

Practise creating graphs in Excel (part 2)

Version: 1.0



Learning intentions

We will be learning more about **creating line graphs and scatterplots** in Excel, specifically,

- how to make **standard changes** to line graphs and scatterplots
- how to plot a **line graph without date value** variables
- how to **add data labels**
- how to amend **data points**

Background

Creating graphs is a useful tool when looking for patterns as part of the analysis steps.

We have looked at design choices and how to create basic graphs in Excel.

Now we are going to look at ways to **amend line graphs and scatterplots.**



Data understanding

Data cleansing

Data manipulation

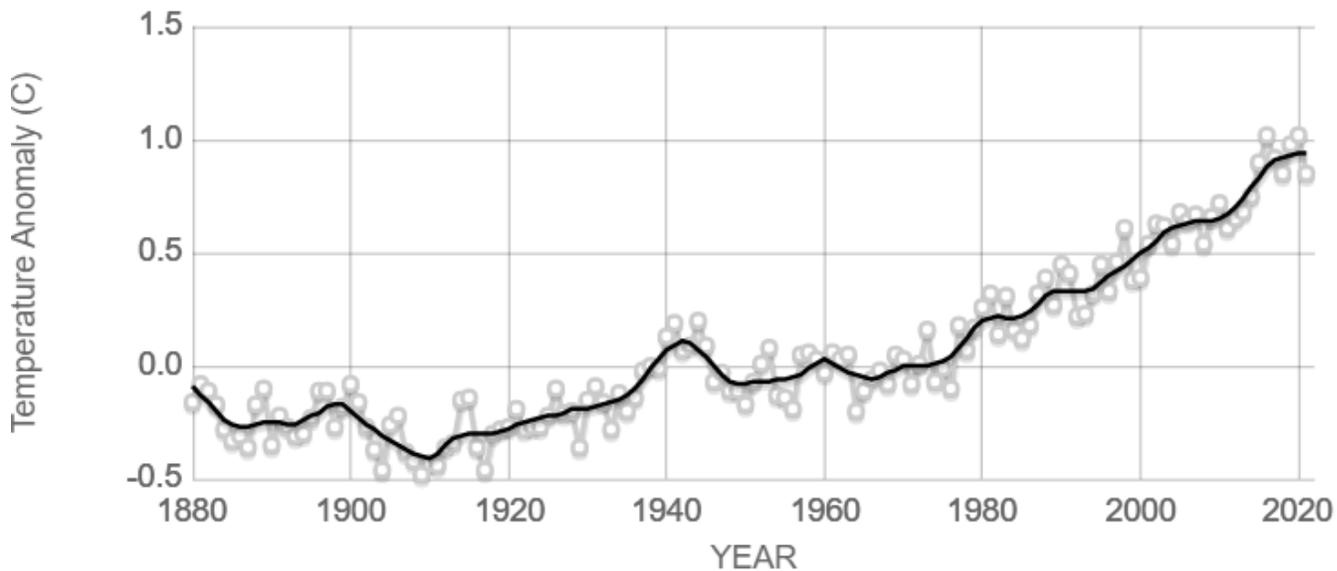
Identifying patterns

Extracting insights

Show me...



Below are examples of line graphs and scatterplots used by NASA and data-to-viz.com



Source: climate.nasa.gov

<https://climate.nasa.gov/vital-signs/global-temperature/>



<https://www.data-to-viz.com/graph/scatter.html>

Creating a line graph in Excel

In this lesson we are going to use a dataset that shows the number of weddings in Scotland by year since 1900.

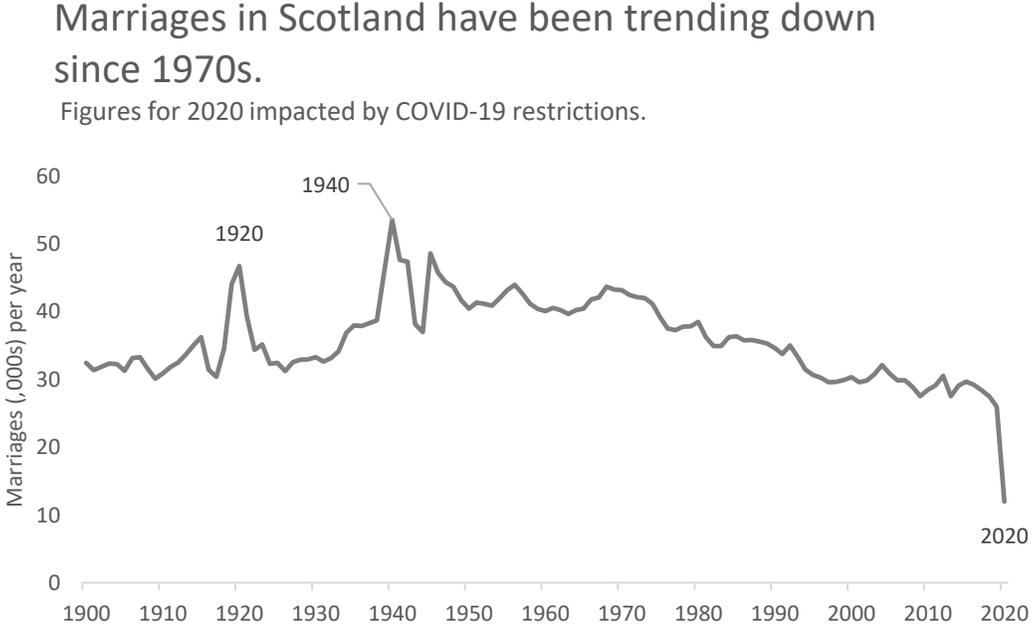
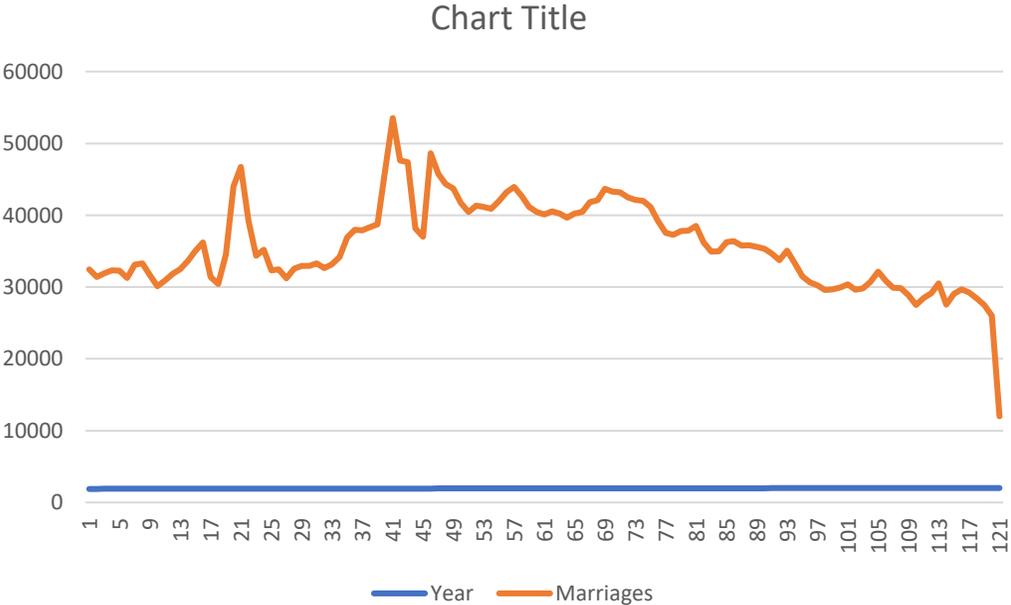
We will start with the default graph and amend it.

The instructions in this lesson are for Microsoft Excel on a Windows based machine. If you are using something different (e.g. Mac) the position of menus/options might be different.



Before vs. after line charts

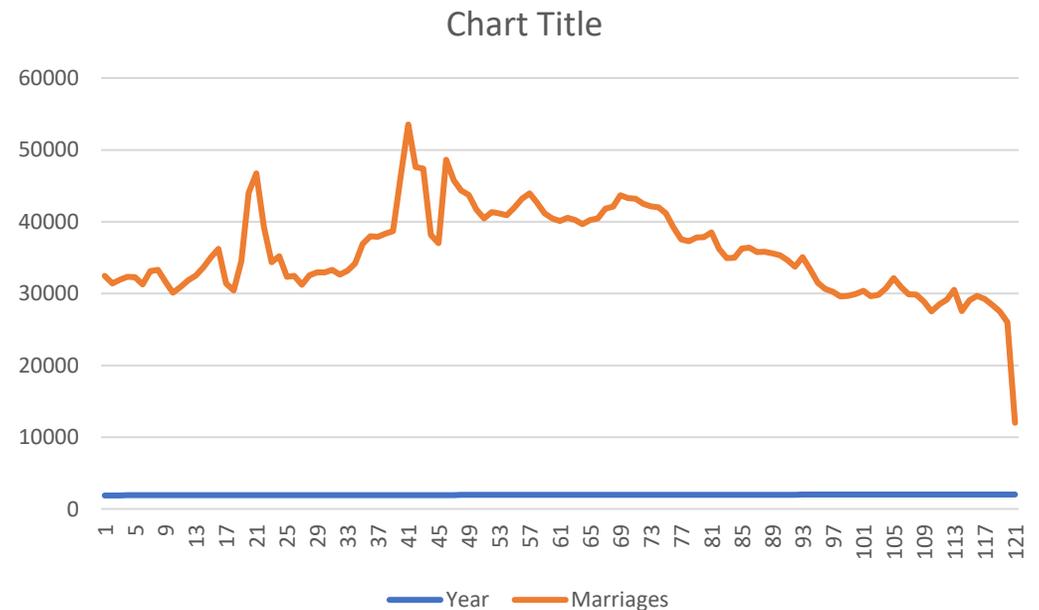
Below are graphs created using the wedding dataset. The first is the default graph created, and the other is after it has been amended. We will now work through how to create the second graph in Excel.



