Microsoft Excel

Download

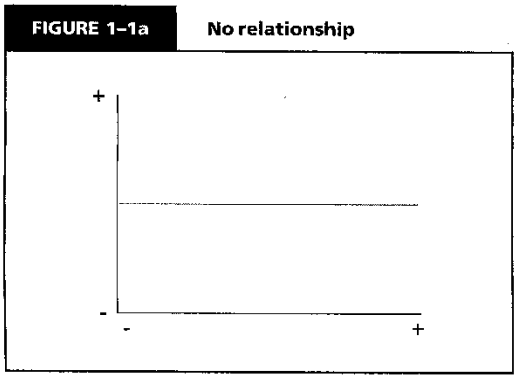
1. Go to the Course Website on BlackBoard
2. In Lectures, right click and save to your desktop “Session3 Excel 365.xlsx

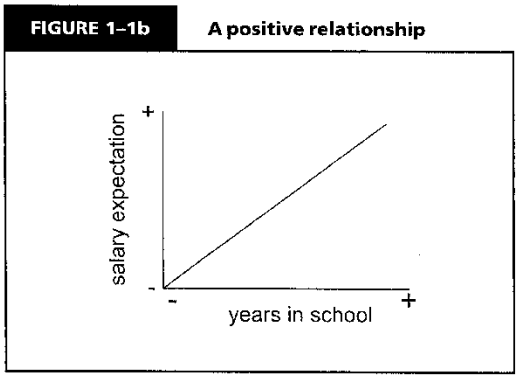
Session 3.1: Modeling Background (Correlation & Regression)

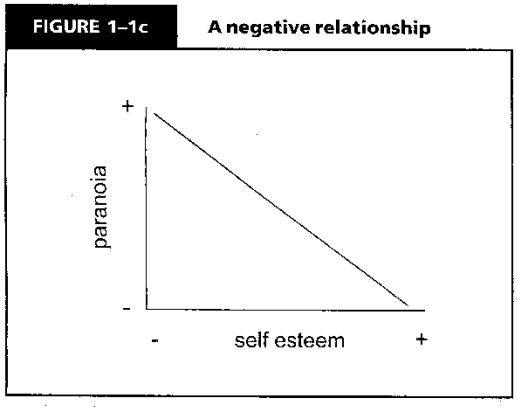
Identifying data relationships is key to modeling behavior of customer, student, and corporate data. First, let’s consider two variables and the relationships between them. When comparing two data variables, you can have:

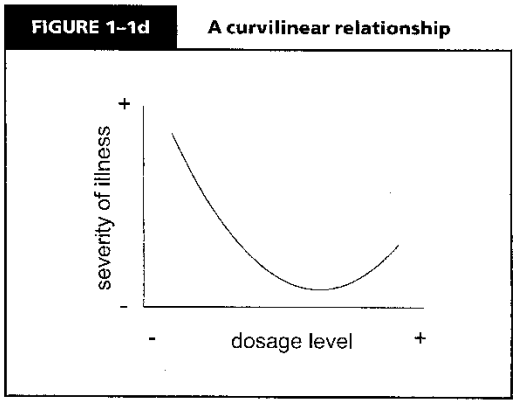
1. No relationship between the variables
2. A positive relationship (when one variable goes up, the other goes up)
3. A negative relationship (when one variable goes up, the other goes down)
4. A curvilinear relationship (a non-linear relationship

Examples of these, from The Research Methods Knowledge Base by Trochim & Donnelly (2007):



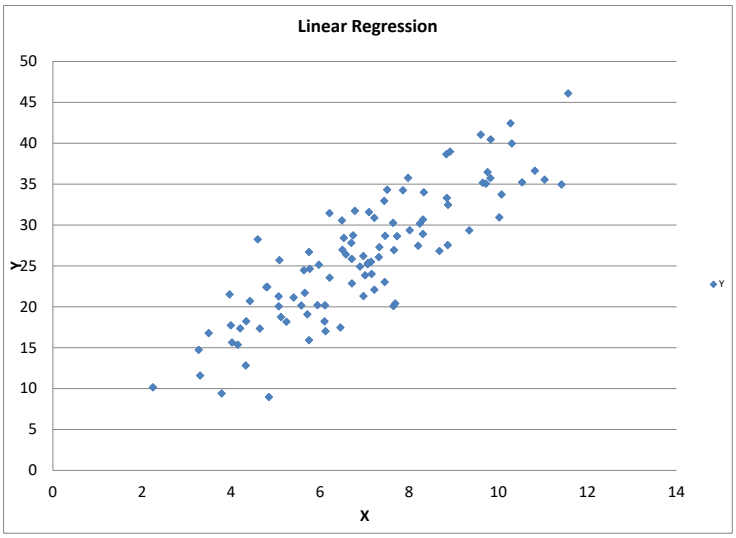




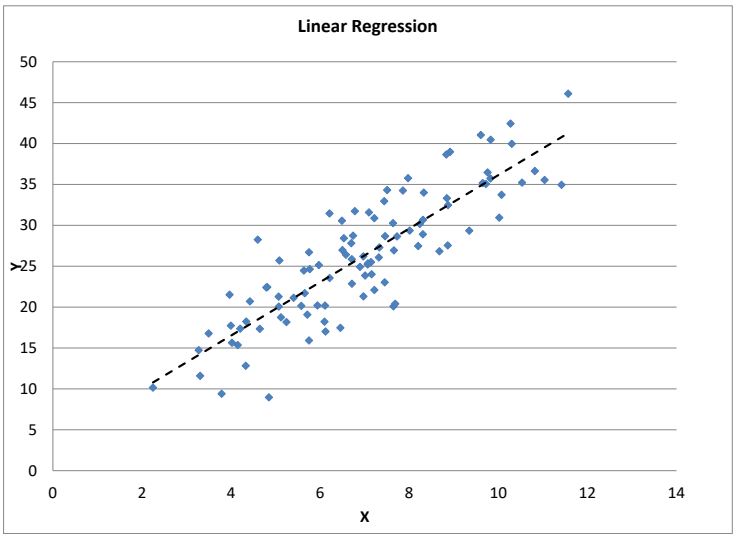


Regression

Linear regression is a technique that calculates the relationship between a dependent variable Y and one or more independent variables, or X's. Assume that you have data like the picture below.



You can calculate a regression trend line based on the data. This dashed line represents $\hat{Y}$ which is the estimate of the Y equation.

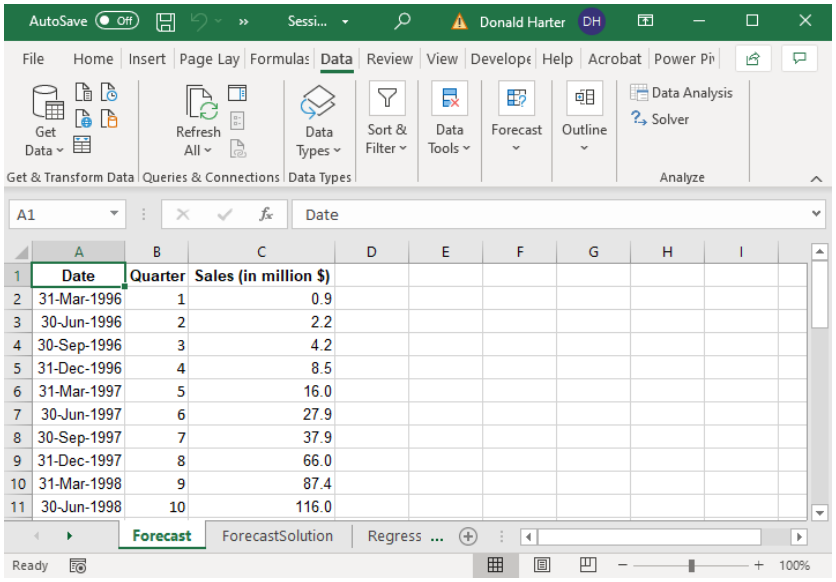


The vertical distance between the line and the data point is called the residual or error term.

Session 3.2: Forecasting

Regression models describe how changes in explanatory variables affect the outcome over the period of the data. If you want to forecast a trend into the future, then the new forecast feature initially offered in Excel 2016 is available. Forecasts use one date variable and one outcome variable. Note that forecasting outcomes over time can be risky because there is no guarantee that the trend will continue.

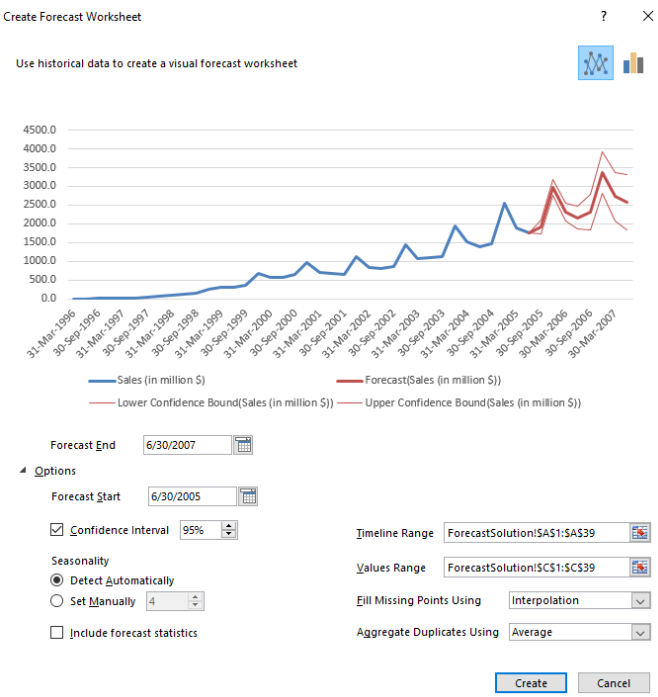
1. Use the Forecast spreadsheet for this exercise. The data below is the quarterly sales data for Amazon.com. Recall that the data displayed seasonality.



