**LAIKIPIA UNIVERSITY COLLEGE**

**NYAHURURU CAMPUS**

**DIBM 0213: BUSINESS STATISTICS**

**TIME: 2 HRS**

**ATTEMPT QUESTION 1 AND CHOOSE TWO OTHER QUESTIONS**

1. As the new manager of a small shop, you want to understand the shopping patterns of your customers. You randomly sample 20 purchases from yesterday’s records (all purchases in Kenya shillings):

|  |  |  |  |
| --- | --- | --- | --- |
| 39 | 34 | 47 | 32 |
| 27 | 64 | 11 | 75 |
| 32 | 51 | 21 | 74 |
| 47 | 56 | 40 | 47 |
| 37 | 81 | 38 | 65 |

1. Make a stem-and-leaf plot of the data. (4 mks)
2. Construct a histogram, using a bar width of Ksh. 10. (4 mks)
3. Calculate, from the data; (6 mks)
   1. The arithmetic mean.
   2. The geometric mean.
   3. The Harmonic mean.
4. Why would it be inappropriate to present the data above in a pie chart? (4 mks)
5. Define the following terms; (4 mks)
   1. Categorical variables.
   2. Quantitative variables.
   3. Outliers.
   4. Inferential statistics.
6. “Bad sample designs yield worthless data.” Discuss four sampling forms that produce BAD or WORTHLESS data. (8 mks).
7. Consider the following data for insurance.

|  |  |
| --- | --- |
| Number of sales people working (X) | Sales (Ksh. Millions) |
| 2 | 10 |
| 3 | 11 |
| 7 | 13 |
| 9 | 14 |
| 10 | 18 |
| 10 | 20 |
| 12 | 20 |
| 15 | 22 |
| 16 | 22 |
| 20 | 26 |

1. Prepare a scatter plot of *Sales (Y-axis)* against *Number of sales people* working (x-axis).(4 mks)
2. What can you say about the direction of the association? (2 mks)
3. Compute the correlation coefficient from the data and interpret it. (4 mks)
4. Assume there is a linear relationship between Y and X, derive the regression equation and interpret the coefficients (6 mks)
5. From the regression equation above, what would be the sales if the company engaged 50 workers? Is this answer necessarily accurate? (4 mks)

*3.*

1. For each of the following, write out the alternative hypothesis, and indicate whether it is one-sided or two-sided.
2. Consumer Reports discovered that 20% of a certain computer model had warranty problems over the first three months. From a random sample, the manufacturer wants to know if a new model has improved that rate. (3 mks)
3. The last time a philanthropic agency requested donations, 4.75% of people responded. From a recent pilot mailing, they wonder if that rate has increased. (3 mks)
4. The average age of a customer of a clothing store is 35.2 years. The company wants to know if customers who use their website are younger on average. (3 mks)
5. A student wants to know if other students on her campus prefer Coke or Pepsi. (3 mks)
6. A company is interested in estimating the costs of lunch in their cafeteria. After surveying employees, the staff calculated that a 95% confidence interval for the mean amount of money spent for lunch over a period of six months is (Ksh.780, Ksh.920). Now the organization is trying to write its report and considering the following interpretations. Comment on each.
7. 95% of all employees pay between Ksh.780 and Ksh.920 for lunch. (2 mks)
8. 95% of the sampled employees paid between Ksh.780 and Ksh.920 for lunch. (2 mks)
9. We’re 95% sure that employees in this sample averaged between Ksh.780 and Ksh.920 for lunch. (2 mks)
10. 95% of all samples of employees will have average lunch costs between Ksh.780 and Ksh.920. (2 mks)

4.

A supermarket chain wants to know if their “buy one, get one free” campaign increases customer traffic enough to justify the cost of the program. For each of 10 stores they select two days at random to run the test. For one of those days (selected by a coin flip), the program will be in effect. They want to test the hypothesis that there is no mean difference in traffic against the alternative that the program increases the mean traffic. Here are the results in number of customer visits to the 10 stores:

|  |  |  |
| --- | --- | --- |
| Store number | With program | Without program |
| 1 | 140 | 136 |
| 2 | 233 | 235 |
| 3 | 110 | 108 |
| 4 | 42 | 35 |
| 5 | 332 | 328 |
| 6 | 135 | 135 |
| 7 | 151 | 144 |
| 8 | 33 | 39 |
| 9 | 178 | 170 |
| 10 | 147 | 141 |

1. Compute the mean difference.(2 mks)
2. Compute the standard deviation of the differences.(2 mks)
3. Compute the standard error of the mean difference.(3 mks)
4. Find the value of the *t-*statistic.(3 mks)
5. How many degrees of freedom does the *t*-statistic have? (1 mk)
6. Is the alternative one- or two-sided? Explain.(2 mks)
7. What is the P-value associated with this *t-*statistic? (3 mks)

(Assume that the other assumptions and conditions for inference are met.)

1. At α=0.05, what do you conclude? (4 mks)

5. A national survey indicated that 30% of adults conduct their banking online. It also found that 40% are under the age of 50, and that 25% are under the age of 50 and conduct their banking online.

a. W hat percentage of adults do not conduct their banking online? (3 mks)

b. What type of probability is the 25% mentioned above? (2 mks)

c. Construct a contingency table showing all joint and marginal probabilities. (6 mks)

d. What is the probability that an individual conducts banking online given that the individual is under the age of 50? (5 mks)

e. Are *Banking online* and *Age* independent? Explain. (4 mks)