

3.4 Graphical methods for presenting data

So far we have looked at

- different types of **data**,
- and how to summarise data in a **frequency table**

We now look at how to display data **graphically**.

Allows us to see the **main features** of a dataset quickly and clearly.

- Is the distribution symmetric or asymmetric?
- Are the data spread out or tightly concentrated?
- Do there appear to be groupings/clusters within the data?

3.4 Graphical methods for presenting data

So far we have looked at

- different types of **data**,
- and how to summarise data in a **frequency table**

We now look at how to display data **graphically**.

Allows us to see the **main features** of a dataset quickly and clearly.

- Is the distribution symmetric or asymmetric?
- Are the data spread out or tightly concentrated?
- Do there appear to be groupings/clusters within the data?

3.4 Graphical methods for presenting data

So far we have looked at

- different types of **data**,
- and how to summarise data in a **frequency table**

We now look at how to display data **graphically**.

Allows us to see the **main features** of a dataset quickly and clearly.

- Is the distribution symmetric or asymmetric?
- Are the data spread out or tightly concentrated?
- Do there appear to be groupings/clusters within the data?

3.4 Graphical methods for presenting data

So far we have looked at

- different types of **data**,
- and how to summarise data in a **frequency table**

We now look at how to display data **graphically**.

Allows us to see the **main features** of a dataset quickly and clearly.

- Is the distribution symmetric or asymmetric?
- Are the data spread out or tightly concentrated?
- Do there appear to be groupings/clusters within the data?

3.4 Graphical methods for presenting data

So far we have looked at

- different types of **data**,
- and how to summarise data in a **frequency table**

We now look at how to display data **graphically**.

Allows us to see the **main features** of a dataset quickly and clearly.

- Is the distribution symmetric or asymmetric?
- Are the data spread out or tightly concentrated?
- Do there appear to be groupings/clusters within the data?

3.4 Graphical methods for presenting data

So far we have looked at

- different types of **data**,
- and how to summarise data in a **frequency table**

We now look at how to display data **graphically**.

Allows us to see the **main features** of a dataset quickly and clearly.

- Is the distribution symmetric or asymmetric?
- Are the data spread out or tightly concentrated?
- Do there appear to be groupings/clusters within the data?

3.4 Graphical methods for presenting data

So far we have looked at

- different types of **data**,
- and how to summarise data in a **frequency table**

We now look at how to display data **graphically**.

Allows us to see the **main features** of a dataset quickly and clearly.

- Is the distribution symmetric or asymmetric?
- Are the data spread out or tightly concentrated?
- Do there appear to be groupings/clusters within the data?

3.4 Graphical methods for presenting data

So far we have looked at

- different types of **data**,
- and how to summarise data in a **frequency table**

We now look at how to display data **graphically**.

Allows us to see the **main features** of a dataset quickly and clearly.

- Is the distribution symmetric or asymmetric?
- Are the data spread out or tightly concentrated?
- Do there appear to be groupings/clusters within the data?

3.4.1 Stem and leaf plots

Stem and leaf plots are a quick and easy way of representing data graphically.

They can be used with both **discrete** and **continuous** data.

A bit like constructing a frequency table – but carries visual information about the shape of the distribution.

Best seen via an example!

3.4.1 Stem and leaf plots

Stem and leaf plots are a quick and easy way of representing data graphically.

They can be used with both **discrete** and **continuous** data.

A bit like constructing a frequency table – but carries visual information about the shape of the distribution.

Best seen via an example!

3.4.1 Stem and leaf plots

Stem and leaf plots are a quick and easy way of representing data graphically.

They can be used with both **discrete** and **continuous** data.

A bit like constructing a frequency table – but carries visual information about the shape of the distribution.

Best seen via an example!

3.4.1 Stem and leaf plots

Stem and leaf plots are a quick and easy way of representing data graphically.

They can be used with both **discrete** and **continuous** data.

A bit like constructing a frequency table – but carries visual information about the shape of the distribution.

Best seen via an example!

