Assignment Name: HW3

Student Name: Mo Pei

3.4 ***Adjusting R^2***. Decide if the following statements are true or false, and explain why:

a. For a multiple regression problem, the adjusted coefficient of determination, R^2_{adj}, will always be smaller than the regular, unadjusted R^2.

True:

R^2= 1-(SSE/SST)

R^2 adj = 1-(SSE/(N-K-1)/SST/(N-K))= 1 – (SSE/SST)*(N-1)/(N-K-1)

So, N >0, K>0, then N-1 > N-K-1

(N-1)/(N-K-1)>1

(SSE/SST)*(N-1)/(N-K-1)> (SSE/SST)

-(SSE/SST)*(N-1)/(N-K-1) < - (SSE/SST)

1-(SSE/SST)*(N-1)/(N-K-1)<1- (SSE/SST)

R^2 adj < R^2

Another way is even if we throw one very useful predictor, the R^2 will go higher but meanwhile R^2 adj is smaller because of taking account of penalty for adding new predictor.

b. If we fit a multiple regression model and then add a new predictor to the model, the adjusted coefficient of determination, R^2_{adj}, will always increase.

False, R^2adj has a penalty if we add a new predictor. So when R square increases, the adjusted R square might decrease.

Also, from the formula, adding one predictor the total sum of square will go up but SSE might be larger than degree freedom increase. SSE/(n-k-1).

So SSE/SST is larger, the R^2 adj is smaller.

3.11 ***Active pulse rates.*** The computer output below comes from a study to model *Active* pulse rates (after climbing several flights of stairs) based on resting pulse rate (*Rest* in beats per minute), weight (*Wgt* in pounds), and amount of *Exercise* (in hours per week). The data were obtained from 232 students taking Stat2 courses in past semesters.

```
The regression equation is Active = 11.8 + 1.12 Rest + 0.0342 Wgt - 1.09 Exercise

Predictor      Coef  SE Coef      T      P
Constant      11.84    11.95   0.99  0.323
Rest         1.1194   0.1192   9.39  0.000
Wgt         0.03420  0.03173   1.08  0.282
Exercise     -1.085    1.600  -0.68  0.498

S = 15.0452   R-Sq = 36.9%   R-Sq(adj) = 36.1%
```

a. Test the hypotheses that $\beta_2 = 0$ versus $\beta_2 \neq 0$ and interpret the result in the context of this problem. You may assume that the conditions for a linear model are satisfied for these data.

So B2 is Wgt

H0: B2=0 H1: B2 not equal 0

Decision: since t=1.08 and P value = 0.282 which is greater than 5%, so we cannot reject H0

Conclusion: the data does not suggest that there is a relationship between average Active pulse rates and Wgt, after accounting Rest Pulse Rate and Exercise.

Test the hypotheses that $\beta_1 = 0$ versus $\beta_1 \neq 0$ and interpret the result in the context of this problem. You may assume that the conditions for a linear model are satisfied for these data.

So B1 is Rest

H0: B2=0 H1: B2 not equal 0

Decision: since P value = 0 which is smaller than 5%, so we can reject H0

Conclusion: the data does not suggest that there is a relationship between average Active pulse rates and Resting pulse rate, after accounting Weight and Exercise.

b. What active pulse rate would this model predict for a 200-pound student who exercises 7 hours per week and has a resting pulse rate of 76 beats per minute?

Active^=11.8 + 1.12*Rest + 0.0342*Wgt – 1.09*Exercise

= 11.8 + 1.12*76 + 0.0342*200 - 1.09*7

= 96.13

3.6 ***Modeling prices to buy a car.*** An information technology specialist used an interesting method for negotiating prices with used car sales representatives. He collected data from the entire state for the model of car that he was interested in purchasing. Then he approached the salesmen at dealerships based on the residuals of their prices in his model.

a. Should he pick dealerships that tend to have positive or negative residuals? Why?