Amending the default graph

We will start with the Excel default graph and go through the steps to make the following changes,

1. Fixing common line graph “bug”
2. Change the display format of the numbers
3. Removing the legend
5. Amend the gridlines
6. Changing the major tick marks
7. Adding data labels
8. Change the line

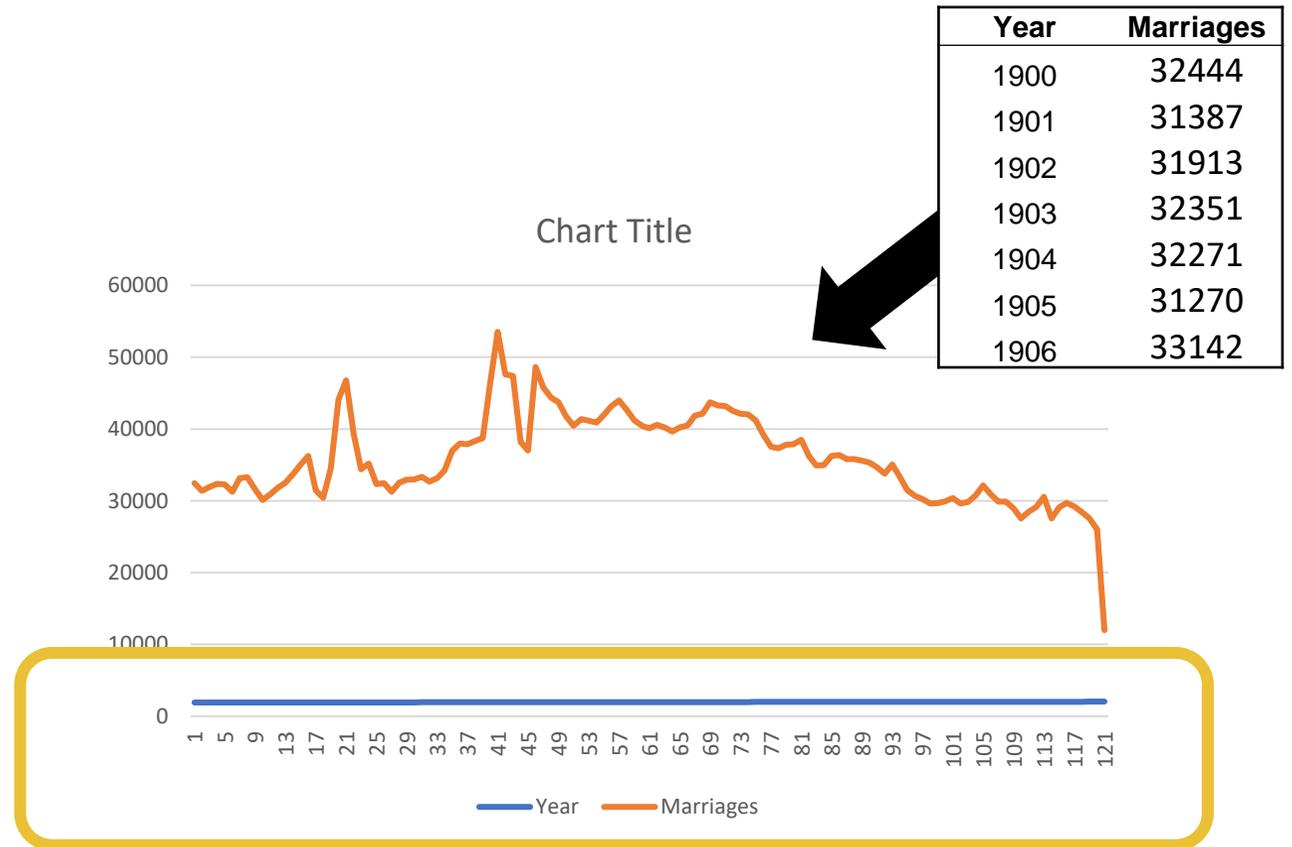


Common “bug” when creating line graphs

Line graphs generally are used to when you have a **date or time variable on the x-axis**.

If you have a **dataset where the variable isn't a full date/time value** (e.g. it's only the month or year part) Microsoft Excel doesn't plot the data as you expect.

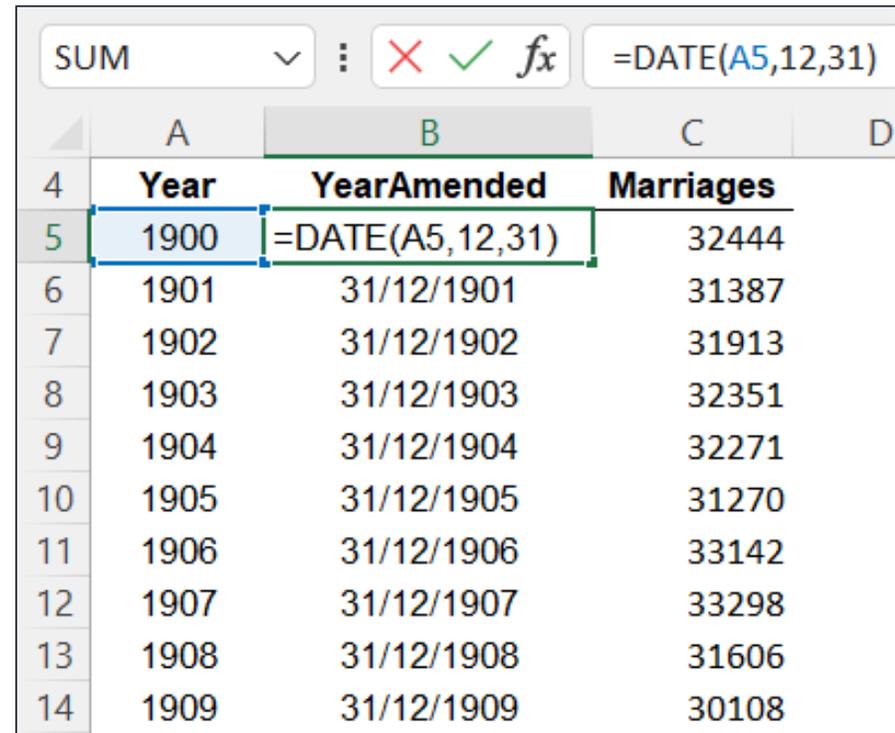
One way of fixing this issue is to convert the variable to a date format.



Create new date value from year

The value in the **Year** variable is seen by Excel as an **integer**. Line graphs need a date value to work correctly.

By using the **DATE function** in Excel you can create a new date variable.



The screenshot shows an Excel spreadsheet with the following data:

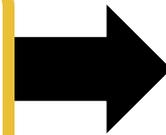
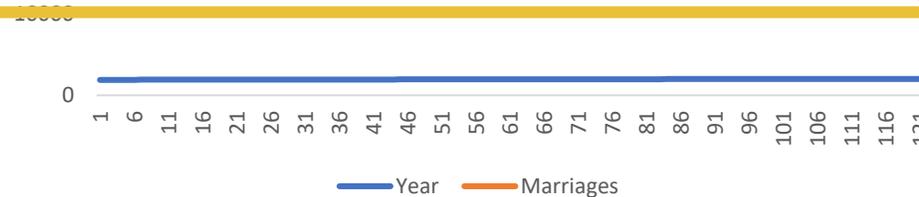
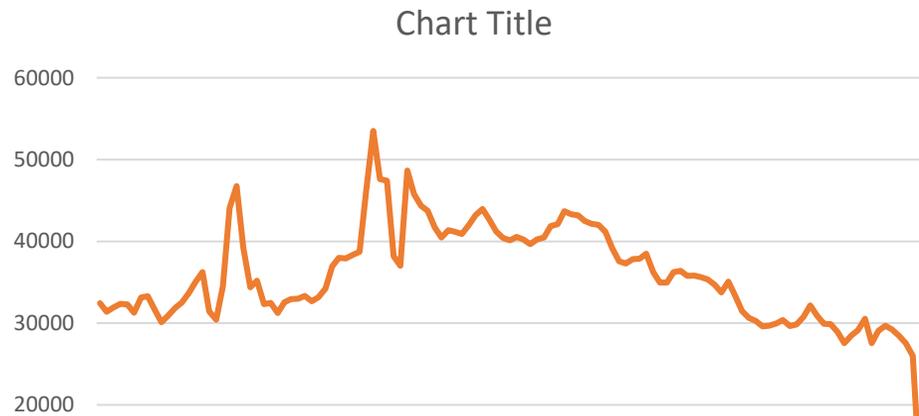
	A	B	C	D
4	Year	YearAmended	Marriages	
5	1900	=DATE(A5,12,31)	32444	
6	1901	31/12/1901	31387	
7	1902	31/12/1902	31913	
8	1903	31/12/1903	32351	
9	1904	31/12/1904	32271	
10	1905	31/12/1905	31270	
11	1906	31/12/1906	33142	
12	1907	31/12/1907	33298	
13	1908	31/12/1908	31606	
14	1909	31/12/1909	30108	

The formula bar at the top shows the formula `=DATE(A5,12,31)` being entered into cell B5. The spreadsheet also shows a status bar with 'SUM' and a formula icon.