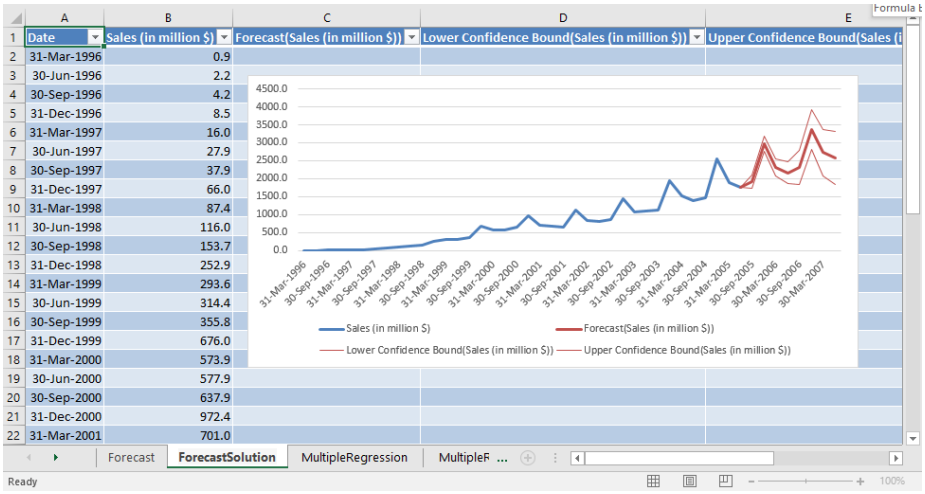
	A	B	C
1	Date	Quarter	Sales (in million $)
2	31-Mar-1996	1	0.9
3	30-Jun-1996	2	2.2
4	30-Sep-1996	3	4.2
5	31-Dec-1996	4	8.5
6	31-Mar-1997	5	16.0
7	30-Jun-1997	6	27.9
8	30-Sep-1997	7	37.9
9	31-Dec-1997	8	66.0
10	31-Mar-1998	9	87.4
11	30-Jun-1998	10	116.0

2. Click on the Data tab, then Forecast Sheet. In this example, Excel automatically identifies the relevant date and trend data.

3. Click on the Options drop down arrow in the lower left corner.



4. You can set the Forecast Start and Forecast End
5. Uncheck the box on Confidence Interval. This removes the confidence interval.
6. Change the Confidence Interval to 50%. What happens?
7. Excel Forecast is generally good at identifying Seasonality. If it has difficulty, you can click on Set Manually and set the seasonality parameter (4 for quarters, 12 for months, 52 for weeks, etc.)
8. If your data has missing data points, you can select Fill Missing Points Using: Interpolation or Zeros. Interpolation is usually better.
9. If there are duplicates in the data, set Aggregate Duplicates Using: Average.
10. In the upper right corner is the option for line versus bar chart. Click each.
11. Click on Create to generate the forecast. New columns with forecasted data are created.



Session 3.3: Multiple regression with categorical variables

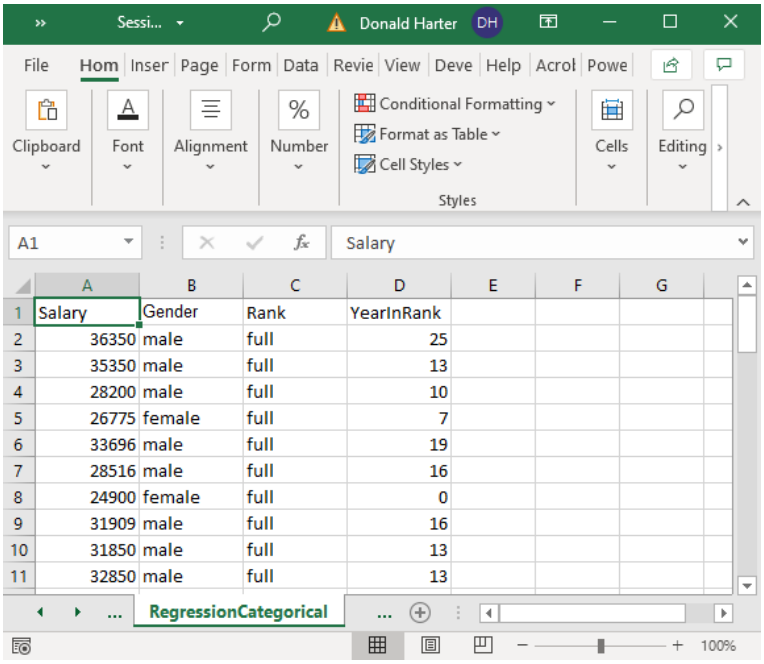
Often variables in a dataset will not be numeric but will be categorical text variables. For this exercise, use the RegressionCategorical spreadsheet.

The following university salary data is from:

> S. Weisberg (1985). *Applied Linear Regression*, Second Edition. New York: John Wiley and Sons. Page 194

The data includes:

Salary	Dollars per year
Gender	male or female
Rank	full, associate, assistant
Years in Rank	number of years in the current position



	A	B	C	D
1	Salary	Gender	Rank	YearInRank
2	36350	male	full	25
3	35350	male	full	13
4	28200	male	full	10
5	26775	female	full	7
6	33696	male	full	19
7	28516	male	full	16
8	24900	female	full	0
9	31909	male	full	16
10	31850	male	full	13
11	32850	male	full	13

1. The Gender and Rank variables are categorical: they are text fields that put the individual in a group. To run a regression on categorical variables, we need to first convert the text values to dummy variables.
2. Dummy variables are variables that take on the values of zero or one. If a categorical field has two values, such as male or female, then you need one dummy variable, for example, Male = 1 for males, Male = 0 for females.
3. In this example there are three ranks for professor. We will create a dummy for full professor, Full = 1 if the person is a full professor, Full = 0 if not. We also create a dummy for associate professor, Associate = 1 if the person is an associate professor, Associate = 0 if not. In general, when you have n categories for variable, you need n-1 dummy variables for that variable.

4. Create the first dummy variable in column E and call it Male.
    a. Enter Male in E1
    b. An if statement tests a cell and returns a value. The format for an if statement is:

        =if(cell=”condition”,1,0)

        where “cell” is the cell address and condition is the value you are testing it against ; if the cell=condition, then it returns a 1, else it returns a 0

    c. In E2, add the following if statement.

        =if(B2=”male”,1,0)

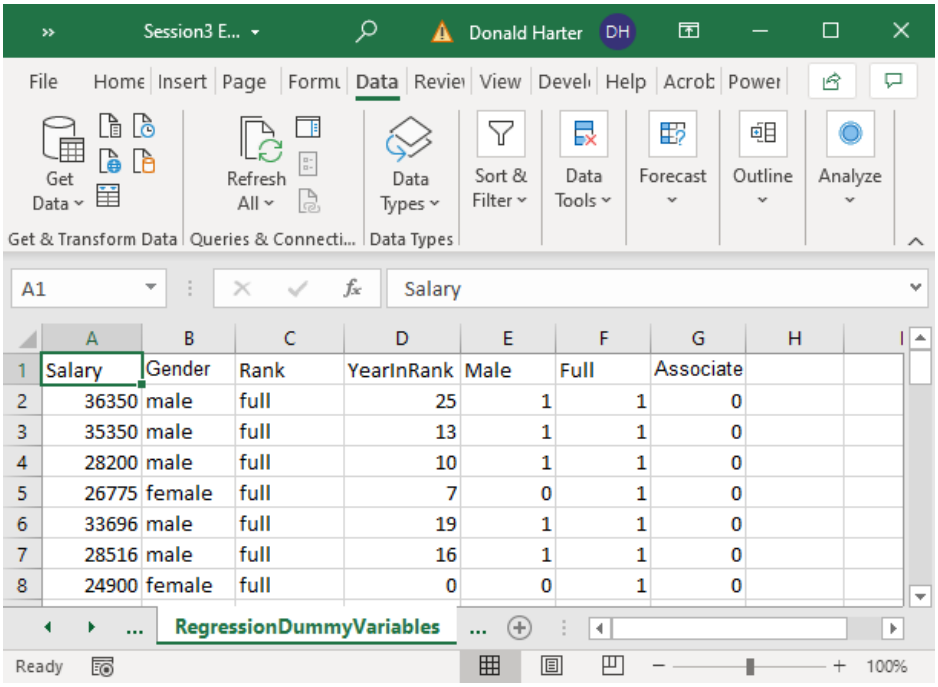
5. Create a dummy for Full Professor.
    a. Enter Full in F1
    b. Enter the following if condition in cell F2

        =if(C2=”full”,1,0)

6. Create a dummy for Associate Professor
    a. Enter Associate in G1
    b. Enter the following if condition in cell G2

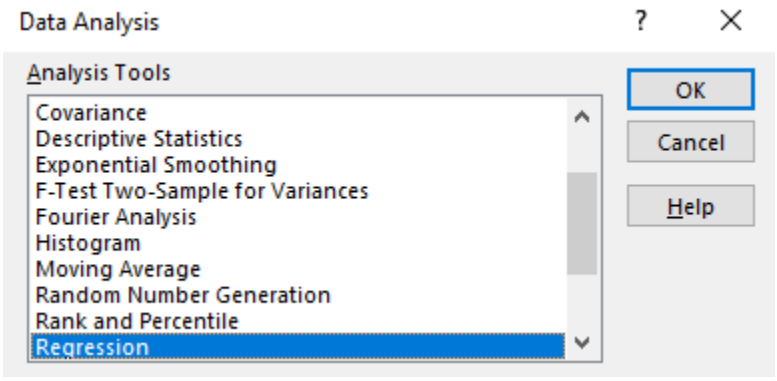
        =if(C2=”associate”,1,0)

