General Mathematics - Data and Statistics Assessment NAME: TOTAL: /42

PART 1 - UNIVARAITE DATA (Chapter 2)

1. A survey was taken of 20 University students as to their level of enjoyment of their course. The possible responses were: “Strong dislike”(SD), “Dislike”(D), “Neutral”(N), “Like”(L) and “Strongly like”(SL). Here is the raw data

D HD D D L HL N D HD L N D D N

L HD N N D D

a) What type of data is this? Discrete, continuous, ordinal or nominal? (1 mark)

b) Construct and complete a frequency table of the data. Make sure to include the count **AND** the frequency percentage. (2 marks)

b) Use the table to find: (2 marks)

i) The number of students that “like” the course.

ii) The percentage of students that “dislike” the course.

c) Construct a bar chart (column graph) for the data. (Use a ruler when measuring the axis out. Don't estimate) (1 mark)

2. The number of points scored for a footy team is given for 15 games. Here is the raw data:

67 92 101 74 81 65 86 88 97 79 54 75 69 94

81

a) Construct and complete a frequency table of the data, using intervals of 10. Make sure to include the count **AND** the frequency percentage. (2 marks)

b) Construct a histogram for the data, including both count **AND** frequency percentage on vertical axis. (Use a ruler when measuring the axis out. Don't estimate) (2 marks)

c) Describe the shape of the histogram (1 mark)

3. The data below shows the time it takes (in min) to get to school at an inner city school and an outer city school.

Inner-city school:

5 14 15 8 21 12 18 32 11 15 25 28

Outer-city school:

18 24 19 34 45 42 31 24 10 12 23 29

a) Draw a back-to-back stem and leaf plot of the data (2 marks)

b) Write a report comparing the two sets of data, describing the centre and spread (there are no outliers) (2 marks)

4. Twelve different fridges have the following prices

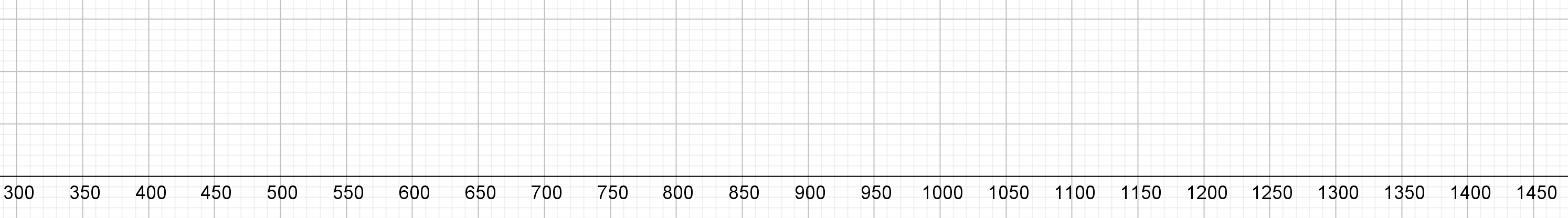
$350 $390 $450 $460 $480 $510 $540 $620 $750 $800 $860 $1400

a) Determine the 5-number summary for the data set (median, Q1, Q3, Minimum, Maximum). (2 marks)

b) Determine the inter-quartile range for the data set (1 mark)

c) Mathematically determine if there are any outliers and state which value(s) are outliers (2 marks)

d) Draw a box plot of the data. (2 marks)



