Chad Huntebrinker’s Homework 10

Chad Huntebrinker

2024-11-17

Problem 9.10 and 9.11

#Chad Huntebrinker  
  
library(leaps)

## Warning: package 'leaps' was built under R version 4.4.2

library(readxl)  
  
excel\_data <- read\_excel("Job\_Proficiency\_Data.xlsx")  
  
#Problem 9.10a  
stem(excel\_data$X1)

##   
## The decimal point is 1 digit(s) to the right of the |  
##   
## 6 | 248  
## 8 | 4671468  
## 10 | 014456902  
## 12 | 0003  
## 14 | 00

stem(excel\_data$X2)

##   
## The decimal point is 1 digit(s) to the right of the |  
##   
## 6 | 37  
## 8 | 135947  
## 10 | 127034789  
## 12 | 01112599

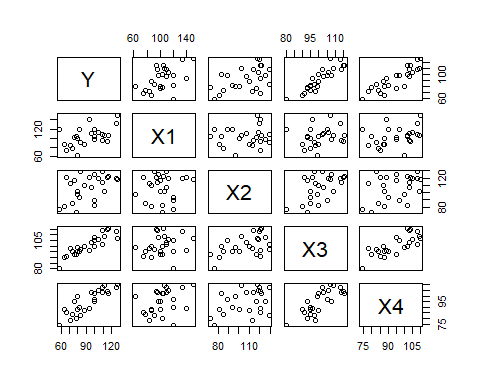
stem(excel\_data$X3)

##   
## The decimal point is 1 digit(s) to the right of the |  
##   
## 8 | 0  
## 9 | 01335556789  
## 10 | 002356789  
## 11 | 3456

stem(excel\_data$X4)

##   
## The decimal point is 1 digit(s) to the right of the |  
##   
## 7 | 48  
## 8 | 03457889  
## 9 | 0557  
## 10 | 0223345889  
## 11 | 0

#Some noteworthy features include:  
#1) All tests have max scores that go over 100, but some max scores are 116 while others  
#are 140  
#2) The distribution of the data for X2 seem to be more on the larger side (greater than  
# 110) while the distribution for X1 are on the less side (less than 110)  
  
#Problem 9.10b  
pairs(excel\_data)



cor(excel\_data)

## Y X1 X2 X3 X4  
## Y 1.0000000 0.5144107 0.4970057 0.8970645 0.8693865  
## X1 0.5144107 1.0000000 0.1022689 0.1807692 0.3266632  
## X2 0.4970057 0.1022689 1.0000000 0.5190448 0.3967101  
## X3 0.8970645 0.1807692 0.5190448 1.0000000 0.7820385  
## X4 0.8693865 0.3266632 0.3967101 0.7820385 1.0000000

cor(excel\_data[,-1])

## X1 X2 X3 X4  
## X1 1.0000000 0.1022689 0.1807692 0.3266632  
## X2 0.1022689 1.0000000 0.5190448 0.3967101  
## X3 0.1807692 0.5190448 1.0000000 0.7820385  
## X4 0.3266632 0.3967101 0.7820385 1.0000000

#It looks like there might be some correlation concerns with Y for X3 and X4  
#There also might be some concerns with X3 and X4  
  
#Problem 9.10c  
model\_1 <- lm(Y~X1 + X2 + X3 + X4, data=excel\_data)  
sum\_of\_model\_1 <- summary(model\_1)  
#Y = -124.38182 + 0.29573X1 + 0.04829X2 + 1.30601X3 + 0.51982X4  
sum\_of\_model\_1

##   
## Call:  
## lm(formula = Y ~ X1 + X2 + X3 + X4, data = excel\_data)  
##   
## 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 0.29573 0.04397 6.725 1.52e-06 \*\*\*  
## X2 0.04829 0.05662 0.853 0.40383   
## X3 1.30601 0.16409 7.959 1.26e-07 \*\*\*  
## X4 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

sum\_of\_model\_1$adj.r.squared

## [1] 0.9554702

#the adjusted R^2 has a good score (0.9554702) and the p-value for these predictor variables  
#seem okay for all except one. X2 has a p-value of 0.40383, suggesting that it might be  
#better to drop X2 from the model  
  
#Problem 9.11a  
ma <- regsubsets(Y~., nbest = 4, data=excel\_data)  
(sma <- summary(ma))

## Subset selection object  
## Call: regsubsets.formula(Y ~ ., nbest = 4, data = excel\_data)  
## 4 Variables (and intercept)  
## Forced in Forced out  
## X1 FALSE FALSE  
## X2 FALSE FALSE  
## X3 FALSE FALSE  
## X4 FALSE FALSE  
## 4 subsets of each size up to 4  
## Selection Algorithm: exhaustive  
## X1 X2 X3 X4   
## 1 ( 1 ) " " " " "\*" " "  
## 1 ( 2 ) " " " " " " "\*"  
## 1 ( 3 ) "\*" " " " " " "  
## 1 ( 4 ) " " "\*" " " " "  
## 2 ( 1 ) "\*" " " "\*" " "  
## 2 ( 2 ) " " " " "\*" "\*"  
## 2 ( 3 ) "\*" " " " " "\*"  
## 2 ( 4 ) " " "\*" "\*" " "  
## 3 ( 1 ) "\*" " " "\*" "\*"  
## 3 ( 2 ) "\*" "\*" "\*" " "  
## 3 ( 3 ) " " "\*" "\*" "\*"  
## 3 ( 4 ) "\*" "\*" " " "\*"  
## 4 ( 1 ) "\*" "\*" "\*" "\*"

sma$adjr2

## [1] 0.7962344 0.7452170 0.2326452 0.2142762 0.9269043 0.8660988 0.7984716  
## [8] 0.7884436 0.9560482 0.9246779 0.8616797 0.8232664 0.9554702

