Creating scatterplots of bivariate data in Excel using annual mean sea level data

This tutorial provides clear step-by-step instructions describing how to:

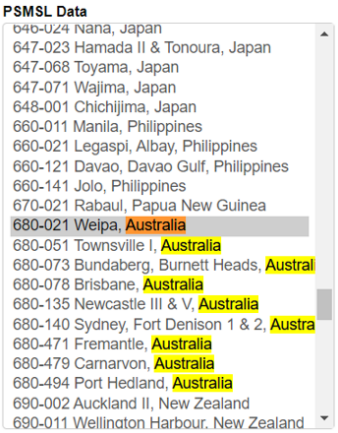
A: Import annual mean sea level data available from the tides and currents website of the US National Oceanic and Atmospheric Administration (NOAA) into MS Excel

B: Cleanup and prepare the data.

C: Create a scatter plot using sea level data and plot the line of best fit (or trendline).

A: Import sea level data into Excel

1. Go to https://tidesandcurrents.noaa.gov/sltrends/sltrends_global.shtml
2. Select a station of interest from the Permanent Service for Mean Sea Level **(PSMSL) Data list**. In this tutorial, we will use Australian sea level data from the **Weipa (QLD)** station. Press **Ctrl F** (i.e. search command) and type 'Weipa' to quickly search for the station within the webpage (Hint: Weipa is station number 680-021 and stations are listed in numerical order).



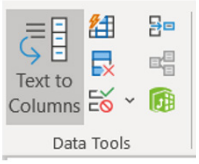
3. A new page will open. Find and click the link Source Data & Additional Metadata below the first plot.
4. You will be redirected to the information page of the **Weipa** station
5. Scroll down to the **Tide Gauge Data** section, and click on Download annual mean sea level data.
6. The data will show on a new page. Press **Ctrl A** followed by **Ctrl C** to select and copy the data. Open Excel, then click on the first cell (**A1**) and paste the data. Remember to save the fil

	A	B
1	1966; 6766;N;000	
2	1967; 6824;N;000	
3	1968;-99999;N;000	
4	1969;-99999;N;000	
5	1970; 6853;N;000	
6	1971; 6976;N;000	
7	1972; 6796;N;000	
8	1973;-99999;N;000	
9	1974;-99999;N;000	
10	1975;-99999;N;000	
11	1976;-99999;N;000	
12	1977;-99999;N;000	
13	1978;-99999;N;000	
14	1979;-99999;N;000	
15	1980;-99999;N;000	
16	1981;-99999;N;000	
17	1982;-99999;N;000	
18	1983;-99999;N;000	

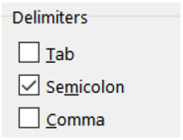
MONASH CLIMATE CHANGE COMMUNICATION RESEARCH HUB
climate extremes
ARC centre of excellence

B: Clean up and prepare the data

7. The data needs some formatting before you can start using it. First you need to split the data into columns using the separator 'semi column': Select Column **A**, by clicking on the header **A**. Then select **Text to Columns** from the **Data** tab.



8. The **Convert Text to Columns Wizard** will open, make sure that Delimited is selected then click Next > . The data is separated with semicolon, so make sure that the **Semicolon** delimiter is checked



Click Next > then Finish .

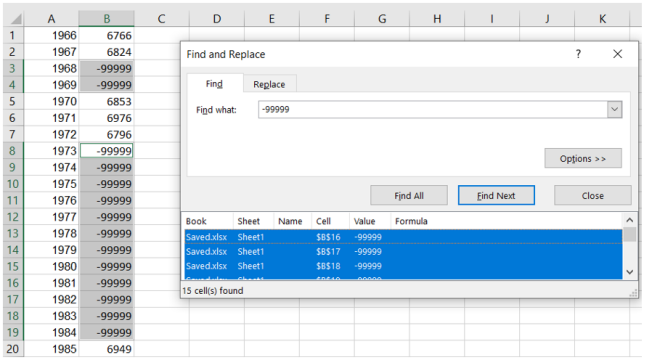
You will see the data split into 4 columns.

	A	B	C	D
1	1966	6766	N	0
2	1967	6824	N	0
3	1968	-99999	N	0
4	1969	-99999	N	0
5	1970	6853	N	0
6	1971	6976	N	0
7	1972	6796	N	0
8	1973	-99999	N	0
9	1974	-99999	N	0
10	1975	-99999	N	0
11	1976	-99999	N	0
12	1977	-99999	N	0
13	1978	-99999	N	0
14	1979	-99999	N	0
15	1980	-99999	N	0
16	1981	-99999	N	0
17	1982	-99999	N	0

9. Your bivariate data consists of years in Column **A** and values of annual mean sea level (in millimeters) in Column **B**. You can now delete the text in Columns **C** and **D** by selecting these columns and pressing **Delete**. Your data has now two columns (Hint: You might like to save the file again now in case you make a mistake in the next step and need to come back to this point).

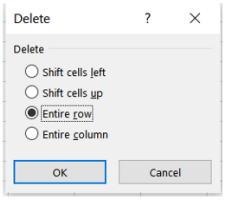
10. The value **-99999** indicates no data. Sometimes the tide gauge device breaks down and fails to take measurement, this is a common issue in monitoring the environment using instruments. You need to remove all rows containing **-99999**: Press **Ctrl A** to select all the data followed by **Ctrl F** to open the **Find and Replace window**. Enter the value **-99999** in the search box then click Find All . This will show a list of all cells containing the value **-99999**.

Click on one of the results that appear below the search box, then press Ctrl A. All the results should be highlighted now. Also, you should notice that each cell in Column **B** containing **-99999** will be selected.



