

Data types in Excel

Text

Numbers

Percentages

Date and time

Data types in Excel

- Unlike Access, we don't have to assign a data type to the data we enter in Excel
- But, Excel does have data types – it assigns them as you enter the data, based on what you enter...in other words, it guesses
- Functions, formulas only work on the correct data type
- Common data types to be aware of are:
 - Numbers
 - Percents
 - Text
 - Dates/times
 - Logical (Boolean)

Numbers

- All numbers are stored as double-precision floating point values
 - Why? Integer math drops decimal places
 - Division of integers produces decimals – Excel can't count on the user to never do division on integer numbers
 - Floating point values work with any mathematical operation – better choice given a spreadsheet's flexibility
 - If you enter an integer, it will display in the cell without decimal places, but internally it's recorded as floating point
- Numbers displayed right-justified by default
- Usually the number entered and the number stored and displayed are the same, aside from rounding
- Exception: percents

Percents are stored as proportions

- Entering 50% causes Excel to:
 - Store the value 0.5 (**not** 50)
 - Set the cell type to Percentage
 - Display 50%
- Entering 0.5 and then setting the display type to Percentage gives same result

The diagram illustrates two methods of entering a percentage in Excel and the resulting cell format and display.

Top Row (Entering 50% directly):

- Left Screenshot:** The 'General' dropdown is selected. The cell contains '50%'.
- Right Screenshot:** The 'Percentage' dropdown is selected. The cell displays '50%'.

Bottom Row (Entering 0.5 and then setting format to Percentage):

- Left Screenshot:** The 'General' dropdown is selected. The cell contains '0.5'.
- Right Screenshot:** The 'Percentage' dropdown is selected. The cell displays '50.00%'.

Arrows indicate the transition from the 'General' format to the 'Percentage' format in both scenarios.

← Display type setting

Setting the display type to percent before entering data

- Setting an empty cell's type to Percentage and then entering 0.5 gives 0.5%
- Changing this back to General shows the number stored is 0.005

The diagram illustrates the effect of changing the display type of a cell in Excel. It shows three states of a spreadsheet with columns G, H, and I, and rows 1 through 5. The first state (top left) shows the 'Percentage' display type selected in the ribbon, with cell H1 containing '50%' and cell H2 containing '50.00%'. The second state (top right) shows the same spreadsheet, but with cell H3 containing '0.50%', indicating that the value 0.5 was entered while the display type was set to Percentage. The third state (bottom) shows the 'General' display type selected in the ribbon, with cell H3 now displaying '0.005', which is the underlying numerical value stored in the cell. Arrows indicate the sequence of changes: from the first state to the second, and then from the second state to the third.

	G	H	I
1		50%	
2		50.00%	
3		0.50%	
4			
5			

	G	H	I
1		50%	
2		50.00%	
3		0.005	
4			
5			

Text

- Any non-numeric entries are stored as text
 - Mixes of numbers and letters (24 g) stored as text
- Left-justified by default
- Text entries won't be used in calculations, but don't cause error messages
 - Text skipped over
- Various text functions
 - len() - length of a text string
 - lower() - converts to lower case
 - proper() - converts to sentence case (first letter capitalized)
 - Substring functions for extracting parts of strings by position
 - trim() - removes extra white spaces from the end of text entries

A number as a number ↓

	123
123	

↑ A number as text

Alphabetical sort order for mixed text and numbers

- We alphabetize words starting with the first letter, then moving to the second when the first ties, etc.
 - Aardvark before Adam
- Shorter strings before longer that start with the shorter
 - an before and
- Numbers come before letters
- If mixed, Excel considers numbers to be text characters, and sorts them alphabetically, not in numeric order
 - a1, a10, a2 instead of a1, a2, a10
 - What about 1a, 2a, 10a?
- To get the numbers to sort properly, use leading zeros (01a, 10a, 02a), or keep letters and numbers in separate columns

Dates and times

- 1/1/17, Jan1, and January 1, 2017 would all be recognized by Excel as 1/1/2017
- When Excel guesses you entered a date, it:
 - Presumed you meant 2017, not 1917 or 2117 (use all four digits as needed)
 - Converted the date internally to a number
 - Sets the display type of the cell to date
- Excel uses 1/0/1900 as the “epoch”
 - All dates stored as number of days since this date
 - Internally, 1/1/17 is 42736
 - 1/1/1900 is the earliest date Excel can store! Earlier dates stored as text
- Be careful, because:
 - What is stored internally is not the same as what is entered
 - What is displayed is not the same as what is stored internally
 - What if Jan1 refers to the first measurement you made on Jan?

Time

- Time is recorded as a fraction of a day
- If the time recorded is 2:00 pm, this is converted as:
 - $14 + 0/60 = 14.00$ ← hour and decimal minutes
 - $14.00/24.00 = 0.5833333333333333$ ← decimal hours and minutes
- If a date is given as well, it is the whole number in the date/time number
 - 2:00 pm on 1/1/17 is 42736.5833333333
 - Note: total of 15 decimal places, so storing a date means storing time to only 10 decimal places
 - Not a huge problem – 1×10^{-10} days is 0.00000864 seconds, or 0.0084 milliseconds (still precise time keeping)

Advantage of storing dates/times as numbers

- Can do math on them
 - Days between two dates, time elapsed between two times easy to calculate
 - Sequences of days easy to generate
- There are special functions that can be applied to date/time data
 - Extracting portions: month(date), day(date), year(date), hour(time), minute(time), second(time)
 - Calculate the weekday, week number of any date: weekday(date), weeknum(date)
- Storing dates as text, splitting year, month, and day into different columns allows any dates to be stored, but loses the advantages of using Excel's numeric date and time format

Logical

- Logical data types can only take one of two values, TRUE or FALSE
- Used for making decisions – results of comparisons produce logicals
- Displayed as text (TRUE or FALSE), internally represented as numbers (1 or 0)
- Logicals are centered in the cell, not left-justified like text, or right-justified like numbers
- Can do math with them (but functions may not work properly with them)

Logicals in a spreadsheet

The screenshot displays an Excel interface. The main worksheet has columns labeled B, C, D, E, and F. Column B is titled 'Logicals' and contains a sequence of logical values: FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, TRUE, TRUE, and FALSE. A secondary window or dropdown menu is open, showing a formula bar with the text '=true'. Below the formula bar, the text 'Entering logicals' is visible, followed by a list of logical values: true, false, TRUE, and FALSE.

A	B	C	D	E
	Logicals		Add zero	
	TRUE		1	
	FALSE		0	
	TRUE		1	
	TRUE		1	
	FALSE		0	
	FALSE		0	
	Text			
	TRUE		#VALUE!	
	FALSE		#VALUE!	
	TRUE		#VALUE!	
	TRUE		#VALUE!	
	FALSE		#VALUE!	
	FALSE		#VALUE!	

Can do math with them!