`=DATE(year,month,day)`

Plot the line graph using new variable

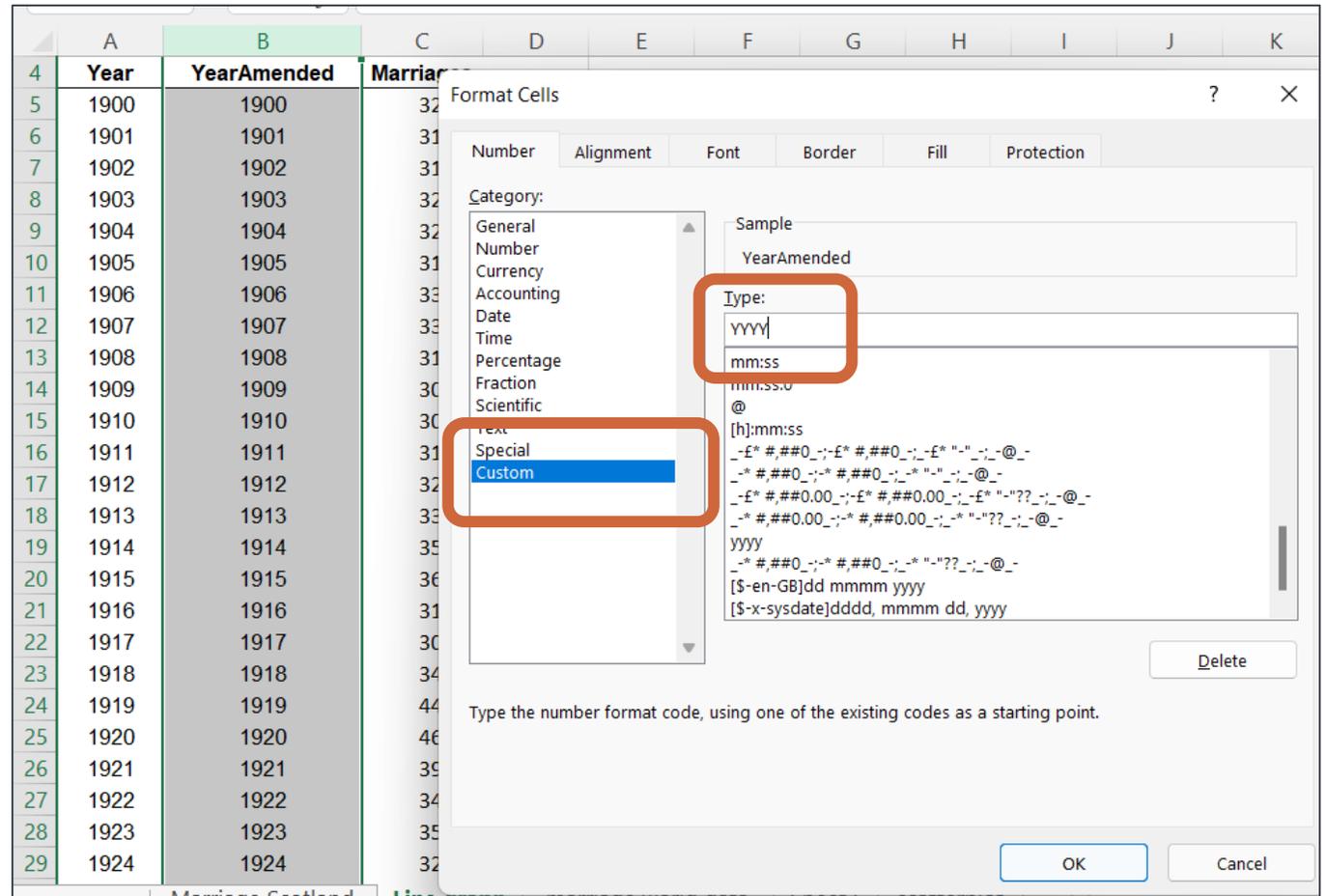
By plotting the graph using the new variable (YearAmended), the x-axis looks how you would expect.



Changing the display format

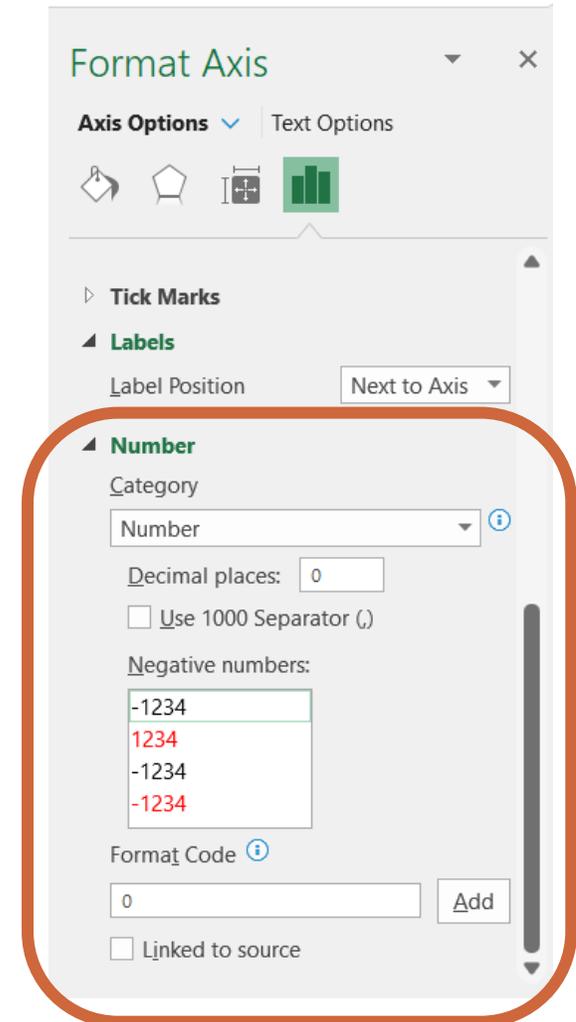
To change the format of new variable (YearAmended) so it appears as just a year, you need to

1. Select the variable in the dataset
2. Open up **Format Cells** window (Ctrl+1).
3. Click on the '**Custom**' category.
4. Type "**yyyy**" into the Type space.



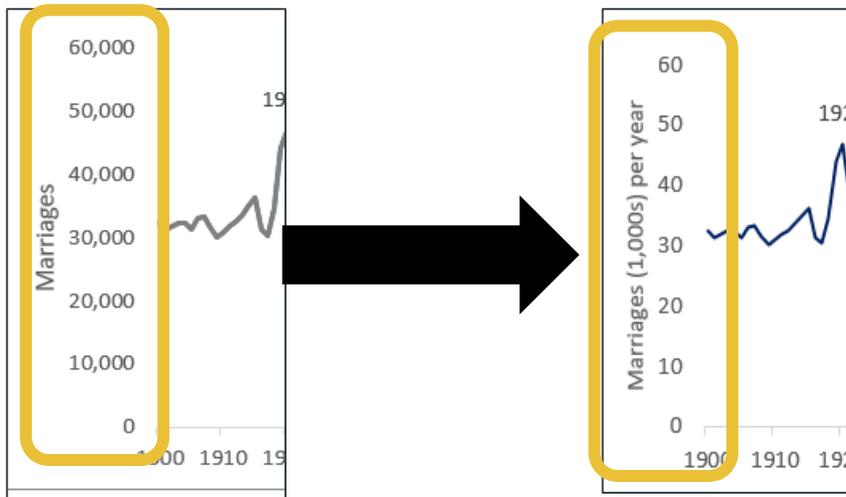
Changing the y-axis display format

As well as changing the display format of the data in the dataset, you can also just change the display format of the numbers in the graph in the **Format Axis** pane.



Format numbers in thousands

To change the data so it is shown in thousands you need add a 'Custom' Type.

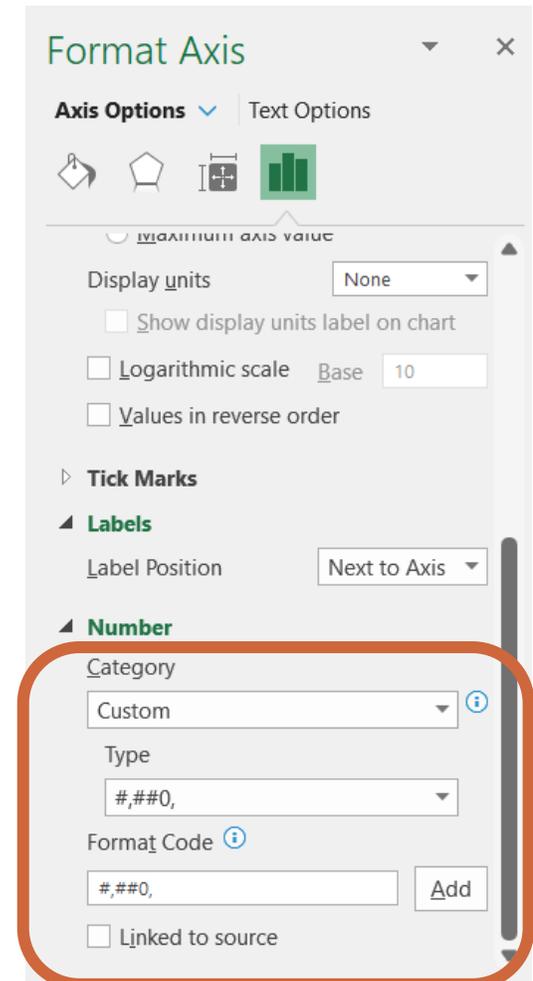


The screenshot shows the 'Format Axis' task pane in Excel. The 'Axis Options' tab is selected. The 'Display units' dropdown is set to 'None'. The 'Logarithmic scale' checkbox is unchecked, and the 'Base' is set to 10. The 'Values in reverse order' checkbox is unchecked. The 'Tick Marks' section is expanded. The 'Labels' section is expanded, and the 'Label Position' dropdown is set to 'Next to Axis'. The 'Number' section is highlighted with an orange box. The 'Category' dropdown is set to 'Custom'. The 'Type' dropdown is set to '#,##0,'. The 'Format Code' field contains '#,##0,' and the 'Add' button is visible. The 'Linked to source' checkbox is unchecked.

