Week 1
Term 3
2023

HAWKER COLLEGE
Engage | Inspire | Achieve

MA2
Univariate data

Goals

Goals:

- Review the statistical investigation process
- Classify a categorical variable as ordinal or nominal and use tables and bar charts to organise and display the data
- Classify a numerical variable as discrete or continuous
- Choose from dot plot, stem plot, bar chart or histogram to display and describe the distribution of a numerical dataset, it's shape (symmetric versus positively or negatively skewed), location and spread, and outliers, and interpret this information in the context of the data

Theoretical Components

Resources:
PDF file: Week 1 Notes and Exercises
Introduction to stem and leaf plots
http://www.youtube.com/watch?v=OaJXJduRiIE
How to construct and use a divided bar graph
http://www.youtube.com/watch?v=oTYJx_mvcug
Interpreting graphs
https://www.youtube.com/watch?v=Sbzw653gW10

Knowledge Checklist

- Numerical and categorical
- Discrete and continuous
- Nominal and ordinal
- Stem and leaf plot
- Histogram
- Bar chart
- Frequency table
- Tally
- Dot plot
- Reading and interpreting graphs

Order

1. Read through the notes and examples
2. Work through the exercises
3. Complete the investigation at the end of the booklet.
4. Complete the reflection at the end of the booklet
5. Come and see your teacher and make sure you are up to date.

Practical Components

Work through the exercises and show the completed tasks to your teacher.

Be sure to ask for help as you need for the successful completion of all tasks.

Remember to regularly check Google Classroom for messages.

Investigation

Complete the task at the end of the booklet and submit your work for checking. ☺

QFO

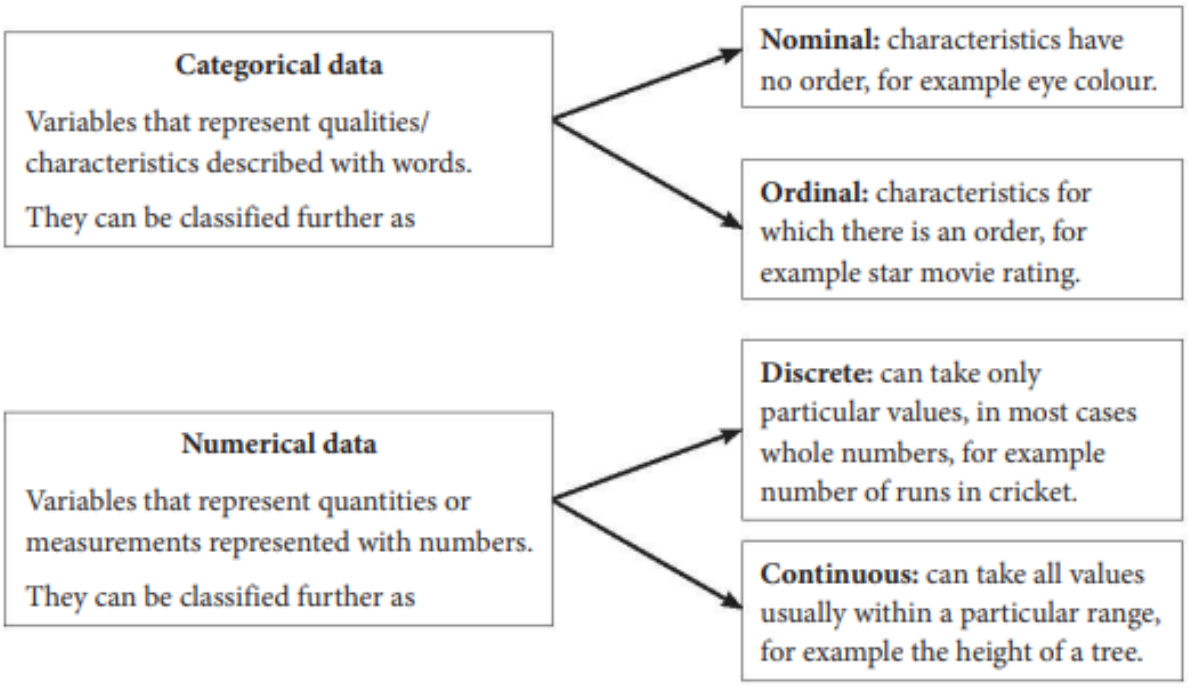
MATHEMATICAL APPLICATIONS 2

WEEK 1 NOTES & EXERCISES

TYPES OF DATA

This era has been called the *information age*. The success of any venture, large or small, relies upon the accurate sourcing, processing and analysing of information. These tasks are assigned to statisticians. In this unit we learn statistical methods which can be used in the processing and analysis of data.

There are several different types of data that may result from a survey or experiment. The type of data will influence how they can be analysed and how useful they might be to a researcher. A distinction is made between *categorical* data and *numerical* data.



