FINAL
ALTERNATIVE ASSESSMENT

(COVER PAGE)

Session	:	**August 2020**		
Programme	:	Diploma in Computer Science (DCS)		
Course	:	**DCS1104: Introduction to Statistics and Data Analytics**		
Date of Examination	:	18 December 2020 (Friday)		
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 in the answer booklet provided. All questions carry equal marks.

Material permitted :

Non-Programmable Scientific Calculator

Materials provided :

Nil

Examiner(s)	:	**Neesha Jothi**
Chief Moderator	:	Vasuky Mohanan

*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 2020 SESSION

Instructions: This paper consists of **FOUR (4)** questions. Answer **ALL FOUR (4)** questions in the answer booklet provided. All questions carry equal marks.

Question 1

(a) Explain the difference between data and information using appropriate examples.
(5 marks)

(b) Based on the table below, draw the following:
i. Pie chart for Company
ii. Line chart for Max temp
iii. Histogram for Height
iv. Bar chart for Weight

(20 marks)

Friend	Max temp (°C)	Weight (kg)	Height (cm)	Gender	Company
Andrew	25	77	175	M	Good
Bernhard	31	110	195	M	Good
Carolina	15	70	172	F	Bad
Dennis	20	85	180	M	Good
Eve	10	65	168	F	Bad
Fred	12	75	173	M	Good
Gwyneth	16	75	180	F	Bad
Hayden	26	63	165	F	Bad
Irene	15	55	158	F	Bad
James	21	66	163	M	Good
Kevin	30	95	190	M	Bad
Lea	13	72	172	F	Good
Marcus	8	83	185	F	Bad
Nigel	12	115	192	M	Good

Question 2

(a) List the **THREE (3)** bivariate analysis methods.

(3 marks)

(b) List any **TWO (2)** examples of two quantitative attributes.

(2 marks)

(c) Based on the table below, calculate the location univariate statistics for sales amount.

(20 marks)

Invoice Number	Invoice Date	Customer	Description	Sales Amount	Tax Code	Bank Code	Account Code
INV0051	01-06-12	DF Manufacturing	Stock Sales	8,000.00	A	B1	IS-100
INV0052	01-06-12	CC Supplies	Stock Sales	15,000.00	A	B1	IS-100
INV0053	04-06-12	IT Solutions	Stock Sales	11,200.00	E	B1	IS-100
INV0054	04-06-12	IQ Bonds	Stock Sales	10,200.00	A	B1	IS-100
INV0055	04-06-12	WC Financial Advisors	Stock Sales	17,000.00	A	B1	IS-100
INV0056	05-06-12	DF Manufacturing	Stock Sales	17,420.00	A	B1	IS-100
INV0057	05-06-12	GP Accountants	Stock Sales	3,000.00	A	B1	IS-100
INV0058	08-06-12	Energy Incorporated	Stock Sales	3,000.00	A	B1	IS-100
INV0059	08-06-12	Energy Incorporated	Stock Sales	22,800.00	A	B1	IS-100
INV0060	09-06-12	CC Supplies	Stock Sales	19,050.00	A	B1	IS-100
INV0061	09-06-12	DF Manufacturing	Stock Sales	34,000.00	A	B1	IS-100
INV0062	09-06-12	WC Financial Advisors	Stock Sales	12,540.00	A	B1	IS-100
INV0063	11-06-12	PTY Consultants	Stock Sales	18,000.00	A	B1	IS-100
INV0064	12-06-12	WW Retail	Stock Sales	15,200.00	A	B1	IS-100
INV0065	12-06-12	EC Estate Agents	Stock Sales	14,000.00	A	B1	IS-100
INV0066	12-06-12	The Paint Shop	Stock Sales	12,970.00	A	B1	IS-100
CN00018	13-06-12	IT Solutions	Stock Sales	3,000.00	E	B1	IS-100

Question 3

(a) Explain the dataset with and without noise with appropriate diagram

(5 marks)

(b) Based on the table below, convert the possible features of the dataset into binary values with suitable justification

(10 marks)

Name	Invoice Date	Customer	Description	Sales Amount	Tax Code	Bank Code	Account Code
INV0051	01-06-12	DF Manufacturing	Stock Sales	8,000.00	A	B1	IS-100
INV0052	01-06-12	CC Supplies	Stock Sales	15,000.00	A	B11	IS-100
INV0053	04-06-12	IT Solutions	Stock Sales	11,200.00	E	B1	IS-100
INV0054	04-06-12	IQ Bonds	Stock Sales	10,200.00	A	B1	IS-100
INV0055	04-06-12	WC Financial Advisors	Invoice	17,000.00	A	B11	IS-100
INV0056	05-06-12	DF Manufacturing	Stock Sales	17,420.00	A	B1	IS-100
INV0057	05-06-12	GP Accountants	Stock Sales	3,000.00	A	B11	IS-100
INV0058	08-06-12	Energy Incorporated	Invoice	8,230.00	A	B1	IS-100
INV0059	08-06-12	Energy Incorporated	Stock Sales	22,800.00	A	B1	IS-100
INV0060	09-06-12	CC Supplies	Stock Sales	19,050.00	A	B1	IS-100
INV0061	09-06-12	DF Manufacturing	Invoice	34,000.00	A	B1	IS-100
INV0062	09-06-12	WC Financial Advisors	Stock Sales	12,540.00	E	B1	IS-100
INV0063	11-06-12	PTY Consultants	Stock Sales	18,000.00	A	B12	IS-100
INV0064	12-06-12	WW Retail	Invoice	15,200.00	A	B1	IS-100
INV0065	12-06-12	EC Estate Agents	Stock Sales	14,000.00	E	B1	IS-100
INV0066	12-06-12	The Paint Shop	Stock Sales	12,970.00	A	B12	IS-100
CN00018	13-06-12	IT Solutions	Stock Sales	-3,100.00	E	B1	IS-100

(c) Explain the concept of dimensionality reduction and state any **FOUR (4)** advantages of dimensionality reduction.

(10 marks)

Question 4

(a) Explain are the differences between irrelevant, inconsistent, and redundant data with appropriate examples.

(10 marks)

(b) Given a scenario whereby the dataset is faced with the curse of dimensionality. Explain if it would be better to increase the number of objects or decrease the number of attributes and provide the **TWO (2)** advantages and disadvantages for each approach.

(10 marks)

(c) Explain in your own words the difference between choosing a model and choosing an algorithm with their respective hyper-parameters.

(5 marks)

~THE END~