Strathmore
UNIVERSITY

STRATHMORE BUSINESS SCHOOL (SBS)
BACHELOR OF FINANCIAL SERVICES (BSF)
END OF SEMESTER EXAMINATION
BFS 4102: ADVANCED OF BUSINESS DATA ANALYTICS

DATE: Monday, 24th July 2023 **TIME**: 2 Hours

INSTRUCTIONS

1. This examination consists of **FIVE** questions.
2. Answer Question **ONE (COMPULSORY)** and any other **TWO** questions.
3. No **MOBILE PHONES** in the exams room.
4. The questions should be answered in **R-MARKDOWN** and a knitted **WORD DOCUMENT** generated. The word document produced should be saved as a **PDF DOCUMENT** and then submitted via **E-LEARNING** for marking.

Question One (30 Marks)

(i) What is customer segmentation? (2 marks)

(ii) Clearly explain the steps involved in hypothesis testing. (5 marks)

(iii) What is the difference between Business intelligence and Bg Data Analytics? (2 marks)

(iv) Explain three types of Business Analytics. (6 marks)

(v) A clinic provides a program to help their clients lose weight and asks a consumer agency to investigate the effectiveness of the program. The agency takes a sample of 15 people, weighing each person in the sample before the program begins and 3 months later to produce the table in below. Formulate the hypothesis that can be tested in this case. (4 marks)

Person	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Before	210	205	193	182	259	239	164	197	222	211	187	175	186	243	246
After	197	195	191	174	236	226	157	196	201	196	181	164	181	229	231

(vi) Based on the data in (iv) can we conclude that the program is effective? Provide a detailed report from your R results (9 marks)

(vii) What is regression analysis? (2 marks)

Question Two (20 Marks)

(i) The following data shows the daily sales of a certain manufacturing company in hundred thousand US dollars.

98	90	85	84	81	79	76	73	69	60
98	90	85	83	80	79	75	72	68	60
93	88	85	82	80	78	75	71	67	59
93	87	84	82	79	77	74	70	64	57
91	86	84	81	79	77	74	70	63	54

(a) Obtain the descriptive statistics in R and interpret each of the output. (10 marks)

(b) Using data in (iii), the sales manager wished to test the following hypothesis.

$$H_0 : \mu = 75 \quad (\text{ no change in the daily sales})$$
$$H_a : \mu > 75 \quad (\text{ there is change in the daily sales})$$

Can the sales person claim that the sales has improved? Perform the analysis in R and give a report of your answer. (10 marks)

Question Three (20 Marks)

(i) Explain four assumptions of linear regression models (4 marks)

(ii) Consider the data 'riders.xlsx' on demand estimation of riders based on various variables. Develop a least-squares regression equation to predict number of weekly riders, based on price per week, population of the city, monthly income of riders and average parking rates per month. Clearly interpret your results and explain if all the four linear regression model assumptions have been met. (16 marks)

Question Four (20 Marks)

(i) Briefly describe what you understand by correlation and multi-collinearity. (6 marks)

(ii) Using the customer data provided in excel, perform correlation analysis by generating a correlation matrix in R and clearly interpret your results. Explain if there is multi-collinearity in the data. (14 marks)

Question Five (20 Marks)

(i) Briefly explain how you will classify the customers (refer to customer data) into various clusters using k-means algorithm. (5 marks)

(ii) Using the Use customer data provided in excel perform the k-means clustering and clearly explain the implication of the results obtained. (15 marks)

END