

CONFIDENTIAL



UNIVERSITI TUN HUSSEIN ONN MALAYSIA

**FINAL EXAMINATION
SEMESTER II
SESSION 2022/2023**

- COURSE NAME : DATA ANALYSIS
- COURSE CODE : BWA 21003
- PROGRAMME CODE : BWA
- EXAMINATION DATE : JULY/AUGUST 2023
- DURATION : 2 HOURS 30 MINUTES
- INSTRUCTIONS :
1. ANSWER ALL QUESTIONS
 2. THIS FINAL EXAMINATION IS CONDUCTED VIA
 - Open book
 - Closed book
 3. STUDENTS ARE **PROHIBITED** TO CONSULT THEIR OWN MATERIAL OR ANY EXTERNAL RESOURCES DURING THE EXAMINATION CONDUCTED VIA CLOSED BOOK

THIS QUESTION PAPER CONSISTS OF **FIVE (5)** PAGES

CONFIDENTIAL

TERBUKA

Q1 A farmer wants to compare papaya's mean pesticide content from two farms, each located at Cameron Highlands and Kundasang. To test for pesticides, the farmer simply passes a swab over the fruit and inserts the swab into the pesticides detector, and waits about 30 seconds. A green light on the face of the device means the pesticide residue is under the Environmental Protection Agency's tolerance; a red light means the opposite. He collects 30 samples from each location.

(a) State the statistical test would the farmer use to make the comparison. Give your reason.

(4 marks)

(b) The farmer expands the study of mean pesticide content in papaya by collecting data from 5 different farms location. Name an appropriate statistical test if the farmer wants to compare mean pesticide content from Cameron Highlands, Kundasang, Rawang, Kota Tinggi, and Bukit Fraser. Write an appropriate hypothesis statement for the test.

(4 marks)

Q2 Emma wants to find out if the time that someone spends in the first round of registration differs between males and females. She has the following set of hypotheses:

H_0 : There is no difference in the time that males spend in the first round of registration compared to the time that females spend.

H_1 : There is a difference in the time that males spend in the first round of registration compared to the time that females spend.

(a) She measures the time that someone spends with an open question where respondents could indicate the number of minutes spent in the registration, and asks respondents whether they are male or female. What kind of test would she use to test the statement?

(2 marks)

(b) If she changes the question that was used to measure the time spent with a closed question with the following answer options: < 10 minutes, 10 –19 minutes, 20-39 minutes, \geq 40 minutes, name the most appropriate test to test the hypotheses statement.

(2 marks)

Q3 A food services manager for a baseball park wants to know if there is a relationship between gender (male or female) and the preferred condiment on a hot dog. **Table Q3.1** summarizes the results.

Table Q3.1

Gender	Condiment		
	Ketchup	Mustard	Mayonnaise
Male	15	23	10
Female	25	19	8

Test the hypothesis with a significance level of 10%.

(15 marks)

Q4 An industrial engineer employed by a soft drink beverage bottler is analyzing the product delivery and service operations for vending machines. He suspects that the time required by a route deliveryman to load and service a machine is related to the number of cases of product delivered. The engineer visits 25 randomly chosen retail outlets having vending machines, and the in-outlet delivery time (in minutes) and the number of products delivered (in cases) are observed for each. The 25 observations on delivery time and delivery volume are listed in **Table Q4.1**.

Table Q4.1

Observations	Delivery Time	Number of Cases	Observations	Delivery Time	Number of Cases
1	16.68	7	14	19.75	6
2	11.50	3	15	24.00	9
3	12.03	3	16	29.00	10
4	14.88	4	17	15.35	6
5	13.75	6	18	19.00	7
6	18.11	7	19	9.50	3
7	8.00	2	20	35.10	17
8	17.83	7	21	17.90	10
9	79.24	30	22	52.32	26
10	21.50	5	23	18.75	9
11	40.33	16	24	19.83	8
12	21.00	10	25	10.75	4
13	13.50	4			

(a) Construct a suitable graphical chart and interpret your findings.

(5 marks)

(b) Compute the measurement to measure the strength of the relationship between the delivery time and the number of cases. Interpret your findings.

(5 marks)



- (c) A linear regression model was fitted to the data and the estimates regression function was obtained. The SPSS output is given in **Table Q4.2**.

Table Q4.2

Model	Unstandardized Coefficients		t stats	Sig.
	β	Std. Error		
Constant	3.321	1.371	2.42	0.024
Cases	2.1762	0.214	17.55	0.000

- (i) Write the estimated regression function. (2 marks)
- (ii) Interpret the slope of the model. Is it possible to interpret the intercept? Explain your answer. (5 marks)
- (iii) Conduct a test that there exists a relationship between the delivery time and the number of cases by using a significant level of 0.05. (5 marks)

Q5 An experiment was conducted to determine if any significant differences exist in the strength of parachutes woven from synthetic fibers from the different brands. The strength of parachutes is measured by placing them in a testing device that pulls on both ends of a parachute until it tears apart. The amount of force required to tear the parachute is measured on a tensile-strength scale where the larger the value the stronger the parachute. The results of this experiment (in terms of tensile strength) are displayed in **Table Q5.1**.

Table Q5.1

Brand A	Brand B	Brand C	Brand D
18.5	26.3	20.6	25.4
24.0	25.3	25.2	19.9
17.2	24.0	20.8	22.6
19.9	21.2	24.7	17.5
18.0	24.5	22.9	20.4

- (a) Conduct an appropriate test whether all four brands' mean tensile strengths are equal. Use 0.05 level of significance. (16 marks)
- (b) Perform the Tukey test to analyze the means of the four different brands of parachutes. (10 marks)



(c) Explain which brand(s) should you choose and avoid.

(5 marks)

- END OF QUESTIONS -

TERBUKA

Small, faint text at the bottom left corner, possibly a footer or page number.