**SECTION A**

1. C 11. C

2. A 12. A   
  
3. A 13. C

4. B 14. B

5. A 15. D

6. C 16. C

7. B 17. C

8. D 18. D

9. B 19. C

10. C 20. B

**SECTION B**

**Question 1.**

(a) (i) Continuous data is measureable data that can on an infinite number of values (in a continuum) while discrete data take on a fixed number of values. **A4**

(ii) Two examples of discrete data: number of overdrawn accounts and the age of bank employees. **A2**

Two examples of continuous data: the time taken to serve a customer and the weight of oil. **A2**.

(b) (i) Average =  **M2**, **A1**

(ii)

|  |  |  |
| --- | --- | --- |
| Classes | Tally marks | Frequency |
| 15 but less than 25  25 but less than 35  35 but less than 45  45 but less than or equal to 55 | |||| |  |||| |||  |||| ||||  |||| || | 6  8  9  7 |

**M1 M2 A1**

1. Using the grouped frequency distribution
2. Calculating the mean,

|  |  |  |  |
| --- | --- | --- | --- |
| Classes | Middle value x, | f | fx |
| 15 but less than 25  25 but less than 35  35 but less than 45  45 but less than or equal to 55 | 20  30  40  50 | 6  8  9  7 | 120  240  360  350 |
| Total |  | 30 | 1070 |

**M1** **M2**

Hence average = , to the nearest whole number. **M1**, **A1**

**(TOTAL : 20 MARKS)**

**Question 2**

(a) (i) Internet and public libraries. **A2**

(ii) The internet is fast resulting in accessibility to a large quantity of data **A2**

Public libraries offer quality data that has been vetted. **A2**

(b) (i) I. A sample is fast and cheaper to collect data from unlike a whole population. **A2**

II. Much more detailed information can be obtained from a sample than a population. **A2**

(ii) Systematic sampling is a type of probability sampling in which sample members from a larger population are selected according to a random starting point and a fixed periodic interval while in stratified sampling members to be studies are selected from groups, called strata, that they belong to and then random selection is done in each group.. **A4**

1. Advantages:
2. It reduces selection bias as all groups or strata are involved.
3. It ensures a high degree of representativeness of groups.

Disadvantage:

I It requires that each item in each of the strata of the population must be known. This may not be possible or difficult in some populations. **A2**

**6 marks**

**(TOTAL : 20 MARKS)**

**Question 3**

1. Regression is a procedure for studying relationships between variables while correlation measures the degree of the relationship between variables. **A4**

(b) (i) Scatter diagram:



**M1** (Scale), **M1** (Labelled axes), **A2** (plotted points)

(ii) Regression line: Let  be the number of visits and  be the numbers of customers attracted. **M1**

Now ,

where  and 

Now 

, 

, **M2**

Then  = , **M2**

and  = , **M1**

Hence  = , **A1**

Plotting the regression line:

When  and when , **M1**

Plotted line **M1, A1**

1. If 8 visits were made, the number of new customers that may be attracted is , to the nearest whole number. **M1**, **A1**

**(TOTAL : 20 MARKS)**

**Question 4**

(a) Scenarios when used:

(i) histogram – when the data is continuous and dealing with one variable. **A2**

(ii) multiple bar chart – when comparing a number of variables. **A2**

(b) (i) Component bar chart – to compare total number of customers across service centres or accounts. **A2**

Multiple bar chart – to compare across service centres or accounts **A2**

(ii) Multiple bar chart:

**M1** (Scale), **M1** (Labeled axes), **M1** (Key), **A3** (Accurate bars)

(c) (i) P(one item of category A and one item of category B)

=  since they are independent, **M1**

= , **M1**, **A1**

(ii) P(one item from each category)

=  since they are independent, **M1**

= , **M1**, **A1**

**(TOTAL : 20 MARKS)**

**Question 5**

(a) (i) A random sample is a set of units selected in such a way that every member of the population has an equal chance of being included. **A2**

(ii) I. Questioning allows further probing into reasons. **A2**

II. Others do not prefer them due to logistics: planning time and venue for the interview – can be demanding. **A2**

(b) (i) Standard deviation:

Mean = , **M1**

Standard deviation,

 , **M2**

= , **M1**, **A1**

(ii) Coefficient of skewness:

Mode = 12, **M1**

The Pearson’s coefficient of skewness is:

, to 2D **M1**, **A1**

Interpretation: The value of 2.85 shows that the data is skewed to the right (i.e. the right tail of the distribution is longer than the left). **A2**

**(TOTAL : 20 MARKS)**