   Residual = observed Y- predicted Y = Y-Y^

   Negative residual means the actual value below the predicted average value. So go with negative residual.

b. Write down a two-predictor regression model that would use just the *Year* of the car and its *Mileage* to predict *Price.*

   Price^= B0 + B1Year+B2Milieage

c. Why might we want to add an interaction term for *Year* × *Mileage* to the model? Would you expect the coefficient of the interaction to be positive or negative? Explain.

   Price^= B0 + B1Year+B2Milieage+B3YearMileage

   Negative. Because either a car has more year driving or more mileage usage would be cheaper. The two predictors interact with each other. The price difference between a new car with more miles and a new car with less miles is a lot greater than an old car with more miles and an old car with less miles. But, the interaction part is to adjust the amount of slope, the slope is negative.

3.14 ***Enrollments in mathematics courses.*** Refer to the model in **Exercise 3.13** to predict *Spring* mathematics enrollments with a two-predictor model based on *Fall* enrollments and academic year (*AYear*) for the data in **MathEnrollment**.

a. What percent of the variability in spring enrollment is explained by the multiple regression model based on fall enrollment and academic year?

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.933[a]	.871	.834	13.367

a. Predictors: (Constant), Fall, Ayear

R^2= 87.1%

The data suggests that 87.1% variation in spring enrollment can be explained by the model

b. What is the size of the typical error for this multiple regression model?

The standard error is 13.367

c. Provide the ANOVA table for partitioning the total variability in spring enrollment based on this model and interpret the associated F-test.

H0: B1=B2=0
H1: at least one slope is not zero
Decision: since p =0.001 smaller than 5% and so we reject null hypothesis
Conclusion: the data suggests there is relationship between the average spring enrollment and predictors fall enrollment and academic year. Together they account significant amount of variability in spring enrollment.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8446.893	2	4223.447	23.638	.001[b]
	Residual	1250.707	7	178.672		
	Total	9697.600	9			

a. Dependent Variable: Spring

b. Predictors: (Constant), Fall, Ayear

d. Are the regression coefficients for *both* explanatory variables significantly different from zero? Provide appropriate hypotheses, test statistics, and p-values in order to make your conclusion.

Yes. For B1 Ayear since t=4.566 and p value is 0.003 and so we reject null hypothesis. There is relationship between academic year and average spring enrollment. After accounting for fall enrollment.

Yes. For B2 Fall enrollment since t= -4.933 and p value is 0.002 and so we reject null hypothesis. There is relationship between fall enrollment and average spring enrollment. After accounting for academic year.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-11715.784	2686.235		-4.361	.003
	Ayear	6.107	1.337	.620	4.566	.003
	Fall	-1.007	.204	-.670	-4.933	.002

a. Dependent Variable: Spring

e. Whether drop any explanatory variable?

Do not drop any variable. Original Adjusted R^2 is 0.834 but it decreases a lot either drop academic year or fall enrollment.

Drop academic year

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.698[a]	.487	.423	24.941

a. Predictors: (Constant), Fall

Drop fall enrollment

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.650[a]	.423	.351	26.453

a. Predictors: (Constant), Ayear

3.18 ***Fish eggs.*** Researchers[7] collected samples of female lake trout from Lake Ontario in September and November 2002–2004. A goal of the study was to investigate the fertility of fish that had been stocked in the lake. One measure of the viability of fish eggs is *percent dry mass(PctDM*), which reflects the energy potential stored in the eggs by recording the percentage of the total egg material that is solid. Values of the *PctDM* for a sample of 35 lake trout (14 in September and 21 in November) are given in **Table 3.6** along with the age (in years) of the fish. The data are stored in three columns in a file called **FishEggs**.

