

Creating graphs in Excel

Version: 1.0



Learning intentions

We will be learning about **creating graphs in Excel**, specifically,

- how creating graphs fits into the analysis steps
- process of creating **bar charts** and **histograms**
- process of creating **line graphs** and **scatterplots**

Background

During the identifying patterns stage of the analysis steps you look for **relationships** and **patterns** in the data.

This can be done through plotting graphs or using calculations designed to measure relationships.

In this lesson, we will be looking at **how to create graphs in Excel** that you can be used to look for relationships and patterns.



Data
understanding

Data cleansing

Data manipulation

Identifying
patterns

Extracting insights

Why identifying patterns is important?



You can **test ideas for relationships** or patterns in the datasets



Identify trends in a dataset



Spot unusual data items that might **need more investigation**

Data
understanding

Data cleansing

Data manipulation

Identifying
patterns

Extracting insights

Graphs in Excel

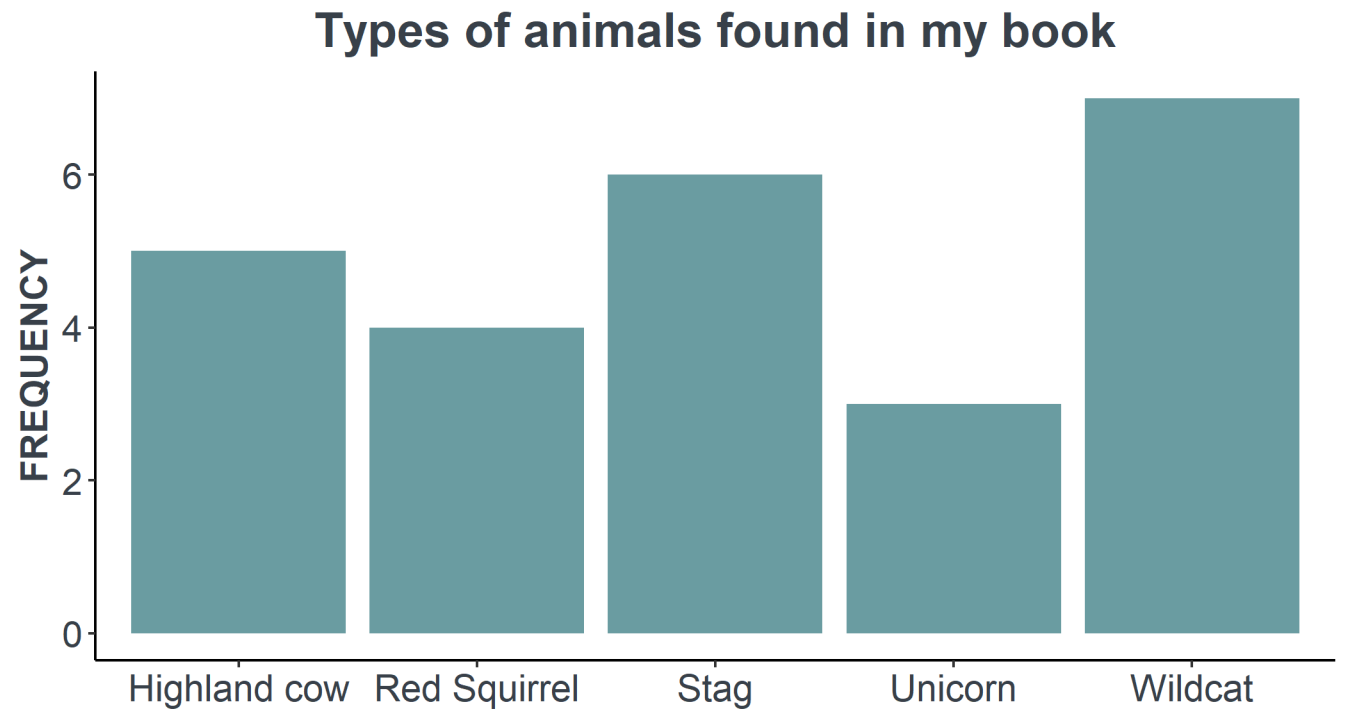
In this lesson we are going to look at **how to create** some of the common graphs used in data science. These are:

- **Bar charts**
- **Histograms**
- **Scatterplots**
- **Line graphs**



When should you use a bar chart?

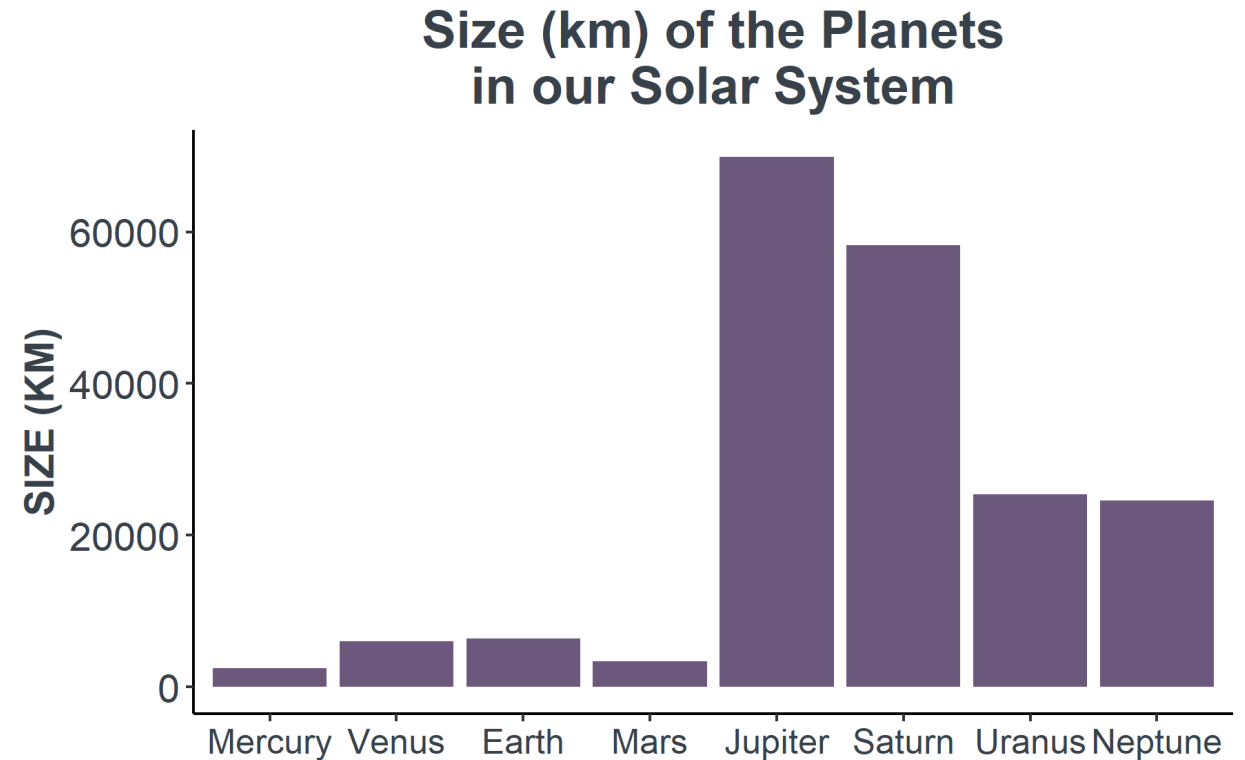
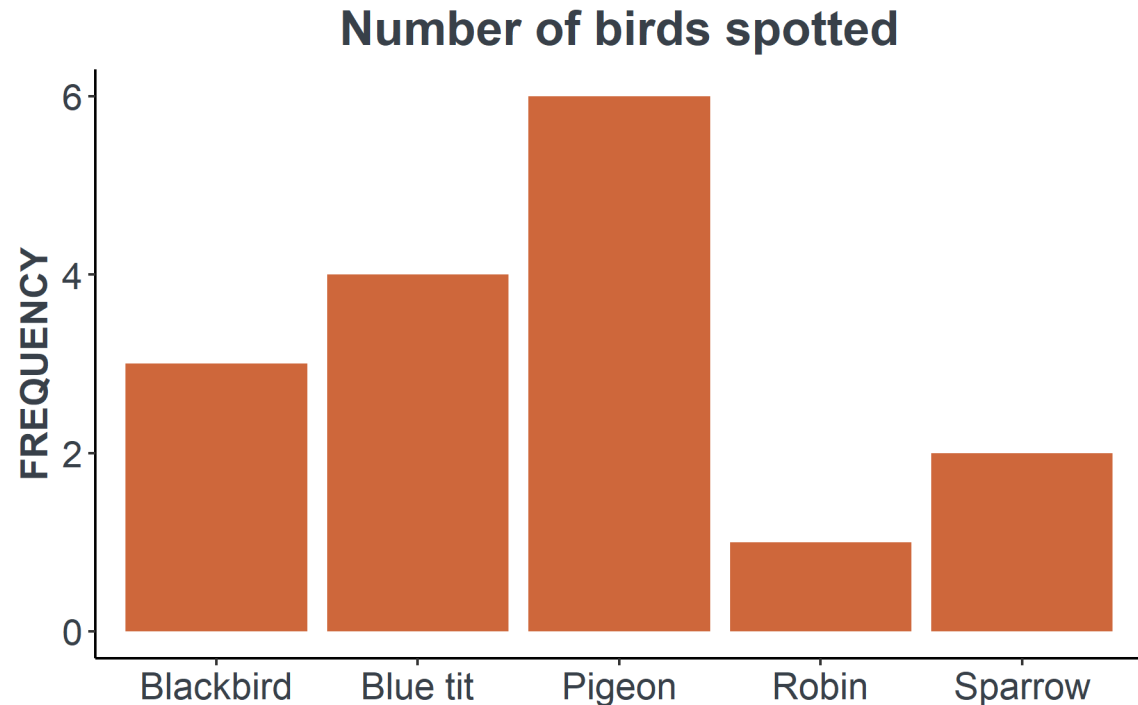
Bar charts should be used if you want to **compare numeric values across different categorical data**.



Show me...



These plots compare a number (frequency or size) against a categorised data item.