Let us consider these questions and their responses:

'Do you own a pet?' The answer is 'yes' or 'no'. **A categorical answer.**

'How many pets do you own?' The answer is a number. **A numerical answer.**

EXERCISE 1

1. State whether the data formed in each of the following situations would be categorical or numerical:

a) The number of matches in each box is counted for a large sample of boxes.

b) The lengths of 40 cod are recorded by a fisheries inspector.

c) The occurrence of hot, warm, mild and cool weather for each day in January is recorded.

d) Cinema critics are asked to judge a film by awarding it a rating of very bad, bad, average, good or very good.

2. State whether the categorical data formed by each of the following situations are nominal or ordinal.

a) Individual members of a group of people are asked to indicate their level of agreement with the statement: 'I support open euthanasia'. Each respondent must select one of the following responses: strongly agree, agree, unsure, disagree, or strongly disagree.

b) Students are asked to write down the name of their favourite rock band.

c) The mode of transport that each student uses to get to school is recorded.

3. State whether the numerical data formed by each of the following situations are discrete or continuous.

a) The heights of 60 tomato plants at a plant nursery

b) The number of jellybeans in each of 50 packets

c) The time taken for each student in a class of 6-year-olds to tie their own shoes

d) The petrol consumption rate of a large sample of cars

4. Sort the following CensusAtSchool question topics according to whether they will yield categorical or numerical data.

Amount of money earned last week	Language mostly spoken at home
Arm span	Foot length
Concentration exercise (seconds)	School post code
Dominant hand reaction time	State/Territory live in
Favourite sport	Travel method to school
Height (cm)	Travel time to school
Hours slept per night	Year level
Birthdate	Opinions on environmental conservation

Categorical (label N for Nominal or O for Ordinal)	Numerical (label D for discrete or C for continuous)

Why is some data that contains numbers, such as post codes and birthdates, considered categorical?

REPRESENTING DATA – FREQUENCY TABLES AND BAR CHARTS

The **frequency** of a particular data value is the number of times the data value occurs. For example, if four students have a score of 80 in mathematics, then the score of 80 is said to have a frequency of 4. The frequency of a data value is often represented by *f*.

A **frequency table** is constructed by arranging collected data values in ascending order of magnitude with their corresponding frequencies.

Example:

Heidi asked a group of her fellow students which night of the week they would prefer to have the school ball. She recorded the responses as T (Thursday), F (Friday), Sa (Saturday) and Su (Sunday). This is her data.

F, F, T, Sa, F, Sa, F, Sa, Sa, F, Sa, T, Sa,
F, F, T, T, Su, F, Sa, T, Sa, F, F, F, F, T, T

Arrange this data in a frequency table.

Solution:

We use the following steps to construct a frequency table:

Step 1: construct a table with three columns. In the first column, write down all the data values in ascending order of magnitude.

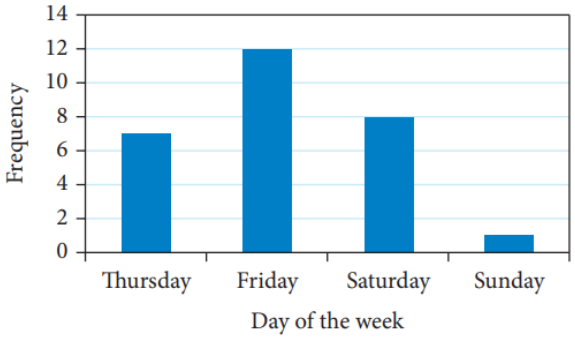
Step 2: to complete the second column, go through the list of data values and place one tally mark at the appropriate place in the second column for every data value. When the fifth tally is reached for a mark, draw a horizontal line through the first four tally marks as shown for in the above frequency table. We continue this process until all data values in the list are tallied.

Step 3: count the number of tally marks for each data value and write it in the third column.

Day of the week	Tally	Frequency
Thursday	𝍸 ||	7
Friday	𝍸 𝍸 ||	12
Saturday	𝍸 |||	8
Sunday	|	1
Total		**28**

A **column graph** can then be used to provide a visual representation of the data.

The column graph for the above example is below:



The horizontal axis will be the days of the week.
The vertical axis will be the frequency.

Columns will have spaces between them.

EXERCISE 2

1. Tina surveyed a selection of homes in her suburb to find out the ages of each child. She recorded the data as B for babies up to 2 years, P for pre-schoolers, I for those who attend infant school and O for children older than 8 years. This is the data:

   B B O I O O B P P P O B I I O O I I I P O B B O I
   P O O P I P P O B I I O P O O B I P B I

   a) Arrange the data in a frequency table.