September							
Age	7	7	7	7	9	9	11
PctDM	34.90	37.00	37.90	38.15	33.90	36.45	35.00
Age	11	12	12	12	16	17	18
PctDM	36.15	34.05	34.65	35.40	32.45	36.55	34.00
November							
Age	7	8	8	9	9	9	9
PctDM	34.90	37.00	37.90	38.15	33.90	36.45	35.00
Age	10	10	11	11	12	12	13
PctDM	36.15	34.05	34.65	35.40	32.45	36.55	34.00
Age	13	13	14	15	16	17	18
PctDM	36.15	34.05	34.65	35.40	32.45	36.55	34.00

Table 3.6: *Percent dry mass of eggs and age for female lake trout*

Ignore the month at first and fit a simple linear regression to predict the *PctDM* based on the *Age* of the fish.

a. Write down an equation for the least squares line and comment on what it appears to indicate about the relationship between *PctDM* and *Age.*

PctDM^=38.702-0.21*B1 because it is single linear regression the B1 is the slope of PctDm and Age. So there is negative linear relationship between PctDM and Age.

b. What percentage of the variability in *PctDM* does *Age* explain for these fish?

The R square is 0.2 and means the model or these age can explain 20% variation in PctDM

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.448[a]	.200	.176	1.42630

a. Predictors: (Constant), Age

c. There evidence that the relationship in (a) is statistically significant? Explain how you know that it is or is not.

Since t = -2.876 and p value is 0.007. It is significant and we can reject null hypothesis.

We can conclude that there is a relationship between average value of PctDm and age.

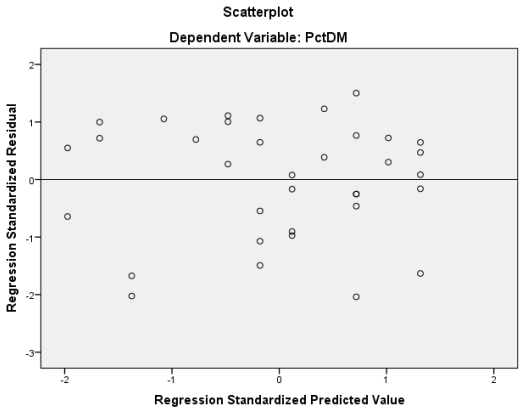
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	38.702	.868		44.596	.000
	Age	-.210	.073	-.448	-2.876	.007

a. Dependent Variable: PctDM

d. Produce a plot of the residuals versus the fits for the simple linear model. Does there appear to be any regular pattern?

Yes, residual plot in right part is more accurate than left part.



e. Modify your plot in (d) to show the points for each *Month* (Sept/Nov) with different symbols or colors. What (if anything) do you observe about how the residuals might be related to the month? Now fit a multiple regression model, using an indicator (*Sept*) for the month and interaction product, to compare the regression lines for September and November.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	39.397	1.074		36.691	.000
	Age	-.218	.089	-.464	-2.440	.021
	Sept	-1.276	1.512	-.404	-.844	.405
	SepAge	-.021	.128	-.082	-.168	.868

a. Dependent Variable: PctDM

PctDM^= 39.397 -0.218*Age-1.276*Sept-0.21*SepAge

September Line: sept = 1

PctDM^=38.121-0.428Age

November Line: sept = 0

PctDM^= 39.397 -0.218*Age

f. Do you need both terms for a difference in intercepts and slopes? If not, delete any terms that aren't needed and run the new model.

Yes. Because this model explains 43% of variation in PctDM. Compare with old model, just 20%. That means without two terms. The simple linear regression with a fixed slope only explain 20% variation. ANOVA is also significant.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.656[a]	.430	.375	1.24212

a. Predictors: (Constant), SeptAge, Age, Sept

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	36.133	3	12.044	7.806	.001[b]
	Residual	47.829	31	1.543		
	Total	83.962	34			

a. Dependent Variable: PctDM

b. Predictors: (Constant), SeptAge, Age, Sept

g. What percentage of the variability in *PctDM* does the model in (f) explain for these fish?

R Square =.43 and means the model can explain 43% variation in PctDM

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.656[a]	.430	.375	1.24212

a. Predictors: (Constant), SeptAge, Age, Sept

h. B0

PctDM^= 39.397 -0.218*Age-1.276*Sept-0.21*SepAge

B0 is the intercept when Sept indicator is 0

3a (Diet)

A team of anthropologists and nutrition experts investigated the influence of protein content in a diet on the relationship between Age in years and Height (HT) in centimeters for New Guinean children. They also recorded whether the children ate a protein rich or protein poor diet. The data are in dataset Diet.