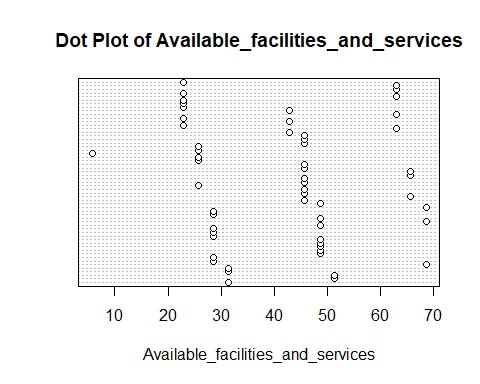
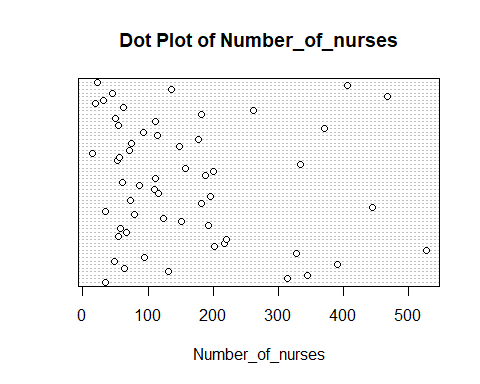
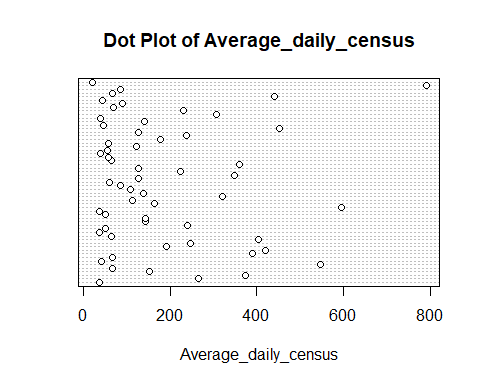
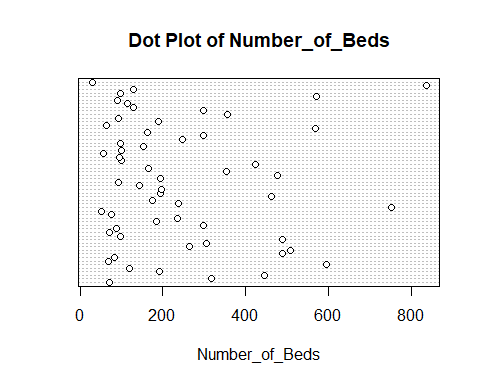
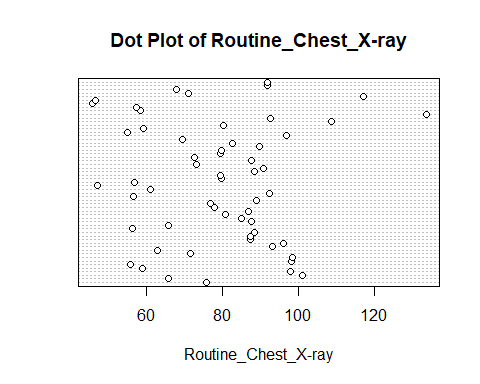
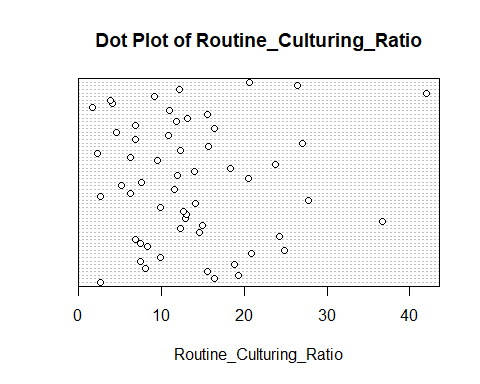
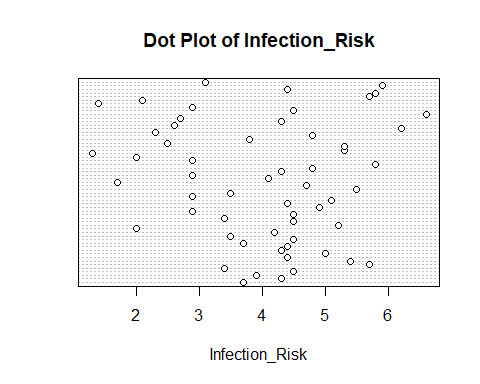
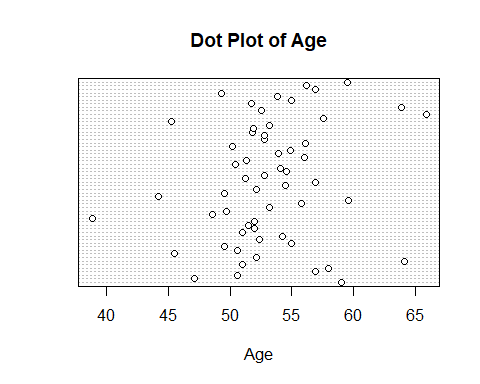
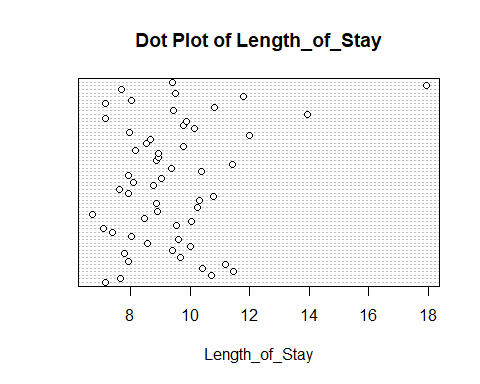
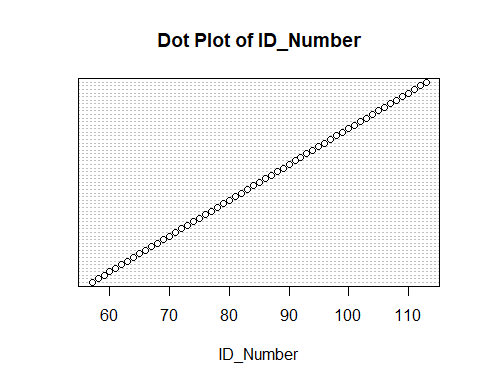
order(sma$adjr2, decreasing = TRUE)[1:4]

