



SPRING SEMESTER 2023
IST 3015 B: BUSSINESS DATA ANALYTICS
INSTRUCTOR: JAPHETH MURSI
DATE: 8TH February, 2023

QUIZZ 1

Total marks (20)

Instructions:

1. Attempt all the questions
2. Show all your workings

Question 1

- a) Distinguish and give practical examples of Prescriptive and Predictive analytics **(3mks)**
- b) It is estimated that 55% of women who applied for nomination for elective posts in Party A are shortlisted. A researcher wants to find out if the number is the same or higher in party B. He commissions a study and finds out that out of 228 women candidates who applied in Party B only 82 were shortlisted. The null hypothesis framed by him was $p = 0.55$, and the alternate hypothesis was $p > 0.55$. Find the p-value for the research. Give a significance level of 5% make conclusion in regards to hypothesis **(5mks)**.
- c) The following table gives the amount of time (in minutes) taken to complete a marathon by 170 athletes.

Time spent in minutes	130-149	150-169	170-189	190-209	210-229
Number of athletes	44	23	52	21	30

- i) Calculate the Median, Mode, Mean and Interquartile Range (IQR) of above data **(6mks)**

Question 2

8 USIU students were asked the total time they took to complete assignment 1. The table below shows the time taken in minutes (x) and their final exam score (y)

Minutes spent to complete a quiz	Final score
45	65
55	75
60	80
35	65
85	55
70	65
75	55
65	85

- a) Calculate and interpret the correlation coefficient r **(3mks)**
- b) Predict the final assignment score of a random student if you know total number minutes spent on assignment 1 **(1mk)**
- c) What is the assignment exam score for a student who spent 2hrs to complete the assignment? **(2mks)**

FORMULAS

Formulas

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

1.

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

2.

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

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} :$$

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

$$a = \bar{Y} - b\bar{X}$$

$$b = \frac{\sum XY - (\sum X \bullet \sum Y) / k}{\sum X^2 - (\sum X)^2 / k}$$

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.

$$M e d i a n = L_m + \left(\frac{\frac{n}{2} - F}{f_m} \right) i$$

.

5.

$$IQR = Q_3 - Q_1$$

6.

$$M o d e = L_{mo} + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) i$$

7.

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

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

Standard Deviation:

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

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

7.

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

8. Regression equation of x on Y

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

9.

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$