Ages	Tally	Frequency
Total		

   b) What category of children contains the largest number of children?

   c) What percentage of children in the survey are babies?

2. Tina also undertook a survey of the number and type of pets owned by the families. She used D for dog, C for cat, M for mice, R for reptile, F for fish, O for other and N for none. The results were:

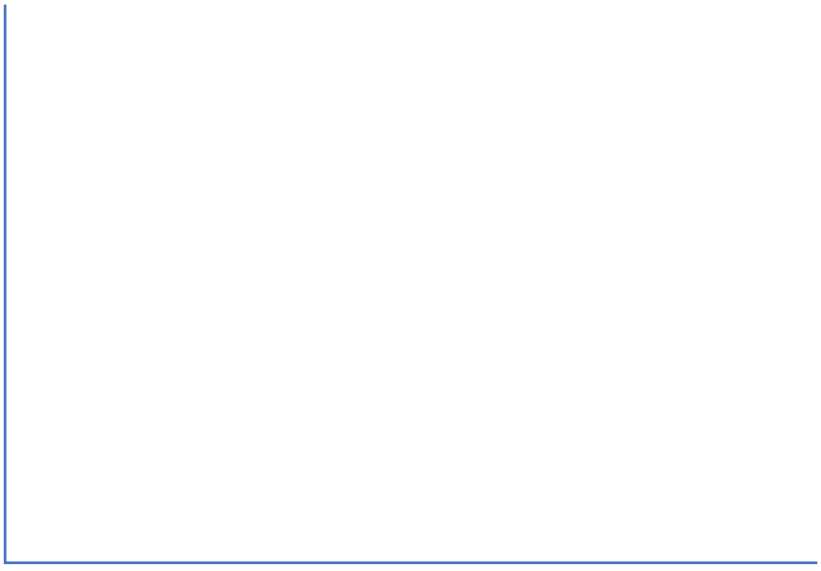
R O F D D	D R C C M	D C N D D	O R C C D
F N O N D	N D D C M	M D M N D	N M C D C

a) Arrange the data in the frequency table.

	Tally	Frequency
Total		

b) What percentage of people had no pets?

c) Construct a bar chart of the data.



STEM-AND-LEAF PLOT

A stem-and-leaf plot, or stem plot for short, is a way of displaying a set of data.

It is best suited to data which contain up to about 50 observations (or records). The following stem plot shows the ages of people attending an advanced computer class

Stem	Leaf
1	6
2	2 2 3
3	0 2 4 6
4	2 3 6 7
5	3 7
6	1

The ages of the members of the class are 16, 22, 22, 23, 30, 32, 34, 36, 42, 43, 46, 47, 53, 57 and 61. A stem plot is constructed by breaking the numerals of a record into two parts — the stem, which in this case is the first digit, and the leaf, which is always the last digit.

Example

The number of cars sold in a week at a large car dealership over a 20-week period is given below.

16 12 8 7 26 32 15 51 29 45 19 11 6 15 32 18 43 31 23 23

Construct a stem plot to display the number of cars sold in a week at the dealership.

Lowest number = 6

Highest number = 51

Use stems from 0–5.

Before we construct an ordered stem plot, construct an unordered stem plot by listing the leaf digits in the order they appear in the data.

Stem	Leaf
0	8 7 6
1	6 2 5 9 1 5 8
2	6 9 3 3
3	2 2 1
4	5 3
5	1

Now rearrange the leaf digits in numerical order to create an ordered stem plot. Include a key so that the data can be understood by anyone viewing the stem plot.

Stem	Leaf
0	6 7 8
1	1 2 5 5 6 8 9
2	3 3 6 9
3	1 2 2
4	3 5
5	1

Key: 2|3 = 23 cars

Stem	Leaf
6	1
6	6 6 7 8 9 9 9
7	0 1 1 2 2 3
7	7

Key: 6|1 = 61

EXERCISE 3

1. In the following, write down all the pieces of data shown on the stem plot.

Stem	Leaf
0	1 2
0	5 8
1	2 3 3
1	6 6 7
2	1 3 4
2	5 5 6 7
3	0 2

2. The stem and leaf plot represents the points scored per match by the Kangaroos in a football season.

Stem	Leaf
4	3 5 9
5	2 0 8 7
6	4 7 8 5 1 2
7	2 9 3 0
8	4 2 9

a) Redraw as an ordered stem and leaf plot.

b) How many matches were played in the season?

c) What was the highest score for a match?

d) What was the lowest score for a match?

3. The money (to the nearest dollar) earned each week by a busker over an 18-week period is shown below. Construct a stem plot for the busker's weekly earnings.