Click Close on the **Find and Replace** window.

Now, press **Ctrl –** to open the **Delete** window.



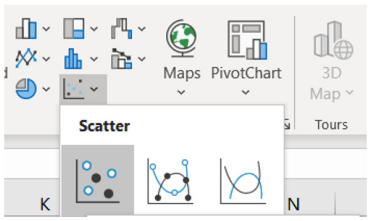
Select the **Entire Row** option, and click OK .

All rows containing **-99999** should be gone now. It's a good idea to save the file again now.

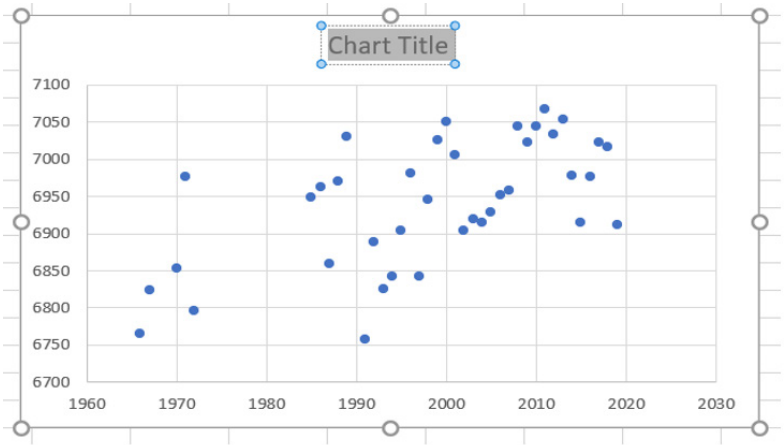
	A	B
1	1966	6766
2	1967	6824
3	1970	6853
4	1971	6976
5	1972	6796
6	1985	6949
7	1986	6962
8	1987	6859
9	1988	6970
10	1989	7031
11	1991	6758
12	1992	6889
13	1993	6825
14	1994	6843
15	1995	6904
16	1996	6981
17	1997	6843

C: Create a scatter plot using sea level data and plot the trendline

11. Select columns **A** and **B**, then in the Insert tab, click on the **Charts** icon. A small window will open, select **Scatter**.

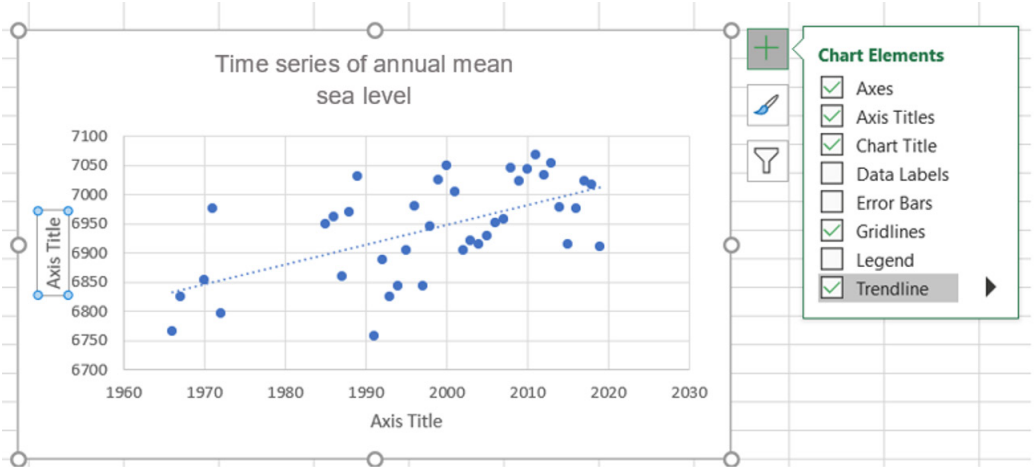


This will show a scatterplot of the data in your spreadsheet. The years will appear on the x-axis and the *Annual mean sea level data* will appear on the y-axis.

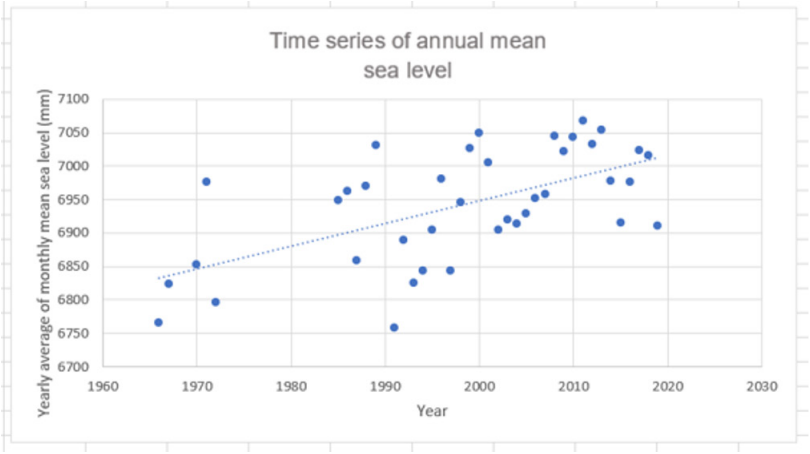


12. Now you need to format your plot.

Press + to add **Axis Titles** and a **Trendline**.



Rename the **Axis titles.**



13. To add the equation of the trendline: Select the trendline to see the **Format Trendline** pane.

Press , then check Display Equation on chart / Display R-squared value on chart to see the equation of the trendline and the coefficient of determination R^2 (optional).

Click on the trendline equation and move it somewhere in the chart space to make it more visible.