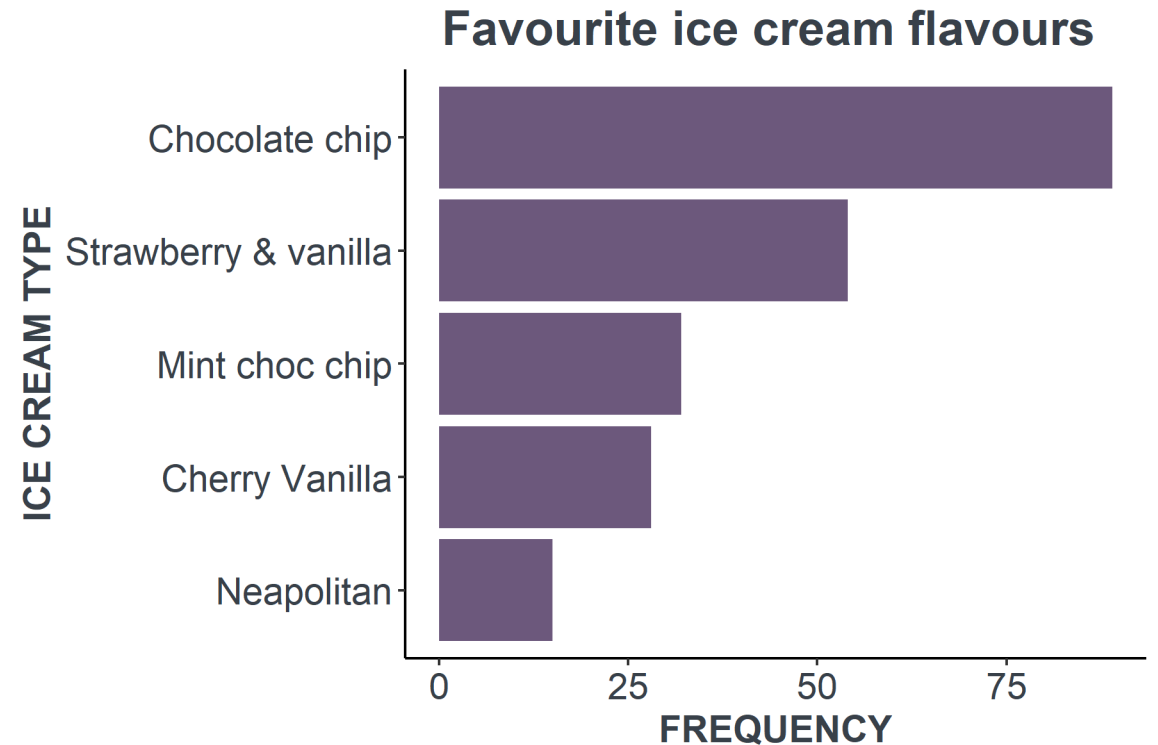
Bar chart best practice

Bars normally show the counts or sizes of categorical data.

Since there is no connection between the bars, they are **shown not touching**.

A **horizontal bar graph** is often a good option when there **are many categories**, or the category labels are long.

It is also possible to reorder the bars, which makes it easier to see the smallest and largest categories.



How to create bar charts

We are now going to look at **how to create bar charts in Excel**.

We are going use a dataset that shows the amount of caffeine in drinks.



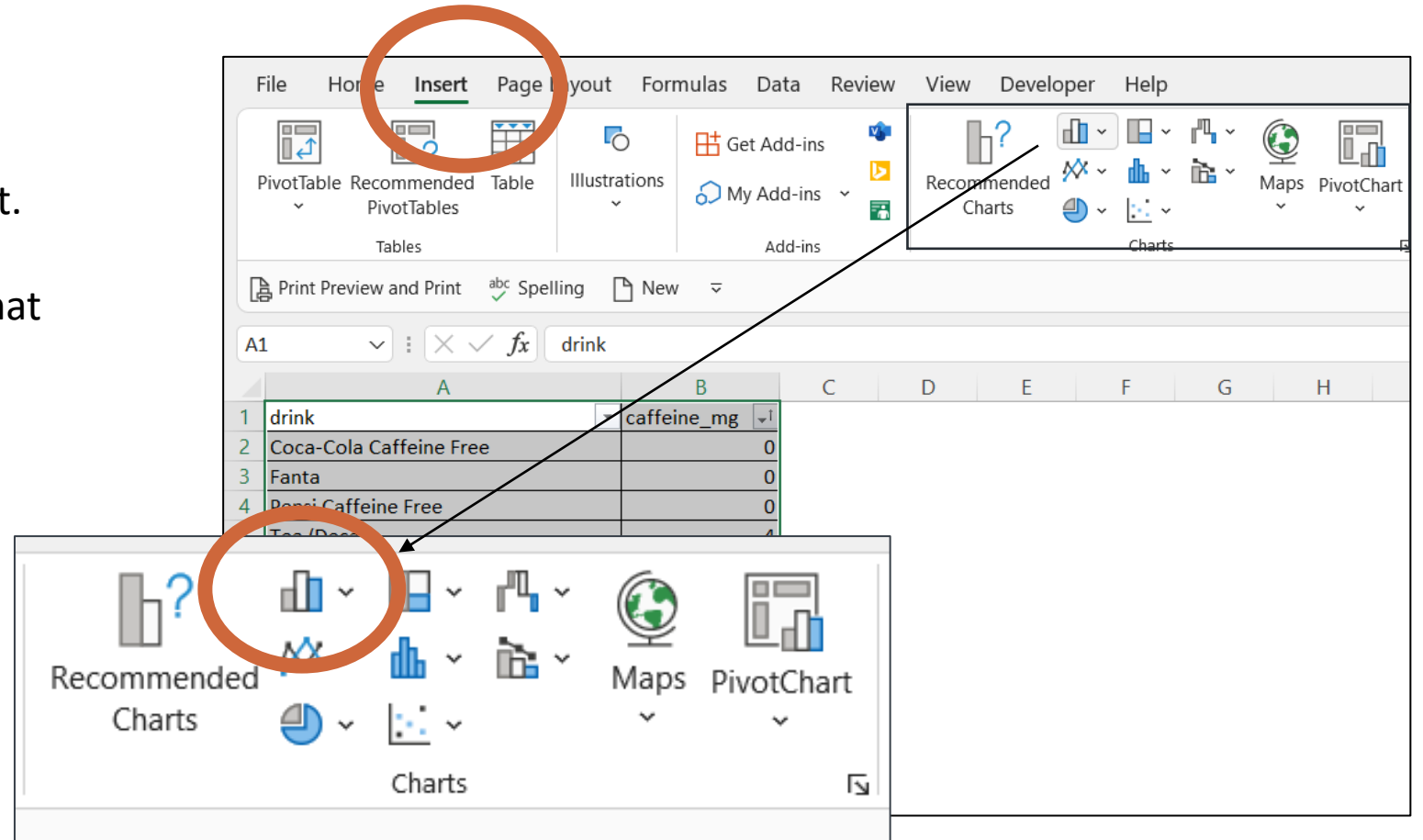
drink	caffeine_mg
Caffe Nero Coffee	160
Chai Tea	50
Coca-Cola Caffeine Free	0
Coca-Cola Classic	34
Costa Coffee	277
Diet Coke	46
Dr Pepper	42
Fanta	0
Irn-bru	30
Lipton Iced Tea	25
McDonalds (McCafe) Mocha	167

Insert of bar chart

Step 1.

Highlight the dataset you are going to plot.

Then in the **Insert** tab, click on the **icon** that looks like a bar chart.

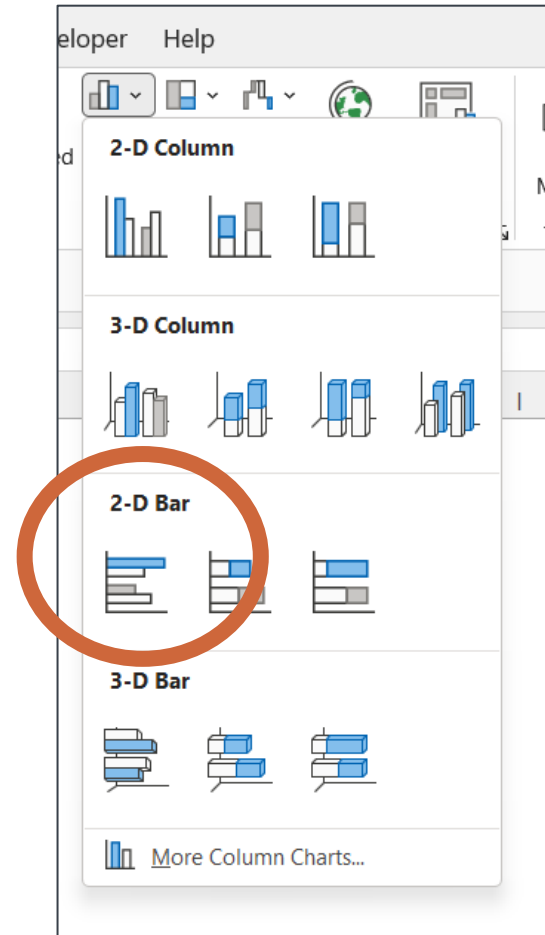


Select the type of bar chart

Step 2.

Select the type of bar chart you would like from the drop down list.

In this case we will use a horizontal bar chart.

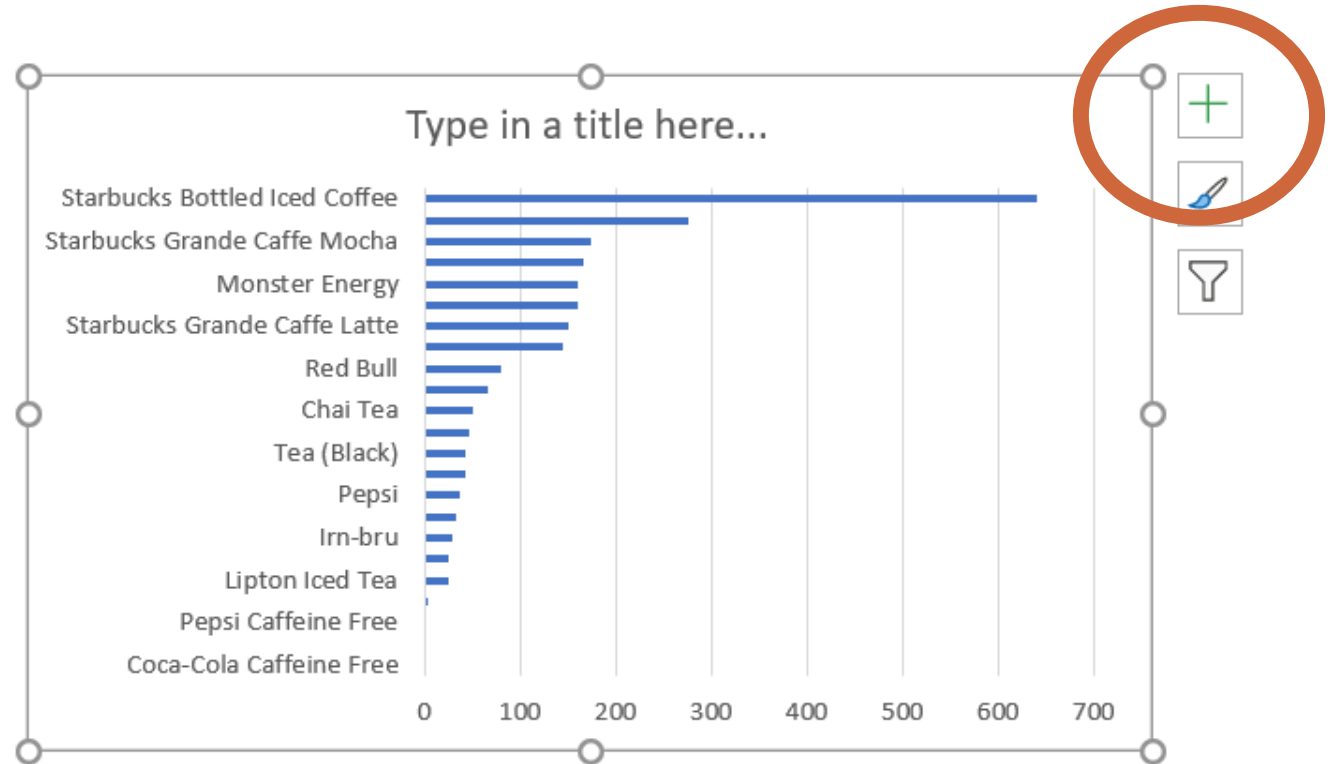


Amend your graph

Step 3.