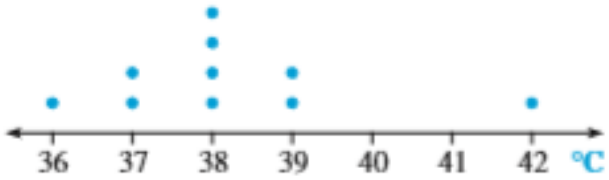
5	19	11	27	23	35	18	42	29
31	52	43	37	41	39	45	32	36

DOT PLOTS

A dot plot is a simplified type of histogram and is a good way of presenting a small amount of data. It is easy to see where clusters of scores occur as well as to identify any outliers (numbers that seem out of line with the other scores). Each score is represented by a symbol, usually a dot.

Example

This dot plot represents the body temperatures (in °C) of hospital patients, where normal body temperature is about 37°C:



The graphical representation allows us to readily see the following points.

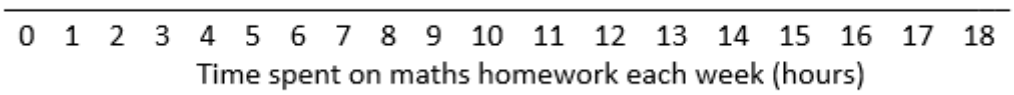
- There were 10 patients
- Most had a temperature of 38˚C
- Most temperatures were between 37˚C and 39˚C
- The 'outlier' was 42˚C (possibly a very ill patient).

EXERCISE 4

1. A Year 11 Maths class was surveyed to find out how many hours each student spent on maths homework each week. The results are shown below.

7	6	8	9	5	10	6	9	9	0	9	8
18	7	5	3	4	9	6	7	8	10	7	8

Draw a dot plot for this data.



2.

Lisa surveyed her group of friends about the amount of money they spent on fuel last week. The answers were rounded to the nearest dollar. This is the data she recorded.

$20 $28 $25 $26 $22 $26 $28 $28 $24 $22 $29 $28

a) Draw a dot plot for this data.

b) How many friends did Lisa survey?

c) What was the most common amount of money spent on fuel?

d) How many of Lisa's friends spent less than $26 on fuel last week?

FREQUENCY HISTOGRAMS AND POLYGONS

A histogram is used to represent quantitative data and is a column graph with no spaces between the columns. The height of each column represents the frequency of the scores. A frequency polygon is a line graph that plots score against frequency and can also be drawn by joining the midpoints of the tops of histogram columns.

Example

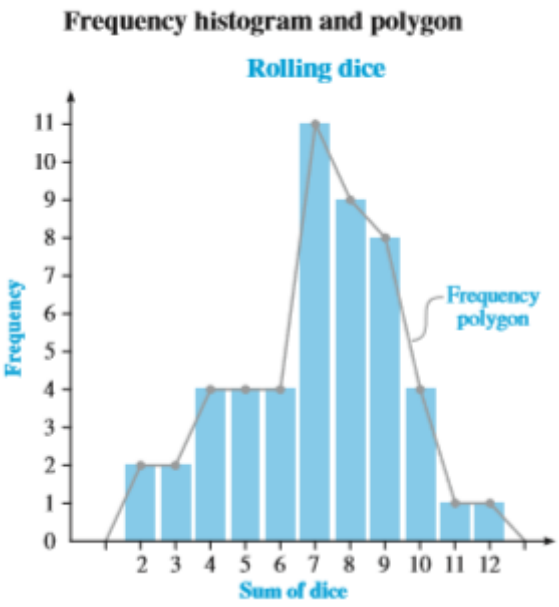
A pair of dice was rolled 50 times and the sum recorded. Construct a frequency distribution table, then draw a frequency histogram and polygon for the data.

7	7	10	12	8	7	9	8	6	10
9	9	8	5	7	3	6	4	5	7
4	4	8	7	10	8	9	7	11	7
5	9	4	6	3	9	2	7	8	8
9	7	5	8	10	6	8	7	9	2

With a large set of numbers, it is necessary to construct a **Frequency Distribution Table**. A Tally column, in which the scores are marked of in order, is useful as it helps reduce the possibility of missing a number. This table makes it easy to construct a histogram. The scores are placed on the horizontal line and the frequencies on the vertical line.

Frequency distribution table

Score	Tally	Frequency
2	||	2
3	||	2
4	||||	4
5	||||	4
6	||||	4
7	𝍸 𝍸 |	11
8	𝍸 ||||	9
9	𝍸 |||	8
10	||||	4
11	|	1
12	|	1
	Total	50



Note that the polygon is drawn through the centre of the columns.

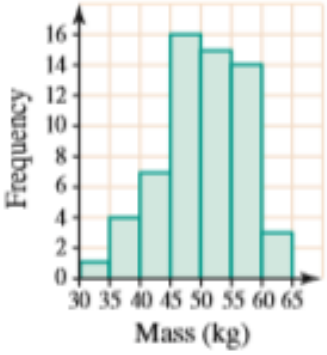
Example

The data below show the distribution of masses (in kilograms) of 60 students in Year 7 at Northwood Secondary College.

