Introductory Spreadsheet Exercise: Excel

Brant Houston
National Institute for Computer-Assisted Reporting
brant@nicar.org

The lesson is an excerpt from Computer-Assisted Reporting: A Practical Guide by Brant Houston with permission of St. Martin's Press. To order the book, call St. Martin's Press at (800) 470-4767, ext. 313 or call the Investigative Reporters and Editors Inc. at 314-882-2042.

Although spreadsheets can be one of the least complex types of software for analyis, they have been used for a variety of stories, both for on-the-beat and for long-term investigations.

Using a spreadsheet, Christopher Schmitt, projects reporter *at The San Jose Mercury News*, was able to pull together a story on dangerously slow reponse times by the city fire department in less than a week when a series of arsons hit the city.

National Law Journal reporters also used spreadsheets in part of their award-winning examination of racism and environmental pollution. *Fortune* magazine used spreadsheets on a major wage study. Many other newspapers used spreadsheets to analyze crime statistics, drug arrests, government corruption, property taxes, and municipal budgets.

Learning addresses

One of the powers of a spreadsheet is that it allows you to do quick calculations and skip repetitive tasks. One of the spreadsheet's tricks is to use the "address" of a number rather than the number itself.

If you have ever looked for a town on a road atlas, you've gone to the index. The index may refer you to page 7 and to the town's location at "D4". You go to Page 7. The map is laid out on a grid with letters across the top and numbers down the sides. You look down column D and then look across on row 4 to find the town.

Perhaps you've played chess and want to replay the game. The instructions for the location of the pieces are also given with letters and numbers defining the squares. A knight, for example, will move to from B1 to C4.

Better yet, you may have played the game of Battleship. In that game, you make a grid with letters and numbers. You block out two or three consecutive squares for a battleship and place mines in individual squares. Your opponent does the same. Then you take turns, saying such things as "A4". Your opponent says "hit" if your opponent's battleship is on that square, "miss" if you hit nothing, and "boom" if you hit a mine. Spreadsheets treat your information as though it's part of a game of Battleship.

Average and Median

Journalists often talk about the average salary, but sometimes it is valuable to know both average and median. Average is the total amount of salaries divided by the number of salaries. Median is the value at which half the salaries are higher and half are lower. If you are looking at a set of salaries and one person is getting an extremely high salary, then the average will be distorted by that high salary. A more accurate and fair way to look at salaries in that case would be to use the median.

The newest versions of spreadsheets make it easy to do that. In this example, we have a fictional collection of salaries received by the cronies of a mayor as shown in table 1-1:

Table 1-1

	A	B	C
1	Name	Last year	This year
2	Mark Forest	$ 15,000	$ 21,000
3	Jane Deed	$ 14,000	$ 19,000
4	Mary Hill	$ 22,000	$ 29,000
5	Joe Smith	$ 30,000	$ 39,000
6	Ed Powell	$ 25,000	$ 30,000
7	Tom Brown	$ 40,000	$ 47,000
8	Julia Jones	$ 50,000	$ 58,000
9	Dee Dale	$ 45,000	$ 52,000
10			

To find the average, move your cursor to A11 and type Average. Then move the cursor to B11 as in Table 13-2 and type =Average(b2:b9) as shown in Table 1-2. The (b2:b9) tells the spreadsheet that you want to calculate a range of numbers from B2 to B9.

Table 1-2

B11 =Average(b2:b9)

	A	B	C
1	Name	Last year	This year
2	Mark Forest	$ 15,000	$ 21,000
3	Jane Deed	$ 14,000	$ 19,000
4	Mary Hill	$ 22,000	$ 29,000
5	Joe Smith	$ 30,000	$ 39,000
6	Ed Powell	$ 25,000	$ 30,000
7	Tom Brown	$ 40,000	$ 47,000
8	Julia Jones	$ 50,000	$ 58,000
9	Dee Dale	$ 45,000	$ 52,000
10			
11	Average	=Average(b2:b9)	
12			

Hit enter to get the average in Table 1-3.

