



WEEK 1 — TEXT FUNCTIONS

The focus this week was on using Excel functions to manipulate and clean text data. We looked at a lot of different functions. This handy quick reference guide includes all functions we looked at, useful keyboard shortcuts and some other functions worth investigating.

TEXT FUNCTIONS AND OPERATORS

CONCATENATE (text1, text2, ...)	Joins two or more text strings into one string
& e.g. =A1&A2&A3	Joins two or more text strings into one string
CONCAT (text1, [text2],...)	Combines the text from multiple ranges and/or strings
TEXTJOIN (delimiter, ignore_empty, text1, [text2], ...)	Combines text from multiple ranges and/or strings, and includes the option to specify a delimiter
LEFT (text, num_chars)	Returns a specified number of characters from the left of a text string.
RIGHT (text, num_chars)	Returns a specified number of characters from the right of a text string.
MID (text, start_num, num_chars)	Returns a specified number of characters from the middle of a text string.
FIND (find_text, within_text, [start_num], [NotFound])	Returns the starting position of one text string within another text string, (is case sensitive)
LEN (text)	Returns the number of characters in a text string
UPPER (text), LOWER (text), PROPER (text)	Change case to uppercase, lowercase and proper case
CLEAN (text)	Removes the first 32 nonprinting characters in the 7-bit ASCII code
TRIM (text)	Removes all spaces from text except for single spaces between words
SUBSTITUTE (text, old_text, new_text, [instance_num])	Substitutes new_text for old_text in a text string (is case sensitive)

FORMULA TIPS

- Start with =
- Function name always followed by bracket
- Close all brackets
- Comma separate arguments
- Text goes in "quotes"
- Nest multiple functions

=UPPER(CONCAT("Hi ", B2))

SHORTCUTS

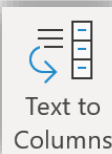
Windows	Mac	
Ctrl+C	⌘+C	Copy
Ctrl+V	⌘+V	Paste
Ctrl+S	⌘+S	Save
Ctrl+Z	⌘+Z	Undo
Ctrl+Y	⌘+Y	Redo
Ctrl+N	⌘+N	New
Ctrl+O	⌘+O	Open

BONUS CONTENT

SEARCH - returns the location of one text string inside another. Unlike FIND, SEARCH allows wildcards, and is not case-sensitive. Syntax: =SEARCH (find_text, within_text, [start_num])

REPLACE - replaces part of a text string, based on the number of characters you specify, with a different text string. Syntax: =REPLACE(old_text, start_num, num_chars, new_text)

You may also want to investigate: **CODE**, **UNICODE**, **CHAR** and **UNICHAR**



Text to Columns

You can take the text in one or more cells, and split it into multiple cells using the Convert Text to Columns Wizard



WEEK 2 — NUMBERS AND DATES

This week we have continued our work cleaning up data. This time we concentrated on converting between numbers and text, and working with dates. These are common issues that you might have to deal with when working with data in Excel.

NUMBER AND DATE FUNCTIONS	
VALUE	Convert text to a number. Gives a #VALUE! error if the text cannot be converted.
TEXT	Format a number as text with a given format. Remember that dates are stored as numbers, so they can be used too.
NOW	Return the current date and time. Updates every time the workbook is calculated.
TODAY	Return the current date. Updates every time the workbook is calculated.
DATE	Create a date given year, month, and day.
DAY	Return the day of a given date.
MONTH	Return the month of a given date.
YEAR	Return the year of a given date.
DAYS (2013)	Returns the number of days between 2 dates.
WORKDAY	Find the next workday before or after a given date. This excludes weekends and holidays.
NETWORKDAYS	Find the number of workdays between 2 dates.
EOMONTH	Find the last day of a month a certain number of months before or after a given date.
EDATE	Move a certain number of months before or after a given date.

SHORTCUTS

Windows	Mac	Description
Ctrl+;	Control+;	Inserts today's date as a fixed value. Note that this is different to the =TODAY() function because this date is fixed and will not change.
Ctrl+Shift+;	⌘+;	Inserts the current time as a fixed value. Note that this is different to the =NOW() function because the time is fixed and will not change.

DATE FORMAT CODES

Assume we are working with July 31, 1980.

d	31	m	7	yy	80
dd	31	mm	07	yyyy	1980
ddd	Thu	mmm	Jul		
dddd	Thursday	mmmm	July		

WORKDAY.INTL and NETWORKDAYS.INTL

WORKDAY and **NETWORKDAYS** both assume that Saturday and Sunday are the weekend days. Using the **INTL** versions of these functions you can specify which days of the week are non-workdays. These functions were introduced in Excel 2010.



WEEK 3 — NAMED RANGES

The focus this week has been on working with Excel's named ranges which make our formulas easier to create and understand and pave the way for better automation. We also had a look at using Data Validation to create drop down lists. Here is a quick reference guide.

TOOLS	
Named_Cell ▾	Use the name box to quickly name a range that does not have a label or consists of multiple columns and rows.
Create from Selection	Use Create from Selection to very quickly name a range or several ranges using the labels in the workbook.
Define Name ▾	Use Define Name to create a name where you need to specify scope.
Name Manager	Use the Name Manager to create, view, modify, and delete defined names.
Use in Formula ▾	Once you have defined some names you can use this to help include names in formulas.
Data Validation ▾	Data Validation allows you to specify rules for what values are allowed in a cell. Named ranges are useful if you have a list of values.

