & 0.1 \end{bmatrix}$$

$$b = (X'X)^{-1}X'Y = \begin{bmatrix} 0.2 & -0.1 \\ -0.1 & 0.1 \end{bmatrix} \begin{bmatrix} 142 \\ 182 \end{bmatrix} = \begin{bmatrix} 10.2 \\ 4 \end{bmatrix}$$

$$e = Y - Xb = \begin{bmatrix} 1.8 \\ -1.2 \\ -1.2 \\ 1.8 \\ -0.2 \\ -1.2 \\ -2.2 \\ 0.8 \\ 0.8 \\ 0.8 \end{bmatrix}, H = X(X'X)^{-1}X' = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0 & 0.2 & -0.1 & 0.1 & 0.2 & 0.1 & 0 & 0.2 \\ 0.1 & 0 & 0.2 & 0 & 0.3 & 0.1 & 0 & 0.1 & 0.2 & 0 \\ 0.1 & 0.2 & 0 & 0.2 & -0.1 & 0.1 & 0.2 & 0.1 & 0 & 0.2 \\ 0.1 & -0.1 & 0.3 & -0.1 & 0.5 & 0.1 & -0.1 & 0.1 & 0.3 & -0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0 & 0.2 & -0.1 & 0.1 & 0.2 & 0.1 & 0 & 0.2 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0 & 0.2 & 0 & 0.3 & 0.1 & 0 & 0.1 & 0.2 & 0 \\ 0.1 & 0.2 & 0 & 0.2 & -0.1 & 0.1 & 0.2 & 0.1 & 0 & 0.2 \end{bmatrix}$$

$$SSE = e'e = 17.6, MSE = \frac{17.6}{8} = 2.2, s^2\{b\} = MSE(X'X)^{-1} = 2.2\begin{bmatrix} 0.2 & -0.1 \\ -0.1 & 0.1 \end{bmatrix} = \begin{bmatrix} 0.44 & -0.22 \\ -0.22 & 0.22 \end{bmatrix}$$

$$\text{when } X_h = 2,\ \hat{Y}_h \ = X_h'b = \begin{bmatrix} 1 & 2 \end{bmatrix}\begin{bmatrix} 10.2 \\ 4 \end{bmatrix} = 18.2, s^2\{\hat{Y}_h\} = X_h's^2\{b\}X_h = 0.44$$

(b)

$$s^2\{b_1\} = 0.22, s\{b_0, b_1\} = -0.22, s\{b_0\} = 0.6633$$

(c)

$$\text{The quadratic form for SSR} = H - \left(\frac{1}{n}\right)J = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & -0.1 & 0.1 & -0.2 & 0 & 0.1 & 0 & -0.1 & 0.1 \\ 0 & -0.1 & 0.1 & -0.1 & 0.2 & 0 & -0.1 & 0 & 0.1 & -0.1 \\ 0 & 0.1 & -0.1 & 0.1 & -0.2 & 0 & 0.1 & 0 & -0.1 & 0.1 \\ 0 & -0.2 & 0.2 & -0.2 & 0.4 & 0 & -0.2 & 0 & 0.2 & -0.2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & -0.1 & 0.1 & -0.2 & 0 & 0.1 & 0 & -0.1 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.1 & 0.1 & -0.1 & 0.2 & 0 & -0.1 & 0 & 0.1 & -0.1 \\ 0 & 0.1 & -0.1 & 0.1 & -0.2 & 0 & 0.1 & 0 & -0.1 & 0.1 \end{bmatrix}$$

Problem 6.2

(a)

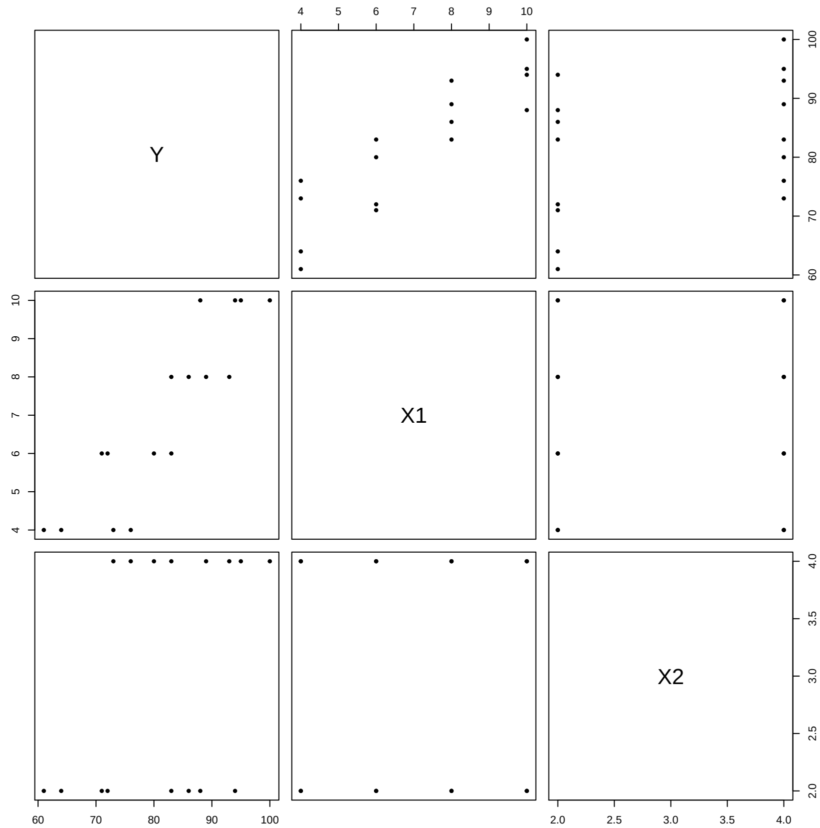
$$X = \begin{bmatrix} X_{11} & X_{12} & X_{11}^2 \\ X_{21} & X_{22} & X_{21}^2 \\ X_{31} & X_{32} & X_{31}^2 \\ X_{41} & X_{42} & X_{41}^2 \\ X_{51} & X_{52} & X_{51}^2 \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

(b)

$$X = \begin{bmatrix} 1 & X_{11} & \log_{10} X_{12} \\ 1 & X_{21} & \log_{10} X_{22} \\ 1 & X_{31} & \log_{10} X_{32} \\ 1 & X_{41} & \log_{10} X_{42} \\ 1 & X_{51} & \log_{10} X_{52} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

Problem 6.5

(a)



Brand Preference Scatterplot Matrix

	Y	X1	X2
Y	1.0000000	0.8923929	0.3945807
X1	0.8923929	1.0000000	0.0000000
X2	0.3945807	0.0000000	1.0000000

Correlation Matrix

There is a strong linear relationship between Y and X1, a moderate linear relationship between Y and X2. There are no outliers for X1 and X2.

