|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Name: | | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | Date: *\_\_\_\_\_\_\_\_\_\_\_* |
| S:\AdminShared\All Staff\1 College Logo's\Baldivis_Logo_colour.jpg | | **Year 11 Mathematics: Applications**  **Investigation 3, 2016**  **Topic – Univariate data analysis and the statistical investigation process**  **Preparation Activities** | | | |
| **Important Information:**  *Although the take-home component is not worth any marks, it is essential in preparation for the in-class component. Knowledge and skills gained will be extended in the in-class validation component. This in-class validation will be completed under test conditions on the day in which this take-home component is due. The take-home component may be used when completing the in-class component. Contact may be made to parent(s) if the take-home component is not available for submission (at the start of the lesson).* | | | | | |
| **Date out:** | *Week \_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_* | | **Date Due:** | *Week \_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_* | |
| **Take home component weighting:** | *0% of the year* | | **In-class component weighting:** | *10% of the Semester, 5% of the Year* | |
| **AIM:** *In this assessment, you will be investigating the statistical investigation* | | | | | |

**Task conditions**

**Students can bring a page of notes summarising the findings of the class investigations to the in-class validation.**

For all of these activities students are expected to work collaboratively to;

* Identify all data displays appropriate for the data collected
* Present their findings in at least one appropriate data display
* Identify which statistics are not appropriate in the situation
* Share their findings with the other students in the class

For Activity 1 and Activity 2, students will also need to:

* Sample the population indicated
* Choose and calculate relevant statistics

**Activity 1**

Plan and carry out an investigation to answer one of the following questions.

1. What are the most popular drinks sold at the canteen during lunchtime?

2. What is the least popular subject chosen by Year 11 students at our school?

3. By what means of transport do our Year 11 students arrive at school?

4. What is the most popular holiday destination for staff in our school?

5. What is the most popular Saturday morning activity for students in Year 11?

Investigating categorical data is the focus for this activity. You will need to decide which data to collect and determine the categories in which the data belong. The numbers in each category are to be compared during this investigation.

**Activity 2**

Select one of the following questions and once you have clarified the question, collect data from a random sample of 50 Year 11 students. Prepare a summary of your data by calculating relevant statistics and preparing two different data displays.

1. On how many days did you watch television last week?

2. How many texts did you receive yesterday?

3. On how many days did you buy food from the canteen in the last fortnight?

4. How many hours of paid work have you done in the past week? [to the nearest hour]

5. How many different sporting grounds have you been to in Perth?

What type of data will be collected in each of these situations?

The data collected could be used to suggest answers to simple problems

e.g., *How many different sporting grounds have been visited by students in Year 11*?

Identify three such problems for which your data could indicate the solution.

**Activity 3**

Some students were given a selection of problems to solve. Five groups each chose a different problem and then in each group the students devised one question to clarify their problem. The questions from each of the five groups are given below.

1. What was the total rainfall in Sydney for each month of the last four years?

2. What was the relative humidity in Brisbane at 9 am each day of the last two months?

3. What were the wind-gust speeds every half hour in Adelaide yesterday?

4. What were the minimum temperatures in Melbourne each month for last four years?

5. What were the maximum temperatures each day in January and February in Hobart last year?

Select one of the questions and use the internet to collect the data indicated.

The Australian Bureau of Meteorology website would have the data for each question.

What type of data have you collected?

Calculate the mean, range and standard deviation.

Draw a histogram to represent the data.

What conclusions can you draw from the data display?