## [1] 9 13 5 10

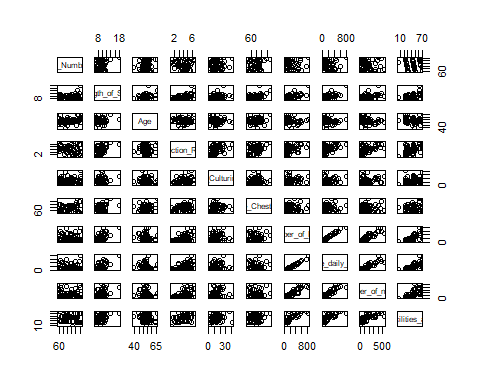
#(1) [X1, X3, X4] = 0.9560482   
#(2) [X1, X2, X3, X4] = 0.9554702  
#(3) [X1, X3] = 0.9269043   
#(4) [X1, X2, X3] = 0.9246779  
  
#9.11b  
#There are a couple of other criteria we could look at to narrow down what model we should use.  
#First, we could use some of the other measurements to see if a model is a good fit (like AIC, SSEp,  
# and PRESSp) to see if they agree. Another thing we could do is look at the residuals of these models  
#to see if they fit the model well. Finally, we can look at the simplicity of the model. In general,  
#it's better to have a model that has less predictive variables than one that has more.

Problem 9.25

#Chad Huntebrinker  
  
library(leaps)  
library(readxl)  
  
excel\_data <- read\_excel("SENIC\_Data.xlsx")  
#Use only rows 57 thru 113 and remove Region and Medical Affiliation  
model\_data <- excel\_data[57:113, ]  
model\_data <- model\_data[, -9]  
model\_data <- model\_data[, -8]  
#Note: I believe we should also remove the ID number from the data we use as this shouldn't have  
#an impact on the patient's length of stay. But we'll keep it in as the instructions didn't say  
#to remove it.  
  
#Problem 9.25a  
for (col\_name in names(model\_data)) {  
 dotchart(model\_data[[col\_name]], main = paste("Dot Plot of", col\_name), xlab = col\_name)  
}



#Variables that have outliers:  
#Number of Beds  
#Average daily census  
#Number of nurses  
  
#Variables that have spaces in the data:  
#Available facilities and services  
  
#Also, ID number is completely linear as patients are assigned them as they are admitted  
  
#Problem 9.25b  
pairs(model\_data)



cor(model\_data)