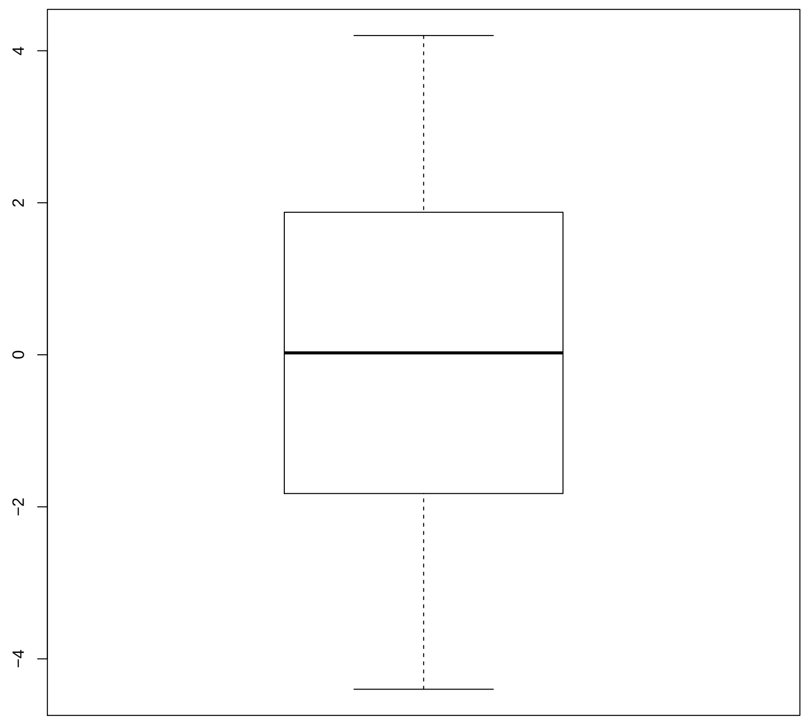
(b)

$$\hat{Y}_i = 37.650 + 4.425X_{i1} + 4.375X_{i2}$$

One unit increase in X_{i1} expects 4.425 increase in Y_i.

(c)

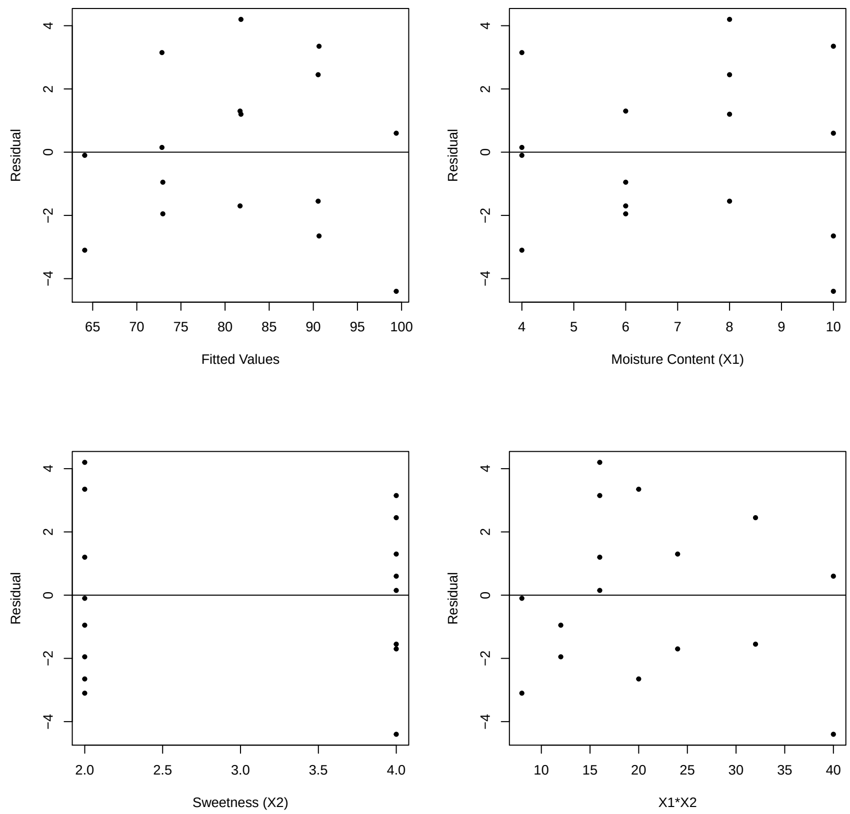
$$Residuals = \begin{bmatrix} -0.10 \\ 0.15 \\ -3.10 \\ 3.15 \\ -0.95 \\ -1.70 \\ -1.95 \\ 1.30 \\ 1.20 \\ -1.55 \\ 4.20 \\ 2.45 \\ -2.65 \\ -4.40 \\ 3.35 \\ 0.60 \end{bmatrix}$$