**Summary**

After investigating students should:

1. Know and be able to use the following terms:

categorical ordinal nominal discrete continuous

modality bimodal unimodal multimodal outlier

skew (positive and negative)

symmetric distribution

statistical investigation process

2. Construct and interpret the following data displays:

bar charts tables dot plots stem plots histograms

3. Calculate and interpret the following statistics:

mean mode median range standard deviation

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name: | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | | | Date: *\_\_\_\_\_\_\_\_\_\_\_* |
| Description: S:\AdminShared\All Staff\1 College Logo's\Baldivis_Logo_colour.jpg | **Year 11 Mathematics: Applications**  **Investigation 3, 2016**  **Topic – Univariate data analysis and the statistical investigation process**  **In Class Component** | | | | | 42  = % |
| **Total Time:** | ***50*** *minutes* | |  | | | |
| **Reading Time:** | *5**minutes* | |
| **Working Time:** | *45**minutes* | |
| **Equipment:** | *SCSA Formula sheets, ClassPad, scientific calculator, 1 x A4 page of notes* | | | | | |
| **Date out:** | | *Week \_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_* | | **Date Due:** | *Week \_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_* | |
| **Take home component weighting:** | | *0% of the year* | | **In-class component weighting:** | *10% of semester, 5% of the year* | |
| **AIM:** *In this assessment, you will be investigating the statistical investigation* | | | | | | |

**Question 1 (6 marks)**

Complete the table by identifying the data described in the first column as being either *numeric or categorical*.

If the data is numerical, then classify it as *discrete or continuous* in the third column.

If the data is categorical then classify it as *ordinal or nominal* in the third column.

|  |  |  |
| --- | --- | --- |
| Data | Numeric  Or  Categorical | If numeric then discrete or continuous  If categorical then ordinal or nominal |
| On the meteorology website there is a bar indicating the amount of rain on the radar. The bar is white or blue if the rain is light, green or yellow if the rain is moderate and red if the rain is heavy. |  |  |
| For the athletics carnival, a three-letter code is written on each student’s hand for identification purposes. The codes are allocated at random and no two students have the same code. |  |  |
| Annie is investigating water storage and has decided to collect data to determine the number of dams in each state. |  |  |
| Plants are affected by the amount of sunshine (hours) each day so Max has collected sunrise and sunset times and calculated the differences. |  |  |
| The air pressure (hectoPascals) is also available on the meteorology website and two readings that Jon wrote down were 1017.0 and 1016.4 |  |  |
| Lee was travelling in remote areas of Australia and collecting data from each weather station to see how often each day the local resident was recording the observations. |  |  |
| Before Tim goes sailing on Sunday, he checks the direction of the wind speed which can be E, NE, W, SW, etc.. |  |  |

**Question 2 (7 marks)**

(a) Nat is trying to decide which player to drop from the team for the finals. She goes through the investigation process which consists of the following steps. The steps are out of order. Using the letters in front of these steps, write down the appropriate order. (2)

**A** Nat prepares a column graph and plots the percentage of games won for each player.

**E** Nat goes to the web and locates the number of games won and lost each week for each player.

**G** Jacki is dropped from the team.

**K** Nat discovers the presence of an outlier; a player with a much lower percentage than all the other players.

**N** Nat calculates the ratio of games won to games lost for each player

**S** For each player Nat determines the games won as a percentage of games played.

**T** Nat decides to look at the number of games that the players have won and lost each week.

**------- ------ ------ ------ ------ ------ ------**

(b) For each set of data described, circle the best type of display to represent the data. (5)

|  |  |  |
| --- | --- | --- |
| Data set | Display 1 | Display 2 |
| Jody had been investigating favourite ice cream flavours for the students in her class and had collected numbers of students favouring five different types of ice cream. | bar chart | stem plot |
| Shoe sizes for the 32 Year 8 students were one of the following sizes: 6, 6.5, 7, 7.5, 8 and 8.5. Data was collected for all the students in the class | stem plot | dot plot |
| The incomes of the basketball players in the local league, about 100 players in all, were recorded so that an interstate comparison could be made. | histogram | bar chart |
| Ruby collected data on rainfall. She has the total rainfall data for each month for the last 100 years. | histogram | dot plot |
| Toby had to prepare a display to show the heights of the 15 students in his Year 11 Maths class | histogram | stem plot |

**Question 3 (9 marks)**

Some people travelling to a concert were arguing about the number of roundabouts they had to pass so Ben started counting the types of intersections they went through.