45.7	34.2	56.3	38.7	52.4	45.7	48.2	52.1	58.7	62.3
45.8	52.4	60.2	48.5	54.3	39.8	36.2	54.3	39.7	46.3
45.9	52.3	44.2	49.6	48.6	42.5	47.2	51.3	43.1	52.4
48.2	51.8	53.8	56.9	53.7	42.9	46.7	51.9	56.2	61.2
48.3	45.7	43.5	43.8	58.7	59.2	58.7	54.6	43.0	48.2
48.4	56.8	57.2	58.3	57.6	53.2	53.1	58.7	56.3	58.3

First construct a frequency table. The lowest data value is 34.2 and the highest is 62.3. Divide the data into class intervals. If we started the first class interval at, say, 30 kg and ended the last class interval at 65 kg, we would have a range of 35. If each interval was 5 kg, we would then have 7 intervals which is a reasonable number of class intervals. The resulting frequency table is shown below. A Tally column would certainly be necessary but has been omitted here for convenience. The corresponding histogram is also shown.

Class interval	Frequency
30–34.9	1
35–39.9	4
40–44.9	7
45–49.9	16
50–54.9	15
55–59.9	14
60–64.9	3
Total	**60**



EXERCISE 5

1. Forty primary school students went on an excursion. Their ages were:

7	9	9	12	11	6	5	12	8	6
11	7	5	9	7	10	5	9	5	5
8	9	5	10	10	7	11	8	5	8
9	10	7	6	7	9	9	8	10	11

a) Draw up a frequency distribution table using scores 5, 6, 7, …, 12.

Score	Tally	Frequency

b) Use your table to construct a frequency histogram and polygon.

2. The number of words per sentence in a magazine article were counted:

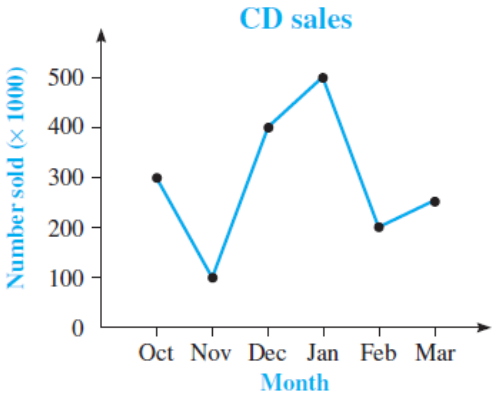
27	22	15	8	14	7	9	25	15	17	5	24
9	11	22	8	5	15	25	28	10	21	24	13
9	14	18	11	9	23	15	19	10	8	14	17

Using class intervals 0−9, 10−19, etc., draw up a frequency table. Draw the corresponding histogram. (Use a ruler)

INTERPRETING GRAPHS

EXERCISE 6

1. The line graph shows the sales of compact discs over a 6-month period.



a) How many CDs were sold in January?

b) What could be a reason for the most discs being sold in January?

c) In which month was the least number of CDs sold?

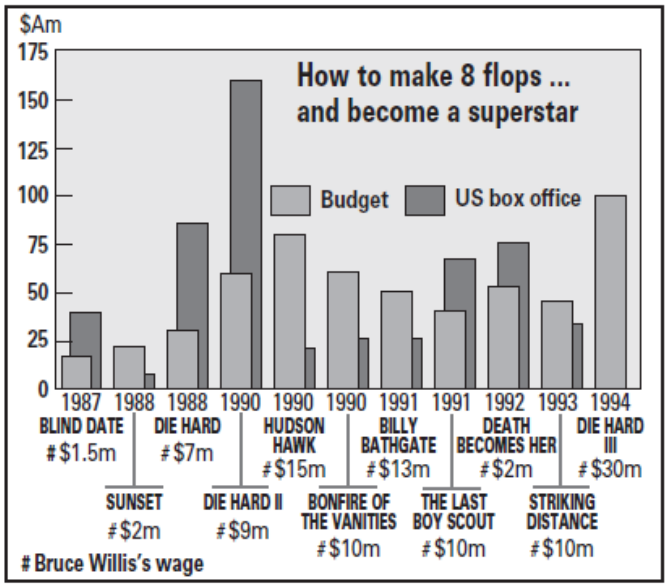
d) In which month were 250 000 CDs sold?

e) Between which 2 months did the greatest drop in sales occur?

f) What were the total sales for the 6 months?

g) What percentage of the CD sales were in November?

2.



a) What is the graph about?