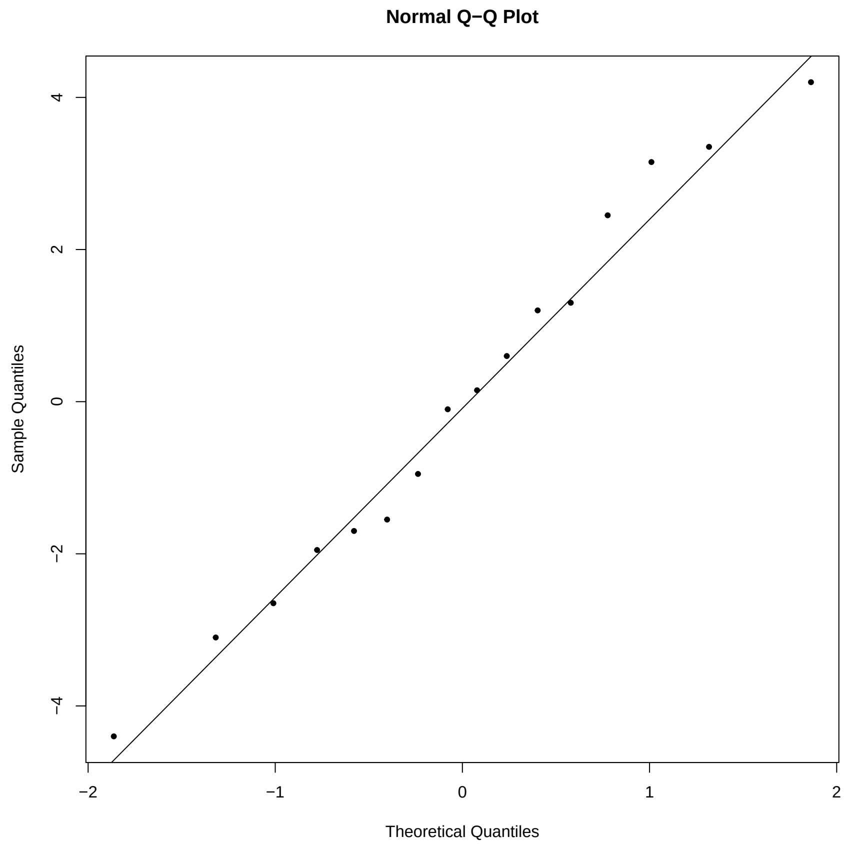
Box Plot of the Residuals

The residual's range is small indicating that the model may be appropriate.

(d)



Residual Plots



Normal Probability Plot

There is no obvious pattern, no obvious change in variability, no obvious departures from normality. The model seems appropriate.

(f)Alternatives:

$$H_0 : E\{Y\} = 37.650 + 4.425X_{i1} + 4.375X_{i2}$$

$$H_a : E\{Y\} \neq 37.650 + 4.425X_{i1} + 4.375X_{i2}$$

$$c - p = 8 - 3 = 5, n - c = 16 - 8 = 8$$

Decisions:

$$\text{If } F^* \leq F(0.99; 5, 8), \text{ conclude } H_0$$

$$\text{If } F^* > F(0.99; 5, 8), \text{ conclude } H_a$$

$$SSPE = 57, SSE = 94.3, SSLF = 37.3, F^* = \frac{37.3}{5} / \frac{57}{8} = 1.0470 < 6.6318$$

We conclude there is no lack of fit.

Problem 6.6

(a)Alternatives:

$$H_0 : \beta_1 \text{and} \beta_2 = 0$$

$$H_a : \text{not both } \beta_1 \text{ and } \beta_2 \text{ equal zero}$$

Decisions:

$$\text{If } F^* \leq F(0.99; 2, 13), \text{ conclude } H_0$$

$$\text{If } F^* > F(0.99; 2, 13), \text{ conclude } H_a$$

$$F^* = \frac{MSR}{MSE} = \frac{936.35}{7.25} = 129.08 > 6.701$$

We conclude there is a regression relation and not both β_1 and β_2 equal zero.

(b)The p-value is 2.6689×10^{-9}.

(c)

$$s^2\{b\} = MSE(X'X)^{-1} \Rightarrow s^2\{b_1\} = 0.0907, s^2\{b_2\} = 0.4537, t(0.99; 13) = 3.3725$$

$$4.425 - 3.3725\sqrt{0.0907} \leq \beta_1 \leq 4.425 + 3.3725\sqrt{0.0907} \Rightarrow 3.41 \leq \beta_1 \leq 5.44$$

$$4.375 - 3.3725\sqrt{0.4537} \leq \beta_2 \leq 4.375 + 3.3725\sqrt{0.4537} \Rightarrow 5.44 \leq \beta_2 \leq 6.65$$

The family confidence coefficient ensures at least 99% that $3.41 \leq \beta_1 \leq 5.44$ and $5.44 \leq \beta_2 \leq 6.65$.

Problem 6.7

(a)

$$SSTO = 1967, SSE = 94.3, R^2 = 1 - \frac{94.3}{1967} = 0.9523$$

When both moisture content (X_1) and sweetness (X_2) are considered, the variation in degree of brand liking is reduced by 95.21%.

(b)

$$cor(Y_i, \hat{Y}_i)^2 = 0.9521 = R^2$$

Problem 6.8

(a)

$$t(0.995; 13) = 3.0123, \hat{Y}_h = 77.275, s^2\{\hat{Y}_h\} = 1.2694$$

$$77.275 - 3.0123\sqrt{1.2694} \leq E\{\hat{Y}_h\} \leq 77.275 + 3.0123\sqrt{1.2694} \Rightarrow 73.88 \leq E\{\hat{Y}_h\} \leq 80.67$$

We are 99% confident that the mean response at X_h lies somewhere between 73.88 and 80.67.

(b)

$$s^2\{\text{pred}\} = 8.5233$$

$$77.275 - 3.0123\sqrt{8.5233} \leq E\{\hat{Y}_h\} \leq 77.275 + 3.0123\sqrt{8.5233} \Rightarrow 68.48 \leq E\{\hat{Y}_h\} \leq 86.07$$

We are 99% confident that the new observation at X_h lies somewhere between 68.48 and 86.07.