He classified the intersection as a *roundabout* (regardless of any other signs), a *give way* (only give-way signs and no other) or a *stop.*

His results for the next 20 intersections are shown in the diagram below.

|  |  |  |
| --- | --- | --- |
| **Give way**    **60%** | **Roundabout**    **30%** | **Stop**    **10%** |

(a) This is categorical data. (4)

(i) Is the data ordinal or nominal? Explain your choice of answer.

(ii) State the number of categories.

(ii) Name each of the categories

(b) Show another display to represent these data. (3)

(c) What feature of the diagram given would make it an accurate representation of the data that Ben collected? (1)

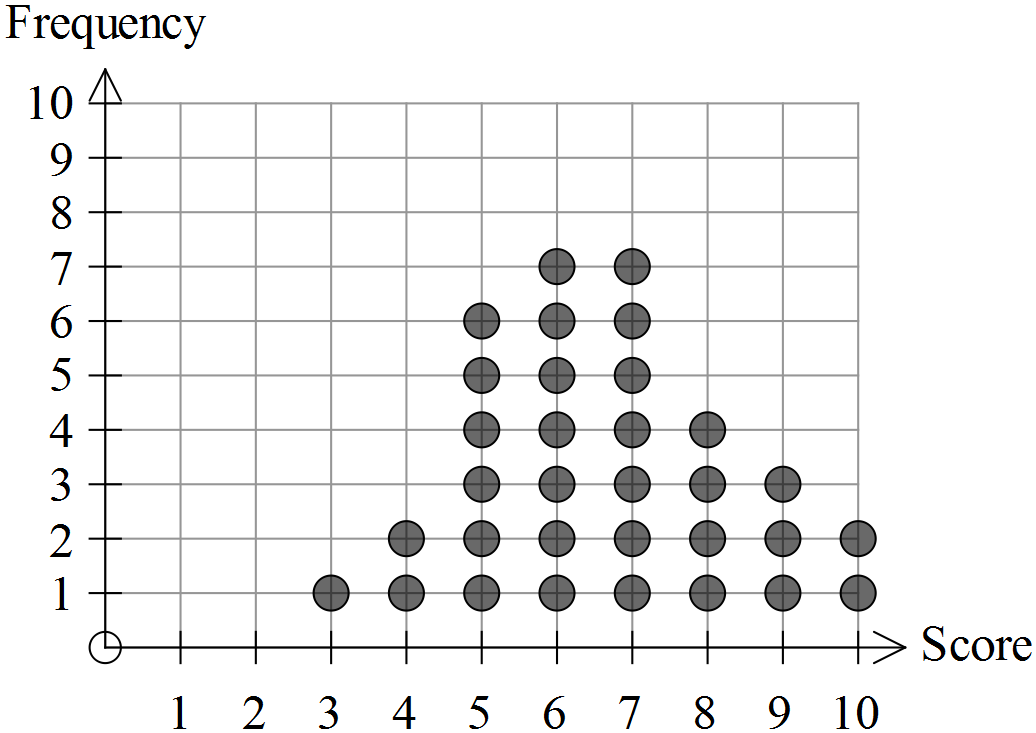
(d) Ben then calculated the mean of the types of intersections as follows: (1)



What does this value represent?

**Question 4 (9 marks)**

The dot plot provided shows the mental maths scores for the 32 students in Year 7.



(a) For this data set, state the (2)

(i) maximum score.

(ii) median score.

(b) Describe the mode of this distribution. (2)

(c) How many students scored less than 5? (1)

(d) What statistic can be calculated using the following process? (1)

