1. Collecting Data

Chapter Objectives
By the end of this chapter students will be able to:
- Categorise different types of data
- Describe some different data collection methods
- Organise data in a frequency table

1.1 Qualitative and quantitative data

Key words
Qualitative - Related to things that are described by words not by numbers
Quantitative - Related to things that are described by numbers
Discrete - Has only a fixed set of values. For example, the ages a group of people
Continuous - The opposite of discrete. Can take any numerical value. For example, height
Variable - Something that changes. A quantity which can take on different values

Practice - Answers
i.
   a. Discrete
   b. Discrete
   c. Continuous
   d. Discrete
   e. Continuous
   f. Discrete

ii. Possible answers:
Discrete variables include: hair colour, eye colour, age in years, gender, number of siblings
Continuous variables include: height, length of arm, leg etc., exact age

iii.
   a. Qualitative
   b. Quantitative
   c. Qualitative
   d. Quantitative
   e. Qualitative
   f. Quantitative
   g. Qualitative
   h. Quantitative
   i. Quantitative
iii. 

a. Discrete  
b. Continuous  
c. Discrete  
d. Discrete  
e. Discrete  
f. Continuous  
g. Discrete  
h. Discrete  
i. Continuous  

iv. Possible answers:

a. The colour is qualitative, the quantity of petrol that can be held in the tank is quantitative 
b. The type of elephant is qualitative, the number of elephants in the herd is quantitative 
c. The ethnicity of the person is qualitative, the age in years of the person is quantitative  

1.2 Sampling

**Key words**

Survey - A general examination of a situation or subject  
Population - The total number of inhabitants in an area  
Census - A sample that includes every member of a population  
Sample - A small group of things that are taken from a larger group of things and studied so that more can be said about the larger group  

**Practice - Answers**

a. Census  
b. Census  
c. Census  
d. Sample  
e. Sample  
f. Sample  
g. Sample  

1.3 Primary and secondary data

**Key words**

Primary data - Data which we collect ourselves  
Secondary data - Data which we use which was collected by another person or organisation  
Source - The place where secondary data comes from  

**Practice - Answers**

i. 

a. Secondary data. Because you could get the information from the school administrator  
b. Secondary data. Because you could ask the teashop for their financial records  
c. Secondary data. Because many books have been written about tourism in Myanmar  
d. Primary data. Because you need to ask people’s opinions directly  
e. Secondary data. Because there are reports available about poverty in African countries
ii. Possible answers:
   c. The internet or the Myanmar tourist office
   e. United Nations website

iii. | Data  | Advantages                        | Disadvantages                      |
     |------|-----------------------------------|-----------------------------------|
     | Secondary | - Cheap to collect | - Data may be old|
     |         | - Easy to collect        | - The data may be inaccurate      |
     | Primary  | - You know how it was collected | - Takes a long time to collect     |
     |         | - Can choose who to collect data from | - Expensive to collect |

iv. If possible, divide the students into small groups and tell them to search the internet using www.google.com to find sources of information. Discuss the answers in the following lesson. (Please note that Google itself is not a source but is used to find sources on other websites.)

1.4 Methods for collecting primary data

Key words

Questionnaire - A set of written questions designed to collect data on a subject from people
Interview - A set of written questions designed to collect data on a subject from people
Observation - Collecting data by going to watch a situation
Experiment - A method for collecting data which involves doing tests
Practice - Answers

i. Possible answers:

First question:
   a. It is difficult to define 'young' and 'old'
   b. It would be better to have categories of ages such as '10-19', '20-29' etc. because the categories given are too general.

Second question:
   a. Hardly anyone is under 1 metre or over 2 metres
   b. People could either write down their actual height or you could use categories again - '1 to 1.2m'

Third question:
   a. If someone answers 'no' then you do not know their real opinion, only that they are not amazing so the information collected is not useful.
   b. More categories and a more specific question would be better, e.g. 'What is your opinion of the standard of teaching in your school? - Very good, good, fair, poor, very poor'. It would also be could to ask for an explanation of the answer, e.g. 'The teaching is good because...'

ii. Ask students to work in pairs to create their questionnaire. The content should focus on what work they would like to do, where they think they will live, choices of family life, etc. After each group has finished their questionnaire, ask them to swap with another group so that they can give feedback on the quality of the group’s questions. Finally, create a list on the board of the best questions by discussing with the students which questions they like and why.

