

**FINAL  
ALTERNATIVE ASSESSMENT**

(COVER PAGE)

Session : August 2021

Programme : Diploma in Computer Science (DCS)

Course : **DCS1104: Introduction to Statistics and Data Analytics**

Date of Examination : 8 December 2021 (Wednesday)

Time : 8.00am – 10.30am Reading Time : Nil

Duration : 2 Hours 30 Minutes

**Special Instructions :**

This paper consists of **FOUR (4)** questions. Answer **ALL FOUR (4)** questions. All questions carry equal marks.

Material permitted : Non-Programmable Scientific Calculator

Materials provided : Nil

Examiner(s) : **Siti Syakirah Sazali, Yogeswari Suppiah**

Chief Moderator : Siti Hajar Khairuddin

*This paper consists of 5 printed pages, including the cover page*

DIPLOMA IN COMPUTER SCIENCE PROGRAMME (DCS)  
DCS1104: INTRODUCTION TO STATISTICS AND DATA ANALYTICS  
FINAL ALTERNATIVE ASSESSMENT: AUGUST 2021 SESSION

**Instructions:** This paper consists of **FOUR (4)** questions. Answer **ALL FOUR (4)** questions. All questions carry equal marks.

**Question 1**

(a) Discuss about Data Analytics including its benefits and challenges. (5 marks)

(b) Data Analytics is widely used in variety of field to support decision making. List and explain **FIVE (5)** fields. (10 marks)

(c) Table below shows a dataset of contacts:

Friend	Max Temp (°C)	Weight (kg)	Height (cm)	Gender	Company
Abu	37	67	163	M	Good
Ben	37	51	154	M	Good
Christine	28	76	182	F	Bad
Dennis	29	68	159	M	Good
Eliza	28	99	151	F	Bad
Farhan	38	105	184	M	Good
Gina	34	51	160	F	Bad
Hua	26	96	156	F	Bad
Iris	25	60	184	F	Bad
John	26	63	184	M	Good
Kelvin	27	72	167	M	Bad
Latifa	26	106	155	F	Good
Maria	36	86	185	F	Bad
Ning	29	105	174	M	Good

Based on the table below, plot any **TWO(2)** charts to represent the following information:

- i. Max Temp
- ii. Weight

(10 marks)

**Question 2**

Based on the table of company invoices below, calculate the location and dispersion univariate statistics for sales amount.

Invoice Number	Invoice Date	Customer	Description	Sales Amount	Tax Code	Bank Code	Account Code
INV0051	01-06-12	DF Manufacturing	Stock Sales	800.00	A	B1	IS-100
INV0052	01-06-12	CC Supplies	Stock Sales	500.00	A	B1	IS-100
INV0053	04-06-12	IT Solutions	Stock Sales	320.00	E	B1	IS-100
INV0054	04-06-12	IQ Bonds	Stock Sales	220.00	A	B1	IS-100
INV0055	04-06-12	WC Financial Advisors	Stock Sales	700.00	A	B1	IS-100
INV0056	05-06-12	DF Manufacturing	Stock Sales	742.00	A	B1	IS-100
INV0057	05-06-12	GP Accountants	Stock Sales	300.00	A	B1	IS-100
INV0058	08-06-12	Energy Incorporated	Stock Sales	300.00	A	B1	IS-100
INV0059	08-06-12	Energy Incorporated	Stock Sales	280.00	A	B1	IS-100
INV0060	09-06-12	CC Supplies	Stock Sales	205.00	A	B1	IS-100
INV0061	09-06-12	DF Manufacturing	Stock Sales	400.00	A	B1	IS-100
INV0062	09-06-12	WC Financial Advisors	Stock Sales	254.00	A	B1	IS-100
INV0063	11-06-12	PTY Consultants	Stock Sales	350.00	A	B1	IS-100
INV0064	12-06-12	WW Retail	Stock Sales	520.00	A	B1	IS-100
INV0065	12-06-12	EC Estate Agents	Stock Sales	400.00	A	B1	IS-100
INV0066	12-06-12	The Paint Shop	Stock Sales	297.00	A	B1	IS-100
CN00018	13-06-12	IT Solutions	Stock Sales	300.00	E	B1	IS-100

- (a) Minimum (1 mark)
- (b) Maximum (1 mark)
- (c) Mode (1 mark)
- (d) Mean (4 mark)
- (e) Median (1 mark)
- (f) First Quartile (1 mark)
- (g) Third Quartile (1 mark)
- (h) Interquartile Range (2 mark)
- (i) Amplitude (2 mark)

- (j) Mean Absolute Deviation (4 mark)
- (k) Variance (4 mark)
- (l) Standard Deviation (3 mark)

**Question 3**

(a) Based on the table in Question 2, convert possible **THREE (3)** attributes of the dataset into binary values with suitable justification. (12 marks)

(b) What is normalization? Differentiate Min-Max scaling and standardization methods. (5 marks)

(c) Table below shows the salary and dinner expenses from a list of contacts:

Friend	Salary (\$)	Dinner Expense (\$)
Matthew	17,000	2,200
Bella	53,500	4,500
Cyde	69,000	6,000
Daniel	72,000	7,100
Harry	125,400	10,800

Perform normalization for the following data using Min-Max scaling method.

(8 marks)

**Question 4**

(a) Clustering, Classification, Association Rule Mining and Regression are the most popular Data Mining techniques. Discuss each technique by giving an example. (16 marks)

(b) Discuss **THREE (3)** different clustering methods with an example for each. (9 marks)

**~THE END~**

*DCS1104 (F)/ August2021 Session/ formatted*