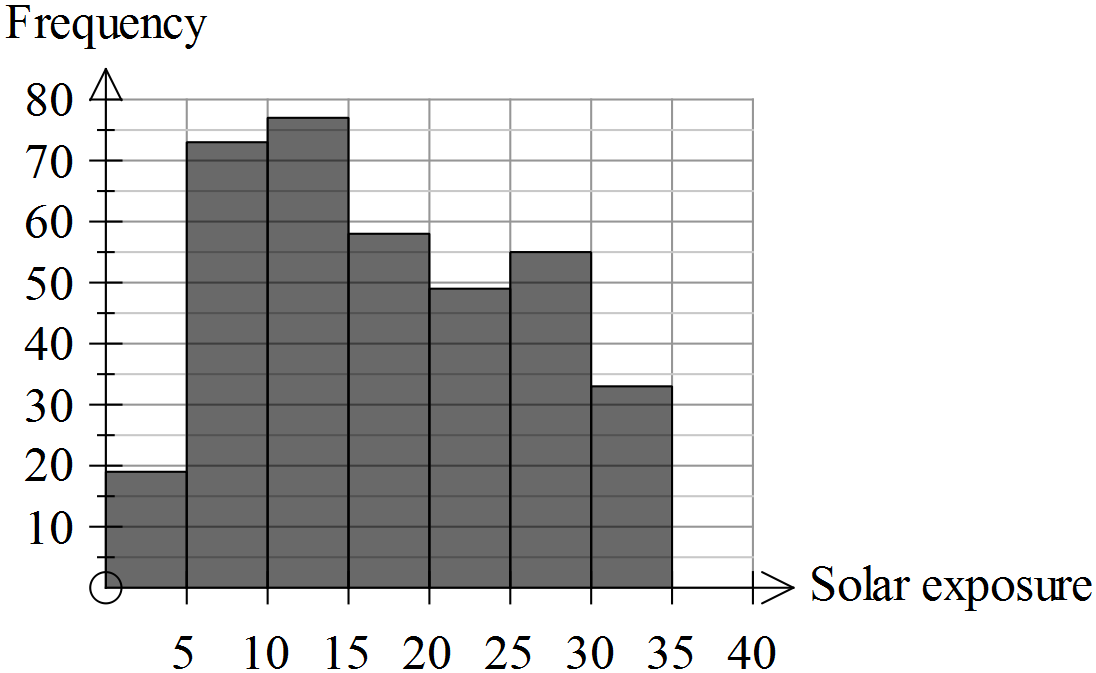


(e) Describe the shape of the distribution and comment on the success of the students with their mental maths skills. (3)

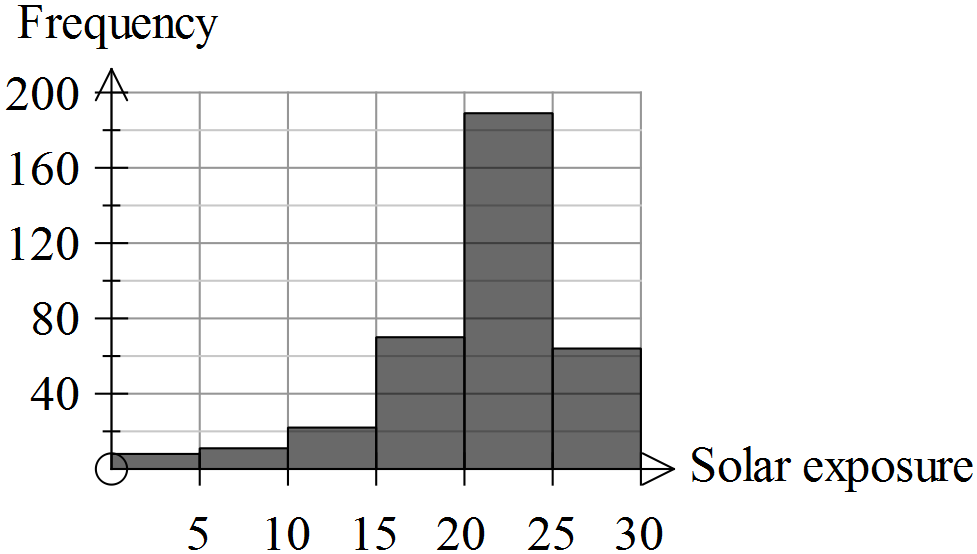
**Question 5 (11 marks)**

The histograms provided show the daily solar exposure (MJ per m2) at weather stations in Darwin and Canberra for 2013. The total solar energy was determined for each day of the year for both capital cities. [Data obtained from the website of the Australian Bureau of Meteorology]

**Canberra**



**Darwin**



For both cities the readings were expressed to 1 decimal place and were grouped in 5 MJ intervals.