Problem 7.3

(a)Analysis of Variance Table

Model 1: Y ~ X1

Model 2: Y ~ X1 + X2

	Res.Df	RSS	Df	Sum of Sq	F	$Pr(>F)$
1	14	400.55				
2	13	94.30	1	306.25	42.219	$2.011e-05***$

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

(b)Alternatives:

$$H_0 : \beta_2 = 0$$
$$H_a : \beta_2 \neq 0$$

Decisions:

$$\text{If } F^* \leq F(0.99; 1, 12), \text{ conclude } H_0$$
$$\text{If } F^* > F(0.99; 1, 12), \text{ conclude } H_a$$

$$F^* = \frac{SSR(X_2|X_1)}{1} / \frac{SSE(X_1, X_2)}{12} = \frac{400.55 - 94.3}{1} / \frac{94.3}{13} = 42.22 > 9.0738,$$

We conclude that X_2 cannot be dropped from the regression model given that X_1 is retained. P-value= 2.011×10^{-5}

Problem 7.12

$$R^2_{Y1} = 1 - \frac{SSE(X_1)}{SSTO} = 1 - \frac{400.55}{1967} = 0.7964$$

The variation in degree of brand liking is reduced by 79.64% with the regression model that only contains moisture content (X_1) .

$$R^2_{Y2} = 1 - \frac{SSE(X_2)}{SSTO} = 1 - \frac{1660.8}{1967} = 0.1595$$

The variation in degree of brand liking is reduced by 15.95% with the regression model that only contains sweetness (X_2) .

$$R^2_{Y12} = 1 - \frac{SSE(X_1, X_2)}{SSTO} = 0.9523$$

The variation in degree of brand liking is reduced by 95.23% with the regression model that contains both moisture content (X_1) and sweetness (X_2).

$$R^2_{Y1|2} = \frac{SSR(X_1|X_2)}{SSE(X_2)} = \frac{1566.5}{1660.8} = 0.9423$$

The error sum of squares of Y is reduced by 94.23% with adding X_1 to the regression model that only contains X_2.

$$R^2_{Y2|1} = \frac{SSR(X_2|X_1)}{SSE(X_1)} = \frac{306.23}{400.55} = 0.7645$$

The error sum of squares of Y is reduced by 76.45% with adding X_2 to the regression model that only contains X_1.

$$R^2 = 0.9523$$

The variation in degree of brand liking is reduced by 95.23% with the regression model that contains both moisture content (X_1) and sweetness (X_2).

Problem 7.24

(a)Since X_1 and X_2 are uncorrelated,

$$\beta_0 = \overline{Y} - \beta_1 \overline{X}_1 = 50.775, \hat{Y} = 50.775 + 4.425X_1$$

(b)β_0 changed but β_1 remained unchanged.

(c)They are not equal but close, $SSR(X_1|X_2) = 1566.5, SSR(X_1) = 1967 - 400.55 = 1566.45$.

(d)If X_1 and X_2 are uncorrelated, β_1 and β_2 remain unchanged for the regression model that contains only X_1 or X_2.

Problem 7.28

(a)

$$SSR(X_5|X_1) = SSE(X_1) - SSE(X_1, X_5)$$
$$SSR(X_3, X_4|X_1) = SSE(X_1) - SSE(X_1, X_3, X_4)$$
$$SSR(X_4|X_1, X_2, X_3) = SSE(X_1, X_2, X_3) - SSE(X_1, X_2, X_3, X_4)$$

(b)

$$SSR(X_5|X_1, X_2, X_3, X_4)$$
$$SSR(X_2, X_4|X_1, X_3, X_5)$$

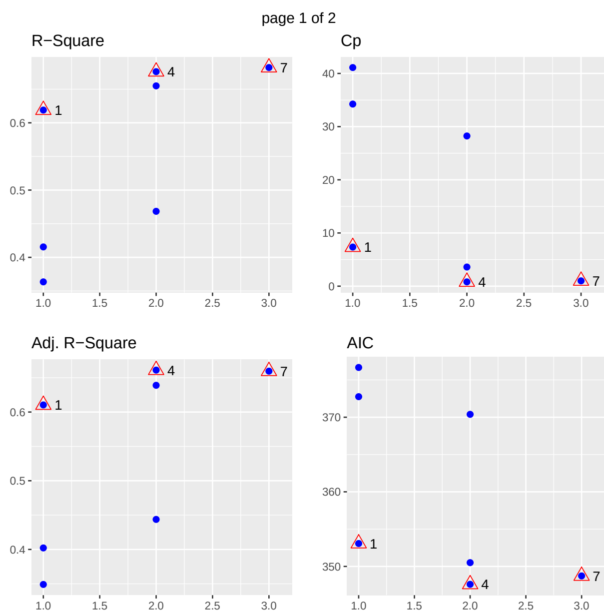
Problem 9.4

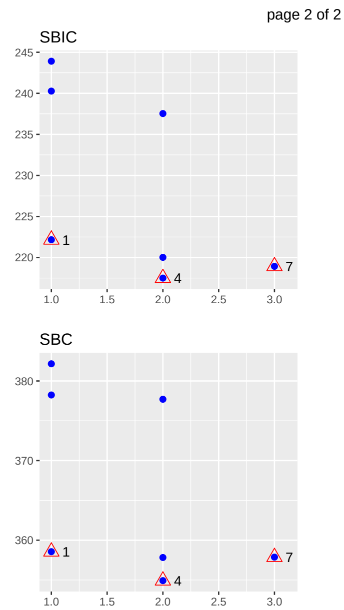
With a relatively small α-to-enter value, the search can be done more quickly. With a relatively large α-to-enter value, the search can be done more accurately, which means the search will be less likely terminated if the p-value of the first step exceeds α.

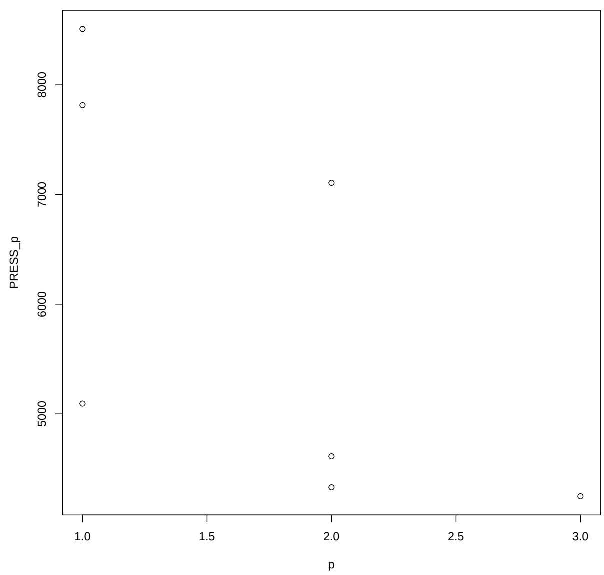
Problem 9.5

Because then there will be a contradiction where α-to-remove value<p-value< α-to-enter value. The corresponding variable is both to be dropped and added.