RULES FOR NAMING RANGES

- Up to 255 characters long
- Must start with a letter or an _
- Can contain letters, numbers, _, and .
- Must not contain spaces
- Must be unique within scope
- Must not be a cell reference

SHORTCUTS

F4	F4	Cycle the formula reference style
F3	F3	Paste Names
Ctrl+F3		Open Name Manager
Ctrl+Shift+F3	⌘+Shift+F3	Create from Selection
	⌘+F3	Open Define Name dialog

FORMULA REFERENCE STYLES

	Column relative	Column absolute
Row relative	① A1 — both column and row will change when copied	④ \$A1 — column will remain fixed and row will change
Row absolute	③ A\$1 — row will remain fixed and column will change	② \$A\$1 — both column and row will remain fixed

Pressing F4 will cycle through ①→②→③→④→①.

BENEFITS OF NAMED RANGES

- They can make formulas easier to understand and create
- Behave like absolute references
- Help with automation



WEEK 4 — TABLES

We have seen this week how we can use Tables in Excel to help structure, format, and automate working with data. You can easily find the data that you are looking for using sorting and filtering. Structured References are a new and powerful way to write formulas.

STRUCTURED REFERENCES

TableName [ColumnName]	This is the name of the table. It can be set in the Table Design tab.
TableName[ColumnName]	This selects all the data in a column from a table.
TableName[@ ColumnName]	This refers to a single cell in the current row in another column.
TableName[#Headers],[ColumnName]	Select the header from a column.
TableName[#All],[ColumnName]	Select the header and data from a column.
TableName[#Totals],[ColumnName]	Select the totals from a column. Returns null if there are no totals for that column.

WHAT MAKES GOOD DATA FOR A TABLE

Your data should be rectangular, with no empty rows or columns. However, it is a good idea to have empty space between your table and any other data on the worksheet. It is useful if you have meaningful column headers to help distinguish each column.

SUMMARY FUNCTIONS FOR THE TOTAL ROW

When you enable the **Total Row** in a table you are given several options for the function to use to summarize each column. The default list comes from the **SUBTOTAL** function, although you can use any function or write your own. Using the **SUBTOTAL** function has the benefit that it will ignore hidden values if you hide any rows of the table, perhaps by using a table filter.

RULES FOR NAMING TABLES

- The same rules as for Named Ranges
- You might want to start names with **tbl_** to differentiate them from Named Ranges

SHORTCUTS

Windows	Mac	Action
Ctrl+T	⌘+T	Insert a Table
Shift+Space	Shift+Space	Select table row
Ctrl+Space	Control+Space	Select table column
Ctrl+A	⌘+A	Select entire table

BONUS CONTENT

When you drag a formula that includes a structured reference to another column in the table, you will notice that the column references behave like relative references — they move with the formula. If you want them to be more like absolute references you can use:

TableName[[ColumnName]:[ColumnName]]



WEEK 5 — LOGICAL & LOOKUP FUNCTIONS

The focus this week was on using Excel's logical and lookup functions to manipulate values, categorise data and match data sets. This handy quick reference guide includes all functions we looked at, a list of logical operators and some other functions worth investigating.

LOGICAL AND LOOKUP FUNCTIONS

IF (logical_test, value_if_true, [value_if_false])	Return different values depending on a test.
AND (logical1, [logical2], ...)	Combine several tests. Will return true only if all the tests are true.
OR (logical1, [logical2], ...)	Combine several tests. Will return true if any of the tests are true.
VLOOKUP (lookup_value, table_array, col_index_num, [range_lookup])	Look up a value in a range of cells. Return the value in the same row but a different column.
INDEX (array, row_num, [column_num])	Returns the value from the array given by position row_num and column_num. Often used with MATCH to find the row and column positions.
MATCH (lookup_value, lookup_array, [match_type])	Look up a value in an array. Returns the position of the value in the array.

XLOOKUP

XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])

XLOOKUP is the newest lookup function in Excel. It only became available in the Office 365 version of Excel in early 2020. It combines the ease-of-use of **VLOOKUP** with the power of **INDEX/MATCH**. One limitation of **VLOOKUP** is that it can only look up values in the leftmost column of the array. With **XLOOKUP** you specify the lookup array and return arrays separately.

There is also a new **XMATCH** function that expands on the **MATCH** function.

LOOKUP TIPS

- Name lookup ranges
- Convert lookup data to tables
- Specify if it is a **Range** lookup or **Exact Match**
- Keep formulas as simple as possible

LOGICAL OPERATORS

=	Equal to
<	Less than
<=	Less than or equal
>	Greater than
>=	Greater or equal
<>	Not equal

BONUS CONTENT

IFS — The **IF** function only allows 1 logical test. We have seen how we can nest **IFs** if you need to perform multiple tests, which can get messy with several tests. Another option is the **IFS** function, introduced in recent versions of Excel. It can perform multiple tests, returning the value associated with the first test that returns true.

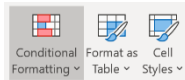
IFERROR/IFNA — sometimes you want to perform a calculation that might produce an error. Using these functions, you can return a specified value if your calculation produces an error. **IFERROR** traps all errors, **IFNA** only traps a **#N/A** error.



WEEK 1 - VISUALISATION TOOLS

This handy quick reference covers the main topics from Week 1: Conditional Formatting, Macros, Sparklines and Custom Number Formats.

CONDITIONAL FORMATS



Conditional formatting changes the way a cell looks if it meets certain criteria. To apply, select the cells, go to the Home ribbon and click Conditional Formatting.

