SPRING SEMESTER 2023
IST 3015 B: BUSSINESS DATA ANALYTICS
INSTRUCTOR: JAPHETH MURSI
DATE: 20TH February, 2023 Duration: 1h45mins

_____

MID SEMESTER EXAMS

Total marks (30)

Instructions:

1. Attempt all the questions

2. Show all your workings

3. Do not WRITE with a PENCIL

Question 1 (8mks)

a) What is the difference between descriptive, predictive, and prescriptive analytics? Give an example of a business scenario where each type of analytics might be used. (4mks)

b) In a sample of 75 athletes, the average time take to complete a 10km race 70 minutes and standard deviation is 8. Assuming the distribution to be normal, find

i) How many athletes completed the race from 73 minutes and above? (2mks)

ii) How many athletes completed the race between 65 and 83 minutes? (2mks)

Question 2 (11mks)

a) A recruitment firm conducted a survey of Data analyst salaries in Kenya and found that the average salary was Kshs. 120,000 with a standard deviation of 30,000. If an employee earns Kshs 190,000, what is their z-score? (3mks)

b) A sample of college students was asked how much they spend to make a phone call per day (to the nearest shillings) as shown in the table below

Monthly Cell Phone Plan Cost	Number of Students
30 – 39	28
40 – 49	26
50 – 59	27
60 – 69	51
70 – 79	25
80 – 89	17
90- 99	30

i) Calculate the Mean, median, and Mode of above data (4mks)

ii) Calculate IQR and Standard Deviation (4mks)

Question 3 (11mks)

a) The table below shows the number of absences X, in IST 3015 course and the final exam grade Y, for seven students

X	1	0	2	7	4	3	3
Y	95	90	90	55	70	80	85

i) Find the correlation coefficient and interpret your result (3mks)

ii) Predict the test score for a student with 5 absences (3mks)

b) A Restaurant manager wants to know if there is a relationship between gender (male or female) and the preferred type of snack. The following table summarizes the results. Test the hypothesis with a significance level of 10% (5mks)

	Chocolate	Chips	Smoothie	Total
Male	30	28	12	70
Female	41	18	37	96
Total	71	46	49	166

Formulas

$$X - \overline{X} = r \frac{\sigma_x}{\sigma_y} (Y - \overline{Y})$$

1.

$$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}}$$

2.

$$\hat{Y} = a + bx$$

3.

$$Q_1 = L_{Q_1} + \left(\frac{\frac{n}{4} - F}{f_{Q_1}}\right)i \qquad Q_3 = L_{Q_3} + \left(\frac{\frac{3n}{4} - F}{f_{Q_3}}\right)i$$

4.

$$\mathsf{Median} = L_m + \left(\frac{\frac{n}{2} - F}{f_m}\right)i$$

.

5.

$$\boxed{IQR = Q_3 - Q_1}$$

6.

$$\mathsf{Mode} = L_{mo} + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)i$$

7.

Population Variance:
$$\sigma^2 = \frac{\sum fx^2 - \frac{\left(\sum fx\right)^2}{N}}{N}$$

Variance for sample data:
$$s^2 = \frac{\sum fx^2 - \frac{\left(\sum fx\right)^2}{n}}{n-1}$$

Standard Deviation:

Population: $\sigma^2 = \sqrt{\sigma^2}$

Sample: $s^2 = \sqrt{s^2}$

8

Finding the y-intercept $b = \frac{\sum x}{n} - m \frac{\sum y}{n}$:

9. Regression equation of x on Y

$$\left(X - \bar{x}\right) = b_{XY}\left(Y - \bar{y}\right)$$