Problem 9.9







Model1=Y∼X1

Model2=Y∼X3

Model3=Y∼X2

Model4=Y∼X1+X3

Model5=Y∼X1+X2

Model6=Y∼X2+X3

Model7=Y∼X1+X2+X3

(a)From the graphs, we see Model4 has the highest $R^2_{a,p}$, smallest AIC_p, closest $C_p = 2.81 < p = 3$, and smallest $PRESS_p$ for $p = 2$ though that of Model7 is the smallest among all. Nevertheless, I still recommend Model4.

(b)No, for $PRESS_p$, Model7 is smaller, but Model4 is the best subset for the other three criteria and small enough for $PRESS_p$. This happens often.

(c)Forward stepwise regression is not necessary here because the data contains only 3 variables.

Problem 9.17

(a)Subset selection object
Call: regsubsets.formula(Y X1 + X2 + X3, data = a, method = "forward", force.in = 3, force.out = 2.9, nvmax = 3)
3 Variables (and intercept)
Forced in Forced out
X3 FALSE FALSE
X1 FALSE TRUE
X2 TRUE FALSE
1 subsets of each size up to 3
Selection Algorithm: forward
X3 X1 X2
2 (1) "*" "*" " "

X1, X3 is the best.
1 ¿ -1.142/sqrt(0.04613854) [1] -5.316602 ¿ -0.442/sqrt(0.24203036) [1] -0.8984364 ¿ -13.470/sqrt(50.4051969) [1] -1.897273 X2

(b)0.1

(c)Subset selection object
Call: regsubsets.formula(Y X1 + X2 + X3, data = a, method = "forward", force.in = 3, force.out = NULL)
3 Variables (and intercept)
Forced in Forced out
X3 FALSE FALSE
X1 FALSE FALSE
X2 TRUE FALSE
1 subsets of each size up to 3
Selection Algorithm: forward
X3 X1 X2
2 (1) "*" "*" " "
3 (1) "*" "*" "*"

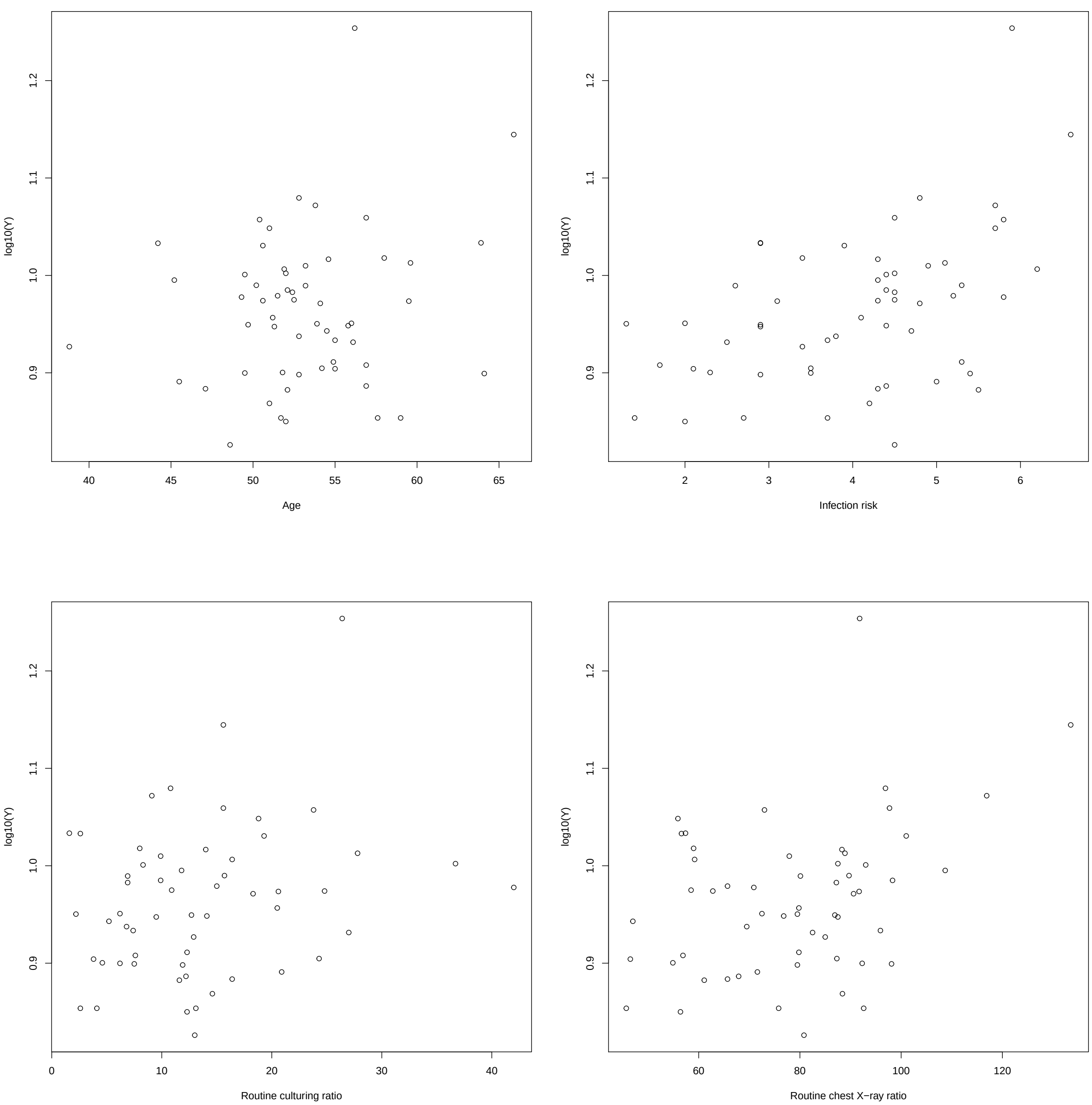
X1 and X3 is the best.

