MGMT 30500: Assignment 5 – Spring 2025

Due Date: March 23, 11:59 pm

PART A. From the textbook.

Chapter 16: 1, 7 (Show Excel table outputs for both problems)

PART B. For each problem, briefly explain why you choose your answer.

For 80 candidates for 3 sales manager positions, applicants are given a subjective Rating by a screening panel. Then, a number of other variables are scored by an HR software program. Below is backward elimination output (with some deletions) for a study of the regressions of Rating vs. these other variables. Use this output for Problems 1-3.

Regression Analysis: Rating versus Leadership, Experience, Technical skills, Communication skills, Adaptability

Backward Elimination of Terms

Candidate terms: Leadership, Experience, Technical skills, Communication skills, Adaptability

	------Step 1------		------Step 2------		------Step 3------	
	Coef	P	Coef	P	Coef	P
Constant	3.342		2.005		2.842	
Adapt	0.00320	0.233	0.00400	0.400		
Tech	0.23	0.197	0.000018	0.167	0.000013	0.019
Commu	0.45	0.195	0.000089	0.200	0.000079	0.117
Experience	0.632	0.070	0.660	0.037	0.633	0.041
Leadership	0.00279	0.432				
S		0.814788		0.798869		0.711174
R-sq		42.81%		_____		_____
R-sq(adj)		31.65%		_____		_____

α to remove = 0.20

1. To go from Step 3 to Step 4:

   A. We would remove Commu.
   B. We would remove Tech.
   C. We would remove Experience and Tech.
   D. Cannot remove any predictors.
   E. Cannot be determined based on the given output.

2. Which step has the lowest multiple R-squared?
   A. Step 1
   B. Step 2
   C. Step 3

3. Which step has the highest adjusted R-squared?
   A. Step 1
   B. Step 2
   C. Step 3 (step 1 -> 2, delete highest p-value 0.541 > 0.32, step 2 -> 3, highest p-value 0.400 > 0.32, deleting ineffective variable, will make adj R^2 increase.)

Below is Best Subsets Regression output from Minitab for a dataset recording a dependent variable (Y) and seven independent variables (V1:V7) for a random sample of 100 observations. Use this output for Problems 4-5.

Best Subsets Regression: Y versus V1:V7

Response is Y

Vars	R-Sq	R-Sq (adj)	R-Sq (pred)	Mallows Cp	S	V1	V2	V3	V4	V5	V6	V7
1	33.3	32.0	28.7	22.5	19.534	X						
1	15.8	14.0	10.0	40.6	21.954					X		
2	39.9	37.3	33.3	17.8	18.749	X					X	
2	38.2	35.6	30.2	19.5	19.005	X			X			
3	47.6	51.6	37.8	11.8	17.686	X			X		X	
3	46.8	43.3	38.1	12.7	17.831	X					X	X
4	55.4	51.4	44.9	5.9	16.507	X			X		X	X
4	49.1	44.6	37.4	12.3	17.628	X		X	X		X	
5	58.1	53.5	46.6	5.1	16.179	X		X	X		X	X
5	56.0	51.0	43.9	7.3	16.582	X			X	X	X	X
6	58.7	53.4	46.0	6.4	16.241	X		X	X	X	X	X
6	58.3	52.5	44.6	6.8	16.318	X	X	X	X		X	X

4. By the criterion of Best Subsets Regression, the best multiple regression model to predict Y from these predictors has how many predictors:
   A. 4
   B. 5 (The model with the highest adjusted R-square of 53.5%.)
   C. 6
   D. 7
   E. None of the above.

5. In the model Y vs. V1, V4, V6, and V7, is the p-value of V7 less than or greater than 0.32?
   A. Greater than. (Adj R^2(V1, V4, V6) = 51.6% and Adj R^2(V1, V4, V6, V7) = 51.4%. Removing V7 increases Adj R^2, so V7 must be an ineffective predictor (p value > 0.32).)
   B. Less than.

C. Cannot be determined from the given output.

(Ln-Ln Model) For each week of 2020, a liquor wholesaler records the number of cases sold of a popular microbrew that has just gone into mass production and the average price of a case for that week. Using Excel, they fit the simple regression model: Cases Sold vs. Price ($). Because the normal linear model conditions are not satisfied for this model, they consider logarithmic transformation of the variables. The model: LnCasesSold vs. LnPrice satisfies the normal linear model conditions. (Note: Ln stands for the natural logarithm; that is, the logarithm with base e.) The Excel output is shown below. Use this output for Problems 6-7.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.942418
R Square	0.888152
Adjusted R	0.885915
Standard E	0.355874
Observatic	52

	Coefficients	*andard Err*	*t Stat*	*P-value*
Intercept	23.83127	0.945641	25.20118	4.43E-30
LnPrice	-6.70521	0.33651	-19.9257	1.95E-25

6. Based on this model, if the weekly average price for a case is $20, the predicted number of cases sold for that week is approximately:
   A. 59.6
   B. 42.3 (Predcited LnCases = 23.83127 – 6.70251 LnPrice = 3.744 when Price = 20. Hence, the predicted Case = exp(3.744) = 42.3.)
   C. 85.7
   D. 291.0

7. Based on this model, for every $1 increase in the price of a case, weekly sales are predicted:
   A. To decrease by 6.7.
   B. To decrease by 1.9.
   C. To decrease by Ln(1).
   D. None of the above. . (Although the relationship between LnCases and LnPrice is strongly linear, the relationship between Cases and Price is nonlinear so the change in Cases when we increase Price actually depends on the current value of Price. That is, it is not constant, unlike linear regression.