UNIVERSITI TEKNOLOGI MARA
ASSESSMENT 2

COURSE	**:**	**STATISTICS FOR BUSINESS AND SOCIAL SCIENCES**
COURSE CODE	**:**	**STA404**
EXAMINATION	**:**	**JAN 2021**
TIME	**:**	**1 HOUR 30 MINUTES**

INSTRUCTIONS TO CANDIDATES

1. This question paper consists of **SIX (6)** questions.
2. Candidates are given 1 hour 30 minutes to accomplish assessment and 15 minutes to submit this assessment via Google Classroom or any suitable platform in form of pdf file.
3. Answer ALL parts of questions in the A4. Start each answer on a new page.
4. Please check to make sure that this examination pack consists of:
   i) the Question Paper
   ii) a two–page Appendix 1
5. Answer ALL questions in English.

NAME:

STUDENT NO:

2	**0**	**1**							

GROUP:

Q1	**/6**
Q2	**/11**
Q3	**/9**
Q4	**/6**
Q5	**/8**
Q6	**/10**
TOTAL	**/50**

PLEASE SUBMIT THIS ASSESSMENT ON THE REQUIRED TIME

This assessment paper consists of 8 printed pages

SELF DECLARATION ASSESSMENT 2:

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. This assessment is my own work.
3. I have not involved, and will not allow anyone, to copy my work with the intention of passing it off as their own work.
4. I acknowledge that copying someone else's work (or part of it) is wrong and declare that my assessment is my own work.

QUESTION 1

In an article in the Journal of Statistics Education describe the mean body temperature for of an adult human is 98.6ºF. Then, a group of researchers collected randomly a data of healthy adults ages 18 to 40. The data collected are as follow:

Adult	Temperatures
1	97.5
2	98.0
3	98.5
4	98.7
5	99.0
6	97.0
7	98.6
8	98.8

At α=0.05, can it be concluded that the mean temperature of adults is 98.6ºF?

(6 marks)

QUESTION 2

MIROS wishes to determine whether age category related to gender among motor vehicle crash deaths. The outputs are shown below:

Gender * Age Crosstabulation

			Age			Total
			16-24	25-34	35-44	
Gender	Male	Count	110	94	73	277
		Expected Count	111.8	92.1	73.1	277.0
	Female	Count	43	32	27	102
		Expected Count	41.2	33.9	26.9	102.0
Total		Count	153	126	100	379
		Expected Count	153.0	126.0	100.0	379.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	**A**	**B**	.879
Likelihood Ratio	.259	2	.879
Linear-by-Linear Association	.062	1	.803
N of Valid Cases	379		

a) Find **A** and **B**.

(5 marks)

b) State the null and alternative hypotheses.

(2 marks)

c) At 5% significance level, what can you conclude about the relationship between gender and their age category among motor vehicle crash deaths?

(4 mark)

QUESTION 3

A music teacher monitored the sight-reading ability of one of her pupils over a 10 week period. At the end of each week, the pupil was given a new piece to sight-read and the teacher noted the number of errors. She also recorded the number of hours that the pupil had practiced each week. The data were recorded and analyzed using IBM SPSS Statistics. The results are as follow.

Pupil	Number of hours	Number of errors
1	12	8
2	15	4
3	7	13
4	11	8
5	1	18
6	8	12
7	4	15
8	6	14
9	9	12
10	3	16

a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	19.356	.484		40.009	.000
	Number_of_hours	-.968	.056	-.987	-17.280	.000

Dependent Variable: Number_of_errors

a) Compute the Pearson's product moment correlation coefficient and explain the value obtained.

(5 marks)

b) Based on given output, write the complete estimated regression equation. Hence interpret the slope in the context of the problem.

(2 marks)

c) Predict the number of errors if the number of hours that the pupil had practiced each week is 17 hours.

(2 marks)

QUESTION 4

The Deputy Rector of Students Affair wants to see whether there is significant difference in ages of resident students and non-resident students. He select a random samples of 40 students from each group. The data were collected and analysed using SPSS. The result illustrated as follow.