Highlight Cell Rules	Value Based Formatting Rules
Greater Than	All values under a specified value
Less Than	All values over a specified value
Between	All values between two specified values (inclusive)
Equal to	All values equal to a specified value
Text that Contains	Text that contains specific characters
A Date Occurring	Yesterday, Today, Tomorrow, In the last 7 days, Last/This/Next week, Last/This/Next Month
Duplicate Values	All values where the value occurs more than once
Top/Bottom Rules	Value Based Formatting Rules
Top 10 Items	The top n items where n is a number you provide
Top 10 %	The top n % of items where n is a number you provide
Bottom 10 Items	The bottom n items where n is a number you provide
Bottom 10 %	The bottom n % of items
Above Average	All numbers below the average
Below Average	All numbers above the average
Data Bars	Trend based format shows a bar that represents the size of the value compared to the rest of the data
Color Scales	Trend based format applies an icon that represents the size of the value compared to the rest of the data
Icon Sets	Trend based format applies a colour that represents the size of the value compared to the rest of the data

Formatting a row with Conditional Formatting:

When applying conditional formatting, you can use formulas that reference other cells. You can use this to format an entire row based on the values in one column. Select all cells where format may apply, select New Rule > Formula and apply a mixed reference, e.g. **\$B3**.

RECORDED MACROS

1. Enable the Developer Tab: File > Options > Customise Ribbon
2. Click Record Macro
3. Enter macro name and optional shortcut
4. Perform actions you want to automate
5. Click Stop recording
6. To run, click Macros > Run (or shortcut)
7. To view/edit code click Visual Basic button or Alt+F11

Start with Sub Sub ACT()
 Comment ' Keyboard Shortcut: Ctrl+Shift+A
 Select a cell Range("K1").Select
 Change value in active cell ActiveCell.FormulaR1C1 = "ACT"
 Range("B4").Select
 Finish with End Sub End Sub

CUSTOM NUMBER FORMATS

Allow us to customise how numbers display. To apply a custom number format, first select cell(s) to be formatted then type Ctrl+1, select Numbers > Custom and type in the new format. Each format can have up to four sections, separated with semi-colons as follows:

1 2 3 4
[Blue]0.00 ; [Color26]-0.00 ; 0"mm" ; @" sales"

This will format different values differently:

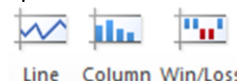
1. Will apply to positive numbers
2. Will apply to negative numbers
3. Will apply to zero values
4. Will apply to text values

Before	After
56.3222	56.32
-56.3222	-56.32
0	0mm
Q1	Q1 sales

You are not limited to these ranges, you can apply your own, e.g.: **[>100][Magenta]♥**; will show values >100 as ♥ and ≤100 as blank cells.

SPARKLINES

Tiny charts that fit in a cell.
3 options on Insert ribbon



SHORTCUTS

	Windows	Mac
Format Cells	Ctrl+1	⌘+1
Macro dialog	Alt+F8	Option+F8
VBA Editor	Alt+F11	Option+F11

