ourdiu.com | Shared by Abrar Hasan Chowdhury

Daffodil International University

Faculty of Science & Information Technology
Department of Computer Science and Engineering
Mid Semester Examination, Spring-2024

Course Code: CSE315 Course Title: Introduction to Data Science
Level: 3 Term: 1

Exam Duration: 1.5 Hours **Marks: 25**

7.96

Answer ALL Questions [~~Optional~~]

[The figures in the right margin indicate the full marks and corresponding course outcomes. All portions of each question must be answered sequentially.]

1. a) [5]

A technology company claims that the average time it takes its software engineers to resolve a specific type of coding issue is 120 minutes. The company's manager decided to test this claim by collecting samples from 25 software engineers who recently worked on the same type of coding issue. The average time to resolve the issue in the sample was found to be 115 minutes, with a standard deviation of 10 minutes. It is assumed that the time to resolve coding issues follows a normal distribution.

Test the company's claim at a 98% confidence level ($\alpha = 0.02$). Formulate the null and alternative hypotheses, perform the hypothesis test, and interpret the results. **Determine** whether there is sufficient evidence to reject the company's claim about the average resolution time for this coding issue.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480

b) As a market researcher employed by a retail company aiming to evaluate customer satisfaction in its diverse store locations, which span urban to suburban areas, your task is to develop an effective sampling strategy. The management places specific emphasis on comprehending customer satisfaction levels across these varied store locations. Outline your proposed sampling strategy, taking into account factors such as geographical location, store size, and customer demographics. **Justify** your selection of a sampling method by providing reasons for its appropriateness in this context. [3

ourdiu.com | Shared by Abrar Hasan Chowdhury

CamScanner

2.	a)	Suppose you are a teacher and have collected exam scores from your class of 40 students. The scores are as follows: 78, 85, 92, 96, 82, 88, 75, 98, 82, 85, 104, 89, 95, 90, 86, 83, 80, 94, 87, 79, 82, 93, 87, 98, 81, 85, 87, 91, 84, 79, 88, 97, 86, 82, 99, 83, 86, 110, 84, 89. I. **Employ** suitable statistical techniques to pinpoint any potential outliers within the dataset of exam scores. Clearly articulate the criteria or method employed for the identification of outliers. II. **Analyze** how the existence of outliers may impact measures such as the mean and standard deviation of the exam scores.	[2.5 +2.5]	CO1
	b)	Suppose you have collected data on the monthly sales (in thousands of dollars) and advertising expenses (in thousands of dollars) for a sample of 20 retail stores over the past year. The dataset is as follows:	[4]	

Store	Monthly Sales	Advertising Expenses
1	150	10
2	120	8
3	100	12
4	220	13
5	240	12
6	200	12
7	150	15
8	110	16

Determine the correlation coefficient between monthly sales and advertising expenses. Explain the significance of the result in relation to the strength and direction of the association between these two variables.

CamScanner

	Let's consider the following dataset to predict whether we can pet an animal or not. All the features have equal importance.	[8]	CO2

ourdiu.com | Shared by Abrar Hasan Chowdhury

Explain the significance of the result in relation to the strength and direction of the association between these two variables.

CamSc

Let's consider the following dataset to predict whether we can pet an animal or not. All the features have equal importance. [8

	Animals	Size of Animal	Body Color	Can we Pet them
0	Dog	Medium	Black	Yes
1	Dog	Big	White	No
2	Rat	Small	White	Yes
3	Cow	Big	White	Yes
4	Cow	Small	Brown	No
5	Cow	Big	Black	Yes
6	Rat	Big	Brown	No
7	Dog	Small	Brown	Yes
8	Dog	Medium	Brown	Yes
9	Cow	Medium	White	No
10	Dog	Small	Black	Yes
11	Rat	Medium	Black	No
12	Rat	Small	Brown	No
13	Cow	Big	White	Yes

3/5

Calculate the prediction for the following test dataset, **test = (Cow, Medium, Black)** to predict the animal will be pet or not.

CamSc