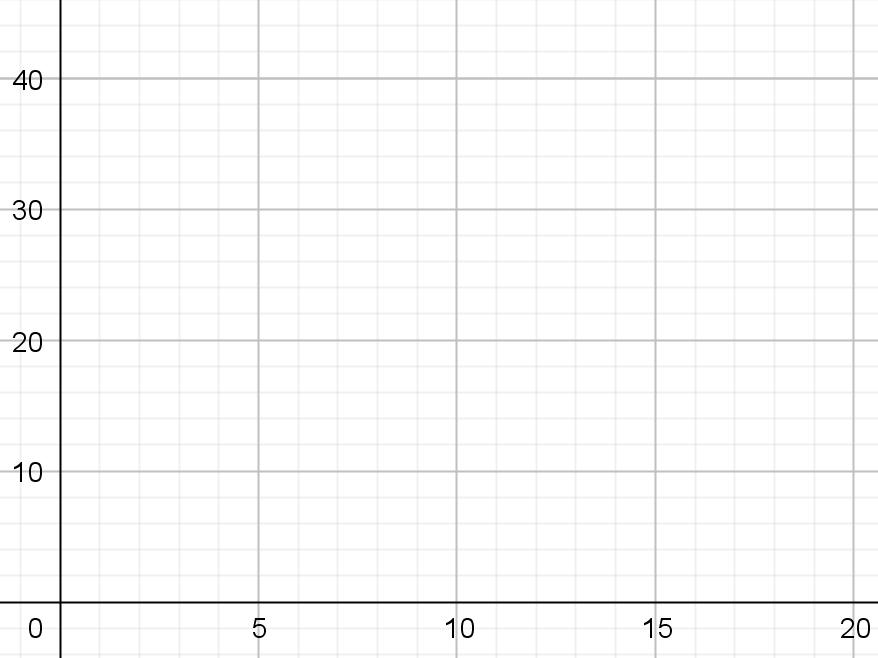
e) What percentage of data values lie between the minimum and Q1? (1 mark)

PART 2 - BIVARAITE DATA (Chapter 7)

1. A manufacturer who is interested in minimising the cost of training gives 15 of his machine operator’s different amounts of training. He then measures the number of machine errors made by each of the operators. The results are shown in the table below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Hours spent training | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 10 | 12 | 13 | 14 | 16 | 18 |
| Number of errors during first week | 43 | 31 | 34 | 16 | 27 | 25 | 24 | 18 | 24 | 22 | 15 | 6 | 8 | 6 | 8 |

1. Identify the explanatory and response variables. (2 marks)
2. Draw a scatterplot of the data by hand using the correct variables for each axis. (2 marks)



1. Describe the data in terms of strength, form and direction (2 marks)
2. Determine the correlation coefficient () for the scatter plot to describe the strength and direction of the plot. How does this compare with your answer from question c)? (3 marks)

2. There was a study to determine if there was a correlation between the hand length of children and their athletic ability. The question proposed was: Does the hand length of children impact their athletic capability? The hand lengths and standing jump height of 30 random students from an all-girls high school was collected.

|  |  |  |
| --- | --- | --- |
| Student | Hand length (cm) | Jump height (cm) |
| 1 | 14.3 | 31 |
| 2 | 15.8 | 48 |
| 3 | 17.9 | 54 |
| 4 | 18.2 | 50 |
| 5 | 16.9 | 45 |
| 6 | 17.1 | 41 |
| 7 | 18.0 | 57 |
| 8 | 19.2 | 61 |
| 9 | 16.3 | 44 |
| 10 | 15.5 | 43 |
| 11 | 14.9 | 39 |
| 12 | 16.2 | 42 |
| 13 | 17.0 | 50 |
| 14 | 17.8 | 45 |
| 15 | 16.4 | 46 |
| 16 | 18.5 | 44 |
| 17 | 17.4 | 44 |
| 18 | 15.8 | 42 |
| 19 | 18.3 | 46 |
| 20 | 18.9 | 51 |
| 21 | 19.1 | 57 |
| 22 | 17.5 | 40 |
| 23 | 15.0 | 35 |
| 24 | 15.6 | 38 |
| 25 | 16.0 | 38 |
| 26 | 15.3 | 38 |
| 27 | 16.9 | 41 |
| 28 | 16.2 | 41 |
| 29 | 16.4 | 46 |
| 30 | 19.3 | 58 |

1. Use geogebra to sketch a scatterplot with a regression line. Provide the equation of the regression line in terms of the two variables. (3 marks)

Equation:

1. Determine the correlation coefficient and describe the strength and direction of the plot. (2 marks)
2. Predict the jumping height of a person with a hand length of 18.5 cm. What could be said about the faith in this prediction? Explain. (2 marks)
3. Predict the jumping height of a person with a hand length of 22.5 cm. What could be said about the faith in this prediction? Explain. (2 marks)

3. The following questions relate to correlation and causality. Provide possible non-causal explanations for them.

1. It was found that there is a strong correlation between the number of churches and the number of deaths each year in cities in a country. Can it be inferred that having mor churches causes more deaths? (2 marks)
2. It was found that there was a correlation between the margarine consumption per capita and the divorce rate in a city. Can it be inferred that increasing margarine consumption will lead to divorce? (2 marks)