



# THE CATHOLIC UNIVERSITY OF EASTERN AFRICA

**A. M. E. C. E. A**

**MAIN EXAMINATION**

**AUGUST - DECEMBER 2018 TRIMESTER**

**FACULTY OF COMMERCE**

**DEPARTMENT OF ACCOUNTING AND FINANCE**

**REGULAR PROGRAMME**

**CMS 211: INTRODUCTION TO BUSINESS STATISTICS**

P.O. Box 62157  
00200 Nairobi - KENYA  
Telephone: 891601-6  
Fax: 254-20-891084  
E-mail: academics@cuea.edu

**Date: DECEMBER 2018**

**Duration: 2 Hours**

**INSTRUCTIONS: Answer Question ONE and ANY OTHER TWO Questions**

- Q1. a) Make a distinction between the following statistical concepts:
- i) Descriptive versus inferential statistics. **(2 marks)**
  - ii) Nominal versus interval data **(2 marks)**
  - iii) Frequency polygon versus ogive **(2 marks)**
  - iv) A positively-skewed versus a negatively-skewed distribution **(2 marks)**
  - v) Coefficient of correlation versus coefficient of determination **(2 marks)**

- b) The following is the number of minutes taken to commute from home to work for a group of automobile executives:

|  |
|--|
| 28 25 48 37 41 19 32 26 16 23 23 29 36 31 26 |
| 21 32 25 31 43 35 42 38 33 28                |

**Required:**

- i) Using Sturge's formula, how many classes would you recommend for a frequency distribution that can be used to summarize the above data? **(2 marks)**
- ii) Compute the class interval that you would suggest for each of the

- classes in the distribution. **(2 marks)**
- iii) What would you recommend as the lower limit (lower boundary) of the first class in the distribution? **(1 mark)**
- iv) Organize the data into a frequency distribution using the exclusive method of data classification. **(3 marks)**
- c) The following are times (in minutes) spent on the Internet by 10 customers of a particular cybercafé:

|                         |
|-------------------------|
| 0 9 22 8 33 7 0 5 14 12 |
|-------------------------|

**Required:**

- (i) Determine the standard deviation of these times. **(3 marks)**
- (ii) Determine the inter-quartile range of these times. **(3 marks)**
- (iii) State one reason you might prefer to use the standard deviation to measure dispersion as opposed to using the range. **(1 mark)**
- d) Interviewers A and B awarded marks and ranks respectively to five candidates who appeared for an interview as follows:

| Candidate | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ |
|-----------|-------|-------|-------|-------|-------|
| Mark by A | 73    | 87    | 76    | 70    | 87    |
| Rank by B | 1     | 4     | 3     | 2     | 5     |

**Required:**

Using Spearman's rank correlation coefficient determine the degree of association between the views of the two interviewers on the candidates and comment on the views based on your computed value. **(5 marks)**

- Q2. a) The daily earnings of a popular music entertainer were tabled as follows:

| Earnings (shs 000s) | 10 - 15 | 15 - 20 | 20 - 25 | 25 - 30 | 30 - 35 | 35 - 40 |
|---------------------|---------|---------|---------|---------|---------|---------|
| No. of weeks        | 8       | 13      | K       | 9       | 6       | 4       |

**Required:**

Given that the median of the data is 22, determine the value of K. Hence compute the mean daily earnings **(8 marks)**

- b) The monthly salaries of a group of twenty five employees was classified as follows:

|                      |       |       |       |       |
|----------------------|-------|-------|-------|-------|
| Salary group (£000s) | 1 - 2 | 2 - 3 | 3 - 4 | 4 - 5 |
| No. of employees     | $f_1$ | 6     | 9     | $f_4$ |

**Required:**

Given that the modal pay is 3500, determine the values of  $f_1$  and  $f_4$ .  
(6 marks)

- c) An experimenter planted particular bean seeds in various plots with a view of finding out the type of bean seeds with the highest yield under particular weather and soil conditions. At the end of the experiment, the data collected was tabled as follows:

| Seed type | No. of plots | Mean yield (kg) | Standard deviation (kg) | Coefficient of variation |
|-----------|--------------|-----------------|-------------------------|--------------------------|
| A         | 10           | ?               | 6.4                     | 16%                      |
| B         | 20           | 65              | ?                       | 14%                      |
| C         | ?            | 25              | 5.0                     | ?                        |

**Required:**

If a total of 35 plots were involved, fill in the missing figures in the places that have a question mark (?). Which is the more consistent seed type far as yield is concerned, and what is your justification for the seed type selected?  
(6 marks)

- Q3. A junior accountant was requested to prepare a frequency distribution of the daily sales of commodity W manufactured by the company. The following distribution is what the accountant prepared:

| Earnings (shs 000s) | 0-5 | 10-15 | 15-20 | 20-25 | 25-30 | 30-35 | 35-40 |
|---------------------|-----|-------|-------|-------|-------|-------|-------|
| No. of days         | 2   | 6     | 8     | 12    | 14    | 5     | 3     |

**Required:**

a) Compute the mean/average deviation of the displayed data. (6 marks)

b) Compute the standard deviation of the displayed data. (6 marks)

- c) It was later discovered that there were some days whose revenues gathered were between shs 40,000 and shs 45,000 or the (40-45) class interval. With this information taken into account the lower quartile of the data was found to be Kshs 14, 375. Determine the number of days whose revenue fell between shs 40,000 and shs 45,000, and hence work out the upper quartile value. (8 marks)

Q4. A university is typically required to prepare budgets well in advance of actually receiving its revenues and incurring the expenditures. An important source of revenue is student tuition fee, which is obviously a function of the number of students enrolled. A university was having problems in preparing accurate budgets because past forecasts of the enrollment, made each February before the start of the academic year in September, were subject to considerable error. One aspect of the problem was determining the relationship between the number of applicants received by February 1, and the actual number of new students joining the university the following September. The data tabulated below was collected on September actual registrations and February 1, applications.

| Year | Number of applicants received by February 1 (hundreds) | Number of new students enrolled in September (hundreds) |
|------|--|---|
| 2008 | 28   | 24  |
| 2009 | 26   | 20  |
| 2010 | 28   | 18  |
| 2011 | 28   | 22  |
| 2012 | 36   | 32  |
| 2013 | 36   | 33  |
| 2014 | 42   | 34  |
| 2015 | 46   | 34  |
| 2016 | 46   | 35  |
| 2017 | 50   | 38  |

Required:

- a) Determine the estimating linear regression equation that can be used to forecast the number of students that will be enrolled in September. **(12 marks)**
- b) Suppose that 4,200 applicants or 42(00s) are received by February 1, what is the best estimate of new students who will be enrolled in the following September, based on the regression model? **(2 marks)**
- c) Calculate the coefficient of correlation and the coefficient of determination. By interpreting these two coefficients, should the university rely on the developed estimating regression model to predict the number of students who will enroll in September? Why or why not? **(6 marks)**

**\*END\***