Example: Percentage returns on a share

The following numbers show the percentage returns on an ordinary share for 23 consecutive months:

0.2	-2.1	1.0	0.1	-0.5	2.4	-2.3	1.5
1.2	-0.6	2.4	-1.2	1.7	-1.3	-1.2	0.9
0.5	0.1	-0.1	0.3	-0.4	0.5	0.9	

- The largest value is 2.4 and the smallest -2.3, and we have lots of decimal values in between
- It seems sensible to have a **stem unit** of 1 and a **leaf unit** of 0.1

Example: Percentage returns on a share

A stem and leaf diagram for this set of returns might look like:

-2	3	1							
-1	3	2	2						
-0	6	5	4	1					
0	1	1	2	3	5	5	9	9	
1	0	2	5	7					
2	4	4							

Stem **Leaf**

$n = 23$, stem unit = 1, leaf unit = 0.1.

Example: Percentage returns on a share

A stem and leaf diagram for this set of returns might look like:

-2		3	1						
-1		3	2	2					
-0		6	5	4	1				
0		1	1	2	3	5	5	9	9
1		0	2	5	7				
2		4	4						

Stem **Leaf**

$n = 23$, stem unit = 1, leaf unit = 0.1.

Example: Production line data

If there is more than one significant figure in the data, the extra digits are **cut**, not **rounded** – e.g. 2.97 would be “cut” to 2.9.

Consider the following data on lengths of items on a production line (in cm):

2.97	3.81	2.54	2.01	3.49
3.09	1.99	2.64	2.31	2.22

The stem and leaf plot for this is as follows:

1		9		
2		0	2	3
2		5	6	9
3		0	4	
3		8		

$n = 10$, stem unit = 1 cm, leaf unit = 0.1 cm.

Example: Production line data

If there is more than one significant figure in the data, the extra digits are **cut**, not **rounded** – e.g. 2.97 would be “cut” to 2.9.

Consider the following data on lengths of items on a production line (in cm):

2.97	3.81	2.54	2.01	3.49
3.09	1.99	2.64	2.31	2.22

The stem and leaf plot for this is as follows:

1		9		
2		0	2	3
2		5	6	9
3		0	4	
3		8		

$n = 10$, stem unit = 1 cm, leaf unit = 0.1 cm.

Example: Production line data

If there is more than one significant figure in the data, the extra digits are **cut**, not **rounded** – e.g. 2.97 would be “cut” to 2.9.

Consider the following data on lengths of items on a production line (in cm):

2.97	3.81	2.54	2.01	3.49
3.09	1.99	2.64	2.31	2.22

The stem and leaf plot for this is as follows:

1		9		
2		0	2	3
2		5	6	9
3		0	4	
3		8		

$n = 10$, stem unit = 1 cm, leaf unit = 0.1 cm.

Example: Production line data

Why do you think we **cut** the extra digits?

So values such as 2.97 do not end up in the wrong row!

Example: Production line data

Why do you think we **cut** the extra digits?

So values such as 2.97 do not end up in the wrong row!

Example 3.3

The observations in the table below are the recorded time it takes to get through to an operator at a telephone call centre (in seconds).

54	56	50	67	55	38	49	45	39	50
45	51	47	53	29	42	44	61	51	50
30	39	65	54	44	54	72	65	58	62

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

Stem **Leaf**

$n = 30$, stem unit = 10, leaf unit = 1

Example 3.3

The observations in the table below are the recorded time it takes to get through to an operator at a telephone call centre (in seconds).

54	56	50	67	55	38	49	45	39	50
45	51	47	53	29	42	44	61	51	50
30	39	65	54	44	54	72	65	58	62

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

Stem **Leaf**

$n = 30$, stem unit = 10, leaf unit = 1

Example 3.3

The observations in the table below are the recorded time it takes to get through to an operator at a telephone call centre (in seconds).

54	56	50	67	55	38	49	45	39	50
45	51	47	53	29	42	44	61	51	50
30	39	65	54	44	54	72	65	58	62

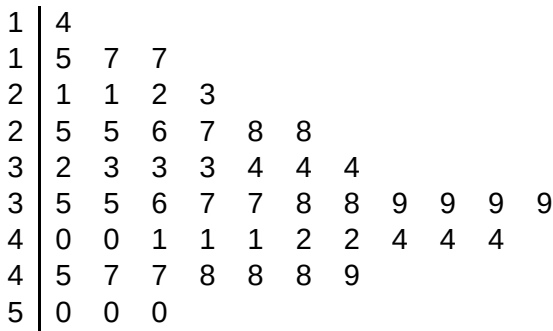
2		9											
3		0	8	9	9								
4		2	4	4	5	5	7	9					
5		0	0	0	1	1	3	4	4	4	5	6	8
6		1	2	5	5	7							
7		2											

Stem Leaf

$n = 30$, stem unit = 10, leaf unit = 1

Example 3.4

The stem and leaf plot below represents the marks on a test for 52 students. Comment on the distribution of these marks.