7. Copy cells E2, F2, and G2 down for all data rows

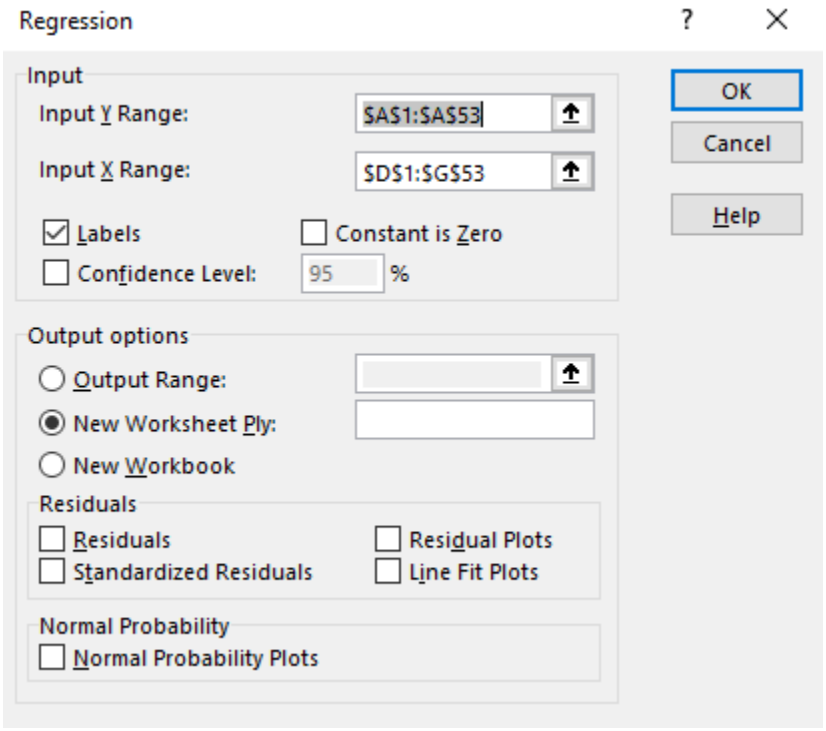


	A	B	C	D	E	F	G
1	Salary	Gender	Rank	YearInRank	Male	Full	Associate
2	36350	male	full	25	1	1	0
3	35350	male	full	13	1	1	0
4	28200	male	full	10	1	1	0
5	26775	female	full	7	0	1	0
6	33696	male	full	19	1	1	0
7	28516	male	full	16	1	1	0
8	24900	female	full	0	0	1	0

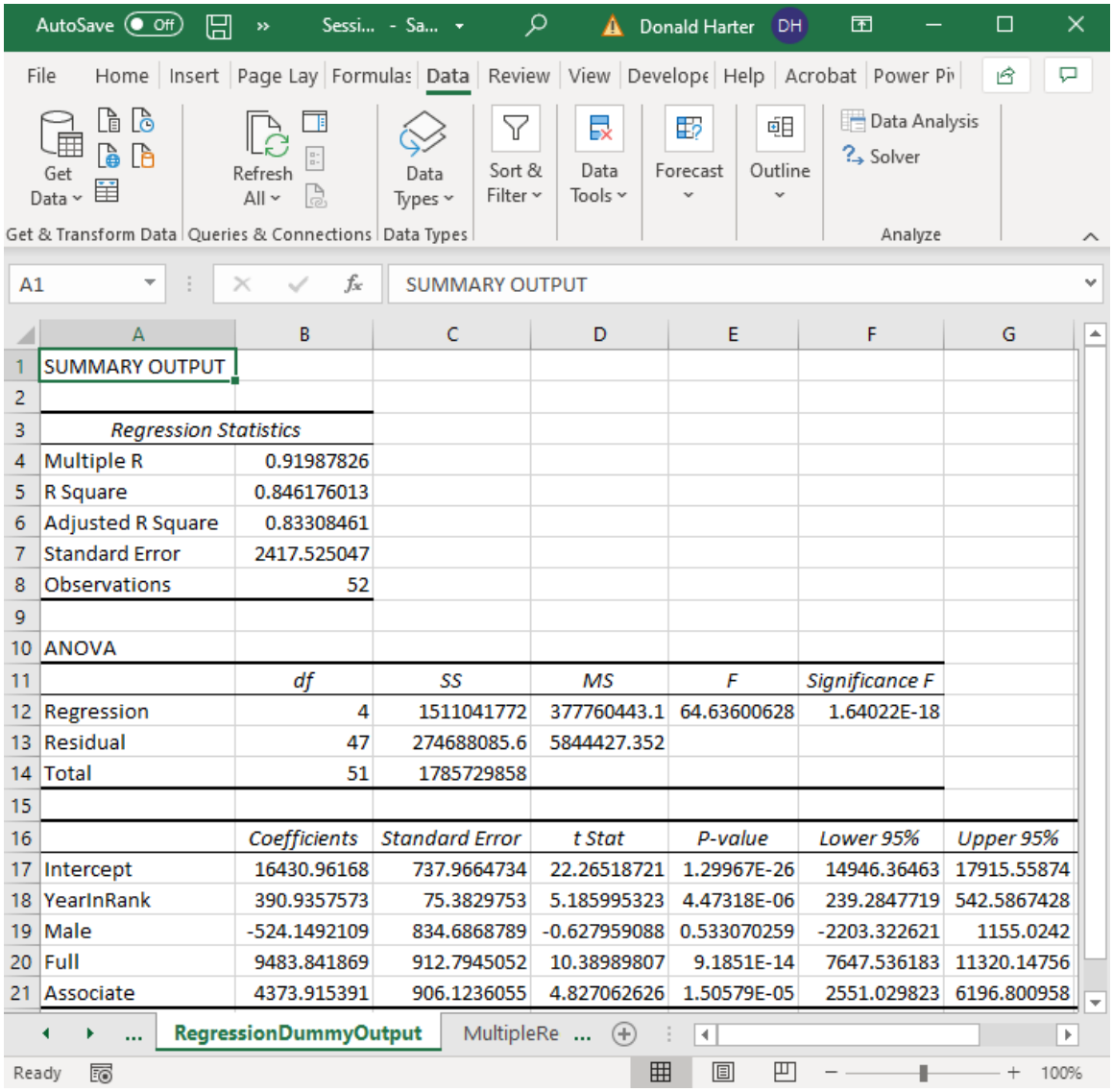
8. Run the regression of the dependent variable Salary on the independent (explanatory) variables YearInRank, Male, Full and Associate.
9. Click on Data at the top of the screen, then Data Analysis, then Regression.



10. In the Regression pop-up menu
    a. For Input Y range, enter A1:A53
    b. For Input X range, enter D1:G53
    c. Check the box for Labels
    d. Click OK



11. Interpret the equation:
    a. Is the equation significant?
    b. What does the R-square mean?
    c. Which coefficients are significant?
    d. What do the coefficients mean?

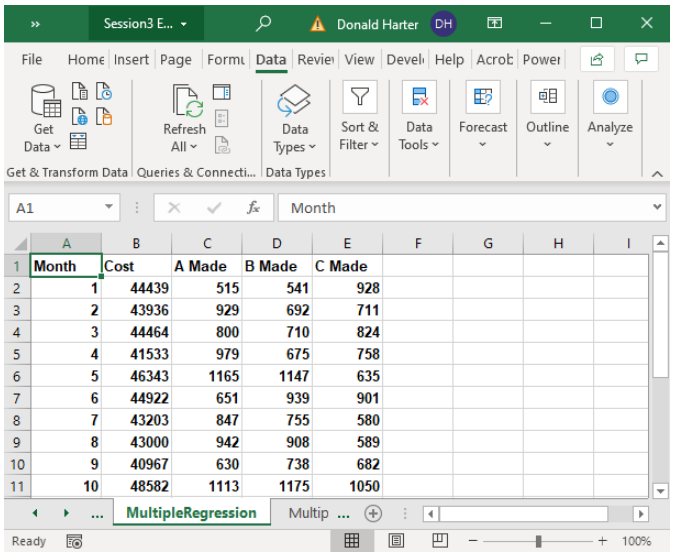


	A	B	C	D	E	F	G
1	SUMMARY OUTPUT						
2							
3	*Regression Statistics*						
4	Multiple R	0.91987826					
5	R Square	0.846176013					
6	Adjusted R Square	0.83308461					
7	Standard Error	2417.525047					
8	Observations	52					
9							
10	ANOVA						
11		*df*	*SS*	*MS*	*F*	*Significance F*	
12	Regression	4	1511041772	377760443.1	64.63600628	1.64022E-18	
13	Residual	47	274688085.6	5844427.352			
14	Total	51	1785729858				
15							
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
17	Intercept	16430.96168	737.9664734	22.26518721	1.29967E-26	14946.36463	17915.55874
18	YearInRank	390.9357573	75.3829753	5.185995323	4.47318E-06	239.2847719	542.5867428
19	Male	-524.1492109	834.6868789	-0.627959088	0.533070259	-2203.322621	1155.0242
20	Full	9483.841869	912.7945052	10.38989807	9.1851E-14	7647.536183	11320.14756
21	Associate	4373.915391	906.1236055	4.827062626	1.50579E-05	2551.029823	6196.800958

12. The categorical variables that are not listed become part of the intercept and are called the base. For example, assistant professor is not listed; its intercept is 16,430.96.
13. The categorical variables that are listed are added to the base intercept. For example, the intercept for Full professors is 16430.96 + 9483.84. The intercept for Associate professors is 16430.96 + 4373.92.
14. Which level of professor makes more?
15. Which gender makes more?