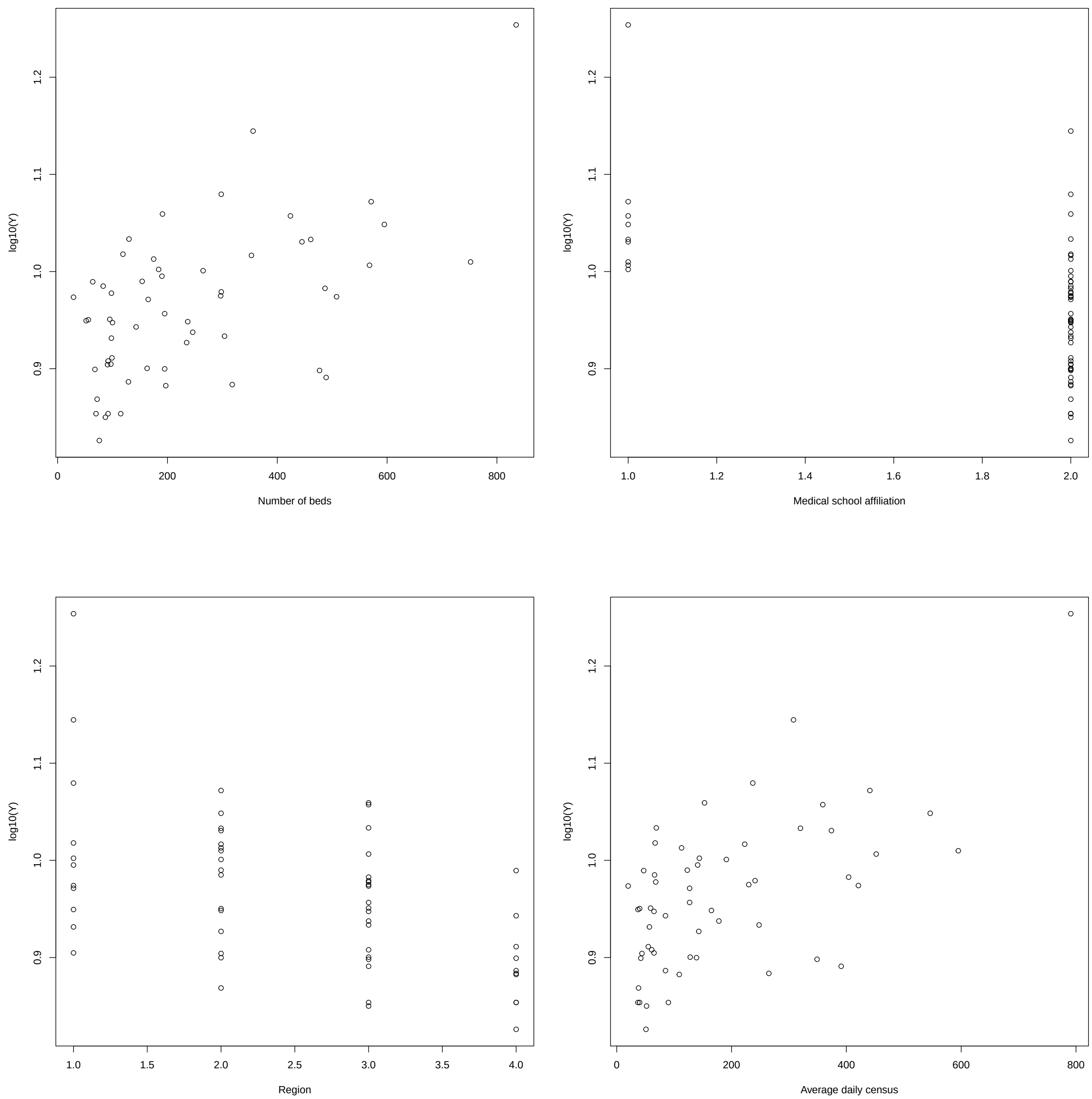
(d)Subset selection object
Call: regsubsets.formula(Y X1 + X2 + X3, data = a, method = "backward", force.out = 2.9, force.in = NULL)
3 Variables (and intercept)
Forced in Forced out
X1 FALSE FALSE
X3 FALSE TRUE

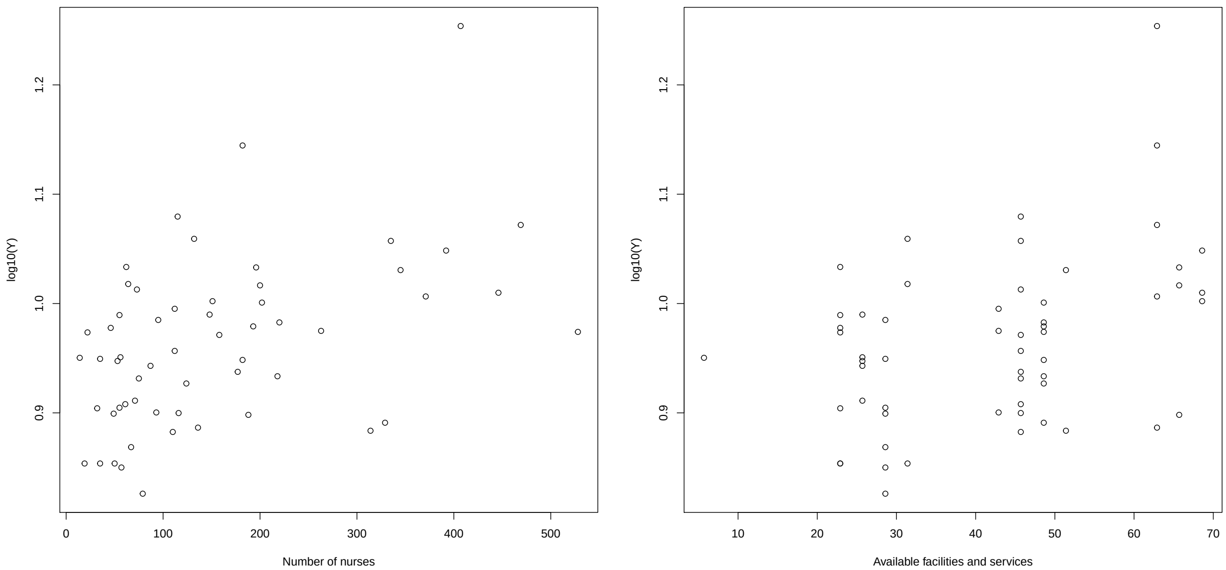
X2 FALSE FALSE
1 subsets of each size up to 3
Selection Algorithm: backward
X1 X3 X2
1 (1) "*" " " " "
2 (1) "*" "*" " "