Creating custom data formats thousands

To create a custom thousands data format,

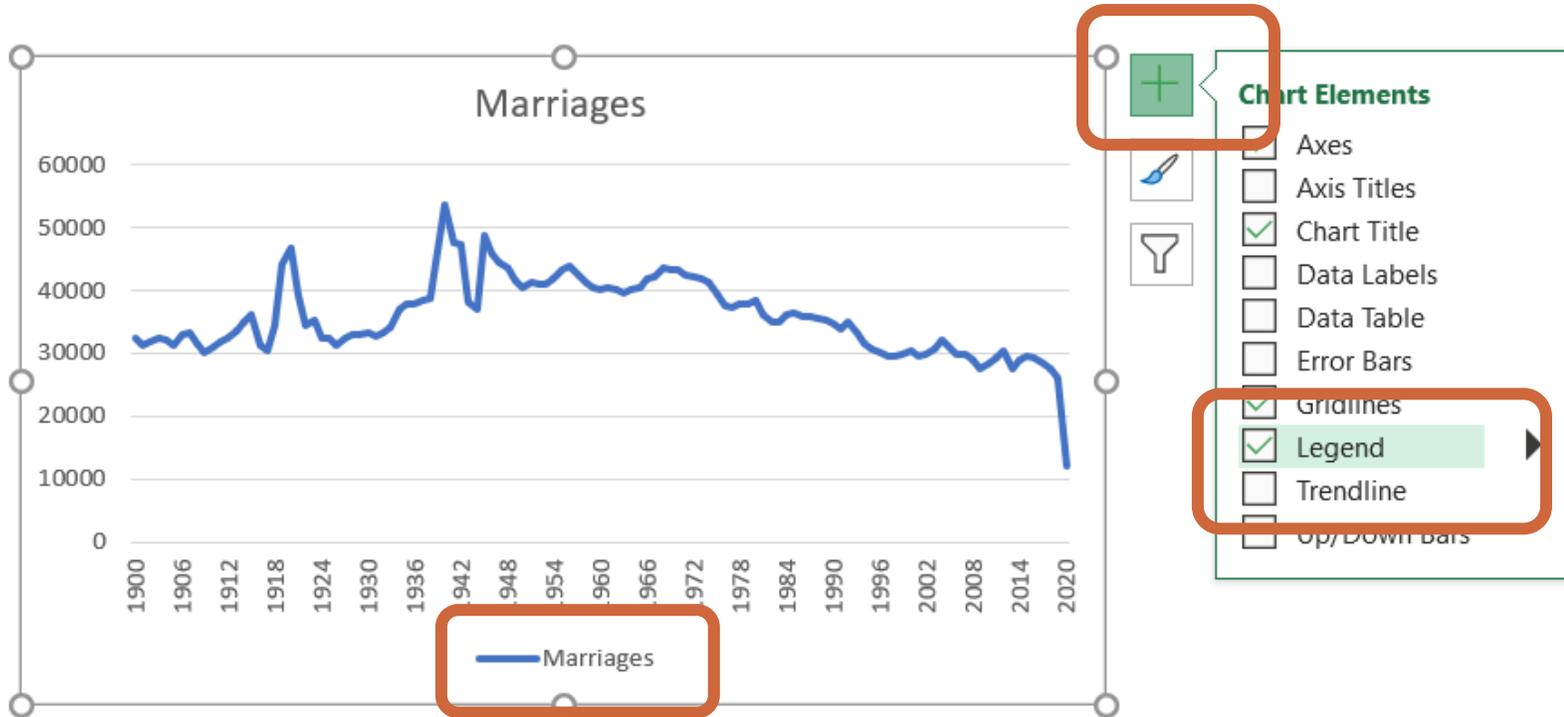
1. Select Custom from the **Category** drop down option
2. Type **#,##0,** into the **Format Code** box and then press “Add”
3. Select the format you have just created (**#,##0,**) from the **Type** drop down option



Removing a Legend

To remove a legend from your graph that you do not need.

1. Click on the + button at the right hand side of the graph
2. Untick the box next to Legend option



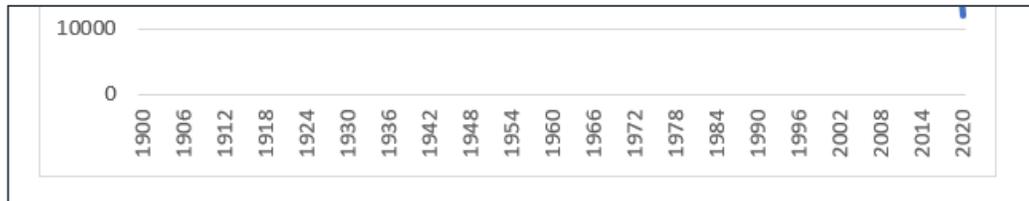
Change the major tick marks

Reminder

Tick marks should be spaced so they make it easier for the reader

You can change how the tick marks are displayed on the x-axis by opening the Format Axis pane.

In this case the major tick marks are going to be shown every 10 years.



The screenshot shows an Excel spreadsheet with a line chart. A context menu is open over the x-axis, with the 'Format Axis...' option highlighted. The 'Format Axis' task pane is also open on the right, with the 'Units' section highlighted. The 'Units' section shows 'Major' ticks set to 10 years and 'Minor' ticks set to 1 year.

Format Axis

Axis Options | Text Options

Axis Options

Axis Type

- Automatically select based on data
- Text axis
- Date axis

Bounds

Minimum: 02/01/1900 Auto

Maximum: 01/01/2020 Auto

Units

Major: 10 Years Reset

Minor: 1 Years Auto

Base: Years Reset

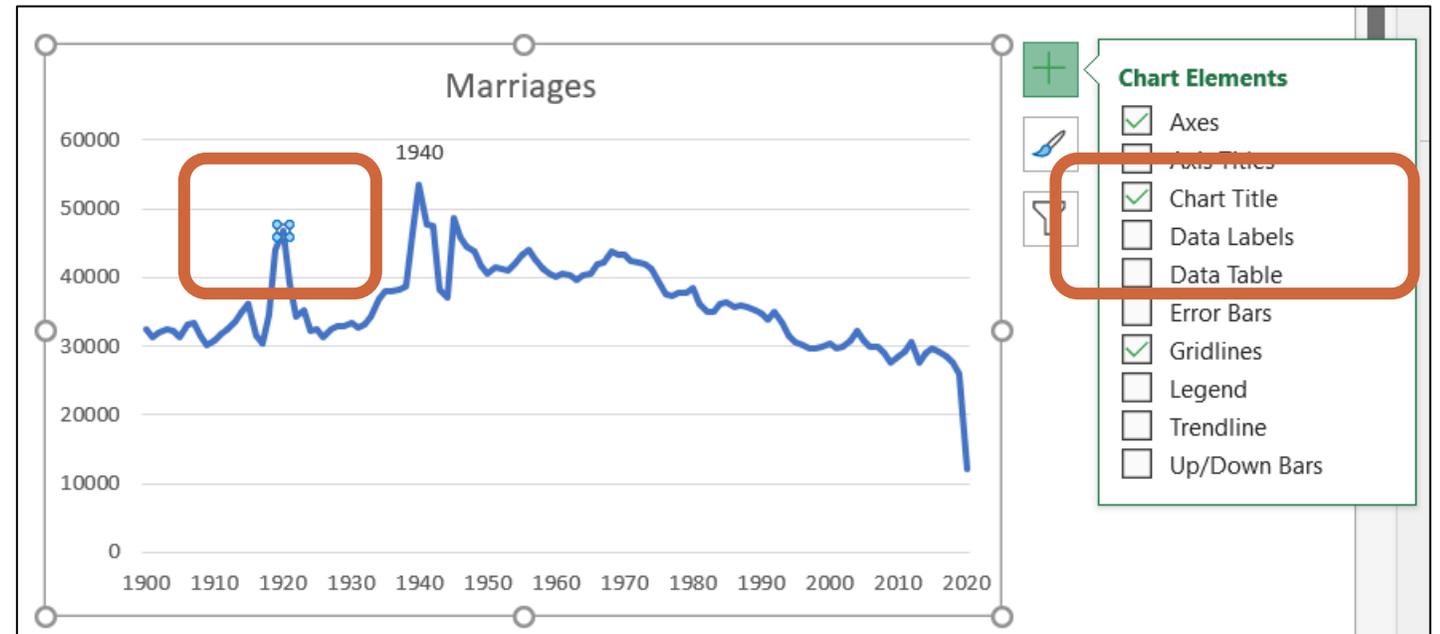
Add data labels

Reminder

Data labels should highlight key information for the reader

If you want to add data labels to specific points on your line graph you need to,

- Click on the data point
- Select 'Data Labels' from the chart elements

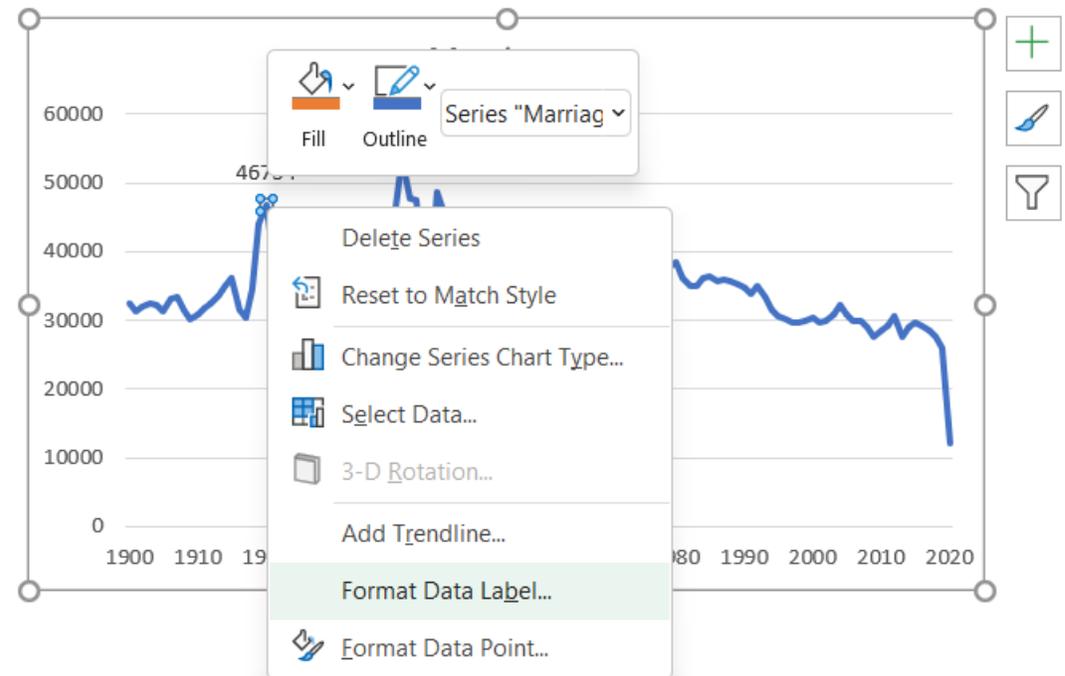


Amend the data labels

When the data label is added to the graph, the default is to show the y-axis value (e.g. number of marriages).