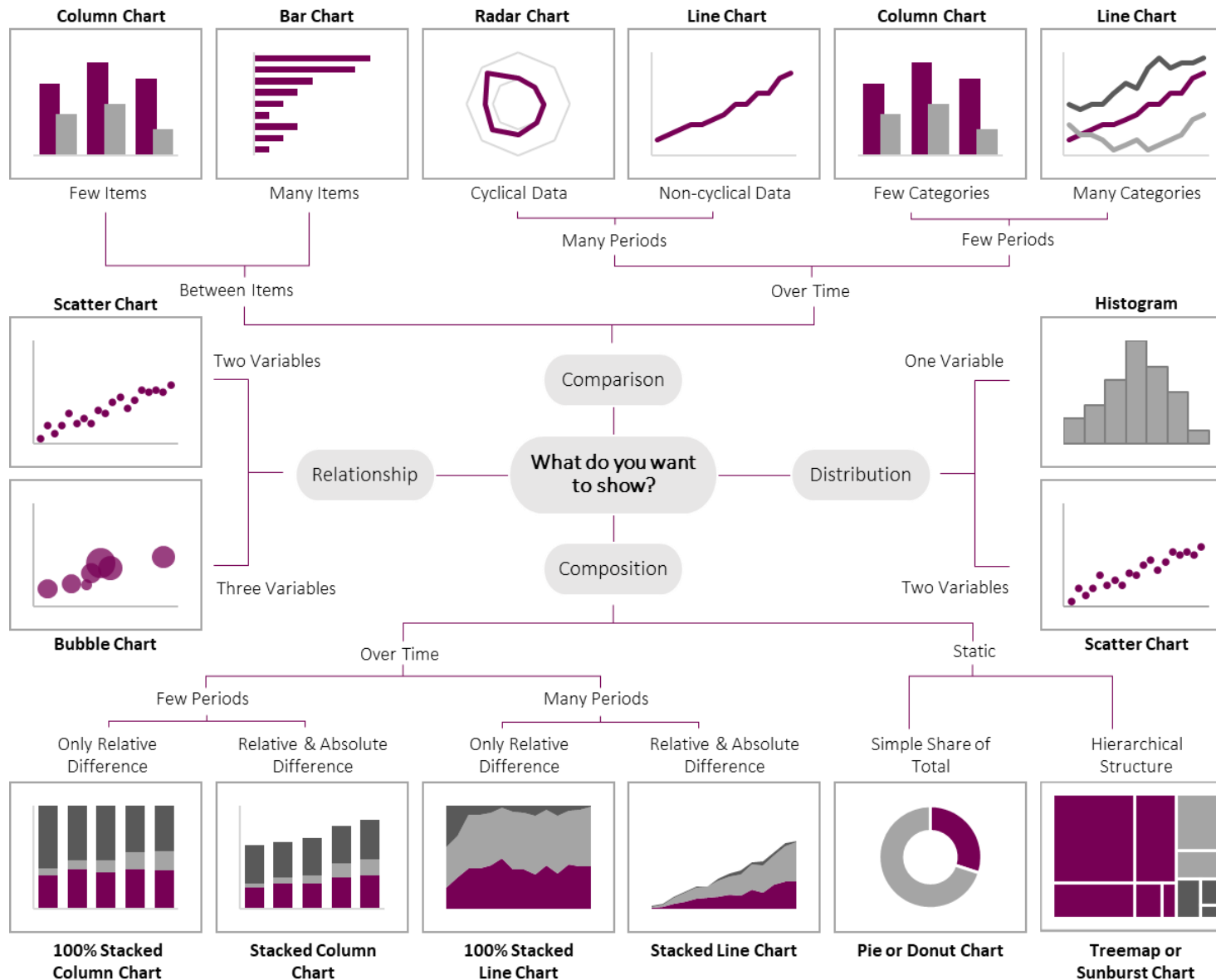
CUSTOM NUMBER COLOUR CODES

1 ▲▼	29 ▲▼
2 ▲▼	30 ▲▼
3 ▲▼	31 ▲▼
4 ▲▼	32 ▲▼
5 ▲▼	33 ▲▼
6 ▲▼	34 ▲▼
7 ▲▼	35 ▲▼
8 ▲▼	36 ▲▼
9 ▲▼	37 ▲▼
10 ▲▼	38 ▲▼
11 ▲▼	39 ▲▼
12 ▲▼	40 ▲▼
13 ▲▼	41 ▲▼
14 ▲▼	42 ▲▼
15 ▲▼	43 ▲▼
16 ▲▼	44 ▲▼
17 ▲▼	45 ▲▼
18 ▲▼	46 ▲▼
19 ▲▼	47 ▲▼
20 ▲▼	48 ▲▼
21 ▲▼	49 ▲▼
22 ▲▼	50 ▲▼
23 ▲▼	51 ▲▼
24 ▲▼	52 ▲▼
25 ▲▼	53 ▲▼
26 ▲▼	54 ▲▼
27 ▲▼	55 ▲▼
28 ▲▼	56 ▲▼



WEEK 2 — MASTERING CHARTING TECHNIQUES

The focus for week 2 and 3 is on working with Excel charts. Below is a chart planner to help determine the best choice of chart(s) to tell your data story.



TRENDLINES



Exponential: When values rise or fall at increasingly higher rates. (No zero or negative values allowed)



Linear: Used for straight lines to show steady increase or decrease.



Logarithmic: Use when rate of change in the data increases or decreases quickly and then levels out.



Polynomial: Use when data fluctuates. (Good for analysing gains and losses over a large data set)



Power: Use with data sets that increase at a specific rate. (No zero or negative values allowed)



Moving average: Tends to smooth out fluctuations in data to show a pattern or trend more clearly.



more accurate

less accurate

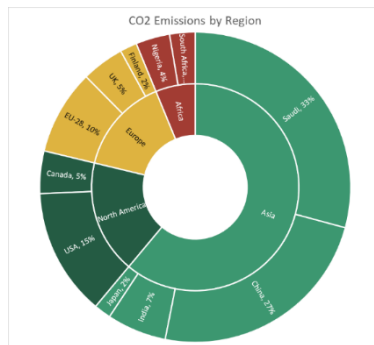
SHORTCUTS

	Windows	Mac
New Chart Sheet	F11	Fn+F11
Embedded Chart	Alt+F1	
Chart Properties	Ctrl+1	⌘+1



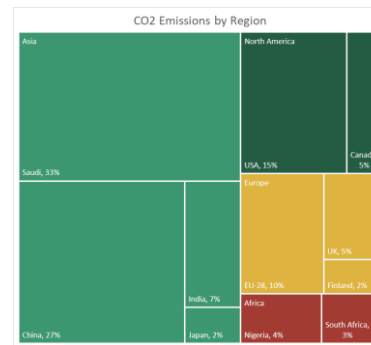
WEEK 3 — SPECIALISED CHARTS

This week we covered some more specialised charts that can display your data in eye-catching ways to help you tell the story of your data.



HIERARCHY CHARTS

We looked at 2 hierarchy charts: **sunburst** and **treemap**. Hierarchy charts are useful when you have higher level categories of data that contain sub-categories. You can visualise how each sub-category contributes to the whole. These were introduced in Excel 2016.



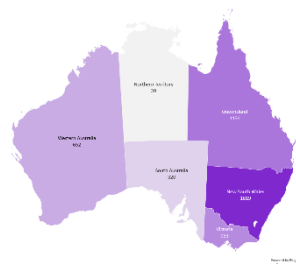
WATERFALL CHARTS



Useful to show a running total and what factors have contributed to that total. Often used to show how income and expenses contribute to a budget, or to show flows between categories. Remember to right-click on the Total columns and **Set as Total**.

Useful to show a running total and what factors have contributed to that total. Often used to show how income and expenses contribute to a budget, or to show flows between categories. Remember to right-click on the Total columns and **Set as Total**.

MAP CHARTS

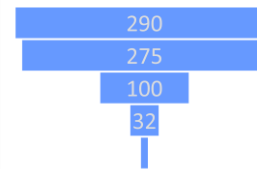


Used to show data that has a geographical component. Excel can work with a variety of geographical regions, like countries, states, counties or postal codes. However, sometimes it needs help to differentiate between

regions. You can include another column with the parent region, for example if you are working with states, include a Country column, even if they are all in the one country. Using proper column headings is also important.

FUNNEL CHARTS

Shows how values change, typically reduce, through stages of a process. Often used in Sales and Marketing to look at how many people enter the sales pipeline compared with how many complete the sale.



GAUGE CHARTS

A custom chart that shows how you can combine Excel's built-in charts in a creative way. Great to show a snapshot of a key metric.