Table 1-3

B11 =AVERAGE(B2:B9)

	A	B	C
1	Name	Last year	This year
2	Mark Forest	$ 15,000	$ 21,000
3	Jane Deed	$ 14,000	$ 19,000
4	Mary Hill	$ 22,000	$ 29,000
5	Joe Smith	$ 30,000	$ 39,000
6	Ed Powell	$ 25,000	$ 30,000
7	Tom Brown	$ 40,000	$ 47,000
8	Julia Jones	$ 50,000	$ 58,000
9	Dee Dale	$ 45,000	$ 52,000
10			
11	Average	$ 30,125	

Repeat the formula for the C column, typing = Average(c2:c9). Then move your cursor to A13 and click. Type Median and then move to B13 and type =Median(b2:b9) as in Table 1-4.

Table 1-4

B13 =Median(b2:b9)

	A	B	C
1	Name	Last year	This year
2	Mark Forest	$ 15,000	$ 21,000
3	Jane Deed	$ 14,000	$ 19,000
4	Mary Hill	$ 22,000	$ 29,000
5	Joe Smith	$ 30,000	$ 39,000
6	Ed Powell	$ 25,000	$ 30,000
7	Tom Brown	$ 40,000	$ 47,000
8	Julia Jones	$ 50,000	$ 58,000
9	Dee Dale	$ 45,000	$ 52,000
10			
11	Average	$ 30,125	$ 36,875
12			
13	Median	=Median(b2:b9)	

Hit enter to get the result in Table 1-5.

Table 1-5

B13 =MEDIAN(B2:B9)

	A	B	C
1	Name	Last year	This year
2	Mark Forest	$ 15,000	$ 21,000
3	Jane Deed	$ 14,000	$ 19,000
4	Mary Hill	$ 22,000	$ 29,000
5	Joe Smith	$ 30,000	$ 39,000
6	Ed Powell	$ 25,000	$ 30,000
7	Tom Brown	$ 40,000	$ 47,000
8	Julia Jones	$ 50,000	$ 58,000
9	Dee Dale	$ 45,000	$ 52,000
10			
11	Average	$ 30,125	$ 36,875
12			
13	Median	$ 27,500	

Repeat the formula for C13.

Now that you have seen how to find the average or median for a column of numbers, take a look at few other calculations.

In this next exercise you will analyze a fictional water and sewer budget shown in Table 1-6:

Table 1-6

	A	B	C
1			
2	Water and sewer division	1994-1995	1995-1996
3			
4	Residential water fees	$ 13,235,122.00	$ 18,405,222.00
5	Residential sewer fees	$ 6,544,344.00	$ 8,324,555.00
6	Commercial water fees	$ 10,882,021.00	$ 11,504,302.00
7	Commercial sewer fees	$ 4,343,123.00	$ 5,662,131.00
8	Investment interest	$ 1,222,494.00	$ 1,445,214.00
9			

You might want to find the total amount of the budget. To do so, type Total in A10 and then move your cursor to B10. In B10, type =SUM(b4:b8) and hit Enter. Then move your cursor to C10 and type =SUM(c4:c8) as show in Table 1-7.

Table 1-7

C10 =SUM(c4:c8)

	A	B	C
1			
2	Water and sewer division	1994-1995	1995-1996
3			
4	Residential water fees	$ 13,235,122.00	$ 18,405,222.00
5	Residential sewer fees	$ 6,544,344.00	$ 8,324,555.00
6	Commercial water fees	$ 10,882,021.00	$ 11,504,302.00
7	Commercial sewer fees	$ 4,343,123.00	$ 5,662,131.00
8	Investment interest	$ 1,222,494.00	$ 1,445,214.00
9			
10	Total	$ 36,227,104.00	=SUM(c4:c8)
11			

Hit the Enter key again to get the result shown in Table 1-8.

