UNIVERSITI TEKNOLOGI MARA
ASSESSMENT 1

COURSE	:	**STATISTICS FOR BUSINESS AND SOCIAL SCIENCES**
COURSE CODE	:	**STA404**
DATE OF EXAMINATION	:	**17TH – 21ST MAY 2021**
DURATION	:	**1 HOUR**

INSTRUCTIONS TO CANDIDATES

1. This question paper consists of **PART A and PART B**.
2. Answer ALL questions in the foolscap paper. Start each answer on a new page.
3. Candidates must accomplish this assessment within 1 hour.
4. Candidates are required to convert their completed answer in one PDF file before submission **(<FULLNAME_GROUP>.pdf**).
5. Candidates are given 30 minutes to email their completed answer to the lecturer.
6. Please check to make sure that this assessment pack consists of :
   i) the Question Paper
   ii) a two–page Appendix 1
7. Answer ALL questions in English.

NAME:

STUDENT NO:

2	**0**								

GROUP:

	N4AM2253A		**N4AM2253C**
	N4AM2253B		**N4AM2253D**

PART A	**10%**	**CLO 1**
Q1	**/10**	
Q2	**/10**	
TOTAL	**/20**	
PART B	**10%**	**CLO 3**
Q1	**/10**	
Q2	**/5**	
Q3	**/5**	
TOTAL	**/20**	

PLEASE SUBMIT THIS ASSESSMENT ON THE REQUIRED TIME

This assessment paper consists of 6 printed page

PART A

QUESTION 1

a) Fill in the blank with an appropriate answer.

i) Inferential statistics is a method of making inferences about a population based on _____.

ii) The two types of quantitative data are _____ and _____.

iii) There are two types of statistics, _____ and _____ statistics.

(5 marks)

b) In the tourism industry, customer satisfaction is a crucial factor affecting number of tourists. The management of Hotel XYZ is interested in determining the level of customer satisfaction with the service provided by all of its branches throughout Malaysia. The hotel has 12 branches all over the country. Six hotels are selected and questionnaires are distributed to the guests who stayed in the hotels for a period of two months. The questionnaire intends to level of customer satisfaction of Hotel XYZ. The level of customer satisfaction survey comprise 25 items that measured in Likert-scale (strongly disagree=1, disagree=2, neutral=3, agree=4, strongly agree=5). Using this situation, answer the following questions.

Answer **TRUE (T)** or **FALSE (F)** based on the above study.

i) The population is all the customers in the six hotels.

ii) The sampling frame for this study is the list of customers' name of Hotel XYZ.

iii) The level of measurement for the interest variable in the study is ordinal.

iv) The most appropriate sampling technique for the above study is stratified sampling.

v) The type of variable for the above study is quantitative.

(5 marks)

QUESTION 2

The following data depicted on the profit earned (in RM) by a sample of petty trader in a day business.

490	452	342	245	187	233	398	401	341	277

a) Compute the mean and standard deviation.

(5 marks)

b) Determine the median of this data.

(2 marks)

c) By using the information in a) and b), calculate Pearson's Coefficient of Skewness.

(2 marks)

d) Based on the value in c), comment on the shape of the distribution.

(1 mark)

PART B

QUESTION 1

The cost (RM'000) of heart disease treatment was investigated for 16 patients in a hospital in Kuala Lumpur for two different heart disease treatment: heart bypass and angioplasty. The data were collected and analyzed using SPSS. The output is given as follow. Assume that the costs (RM'000) of heart disease are normally distributed.

Group Statistics

	Type of Treament	N	Mean	Std. Deviation	Std. Error Mean
cost	Heart Bypass	8	73.2563	11.52009	4.07297
	Angioplasty	8	50.6375	11.15552	3.94407

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means			
						95% Confidence Interval of the Difference	
		F	Sig.	df	Mean Difference	Lower	Upper
cost	Equal variances assumed	.148	.706	14	22.61875	10.45859	34.77891
	Equal variances not assumed			13.986	22.61875	10.45742	34.78008

a) Determine whether the variances of cost (RM'000) of heart disease treatment for the heart diseases are equal. Use α=0.05.

(3 marks)

b) Compute the standard error difference for this study.

(3 marks)

c) Prove the 95% confidence interval for the above study is between 10.45859 and 34.77891.

(3 marks)

d) Based on the values in c), is there any enough evidence to indicate there is significant evidence to indicate that there is difference in the cost (RM'000) of heart bypass and angioplasty? Give a reason to support your answer.

(1 mark)

QUESTION 2

According to the Health Journal, a good heart health can be achievable by burning the calories. Walking is one of the simplest ways to maintain heart health. A group of nutrition researcher wants to determine whether an adult who weight 68 kg and walks a normal pace for 60 minutes can burn more than 250 calories. Hence, the data were collected randomly among ten adults whose ages 25 – 30 years with weight 68 kg. The data collected and recorded as follow. Assume that the calories burned (in 60 minutes) are normally distributed.

	N	Mean	Std. Deviation	Std. Error Mean
Calories Burned (in 60 minutes)	10	217.5000	33.45063	10.57802

Test Value = 250

	t	df	Sig. (2-tailed)	Mean Difference	99% Confidence Interval of the Difference	
					Lower	Upper
Calories Burned (in 60 minutes)	M	9	.013	-32.50000	-66.8768	1.8768

a) Name the statistical technique used in the above study.

(1 mark)

b) Find the M value.

(2 marks)

c) State a 95% confidence interval for the mean calories burned by an adult who weight 68 kg and walks a normal pace for 60 minutes

(2 marks)

QUESTION 3

A teacher randomly selects ten students to participate in week of training designed to improve their tying speed. The number of words typed in a minute is recorded before and after the course, to see if the students can type faster after attending the course. The data analyzed and illustrated in the following output. Assume that the number of typed in a minute are normally distributed.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before	43.800	10	4.4171	1.3968
	After	58.800	10	3.8239	1.2092

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	Before - After	**X**	3.9158	1.2383

a) Find the X value.

(1 mark)

b) Construct a 95% confidence interval for the mean.

(3 marks)

c) Based on the values obtained in b), can it be concluded that there is difference in their typing speed when attending the course? Give a reason to support your answer.

(1 mark)

END OF QUESTION PAPER

SAMPLE MEASUREMENTS

Mean	$\bar{x} = \frac{\sum x}{n}$
Standard deviation	$s = \sqrt{\frac{1}{n-1}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right]}$ or $s = \sqrt{\frac{1}{n-1}\left[\sum (x-\bar{x})^2\right]}$
Coefficient of Variation	$CV = \frac{s}{\bar{x}} \times 100\%$
Pearson's Measure of Skewness	Coefficient of Skewness = $\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$ OR $\frac{\text{mean} - \text{mode}}{\text{standard deviation}}$

CONFIDENCE INTERVAL

Parameter and description	A (1 - α) 100% confidence interval
Mean μ, for large samples, σ^2 unknown	$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$
Mean μ, for small samples, σ^2 unknown	$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$; df = n – 1
Difference in means of two normal distributions, μ_1 - μ_2 $\sigma_1^2 = \sigma_2^2$ and unknown	$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$; df = $n_1 + n_2 - 2$ $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$
Difference in means of two normal distributions, μ_1 - μ_2 , $\sigma_1^2 \neq \sigma_2^2$ and unknown	$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$; $df = \frac{\left[s_1^2/n_1 + s_2^2/n_2 \right]^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1} + \frac{\left(s_2^2/n_2\right)^2}{n_2-1}}$
Mean difference of two normal distributions for paired samples, μ_d	$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}$; df = n – 1 where n is no. of pairs