Assignment Two – Chapter Three (six problems)

#1

Accountants at the firm Walker and Walker believed that several traveling executives submit unusually high travel vouchers when they return from business trips. The accountants took a sample of 200 vouchers submitted from the past year; they then developed the following multiple regression equation relating expected travel cost (Y) to number of days on the road (x_1) and distance traveled (x_2) in miles:

$$\hat{Y} = \$145.00 + \$64.38x_1 + \$1.12x_2$$

The coefficient of correlation computed was 0.67.

(a) If Thomas Williams returns from a 1,056-mile trip that took him out of town for five days, what is the expected amount that he should claim as expenses?
(b) Williams submitted a reimbursement request for $1,850; what should the accountant do?
(c) Comment on the validity of this model. Should any other variables be included? Which ones? Why?

#2

Bus and subway ridership in Washington, D.C., during the summer months is believed to be heavily tied to the number of tourists visiting the city. During the past 12 years, the following data have been obtained:

YEAR	NUMBER OF TOURISTS (1,000,000S)	RIDERSHIP (100,000s)
1	13	18
2	7	15
3	9	11
4	8	22
5	19	30
6	20	38
7	14	34
8	16	32
9	19	39
10	27	52
11	21	41
12	14	28

(a) Develop a regression model. Please write down the regression equation.
(b) What is expected ridership if 35 million tourists visit the city?
(c) If there are no tourists at all, explain the predicted ridership.

#3

The following data give the starting salary for students who recently graduated from a local university and accepted jobs soon after graduation. The starting salary, grade-point average (GPA), and major (business or other) are provided.

SALARY	GPA	MAJOR
$42,000	3.4	Business
$41,500	3.6	Business
$37,800	3.5	Other
$40,500	3.2	Other
$44,000	3.9	Business
$31,500	2.1	Other
$36,200	2.6	Business
$33,200	3.1	Other
$41,200	2.8	Business
$39,200	3.5	Other
$38,700	3.1	Other
$42,400	3.6	Business

(a) Using a computer, develop a regression model that could be used to predict starting salary based on GPA and major. Please write down the regression equation.
(b) Use this model to predict the starting salary for a business major with a GPA of 3.5.
(c) What does the model say about the starting salary for a business compared to a nonbusiness major?
(d) Do you believe this model is useful in predicting the starting salary? Justify your answer, using information provided in the computer output.

#4

The following data give the selling price, square footage, number of bedrooms, and age of houses that have sold in a neighborhood in the past 6 months. Develop three regression models to predict the selling price based upon each of the other factors individually. Which of these is better?

SELLINMG PRICES($)	SQUARE FOOTAGE	BEDROOMS	AGE (YEARS)
247,600	1,675	2	23
258,000	1,765	3	26
278,600	1,927	3	31
287,100	2,187	3	18
289,100	2,321	3	22
294,600	2,395	3	17
352,600	2,464	3	7
354,600	2,723	4	8
387,600	2,627	3	2
392,100	2,657	4	2
393,600	2,666	3	1
404,600	2,497	3	1
403,100	3,111	4	3
413,600	2,787	4	2
525,100	3,242	5	8
420,600	2,846	4	4

#5

Use the data in problem #4 and develop a regression model to predict selling price based on the square footage, number of bedroom, and age. Use this to predict the selling price of a 2-year-old, 2,870-square-foot house with 4 bedrooms.

#6

In 2009, the New York Yankees won 100 baseball games during the regular season. The table on the next page lists the number of victories (W), the earned-run-average (ERA), and the batting average (AVG) of each team in the American League. The ERA is one measure of the effectiveness of the pitching staff, and a lower number is better. The batting average is one measure of effectiveness of the hitters, and a higher number is better.

TEAM	W	ERA	AVG
New York Yankees	100	4.22	0.294
Los Angeles Angels	97	4.34	0.281
Boston Red Sox	94	4.36	0.285
Minnesota Twins	89	4.48	0.275
Texas Rangers	86	4.42	0.272
Detroit Tigers	85	4.31	0.263
Seattle Mariners	84	4.01	0.259
Tampa Bay rays	83	4.36	0.262
Chicago White Sox	78	4.12	0.254
Toronto Blue Jays	74	4.46	0.269
Oakland Athletics	72	4.23	0.274
Cleveland Indians	65	5.01	0.268
Kanas City Royals	64	4.75	0.263
Baltimore Orioles	62	5.01	0.262

(a) Develop a regression model that could be used to predict the number of victories based on the ERA.

(b) Develop a regression model that could be used to predict the number of victories based on the batting average.

(c) Which of the two models is better for predicting the number of victories?

(d) Develop a multiple regression model that includes both ERA and batting average. How does this compare to the previous models?