## ID\_Number Length\_of\_Stay Age  
## ID\_Number 1.000000000 0.2172148 0.20080286  
## Length\_of\_Stay 0.217214831 1.0000000 0.18962190  
## Age 0.200802862 0.1896219 1.00000000  
## Infection\_Risk -0.074891449 0.4700307 0.02518057  
## Routine\_Culturing\_Ratio -0.006311356 0.2577142 -0.10112797  
## Routine\_Chest\_X-ray -0.017645336 0.3639358 0.16098871  
## Number\_of\_Beds -0.059173991 0.5913953 -0.19786822  
## Average\_daily\_census -0.055113659 0.6345872 -0.17221214  
## Number\_of\_nurses -0.147822193 0.4658741 -0.23643184  
## Available\_facilities\_and\_services -0.095798792 0.4009459 -0.16351870  
## Infection\_Risk Routine\_Culturing\_Ratio  
## ID\_Number -0.07489145 -0.006311356  
## Length\_of\_Stay 0.47003068 0.257714155  
## Age 0.02518057 -0.101127969  
## Infection\_Risk 1.00000000 0.447832724  
## Routine\_Culturing\_Ratio 0.44783272 1.000000000  
## Routine\_Chest\_X-ray 0.33395894 0.194817381  
## Number\_of\_Beds 0.49007435 0.167802589  
## Average\_daily\_census 0.50084503 0.203624821  
## Number\_of\_nurses 0.53008986 0.238843918  
## Available\_facilities\_and\_services 0.45333767 0.239544944  
## Routine\_Chest\_X-ray Number\_of\_Beds  
## ID\_Number -0.01764534 -0.05917399  
## Length\_of\_Stay 0.36393576 0.59139525  
## Age 0.16098871 -0.19786822  
## Infection\_Risk 0.33395894 0.49007435  
## Routine\_Culturing\_Ratio 0.19481738 0.16780259  
## Routine\_Chest\_X-ray 1.00000000 0.06682347  
## Number\_of\_Beds 0.06682347 1.00000000  
## Average\_daily\_census 0.08554147 0.99000302  
## Number\_of\_nurses 0.06020464 0.90892888  
## Available\_facilities\_and\_services 0.12832885 0.76447839  
## Average\_daily\_census Number\_of\_nurses  
## ID\_Number -0.05511366 -0.14782219  
## Length\_of\_Stay 0.63458724 0.46587415  
## Age -0.17221214 -0.23643184  
## Infection\_Risk 0.50084503 0.53008986  
## Routine\_Culturing\_Ratio 0.20362482 0.23884392  
## Routine\_Chest\_X-ray 0.08554147 0.06020464  
## Number\_of\_Beds 0.99000302 0.90892888  
## Average\_daily\_census 1.00000000 0.90388584  
## Number\_of\_nurses 0.90388584 1.00000000  
## Available\_facilities\_and\_services 0.72941653 0.70705586  
## Available\_facilities\_and\_services  
## ID\_Number -0.09579879  
## Length\_of\_Stay 0.40094591  
## Age -0.16351870  
## Infection\_Risk 0.45333767  
## Routine\_Culturing\_Ratio 0.23954494  
## Routine\_Chest\_X-ray 0.12832885  
## Number\_of\_Beds 0.76447839  
## Average\_daily\_census 0.72941653  
## Number\_of\_nurses 0.70705586  
## Available\_facilities\_and\_services 1.00000000

#Remove ID number and length of stay  
cor(model\_data[,-2])

## ID\_Number Age Infection\_Risk  
## ID\_Number 1.000000000 0.20080286 -0.07489145  
## Age 0.200802862 1.00000000 0.02518057  
## Infection\_Risk -0.074891449 0.02518057 1.00000000  
## Routine\_Culturing\_Ratio -0.006311356 -0.10112797 0.44783272  
## Routine\_Chest\_X-ray -0.017645336 0.16098871 0.33395894  
## Number\_of\_Beds -0.059173991 -0.19786822 0.49007435  
## Average\_daily\_census -0.055113659 -0.17221214 0.50084503  
## Number\_of\_nurses -0.147822193 -0.23643184 0.53008986  
## Available\_facilities\_and\_services -0.095798792 -0.16351870 0.45333767  
## Routine\_Culturing\_Ratio Routine\_Chest\_X-ray  
## ID\_Number -0.006311356 -0.01764534  
## Age -0.101127969 0.16098871  
## Infection\_Risk 0.447832724 0.33395894  
## Routine\_Culturing\_Ratio 1.000000000 0.19481738  
## Routine\_Chest\_X-ray 0.194817381 1.00000000  
## Number\_of\_Beds 0.167802589 0.06682347  
## Average\_daily\_census 0.203624821 0.08554147  
## Number\_of\_nurses 0.238843918 0.06020464  
## Available\_facilities\_and\_services 0.239544944 0.12832885  
## Number\_of\_Beds Average\_daily\_census  
## ID\_Number -0.05917399 -0.05511366  
## Age -0.19786822 -0.17221214  
## Infection\_Risk 0.49007435 0.50084503  
## Routine\_Culturing\_Ratio 0.16780259 0.20362482  
## Routine\_Chest\_X-ray 0.06682347 0.08554147  
## Number\_of\_Beds 1.00000000 0.99000302  
## Average\_daily\_census 0.99000302 1.00000000  
## Number\_of\_nurses 0.90892888 0.90388584  
## Available\_facilities\_and\_services 0.76447839 0.72941653  
## Number\_of\_nurses  
## ID\_Number -0.14782219  
## Age -0.23643184  
## Infection\_Risk 0.53008986  
## Routine\_Culturing\_Ratio 0.23884392  
## Routine\_Chest\_X-ray 0.06020464  
## Number\_of\_Beds 0.90892888  
## Average\_daily\_census 0.90388584  
## Number\_of\_nurses 1.00000000  
## Available\_facilities\_and\_services 0.70705586  
## Available\_facilities\_and\_services  
## ID\_Number -0.09579879  
## Age -0.16351870  
## Infection\_Risk 0.45333767  
## Routine\_Culturing\_Ratio 0.23954494  
## Routine\_Chest\_X-ray 0.12832885  
## Number\_of\_Beds 0.76447839  
## Average\_daily\_census 0.72941653  
## Number\_of\_nurses 0.70705586  
## Available\_facilities\_and\_services 1.00000000