Click on the **+** symbol next to the graph to add elements to graph such as Axis Titles.

Then right click on any element you would like to change (e.g. the Title, Axis title)

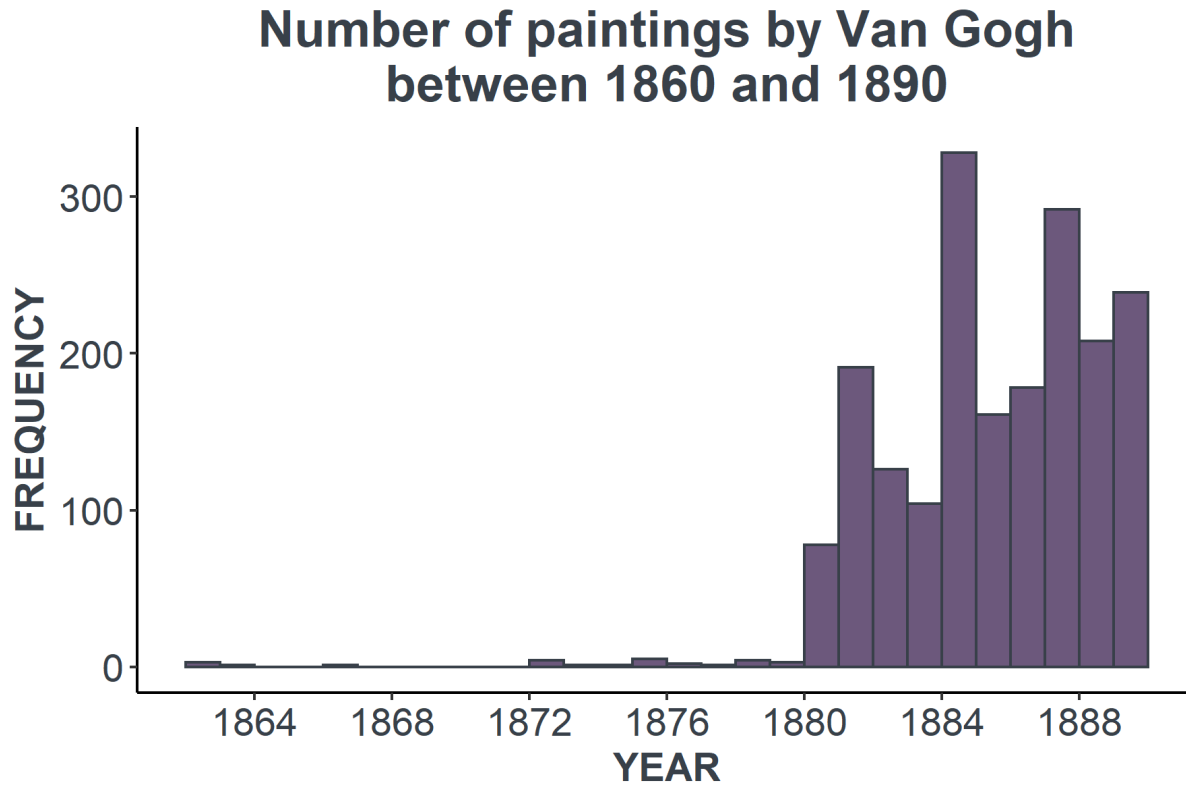


Next steps

Complete **questions 1 to 8**
in **section 1** of the
'Creating graphs in Excel' workbook.

When should you use a histogram chart?

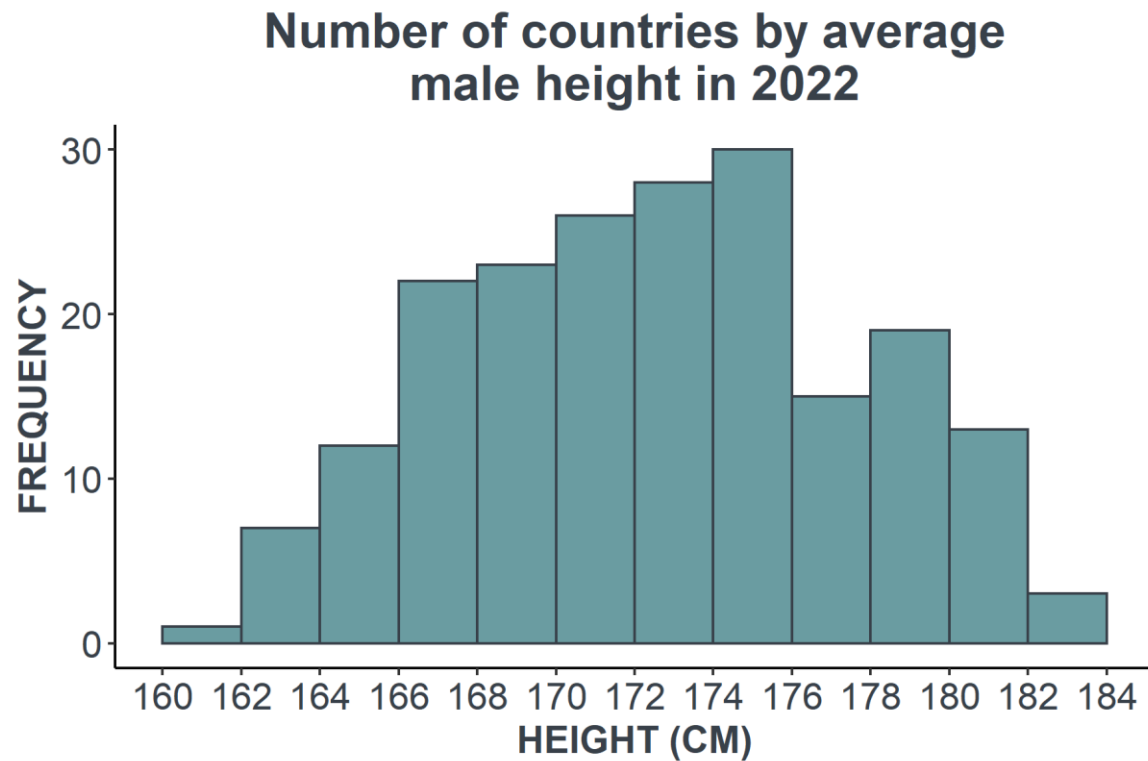
Histograms should be used if you want to **compare numeric** data. The numeric data on the x-axis should be grouped into bins.



Show me...



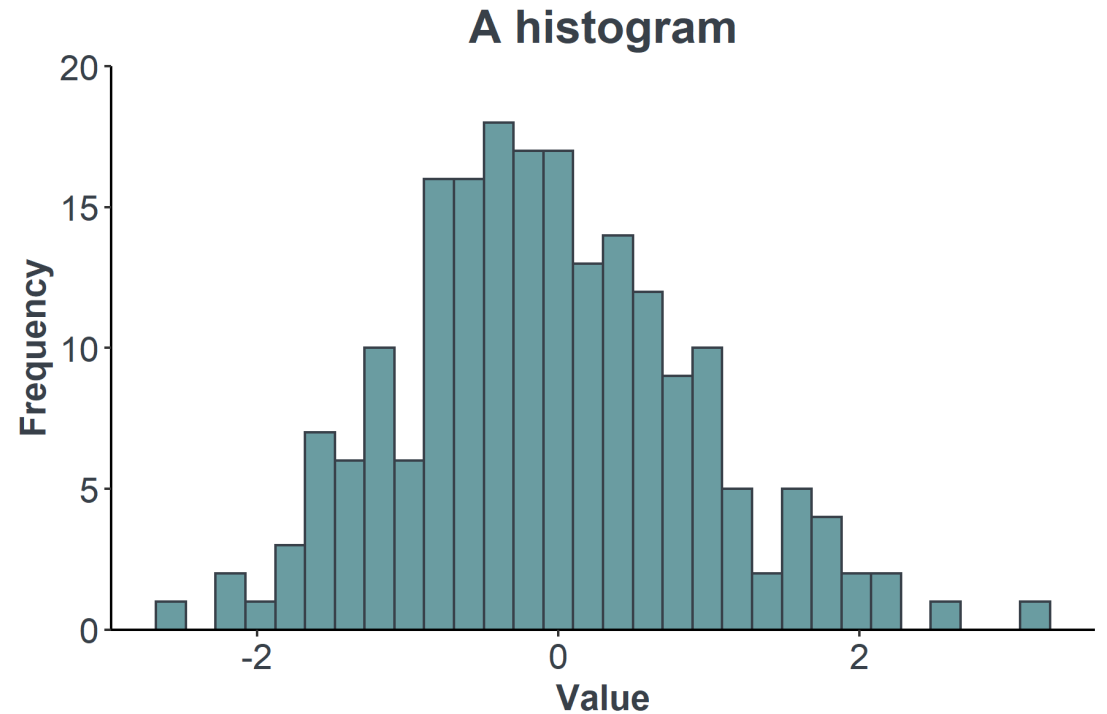
This plot shows the average male height (cm) by country.
Each bar covers a 2cm range (e.g. 164 -166cm)



Histogram best practice

The bars on a histogram should always be **displayed as touching**, since the variable shown on the x-axis is continuous.

If there are gaps between the bars, then that implies missing data in that range.



How to create histogram

We are now going to look at **how to create histograms in Excel**.

We are going use a dataset that shows the average female height by country.