Table 1-8

C10 =SUM(C4:C8)

	A	B	C
1			
2	Water and sewer division	1994-1995	1995-1996
3			
4	Residential water fees	$ 13,235,122.00	$ 18,405,222.00
5	Residential sewer fees	$ 6,544,344.00	$ 8,324,555.00
6	Commercial water fees	$ 10,882,021.00	$ 11,504,302.00
7	Commercial sewer fees	$ 4,343,123.00	$ 5,662,131.00
8	Investment interest	$ 1,222,494.00	$ 1,445,214.00
9			
10	Total	$ 36,227,104.00	$ 45,341,424.00
11			

In looking at a budget, you usually want to see how much it has decreased or increased. Move your cursor to D2 and type Change. Then move your cursor to D4 and type =(c4-b4) as shown Table 1-9 to calculate the different residential water fees from the 1995 budget to the 1996 budget.

Table 1-9

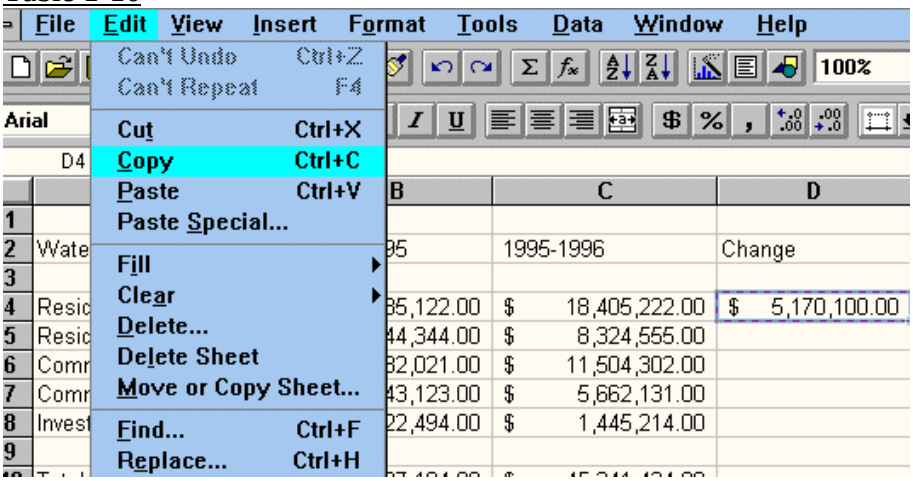
D4 =(c4-b4)

	A	B	C	D
1				
2	Water and sewer division	1994-1995	1995-1996	Change
3				
4	Residential water fees	$ 13,235,122.00	$ 18,405,222.00	=(c4-b4)
5	Residential sewer fees	$ 6,544,344.00	$ 8,324,555.00	
6	Commercial water fees	$ 10,882,021.00	$ 11,504,302.00	
7	Commercial sewer fees	$ 4,343,123.00	$ 5,662,131.00	
8	Investment interest	$ 1,222,494.00	$ 1,445,214.00	
9				
10	Total	$ 36,227,104.00	$ 45,341,424.00	

Tap Enter to get your result. You may get back a series of hash marks sucha ###. That means the column is too narrow. To widen the column, go to the grey area just above the column D and place your cursor on the far right line until you get a symbol of arrows pointing left and right. Click down and drag the cursor to the right. As you do, you will see the numbers appear.

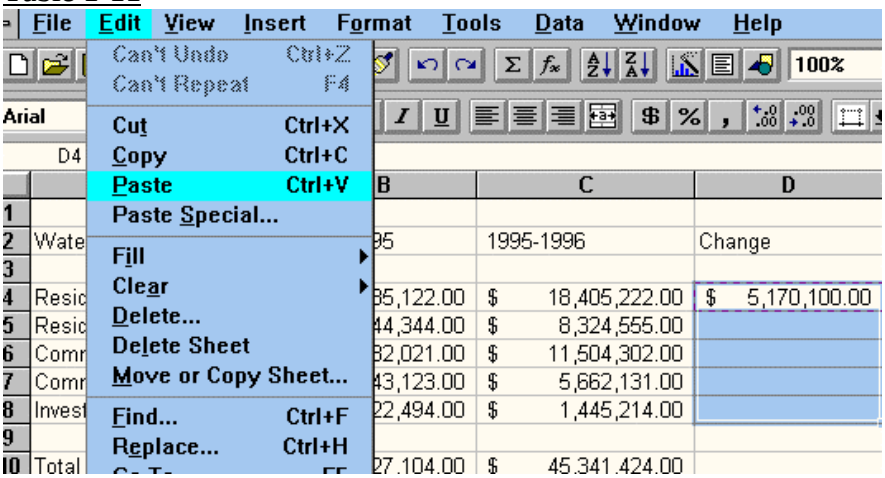
Rather than type the formula for each row, take the shortcut a spreadsheet offers when it allows you to copy a formula. Place your cursor on the Highlight D4, click, and then move your cursor to Edit on the menu at the top of the screen. Click on Edit and then Click on Copy as shown in Table 1-10.

