SECI2143: PROBABILITY & STATISTICAL DATA ANALYSIS
2020/2021 – SEMESTER 2

ASSIGNMENT 3

This assignment has 10 questions (130 marks), which contribute **10%** of overall course marks. Submission due date: **3rd July 2021**

QUESTION 1 (10 MARKS)

(a) In a study of women's 100-meter performance of 25 university students, mean time is 19.5 seconds with standard deviation of 9.88 seconds were recorded. Construct a 90% confidence interval to estimate the mean.

(b) In a random sample of 360 food stores, it was found that 160 of these food stores offers special promotions. Construct a 99% confidence interval to estimate the population of food stores that offers promotion.

QUESTION 2 (15 MARKS)

(a) The company ABC introduce a new drug test method, 350 subjects were tested and results from 33 subjects were incorrect (either false positive or false negative). The company claims that the incorrect test results will be less than 10%. Test this claim using α=0.05.

(b) A random sample of 16 steel beams has a mean compressive strength of 58400 psi (pound per square inch) and population standard deviation of 652 psi. Using 99% significance level, test if the true average compressive strength of steel is 58000 psi.

QUESTION 3 (15 MARKS)

(a) The content volume of 10 car lubricants from Brand X were recorded as in Table 1. Using α=0.01, test the hypothesis that the average content volume of the Brand X car lubricants is 10 liters.

Table 1: Content of containers for 10 car lubricants

10.3	9.9	10.2	10.1	9.7	9.9	9.8	10.3	10.0	10.4

(b) **Table 2** shows the birth weights (in Kgs) of babies born from fourteen (14) mothers who taking a special vitamin supplement. In general, the standard deviation of the birth weights of babies is 0.44kg. At 95% significance level, test if the vitamin supplement has any effects on the standard deviation of the birth weight.

Table 2: Birth weights of babies

3.70	4.31	3.73	4.33	3.33	2.58	4.47
3.55	4.66	3.68	3.02	4.09	2.36	3.35

QUESTION 4 **(15 MARKS)**

In general, an experienced quality control inspector of measurement of sheet metal stamping would have a variance of 0.18 $(\text{inch})^2$. A new joint inspector measures 101 stampings with variance of 0.165 $(\text{inch})^2$. Assume data is normally distributed, by using 95% significance level, test whether the new joint inspector is making satisfactory measurements.

QUESTION 5 **(15 MARKS)**

A marketing study was done on consumer from two age groups for their spending on 20 categories of consumer items. Based on 15 respondents in the 18 – 34 age group, the average spending is RM76.40 with a variance of RM25.30. Meanwhile, based on 18 respondents in the 35+ group, the mean and variance were RM71.20 and RM22.20 respectively. By assuming the population variance are not equal, test if there is any difference in the mean spending between these two populations in 95% significance level.

QUESTION 6 **(15 MARKS)**

Table 3 shows the information regarding the thermal performance of two (2) batches of computer processor under stress test.

Table 3: Thermal performance of two batches of processors.

	Batch 1	**Batch 2**
No. of observations	16	21
Mean	$93.6°C$	$98.5°C$
Standard deviation	$3.6°C$	$2.5°C$

Using 95% significance level, test whether the standard variation of processors from batch 1 is greater than the standard deviation of batch 2.

QUESTION 7 **(15 MARKS)**

Table 4 shows the number of words spelled correctly by 10 participants of an international spelling bee competition before and after training. Using 0.05 significance level, is there sufficient evidence to conclude that training can increase the number of words spelled correctly?

Table 4: Number of words spelled correctly before and after training.

Before Training	203	390	389	279	333	213	410	364	470	464
After Training	225	410	402	285	355	240	444	370	501	490

QUESTION 8 **(10 MARKS)**

A local authority is going to study the relationship between the size of household and the daily plastic usage consumed by them. The results of a sampling given below in **Table 5** indicate some values of X (weight of plastic usage) and Y (the size of household).

Table 5: Weight of plastic usage (X) and size of household (Y)

X	0.27	1.41	2.19	2.83	2.19	1.81	0.85	3.05
Y	2	3	3	6	4	2	1	5

i) From the above sampling, find the value of correlation coefficient r.

ii) Using the same sample data set above, conduct the hypothesis testing to know whether the variable X and Y are really correlates using 95% confidence level.

iii) Find out if the decision in (ii) may change if you increase the confidence level to 99%.

QUESTION 9 **(10 MARKS)**

An Automotive Association (AA) analysed the relationship between engine volume and mileage, given the same volume of petrol. **Table 6** is tabulated for nine standard-transmission, four cylinders, petrol-fuelled cars from different car manufacturer. The engine volume is in total cubic kilometre (km).

Table 6: Engine volume and mileage of different car manufacturer

Car manufacturer	**Engine volume, *x* *(cubic km)***	**Mileage, *y* *(km)***
Honda	97	24
Datsun	85	29
Chevrolet	98	26
Mazda	105	24
Proton	120	24
Perodua	151	22
Toyota	140	23
Nissan	134	23
Kia	146	21

Based on the given data,

i) Find the least squares equation (i.e., linear regression) for the data.

ii) Use the least-square equation to estimate the mileage for a car which has 125 cubic km engine volume.

QUESTION 10 **(10 MARKS)**

Table 7 show a test scores (%) for three different groups of athletes that received particular treatment before they begin the performance test.

Table 7: Performance test scores

Memory booster	Placebo	Without Treatment
70	37	3
77	43	10
83	50	17
90	57	23
97	63	30

Conduct an analysis using a one-way ANOVA to test whether these data provide evidence to support the claim that the treatments will have different effects. Use a significance level of 0.01