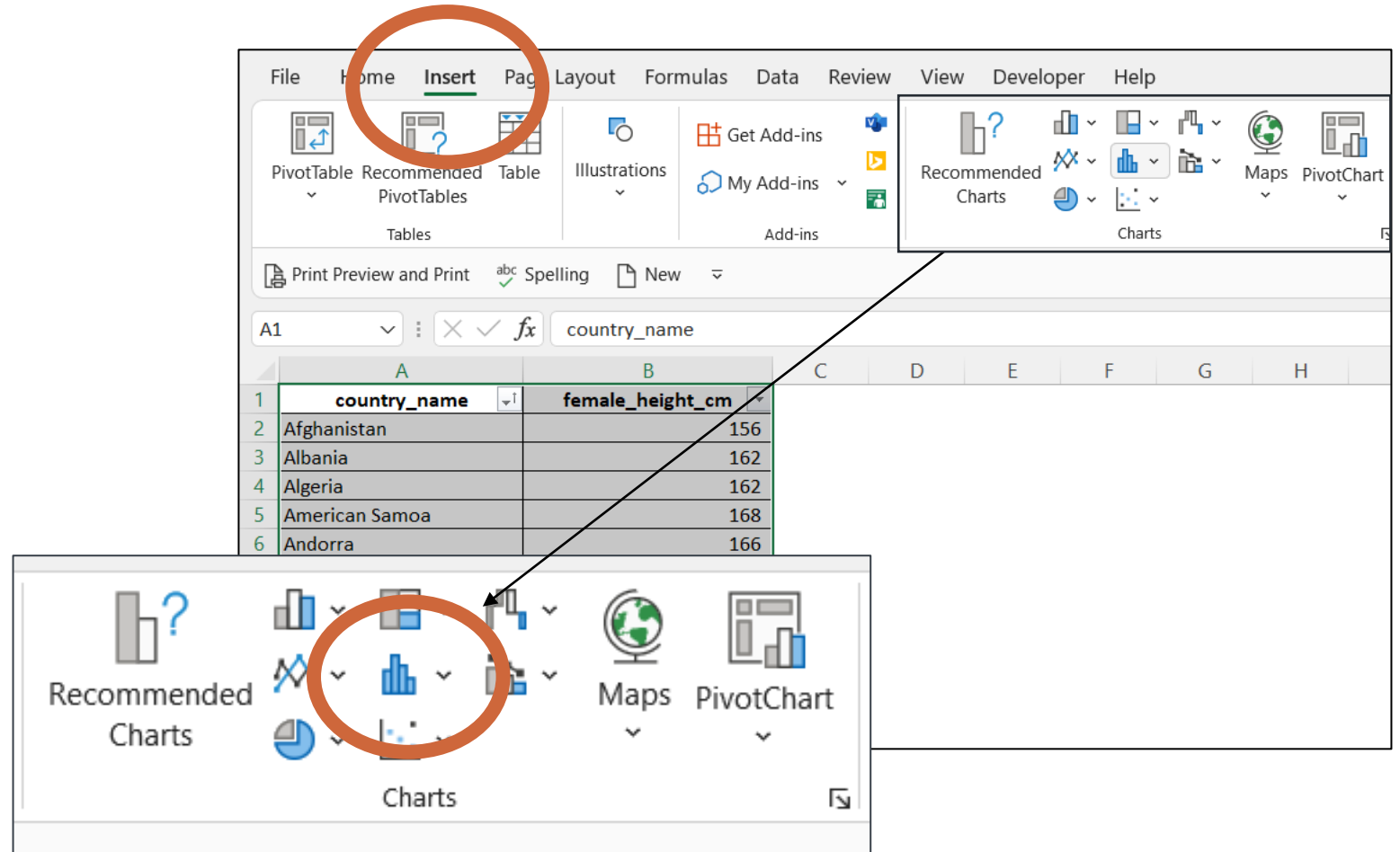
country_name	female_height_cm
Afghanistan	156
Albania	162
Algeria	162
American Samoa	168
Andorra	166
Angola	158
Antigua and Barbuda	166
Argentina	161
Armenia	160
Australia	165
Austria	167
Azerbaijan	161
Bahamas	163
Bahrain	158
Bangladesh	152
Barbados	166
Belarus	167

Insert a histogram

Step 1.

Highlight the dataset you are going to plot.

Then in the **Insert** tab, click on the **icon** that looks like a histogram chart.

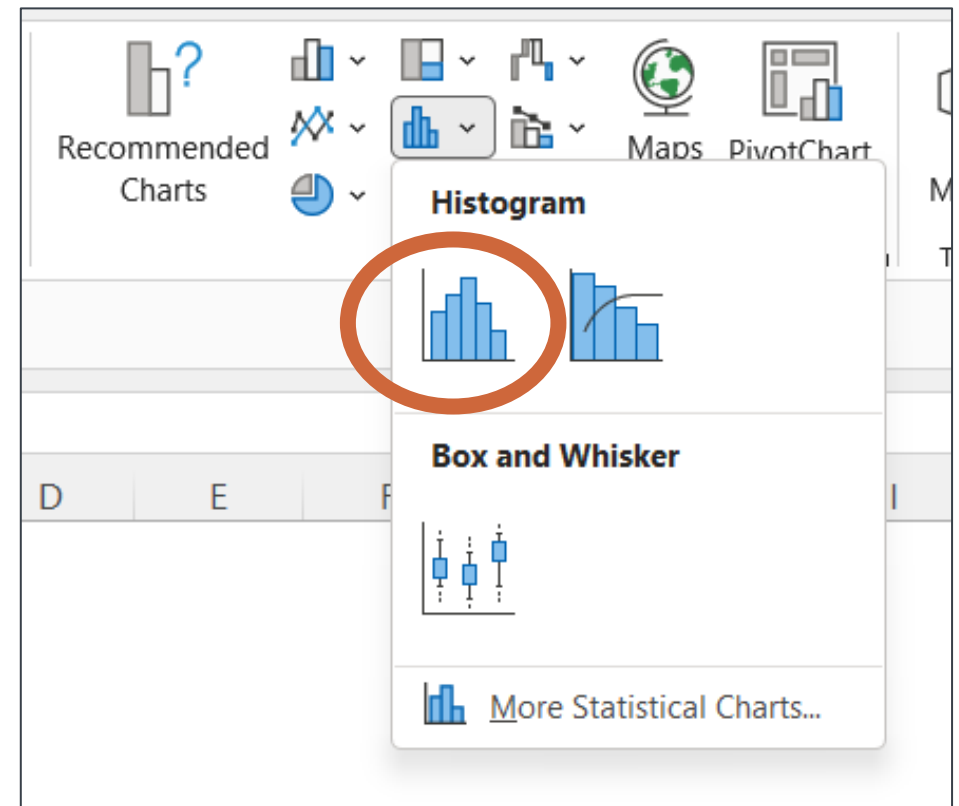


Select the appropriate histogram

Step 2.

Click on the **standard histogram icon** in the drop down list.

The graph will now appear in the workbook.



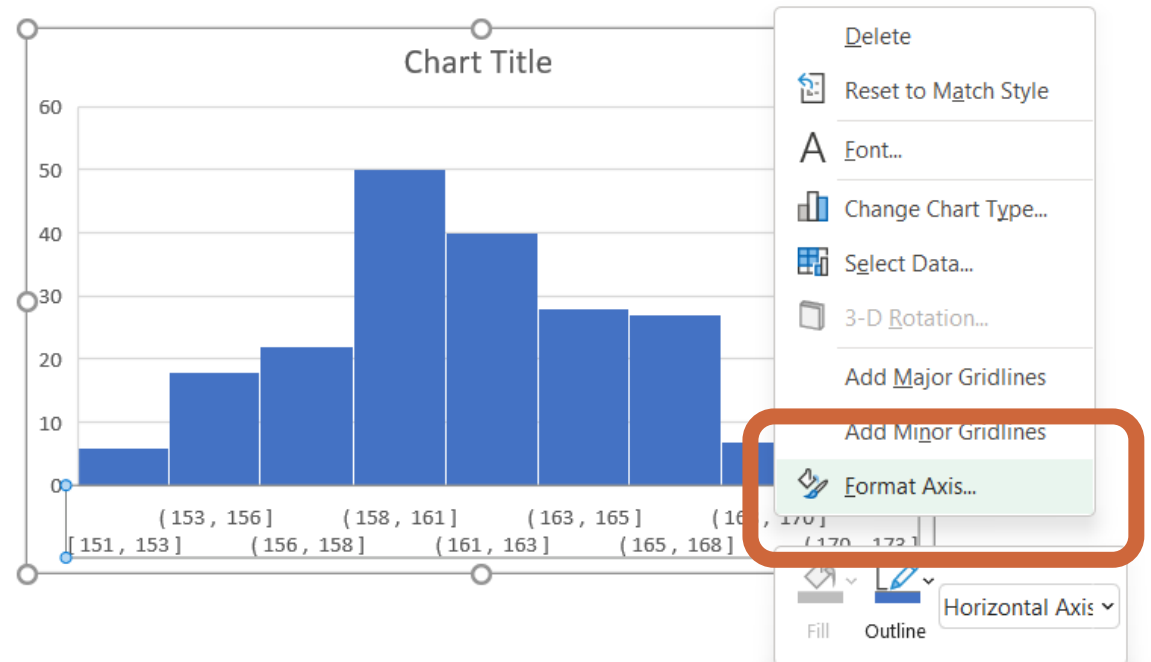
Amend the size of the bins

Step 3.

Excel will automatically decide how to group your data into bins.

Often the size of the bins will need to be amended.

Right click on the values in the x-axis, then select **Format Axis**.



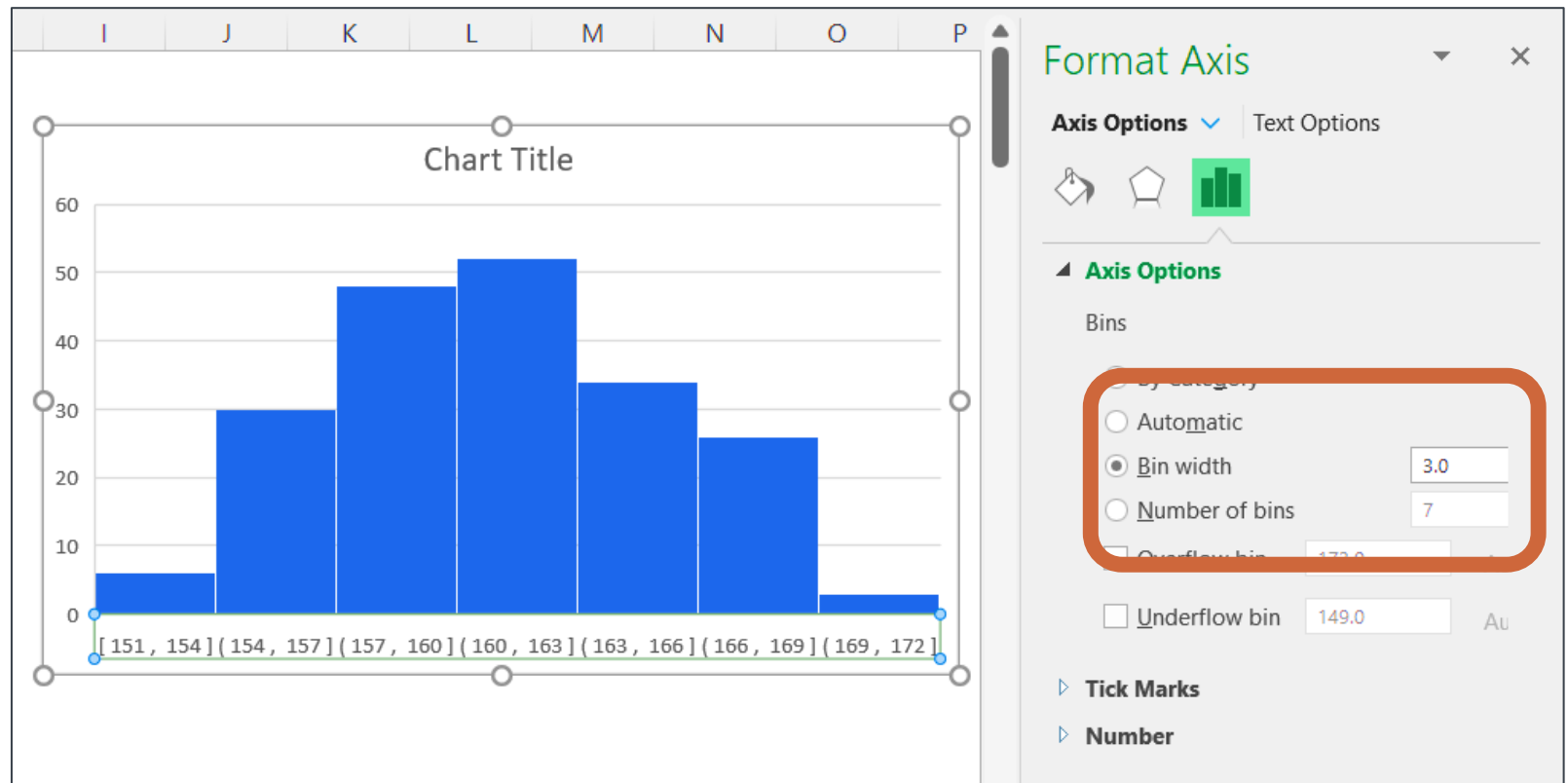
Amend the size of the bins

Step 4.

