UNIVERSITI TEKNOLOGI MARA
ASSESSMENT 1
(INDIVIDUAL ASSIGNMENT)

COURSE	:	**INTRODUCTION TO STATISTICS**
COURSE CODE	:	**QMT181/STA104**
DATE	:	**20TH MAY 2022**
TIME	:	**9.00 – 11.00 AM (120 MINUTES)**

Please read these instructions:

1) This assessment paper consists of **FIVE (5)** questions.
2) Answer **ALL** questions.
3) The assessment must be taken completely **alone**. Showing it or discussion with anyone is forbidden.
4) Please write your answer on your own papers using **pen**.
5) Student must ensure that their test papers are **readable**. Ensure that your answers are **written clearly** with your name, group and student ID are provided.
6) Student must prepare their answer in **pdf format** and submit via **Google Classroom** or any other platform used by the lecturer. **(FULLNAME_GROUP.pdf)**
7) Keep close track of your allocated time. Due to internet connectivity, students are given no more than 20 minutes (11.00 - 11.20am) to submit their works electronically.
8) **Late submission (after 11.20 am) will not be accepted.**

ALL THE BEST

DO NOT TURN THIS PAGE UNTIL YOU ARE TOLD TO DO SO

This assessment paper consists of 5 pages

QUESTION 1

a)   Identify the type of variable and scale of measurement for the following variable.

   i)   Socioeconomic status

   ii)   The number of absences per semester

   iii)   Sizes of the t-shirt

(6 marks)

b)   State whether each statement below is TRUE (T) or FALSE (F).

   i)   A census observes or measures every member of a population.

   ii)   Ratio is similar to the ordinal data, with extra property that division may be carried out on ratio data and there exist natural zero for the ratio data.

   iii)   One of disadvantage of cluster sampling as it may not be representative of the whole population but may be too alike.

   iv)   Quota sampling procedure is similar to convenience sampling except for the number allocated for each group of respondents with specific characteristics are based on population statistic.

(4 marks)

QUESTION 2

A survey was taken from 125 shop owners at the Shopping Mall 3A for the purpose of obtaining the advertisement techniques used by them to promote their own products.

a)   State the population for the survey.

(1   mark)

b)   What is the most appropriate sampling technique to be used in this survey? Give ONE(1) advantage for using this technique.

(2   marks)

c)   What is the most appropriate method of data collection that can be used in this survey? State ONE reason.

(2 marks)

CONFIDENTIAL

QUESTION 3

The following data represent the serum HDL cholesterol level of a randomly selected 15 patients in their 20's.

33	47	41	48	51
48	36	39	28	35
45	30	38	52	45

Develop a frequency distribution for the above data with the lower limit of the first class is 28. (Hint: use Sturges' Rule to decide on the appropriate number of classes)

(3 marks)

QUESTION 4

The following data represent the dividend yields (in percent) of a random sample of 12 publicly traded stocks with a value of at least $5 billion.

1.7	0	1.15	0.62	1.06	2.45
2.83	2.16	1.12	1.22	1.68	2.38

a)   Compute the five-number summary.

(5   marks)

b)   Draw a box and whisker plot of the data and give your interpretation.

(2 marks)

QUESTION 5

IQ Scores of a random sample of 200 seventh-grade students in School A are tabulated as follows:

IQ Test Scores	Frequency
60-80	8
80-100	25
100-120	48
120-140	98
140-160	21

a)   Calculate the mean and standard deviation for the above data.

(5 marks)

b)   Draw a histogram and estimate the modal value for the IQ scores of the seventh-grade students from the graph.

(5 marks)

CONFIDENTIAL

c)   Determine the skewness of the data using an appropriate measurement.

(2 marks)

d)   Given the mean and variance of the seventh-grade students of School B is 110.8 and 127.13, compare the variations between both schools and determine which school is more consistent.

(3 marks)

END OF QUESTION PAPER

FORMULA LIST

Sample Measurements

1. Mean, $\bar{x} = \dfrac{\sum x}{n}$ or $\dfrac{\sum fx}{n}$

2. Median, $\tilde{x} = L_m + \left[\dfrac{\dfrac{n}{2} - \sum f_{m-1}}{fm}\right] \times C$

3. Mode, $\hat{x} = L_{mo} + \left\{\dfrac{\Delta_1}{\Delta_1 + \Delta_2}\right\} \times C$

4. Standard Deviation, $s = \sqrt{\dfrac{1}{n-1}\left[\sum(x - \bar{x})^2\right]}$ or $\sqrt{\dfrac{1}{n-1}\left[\sum x^2 - \dfrac{(\sum x)^2}{n}\right]}$

$$\sqrt{\dfrac{1}{n-1}\left[\sum f(x - \bar{x})^2\right]} \text{ or } \sqrt{\dfrac{1}{n-1}\left[\sum fx^2 - \dfrac{(\sum fx)^2}{n}\right]}$$

5. Coefficient of Variation, $CV = \dfrac{s}{\bar{x}} \times 100$

6. Pearson's Measure of Skewness $= \dfrac{Mean - Mode}{Standard Deviation}$ or $\dfrac{3(Mean - Median)}{Standard Deviation}$

where

n	:	total frequency
L_m		lower median class boundary
L_{mo}	:	lower modal class boundary
$\sum f_{m-1}$	:	cumulative frequencies for the classes before the median class
f_m	:	median class frequency
Δ_1	:	(modal class frequency) – (frequency for the class before the modal class)
Δ_2	:	(modal class frequency) – (frequency for the class after the modal class)
C	:	class size

           CONFIDENTIAL