Group Statistics

Student		N	Mean	Std. Deviation	Std. Error Mean
age	Resident	40	22.08	3.331	.527
	Non-resident	40	23.03	3.340	.528

a) Calculate the mean difference of age between resident and non-resident students.

(1 mark)

b) State the null and alternative hypotheses for the above study.

(1 mark)

c) Show the value of Z-statistic is -1.274.

(2 marks)

d) Using the information in b) and c), can the researcher reject the claim that there is no difference in the ages of the two groups? Use $\alpha = 0.05$.

(2 marks)

QUESTION 5

An English Language lecturer at University of A_Star wishes to see whether a new smartphone app will reduce the number of grammatical errors her students make when writing the essay. She randomly selects five students, and the data are shown below.

Student	A	B	C	D	E
Error before	9	12	4	0	3
Error after	6	9	2	1	3

a) What is the most appropriate statistical method used for the above data?

(1 mark)

b) Based on the above data, at 5% significance level, test whether a new smartphone app has reduced the number of grammatical errors. Given the standard deviation difference is 1.8166.

(7 marks)

QUESTION 6

A researcher wishes to see whether there is any difference in the weight gains (in kg) of athletes following one of three special diet programs. Athletes are randomly assigned to three groups and placed on diet for 6 weeks. The weight gains are shown below.

Diet A	Diet B	Diet C
4	14	5
3	10	8
7	11	2
6	12	3

ANOVA

Weight gain

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	**W**	2	65.583	**Y**	.001
Within Groups	39.750	**X**	4.417		
Total	170.917	11			

a) Calculate the value of **W** using the SSB formula. Hence, find **X** and **Y** values.

(6 marks)

b) At $\alpha = 0.05$, can the researcher conclude that there is a difference in the diets? Show the relevant steps.

(4 marks)

END OF QUESTION PAPER

APPENDIX 1 (1)

HYPOTHESIS TESTING

Null Hypothesis	Test statistic
$H_0 : \mu = \mu_0$ σ^2 unknown, small samples	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$; df = n – 1
$H_0 : \mu_1 - \mu_2 = 0$ $\sigma_1^2 = \sigma_2^2$ and unknown	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$; df = $n_1 + n_2 - 2$ $s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$
$H_0 : \mu_1 - \mu_2 = 0$ $\sigma_1^2 \neq \sigma_2^2$ and unknown	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $df = \frac{\left[s_1^2/n_1 + s_2^2/n_2\right]^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$
$H_0 : \mu_d = 0$	$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}}$; df = n – 1, where n is no. of pairs
Hypothesis for categorical data	$\chi^2 = \sum \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$

APPENDIX 1 (2)

ANALYSIS OF VARIANCE FOR A COMPLETELY RANDOMIZED DESIGN

Let:

k = the number of different samples (or treatments)

n_i = the size of sample i

T_i = the sum of the values in sample i

n = the number of values in all samples

= $n_1 + n_2 + n_3 + \ldots$

$\sum x$ = the sum of the values in all samples

= $T_1 + T_2 + T_3 + \ldots$

$\sum x^2$ = the sum of the squares of values in all samples

Degrees of freedom for the numerator = k – 1
Degrees of freedom for the denominator = n – k

Total sum of squares: $SST = \sum x^2 - \frac{(\sum x)^2}{n}$

Between-samples sum of squares:

$$SSB = \left(\frac{T_1^2}{n_1} + \frac{T_2^2}{n_2} + \frac{T_3^2}{n_3} + \ldots \right) - \frac{(\sum x)^2}{n}$$

Within- samples sum of squares = SST - SSB

Variance between samples: $MSB = \frac{SSB}{(k-1)}$

Variance within samples: $MSW = \frac{SSW}{(n-k)}$

Test statistic for a one-way ANOVA test: $F = \frac{MSB}{MSW}$