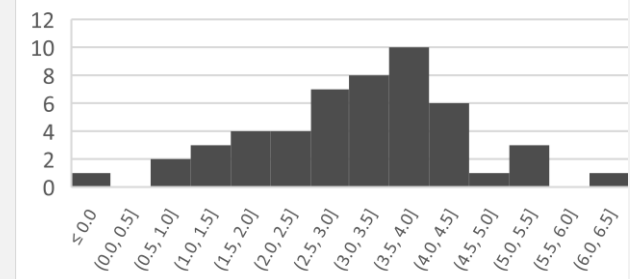


- Change angle of first slice to 270°, set to transparent
- Add pie on the second axis, change angle to 270°, everything except the pointer transparent

Data		Gauge
Value to show	78	2 Gauge section 1
Pointer	1	1 Gauge section 2
Gauge max – val+ptr	=100-(78+1)	1 Gauge section 3
Hidden half	100	=2+1+1 Hidden half

BONUS CONTENT — HISTOGRAMS

Bar and column charts are useful to show the number of items in categories, but they do not work with continuous (numeric) data on the x-axis. Histograms are useful here.



The numeric data is grouped into “bins” and then the number of values that fall into each bin is displayed on the chart. The bins can be adjusted in the Format Pane.

Histograms are a standard chart in Excel 2016 and newer. They can be created in older versions using the **Analysis ToolPak addin**.

MACROS — CODE SNIPPETS

```
Sub ChartAnimate()  
    Dim nCount As Integer  
    For nCount = 1996 To 2019  
        DoEvents  
        ' start the subroutine  
        ' create a variable for the year  
        ' loop over years 1996 to 2019  
        ' pause the macro quickly  
        ' set the dropdown cell to the year of interest  
        ActiveWorkbook.Worksheets(6).Range("I3") = nCount  
        DoEvents  
        ' pause the macro again  
        Next  
        ' go to the next year  
    End Sub  
    ' end the subroutine
```




WEEK 4 — PIVOTTABLES & CHARTS

This week we covered PivotTables and PivotCharts. These are quick, easy, and powerful ways to investigate your data.

To create a PivotTable, click in your data (preferably a Table), **Insert > PivotTable**.

The 4 “quadrants” or areas of a PivotTable are **Filters**, **Rows**, **Columns**, and **Values**.

Year	2019		
Sum of Electricity Capacity (MW)	Column Labels		
Row Labels	Non-Renewable	Renewable	Grand Total
North America	973,689.73	391,240.68	1,364,930.41
Africa	185,620.44	48,442.54	234,062.98
Eurasia	253,311.18	105,799.62	359,110.80
Europe	640,201.58		92
Middle East	297,772.17		98
Asia	2,245,950.87	1	88
Oceania	49,251.53	39,722.98	88,974.51
South and Central America	136,842.23	236,702.16	373,544.39
Grand Total	4,782,639.73	2,536,853.14	7,319,492.87

SORTING AND FILTERING

The sorting and filtering dropdown allows you to sort by that field or apply filters. Different filters are available depending on the type of data. You can also change the order of items manually by dragging them to a new position.

Add a Field to the Filter quadrant to have it affect the entire table. Use Report Filter Pages to create a sheet for each value in a Filter.

CALCULATED FIELDS AND ITEMS

Rather than add a column to your data, you can do the calculation in your PivotTable. Calculated Fields adds a new Field to your data. Calculated Items add a new item (or several) to a Field. You can find these on the **PivotTable Analyze** ribbon in the **Calculations** area.

You can drag fields from the Field List to the quadrants, or between quadrants, to rearrange the PivotTable. A field can appear multiple times in the **Values** quadrant with a different summary function or calculation. Click the drop-down next to the field in the **Values** quadrant to access the **Value Field Settings** or right-click the field in the PivotTable.

GROUPING ITEMS

When you add a Date field to your PivotTable, it will get automatically grouped. You can change or remove the groups, or create groups for other fields, on the right-click menu or on the **PivotTable Analyze** ribbon in the **Group** area.

PIVOTCHARTS

A PivotChart is a chart based on data in a PivotTable. It will update if the data or structure of the PivotTable changes. You can use the Field Buttons to sort and filter the chart or turn them off for a better look when printed or in a dashboard.

REFRESH

If the data behind the PivotTable changes you can click Refresh to update the table with the new data.

SHOW VALUE AS

There are 4 main types of calculations you can **Show Value As**:

% of — the options here include **% of Grand Total**, **% of Column Total**, **% of Row Total**, and others. In the example to the left, **% of Row Total** would be the percent of energy generated by renewable and non-renewable sources per region.

Difference From — show the difference between each item and a base item. Here we might want to set Europe as the base item, then all the other rows will show the difference between Europe and that region.

Running Total In — also known as the cumulative sum. Add each element to the running total to get the overall total in the last cell.

Rank — either rank smallest to largest or largest to smallest.

CONDITIONAL FORMATTING

Don't forget about conditional formatting! All the conditional formatting tools available to you in normal Tables are still available in PivotTables.

Number formats should be set in the **Field Settings** rather than using the tools on the **Home** ribbon.

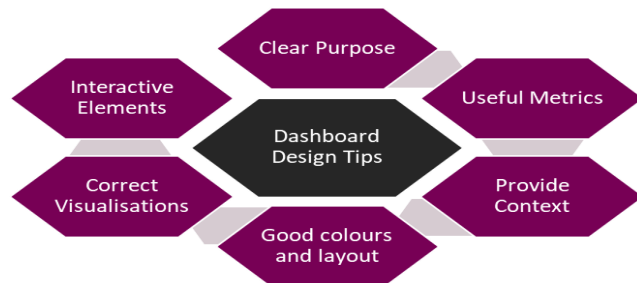


WEEK 5 — CREATE AN INTERACTIVE DASHBOARD

To wrap up the course we created an interactive dashboard to show key metrics at a glance and help make informed business decisions quickly.