#Major concern with linear pairwise associations:  
#number of beds and average daily census = 0.99000302  
#average daily census and number of nurses 0.90388584  
#number of beds and number of nurses = 0.90892888  
  
#Minor concern:  
#available facilities and services and number of nurses = 0.7070559  
#available facilities and services and number of beds = 0.7644784  
#average daily census and available facilities and services = 0.7294165  
  
#Problem 9.25c  
model\_subsets <- regsubsets(log(Length\_of\_Stay)~., nbest = 12, data=model\_data)  
(sum\_of\_model\_subsets <- summary(model\_subsets))

## Subset selection object  
## Call: regsubsets.formula(log(Length\_of\_Stay) ~ ., nbest = 12, data = model\_data)  
## 9 Variables (and intercept)  
## Forced in Forced out  
## ID\_Number FALSE FALSE  
## Age FALSE FALSE  
## Infection\_Risk FALSE FALSE  
## Routine\_Culturing\_Ratio FALSE FALSE  
## `Routine\_Chest\_X-ray` FALSE FALSE  
## Number\_of\_Beds FALSE FALSE  
## Average\_daily\_census FALSE FALSE  
## Number\_of\_nurses FALSE FALSE  
## Available\_facilities\_and\_services FALSE FALSE  
## 12 subsets of each size up to 8  
## Selection Algorithm: exhaustive  
## ID\_Number Age Infection\_Risk Routine\_Culturing\_Ratio  
## 1 ( 1 ) " " " " " " " "   
## 1 ( 2 ) " " " " " " " "   
## 1 ( 3 ) " " " " "\*" " "   
## 1 ( 4 ) " " " " " " " "   
## 1 ( 5 ) " " " " " " " "   
## 1 ( 6 ) " " " " " " " "   
## 1 ( 7 ) " " " " " " "\*"   
## 1 ( 8 ) "\*" " " " " " "   
## 1 ( 9 ) " " "\*" " " " "   
## 2 ( 1 ) " " " " " " " "   
## 2 ( 2 ) " " "\*" " " " "   
## 2 ( 3 ) " " " " " " " "   
## 2 ( 4 ) "\*" " " " " " "   
## 2 ( 5 ) " " "\*" " " " "   
## 2 ( 6 ) " " " " " " " "   
## 2 ( 7 ) " " " " "\*" " "   
## 2 ( 8 ) " " " " " " " "   
## 2 ( 9 ) " " " " " " "\*"   
## 2 ( 10 ) "\*" " " " " " "   
## 2 ( 11 ) " " " " "\*" " "   
## 2 ( 12 ) " " " " " " " "   
## 3 ( 1 ) "\*" " " " " " "   
## 3 ( 2 ) " " "\*" " " " "   
## 3 ( 3 ) " " " " " " " "   
## 3 ( 4 ) " " " " " " " "   
## 3 ( 5 ) " " "\*" " " " "   
## 3 ( 6 ) "\*" " " " " " "   
## 3 ( 7 ) "\*" "\*" " " " "   
## 3 ( 8 ) " " " " "\*" " "   
## 3 ( 9 ) " " " " " " " "   
## 3 ( 10 ) " " " " " " "\*"   
## 3 ( 11 ) " " "\*" "\*" " "   
## 3 ( 12 ) " " "\*" " " "\*"   
## 4 ( 1 ) "\*" "\*" " " " "   
## 4 ( 2 ) "\*" " " " " " "   
## 4 ( 3 ) " " "\*" " " " "   
## 4 ( 4 ) "\*" " " " " " "   
## 4 ( 5 ) " " "\*" " " " "   
## 4 ( 6 ) "\*" " " "\*" " "   
## 4 ( 7 ) " " "\*" " " "\*"   
## 4 ( 8 ) "\*" "\*" " " " "   
## 4 ( 9 ) "\*" " " " " "\*"   
## 4 ( 10 ) "\*" " " " " " "   
## 4 ( 11 ) " " "\*" "\*" " "   
## 4 ( 12 ) " " "\*" " " " "   
## 5 ( 1 ) "\*" "\*" " " " "   
## 5 ( 2 ) "\*" "\*" " " "\*"   
## 5 ( 3 ) "\*" "\*" " " " "   
## 5 ( 4 ) "\*" "\*" "\*" " "   
## 5 ( 5 ) "\*" "\*" " " " "   
## 5 ( 6 ) "\*" " " "\*" " "   
## 5 ( 7 ) "\*" " " "\*" " "   
## 5 ( 8 ) "\*" " " " " " "   
## 5 ( 9 ) " " "\*" " " "\*"   
## 5 ( 10 ) " " "\*" "\*" " "   
## 5 ( 11 ) "\*" " " " " "\*"   
## 5 ( 12 ) "\*" " " " " "\*"   
## 6 ( 1 ) "\*" "\*" "\*" " "   
## 6 ( 2 ) "\*" "\*" "\*" " "   
## 6 ( 3 ) "\*" "\*" " " "\*"   
## 6 ( 4 ) "\*" "\*" " " " "   
## 6 ( 5 ) "\*" "\*" " " "\*"   
## 6 ( 6 ) "\*" "\*" " " "\*"   
## 6 ( 7 ) "\*" "\*" " " " "   
## 6 ( 8 ) "\*" " " "\*" " "   
## 6 ( 9 ) "\*" "\*" "\*" " "   
## 6 ( 10 ) "\*" "\*" "\*" "\*"   
## 6 ( 11 ) "\*" "\*" " " " "   
## 6 ( 12 ) "\*" " " "\*" " "   
## 7 ( 1 ) "\*" "\*" "\*" " "   
## 7 ( 2 ) "\*" "\*" "\*" "\*"   
## 7 ( 3 ) "\*" "\*" " " "\*"   
## 7 ( 4 ) "\*" "\*" "\*" " "   
## 7 ( 5 ) "\*" "\*" " " "\*"   
## 7 ( 6 ) "\*" "\*" "\*" "\*"   
## 7 ( 7 ) "\*" "\*" "\*" " "   
## 7 ( 8 ) "\*" "\*" " " "\*"   
## 7 ( 9 ) "\*" "\*" " " " "   
## 7 ( 10 ) "\*" "\*" "\*" "\*"   
## 7 ( 11 ) "\*" " " "\*" "\*"   
## 7 ( 12 ) "\*" " " "\*" " "   
## 8 ( 1 ) "\*" "\*" "\*" "\*"   
## 8 ( 2 ) "\*" "\*" "\*" " "   
## 8 ( 3 ) "\*" "\*" "\*" "\*"   
## 8 ( 4 ) "\*" "\*" " " "\*"   
## 8 ( 5 ) "\*" "\*" "\*" "\*"   
## 8 ( 6 ) "\*" " " "\*" "\*"   
## 8 ( 7 ) "\*" "\*" "\*" "\*"   
## 8 ( 8 ) " " "\*" "\*" "\*"   
## 8 ( 9 ) "\*" "\*" "\*" "\*"   
## `Routine\_Chest\_X-ray` Number\_of\_Beds Average\_daily\_census  
## 1 ( 1 ) " " " " "\*"   
## 1 ( 2 ) " " "\*" " "   
## 1 ( 3 ) " " " " " "   
## 1 ( 4 ) " " " " " "   
## 1 ( 5 ) " " " " " "   
## 1 ( 6 ) "\*" " " " "   
## 1 ( 7 ) " " " " " "   
## 1 ( 8 ) " " " " " "   
## 1 ( 9 ) " " " " " "   
## 2 ( 1 ) "\*" " " "\*"   
## 2 ( 2 ) " " " " "\*"   
## 2 ( 3 ) "\*" "\*" " "   
## 2 ( 4 ) " " " " "\*"   
## 2 ( 5 ) " " "\*" " "   
## 2 ( 6 ) " " "\*" "\*"   
## 2 ( 7 ) " " " " "\*"   
## 2 ( 8 ) " " " " "\*"   
## 2 ( 9 ) " " " " "\*"   
## 2 ( 10 ) " " "\*" " "   
## 2 ( 11 ) " " "\*" " "   
## 2 ( 12 ) " " " " "\*"   
## 3 ( 1 ) "\*" " " "\*"   
## 3 ( 2 ) "\*" " " "\*"   
## 3 ( 3 ) "\*" " " "\*"   
## 3 ( 4 ) "\*" "\*" "\*"   
## 3 ( 5 ) "\*" "\*" " "   
## 3 ( 6 ) "\*" "\*" " "   
## 3 ( 7 ) " " " " "\*"   
## 3 ( 8 ) "\*" " " "\*"   
## 3 ( 9 ) "\*" " " "\*"   
## 3 ( 10 ) "\*" " " "\*"   
## 3 ( 11 ) " " " " "\*"   
## 3 ( 12 ) " " " " "\*"   
## 4 ( 1 ) "\*" " " "\*"   
## 4 ( 2 ) "\*" "\*" "\*"   
## 4 ( 3 ) "\*" " " "\*"   
## 4 ( 4 ) "\*" " " "\*"   
## 4 ( 5 ) "\*" "\*" "\*"   
## 4 ( 6 ) "\*" " " "\*"   
## 4 ( 7 ) "\*" " " "\*"   
## 4 ( 8 ) "\*" "\*" " "   
## 4 ( 9 ) "\*" " " "\*"   
## 4 ( 10 ) "\*" " " "\*"   
## 4 ( 11 ) "\*" " " "\*"   
## 4 ( 12 ) "\*" " " "\*"   
## 5 ( 1 ) "\*" "\*" "\*"   
## 5 ( 2 ) "\*" " " "\*"   
## 5 ( 3 ) "\*" " " "\*"   
## 5 ( 4 ) "\*" " " "\*"   
## 5 ( 5 ) "\*" " " "\*"   
## 5 ( 6 ) "\*" " " "\*"   
## 5 ( 7 ) "\*" "\*" "\*"   
## 5 ( 8 ) "\*" "\*" "\*"   
## 5 ( 9 ) "\*" " " "\*"   
## 5 ( 10 ) "\*" " " "\*"   
## 5 ( 11 ) "\*" " " "\*"   
## 5 ( 12 ) "\*" "\*" "\*"   
## 6 ( 1 ) "\*" " " "\*"   
## 6 ( 2 ) "\*" "\*" "\*"   
## 6 ( 3 ) "\*" " " "\*"   
## 6 ( 4 ) "\*" "\*" "\*"   