1.5 Recording data in tables

Key words
Table - A set of data presented in rows and columns. Choosing one value in the table enables another connected value to be read
Tally - A simple way of counting things in groups of five using lines
Frequency - How often something which we are studying occurs
Frequency distribution - A table which presents the frequencies of different events we are studying
Class intervals - The groups which we use to organise continuous data
Think
a. 4 (the students should write 4 in the ‘frequency column’)
b. On Sunday 11 students were born
c. On Monday and Saturday 7 students were born
d. To find this figure the students should complete the ‘frequency’ column and then add all the numbers to make 52

Think
a. 30-39 years
b. 90+
c. 4,088,469 + 4,172,971 = 8,261,440
d. This class interval is different because not many people will be over 90 years old

Practice - Answers
i. a.

<table>
<thead>
<tr>
<th>Job</th>
<th>Tally</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Doctor</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Musician</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Soldier</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Nurse</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Translator</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>24</td>
</tr>
</tbody>
</table>

b. 24
c. Teacher
d. It is much easier to interpret and analyse the data when it is in a table
ii.

a.

<table>
<thead>
<tr>
<th>Age</th>
<th>Tally</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 - 19</td>
<td>HHHH</td>
<td>8</td>
</tr>
<tr>
<td>20 - 29</td>
<td>HHHHH</td>
<td>12</td>
</tr>
<tr>
<td>30 - 39</td>
<td>HHH</td>
<td>8</td>
</tr>
<tr>
<td>40 - 49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 - 59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>40</td>
</tr>
</tbody>
</table>

b. 10 years

c. 8

d. 12
2. Analysing Data

Chapter Objectives
By the end of this chapter students will be able to:

- Calculate the mean, mode and median of discrete and continuous data
- Calculate the range and interquartile range of discrete and continuous data
- Draw a scatter diagram from a table of data
- Describe the relationship between two sets of data by reading a scatter diagram

2.1 Mean, mode and median

Key words

Average - A number which can be used to represent a set of data
Mean - One kind of average. The mean is calculated by adding up all the values and dividing by the total number of values
Mode - One kind of average. The mode is the value which occurs most often in a data set
Median - One kind of average. The median is found by ordering the data from smallest to largest and finding the middle value

Think
The mean of a set of data is the sum of the values divided by the number of values.
The median is the middle value when the data is arranged in order of size.
The mode of a set of data is the value which occurs most often.

Practice - Answers
i.

a. 34
b. \((28 + 29)/2 = 57/2 = 28.5\)
c. 26.4
ii.
   a. There is no mode because each value occurs only once
   b. 3,839,000
   c. 4,263,328

iii.
   a. 6,471,000
   b. twelve million and eighty thousand

iv.
   a. 9,951,200
   b. The answer is that there is no mode because each value occurs only once. Explain this to the students if nobody thinks of it themselves
Think
The set has 12 numbers so, \( n = 12 \)
The total of the set is 36 so, \( \Sigma X = 36 \)
Mean = \( \frac{\Sigma X}{n} = \frac{36}{12} = 3 \)

2.2 Choosing an appropriate average

2.3 The quartiles

Key words

Quartiles - Numbers which divide a set of data into 4 intervals, each containing 25% of the data
Lower quartile - The number which is one quarter or 25% into the data set when it is arranged in numerical order
Upper quartile - The number which is three quarters or 75% into the data set when it is arranged in numerical order
Life expectancy - The number of years a person is predicted (expected) to live based on statistical analysis of a population
<table>
<thead>
<tr>
<th>Quartile</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower quartile</td>
<td>( \left( \frac{n + 1}{2} \right) ) th Value</td>
</tr>
<tr>
<td>Median</td>
<td>( \frac{3(n + 1)}{4} ) th Value</td>
</tr>
<tr>
<td>Upper quartile</td>
<td>( \left( \frac{n + 1}{4} \right) ) th Value</td>
</tr>
</tbody>
</table>
In order the populations are:

1,145,000  1,581,082  3,083,000  3,839,000  4,082,000  5,882,000  10,231,271

a. Lower quartile = \((n + 1)/4\) th value = \(8/4 = 2\text{nd value} = 1,581,082\)
b. Upper quartile = \(3(n + 1)/4\) th value = \(24/4 = 6\text{th value} = 5,882,000\)

### 2.4 The range and interquartile range

**Key words**

Range - The difference between the largest and smallest pieces of a data set
Interquartile range - The difference between the upper quartile and lower quartile of a data set

**Practice - Answers**

i. The lowest value is 63.1 and the highest is 76.8 so the range = 76.8 - 63.1 = 13.7 years
The lower quartile is 71 years and the upper quartile is 75.7 years so the Interquartile range = 4.7 years

ii. The lowest value is 1,145,000 and the highest is 10,231,271 so:
\[
\text{Range} = 10,231,271 - 1,145,000 = 9,086,217
\]
The lower quartile is 1,581,082 and the upper quartile is 5,882,000 so:
\[
\text{Interquartile range} = 5,882,000 - 1,581,082 = 4,300,918
\]
2.5 Averages from frequency distributions

i.

a.