1. Fit the regression model predicting height using a child's age, and test whether there is a significant effect.

Since t = 7.845 and p value is =0 and so it is significant. We reject null hypothesis. The data suggests that there is a relationship between age and average height.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	51.060	2.871		17.787	.000
	AGE	12.045	1.535	.843	7.845	.000

a. Dependent Variable: HT

2. Fit a model that produces parallel regression lines for children with protein rich and poor diets.

HT^=45.661 +12.058*Age+11.166*DietIndicator

Rich line: DietIndicator=1

HT^=45.661 +12.058*Age+11.166*1=56.827+12.058*Age

Poor line: DietIndicator = 0

HT^=45.661 +12.058*Age+11.166*1=45.661 +12.058*Age

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	45.661	1.829		24.966	.000
	AGE	12.058	.890	.844	13.554	.000
	DietIndicator	11.166	1.572	.442	7.103	.000

a. Dependent Variable: HT

3. Test the null hypothesis that a signal regression line adequately describes these data against the alternative that two parallel lines are needed.

   The DietIndicator p value is 0 and t = 7.103 and so it is significant and we can reject null hypothesis. The data indicates there is intercept change from “Rich” to “Poor.” So two parallel lines are needed.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	45.661	1.829		24.966	.000
	AGE	12.058	.890	.844	13.554	.000
	DietIndicator	11.166	1.572	.442	7.103	.000

a. Dependent Variable: HT

4. Write the prediction equation using all the necessary terms.

HT^=51.225 + Age*8.686-0.901*DietIndicator+7.323*AgeDiet

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	51.225	1.271		40.299	.000
	AGE	8.686	.676	.608	12.845	.000
	DietIndicator	-.901	1.862	-.036	-.484	.633
	AgeDiet	7.323	.996	.591	7.349	.000

a. Dependent Variable: HT

3.30 ***2008 U.S. presidential polls.*** The file **Pollster08** contains data from 102 polls that were taken during the 2008 U.S. presidential campaign. These data include all presidential polls reported on the Internet site pollster.com that were taken between August 29, when John McCain announced that Sarah Palin would be his running mate as the Republican nominee for vice president, and the end of September. The variable *MidDate* gives the middle date of the period when the poll was "in the field" (i.e., when the poll was being conducted). The variable *Days* measures the number of days after August 28 (the end of the Democratic convention) that the poll was conducted. The variable *Margin* shows Obama%–McCain% and is a measure of Barack Obama's lead. *Margin* is negative for those polls that showed McCain to be ahead.

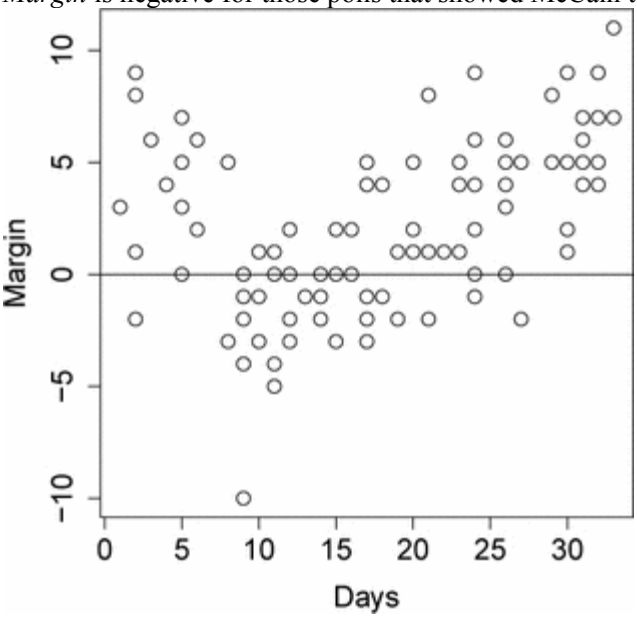


Figure 3.24: ***Obama–McCain margin in 2008 presidential polls***

The scatterplot in **Figure 3.24** of *Margin* versus *Days* shows that Obama's lead dropped during the first part of September but grew during the latter part of September. A quadratic model might explain the data. However, two theories have been advanced as to what caused this pattern, which you will investigate in this exercise.