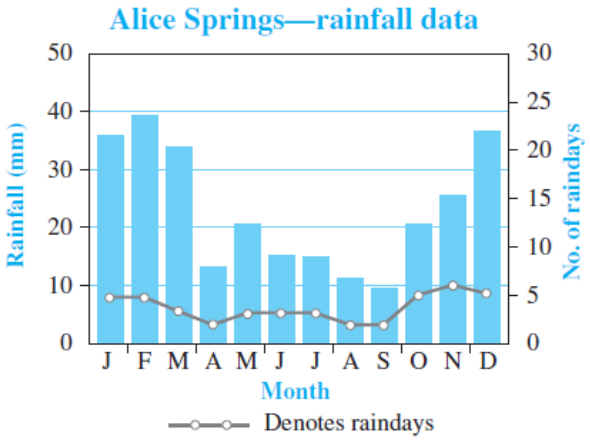
b) What was the budget for *Die Hard*, and how much did it make at the US box office?

d) Which movie grossed the most money at the US box office?

e) Which movies were flops? Explain how you know this.

(f) Which movie made the greatest percentage profit of takings at the box office compared to the budget for making the movie?

3. This graph from the Bureau of Meteorology gives rainfall information for Alice Springs.



a) Which month had the largest number of rainy days?

b) Which month had the highest rainfall?

c) What was the rainfall for:

(i) September (ii) April?

d) How many rain days were in:

(i) February (ii) October?

e) When would be the best time to travel to Alice Springs? Why?

f) Without using numbers this graph clearly shows a relationship between the number of rainy days in a month and the amount of rainfall. What is this relationship?

WEEK 1 INVESTIGATION

Olympic Records

Below are ten graphs showing how Olympic records have changed over time in ten athletic events.

The list of the ten events is given, in no particular order, to help you deduce which event belongs to each graph.

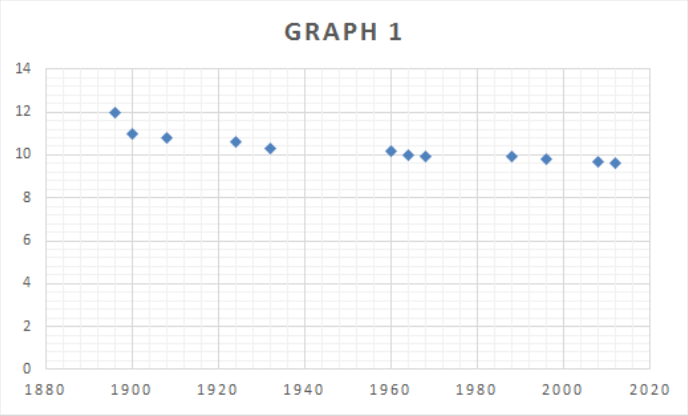
1. Men's High Jump
2. Women's 10,000 metres
3. Men's 100 metres
4. Women's High Jump
5. Men's Long Jump
6. Men's Decathlon
7. Men's 4 x 100 metres relay
8. Women's 100 metres
9. Men's Javelin
10. Women's 1500 metres

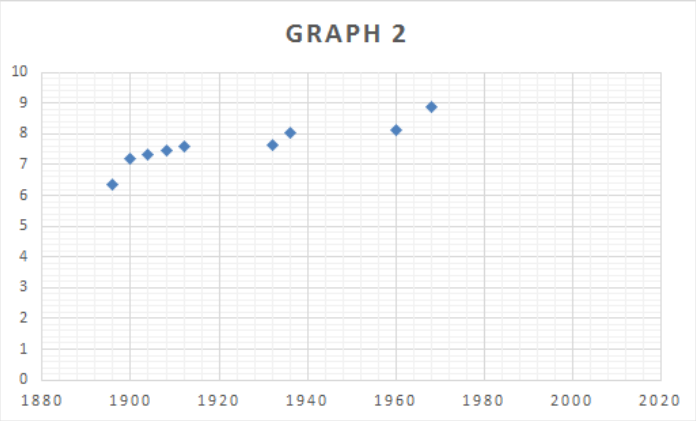
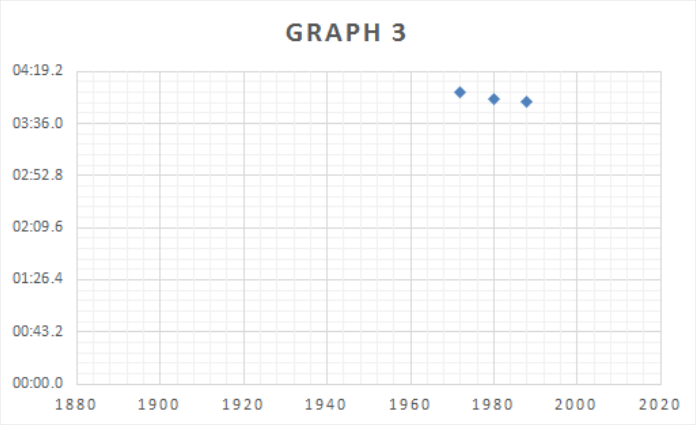
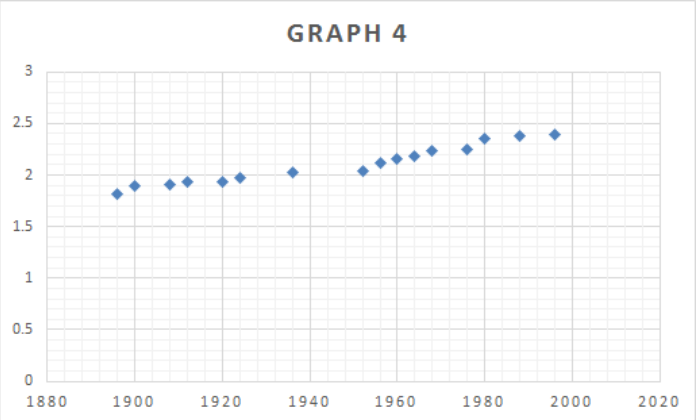
Can you deduce which event each graph represents?

