

Filtering Data In Excel

Copyright 2009 Peter R. Douglass

All rights reserved

Example Problem

- You are given a spreadsheet like the one on the next slide
- You are asked to find the average QPA for the MA residents and the average QPA for the non-MA residents.

Example Problem

The screenshot shows the Microsoft Excel interface. The title bar reads "Book1 - Microsoft Excel". The ribbon includes "Home", "Insert", "Page Layout", "Formulas", "Data", "Review", "View", and "Acrobat". The Acrobat ribbon has three main groups: "Create PDF" (with "Create Adobe PDF"), "Create and Attach to Email" (with "Create and Email"), and "Create and Send For Review" (with "Review And Comment"). The formula bar shows "F12" and a function icon. The spreadsheet has columns A, B, C, D, E, and F. Column F is highlighted in orange. The data is as follows:

	A	B	C	D	E	F
1	Student	State	GPA			
2	Tom Smith	MA	3.675			
3	Betty Jones	VT	3.829			
4	Amy Kane	MA	3.506			
5	Will Bentley	MA	3.102			
6	Jesse Liu	NH	3.441			
7	Nancy Sims	MA	2.987			
8	John Prince	CT	3.378			
9						

The status bar at the bottom shows "Ready", "Sheet1", "Sheet2", "Sheet3", and a zoom level of "100%".

Example Problem

- What we will do is filter the qpa data into columns based upon the state information.
- We want to make sure that the given data is hard-coded in only one place, so that if the data changes, our answers will change appropriately
- We will use excel's `if` formula to achieve our goals

Data Filtered into Columns

The screenshot shows the Microsoft Excel interface with a filtered table. The ribbon is set to the 'Acrobat' tab. The table has 9 rows and 6 columns. The first row contains the headers: Student, State, GPA, MA, and non-MA. The data rows are: Tom Smith (MA, 3.675), Betty Jones (VT, 3.829), Amy Kane (MA, 3.506), Will Bentley (MA, 3.102), Jesse Liu (NH, 3.441), Nancy Sims (MA, 2.987), and John Prince (CT, 3.378). The 'MA' and 'non-MA' columns are filtered to show only rows where the GPA is greater than 3.0.

	A	B	C	D	E	F	G
1	Student	State	GPA		MA	non-MA	
2	Tom Smith	MA	3.675		3.675		
3	Betty Jones	VT	3.829			3.829	
4	Amy Kane	MA	3.506		3.506		
5	Will Bentley	MA	3.102		3.102		
6	Jesse Liu	NH	3.441			3.441	
7	Nancy Sims	MA	2.987		2.987		
8	John Prince	CT	3.378			3.378	
9							

How it is done

The screenshot shows Microsoft Excel with a spreadsheet containing the following data:

	A	B	C	D	E	F
1	Student	State	GPA		MA	non-MA
2	Tom Smith	MA	3.675		=IF(B2=\$E\$1,C2,"")	=IF(B2=\$E\$1,"",C2)
3	Betty Jones	VT	3.829		=IF(B3=\$E\$1,C3,"")	=IF(B3=\$E\$1,"",C3)
4	Amy Kane	MA	3.506		=IF(B4=\$E\$1,C4,"")	=IF(B4=\$E\$1,"",C4)
5	Will Bentley	MA	3.102		=IF(B5=\$E\$1,C5,"")	=IF(B5=\$E\$1,"",C5)
6	Jesse Liu	NH	3.441		=IF(B6=\$E\$1,C6,"")	=IF(B6=\$E\$1,"",C6)
7	Nancy Sims	MA	2.987		=IF(B7=\$E\$1,C7,"")	=IF(B7=\$E\$1,"",C7)
8	John Prince	CT	3.378		=IF(B8=\$E\$1,C8,"")	=IF(B8=\$E\$1,"",C8)
9						

The formula bar shows the active cell E4 containing the formula: `=IF(B4=E1,C4,"")`

Now Calculating Averages

The screenshot shows the Microsoft Excel interface with the following data and formula:

	A	B	C	D	E	F	G	H
1	Student	State	GPA		MA	non-MA		
2	Tom Smith	MA	3.675		3.675			
3	Betty Jones	VT	3.829			3.829		
4	Amy Kane	MA	3.506		3.506			
5	Will Bentley	MA	3.102		3.102			
6	Jesse Liu	NH	3.441			3.441		
7	Nancy Sims	MA	2.987		2.987			
8	John Prince	CT	3.378			3.378		
9								
10	Average		3.417		3.3175	3.549333		

The formula bar shows the formula for cell E10: `=AVERAGE(E2:E8)`. The value 3.3175 in cell E10 is highlighted with a black border.

Averagelf

- The technique shown in these slides is of *general* use.
- For the *particular* problem shown in these slides, there is another solution, which is to use the `averageif` function. (There is also a `sumif` and a `countif` function)
- There are many cases where the technique shown here is necessary, because there are no special functions to do job.