On this graph we want to label the year instead.

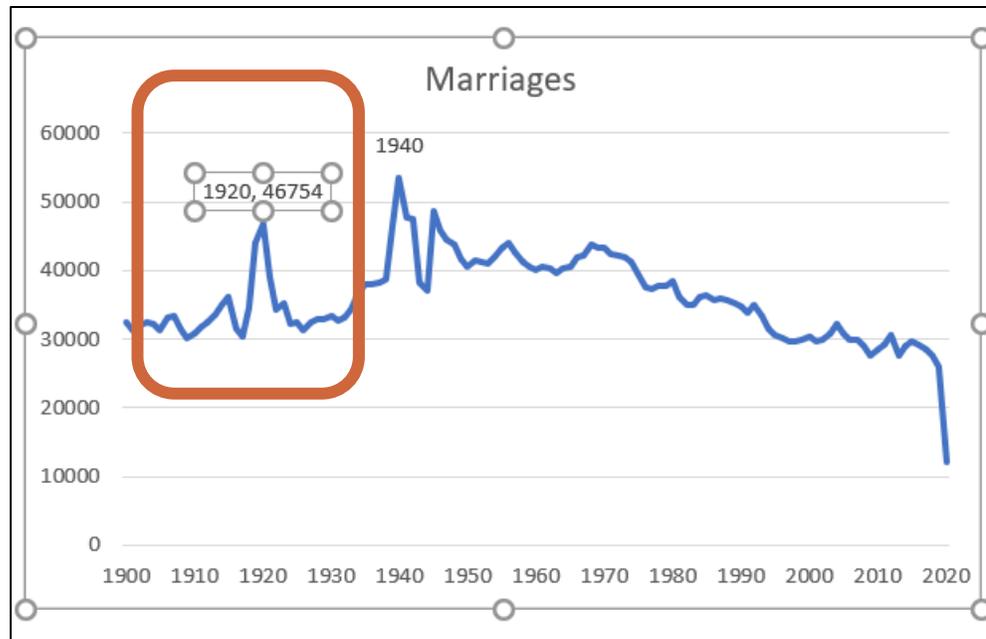
Right click on the data point and select '**Format Data Label**'.



Amend data labels

The label options allows you to change what you want to be shown on the graph.

We want to show the **year** (from the x-axis) which Excel calls the 'Category Name'.



The screenshot shows the 'Format Data Label' task pane. The 'Label Options' section is highlighted with an orange rounded rectangle. The 'Label Contains' section has the following options:

- Series Name
- Category Name
- Value
- Show Leader Lines
- Legend key

The 'Separator' is set to a comma (,). There is a 'Reset Label Text' button. The 'Label Position' is set to 'Center'.

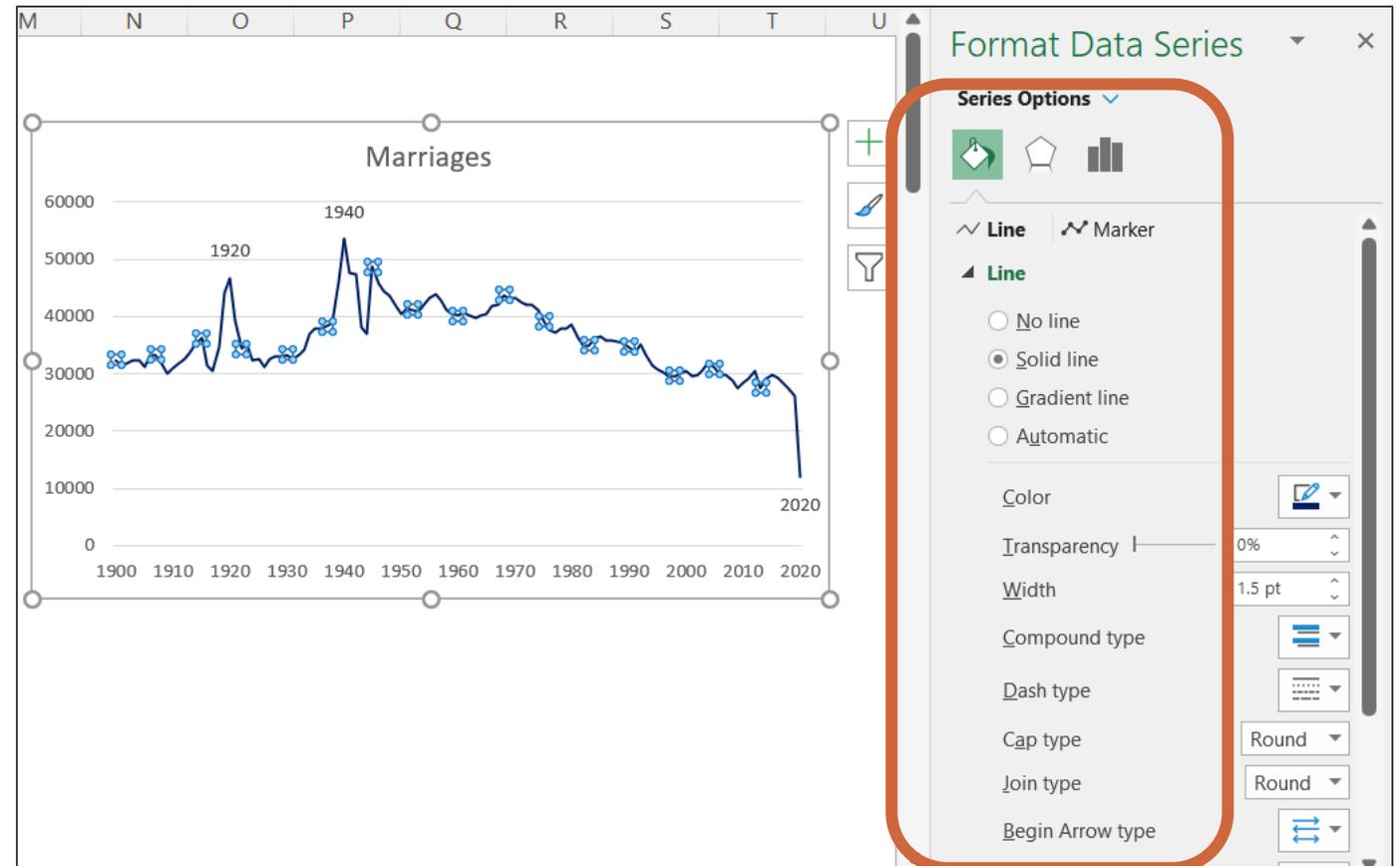
Change the line style

Reminder

Design with accessibility (e.g. colour-blindness) in mind

To change the line style or colour, right click on the line and choose '**Format Data Series**'.

This will open up the Format Data Series pane where you edit the style of the line.

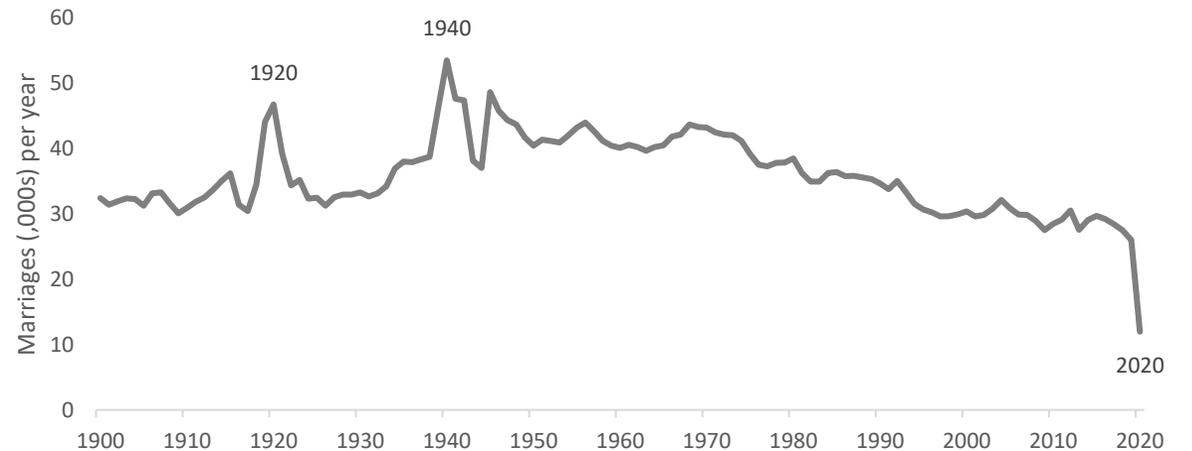


Change the axis and chart titles

Now you have amended the elements of the graph you **can change/add axis and chart titles** in the same way as you did in previous graphs.

Number of marriages in Scotland have been trending down since 1970s.

Figures for 2020 impacted by COVID-19 restrictions.