<table>
<thead>
<tr>
<th>Number of goals (x)</th>
<th>Frequency (f)</th>
<th>fx</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

\[ \sum f = 31 \quad \sum fx = 56 \]

b. Using the formula the mean = \( \frac{56}{31} = 1.81 \) goals per game

ii.

<table>
<thead>
<tr>
<th>Number of people (x)</th>
<th>Frequency (f)</th>
<th>fx</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>18</td>
</tr>
</tbody>
</table>

\[ \sum f = 36 \quad \sum fx = 153 \]

The mean = \( \frac{153}{36} = 4.25 \) people per household
2.6 Averages from grouped data

**Practice - Answers**

i. 

a. 

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
<th>Middle value</th>
<th>$fx$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 9</td>
<td>2</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>10 - 19</td>
<td>4</td>
<td>14.5</td>
<td>58</td>
</tr>
<tr>
<td>20 - 29</td>
<td>12</td>
<td>24.5</td>
<td>294</td>
</tr>
<tr>
<td>30 - 39</td>
<td>5</td>
<td>34.5</td>
<td>172.5</td>
</tr>
<tr>
<td>40 - 49</td>
<td>2</td>
<td>44.5</td>
<td>89</td>
</tr>
<tr>
<td><strong>Total (Σf)</strong></td>
<td><strong>25</strong></td>
<td><strong>Total (Σfx)</strong></td>
<td><strong>622.5</strong></td>
</tr>
</tbody>
</table>

b. The mean $= \frac{\Sigma x}{\Sigma f} = \frac{622.5}{25} = 24.9$
ii.

a.

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency (f)</th>
<th>Middle value (x)</th>
<th>fx</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 5</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>6 - 10</td>
<td>9</td>
<td>8</td>
<td>72</td>
</tr>
<tr>
<td>11 - 15</td>
<td>3</td>
<td>13</td>
<td>39</td>
</tr>
<tr>
<td>16 - 20</td>
<td>1</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td><strong>Total (∑f)</strong></td>
<td><strong>15</strong></td>
<td><strong>Total (∑fx)</strong></td>
<td><strong>135</strong></td>
</tr>
</tbody>
</table>

The mean = ∑f x / ∑f = 135 / 15 = 9

b.

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency (f)</th>
<th>Middle value (x)</th>
<th>fx</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 - 19</td>
<td>8</td>
<td>14.5</td>
<td>116</td>
</tr>
<tr>
<td>20 - 29</td>
<td>11</td>
<td>24.5</td>
<td>269.5</td>
</tr>
<tr>
<td>30 - 39</td>
<td>13</td>
<td>34.5</td>
<td>448.5</td>
</tr>
<tr>
<td>40 - 49</td>
<td>9</td>
<td>44.5</td>
<td>400.5</td>
</tr>
<tr>
<td>50 - 59</td>
<td>7</td>
<td>54.5</td>
<td>381.5</td>
</tr>
<tr>
<td><strong>Total (∑f)</strong></td>
<td><strong>48</strong></td>
<td><strong>Total (∑fx)</strong></td>
<td><strong>1616</strong></td>
</tr>
</tbody>
</table>

The mean = ∑f x / ∑f = 1616 / 48 = 33.7

c.

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency (f)</th>
<th>Middle value (x)</th>
<th>fx</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 - 12</td>
<td>1</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>12 - 14</td>
<td>5</td>
<td>13</td>
<td>65</td>
</tr>
<tr>
<td>14 - 16</td>
<td>12</td>
<td>15</td>
<td>180</td>
</tr>
<tr>
<td>16 - 18</td>
<td>3</td>
<td>17</td>
<td>51</td>
</tr>
<tr>
<td>18 - 20</td>
<td>0</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total (∑f)</strong></td>
<td><strong>21</strong></td>
<td><strong>Total (∑fx)</strong></td>
<td><strong>307</strong></td>
</tr>
</tbody>
</table>

The mean = ∑f x / ∑f = 307 / 21 = 14.6

### 2.7 Scatter diagrams

**Key words**

**Scatter diagram** - A graph which is used to present statistical data about two variables. The graph can be used to find relationships between the two variables.

**Correlation** - A measure of the relationship between two sets of data.

**Positive correlation** - If the values in two sets of data increase or decrease at the same time then they have a positive correlation.