The **Pollster08** datafile contains a variable *Charlie* that equals 0 if the poll was conducted before the telecast of the first ABC interview of Palin by Charlie Gibson (on September 11) and 1 if the poll was conducted after that telecast. The variable *Meltdown* equals 0 if the poll was conducted before the bankruptcy of Lehman Brothers triggered a meltdown on Wall Street (on September 15) and 1 if the poll was conducted after September 15.

a. Fit a quadratic regression of *Margin* on *Days*. What is the value of R^2 for this fitted model? What is the value of SSE?

Adjusted R square is 0.336 and standard error is 3.014

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.591[a]	.349	.336	3.014

a. Predictors: (Constant), Days, DaysSquare

Margin^=4.478 -0.604*Days+0.021*Days^2

Yes, because DaysSquare(Quadratic term) t value is -4.361 and it is significant. There is relationship between average margin and DaysSquare term.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.478	1.096		4.087	.000
	DaysSquare	.021	.004	1.960	5.595	.000
	Days	-.604	.139	-1.528	-4.361	.000

a. Dependent Variable: Margin

b. Fit a regression model in which *Margin* is explained by *Days* with two lines: one line before the September 11 ABC interview (i.e., *Charlie* = 0) and one line after that date (*Charlie* = 1). What is the value of R^2 for this fitted model? What is the value of SSE?

R square is 0.146 and standard error is 3.454

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.382[a]	.146	.129	3.454

a. Predictors: (Constant), Charlie, Days

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.313	.787		-.398	.692
	Days	.122	.068	.309	1.804	.074
	Charlie	.636	1.304	.084	.488	.627

a. Dependent Variable: Margin

c. Fit a regression model in which *Margin* is explained by *Days* with two lines: one line before the September 15 economic meltdown (i.e., *Meltdown* = 0) and one line after September 15 (*Meltdown* = 1). What is the value of R^2 for this fitted model? What is the value of SSE?

Adjusted R Square = 0.174 and Standard Error is 3.363

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.436[a]	.190	.174	3.363

a. Predictors: (Constant), Meltdown, Days

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.735	.881		.834	.406
	Days	.001	.072	.004	.020	.984
	Meltdown	3.187	1.341	.433	2.377	.019

a. Dependent Variable: Margin

d. Compare your answers to parts (a–c). Which of the three models best explains the data?

Pick up the highest Adjusted R Square and lowest standard error

Quadratic days model. So the Quadratic days model with Adjusted R Square of .336 and standard error of 3.014. So this model is the best model to explain data.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.591[a]	.349	.336	3.014

a. Predictors: (Constant), Days, DaysSquare

Charlie Binary model

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate

1	.382[a]	.146	.129	3.454

b. Predictors: (Constant), Charlie, Days

Meltdown Binary model

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.436[a]	.190	.174	3.363

a. Predictors: (Constant), Meltdown, Days

3.31 ***Metropolitan doctors.*** In Example 1.6, we considered a simple linear model to predict the number of doctors (*NumMDs*) from the number of hospitals (*NumHospitals*) in a metropolitan area. In that example, we found that a square root transformation on the response variable, *sqrt*(*NumMDs*), produced a more linear relationship. In this exercise, use this transformed variable, stored as *SqrtMDs* in **MetroHealth83**, as the response variable.

1. Either the number of hospitals (*NumHospitals*) or number of beds in those hospitals (*NumBeds*) might be good predictors of the number of doctors in a city. Find the correlations between each pair of the three variables, *SqrtMDs, NumHospitals, NumBeds.* Based on these correlations, which of the two predictors would be a more effective predictor of *SqrtMDs* in a simple linear model by itself?

