Strathmore UNIVERSITY

STRATHMORE BUSINESS SCHOOL

BACHELOR OF SCIENCE IN FINANCIAL SERVICES

END OF SEMESTER EXAMINATION

BFS 4102: ADVANCED BUSINESS DATA ANALYTICS

DATE: Thurs, 1st Aug 2024 **TIME:** 10:30 – 12:30

Instructions

1. This examination consists of **FIVE** questions.
2. Answer **Question ONE (COMPULSORY)** and any other **TWO** questions.

Question One [30 Marks]

Below is a sample of part of 1.2 million records of a retail store collected over a period of 5 years. Study it and answer the questions that follow.

Invoice ID	City	Branch Name	Customer type	Gender	Product line	Unit price	Quantity	Tax 5%	Date	Time	Payment
750-67-8428	Nairobi	Kileleshwa	Member	Female	Health and beauty	2145.00	7	26.14	1/5/2019	13:08	Ewallet
226-31-3081	Nairobi	Mathare	Normal	Female	Electronic accessories	50.00	1	3.82	3/8/2019	10:29	Cash
631-41-3108	Nairobi	Kileleshwa	Normal	Male	Home and lifestyle	200000.00	1	16.22	3/3/2019	13:23	Credit card
123-19-1176	Nairobi	Kahawa	Member	Male	Health and beauty	500.00	28	23.29	1/27/2019	20:33	Ewallet
373-73-7910	Nairobi	Kahawa	Normal	Male	Sports and travel	1000.00	10	30.21	0/8/2019	10:37	Ewallet
699-14-3026	Nairobi	Kileleshwa	Normal	Male	Electronic accessories	50000.00	10	29.89	3/25/2019	18:30	Ewallet
355-53-5943	Nairobi	Mathare	Member	Femal	Electronic accessories	2000.00	1	20.65	2/25/2019	14:36	Ewallet
315-22-5665	Nairobi	Kileleshwa	Normal	Female	Home and lifestyle	70000.00	10	36.78	2/24/2019	11:38	Ewallet
665-32-9167	Nairobi	Kahawa	Member	Female	Health and beauty	30000.00	2	13.63	1/10/2019	17:15	Credit card
351-62-0822	Nairobi	Mathare	Member	Female	Fashion accessories	1500.00	4	12.90	2/6/2019	18:07	Ewallet
529-56-3974	Nairobi	Kahawa	Member	Male	Electronic accessories	10000.00	4	15.10	3/9/2019	17:03	Cash
365-64-0515	Nairobi	Kahawa	Normal	Female	Electronic accessories	3000.00	5	11.74	2/12/2019	10:25	Ewallet
252-56-2699	Nairobi	Mathare	Normal	Males	Food and beverages	500.00	10	2.60	2/7/2019	16:48	Ewallet
829-34-3910	Nairobi	Mathare	Normal	Female	Health and beauty	1000.00	10	35.69	3/29/2019	19:21	Cash
299-46-1805	Nairobi	Kahawa	Member	Female	Sports and travel	2000.00	6	28.12	0/15/2019	16:19	Cash
656-95-9349	Nairobi	Kileleshwa	Member	Female	Health and beauty	60000.00	7	24.13	3/11/2019	11:03	Credit card
765-26-6951	Nairobi	Mathare	Normal	Male	Sports and travel	2000.00	6	21.78	1/1/2019	10:39	Credit card
329-62-1586	Nairobi	Kileleshwa	Normal	Male	Food and beverages	20000.00	3	2.20	1/21/2019	18:00	Credit card

Required;

i. Outline **two** possible data cleaning procedures that you could apply to have a data analytics ready dataset. **[4 marks]**

ii. Discuss and provide a summary of four data description approaches that you could apply to this dataset **[12 marks]**

iii. Discuss **four** data diagnoses for the trends you see in the answer in (b) above **[8 marks]**

iv. Discuss **two** possible data science models that can be deployed to predict sales for the "Kileleshwa" branch. **[4 marks]**

v. Explain **two** ways the business can leverage analytical outcomes from question (b, c, and d) to gain market advantage. **[2 marks]**

Question Two [15 Marks]

1. Explain the following terms
   a) Data **[1 mark]**
   b) Information **[1 mark]**
   c) Visualization **[1 mark]**

2. Differentiate the following terms in the context of data analytics
   a) Business intelligence **[2 marks]**
   b) Business strategy **[2 marks]**

3. Contrast descriptive analytics and diagnostics analytics **[8 marks]**

Question Three [15 Marks]

The number of bottles of red wine sold by a local supermarket over a two-week period is shown below.

22, 14, 11, 33, 32, 45, 4, 12, 13, 20, 27, 44, 30, 15

a) Calculate the;
   (i) Mean **[2 marks]**
   (ii) Standard deviation **[2 marks]**
   (iii) Lower quartiles **[2 marks]**
   (iv) Upper quartile **[2 marks]**

b) Draw a suitably labelled box plot for this data **[6 marks]**

Question Four [15 Marks]

In a supply chain business, the daily commuting distances of 125 tracks rounded to the nearest kilometres are summarised in the table below.

Distance	**Frequency**
0 – 9	12
10 –19	22
20 – 29	48
30 – 39	26
40 – 49	8
50 – 59	5

Questions

1. Draw an accurate histogram to represent this data **[6 marks]**
2. Determine with justification the skewness of the data **[3 marks]**
3. Use linear interpolation to estimate the value of the median **[6 marks]**

Question Five [15 Marks]

Read this description and answer the questions that follow:

Your task is to develop a tailored smart advisory data driven system for the "Sunflower Cooking Oil" product that has been in the market for 10 years. Here is a sample of the data and variables.

store_id	record_id	date	country	county	sub-county	product	quantity	expected_retail_price
NRB_001	AYU091	01-10-2015	Kenya	Nairobi	Dagoretti	Sunflower Cooking Oil	2 liters	Kes. 5000
NKR_009	REY056	05-2016	Kenya	Nakuru	West	Sunflower Cooking Oil	1 liters	USD 25
NRB_009	SED205	10/12/16	Kenya	Nairobi	Dagoretti	Sunflower Cooking Oil	0.5 liters	USD 5

Questions

a) Discuss **three** possible exploratory data analysis approaches to summarize the main characteristics of the data. **[6 marks]**

b) Design a comprehensive executable analytical workflow that you would use to develop a tailored and business-smart advisory system for future product releases using an appropriate regression model. **[9 marks]**