(a) Approximate the number of readings in each of these groups for the city of Darwin and enter these in the table provided. (2)

|  |  |
| --- | --- |
| Solar exposure (MJ per m2) | Frequency (days) |
| 0 – 4.9 |  |
| 5 – 9.9 |  |
| 10 – 14.9 |  |
| 15 – 19.9 |  |
| 20 – 24.9 |  |
| 25 – 29.9 |  |

(b) Which city had the highest reading for solar exposure? (2)

Justify your selection.

(c) For which city would the standard deviation be lower? (2)

Justify your selection.

(d) For which city would the mean be higher? (2)

Justify your selection

(e) Describe the shape of the data distribution for Darwin. (3)

**End of questions**

**Statistical investigations**

**Extended investigation Part 2:** **In-class validation**

**Solutions and marking key**

**Question 1**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Solution | Marking key/mathematical behaviours | Marks |
| (a) | Categorical – ordinal  Categorical – nominal  Numeric – discrete  Numeric – continuous  Numeric – continuous  Numeric – discrete  Categorical - nominal | Selects numeric or categorical correctly for:   * all seven * five or six * three or four   Selects discrete, continuous, ordinal or nominal correctly for   * all seven * five or six * three or four | 3  2  1  3  2  1 |

**Question 2**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Solution | Marking key/mathematical behaviours | Marks |
| (a) | T E N S A K G or  T E S N A K G | Lists events in correct order  [1 mark if one out of order] | 2 |
| (b) | Bar chart Dot plot Histogram Histogram Stem plot | Identifies most appropriate graph in the each situation | 5 |

**Question 3**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Solution | Marking key/mathematical behaviours | Marks |
| (a) | (i) nominal  They are not in order of size or importance etc... They are named  (ii) 3  (iii) Give way, roundabout, stop | * Indicates (1) and justifies type of categorical data * States number of categories * Names categories | 2  1  1 |
| (b) |  | * Provides categories * States frequencies * Represents categories and frequencies accurately in a table or graph | 1  1  1 |
| (c) | The sections are in proportion to their frequencies | Understands the importance of scale | 1 |
| (d) | It does not represent anything useful | Identifies mean for categorical data is not appropriate | 1 |

**Question 4**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Solution | Marking key/mathematical behaviours | Marks |
| (a) | (i) 10  (ii) 6.5 | * Identify maximum score * Determines the median | 1  1 |
| (b) | Bi-modal. 6 and 7 have equal (and maximum) frequencies | * Identifies bi-modal * States the mode(s) | 1  1 |
| (c) | 3 | * Determines frequency | 1 |
| (d) | Mean | * Identifies process for calculating mean | 1 |
| (e) | Skewed negatively – most students score 5 or more, very few below 5  Hard to determine success unless you know what the test is out of.  (No outliers)  (Not symmetric) | * Identifies data is skewed * Identifies direction of skew * Recognises success depends on maximum score possible (cf achieved) | 1  1  1 |

**Question 5**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Solution | Marking key/mathematical behaviours | Marks |
| (a) | |  |  | | --- | --- | | Solar exposure  (MJ per m2) | Frequency  (days) | | 0 – 4.9 | 8 | | 5 – 9.9 | 12 | | 10 – 14.9 | 21 | | 15 – 19.9 | 70 | | 20 – 24.9 | 190 | | 25 – 29.9 | 62 | | * Approximates 6 values * Approximates no more than 4 values (1) | 2 |
| (b) | There are data in the 30-35 class for Canberra but not for Darwin. | * Locates maximum reading * Compares cities | 1  1 |
| (c) | Canberra  The data are more similar to each other but for Darwin the variation is higher | * Identifies graph with lower standard deviation * Explains why the deviation is lower | 1  1 |
| (d) | Darwin  About two thirds of the readings are over 20 but for Canberra only about one third are over 20. | * Identifies graph with lower mean * Justifies choice of city | 1  1 |
| (e) | Skewed negatively  Unimodal / One class has very high proportion of readings | * Identifies data is skewed * Identifies direction of skew * Describes modality | 1  1  1 |