## 6 ( 5 ) "\*" "\*" "\*"   
## 6 ( 6 ) "\*" " " "\*"   
## 6 ( 7 ) "\*" "\*" "\*"   
## 6 ( 8 ) "\*" "\*" "\*"   
## 6 ( 9 ) "\*" " " "\*"   
## 6 ( 10 ) "\*" " " "\*"   
## 6 ( 11 ) "\*" " " "\*"   
## 6 ( 12 ) "\*" " " "\*"   
## 7 ( 1 ) "\*" "\*" "\*"   
## 7 ( 2 ) "\*" " " "\*"   
## 7 ( 3 ) "\*" "\*" "\*"   
## 7 ( 4 ) "\*" " " "\*"   
## 7 ( 5 ) "\*" " " "\*"   
## 7 ( 6 ) "\*" "\*" "\*"   
## 7 ( 7 ) "\*" "\*" "\*"   
## 7 ( 8 ) "\*" "\*" "\*"   
## 7 ( 9 ) "\*" "\*" "\*"   
## 7 ( 10 ) "\*" " " "\*"   
## 7 ( 11 ) "\*" "\*" "\*"   
## 7 ( 12 ) "\*" "\*" "\*"   
## 8 ( 1 ) "\*" "\*" "\*"   
## 8 ( 2 ) "\*" "\*" "\*"   
## 8 ( 3 ) "\*" " " "\*"   
## 8 ( 4 ) "\*" "\*" "\*"   
## 8 ( 5 ) "\*" "\*" "\*"   
## 8 ( 6 ) "\*" "\*" "\*"   
## 8 ( 7 ) "\*" "\*" " "   
## 8 ( 8 ) "\*" "\*" "\*"   
## 8 ( 9 ) " " "\*" "\*"   
## Number\_of\_nurses Available\_facilities\_and\_services  
## 1 ( 1 ) " " " "   
## 1 ( 2 ) " " " "   
## 1 ( 3 ) " " " "   
## 1 ( 4 ) "\*" " "   
## 1 ( 5 ) " " "\*"   
## 1 ( 6 ) " " " "   
## 1 ( 7 ) " " " "   
## 1 ( 8 ) " " " "   
## 1 ( 9 ) " " " "   
## 2 ( 1 ) " " " "   
## 2 ( 2 ) " " " "   
## 2 ( 3 ) " " " "   
## 2 ( 4 ) " " " "   
## 2 ( 5 ) " " " "   
## 2 ( 6 ) " " " "   
## 2 ( 7 ) " " " "   
## 2 ( 8 ) "\*" " "   
## 2 ( 9 ) " " " "   
## 2 ( 10 ) " " " "   
## 2 ( 11 ) " " " "   
## 2 ( 12 ) " " "\*"   
## 3 ( 1 ) " " " "   
## 3 ( 2 ) " " " "   
## 3 ( 3 ) "\*" " "   
## 3 ( 4 ) " " " "   
## 3 ( 5 ) " " " "   
## 3 ( 6 ) " " " "   
## 3 ( 7 ) " " " "   
## 3 ( 8 ) " " " "   
## 3 ( 9 ) " " "\*"   
## 3 ( 10 ) " " " "   
## 3 ( 11 ) " " " "   
## 3 ( 12 ) " " " "   
## 4 ( 1 ) " " " "   
## 4 ( 2 ) " " " "   
## 4 ( 3 ) "\*" " "   
## 4 ( 4 ) "\*" " "   
## 4 ( 5 ) " " " "   
## 4 ( 6 ) " " " "   
## 4 ( 7 ) " " " "   
## 4 ( 8 ) " " " "   
## 4 ( 9 ) " " " "   
## 4 ( 10 ) " " "\*"   
## 4 ( 11 ) " " " "   
## 4 ( 12 ) " " "\*"   
## 5 ( 1 ) " " " "   
## 5 ( 2 ) " " " "   
## 5 ( 3 ) "\*" " "   
## 5 ( 4 ) " " " "   
## 5 ( 5 ) " " "\*"   
## 5 ( 6 ) "\*" " "   
## 5 ( 7 ) " " " "   
## 5 ( 8 ) "\*" " "   
## 5 ( 9 ) "\*" " "   
## 5 ( 10 ) "\*" " "   
## 5 ( 11 ) "\*" " "   
## 5 ( 12 ) " " " "   
## 6 ( 1 ) "\*" " "   
## 6 ( 2 ) " " " "   
## 6 ( 3 ) "\*" " "   
## 6 ( 4 ) "\*" " "   
## 6 ( 5 ) " " " "   
## 6 ( 6 ) " " "\*"   
## 6 ( 7 ) " " "\*"   
## 6 ( 8 ) "\*" " "   
## 6 ( 9 ) " " "\*"   
## 6 ( 10 ) " " " "   
## 6 ( 11 ) "\*" "\*"   
## 6 ( 12 ) "\*" "\*"   
## 7 ( 1 ) "\*" " "   
## 7 ( 2 ) "\*" " "   
## 7 ( 3 ) "\*" " "   
## 7 ( 4 ) "\*" "\*"   
## 7 ( 5 ) "\*" "\*"   
## 7 ( 6 ) " " " "   
## 7 ( 7 ) " " "\*"   
## 7 ( 8 ) " " "\*"   
## 7 ( 9 ) "\*" "\*"   
## 7 ( 10 ) " " "\*"   
## 7 ( 11 ) "\*" " "   
## 7 ( 12 ) "\*" "\*"   
## 8 ( 1 ) "\*" " "   
## 8 ( 2 ) "\*" "\*"   
## 8 ( 3 ) "\*" "\*"   
## 8 ( 4 ) "\*" "\*"   
## 8 ( 5 ) " " "\*"   
## 8 ( 6 ) "\*" "\*"   
## 8 ( 7 ) "\*" "\*"   
## 8 ( 8 ) "\*" "\*"   
## 8 ( 9 ) "\*" "\*"