DASHBOARD DESIGN

A dashboard should show the most important information on one screen. Choose a few key items to tell your story.



Start your design by wireframing using pencil and paper, markers and whiteboard or even sticky notes.

DASHBOARD LAYOUT

Dashboards are usually laid out using a series of tiles, these can be created by changing the colour of groups of cells but using shapes can give you greater control over where each element on the page lives. There are a wide variety of different shapes at **Insert > Shapes**.

THEMES AND COLOURS

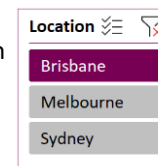
With themes, you can change your dashboard colour scheme by selecting a different theme. However, themes can also change the font, which can affect the layout. Instead, you can just change the colour scheme using the **Colors** dropdown on the **Page Layout** Ribbon.

HYPERLINKS

You can use hyperlinks to link to a different place within an Excel workbook, allowing for easy navigation. You can link to a cell reference or defined name. To create a hyperlink, **right-click > Link** or **Insert > Link**.

SLICERS

A slicer allows you to filter one or more PivotTables or PivotCharts at once. You can use any fields in the data set. There is also a special type of slicer called a timeline slicer that filters by date.

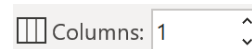


To create a slicer, click in the PivotTable or PivotChart and select **Insert Slicer** from the **PivotTable/Chart Analyze** Ribbon. You can create as many slicers as you have fields.



On the **Slicer** Ribbon select from the style options to change the colours of your slicer or you can create a custom style to match your dashboard by either duplicating an existing style and modifying it or create a new style.

Use **Slicer Settings** to display or hide the header and items with no data, among others. On the slicer ribbon, you can also adjust the number of columns in a slicer.



Use **Report Connections** to connect the slicer to several PivotTables.

LINKED IMAGES

Some items, like SparkLines, cannot be put into Shapes. Use **Paste Special > Linked Picture** to show a live snapshot that behaves just like the real thing.



DESIGN BEST PRACTICE

Choosing where to place your metrics and visualisations is an art, but there are some guidelines to help.

- If most of your readers use a left-to-right language, then they will probably start looking at your dashboard from the top-left, move right, and then move down (like the letter Z). Position your key metrics in the top row. This is called Z-pattern design. There is a similar pattern called the F-pattern. Both emphasise the top row being important.
- Group related items together.
- Use slicers or dropdown lists to allow your users to interact with the dashboard to focus on and highlight the areas that are important to them.
- Move detailed metrics to a separate sheet so they do not clutter the main dashboard. Use hyperlinks to allow users to move between the sheets.

CODE SNIPPETS

```
Sub RefreshAll()  
    ActiveWorkbook.RefreshAll  
End Sub
```

OR to only refresh PivotTables in the active sheet:

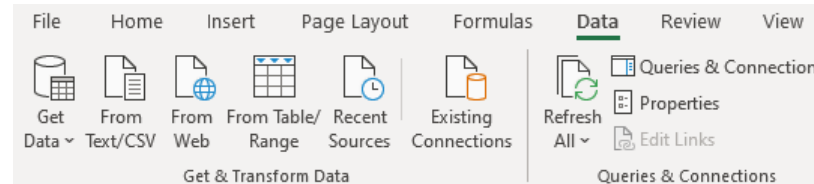
```
Sub RefreshPivotTables() 'Only refresh pivot tables  
    Dim pivotTable As PivotTable  
    For Each pivotTable In ActiveSheet.PivotTables  
        pivotTable.RefreshTable  
    Next  
End Sub
```



WEEK 1 — GET AND TRANSFORM (POWER QUERY) — PART 1

Get and Transform is a new and powerful way to import and manipulate data in Excel. This week is just the beginning.

You access **Get and Transform** on the **Data** Ribbon (it looks slightly different in older versions of Excel). Under **Get Data** you can select the source of your data. Common sources include a Table in the current workbook, an Excel or CSV file, or a database. You can also import all files in a folder.



CREATE A TABLE

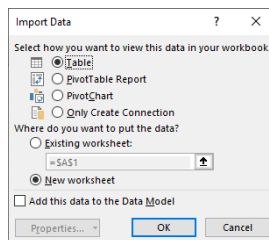
When creating a query from data in the same Excel file it will automatically be converted to a table. It is good practice to create the table yourself by going to **Insert > Table** or **Ctrl+T**. Rename the table on the **Table Design** ribbon, then create the query. The name will carry through to become the name of the query. If you do not wish to convert the data to a table, connect to it from a different workbook.

FORMULA BAR

Like the Excel workbook formula bar, the Power Query Editor formula bar allows you to view and edit the M code for the current formula. Make it visible by going to **View > Formula Bar**. You can use the f_x button to add a new custom step.

LOAD TO

When you import data, you can choose whether to include it as a table in your workbook or just create a connection. A connection is useful if the query is used as input to other queries, but you do not need to view it directly. In O365 you can also load the data directly to a PivotTable or PivotChart.



APPEND QUERIES



Append queries allow you to stack datasets on top of each other. Ensure that the tables have the same column names before appending them, otherwise Excel will not append correctly. The columns do not need to be in the same order, but the names must match.

MERGE QUERIES



You use Merge when you want to add tables side-by-side (add columns from one table to another). When you Merge, you specify which columns contain the values used to match between the tables, often ID or Name. You can match multiple columns by holding down Ctrl.