Table 1-10



Now highlight from D4 to D8 by clicking on D4 and holding it down until you reach D8. Then return to Edit and click on Paste as shown in Table 1-11.

Table 1-11



You also can quickly copy a formula by moving your cursor to the lower right hand corner of D4 until you see a thin cross. Once you have that, you can click and drag to the D8. When you let go, the formula will be copied. In any case, your result should look like Table 1-12.

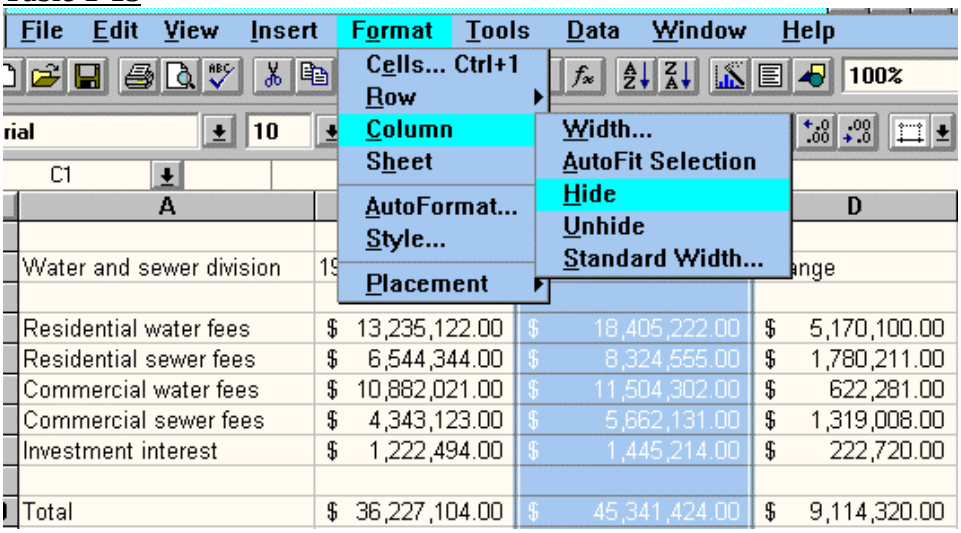
Table 1-12

D4 =(C4-B4)

	A	B	C	D
	Water and sewer division	1994-1995	1995-1996	Change
	Residential water fees	$ 13,235,122.00	$ 18,405,222.00	$ 5,170,100.00
	Residential sewer fees	$ 6,544,344.00	$ 8,324,555.00	$ 1,780,211.00
	Commercial water fees	$ 10,882,021.00	$ 11,504,302.00	$ 622,281.00
	Commercial sewer fees	$ 4,343,123.00	$ 5,662,131.00	$ 1,319,008.00
	Investment interest	$ 1,222,494.00	$ 1,445,214.00	$ 222,720.00
	Total	$ 36,227,104.00	$ 45,341,424.00	

Sometimes you may want to see only a few of the columns in a spreadsheet. A quick way to "hide" one or more columns is to use the Hide command. To hide a column, you can highlight it by clicking in the grey bar just above the spreadsheet. In this case, move the cursor to the C column and click. That will highlight the entire column. Then move your cursor to Format and click, and then click on Column, and click on Hide as show in Table 1-13.

Table 1-13



After you click on Hide, you will be left with three columns. Notice that the C column is missing in Table 1-14. If you want to Unhide the column later, highlight the *both* B column and the D column and then return to Format, Column, Unhide.