   NumberofBed because the correlation between NumBeds and SqrtMDs is 0.946

Correlations

		NumHospitals	NumBeds	SqrtMDs
NumHospitals	Pearson Correlation	1	.942**	.904**
	Sig. (2-tailed)		.000	.000
	N	83	83	83
NumBeds	Pearson Correlation	.942**	1	.946**
	Sig. (2-tailed)	.000		.000
	N	83	83	83
SqrtMDs	Pearson Correlation	.904**	.946**	1
	Sig. (2-tailed)	.000	.000	

N	83	83	83

**. Correlation is significant at the 0.01 level (2-tailed).

2. How much of the variability in the *SqrtMDs* values is explained by *NumHospitals* alone? How much by *NumBeds* alone?

   For *NumHospitals* alone, the data suggests the model explains 81.7% variation of SqrtMDs

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.904[a]	.817	.815	8.8504024

a. Predictors: (Constant), NumHospitals

   For *NumBeds* alone, the data suggests the model explains 89.5% variation of SqrtMDs

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.946[a]	.895	.894	6.7083094

a. Predictors: (Constant), NumBeds

3. How much of the variability in the *SqrtMDs* values is explained by using a two-predictor multiple regression model with both *NumHospitals* and *NumBeds?*

   The R Square is 0.896. The model explains 89.6% variation of SqrtMDs.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate

1	.947[a]	.896	.894	6.7053905

a. Predictors: (Constant), NumHospitals, NumBeds

4. Based on the two separate simple linear models (or the individual correlations), which of *NumHospitals* and/or *NumBeds* have significant relationship(s) with *SqrtMDs?*

   They are both significant at p value of 0. So the relationships both exist.

   For NumHospitals alone,

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	14.033	1.469		9.555	.000
	NumHospitals	2.915	.153	.904	19.036	.000

a. Dependent Variable: SqrtMDs

   For NumBeds alone,

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	15.259	1.052		14.508	.000
	NumBeds	.013	.000	.946	26.282	.000

a. Dependent Variable: SqrtMDs

5. Which of these two predictors are important in the multiple regression model? Explain what you use to make this judgment.

   Only NumBeds,

   The correlation between NumHospitals and NumBeds is 0.942. So it is Multicollinearity.

   One predictor is redundant.

Correlations

		NumHospitals	NumBeds
NumHospitals	Pearson Correlation	1	.942**
	Sig. (2-tailed)		.000
	N	83	83
NumBeds	Pearson Correlation	.942**	1
	Sig. (2-tailed)	.000	
	N	83	83

**. Correlation is significant at the 0.01 level (2-tailed).

To see which one is better, pick up the one left least affect adjusted R Square when take the other one out.

So when I take out NumHospitals and left NumBeds, the adjusted R Squre still 0.894 no change.

However, when I take out NumBeds and left NumHospitals adjusted R Square goes down from 0.984 to 0.815.

Model with both predictors

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.947[a]	.896	.894	6.7053905

a. Predictors: (Constant), NumHospitals, NumBeds

First try to NumBeds left and observe Adjusted R Square

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.946[a]	.895	.894	6.7083094

a. Predictors: (Constant), NumBeds

Second try to take NumHosptials left and observe Adjusted R Square

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.904[a]	.817	.815	8.8504024

a. Predictors: (Constant), NumHospitals

6. The answers to the last two parts of this exercise might appear to be inconsistent with each other. What might account for this? *Hint:* Look back at part (a).

NumBeds and *NumHospitals* are strongly related with correlation of 0.942. They are fighting with each other when respond to response variable even the overall model is significant. But each of the two are not both significant. Also, when test the slope of individual beta of multiple regression model, we assume to account of the other variable. Because the two variables are strong correlated. So when we know one variable, we already know the variation of response variable. So individual beta is not significant when we use two correlated variables together.

Correlations

		NumHospitals	NumBeds
NumHospitals	Pearson Correlation	1	.942**
	Sig. (2-tailed)		.000
	N	83	83
NumBeds	Pearson Correlation	.942**	1
	Sig. (2-tailed)	.000	
	N	83	83

**. Correlation is significant at the 0.01 level (2-tailed).