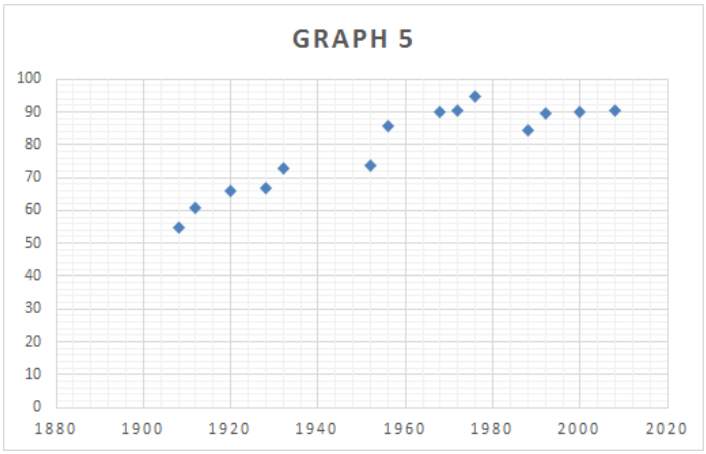
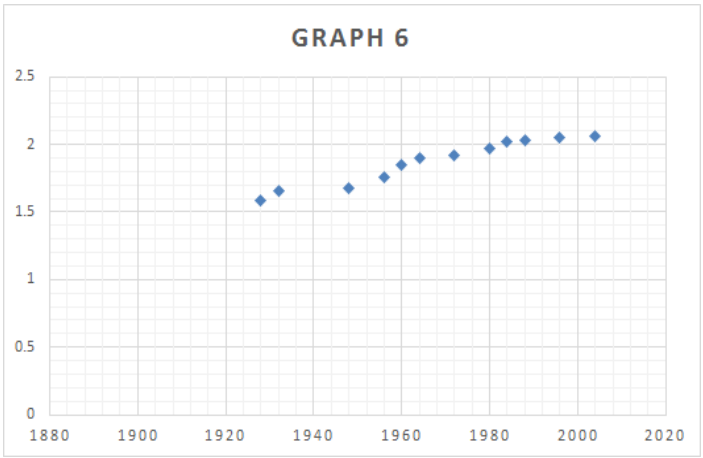
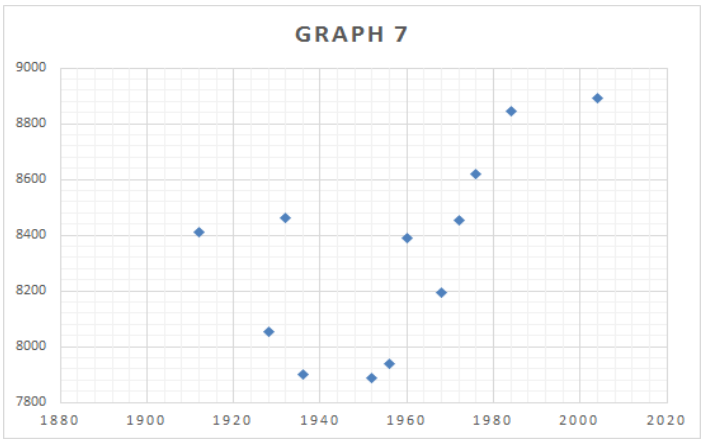
- Determine what the units might be on the vertical axis for each graph.
    - The x-axis represents the year of the event
- Some graphs show a decreasing trend and some an increasing trend. Identify those graphs and describe the trend.
- Are there any unusual features on any of the graphs? What might be a plausible explanation for them?