Session 3.4: Multiple regression with prediction and sensitivity analysis

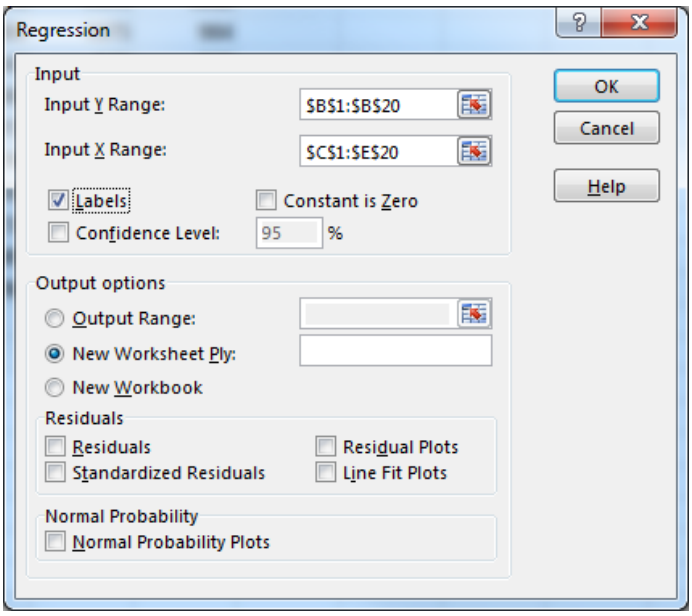
Multiple regression includes several independent variables. Use the Multiple Regression spreadsheet.

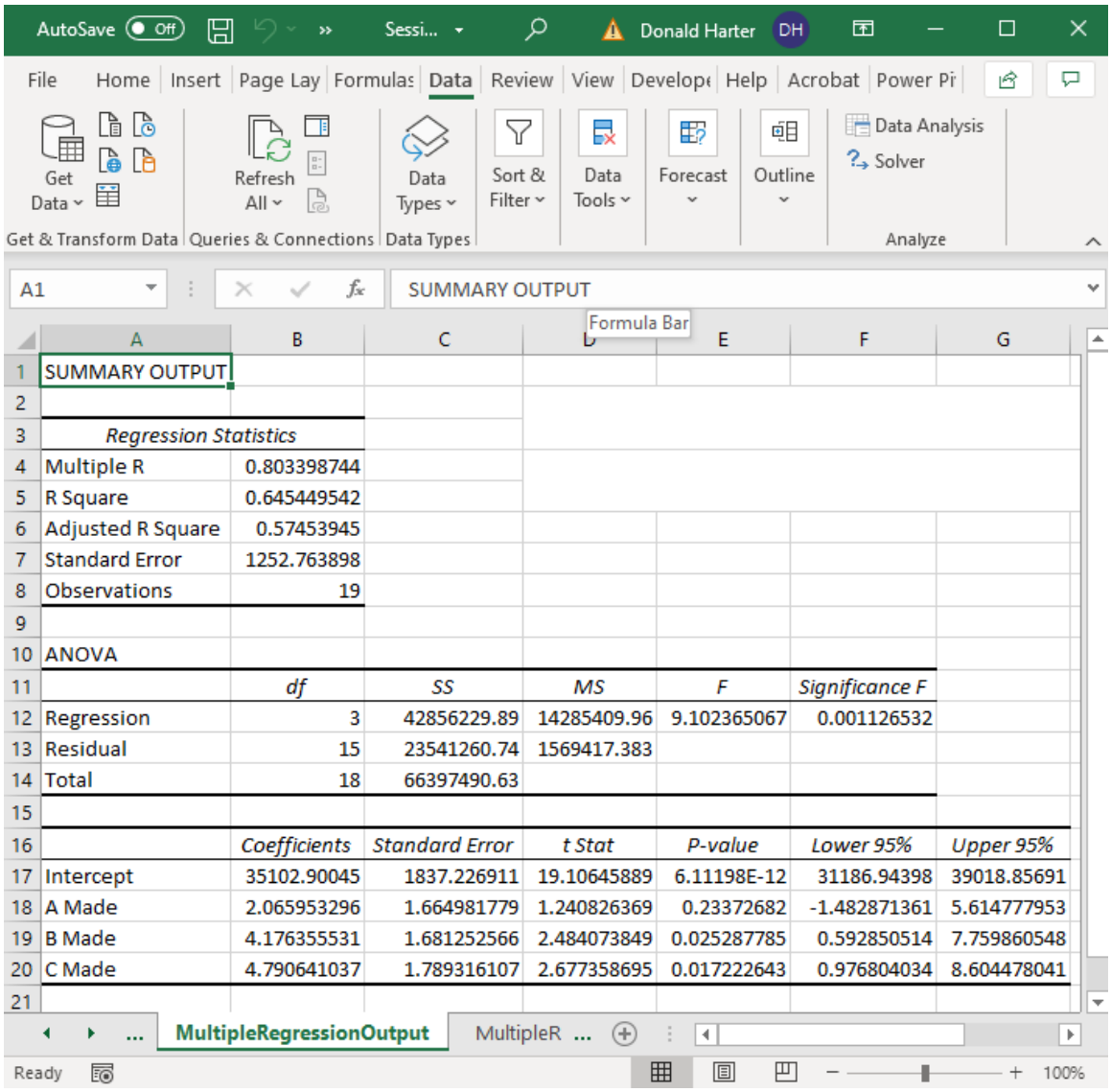


	A	B	C	D	E
1	Month	Cost	A Made	B Made	C Made
2	1	44439	515	541	928
3	2	43936	929	692	711
4	3	44464	800	710	824
5	4	41533	979	675	758
6	5	46343	1165	1147	635
7	6	44922	651	939	901
8	7	43203	847	755	580
9	8	43000	942	908	589
10	9	40967	630	738	682
11	10	48582	1113	1175	1050

To run a multiple regression:

1. Click on the data tab, data analysis, regression, then OK.
2. For the Y-range, highlight the values in the B column for cost
3. For the X-range, highlight the values in the C, D, and E columns.
4. If you included the headings at the top of the columns, click labels.
5. Click OK.





	A	B	C	D	E	F	G
1	SUMMARY OUTPUT						
2							
3	*Regression Statistics*						
4	Multiple R	0.803398744					
5	R Square	0.645449542					
6	Adjusted R Square	0.57453945					
7	Standard Error	1252.763898					
8	Observations	19					
9							
10	ANOVA						
11		*df*	*SS*	*MS*	*F*	*Significance F*	
12	Regression	3	42856229.89	14285409.96	9.102365067	0.001126532	
13	Residual	15	23541260.74	1569417.383			
14	Total	18	66397490.63				
15							
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
17	Intercept	35102.90045	1837.226911	19.10645889	6.11198E-12	31186.94398	39018.85691
18	A Made	2.065953296	1.664981779	1.240826369	0.23372682	-1.482871361	5.614777953
19	B Made	4.176355531	1.681252566	2.484073849	0.025287785	0.592850514	7.759860548
20	C Made	4.790641037	1.789316107	2.677358695	0.017222643	0.976804034	8.604478041
21							

6. Is the equation any good? How do you know?
7. What does the R-square mean in this example in business terms?
8. Which coefficients are significant?
9. What do the coefficients mean in business terms?

Session 3.5: Prediction Models

A prediction model allows you to enter values for each of the inputs (independent variables or X variables) and make a prediction of the outcome (Y variable or dependent variable).

The general form of the equation is:

$$Y = \beta_0 + \beta_1*X_1 + \beta_2*X_2 + \beta_3*X_3 + \ldots.$$

For the regression results above, we have:

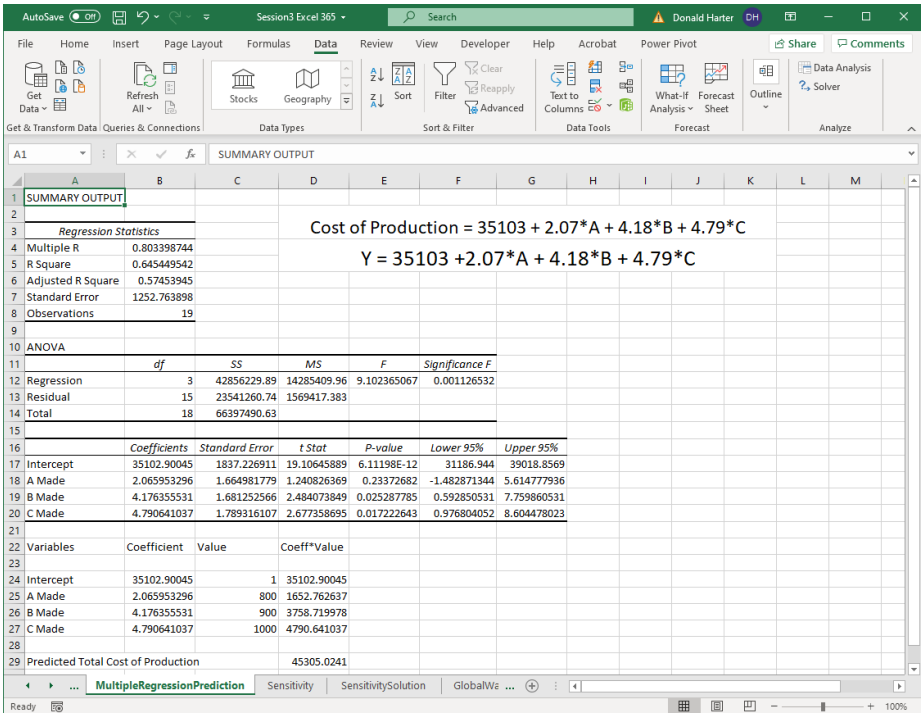
$$Y = 35103 + 2.07*X_1 + 4.18*X_2 + 4.79*X_3$$