X1 and X3 is the best. Problem 9.25

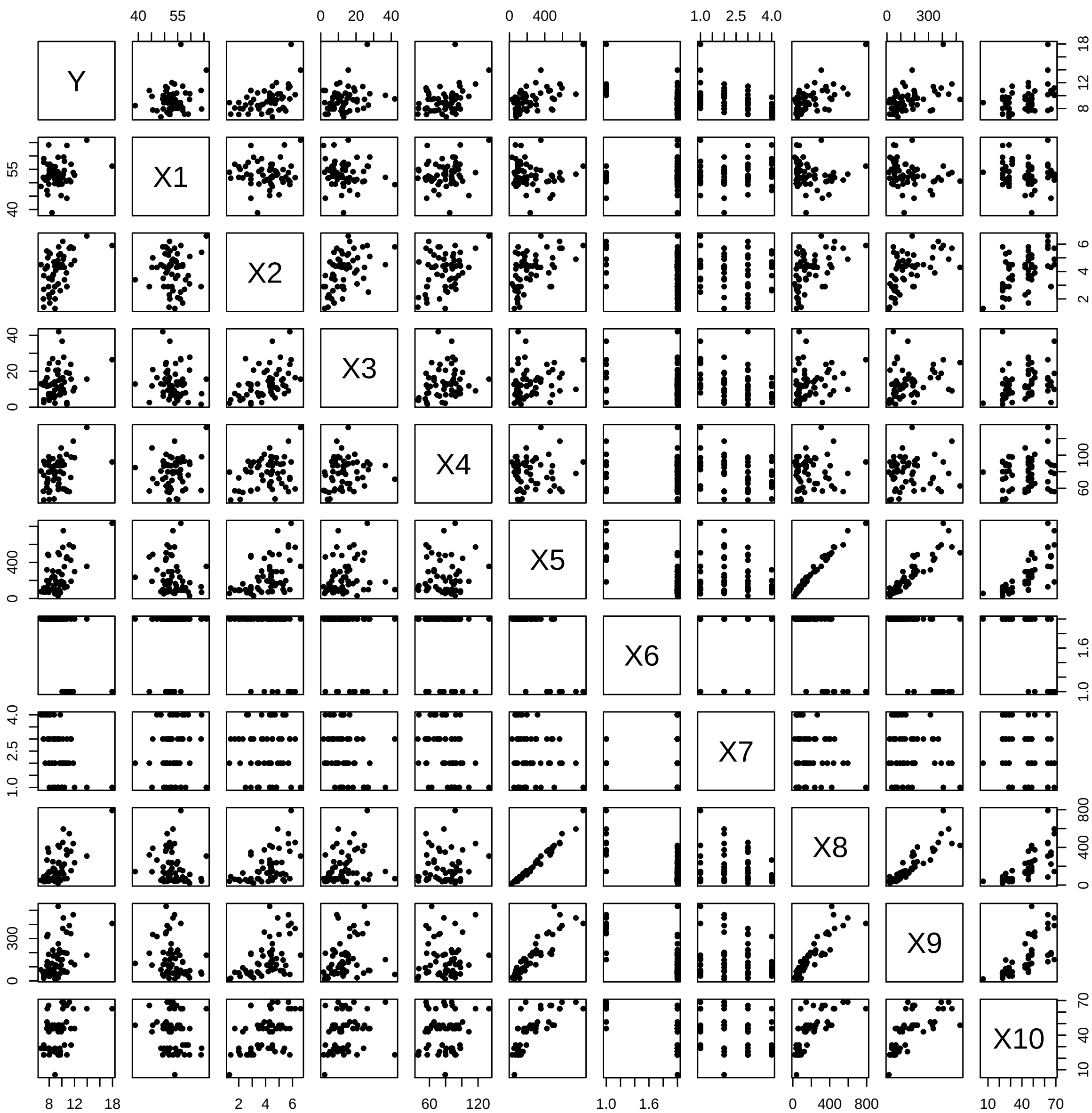
(a)