The Format Axis editor will now be open.

You can change the size of the bins either based on the **width** or the **number of bins**.

In this case the size of the bin has been changed to 3cm (e.g. 157 – 160cm)



Histogram vs. bar chart

Histograms and bar charts can look similar, but there are some key differences that can help you decide which one to use.

	Bar chart	Histogram
Show categorical variables	✓	✗
Reorder bars	✓	✗
Touching bars	✗	✓

Your turn...



This dataset shows the names given to baby girls born in Scotland in 2020.

If you wanted to plot a graph of this data, should you use a **histogram** or a **bar chart**?

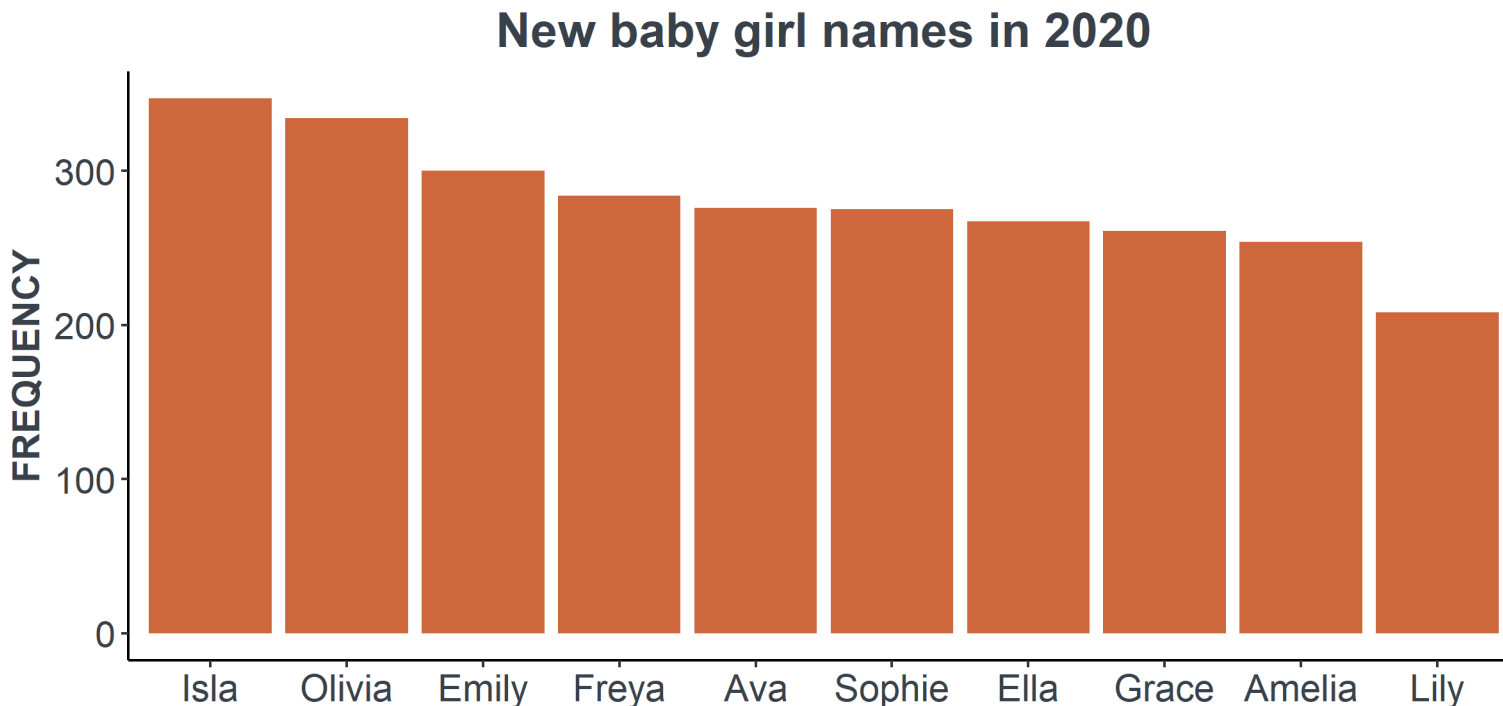


name	number_babies
Isla	347
Olivia	334
Emily	300
Freya	284
Ava	276
Sophie	275
Ella	267
Grace	261
Amelia	254
Lily	208

Your turn...



The graph needs to plot numeric data against different categorical values , so it needs to be a **bar chart**.



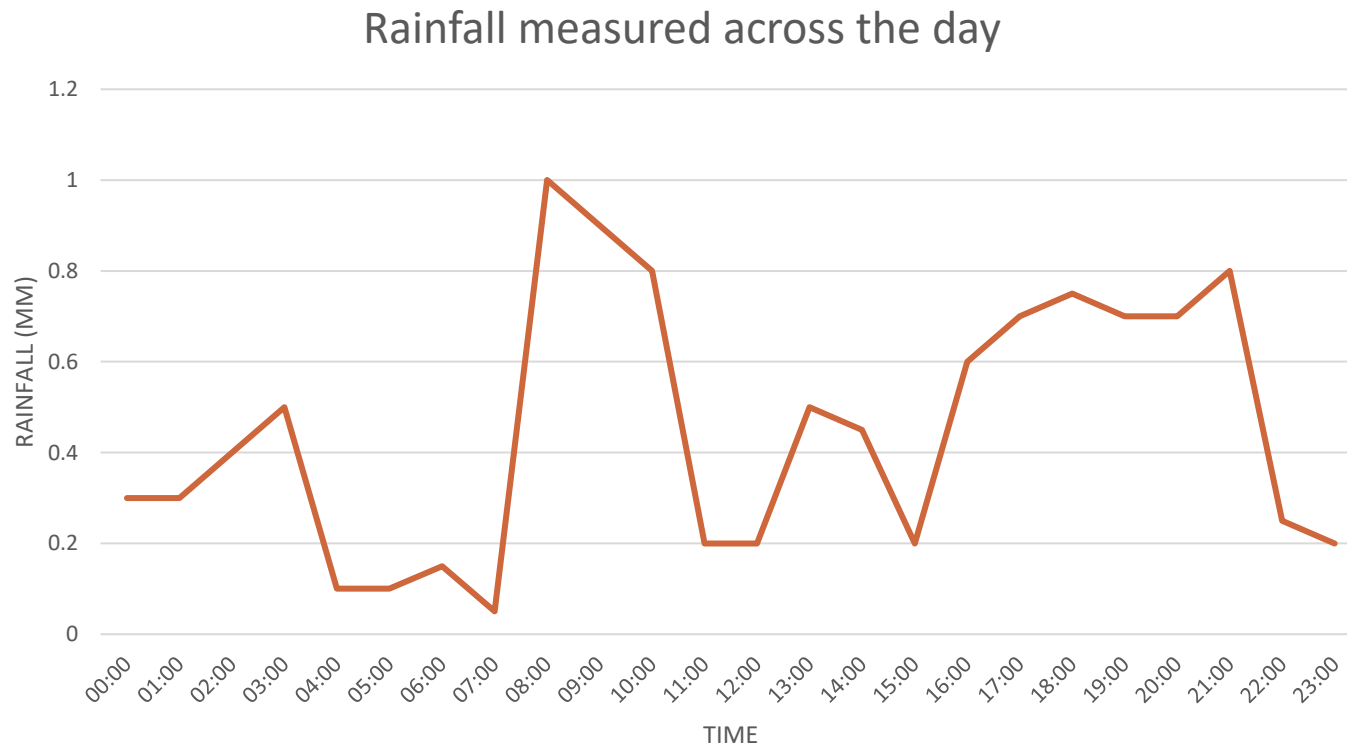
name	number_babies
Isla	347
Olivia	334
Emily	300
Freya	284
Ava	276
Sophie	275
Ella	267
Grace	261
Amelia	254
Lily	208

Next steps

Complete **questions 1 to 5**
in **section 2** of the
'Creating graphs in Excel' workbook.

When should you use a line graphs?

Line graphs should be used to show the **change of a numerical value** as another quantity varies.

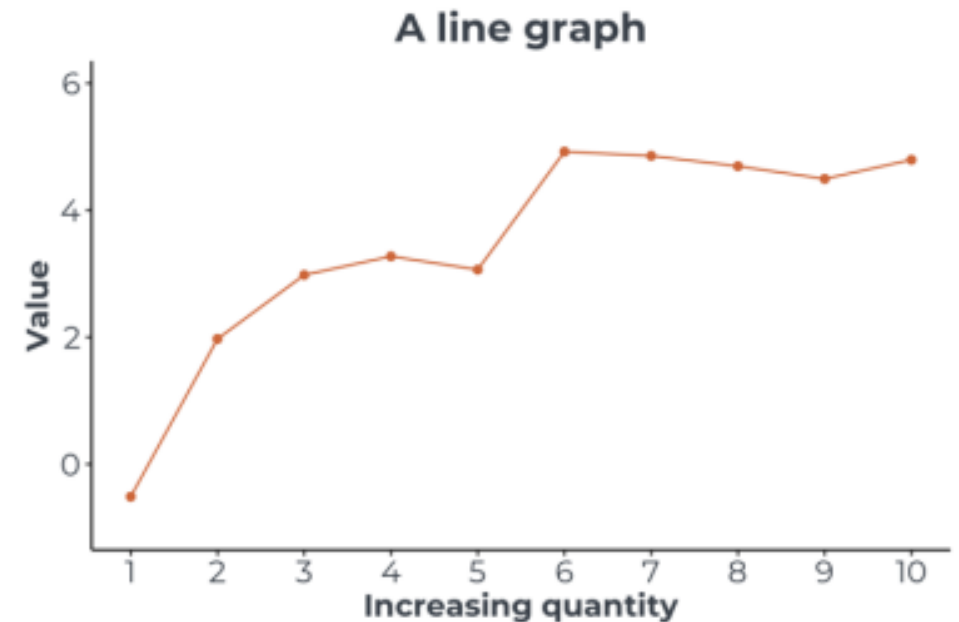


When should you use line graphs?