Or:

$$\text{Cost of Production} = 35103 + 2.07*A + 4.18*B + 4.79*C$$

Let's now build the prediction model:

1. In cell A22, enter Variables
2. In cells A24:A27 enter Intercept, A Made, B Made, C Made
3. In cells B22, enter Coefficient
4. In cells B24:B27, copy the coefficients from the regression
5. In cell C22, enter Values
6. In cells C24:C27, enter 1, 800, 900, 1000. Note that the value for the intercept should be one; the values for the X variables must be in the range of the original data
7. In cell D22, enter Coeff*Value
8. In cell D24, enter the formula =B24*C24
9. Copy the formula from D24 to D25:D27
10. In cell A29, enter Predicted Total Cost of Production
11. In cell D29, enter the formula =sum(D24:D27)



Cost of Production = 35103 + 2.07*A + 4.18*B + 4.79*C

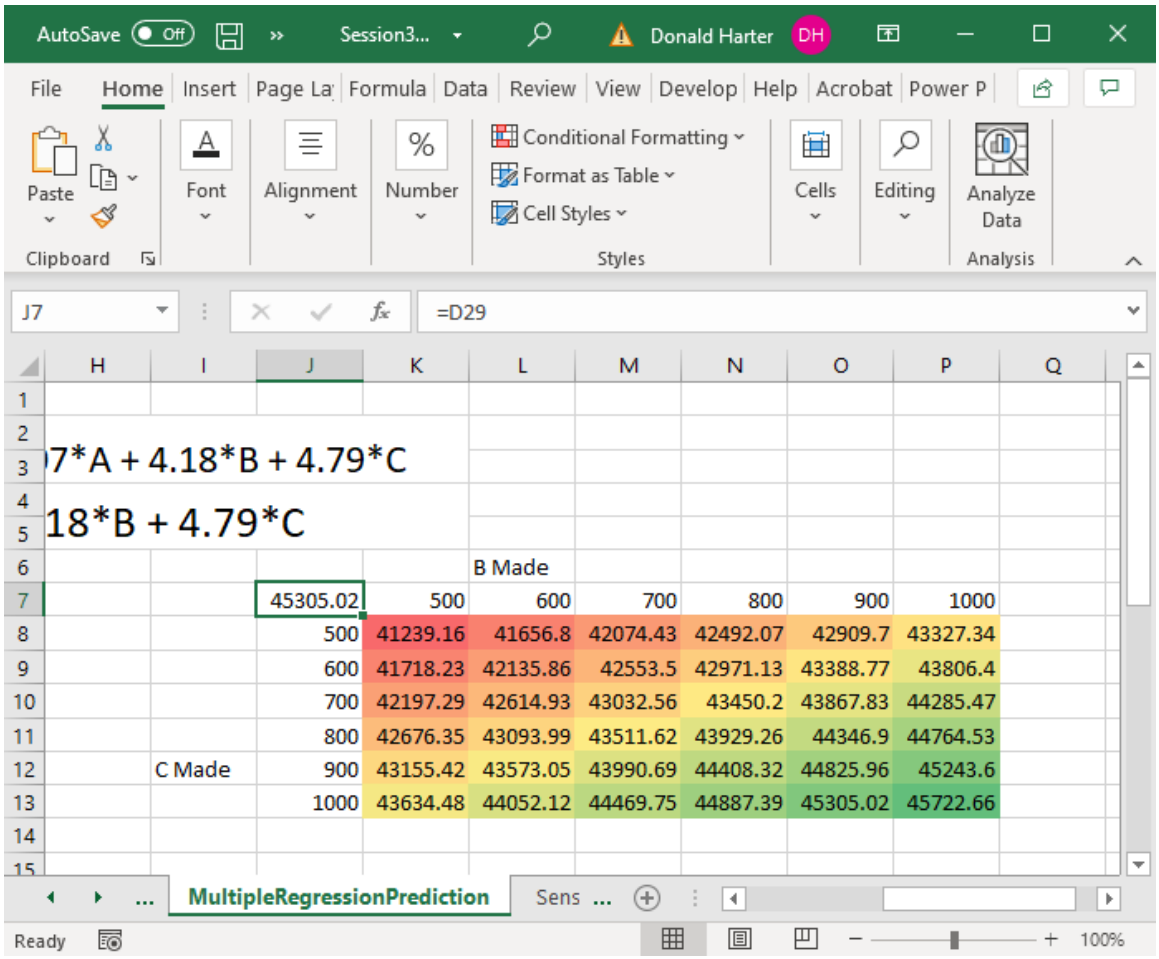
Y = 35103 +2.07*A + 4.18*B + 4.79*C

	A	B	C	D	E	F	G
1	SUMMARY OUTPUT						
2							
3	*Regression Statistics*						
4	Multiple R	0.803398744					
5	R Square	0.645449542					
6	Adjusted R Square	0.57453945					
7	Standard Error	1252.763898					
8	Observations	19					
9							
10	ANOVA						
11		*df*	*SS*	*MS*	*F*	*Significance F*	
12	Regression	3	42856229.89	14285409.96	9.102365067	0.001126532	
13	Residual	15	23541260.74	1569417.383			
14	Total	18	66397490.63				
15							
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
17	Intercept	35102.90045	1837.226911	19.10645889	6.11198E-12	31186.944	39018.8569
18	A Made	2.065953296	1.664981779	1.240826369	0.23372682	-1.482871344	5.614777936
19	B Made	4.176355531	1.681252566	2.484073849	0.025287785	0.592850531	7.759860531
20	C Made	4.790641037	1.789316107	2.677358695	0.017222643	0.976804052	8.604478023
21							
22	Variables	Coefficient	Value	Coeff*Value			
23							
24	Intercept	35102.90045	1	35102.90045			
25	A Made	2.065953296	800	1652.762637			
26	B Made	4.176355531	900	3758.719978			
27	C Made	4.790641037	1000	4790.641037			
28							
29	Predicted Total Cost of Production			45305.0241			

Session 3.6: Sensitivity Analysis of Regression Results

After you build the prediction model, you can create a sensitivity analysis of the regression results.

1. In cell L6, enter B Made
2. In cells K7:R7, enter 600, 700, …, 1000
3. In cell I12, enter C Made
4. In cells J8:J15, enter 600, 700, …, 1000
5. In cell J7, enter a formula that points to Predicted Cost, =D29
6. Highlight the data, cells J7:R15
7. Click on the Data tab, What if analysis, Data table
8. For Row input cell, enter C26, the value for B Made
9. For Column input cell, enter C27, the value for C Made
10. Click OK
11. To add conditional formatting, highlight the production costs, cells K8:R15
12. Click on the Home tab
13. Click on Conditional Formatting, Color Scales, your choice of color



	H	I	J	K	L	M	N	O	P	Q
1										
2										
3	7*A + 4.18*B + 4.79*C									
4										
5	18*B + 4.79*C									
6					B Made					
7			45305.02	500	600	700	800	900	1000	
8			500	41239.16	41656.8	42074.43	42492.07	42909.7	43327.34	
9			600	41718.23	42135.86	42553.5	42971.13	43388.77	43806.4	
10			700	42197.29	42614.93	43032.56	43450.2	43867.83	44285.47	
11			800	42676.35	43093.99	43511.62	43929.26	44346.9	44764.53	
12		C Made	900	43155.42	43573.05	43990.69	44408.32	44825.96	45243.6	
13			1000	43634.48	44052.12	44469.75	44887.39	45305.02	45722.66	
14										
15										

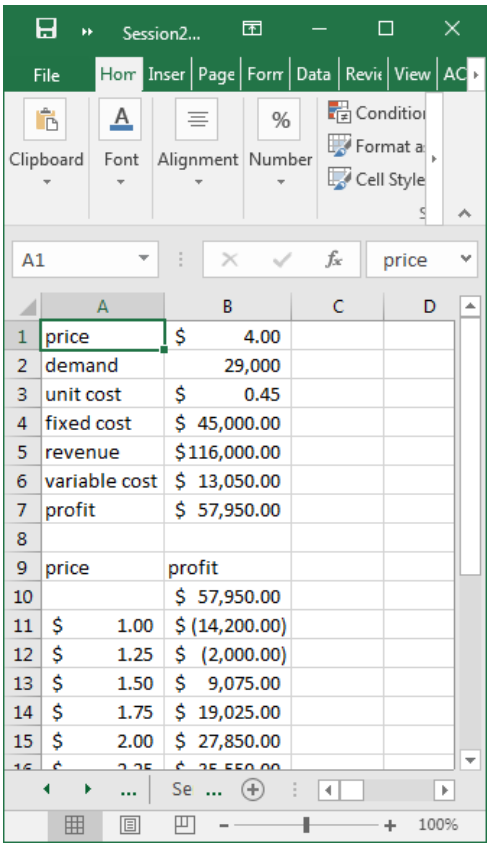
Session 3.7: Sensitivity Analysis – One-way and Two-way Sensitivity

Sensitivity analysis allows you to see the effect of a changing variable on an outcome variable. In the example below, use the spreadsheet tab Sensitivity.