Sometimes the percentage increase or decrease in a budget is more important than the actual changes in amounts. To calculate the percentage difference, or in this case, the percentage increase from the 1994-1995 budget to the 1995-1996, you type =d4/b4 in E4 as shown in Table 1-14.

Table 1-14

E4 =d4/b4

	A	B	D	E
1				
2	Water and sewer division	1994-1995	Change	Percent Change
3				
4	Residential water fees	$ 13,235,122.00	$ 5,170,100.00	=d4/b4
5	Residential sewer fees	$ 6,544,344.00	$ 1,780,211.00	
6	Commercial water fees	$ 10,882,021.00	$ 622,281.00	
7	Commercial sewer fees	$ 4,343,123.00	$ 1,319,008.00	
8	Investment interest	$ 1,222,494.00	$ 222,720.00	
9				
10	Total	$ 36,227,104.00	$ 9,114,320.00	

You hit enter and then copy the formula to E10. You can get rid of the garbage character in E9 by highlighting E9 and hitting the space bar. Your result is shown in Table 1-15.

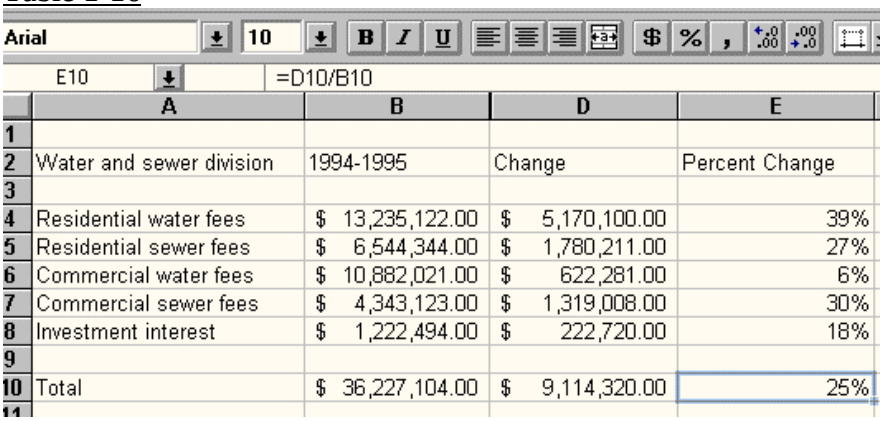
Table 1-15

E10 =D10/B10

	A	B	D	E
1				
2	Water and sewer division	1994-1995	Change	Percent Change
3				
4	Residential water fees	$ 13,235,122.00	$ 5,170,100.00	0.390634858
5	Residential sewer fees	$ 6,544,344.00	$ 1,780,211.00	0.272022834
6	Commercial water fees	$ 10,882,021.00	$ 622,281.00	0.057184323
7	Commercial sewer fees	$ 4,343,123.00	$ 1,319,008.00	0.303700356
8	Investment interest	$ 1,222,494.00	$ 222,720.00	0.182184943
9				
10	Total	$ 36,227,104.00	$ 9,114,320.00	0.251588424
11				

Unfortunately, the decimal numbers are not too helpful, but you can change them quickly to percentages. To change the decimal numbers to percentages, highlight the column by clicking on E in the grey bar and then click on the percentage sign in the bar above that. Your result should look like Table1-16.