**Negative correlation** - If the value of one set of data decreases as the other increases then the two sets of data have a negative correlation.
Practice - Answers

i. The answer is quite easy: More drinks are sold when it is hotter because people are hotter!

ii. Yes, there is a relationship. The longer Chandra drives the less distance is remaining.

iii.

a.

b. The scatter diagram doesn’t show a relationship between the temperature and the amount of rain.
Think

i. There is a positive correlation between the average daily temperature and the number of cold drinks sold, because as the temperature increases the number of cold drinks sold increases. There is a negative correlation between the time spent driving and the distance remaining, because as the time increases the distance remaining decreases.

Practice - Answers

i.

a. A comparison of maximum temperature and number of hours of sunshine

b. There is a positive correlation between the hours of sunshine and the maximum temperature, because as the hours increase the temperature increases.
ii.

a. Check the students’ scatter diagrams. Make sure the students label the axes and give the graph a title.

b. Ask the students whether there is a relationship between the area and population of a country. The correct answer is that there is no relationship.
3. Presenting Data

Chapter Objectives

By the end of this chapter students will be able to:

- Draw pie charts and bar graphs to present discrete data
- Extract information from pie charts and bar graphs to provide information about data
- Draw histograms and cumulative frequency polygons to present continuous data
- Extract information from histograms and cumulative frequency polygons to provide information about data
- Calculate the range and interquartile range of data by reading a cumulative frequency polygon

3.1 Introduction

Key words

Diagram - A picture which is designed to show how something works or how the relationship between the parts works

Pie charts - A way of showing information using different sized sectors of a circle. The sectors look like slices of a pie

Bar graph/bar chart - A diagram which uses horizontal or vertical bars of equal width to represent frequency

Histogram - A type of bar graph which represents grouped continuous data

Cumulative frequency - The number of occurrences of something at or before a given point

Cumulative frequency graph - A graph which shows the cumulative frequency plotted against values of another variable

Think

a. Ask students to make a list. If they can’t think of anything ask them to look around their environment after school. Ask students to explain the diagrams and what was being shown.

b. Discuss students ideas on why we use diagrams to present data. The most obvious answer is that they are easy to look at and understand compared to lists of unorganised data.
3.2 Pie Charts

Practice - Answers

i.

a.

<table>
<thead>
<tr>
<th>Type of vehicle</th>
<th>Number of vehicles</th>
<th>Calculation</th>
<th>Degrees of circle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cars</td>
<td>110</td>
<td>((110/240)*360)</td>
<td>165</td>
</tr>
<tr>
<td>Motorbikes</td>
<td>80</td>
<td>((80/240)*360)</td>
<td>120</td>
</tr>
<tr>
<td>Vans</td>
<td>40</td>
<td>((40/240)*360)</td>
<td>60</td>
</tr>
<tr>
<td>Buses</td>
<td>10</td>
<td>((10/240)*360)</td>
<td>15</td>
</tr>
</tbody>
</table>

b.

![Pie chart of types of vehicles passing the town hall](image1)

ii.

![Pie chart of student grades](image2)

iii. a. Bus  
      b. One quarter  
      c. 6 x 4 = 24  
      d. 2  
      e. 4
3.3 Bar Graphs

Practice - Answers

i.
   a. Possible answer: ‘Number of peas per pod against frequency’.
   b. The modal value is 6 as this is the number of peas in a pod with the highest frequency

ii.

iii.
   a. The data is discrete as animals are counted by whole numbers only.
   b.

<table>
<thead>
<tr>
<th>Number of pets</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

c.

   d. 6 out of 20 households had no pets. This is 30 %. 3 out of 20 households had 3 or more pets. This is 15 %

iv. a. 50       b. 15       c. 36%

Answers continued on the next page.
v.

a.

<table>
<thead>
<tr>
<th>Number of peas in a pod</th>
<th>Tally</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b.

![Graph of Number of peas in a pod against frequency](image)

c. If we compare this graph with the graph on page 23 we can say that fertiliser increases the number of peas for several reasons: the mode is higher, the minimum number of peas in a pod is higher and the highest number of peas in a pod is higher.

### 3.4 Multiple bar graphs

**Key words**

Multiple bar graph - A bar graph which shows two or more sets of data together so that they can be compared

**Think**

- a. Subjects studied by first year students
- b. 64
- c. 74
- d. 38
- e. Arts
- f. 360
i.

a. August  
b. September  
c. October  
d. 35  
e. 40  
f. September  
g. 190

ii.

a. Males

<table>
<thead>
<tr>
<th>Age in years</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Females

<table>
<thead>
<tr>
<th>Age in years</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

b.

c. Possible answers:
The dark columns represent ___________. The age with the highest number of patients for boys was ________. In total there were ________ girls. There were more ________ than ________.