The spreadsheet provides the demand curve, costs, revenue, and profit for a widget. By changing the price, you can observe the effect on sales and profitability. This approach only allows you to look at one price at a time.

The following steps build a table to examine the effect price has on profits, revenue and costs.

1. To set up a one-way table, enter the labels of your one input variable and all of the output variables in a row. In this example, enter price in A9 and profit in B9.
2. Next, enter a column of prices under price **but skip the first cell below the word price**. In this example, enter a column for price with price ranging from $1 to $4 by $0.25 increments.
3. Immediately under profit in cell B10, enter =B7, so the profit calculation is filled in.
4. Finally, we will create a data table. Highlight the fields immediately under the labels and continue until all of the prices are highlighted. In this case, we highlight A10:B23. Click on the Data tab at the top of the page, click on What-If Analysis, Data Table, then enter in the column input cell the cell reference for variable that you want to vary (i.e., B1, which is price), then click OK. This should populate the data table.



	A	B	C	D
1	price	$ 4.00		
2	demand	29,000		
3	unit cost	$ 0.45		
4	fixed cost	$ 45,000.00		
5	revenue	$116,000.00		
6	variable cost	$ 13,050.00		
7	profit	$ 57,950.00		
8				
9	price	profit		
10		$ 57,950.00		
11	$ 1.00	$ (14,200.00)		
12	$ 1.25	$ (2,000.00)		
13	$ 1.50	$ 9,075.00		
14	$ 1.75	$ 19,025.00		
15	$ 2.00	$ 27,850.00		

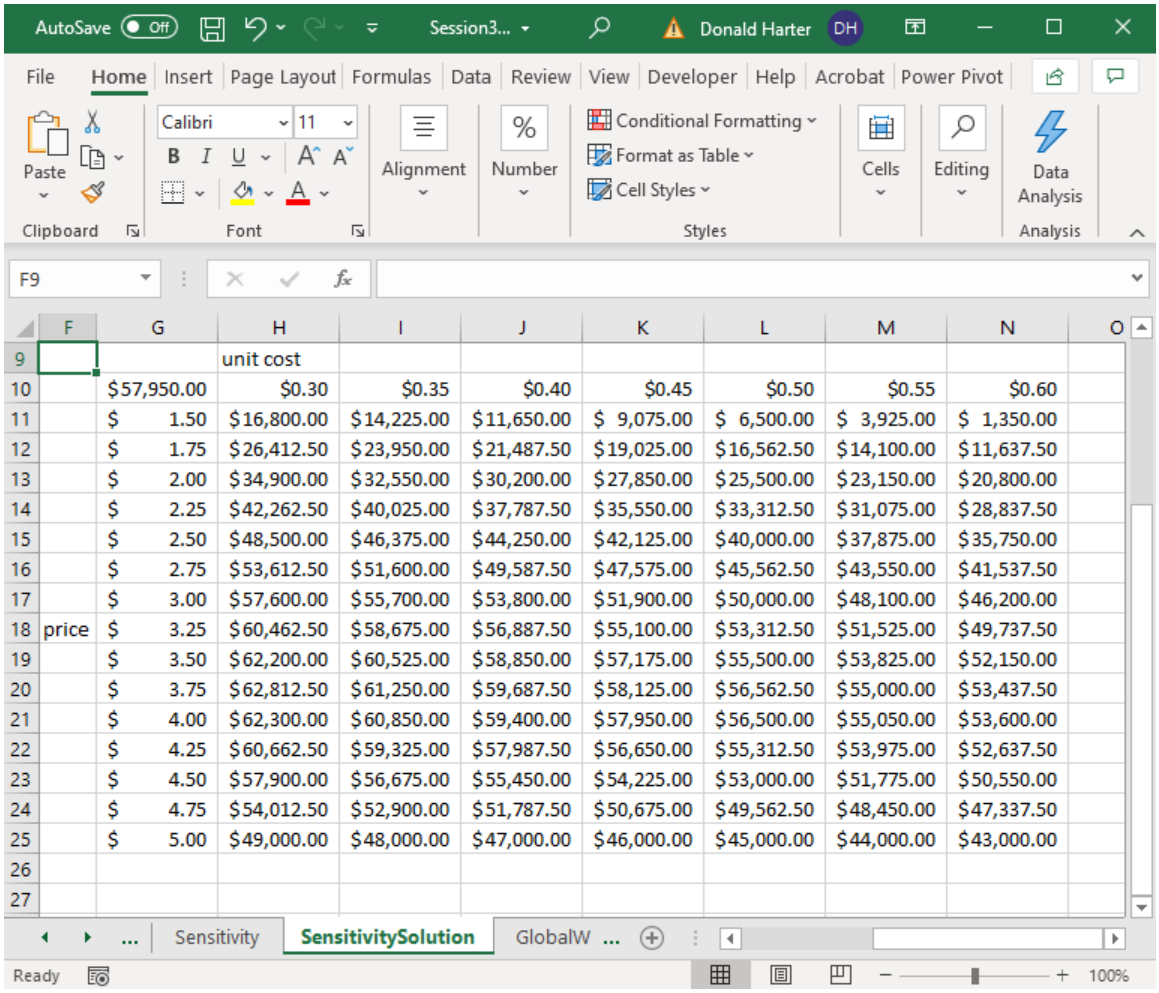
What you really did: The price column represents on which prices you would like to perform a what-if analysis, and the three formulas are the outcomes that you want to analyze when price

changes. The first step was to set up the price points and the formulas. When you created the data table, you identified what variable changes, but used an absolute address.

If you want to change two variables at once, we create a two-way table. In this example, let's change both price and unit cost for our widgets. The result that we want will look like the picture below.

Note that in a two-way table, you can only have one outcome variable. In this case, let's use profit as the outcome variable. The following steps create a two-way what-if analysis:

1. As before, enter a column for prices. In the picture, the prices are in cells G11:G23. You can type a label of price to the left of any cell. Enter $1.50 to $5.00 in $.25 increments.
2. Enter variable costs in H10:N10 from $.30 to $.60 in increments of $.05. You can enter a label for variable costs above the first entry (cell H9).
3. The price column and variable cost row will form the boundary of the table.
4. Now enter the formula that you want to calculate where the column and row intersect, in this case, cell G10. The formula we want to use is for profit; enter =B7.
5. Now highlight the entire table of prices and variable costs G10:N25. Click on the Data tab at the top of the spreadsheet, then What-if Analysis and Data Table. Since the variable that we are varying across the row is unit cost, and unit cost is a named variable, enter unit cost for the row variable. Likewise, since price varies down the column and is named variable, enter price as the column variable. Click OK.

