ourdiu.com | Shared by Imtiaz Deepto 1680

Daffodil International University
Department of Software Engineering
Faculty of Science & Information Technology
Final Examination, Spring 2024
Course Code: DS411; Course Title: Statistical Data Analysis
Sections & Teachers: A (MHS)

Time: 2:00 Hours **Marks: 40**

Answer All Questions

[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)* The height (sidewalk to roof) of notable tall buildings in America is compared to the number of stories of the building (beginning at street level). **10** **CLO-1** ***Level-3***

Height (Feet)	Stories	Height (Feet)	Stories
1050	57	790	60
428	28	401	22
362	26	380	38
529	40	1454	110
1127	100	700	46

   1. Using "stories" as the independent variable and "height" as the dependent variable, make a scatter plot of the data.
   2. Does it appear from inspection that there is a relationship between the variables?
   3. Calculate the least squares line. Put the equation in the form of: $\hat{y} = a + bx$
   4. Calculate the correlation coefficient. Is it significant?
   5. Calculate the estimated heights for 32 stories and for 94 stories.
   6. Based on the data, is there a linear relationship between the number of stories in tall buildings and the height of the buildings?
   7. Are there any outliers in the data? If so, which point(s)?
   8. Calculate the estimated height of a building with six stories? Does the least squares line give an accurate estimate of height? Explain why or why not.
   9. Based on the least squares line, adding an extra story is predicted to add about how many feet to a building?
   10. Calculate the slope of the least squares (best-fit) line? Interpret the slope.

2. *a)* Suppose that you take a random sample of 10,000 Americans and find that 1,111 are left-handed. You perform a test of significance to assess whether the sample data provide evidence that more than 10% of all Americans are left-handed, and you calculate a test statistic of 3.70 and a p-value of .0001. Furthermore, you calculate a 99% confidence interval for the proportion of left-handers in America to be (.103,.119). Consider the following statements: **5** **CLO-2** ***Level-4***

ourdiu.com | Shared by Imtiaz Deepto 1680

1. The sample provides strong evidence that more than 10% of all Americans are left-handed.
2. The sample provides evidence that the proportion of left-handers in America is much larger than 10%.

Investigate the statistical results and figure out which of these two statements is the more appropriate conclusion to draw? Explain your answer based on the results of the significance test and confidence interval.

b) A university is investigating the proportion of students who have participated in extracurricular activities. A random sample of 2,000 students is taken, and 640 of them are found to have participated in at least one extracurricular activity. You perform a test of significance to assess whether the sample data provide evidence that more than 30% of all students participate in extracurricular activities, and you calculate a test statistic of 4.50 with a p-value of 0.00001. Additionally, you calculate a 95% confidence interval for the proportion of students participating in extracurricular activities, yielding the range (0.31, 0.35). Based on these results, which of the following conclusions is more appropriate? 5

1. The sample provides strong evidence that more than 30% of all students participate in extracurricular activities.
2. The sample provides evidence that the proportion of students participating in extracurricular activities is much higher than 30%.

Investigate the statistical results and figure out which of these two statements is the more appropriate conclusion to draw? Explain your answer based on the results of the significance test and confidence interval.

3. *a)* As the person in charge of transportation, you research new transportation techniques for Daffodil International University and find that there is a low cost alternative to the Rajanigandha buses. Unfortunately, there are some troubling safety issues related to this alternative but you find several other universities that have implemented this new technique. You ask each university for data on the safety numbers of the new technique. You analyze the data and realize that if you eliminate one of the universities there is no difference in the safety of the old vs. new techniques. And, your budget will be balanced if you implement the new technique. 5 **CLO-3** ***Level-5***

Assess the situation and recommend a course of action according to the American Statistical Association Ethical Guidelines.

b) How does sleep deprivation affect your ability to drive? A recent study measured the effects on 19 professional drivers. Each driver participated in two experimental sessions: one after normal sleep and one after 27 hours of total sleep deprivation. The treatments were assigned in random order. In each session, performance was measured on a variety of tasks including a driving simulation. Assess the design of this experiment. 5

ourdiu.com

ourdiu.com | Shared by Imtiaz Deepto 1680

A recent study examined the impact of different types of physical exercise on cognitive function in older adults. The study involved 200 participants aged 65-80, split into four groups: aerobic exercise, strength training, combined aerobic and strength training, and a control group that did not engage in any structured exercise. Over a period of 12 months, participants completed a variety of cognitive tests to assess memory, attention, and processing speed. The results showed that participants in the combined aerobic and strength training group demonstrated significant improvements in memory and attention compared to the control group. Additionally, the aerobic exercise group showed notable gains in processing speed. The strength training group also showed some improvement, but it was not statistically significant. The findings suggest that combining aerobic and strength training may be the most effective way to enhance cognitive function in older adults.

10 CLO-4 Level-4

Analyze the given abstract to answer the following questions:

ImtiazKamal

1. Research Design and Methodology: What type of research design does this study likely use, and why do you think this design was chosen for this research? Can you identify any flaw in the methodology?
2. Variable Identification: Identify the independent variables and the dependent variables in this study.
3. Results Interpretation: What do the results indicate about the effectiveness of different types of exercise on cognitive function? Which group(s) showed statistically significant improvements, and in which areas?
4. Implications and Further Research: Based on the abstract, suggest one implication for further research or practical application.
5. How can you extend this study? What might be the next step in exploring this topic further, or how could these findings be used in a real-world context?