sum\_of\_model\_subsets$cp

## [1] 19.388893 23.960723 36.323274 36.465548 43.103459 46.642320 54.381254  
## [8] 57.490685 58.494848 10.144822 12.416413 13.747055 15.396149 16.301905  
## [15] 16.993418 17.121005 17.422410 19.331957 19.940133 20.509882 21.004753  
## [22] 5.935733 6.218611 8.690994 9.309078 9.367642 9.487228 10.799968  
## [29] 11.160453 11.244932 11.441974 11.564080 11.716168 4.032107 5.380489  
## [36] 6.182477 6.192455 6.606737 6.665581 7.011539 7.204526 7.263889  
## [43] 7.372622 7.562821 7.628945 4.414536 4.951662 4.953849 5.118260  
## [50] 5.632481 6.027037 6.133510 6.418307 6.530534 6.704149 7.144378  
## [57] 7.160537 5.385926 5.478069 5.543931 5.788681 5.861514 6.376906  
## [64] 6.410640 6.483551 6.542487 6.593056 6.718816 7.556996 6.336248  
## [71] 6.779775 6.950371 7.020266 7.177946 7.329563 7.439387 7.788274  
## [78] 7.788467 7.930008 8.438062 8.461812 8.085196 8.308118 8.343960  
## [85] 8.881629 9.242676 10.397829 10.966718 11.188088 13.125181

order(sum\_of\_model\_subsets$cp, decreasing = FALSE)[1:3]

## [1] 34 46 47

sum\_of\_model\_subsets$which[34,]

## (Intercept) ID\_Number   
## TRUE TRUE   
## Age Infection\_Risk   
## TRUE FALSE   
## Routine\_Culturing\_Ratio `Routine\_Chest\_X-ray`   
## FALSE TRUE   
## Number\_of\_Beds Average\_daily\_census   
## FALSE TRUE   
## Number\_of\_nurses Available\_facilities\_and\_services   
## FALSE FALSE

sum\_of\_model\_subsets$which[46,]

## (Intercept) ID\_Number   
## TRUE TRUE   
## Age Infection\_Risk   
## TRUE FALSE   
## Routine\_Culturing\_Ratio `Routine\_Chest\_X-ray`   
## FALSE TRUE   
## Number\_of\_Beds Average\_daily\_census   
## TRUE TRUE   
## Number\_of\_nurses Available\_facilities\_and\_services   
## FALSE FALSE

sum\_of\_model\_subsets$which[47,]

## (Intercept) ID\_Number   
## TRUE TRUE   
## Age Infection\_Risk   
## TRUE FALSE   
## Routine\_Culturing\_Ratio `Routine\_Chest\_X-ray`   
## TRUE TRUE   
## Number\_of\_Beds Average\_daily\_census   
## FALSE TRUE   
## Number\_of\_nurses Available\_facilities\_and\_services   
## FALSE FALSE

#(1) 4.032107  
#(Intercept), ID\_Number, Age, Routine\_Chest\_X-ray, Average\_daily\_census  
5 - 4.032107

## [1] 0.967893

#(2) 4.414536  
#(Intercept), ID\_Number, Age, Routine\_Chest\_X-ray, Number\_of\_Beds, Average\_daily\_census  
6 - 4.414536

## [1] 1.585464

#(3) 4.951662  
#(Intercept), ID\_Number, Age, Routine\_Culturing\_Ratio, Routine\_Chest\_X-ray, Average\_daily\_census  
6 - 4.951662

## [1] 1.048338

#The first model (the one with with the predictor values of Intercept, ID\_Number,   
#Age, Routine\_Chest\_X-ray, Average\_daily\_census) has the least bias as the Cp score is closest to the  
#number of predictor variables it uses.