SPLIT AND MERGE COLUMNS

A common issue is having several pieces of information in 1 column. You can use Split Column to extract these into separate columns. There are many ways to split the data, depending on how your data is arranged. The opposite of this is Merge Columns, where you can take the contents of 2 or more columns and join them together with the separator of your choice. Like TEXTJOIN in Excel.

JOIN KINDS

When merging tables, you can choose what to do when there are rows that do not match between the tables. This is called the Join Kind. In older versions of Excel, you need to edit the M code of the merge to choose the Join Kind.

Left Outer



All rows from the first table and matching rows from the second table.

M: JoinKind.LeftOuter

Right Outer



All rows from the second table and matching rows from the first table.

M: JoinKind.RightOuter

Inner



Only rows that match in both tables.

M: JoinKind.Inner

Full Outer



All rows from both tables.

M: JoinKind.FullOuter

Left Anti



All rows in the first table which do not match the second table.

M: JoinKind.LeftAnti

Right Anti



All rows in the second table which do not match the first table.

M: JoinKind.RightAnti

RENAME A QUERY

Just like with tables, it is good practice to give your query a suitable name. You can change it in the Query Settings when you are creating it, or right-click > Rename in the Queries & Connections panel.



WEEK 2 — GET AND TRANSFORM (POWER QUERY) — PART 2

Now that we have imported some data, we can start to clean and organise it to make it easier to analyse later.



KEEP AND REMOVE ROWS



The **Keep Rows** and **Remove Rows** buttons on the **Home** ribbon contain several options for choosing which rows to keep or remove. Often data will have extra rows at the top or a total row at the bottom, or blank rows, that you need to remove before analysis. **Keep Errors** and **Keep Duplicates** can be useful when you are troubleshooting your query.



PROMOTE HEADERS

Usually the first row of your data will contain the column headers. You can **Promote Headers** to get Power Query to recognise these as column headers.

FILTER AND SORT

To remove values from a column, click on the Filter dropdown to the right of the column header. Like in Excel, you can tick/untick specific values or you can use filters based on the data type of the column. You can also sort from this menu. Unlike Excel, sorting by a second column adds to the sort.

SPLIT COLUMN

Sometimes a column will contain multiple pieces of information and you need to extract them into separate columns. The **Split Column** tools can help. You can find them on the **Home** and **Transform** tabs.

If there is a character or characters that separate values, you can use **Split by delimiter**. Newer versions of Power Query allow you to split on lowercase to uppercase, or digit to non-digit.

REPLACE VALUES

You can use **Replace Values** on the **Home** or **Transform** ribbon to replace a value with another. In the **Advanced options** you can choose some special characters. You might need to replace nulls with 0 if you want to perform calculations with your data.



COLUMN FROM EXAMPLES

A powerful addition to recent versions of Power Query is **Column from Examples**. This is like **Flash Fill** in Excel but much smarter. Start typing what you want your new column to look like and Power Query will figure out how to make it from the existing columns. You sometimes need to give several examples to get the right transformation.



ADD CUSTOM AND CONDITIONAL COLUMNS

To create a column based on calculations you can add a **Custom Column**. Here you have all the power of the M language to create exactly what you want. **Conditional Columns** assist you to create columns based on a series of logical tests. You can use a **Custom Column** with **if/then/else** statements if you do not have this feature.



DATA TYPES



When you import some data, Power Query will try to guess what type of variable each column contains. You can also change them manually if you need to by clicking to the left of the column name or on the **Transform** ribbon.



PIVOTING DATA



There are 2 main ways to organise data in spreadsheets: values across the columns (*wide*) and values down the rows (*long*). For example, you might organise sales data with each month having its own column and the departments in the rows. This is a *wide* format. The equivalent *long* format would have columns Department, Month and Sales.

The preferred layout for data analysis is the *long* format but for reports often *wide* is easier to read. In Power Query, going from *long* to *wide* is **Pivoting**. Going from *wide* to *long* is **Unpivoting**. When you **Pivot** you have the choice of how to aggregate the values. **Sum** and **Count** are commonly used, and there is also the **Don't Aggregate** option, which keeps all the values.

Long

Market Type	Month	Total Sales
Farm Stand	January	1994.5
Mobile Market	January	814.15
Farm Stand	February	2606.25
Mobile Market	February	707.9
Farm Stand	March	3076.75
Mobile Market	March	1008.25

Wide

Market Type	January	February	March
Farm Stand	1994.5	2606.25	3076.75
Mobile Market	814.15	707.9	1008.25

TRANSFORM/ADD COLUMN RIBBON TABS

With a bit of exploration, you can see that the **Transform** and **Add Column** ribbon tabs contain some common items. The difference is that the **Transform** tools will modify the current column whereas the **Add Column** tools will create a new column and keep the existing column.



WEEK 3 — POWER PIVOT

Power Pivot and the data model bring big datasets and database-like modelling capability to Excel.

ENABLE THE POWER PIVOT ADDIN

Power Pivot comes with recent versions of Excel but is not enabled by default. To enable it, go to **File** ► **Options** ► **Add-ins**. In the **Manage** drop-down box, select **COM Add-ins**, and click **Go**. Tick **Microsoft Power Pivot for Excel**. Click OK. You should see a new **Power Pivot** ribbon tab.

THE DATA MODEL

The Data Model is where Excel stores big data and tables with relationships. Add data by clicking into your table then going to **Power Pivot** ► **Add to Data Model**, or in Power Query when you load the query: **Load To** ► **Add this data to the Data Model**.

DATA AND DIAGRAM VIEWS

In Power Pivot on the **Home** tab, you can switch between the **Data** and **Diagram** views. The **Data** view looks like a normal Excel workbook and is where you define calculated columns and measures. The **Diagram** view is where you define relationships between tables.