Line graphs compare two numeric values.

The x-axis is often **time** but could be another varying quantity such as **temperature or distance**.

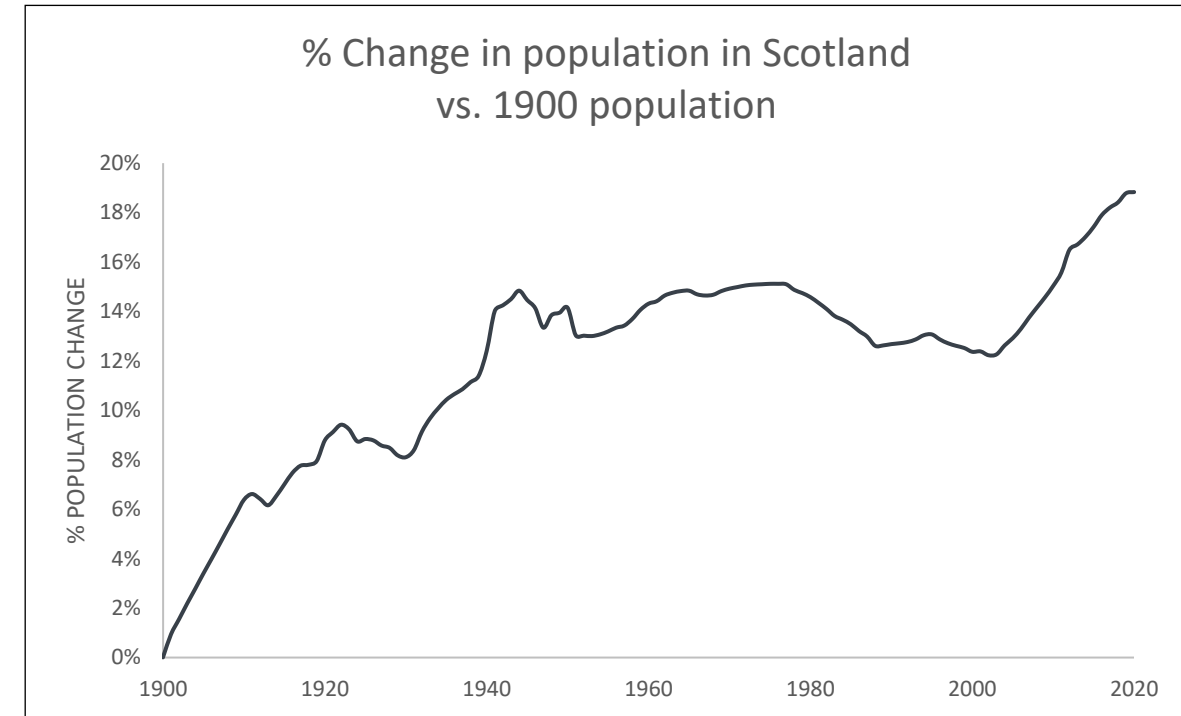
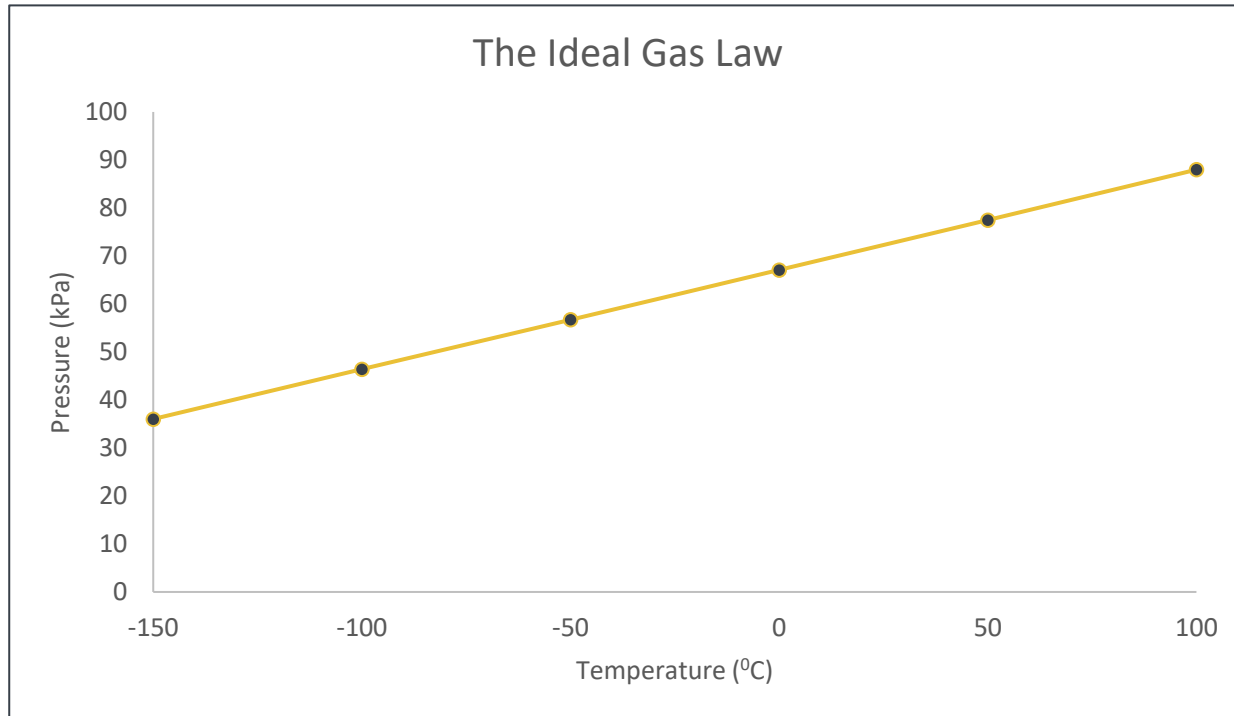
The data points in a line graph are joined sequentially by lines.



Show me...



These lines graphs shows numeric change compared with another numeric variable.



How to create a line graph

We are now going to look at **how to create a line graph in Excel**.

We are going use a dataset that shows the male life expectancy in Scotland by year from 1980 to 2018.



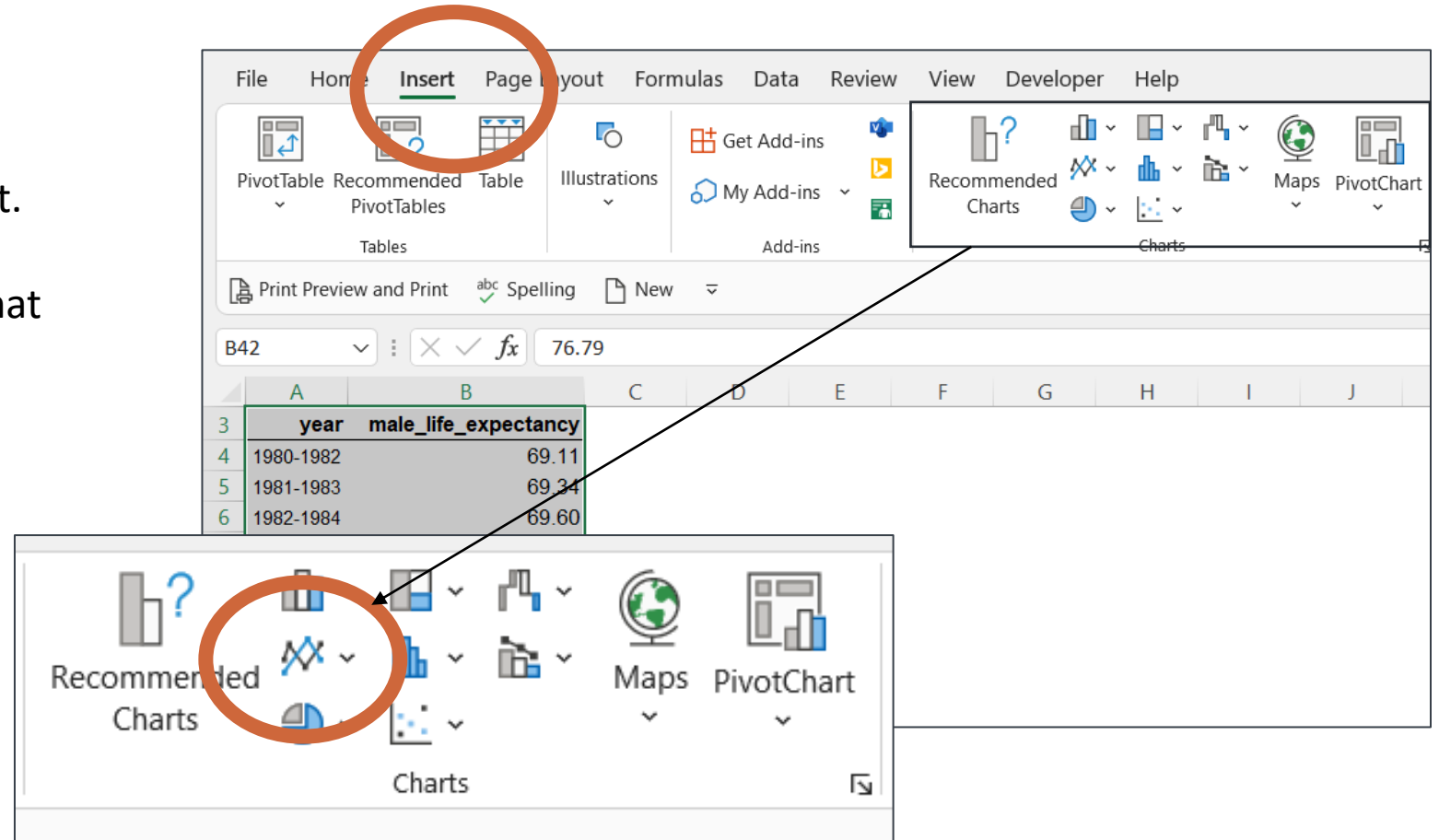
year	male_life_expectancy
1982	69.11
1983	69.34
1984	69.60
1985	69.87
1986	70.01
1987	70.21
1988	70.35
1989	70.55
1990	70.76

Insert a line graph

Step 1.

Highlight the dataset you are going to plot.

Then in the **Insert** tab, click on the **icon** that looks like a line graph.