3.5 Histograms
Practice - Answers

i. Possible answer: Heights of people in Verti village

ii. .
iii.
   a. Weight is a continuous measurement as it can take any value: 1, 1.5, 1.55, 1.555 etc.
   b. 
<table>
<thead>
<tr>
<th>Weight (kg) (W)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 ≤ W &lt; 1</td>
<td>2</td>
</tr>
<tr>
<td>1 ≤ W &lt; 2</td>
<td>5</td>
</tr>
<tr>
<td>2 ≤ W &lt; 3</td>
<td>4</td>
</tr>
<tr>
<td>3 ≤ W &lt; 4</td>
<td>2</td>
</tr>
<tr>
<td>4 ≤ W &lt; 5</td>
<td>4</td>
</tr>
</tbody>
</table>

c. 

d. 
   a. 
<table>
<thead>
<tr>
<th>Height (cm) (h)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>110 ≤ h &lt; 115</td>
<td>3</td>
</tr>
<tr>
<td>115 ≤ h &lt; 120</td>
<td>4</td>
</tr>
<tr>
<td>120 ≤ h &lt; 125</td>
<td>5</td>
</tr>
<tr>
<td>125 ≤ h &lt; 130</td>
<td>7</td>
</tr>
<tr>
<td>130 ≤ h &lt; 135</td>
<td>10</td>
</tr>
</tbody>
</table>

3.6 Cumulative frequency
### Practice - Answers

**a.**

<table>
<thead>
<tr>
<th>Time listening to the radio (hours)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 3</td>
<td>3</td>
</tr>
<tr>
<td>0 - 7</td>
<td>8</td>
</tr>
<tr>
<td>0 - 11</td>
<td>16</td>
</tr>
<tr>
<td>0 - 15</td>
<td>19</td>
</tr>
<tr>
<td>0 - 18</td>
<td>20</td>
</tr>
</tbody>
</table>

**b.**

<table>
<thead>
<tr>
<th>Number of students in the class</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 5</td>
<td>8</td>
</tr>
<tr>
<td>0 - 10</td>
<td>15</td>
</tr>
<tr>
<td>0 - 15</td>
<td>24</td>
</tr>
<tr>
<td>0 - 20</td>
<td>31</td>
</tr>
<tr>
<td>0 - 25</td>
<td>40</td>
</tr>
</tbody>
</table>

**c.**

<table>
<thead>
<tr>
<th>Age of mother at birth of baby (years)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 - 20</td>
<td>3</td>
</tr>
<tr>
<td>16 - 25</td>
<td>9</td>
</tr>
<tr>
<td>16 - 30</td>
<td>26</td>
</tr>
<tr>
<td>16 - 35</td>
<td>52</td>
</tr>
<tr>
<td>16 - 40</td>
<td>63</td>
</tr>
<tr>
<td>16 - 45</td>
<td>65</td>
</tr>
</tbody>
</table>

**d.**

<table>
<thead>
<tr>
<th>Daily temperature (°C)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10 ≤ t &lt; 0</td>
<td>12</td>
</tr>
<tr>
<td>-10 ≤ t &lt; 10</td>
<td>98</td>
</tr>
<tr>
<td>-10 ≤ t &lt; 20</td>
<td>283</td>
</tr>
<tr>
<td>-10 ≤ t &lt; 30</td>
<td>362</td>
</tr>
<tr>
<td>-10 ≤ t &lt; 40</td>
<td>365</td>
</tr>
</tbody>
</table>
3.7 Cumulative frequency graphs

**Practice - Answers**

a. 
![Graph: Number of people against time listening to the radio]

b. 
![Graph: Number of students in each class against number of classes]

c. 
![Graph: Age of mother at baby's birth against number of births]

d. 
![Graph: Daily temperature against number of days]
3.8 Spread from cumulative frequency graphs
Practice - Answers

i. 1

a. 

<table>
<thead>
<tr>
<th>Number of particles</th>
<th>Cum. Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 50</td>
<td>10</td>
</tr>
<tr>
<td>0 - 100</td>
<td>26</td>
</tr>
<tr>
<td>0 - 150</td>
<td>39</td>
</tr>
<tr>
<td>0 - 200</td>
<td>50</td>
</tr>
<tr>
<td>0 - 250</td>
<td>57</td>
</tr>
<tr>
<td>0 - 300</td>
<td>60</td>
</tr>
</tbody>
</table>

c. 115
d. 66 and 177

e. 111

2.

a.