Next steps

Complete **questions 1 to 10**
in **section 1** of the
'Practise creating graphs
in Excel (part 1)' workbook.

Then complete questions 1 and 2 in **section 2**.

Creating a scatterplot in Excel

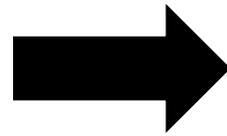
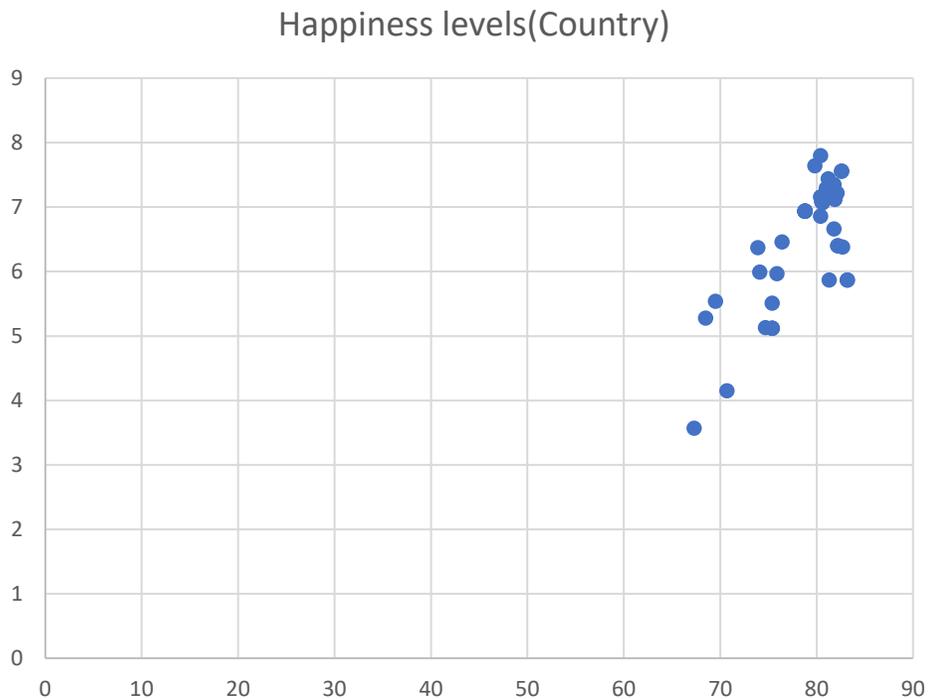
Now, we are going to use a dataset that has the Happiness index and Life Expectancy of cities to create a scatterplot.

Again we will start with the default Excel graph and then amend it.



Before vs. after scatterplots

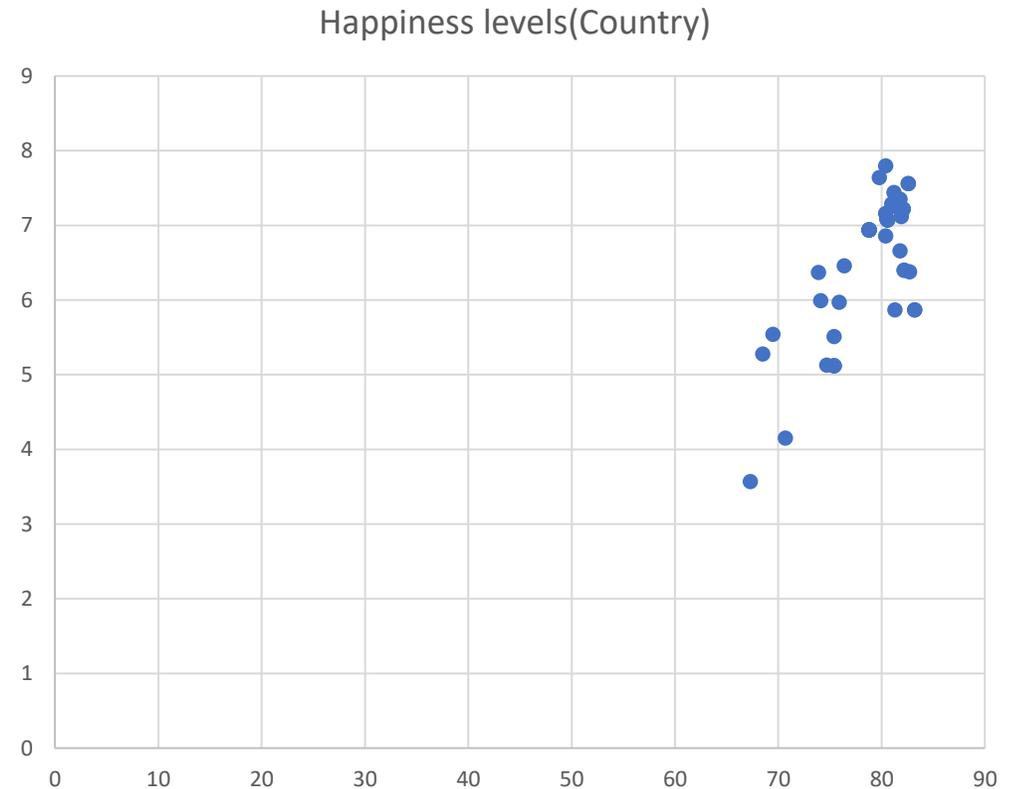
Below are graphs created using the happiness vs. life expectancy dataset. The first is the default graph created, and the other is after it has been amended.



Amending the default graph

We will start with the Excel default graph and go through these steps to amend the graph,

1. Add axis title
2. Amend the gridlines
3. Change the axis range
4. Format data points
5. Add a trendline
6. Change the chart title

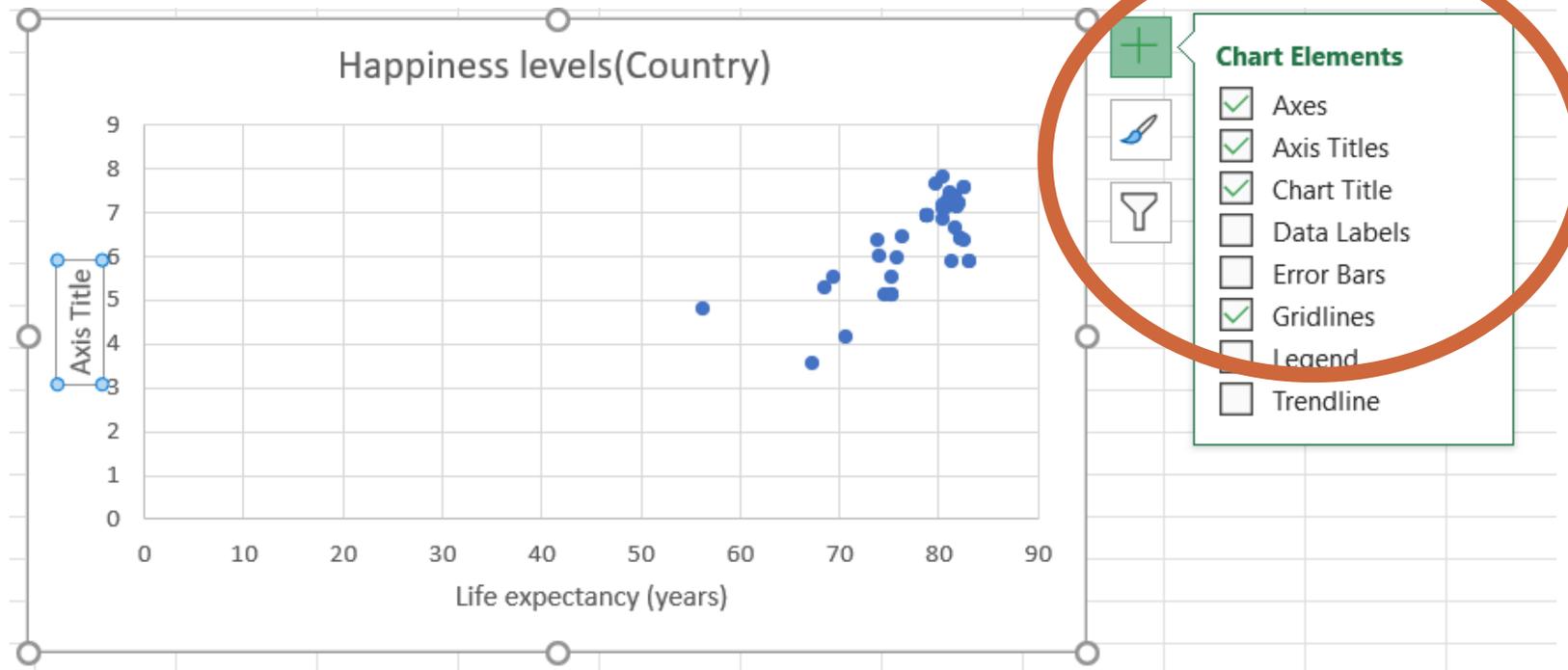


Adding axis titles

Reminder

Axis titles avoid confusion around what is being shown

To add axis or chart titles to a graph, click on the graph then **press + button** on the right hand side and select chart elements you need.