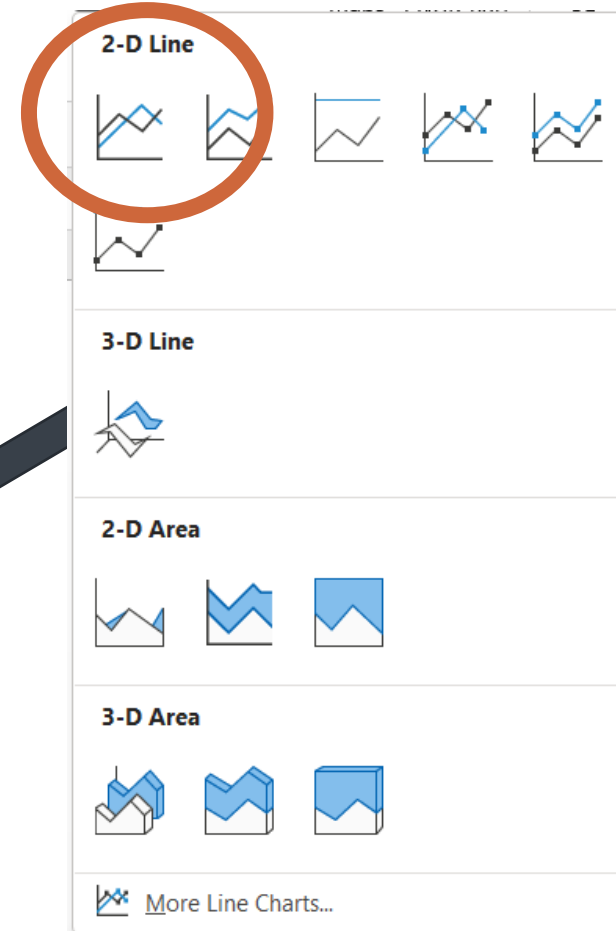
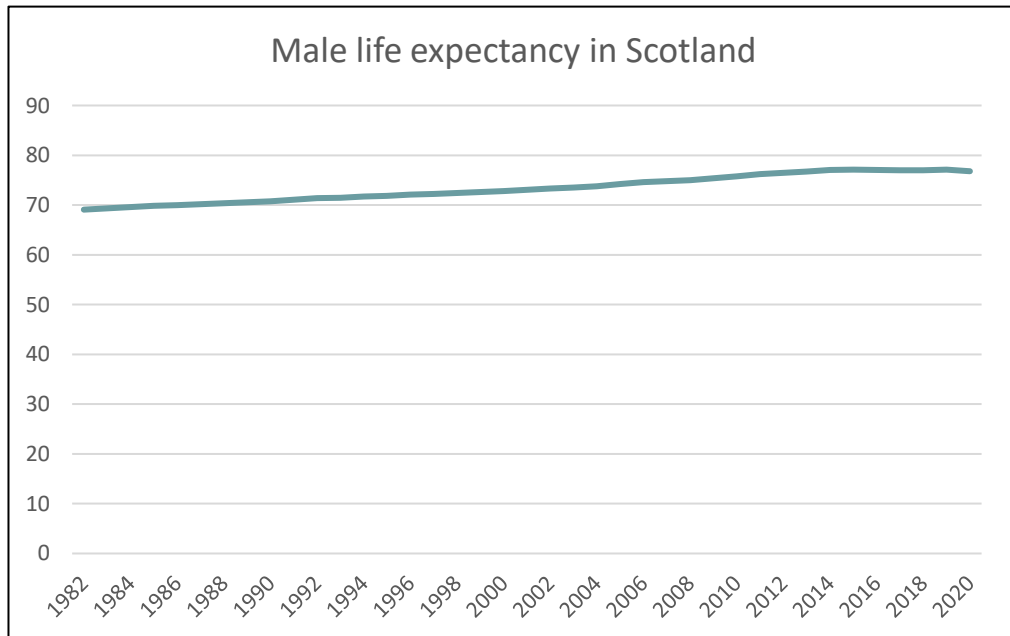


Select the type of line graph

Step 2.

Select the type of line graph you would like from the drop down list.

The graph will now appear in the workbook.

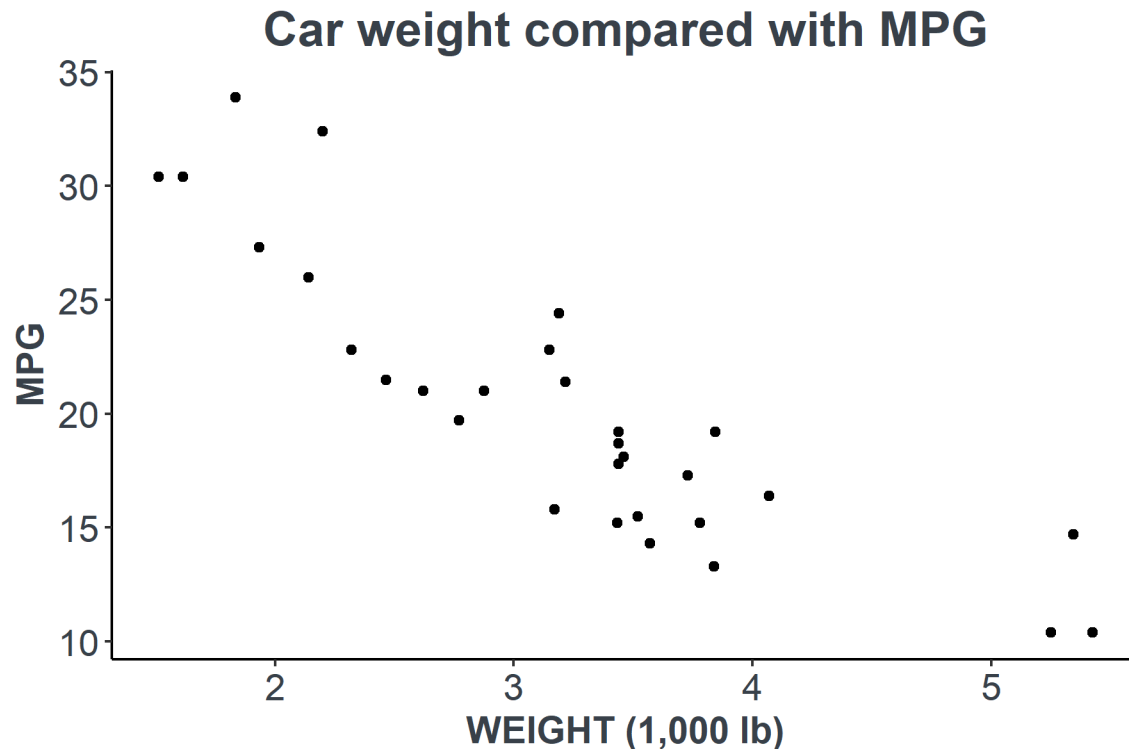


Next steps

Complete **questions 1 to 5**
in **section 3** of the
'Creating graphs in Excel' workbook.

When should you use a scatterplots?

Scatterplots should be used when you want to **show the relationship between two numeric variables**.

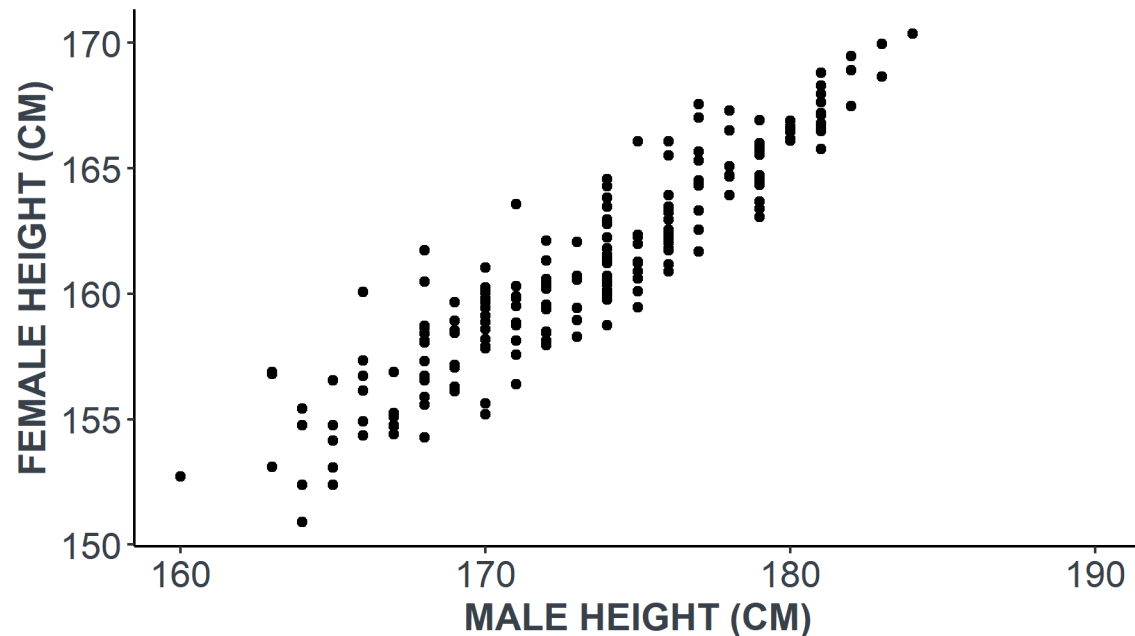


Show me...

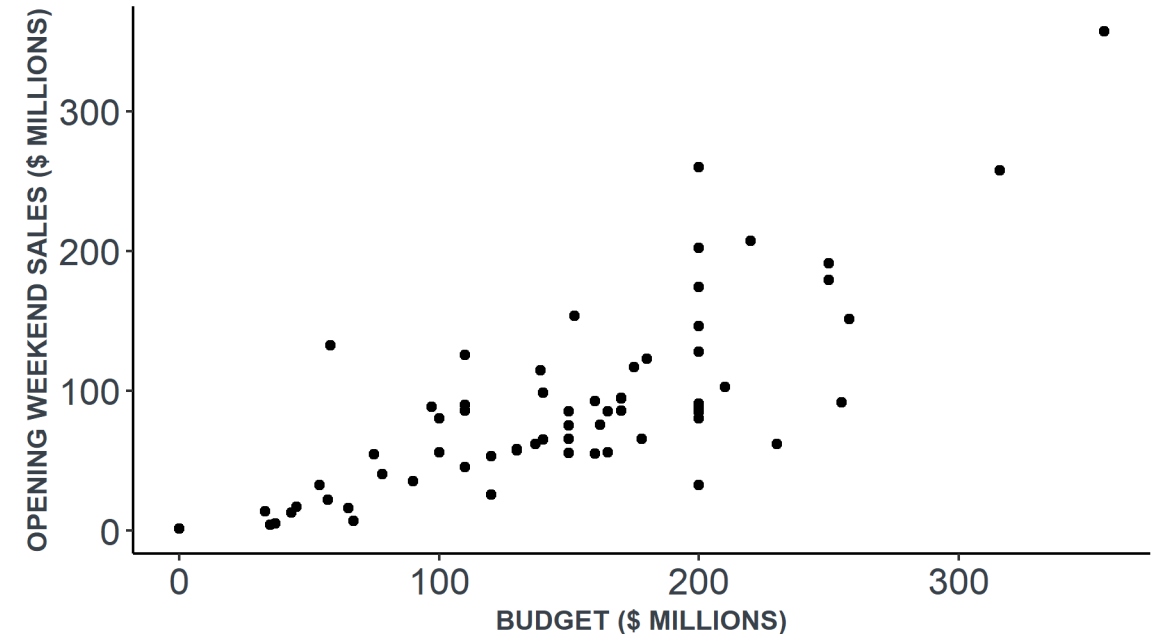


These scatterplots show the relationship between two numeric variables.