<table>
<thead>
<tr>
<th>Age of company employee (years)</th>
<th>Cum. Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 &lt; a ≤ 20</td>
<td>6</td>
</tr>
<tr>
<td>16 &lt; a ≤ 25</td>
<td>15</td>
</tr>
<tr>
<td>16 &lt; a ≤ 30</td>
<td>29</td>
</tr>
<tr>
<td>16 &lt; a ≤ 35</td>
<td>33</td>
</tr>
<tr>
<td>16 &lt; a ≤ 40</td>
<td>35</td>
</tr>
<tr>
<td>16 &lt; a ≤ 45</td>
<td>36</td>
</tr>
</tbody>
</table>

c. 26 years
d. 22 years and 29 years
e. 7 years
ii.
To answer this question students should draw a cumulative frequency table and graph. They can then use the graph to find the answers:

![Cumulative Frequency Graph](image)

- **a.** 6.75° C
- **b.** 0.65° C
- **c.** 86 people
4. Probability

Chapter Objectives

By the end of this chapter students will be able to:

- Describe the probability of an event occurring in words
- Calculate the probability of a single event using a formula
- Calculate the probability of more than one event using a formula
- Draw a sample space to show all possible outcomes of events involving more than one object
- Calculate probabilities by reading information in a sample space
- Calculate probabilities by reading probability trees
- Draw probability trees to show all possible outcomes of two or more independent events
- Calculate probabilities of two or more dependent events using probability trees

4.1 Finding probabilities

Key words

**Probability** - A measure of how likely something is to happen. Usually represented as a number between 0 and 1

**Event** - Something which may or may not happen

**Impossible** - Describes something which definitely will not happen

**Certain** - Describes something which will definitely happen

**Likely** - Describes something which has a high probability (chance) of happening

**Unlikely** - Describes something which has a low probability (chance) of happening

Practice - Answers

i. There are an infinite number of answers to this question. Tell students they can write anything provided they can give a reason for the event being impossible, certain or in between.

ii.

a. certain

b. impossible

c. unlikely
Practice - Answers

i.

a. $1/6$

b. $3/6$

c. $2/6$

d. $5/6$

ii.

a. $3/10$

b. $8/10$

iii.

a. $26/52$

b. $26/52$

c. $13/52$

d. $4/52$

e. $2/52$
iv.

a. 3

b. red

c. There are 2 chances of getting red, whereas there is only 1 blue and 1 yellow chance. Using probability we have \( P(\text{red}) = \frac{2}{4}, P(\text{blue}) = P(\text{yellow}) = \frac{1}{4} \). The probability of getting red is higher so it is better to choose red.

v.

<table>
<thead>
<tr>
<th>Event</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fraction</td>
</tr>
<tr>
<td>A newborn baby is a boy</td>
<td>1/2</td>
</tr>
<tr>
<td>Rolling a dice and getting an even number</td>
<td>3/6</td>
</tr>
<tr>
<td>Spinning the spinner in iv. and getting blue</td>
<td>1/4</td>
</tr>
<tr>
<td>Pulling a red card from a pack of cards</td>
<td>26/52</td>
</tr>
</tbody>
</table>

4.2 More than one event

**Key words**

**Mutually exclusive** - Events which cannot happen at the same time are said to be mutually exclusive

**Sample space** - The set of all possible outcomes of experiments involving more than one object
Think

a. P(green) = 3/10 because there are 10 counters in total and 3 of them are green. The probability of getting green is 3 out of 10 or 3/10.

b. It is not possible to choose a red counter and a green counter at the same time.

c. There are only 3 different colours so if the counter is not yellow then it also has to be either green or red, meaning P(not yellow) = P(red or green).

d. The total probability is equal to 1 and P(not yellow) + P(yellow) includes all possible outcomes, so P(not yellow) + P(yellow) = 1 which is the same as P(not yellow) = 1 - P(yellow).

Practice - Answers

a. There are 52 cards in a pack and there are 4 tens and 4 aces so P(ace or ten) = 8/52

b. There are 52 cards in a pack and there are 26 black cards and 26 red cards

c. P(black or two) = 52/52 = 1 There are 52 cards in a pack and there are 4 aces, 4 tens and 4 nines so P(ace or ten or nine) = 12/52

d. There are 52 cards in a pack and there are 2 black kings and 2 red jacks so P (black king or red jack) = 4/52
Practice - Answers

i.

a. 4
b. \( P(2 \text{ girls}) = 0.25 \)
c. \( P(2 \text{ boys}) = 1/4 \)
d. \( P(1 \text{ girl and 1 boy}) = 2/4 \)
e. Complete the sentence:
A woman is more likely to have 1 girl and 1 boy than 2 boys or 2 girls.
f. \( P(\text{Twin} \text{s are the same sex}) = P(2 \text{ girls}) + P(2 \text{ boys}) = 2/4 \)

ii.