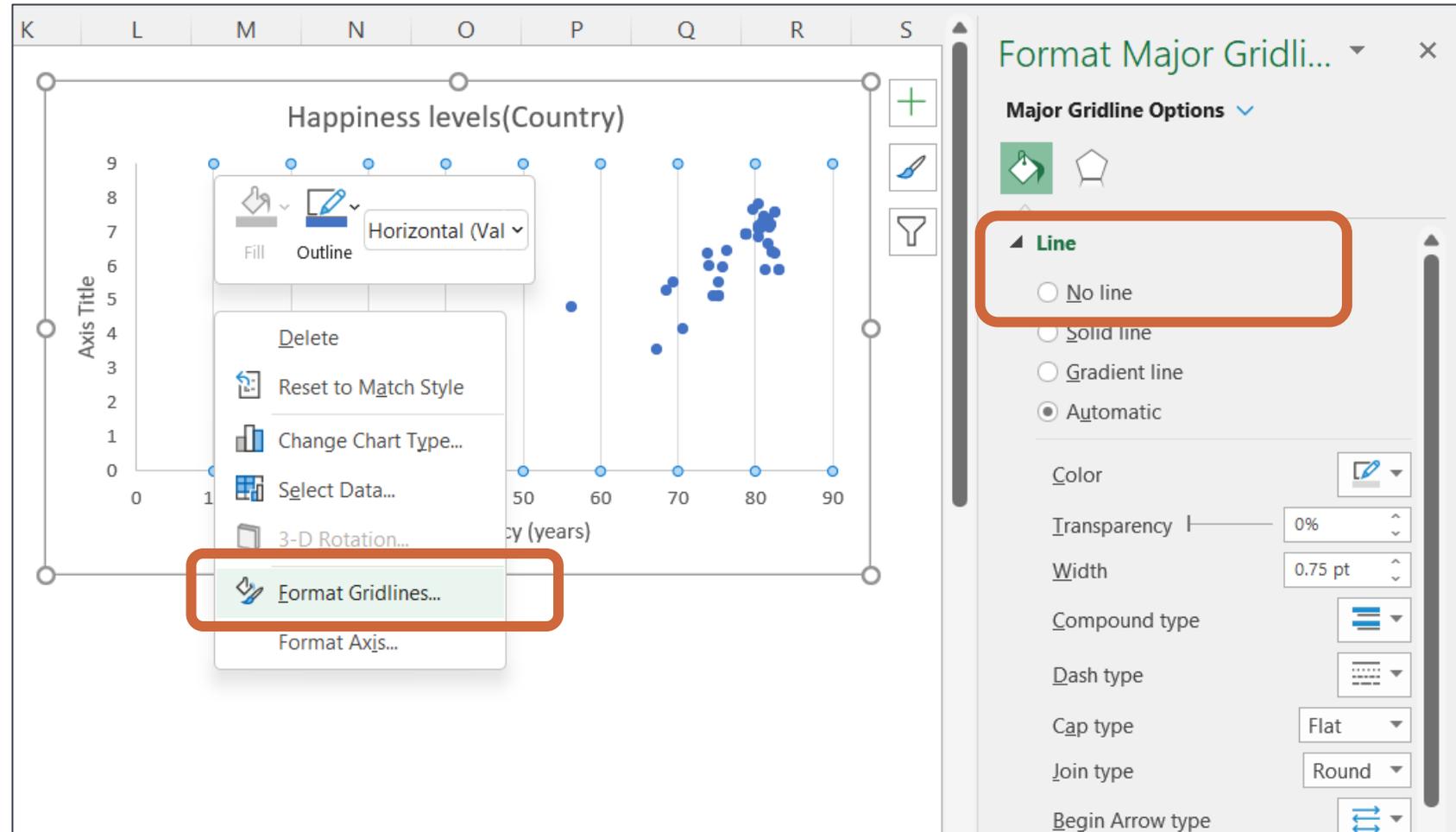
Amend the gridlines

Reminder

Gridlines should not distract from your data

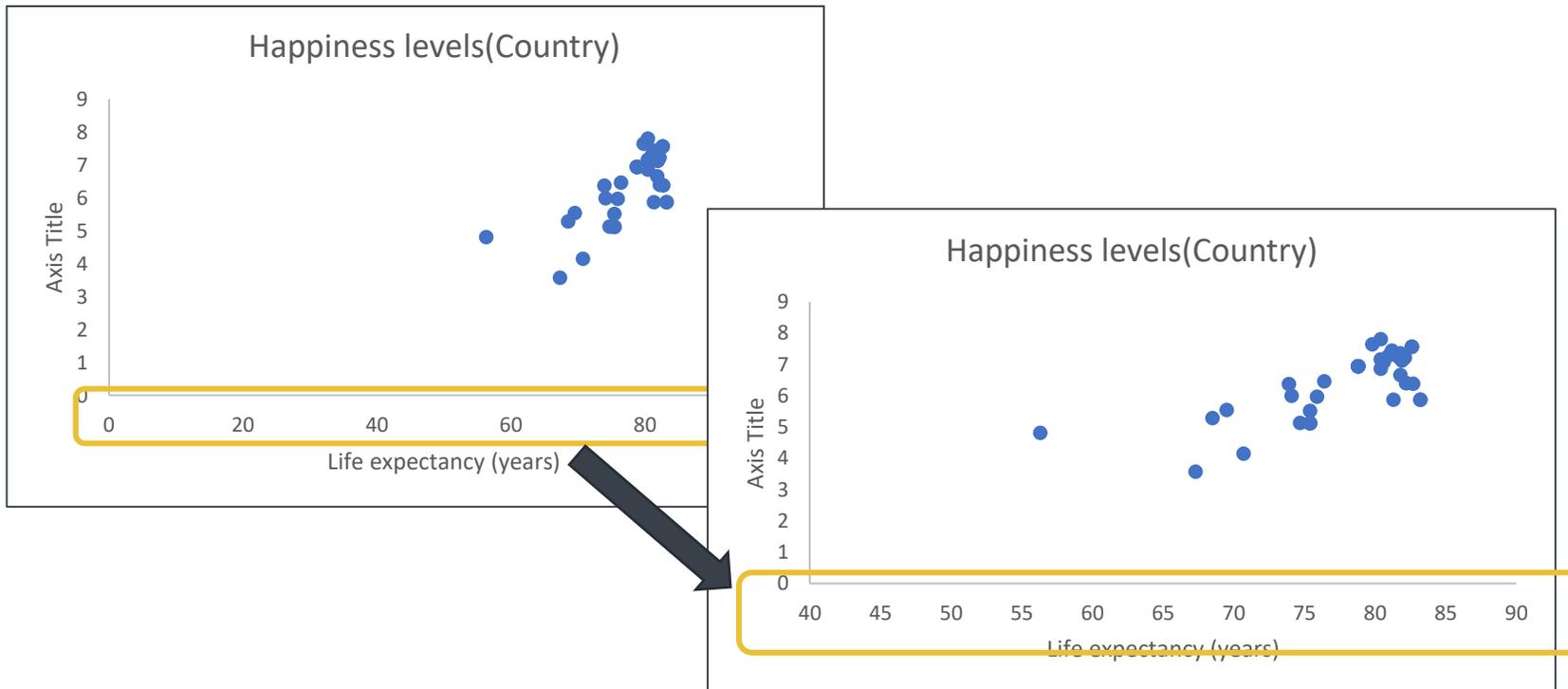
To amend the gridlines, right click on the gridlines and then select Format Gridlines.

For this graph, the gridlines have been removed by selecting the No line option.



Changing the axis range

In the Format Axis pane, click on the **Axis Options**. You can then change the **minimum and maximum** range of the axis in the **Bounds** section and the major and minor units.



The screenshot shows the "Format Axis" pane with the "Axis Options" tab selected. The "Bounds" section is highlighted with a yellow box. The "Minimum" value is set to 40.0 and the "Maximum" value is set to 90.0. The "Units" section shows "Major" set to 5.0 and "Minor" set to 1.0. The "vertical axis crosses" section has "Automatic" selected. The "Display units" dropdown is set to "None". The "Logarithmic scale" checkbox is unchecked, and the "Values in reverse order" checkbox is also unchecked.

