Business Analytics (108-1)

Assignment 3

1.

You are trying to determine the effects of three packaging displays (A, B, and C) on sales of toothpaste. The number of cases toothpaste sold for 9 consecutive weeks is listed below. The type of store (GR = grocery, DI = discount, and DE = department store) and store location (U = urban, S = suburban, and R = rural) are also listed.

(a) Run a multiple regression to determine how the type of store, display, and store location influence sales. Which potential explanatory variables should be included in the equation? Explain your rationale for including and excluding variables.

(b) What type of store, store location, and display appears to maximize sales?

(c) For the type of store in (b), estimate the probability that 80 or more toothpaste will be sold during a week.

(d) Does multicollinearity seem to be a problem?

Week	Store Location	Store Type	Display	Sales
1	U	GR	C	51
2	U	DI	A	59
3	U	DE	B	67
4	S	GR	A	32
5	S	DI	B	66
6	S	DE	C	48
7	R	GR	B	37
8	R	DI	C	52
9	R	DE	A	37

2.

A farmer wants to conduct an experiment to compare eight varieties of oats. The farmer knows that the growing area is heterogeneous so he decided to group the area into five blocks. He randomly plants each variety of oats in each block and records the yields accordingly. Dataset oatvar is the experiment result.

(a) What kind of experiment is this farmer using?

(b) Is there any interaction effect between the variety of oats and the growing area block?

(c) Conduct a hypothesis test to determine whether the yield of oats is affected by different varieties at 5% significance level.

(d) Check the diagnostics. Is there any unusual findings?

(e) Is it necessary to perform multiple comparisons? If yes, carry out the comparison following the structure and procedure presented in the lecture.

3.

The following experiment (fabric data file) was designed to found out to what extent a particular type of fabric gave homogeneous results over its surface for a standard wear test. In a single run the test machine could accommodate four samples of fabric, at positions 1, 2, 3, and 4. On a large sheet of the fabric four areas A, B, C, and D were marked out at random at different places over the surface. From each area 4 samples were taken, and the 16 samples thus obtained were compared in the machine with the following results, given in milligrams of wear.

POSITION	RUN			
	1	2	3	4
1	A = 21	B = 26	D = 20	C = 25
2	D = 23	C = 26	A = 20	B = 27
3	B = 15	D = 13	C = 16	A = 16
4	C = 17	A = 15	B = 20	D = 20

(a) What kind of experiment is there using?

(b) Write down the linear model equation of this experiment design?

(c) Conduct an appropriate analysis for this experiment and what is the conclusion?