<table>
<thead>
<tr>
<th></th>
<th>+</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

b. The total number of outcomes is 16. The number of outcomes with score 11 is 3 so \( P(11) = 3/16 \)
c. The total number of outcomes is 16. The number of outcomes with score more than 10 is 6 so \( P(11) = 6/16 \)
d. The total number of outcomes is 16. The number of outcomes with a prime number score is 11 so \( P(\text{prime number}) = 11/16 \)
e. The total number of outcomes is 16. The number of outcomes with score which is a multiple of 3 is 6 so \( P(\text{multiple of 3}) = 6/16 \)

4.3 Tree diagrams

Key words
Tree diagrams - A type of diagrams used to show the different outcomes that can happen as a result of a sequence of events.
Practice - Answers

i.

a. 4
b. b. 1/4
c. c. 1/4
d. d. 2/4 = 1/2

ii.  

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>HH</td>
<td>TH</td>
</tr>
<tr>
<td>T</td>
<td>HT</td>
<td>TT</td>
</tr>
</tbody>
</table>

iii.  

iv.

a.  

b. 8
c. 1
d. 1/2 * 1/2 * 1/2 = 1/8
e. (1/2 * 1/2 * 1/2) + (1/2 * 1/2 * 1/2) + (1/2 * 1/2 * 1/2) = 3/8
g. 1/2
h. 1/4
i. 1/8
j. The pattern is that the denominator doubles with each flip of the coin (because it is multiplied by 2). The probability of getting 4 heads in four flips is 1/16.

4.4 Dependent and Independent events
Key words

Dependent event - An event whose outcome depends on the outcome of previous events
Independent event - An event whose outcome does not depend on the outcome of previous events.

Practice - Answers

i.

a. \( \frac{3}{8} \times \frac{3}{8} = \frac{9}{64} \)

b. \( \frac{3}{8} \times \frac{2}{7} = \frac{6}{56} \)

ii.

a. 

\[ \text{First choice} \quad \text{Second choice} \]

\[ \text{green} \quad \frac{3}{10} \quad \text{red} \quad \frac{7}{10} \]

\[ \text{green} \quad \frac{3}{10} \quad \text{red} \quad \frac{7}{10} \]

b. \( \frac{6}{10} \times \frac{5}{9} = \frac{30}{90} \)

iii.

a. \( \frac{2}{7} \times \frac{1}{6} \times 0 = 0 \) (there are only two male cats)

b. \( \frac{5}{7} \times \frac{4}{6} \times \frac{3}{5} = \frac{60}{210} \)

c. \( P(2 \text{ males}) = P(\text{MMF}) + P(\text{FMM}) + P(\text{MFM}) = \left( \frac{2}{7} \times \frac{1}{6} \times \frac{5}{5} \right) + \left( \frac{5}{7} \times \frac{2}{6} \times \frac{1}{5} \right) + \left( \frac{2}{7} \times \frac{5}{6} \times \frac{1}{5} \right) = \frac{10}{210} + \frac{10}{210} + \frac{10}{210} = \frac{30}{210} \)
Glossary of Keywords

The Glossary in the Student’s Book is a list of all mathematical words that appear in the module. They are given in the order that they appear.

The following short activities are added to this guide to help students remember mathematical vocabulary. They can be used in several ways: to test prior knowledge of a topic, as warm-up activities at the beginning of a lesson or to review what has been learnt at the end of a topic.

Activity 1 - Discuss questions in pairs.
Students are given questions to discuss that relate to a topic.
Example questions -
What is an improper fraction?
How do I change from milligrams to tonnes?
How do I find the perimeter of a square?
What is the commutative law?
What is the order of operations?

Activity 2 - True or false.
Students work in pairs to decide if statements about a topic are true or false.
Example for fractions -
The denominator is the top number in a fraction.
The numerator is less than the denominator in an improper fraction.
Equivalent fractions have the same numerator.

Activity 3 - Give an explanation.
Students work in pairs to prepare a short explanation to questions. Ask some students to give their explanation to the class.
Examples -
Explain how to change from a mixed number to an improper fraction.
Explain how to calculate: \((2 + 3) \times (7 - 42)\)
Explain the mistake in this statement.
Explain what a negative number is.

Activity 4 - Brainstorming
Write a topic on the board and ask students what they know about the topic. Write their answers on the board.

Activity 5 - What’s the topic?
Write words linked to a topic on the board and ask students if they can guess the topic.
Assessment

This is assessment covers most of the topics in this module and should give you an idea of how much the students have understood. It is recommended that you give it as a class test, with some time for review and revision beforehand.

Students will need a protractor to answer question 5 and 12 in part 2.

The total mark for each question is given on the right hand side of the page.

