

The goal of this exercise is to gain familiarity with some basic functions of Excel that can be used for data manipulation and statistical applications. We will use a small data set (N=10). As we work through the exercise, can you explain to someone else what we are doing on each step?

<b>Student Number</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
<b>Age (X)</b>	:	18	17	19	22	21	17	20	18	19	19
<b>Self Actualization (Y)</b>	:	50	43	65	89	75	60	71	57	63	58

1. Call up Excel. Prepare the settings. Under the 'View' tab, check 'Formula Bar' and under 'View' put the cursor on 'Toolbars' and check 'Standard' and 'Formatting.'
2. Enter **Student** in Cell A1, **Age(X)** in Cell B1, and **SelfAct(Y)** in Cell C1.
3. Enter the data in these three columns. Hint: after you enter the first two numbers in the Student column, an Excel shortcut can fill in the rest of the column. Highlight the two cells (A2 and A3), move your cursor to the little black square at the bottom-right of Cell A3 and the cursor becomes a + sign, hold the left button and drag downward through 10.
4. Find the sum of X and the sum of Y. In Cell A12 type **Sum**. Move the cursor to Cell B12 and type **=sum(B2:B11)**. The "=" sign tells Excel to interpret the entry as a formula rather than text. (Check  $\Sigma X=190$ ). We can also use the formula menu: Click on the empty Cell C12, click on the ' $f_x$ ' that is located above the data to open the Insert Function window, select category 'All,' select 'SUM' and click OK. Now highlight the Y data from Cell C2 through C11 and click OK.
5. Now find the means of X and Y. In Cell A13 type **Mean**. Move the cursor to Cell B13 and type **=average(b2:b11)**. We can also use the formula bar as we did in Step 4. Move the cursor to cell C13 first, but this time look for the 'AVERAGE' function.
6. Next find the sums of squares. Put labels on the columns, **X^2** in Cell D1 and **Y^2** in Cell E1. Move the cursor to Cell D2. Type **=B2^2** and press Enter. We can copy this formula into the rest of the column. Click on Cell D2, move the cursor to the little black square, hold the left mouse button, and drag down through Cell D11. We can easily copy these formulas to the next column. Highlight the block of cells from D2 through D11, move the cursor to the little black square, hold the left mouse button and drag to the right one column. Now find the sum of squares by dragging the summation formula in Cell C12 into cells D12 and E12.
7. Compute the sum of cross products. In Cell F1 type **XY**. In Cell F2 type **=B2\*C2**. Drag this formula from Cell F3 through Cell F11. Drag the sum formula from E12 to F12.
8. For illustration, we will calculate the standard deviation of X three different ways. To prepare, type **First Sx =** into Cell A15, **Second Sx =** into Cell A16, and **Third Sx =** into Cell A17.

First, let's compute Sx using deviation scores with the formula:

$$s_x = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}}$$

In Cell G1 type **(X-Xbar)**. In Cell G2 type **=B2-B\$13**. Copy this

formula into cells G3 through G11. Note: Click on Cell G3 and see that the formula in this cell is B3-B\$13. When we moved the formula B2-B\$13 from row 2 to row 3, the B2 changed to B3 to retain the relative position between cell with the formula and the data it refers to. The \$ sign in front of the 13 fixed that number to remain at 13 even when we moved the formula to a new cell.

9. In Cell H1 type  $(X - \bar{X})^2$ . In Cell H2 type  $=G2 * G2$ . Drag this formula into cells H3 through H11. Drag the summation formula from Cell F12 into G12 and H12.

10. Move the cursor to Cell B15 and enter the formula  $=SQRT(H12/9)$ . Explain to Bumble each step of this calculation. Hint:  $S_x = 1.633$ .

11. Next compute  $S_x$  using the formula 
$$S_x = \sqrt{\frac{\sum X^2 - (\sum X)^2 / N}{N - 1}}$$
 Move the cursor to Cell B16 and Type  $=sqrt((D12 - (B12^2)/10)/9)$ . Explain this to Bumble.

12. Now let's do it the easy way. Move the cursor to the empty Cell B17. Click on the  $f_x$ , find the formula STDEV and double-click it; when the Function Arguments window opens, highlight the X data from cells B2 through B11 and click OK.

13. Let's compute  $S_y$  using the easy method. Type  $S_y =$  in Cell A18, put the cursor in Cell B18, and repeat Step 12 using the Y data instead of the X data. Hint:  $S_y = 13.025$ .

14. Compute the covariance of X and Y using the formula 
$$COV(X, Y) = \frac{\sum XY - (\sum X \sum Y) / N}{N - 1}$$
 Type  $COV(X, Y) =$  in Cell A19, move the cursor to Cell B19, and enter  $=((F12 - (B12 * C12) / 10) / 9)$   
**Note:** The Excel formula for COV is for a population rather than for a sample estimate of a population value. Thus, the denominator is N rather than (N-1). If you use the Excel formula for COV, you can convert it to a sample estimate of a population value by multiplying by  $N / (N - 1)$ .

15. Calculate the correlation two ways. Type **First r =** in Cell D15 and **Second r =** in Cell D16.

We will use the formula  $r_{xy} = s_{xy} / (s_x s_y)$  and the built-in Excel formula. First the hard way: Move the cursor to Cell E15 and type  $=b19 / (b17 * b18)$ . Explain to Bumble. Now move the cursor to Cell E16, click the  $f_x$  to call up formulas, find CORREL and click OK, highlight the X data as Array 1 and the Y data as Array 2 and click OK. Hint:  $r = .914$ .

**Summary:**

Excel is very useful for simple data manipulation and calculations and it can do some sophisticated analyses and graphics. There is no merit in doing tedious calculations by hand when you can use a computer. However, the user must be very careful to check that Excel is doing what is intended. It is EXTREMELY easy to enter a formula incorrectly, miss a few numbers from an array, enter the wrong formula, or even use what appears to be a correct formula but not get what you intend (such as if you had used the built-in COV on Step 14 of our exercise). A limitation is that a printout does not show the formulas, so errors are often hidden.

**Lesson: User Beware! Verify computer analyses carefully.**