

CONFIDENTIAL



UNIVERSITI TUN HUSSEIN ONN MALAYSIA

**FINAL EXAMINATION
SEMESTER II
SESSION 2015/2016**

COURSE NAME : DATA ANALYSIS
COURSE CODE : BWA 21003
PROGRAMME CODE : BWA
EXAMINATION DATE : JUNE / JULY 2016
DURATION : 3 HOURS
INSTRUCTION : ANSWER ALL QUESTIONS

THIS QUESTION PAPER CONSISTS OF SIX (6) PAGES

CONFIDENTIAL

- Q1** (a) Explain two differences between primary data and secondary data. (4 marks)
- (b) List five steps of statistical analysis. (5 marks)
- (c) State four scales of measurement. Give one example of each scale. (8 marks)
- (d) Given the SPSS output of 'mpg' as presented in **Table Q1(d)** below.

Table Q1(d): Descriptive summary of 'mpg'

		Statistic	Std. Error	
mpg	Mean	20.9231	.93302	
	95% Confidence Interval for Mean	Lower Bound	19.0015	
		Upper Bound	22.8447	
	5% Trimmed Mean	20.6026		
	Median	21.0000		
	Variance	22.634		
	Std. Deviation	4.75750		
	Minimum	14.00		
	Maximum	35.00		
	Range	21.00		
	Interquartile Range	6.25		
	Skewness	.935	.456	
	Kurtosis	1.793	.887	

- (i) Interpret the skewness of 'mpg'. (2 marks)
- (ii) Interpret the 95% confidence interval of mean. (2 marks)

- Q2** A group of researcher wants to investigate the effect of temperature causes by haze. The temperature data were collected for 12 months consists of minimum and maximum measurement.

Table Q2: Temperature data

Month	Min Temperature	Max Temperature
1	38	61
2	43	65
3	50	72
4	56	76
5	64	83
6	68	90
7	72	91
8	73	90
9	69	87
10	55	59
11	48	54
12	56	67

Without conducting any calculation, based on the above scenario, propose your idea on how to analyze the data. Your proposal should includes

- (i) two research objectives
- (ii) the appropriate descriptive statistics including graphical charts.
- (iii) the analysis on inferential statistics to achieve the research objectives.
[Hint: Name all the statistical tools, graphical charts and/or testing hypothesis (if applicable)].

(10 marks)

- Q3** An auto manufacturing company wanted to investigate how the price of one of its car models depreciates with age. They took a sample of twelve cars of this model and collected the following information on the ages (in years) and prices (in RM) of these cars. Assume the age and the price are normally distributed as given in **Table Q3**.

Table Q3: Information of Age and Price

Age	Price (in thousand)
6	48
3	96
8	15
2	120
6	52
4	89
8	18
5	66
9	20
3	95
4	81
5	67

- (a) Construct a scatter diagram for the above data. What can you conclude from the diagram?
(5 marks)
- (b) Find the estimated regression line by using the method of least squares.
(5 marks)
- (c) If the car is 7 years old, what is the depreciated price?
(2 marks)
- (d) Do the data support the existence of a linear relationship between age and price? Test using $\alpha = 0.05$.
(8 marks)
- (e) Find the Pearson correlation coefficient. What can you infer from the value?
(4 marks)

Q4 (a) The grades in a statistics course for a particular semester were as follows.

Table Q4(a): Grade results

Grade	Frequency
A	14
B	18
C	32
D	20
F	16

Test the hypothesis that the distribution of the grades is following uniform distribution at 0.05 level of significance.

(10 marks)

(b) In an experiment to study the dependence of hypertension on smoking habits, the following data were taken on 180 individuals.

Table Q4(b): Smoking habits

	Nonsmokers	Moderate smokers	Heavy smokers
Hypertension	21	36	30
No hypertension	48	26	19

Test the hypothesis that the presence or absence of hypertension is independent of smoking habits. Use 0.05 level of significance.

(10 marks)

Q5 The data in the following table represent the number of hours of relief provided by 5 different brands of headache tablets administered to 25 subjects experiencing fevers of $38^{\circ}C$ or more.

Table Q5: Brands of tablets

A	B	C	D	E
5.2	9.1	3.2	2.4	7.1
4.7	7.1	5.8	3.4	6.6
8.1	8.2	2.2	4.1	9.3
6.2	6.0	3.1	1.0	4.2
3.0	9.1	7.2	4.0	7.6

- (a) Construct an ANOVA table. (10 marks)

- (b) From result in (a), test the hypothesis that the mean number of hours of relief provided by the tablets is same for all 5 brands. Use 0.05 level of significance. (7 marks)

- (c) Perform the Duncan's multiple range-test to analyze the means of the 5 different brands of headache. Use 0.05 level of significance. (8 marks)

- END OF QUESTION -

UNIVERSITY OF MALAYA
 FACULTY OF SCIENCE
 DEPARTMENT OF STATISTICS
 40100 KUALA LUMPUR
 MALAYSIA