WORKING WITH RELATIONSHIPS

Excel may offer to auto-detect relationships or you can add them yourself. In the **Diagram** view, click on the field in one table that you want to link and drag to the field in the other table. Hover over a link to see which fields are related or double-click to see a detailed view. To delete, select the link and press **Delete**, or right-click and select **Delete**. Auto-detection will only work if the field names match.

HIERARCHIES

Hierarchies occur when a field is nested within another one. For example, day within month within year, or city within state within country. In the **Diagram** view, click **Create Hierarchy** at the top-right of a table panel. Drag columns into this to create levels in the hierarchy. Use as normal in a PivotTable. The **Expand** and **Collapse** buttons show more levels while **Drill Down** goes into the current level.

CALCULATED COLUMNS AND MEASURES

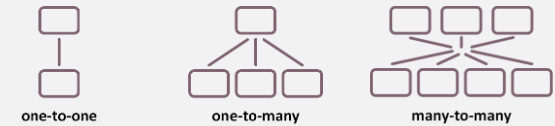
Both calculated columns and measures are created using DAX, but calculated columns add a new column to the dataset whereas measures are used when the table is summarised, for example in a PivotTable.

Implicit measures are used when you add a field to a PivotTable. Measures that you create are *explicit* measures and you have more control over how they are calculated and displayed.

DATE TABLES

If you want to do more complicated calculations with dates, for example if you want to use the DAX time intelligence functions, you will need a date table. Tools for working with these in Power Pivot are on the **Design** tab under **Date Table**. **New** will create one that you can then modify, or you can import one that you have already created.

TYPES OF RELATIONSHIPS



There are 3 main ways that data fields can relate to each other.

One-to-one: each UN representative country has 1 vote in the UN General Assembly.

One-to-many: one company can have many branch offices.

Many-to-many: a book can have many authors and an author can write many books.

When we are linking tables, generally we create one-to-many relationships. The field that is used to link between tables is called the *key*. In the “one” table this is the *primary key*, in the “many” table it is a *foreign key*.

DAX vs M

DAX (data analysis expression) is used in Power Pivot and Power BI to create calculated columns and measures. These calculations look like Excel workbook functions in tables, but the available functions are different.

M is used in Power Query and Power BI to filter and combine data from multiple sources. It does not look anything like Excel workbook functions.



WEEK 4 — POWER BI

Power BI is Microsoft's Business Intelligence tool that includes Power Query, DAX, and powerful visualisation tools.

GETTING AROUND IN POWER BI

There are 3 main views in Power BI: Report, [Data](#), and [Model](#). Report view is where you use visualisations to display your data. The Data view is an Excel-like view of the raw data. Model view allows you to review and modify the relationships between the data tables. When importing data, you have the familiar Power Query interface.

CALENDAR TABLES

To work with time intelligence functions, you need a calendar table. You can import one with Power Query, generate one using M, or create it with DAX. Remember to mark it as a date table and create relationships to the other tables.


CREATING VISUALISATIONS

With nothing selected in the report page, if you tick a field or measure the default visual for that field is added to the report. Or click on the visual you want from the Visualizations pane. To change a visual, select the visual in the report then click on the required visual in the Visualizations pane.

VISUALISING NUMBERS

If you want to display numbers rather than graphs, you can use the [Card](#), [Table](#), or [Matrix](#) visuals. A Card will display a single number. Tables display multiple columns of data that is not aggregated. A Matrix is like a PivotTable in Excel, so it will aggregate by category, and you can drill down into categories.


FORMATTING VISUALISATIONS


 On the Visualizations pane you can access the Format tools. Here you can change the legend, titles, colours, font size, and other settings.

CONDITIONAL FORMATTING

One of the formatting options for the Table and Matrix visualisations is Conditional formatting. Like Excel you can change colours and add data bars or icons. Use Advanced controls for more options.

FILTERS AND SLICERS

 On the Filters pane you can work with the [filters](#) applied to the current visual, the current page, and the entire report. Fields used to create the visual will show automatically and you can add more fields as needed. Expand each filter to see more options. Available options will depend on the field data type.

 [Slicers](#) are a type of visual that can be added to a report and act like a filter. By default, it will act on all visuals on a page. Go to **Format ► Edit interactions** to choose which visuals are affected. In **View ► Sync slicers** you can show and link a slicer to multiple pages in the report.

SHARING REPORTS

You can export the report to a PDF or publish it to the Power BI service. **My Workspace** is just for you, use a shared workspace so that other people can see your reports. You can export to PowerPoint from the Power BI service

DAX FUNCTIONS

The DAX functions that we have looked at are:

[CALCULATE](#) – evaluate an expression with the specified filters.

[SUM](#) – adds the values.

[DIVIDE](#) – you can divide using the / operator, however you will get an error if you divide by 0. This function allows you to specify what happens in this case.

[COUNT](#), [DISTINCTCOUNT](#) – counts values, either all values or only distinct values.

[MAX](#), [MIN](#) – find the maximum or minimum value.

[YEAR](#), [MONTH](#) – extract the specified part of a date as a number. Use **FORMAT** if you want the month name or the year in a particular format.

[FORMAT](#) – converts a value (number/date) to the specified format.

[VALUE](#) – convert text to a number.

[CALENDAR](#), [CALENDARAUTO](#) – create a table with a Date column. CALENDARAUTO uses the dates in the model; with CALENDAR you specify the dates.

[PREVIOUSMONTH](#) – returns all the dates from the previous month.

[SAMEPERIODLASTYEAR](#) – returns the dates one year earlier than the given dates.