Format data points

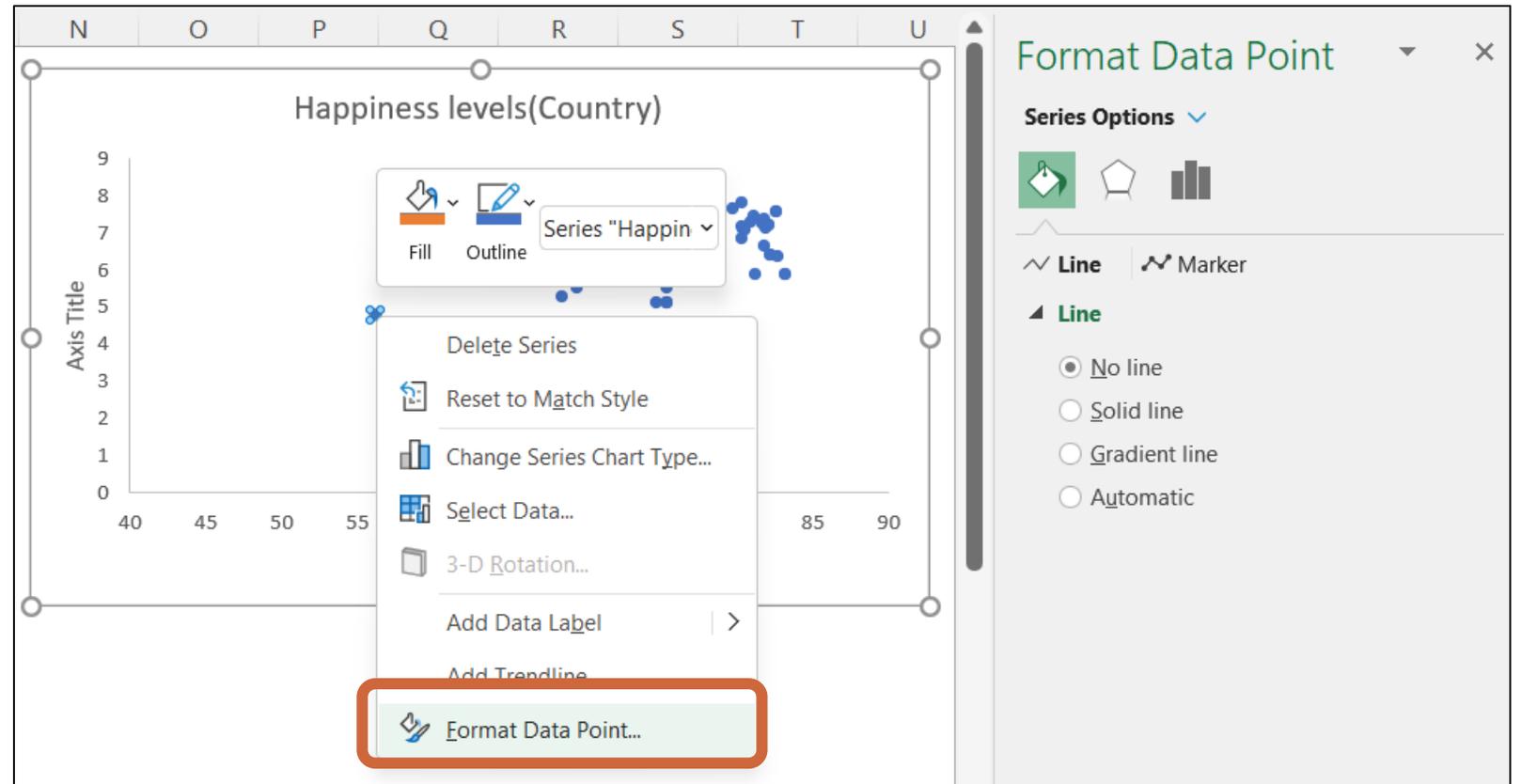
Reminder

Design with accessibility (e.g. colour-blindness) in mind

You can change the following about the marker,

- Shape
- Colour to fill
- Colour around the border of the shape

To format the data points click on any of the points then right click and select **'Format Data Point..'**



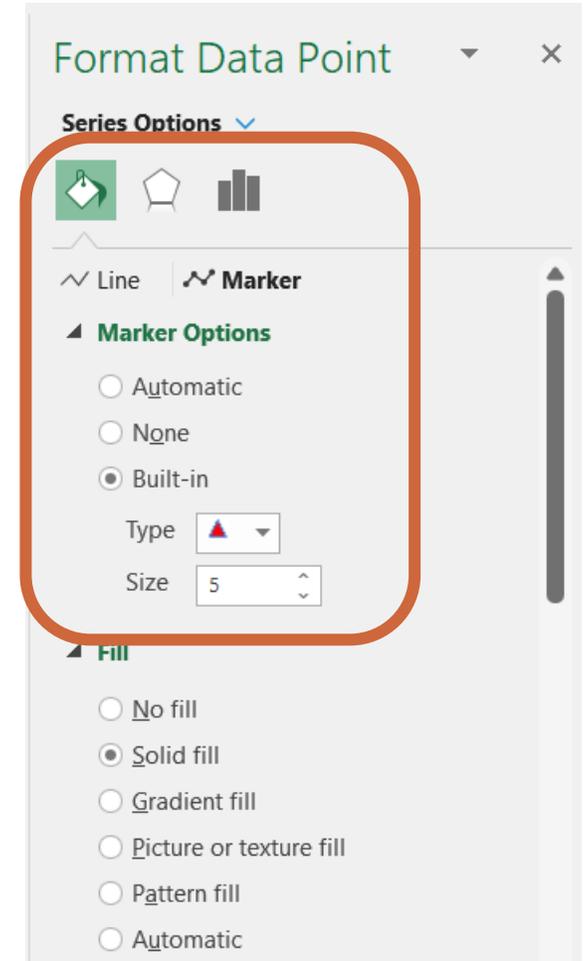
Format data points

You will now have the Format Data Point pane open.

To amend the marker,

1. Click on the **Paint Pot symbol**
2. Click on **Marker**
3. Click on **Marker Options**

You can now change style and colour of the markers.

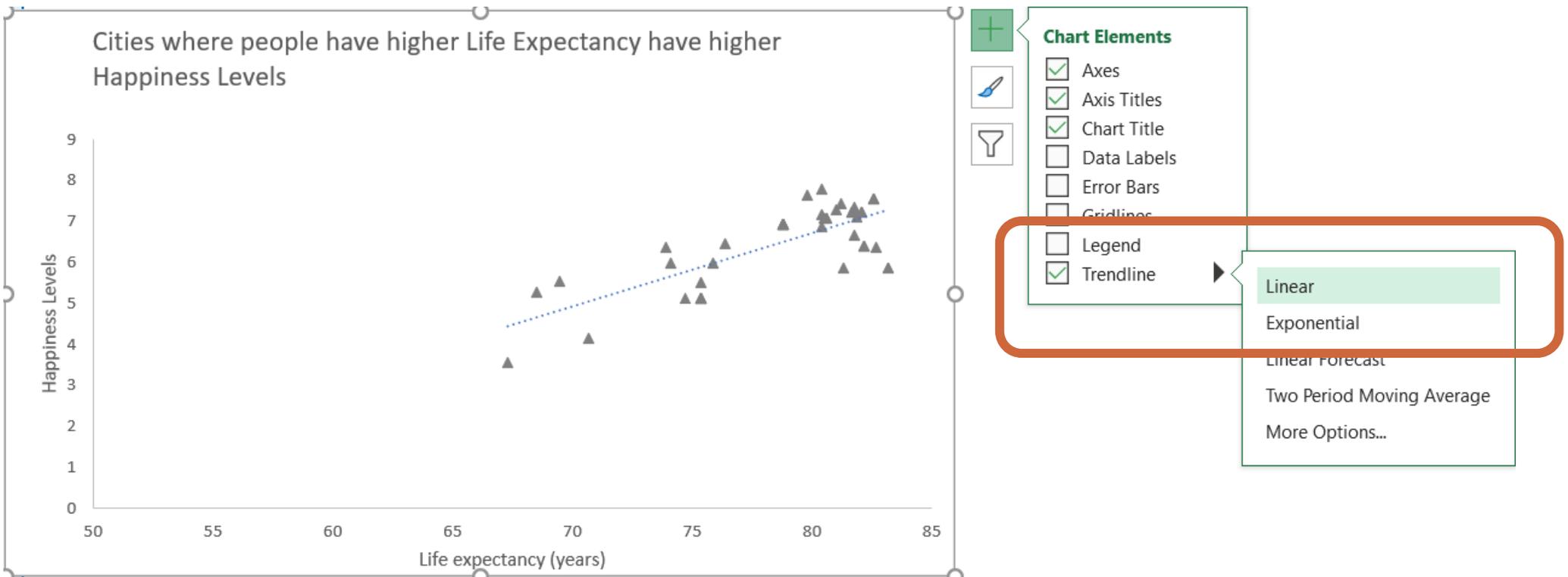


Add a trend line

Reminder

A trendline shows you the pattern of the data

To add a trend line to a scatterplot you can **click on the Trendline box in the Chart Elements**. The colour and width of the line can be amended in the same way as the line on a line graph.



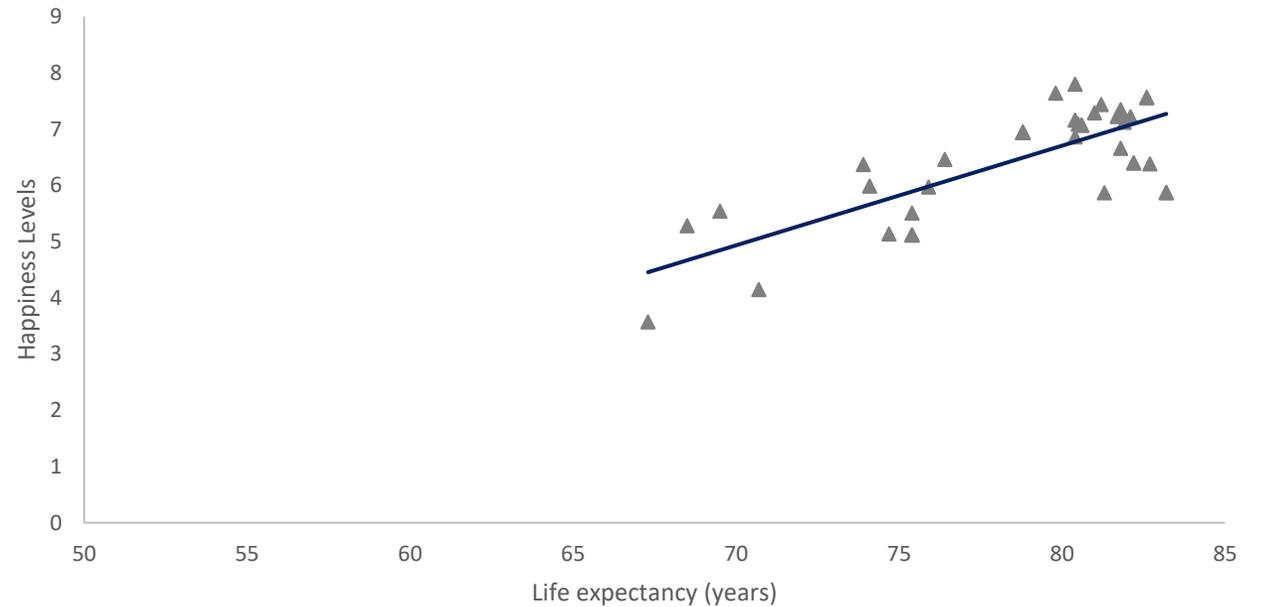
Amend the chart title

Reminder

Descriptive graph titles allow key messages to be easily understood

As with all the Excel graphs you can amend the chart title by clicking on the title and typing in your new title.

Cities where people have higher Life Expectancy also have higher Happiness Levels



Next steps

Complete **questions 1 to 8**
in **section 3** of the
'Practise creating graphs
in Excel (part 2)' workbook.

Then complete questions 1 and 2 in **section 4**.

Then complete the **Extension**.

Learning checklist

I can *change* the font, colour and display format of graphs in Excel

I can *add/remove* the gridlines and legends on graphs in Excel

I can *plot* line graphs without a date value variable

I can *add* data labels to a line graph

I can *change* the style and colour of data points on a scatterplot

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