Table 1-16



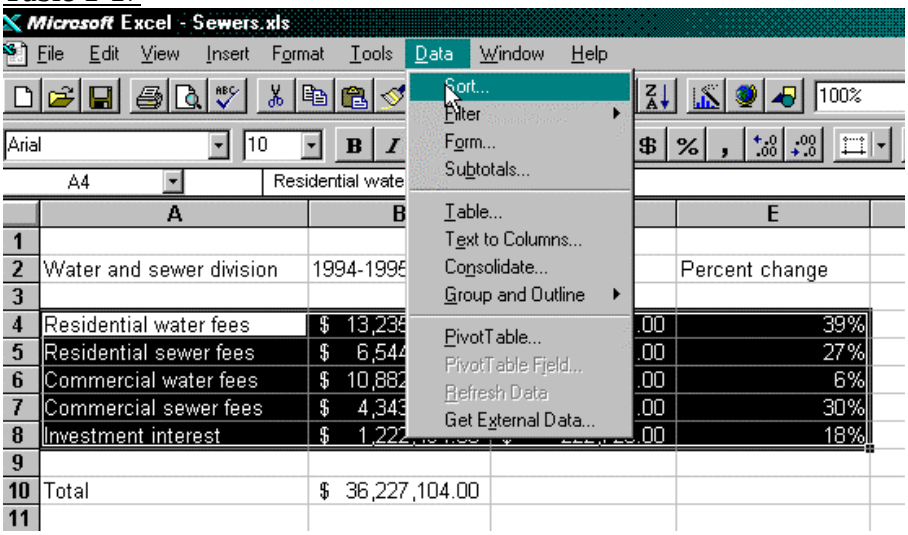
Arial 10 B I U

E10 =D10/B10

	A	B	D	E
1				
2	Water and sewer division	1994-1995	Change	Percent Change
3				
4	Residential water fees	$ 13,235,122.00	$ 5,170,100.00	39%
5	Residential sewer fees	$ 6,544,344.00	$ 1,780,211.00	27%
6	Commercial water fees	$ 10,882,021.00	$ 622,281.00	6%
7	Commercial sewer fees	$ 4,343,123.00	$ 1,319,008.00	30%
8	Investment interest	$ 1,222,494.00	$ 222,720.00	18%
9				
10	Total	$ 36,227,104.00	$ 9,114,320.00	25%
11				

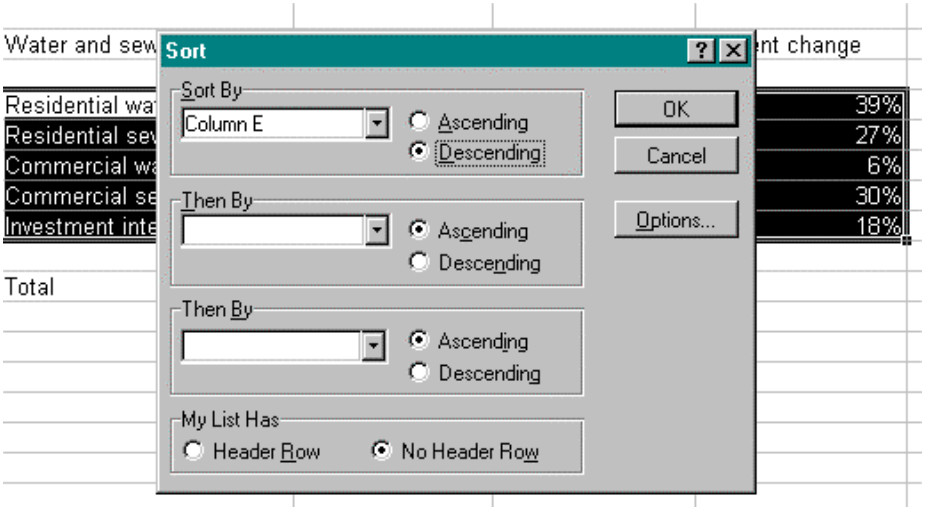
If you want to sort the percentages from highest to lowest, highlight A4 to E8. Then click on Data in the bar at the top of the screen shown on Table 1-17.

Table 1-17



Then click on Sort and you get a choice of column to sort by and whether to sort it from highest to lowest (descending) or lowest to highest (ascending) as shown in Table 1-18:

Table 1-18



Choose Column E and Descending. Then click on OK. The budget should rearrange itself by the percentage column as shown in Table 1-19:

Table 1-19

Residential water fees	$ 13,235,122.00	$ 5,170,100.00	39%
Commercial sewer fees	$ 4,343,123.00	$ 1,319,008.00	30%
Residential sewer fees	$ 6,544,344.00	$ 1,780,211.00	27%
Total	$ 24,122,589.00	$ 8,269,319.00	34%
Investment interest	$ 1,222,494.00	$ 222,720.00	18%
Commercial water fees	$ 10,882,021.00	$ 622,281.00	6%