Comparison of average male and female height



Marvel movie budgets vs. opening weekend sales



How to create a scatterplot

We are now going to look at **how to create a scatterplot in Excel**.

We are going use a dataset that shows the percentage of empty and occupied houses by council area in Scotland.



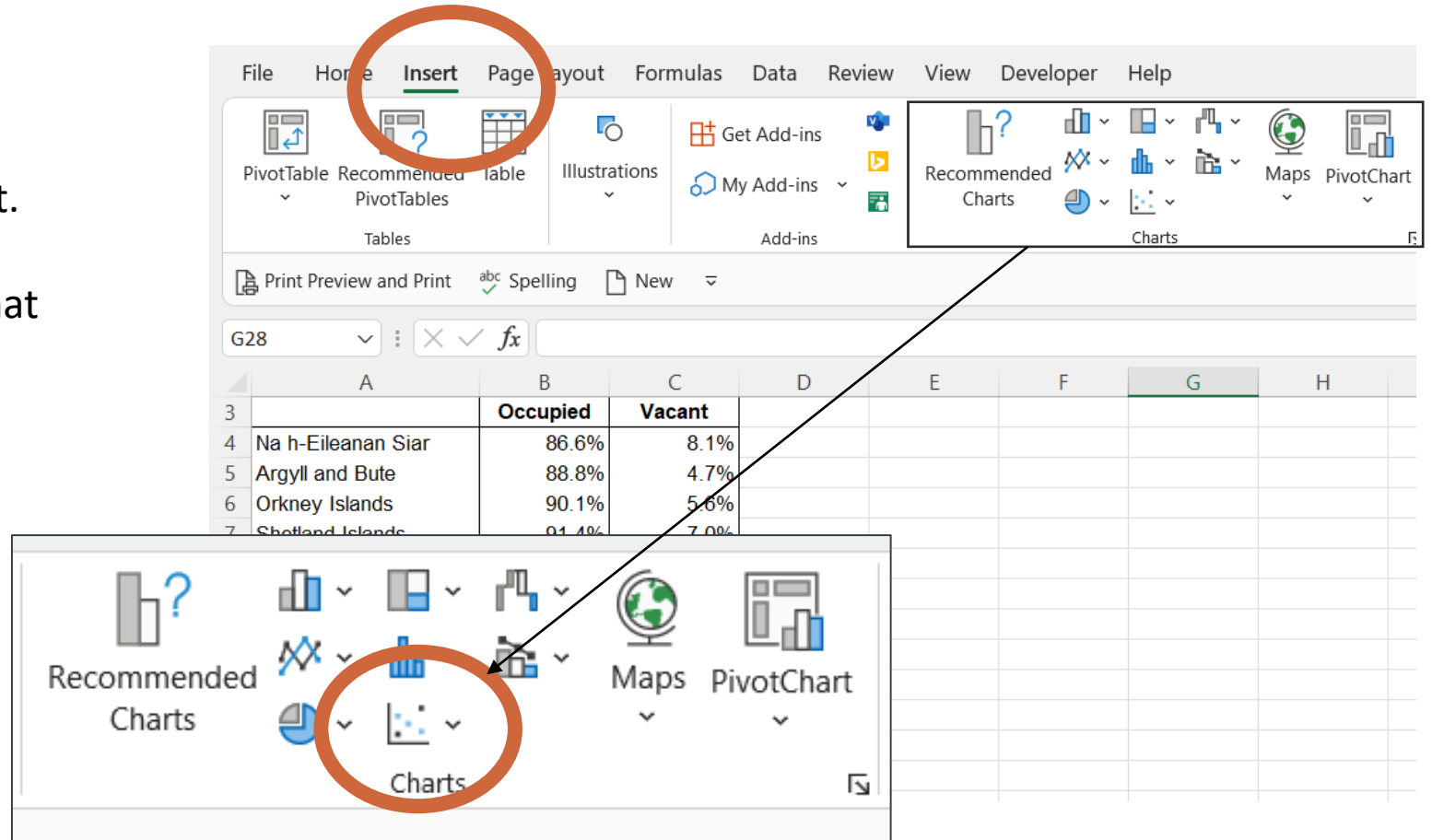
Council_Area	Occupied	Vacant
Na h-Eileanan Siar	86.6%	8.1%
Argyll and Bute	88.8%	4.7%
Orkney Islands	90.1%	5.6%
Shetland Islands	91.4%	7.0%
Highland	93.1%	3.8%
Scottish Borders	93.4%	4.9%
Aberdeen City	93.7%	5.6%
Dumfries and Galloway	93.8%	4.3%
Moray	94.2%	4.1%
Aberdeenshire	94.3%	4.7%
North Ayrshire	94.4%	3.4%

Insert a scatterplot

Step 1.

Highlight the dataset you are going to plot.

Then in the **Insert** tab, click on the **icon** that looks like a scatterplot.

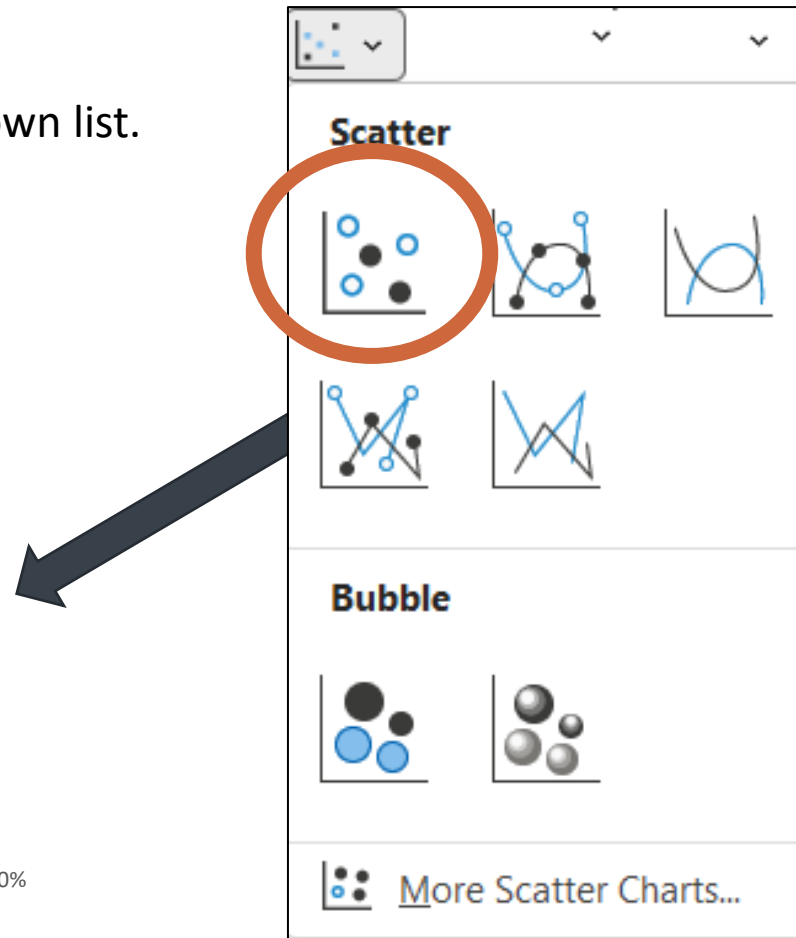
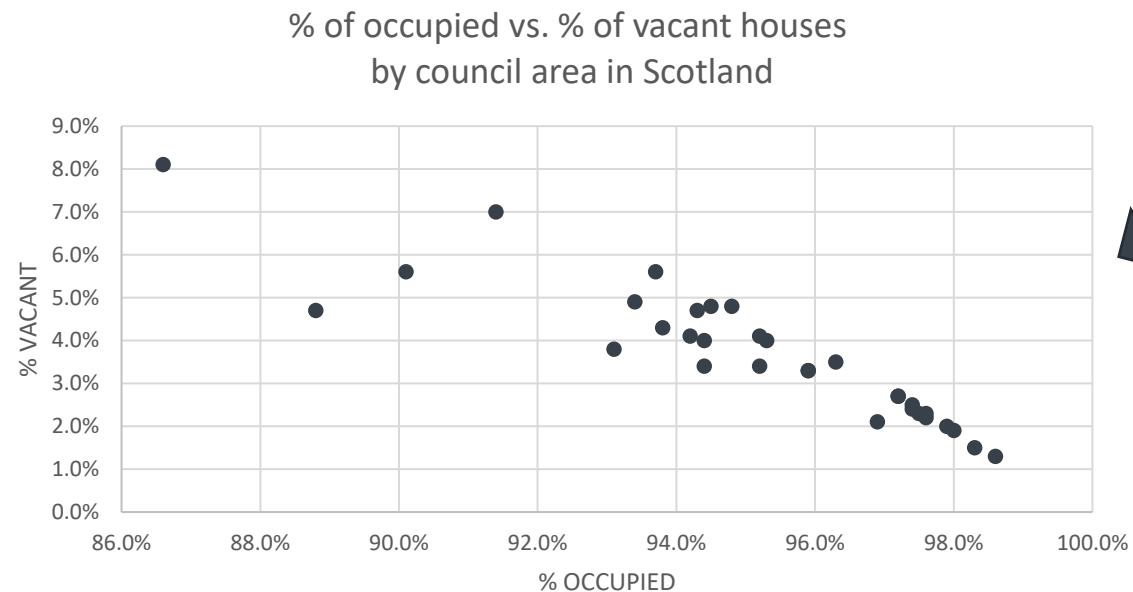


Select the type of scatterplot

Step 2.

Select the type of scatterplot you would like from the drop down list.

The graph will now appear in the workbook.



Next steps

Complete **questions 1 to 5**
in **section 4** of the
'Creating graphs in Excel' workbook.

Learning checklist

I can *describe* how creating graphs is part of the analysis steps.

I can *create* bar charts in Excel.

I can *create* histograms in Excel.

I can *create* line graphs in Excel.

I can *create* scatterplots in Excel.

How you can use this lesson



You are free to:

- **Share** – copy and redistribute the material in any medium or format
- **Adapt** – remix, transform and build upon the material

Under the following terms:

- **Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **NonCommercial** — You may not use the material for [commercial purposes](#).
- **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.

© 2022. This work is licensed under a [CC BY-NC-SA 4.0 license](#).

Created by Effini in partnership with Data Education in Schools, The Data Lab and Data Skills for Work, with funding from the Scottish Government.



Alternative format

If you require this document in an alternative format, such as large print or a coloured background, please contact

hello@effini.com

or

**4th Floor, The Bayes Centre
47 Potterrow
Edinburgh
EH8 9BT**