**Part 1 - Answers**

Each question in part 1 is worth 1 mark

- a. continuous
- b. secondary
- c. median
- d. mode
- e. scatter diagram
- f. correlation
- g. discrete
- h. probability
- i. certain
- j. independent

**Total for part 1: 10 marks**

**Part 2 - Answers**

1. 
   - a. 40
   - b. 25
   - c. There are 100 people in total so the number of people who said the UK is
     \[100 - (40 + 25 + 5 + 10) = 20\]
   - d. Check the students bar graphs for accuracy 

**6 marks**
2.

<table>
<thead>
<tr>
<th>Word</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impossible</td>
<td>0</td>
</tr>
<tr>
<td>Likely</td>
<td>0.75</td>
</tr>
<tr>
<td>Certain</td>
<td>1</td>
</tr>
<tr>
<td>Unlikely</td>
<td>0.25</td>
</tr>
<tr>
<td>Even chance</td>
<td>0.5</td>
</tr>
</tbody>
</table>

3.

a. 13
b. \((6 + 8)/2 = 7\)
c. 7.66
d. No because no data value occurs more than once

4. Check the graphs for accuracy. They lose a mark if they didn’t give the graph a title.
5.
   a. 1/6
   b. 3/6
   c. 

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>6</td>
<td>8</td>
<td>12</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td>16</td>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
</tr>
</tbody>
</table>

d. There are 36 possible outcomes and 9 outcomes which have a score of 15 or more. So, 
P(15 or more) = 9/36 = 1/4

6.
   a. Check the students diagrams for accuracy. They lose a mark if they didn’t label the
      axes and give the graph a title.
   b. There is no correlation between the two sets of data.
7. 
   a. 

<table>
<thead>
<tr>
<th>Type of school</th>
<th>Number</th>
<th>Angle on Pie Chart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>63</td>
<td>168</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>45</td>
<td>120</td>
</tr>
<tr>
<td>Grammar</td>
<td>18</td>
<td>48</td>
</tr>
<tr>
<td>Others</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>135</td>
<td>360</td>
</tr>
</tbody>
</table>

b. Check the students pie charts for accuracy. They lose a mark if they didn’t give the chart a title.  
   3 marks

8. Check the students bar charts for accuracy. They lose a mark if they didn’t give the chart a title.  
   4 marks
9.

a.

<table>
<thead>
<tr>
<th>Time</th>
<th>Frequency (f)</th>
<th>Middle value (x)</th>
<th>fx</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 &lt; t ≤ 20</td>
<td>6</td>
<td>10</td>
<td>60</td>
</tr>
<tr>
<td>20 &lt; t ≤ 40</td>
<td>18</td>
<td>30</td>
<td>540</td>
</tr>
<tr>
<td>40 &lt; t ≤ 60</td>
<td>30</td>
<td>50</td>
<td>1500</td>
</tr>
<tr>
<td>60 &lt; t ≤ 80</td>
<td>9</td>
<td>70</td>
<td>610</td>
</tr>
<tr>
<td>80 &lt; t ≤ 100</td>
<td>12</td>
<td>90</td>
<td>1080</td>
</tr>
<tr>
<td><strong>Total (Σf)</strong></td>
<td><strong>75</strong></td>
<td></td>
<td><strong>3790</strong></td>
</tr>
</tbody>
</table>

So, mean = Σfx / Σf = 3790 / 75 = 50.5 minutes or 50 minutes and 30 seconds

b. Check the students histograms for accuracy. They lose a mark if they didn’t give the chart a title.

5 marks
10.

a. Check the students polygons for accuracy. They lose a mark if they didn’t give the chart a title.

b. 80 or 80.5

c. The lower quartile is around 8 and the upper quartile is 17.5 so the interquartile range is 9.5. (Remember that the answers to b) and c) are estimates so small differences to the answers here are acceptable.) 

7 marks
11.  
   a. The gaps on the graph should be completed with Wet = 1/4, Dry = 3/4, fails to reach the top on a dry day = 1/5 and fails to reach the top on a wet day = 9/10  
   b. b) 1/4 x 1/10 = 1/40  
   c. P(reaching the top on a random day) = P(reaching the top on a wet day) + P(reaching the top on a dry day) = (1/4 x 1/10) + (3/4 x 4/5) = 1/40 + 12/20  
      = 25/40 = 5/8  

6 marks

12. Students should create the table below and then use it to draw a pie chart. Check the pie charts for accuracy. They lose a mark if they didn’t give the chart a title.

<table>
<thead>
<tr>
<th>Type of school</th>
<th>Number</th>
<th>Angle on Pie Chart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defenders</td>
<td>9</td>
<td>54</td>
</tr>
<tr>
<td>Midfielders</td>
<td>12</td>
<td>72</td>
</tr>
<tr>
<td>Attackers</td>
<td>39</td>
<td>234</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>360</td>
</tr>
</tbody>
</table>

3 marks