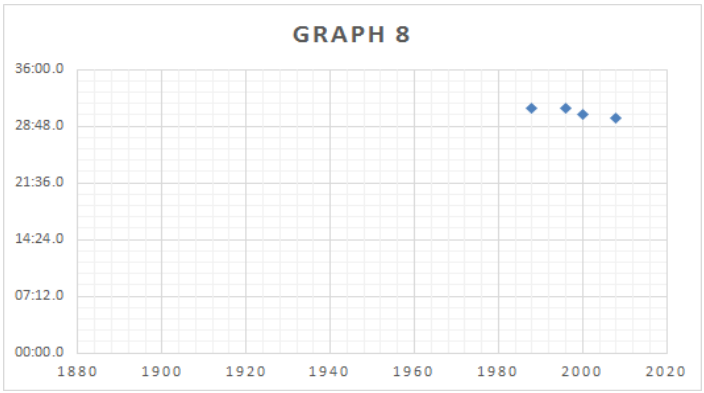
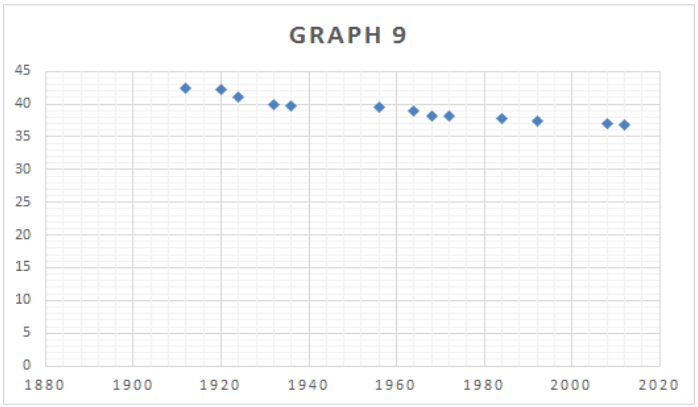
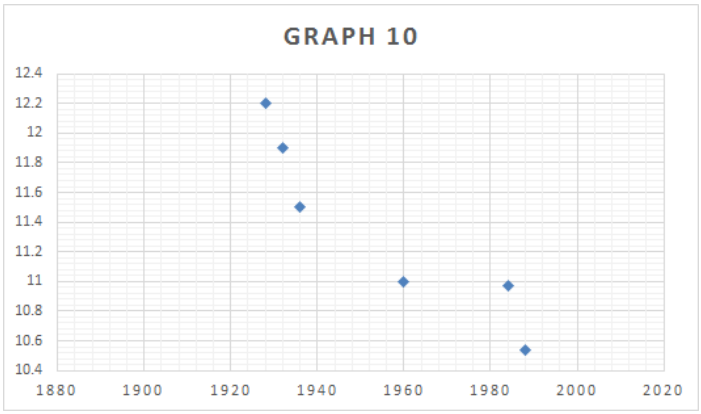
Be sure to write up your solutions next to each graph listing the **event**, labelling the **vertical axis with the appropriate units**, and discussing any increasing or decreasing trends and any **unusual features** (with explanations).

The first one has been done for you.

Graph	Description
	Event: *Men's 100 metres* Vertical Axis and Units: *Time - seconds* Features: • *Decreasing (ie. getting faster)* • *The Olympic record of 9.63 seconds was set in 2012.*

	Event: Vertical Axis and Units: Features:
	Event: Vertical Axis and Units: Features:
	Event: Vertical Axis and Units: Features:

	Event: Vertical Axis and Units: Features:
	Event: Vertical Axis and Units: Features:
	Event: Vertical Axis and Units: Features:

	Event: Vertical Axis and Units: Features:
	Event: Vertical Axis and Units: Features:
	Event: Vertical Axis and Units: Features:

MARKING RUBRIC

CRITERIA	EXPECTATIONS	POSS	MULT	GIVEN	TOTAL
Practical	Student completes practical work of the brief to an acceptable standard set by the teacher.	**2**	**3**		**/6**
Investigation	Student completes the investigation of the brief to an acceptable standard set by the teacher.	**2**	**2**		**/4**
Reasoning and communications	Student responses are accurate and appropriate in presentation of mathematical ideas in different contexts, with clear and logical working out shown.	**4**	**-**		**/4**
Concepts and techniques	Student submitted work selects and applies appropriate mathematical modelling and problem solving techniques to solve practical problems, and demonstrates proficiency in the use of mathematical facts, techniques and formulae.	**4**	**-**		**/4**
	Submission Guidelines				
Timeliness	Student submits the exercises and portfolio task by the set deadline. See scoring guidelines for specific details.	**2**	**-**		**/2**
		FINAL			**/20**

Student Reflection:
How did you go with this week's work?

What was interesting?

What did you find easy?

What do you need to work on?