$n = 52$, stem unit = 10, leaf unit = 1.

Solution to Example 3.4

- **Asymmetric distribution** of marks – in fact, marks are **negatively skewed**
- Majority of students achieve marks in the 30s/40s
- The **modal class** is 35–39
- Large **spread** of marks

Solution to Example 3.4

- **Asymmetric distribution** of marks – in fact, marks are **negatively skewed**
- Majority of students achieve marks in the 30s/40s
- The **modal class** is 35–39
- Large **spread** of marks

Solution to Example 3.4

- **Asymmetric distribution** of marks – in fact, marks are **negatively skewed**
- Majority of students achieve marks in the 30s/40s
- The **modal class** is 35–39
- Large **spread** of marks

Solution to Example 3.4

- **Asymmetric distribution** of marks – in fact, marks are **negatively skewed**
- Majority of students achieve marks in the 30s/40s
- The **modal class** is 35–39
- Large **spread** of marks

Solution to Example 3.4

- **Asymmetric distribution** of marks – in fact, marks are **negatively skewed**
- Majority of students achieve marks in the 30s/40s
- The **modal class** is 35–39
- Large **spread** of marks

3.4.2 Bar charts

Bar charts are a commonly-used and clear way of presenting categorical data or ungrouped discrete frequency observations.

See **section 2.3.2** of the **summer revision booklet**.

Recall the example on students' modes of transport:

Student	Mode	Student	Mode	Student	Mode
1	Car	11	Walk	21	Walk
2	Walk	12	Walk	22	Metro
3	Car	13	Metro	23	Car
4	Walk	14	Bus	24	Car
5	Bus	15	Train	25	Car
6	Metro	16	Bike	26	Bus
7	Car	17	Bus	27	Car
8	Bike	18	Bike	28	Walk
9	Walk	19	Bike	29	Car
10	Car	20	Metro	30	Car

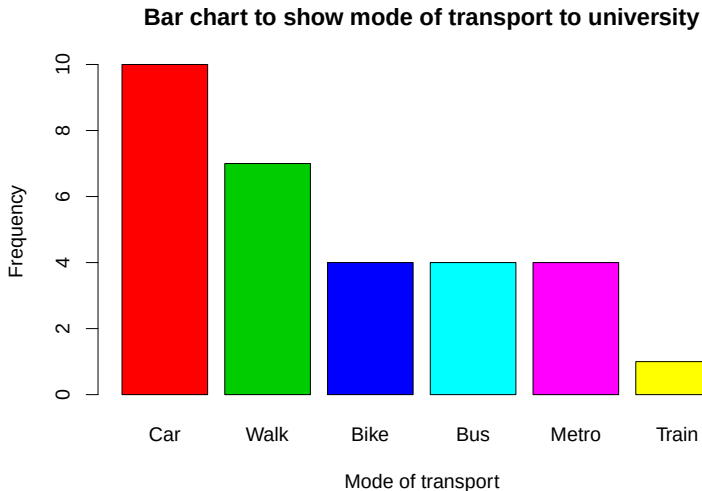
3.4.2 Bar charts

Bar charts are a commonly-used and clear way of presenting categorical data or ungrouped discrete frequency observations. See **section 2.3.2** of the **summer revision booklet**.

Recall the example on students' modes of transport:

Student	Mode	Student	Mode	Student	Mode
1	Car	11	Walk	21	Walk
2	Walk	12	Walk	22	Metro
3	Car	13	Metro	23	Car
4	Walk	14	Bus	24	Car
5	Bus	15	Train	25	Car
6	Metro	16	Bike	26	Bus
7	Car	17	Bus	27	Car
8	Bike	18	Bike	28	Walk
9	Walk	19	Bike	29	Car
10	Car	20	Metro	30	Car

3.4.2 Bar charts



3.4.3 Pie charts

These charts are commonly used within industry to communicate simple ideas, for example market share.

They are used to show the proportions of a whole.

They are best used when there are only a handful of categories to display.

See [Section 2.3.3](#) of the [summer revision booklet](#) for more details.