log10(Y)
Number of beds
Medical school affiliation
Region
Average daily census



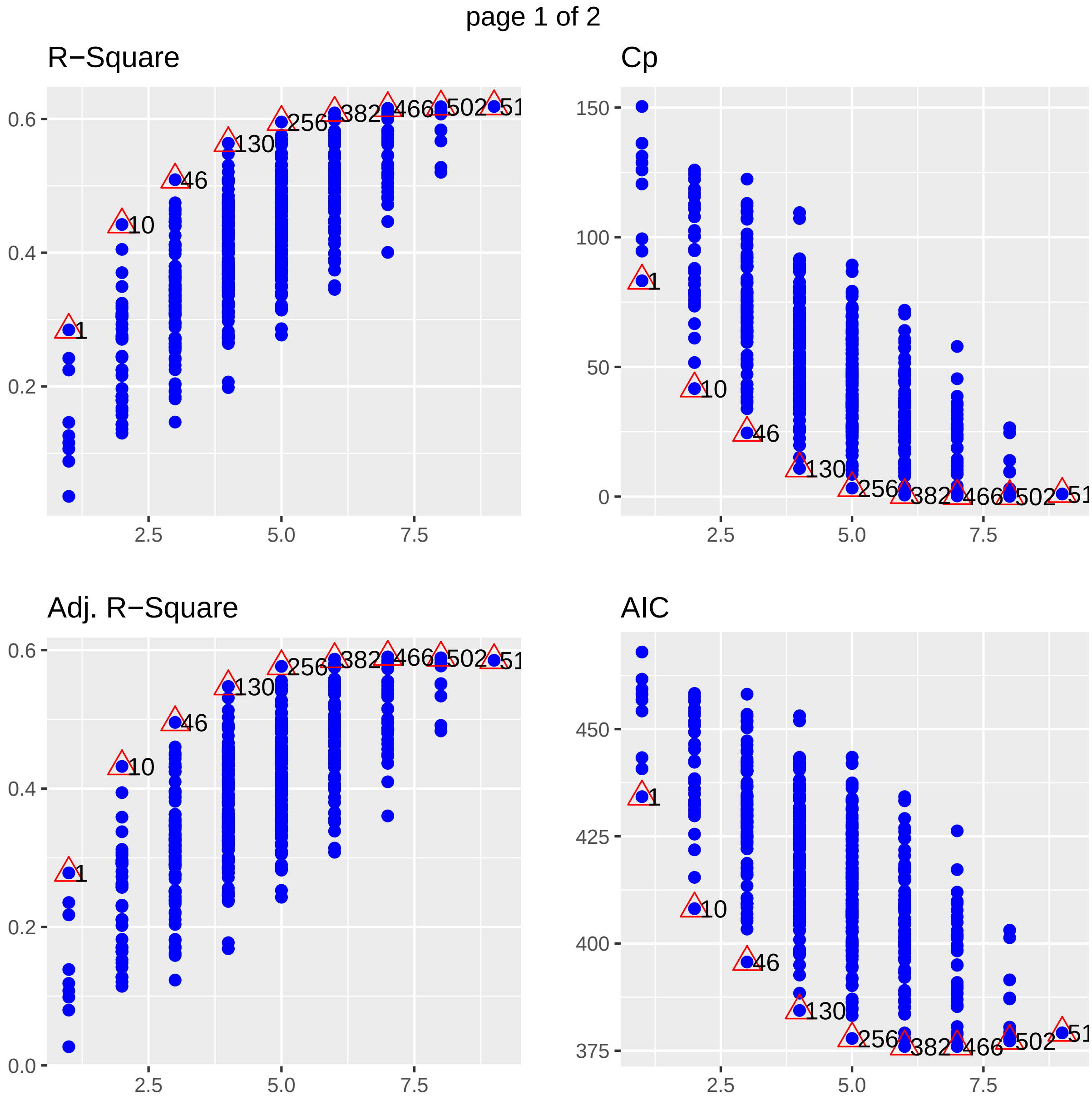
No obvious patterns in all plots but the one with region shows linear relation with Y. (b)



	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Y	1.0000000	0.18962190	0.47003068	0.2577142	0.36393576	0.59139525	-0.52689028	-0.49121696	0.63458724	0.46587415	0.4009459
X1	0.1896219	1.00000000	0.02518057	-0.1011280	0.16098871	-0.19786822	0.15819893	0.09307473	-0.17221214	-0.23643184	-0.1635187
X2	0.4700307	0.02518057	1.00000000	0.4478327	0.33395894	0.49007435	-0.34628809	-0.05882656	0.50084503	0.53008986	0.4533377
X3	0.2577142	-0.10112797	0.44783272	1.0000000	0.19481738	0.16780259	-0.24127178	-0.30577836	0.20362482	0.23884392	0.2395449
X4	0.3639358	0.16098871	0.33395894	0.1948174	1.00000000	0.06682347	-0.02379137	-0.34588311	0.08554147	0.06020464	0.1283289
X5	0.5913953	-0.19786822	0.49007435	0.1678026	0.06682347	1.00000000	-0.68689148	-0.23874029	0.99000302	0.90892888	0.7644784
X6	-0.5268903	0.15819893	-0.34628809	-0.2412718	-0.02379137	-0.68689148	1.00000000	0.22177719	-0.69224944	-0.64569908	-0.5667642
X7	-0.4912170	0.09307473	-0.05882656	-0.3057784	-0.34588311	-0.23874029	0.22177719	1.00000000	-0.25493552	-0.20555479	-0.2737778
X8	0.6345872	-0.17221214	0.50084503	0.2036248	0.08554147	0.99000302	-0.69224944	-0.25493552	1.00000000	0.90388584	0.7294165
X9	0.4658741	-0.23643184	0.53008986	0.2388439	0.06020464	0.90892888	-0.64569908	-0.20555479	0.90388584	1.00000000	0.7070559
X10	0.4009459	-0.16351870	0.45333767	0.2395449	0.12832885	0.76447839	-0.56676422	-0.27377784	0.72941653	0.70705586	1.0000000

It shows strong linear relation bewteen X5 and X8,X5 and X9,X5 and X10, X8 and X9, X8 and X10, X9 and X10.

(c)



Model480 X1 X2 X3 X7 X8 X9 X10
Model502 X1 X2 X3 X4 X6 X7 X8 X9
Model511 X1 X2 X3 X4 X6 X7 X8 X9 X10
Model511 appears to have the smallest bias.

Problem 9.27

(a)Model480V: (Intercept)

	X1	X2	X3	X7	X8	X9	X10
4.298160	0.069521	0.605099	0.012465	-0.601589	0.009401	-0.006426	-0.009427

Model480: (Intercept)

(Intercept)	X1	X2	X3	X7	X8	X9	X10
4.309040	0.097872	0.404608	-0.001677	-0.726276	0.011439	-0.007183	-0.022343

Model502V: (Intercept)

(Intercept)	X1	X2	X3	X4	X6	X7	X8	X9
2.319181	0.074137	0.493366	0.009625	0.014410	0.331678	-0.561383	0.009655	-0.006866

Model502: (Intercept)

(Intercept)	X1	X2	X3	X4	X6	X7	X8	X9
5.028361	0.096057	0.331179	-0.005930	0.009886	-0.957489	-0.615381	0.009280	-0.007817

Model511V: (Intercept)

(Intercept)	X1	X2	X3	X4	X6	X7	X8	X9	X10
2.529723	0.073823	0.508292	0.008505	0.014219	0.315916	-0.571376	0.009792	-0.006512	-0.006062

Model511: (Intercept)

(Intercept)	X1	X2	X3	X4	X6	X7	X8	X9	X10
6.072872	0.093777	0.361420	-0.005569	0.010288	-1.065482	-0.651646	0.010213	-0.007170	-0.025497