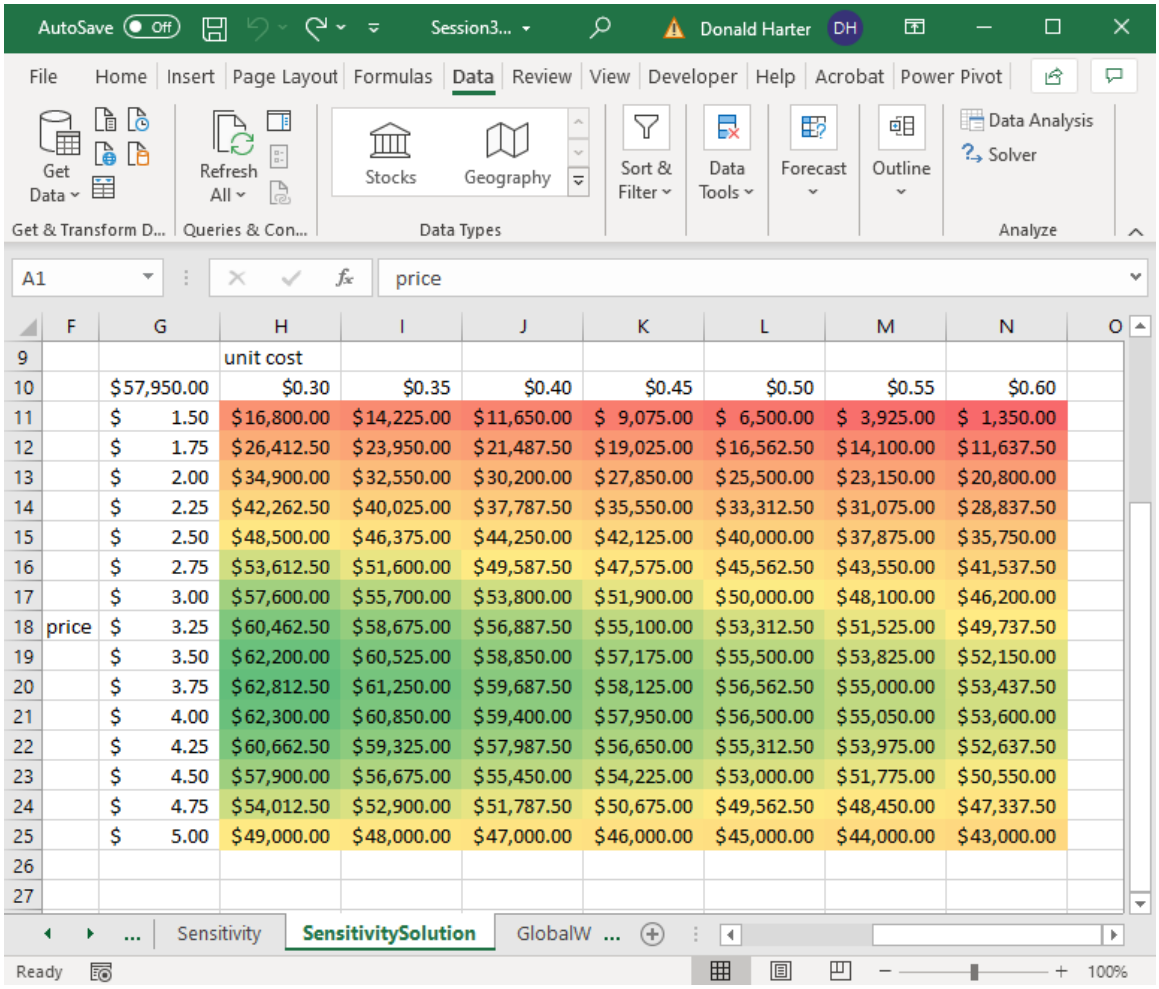


	F	G	H	I	J	K	L	M	N	O
9			unit cost							
10		$57,950.00	$0.30	$0.35	$0.40	$0.45	$0.50	$0.55	$0.60	
11		$ 1.50	$16,800.00	$14,225.00	$11,650.00	$ 9,075.00	$ 6,500.00	$ 3,925.00	$ 1,350.00	
12		$ 1.75	$26,412.50	$23,950.00	$21,487.50	$19,025.00	$16,562.50	$14,100.00	$11,637.50	
13		$ 2.00	$34,900.00	$32,550.00	$30,200.00	$27,850.00	$25,500.00	$23,150.00	$20,800.00	
14		$ 2.25	$42,262.50	$40,025.00	$37,787.50	$35,550.00	$33,312.50	$31,075.00	$28,837.50	
15		$ 2.50	$48,500.00	$46,375.00	$44,250.00	$42,125.00	$40,000.00	$37,875.00	$35,750.00	
16		$ 2.75	$53,612.50	$51,600.00	$49,587.50	$47,575.00	$45,562.50	$43,550.00	$41,537.50	
17		$ 3.00	$57,600.00	$55,700.00	$53,800.00	$51,900.00	$50,000.00	$48,100.00	$46,200.00	
18	price	$ 3.25	$60,462.50	$58,675.00	$56,887.50	$55,100.00	$53,312.50	$51,525.00	$49,737.50	
19		$ 3.50	$62,200.00	$60,525.00	$58,850.00	$57,175.00	$55,500.00	$53,825.00	$52,150.00	
20		$ 3.75	$62,812.50	$61,250.00	$59,687.50	$58,125.00	$56,562.50	$55,000.00	$53,437.50	
21		$ 4.00	$62,300.00	$60,850.00	$59,400.00	$57,950.00	$56,500.00	$55,050.00	$53,600.00	
22		$ 4.25	$60,662.50	$59,325.00	$57,987.50	$56,650.00	$55,312.50	$53,975.00	$52,637.50	
23		$ 4.50	$57,900.00	$56,675.00	$55,450.00	$54,225.00	$53,000.00	$51,775.00	$50,550.00	
24		$ 4.75	$54,012.50	$52,900.00	$51,787.50	$50,675.00	$49,562.50	$48,450.00	$47,337.50	
25		$ 5.00	$49,000.00	$48,000.00	$47,000.00	$46,000.00	$45,000.00	$44,000.00	$43,000.00	
26										
27										

Conditional formatting

Conditional formatting allows you to display data differently based on the values of the data. For example, you can change the font color, background color, or other characteristics to make the data jump out at you when it satisfies certain conditions. Let us convert the data in the Sensitivity spreadsheet to something more visual.

1. Continuing to use the Sensitivity data, highlight cells H11:N25.
2. Click on the Home tab, Conditional Formatting, Color Scales, then select a color scheme from the right.

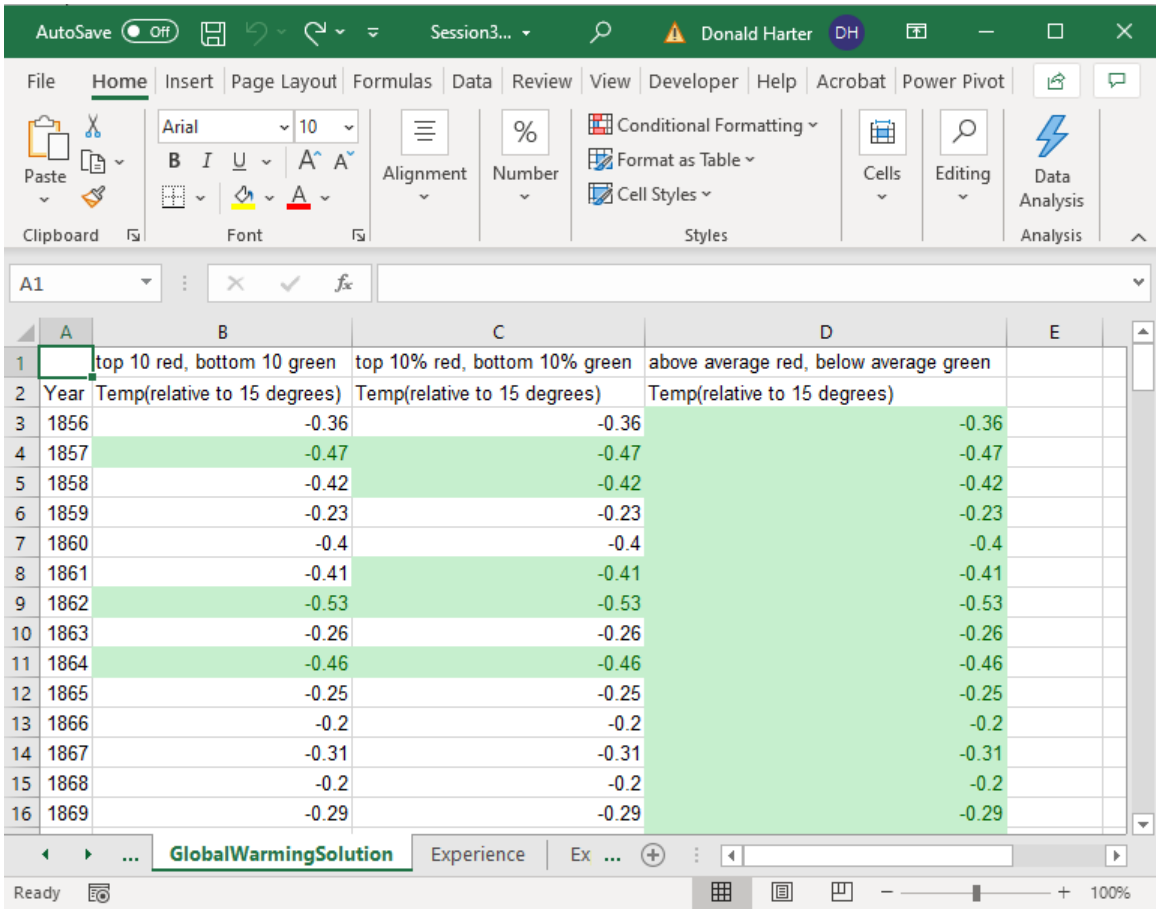


	F	G	H	I	J	K	L	M	N	O
9			unit cost							
10		$57,950.00	$0.30	$0.35	$0.40	$0.45	$0.50	$0.55	$0.60	
11		$ 1.50	$16,800.00	$14,225.00	$11,650.00	$ 9,075.00	$ 6,500.00	$ 3,925.00	$ 1,350.00	
12		$ 1.75	$26,412.50	$23,950.00	$21,487.50	$19,025.00	$16,562.50	$14,100.00	$11,637.50	
13		$ 2.00	$34,900.00	$32,550.00	$30,200.00	$27,850.00	$25,500.00	$23,150.00	$20,800.00	
14		$ 2.25	$42,262.50	$40,025.00	$37,787.50	$35,550.00	$33,312.50	$31,075.00	$28,837.50	
15		$ 2.50	$48,500.00	$46,375.00	$44,250.00	$42,125.00	$40,000.00	$37,875.00	$35,750.00	
16		$ 2.75	$53,612.50	$51,600.00	$49,587.50	$47,575.00	$45,562.50	$43,550.00	$41,537.50	
17		$ 3.00	$57,600.00	$55,700.00	$53,800.00	$51,900.00	$50,000.00	$48,100.00	$46,200.00	
18	price	$ 3.25	$60,462.50	$58,675.00	$56,887.50	$55,100.00	$53,312.50	$51,525.00	$49,737.50	
19		$ 3.50	$62,200.00	$60,525.00	$58,850.00	$57,175.00	$55,500.00	$53,825.00	$52,150.00	
20		$ 3.75	$62,812.50	$61,250.00	$59,687.50	$58,125.00	$56,562.50	$55,000.00	$53,437.50	
21		$ 4.00	$62,300.00	$60,850.00	$59,400.00	$57,950.00	$56,500.00	$55,050.00	$53,600.00	
22		$ 4.25	$60,662.50	$59,325.00	$57,987.50	$56,650.00	$55,312.50	$53,975.00	$52,637.50	
23		$ 4.50	$57,900.00	$56,675.00	$55,450.00	$54,225.00	$53,000.00	$51,775.00	$50,550.00	
24		$ 4.75	$54,012.50	$52,900.00	$51,787.50	$50,675.00	$49,562.50	$48,450.00	$47,337.50	
25		$ 5.00	$49,000.00	$48,000.00	$47,000.00	$46,000.00	$45,000.00	$44,000.00	$43,000.00	
26										
27										

Session 3.8: Conditional formatting rules and options

For the next example, use the Global Warming spreadsheet. In the column labeled “top ten red…”, we will use conditional formatting to change the background based on the values in the column. To do this, use the following steps:

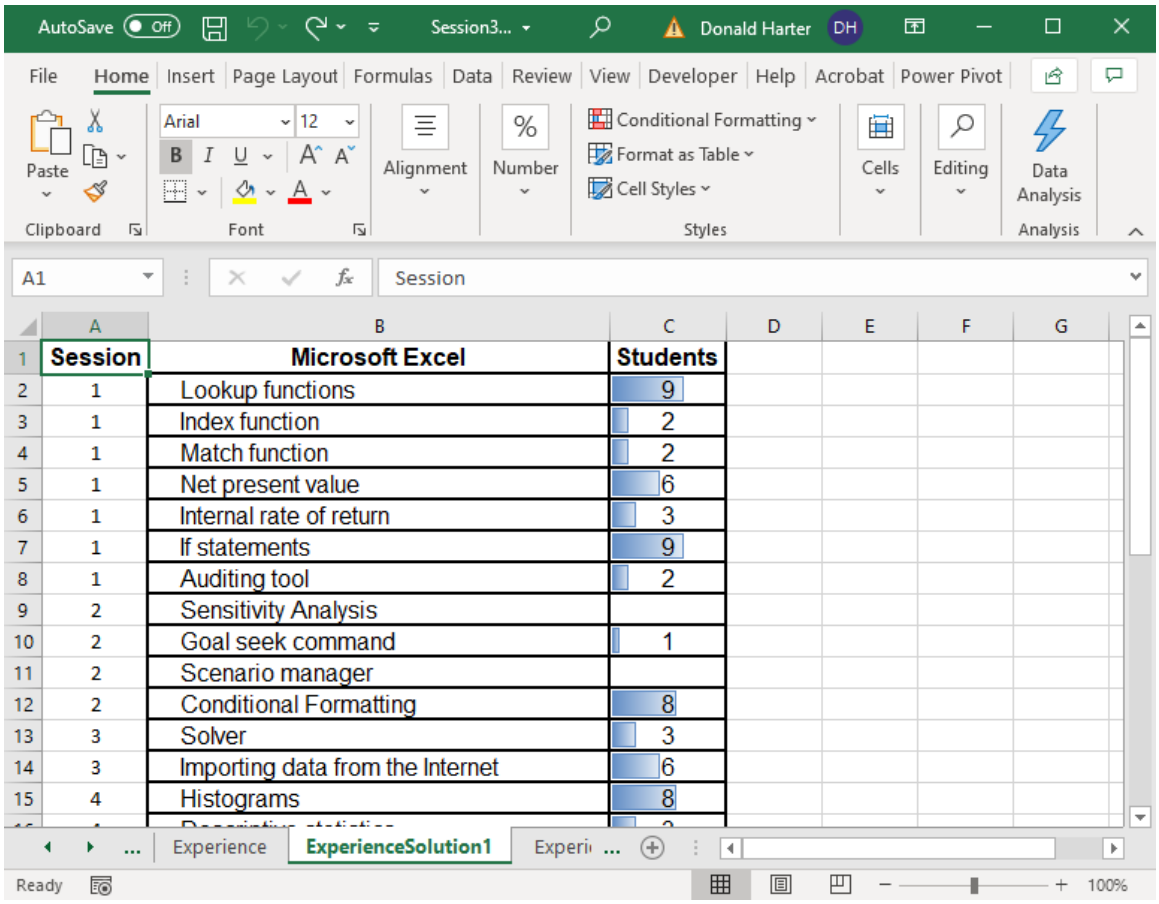
1. Highlight the values B3:B152
2. Click on the Home tab at the top of your spreadsheet.
3. Click on Conditional Formatting, Top/Bottom Rules, Top 10. Note that in the next screen, you can change top 10 to top anything. Select red as the color, then OK.
4. Click on Conditional Formatting a second time, Top/Bottom rules, Bottom 10. Change the font and background color to match the column heading.
5. For the next column, use the same approach but select the top and bottom 10%.
6. For the last column, use the same approach but select above and below average.



	A	B	C	D	E
1		top 10 red, bottom 10 green	top 10% red, bottom 10% green	above average red, below average green	
2	Year	Temp(relative to 15 degrees)	Temp(relative to 15 degrees)	Temp(relative to 15 degrees)	
3	1856	-0.36	-0.36	-0.36	
4	1857	-0.47	-0.47	-0.47	
5	1858	-0.42	-0.42	-0.42	
6	1859	-0.23	-0.23	-0.23	
7	1860	-0.4	-0.4	-0.4	
8	1861	-0.41	-0.41	-0.41	
9	1862	-0.53	-0.53	-0.53	
10	1863	-0.26	-0.26	-0.26	
11	1864	-0.46	-0.46	-0.46	
12	1865	-0.25	-0.25	-0.25	
13	1866	-0.2	-0.2	-0.2	
14	1867	-0.31	-0.31	-0.31	
15	1868	-0.2	-0.2	-0.2	
16	1869	-0.29	-0.29	-0.29	

A feature of Excel is data bars to enhance data display. Data bars, color scales and icon sets can be used to enhance tabular data. Using the spreadsheet Experience, the steps to implement these new features are listed below.

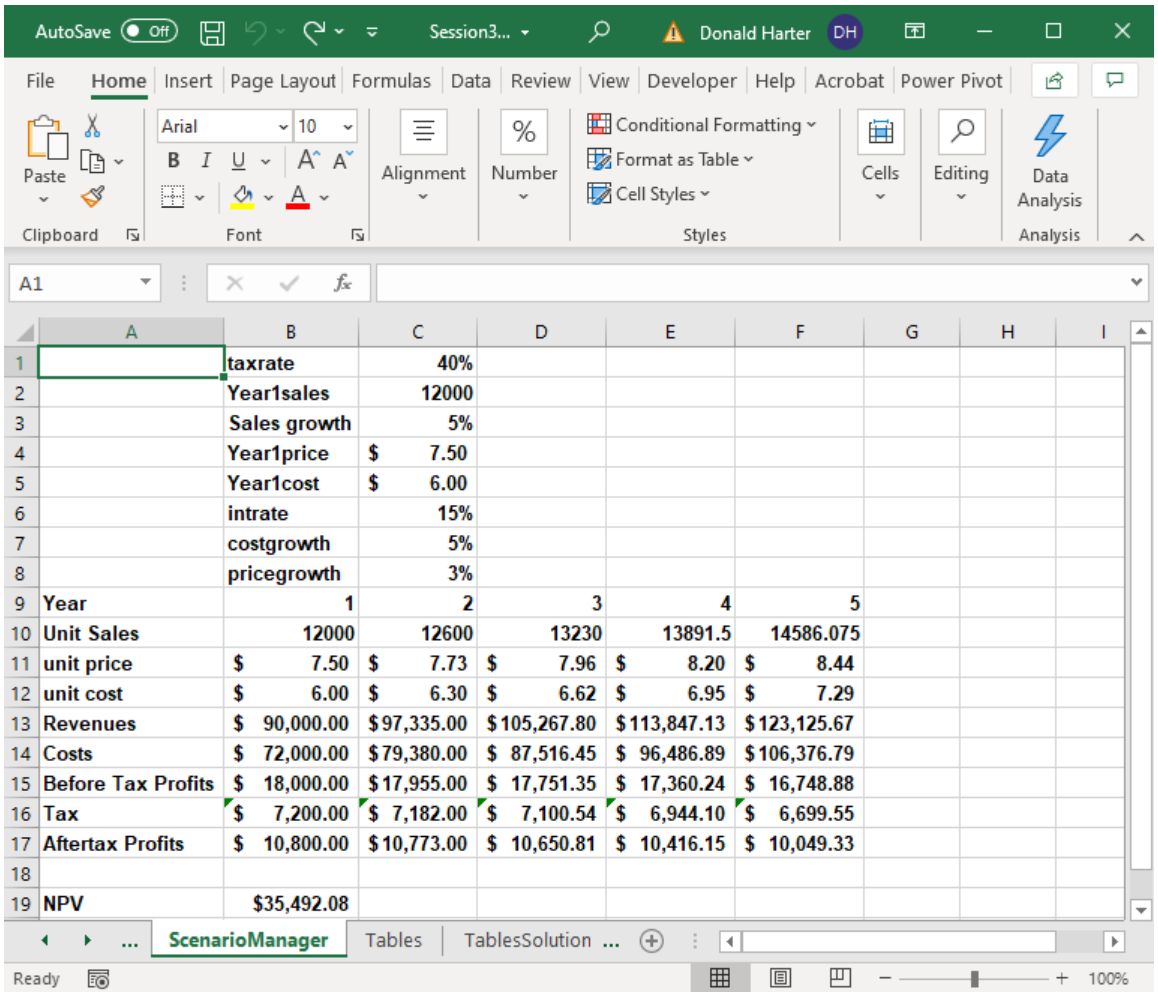
1. Select the range of values under the title students.
2. Click on Conditional Formatting.
3. First try Data Bars and place your cursor over the options that appear on the right.
4. Next try Color Scales, placing your cursor over the options on the right.
5. Finally try Icon Sets, again placing your cursor over the options listed.
6. For each of these features, you can click on more rules and specify additional combinations.



Session	Microsoft Excel	Students
1	Lookup functions	9
1	Index function	2
1	Match function	2
1	Net present value	6
1	Internal rate of return	3
1	If statements	9
1	Auditing tool	2
2	Sensitivity Analysis	
2	Goal seek command	1
2	Scenario manager	
2	Conditional Formatting	8
3	Solver	3
3	Importing data from the Internet	6
4	Histograms	8

Session 3.9: Scenario Manager

The scenario manager expands on the previous examples and allows you to vary up to 32 variables simultaneously. We will use the ScenarioManager spreadsheet shown below



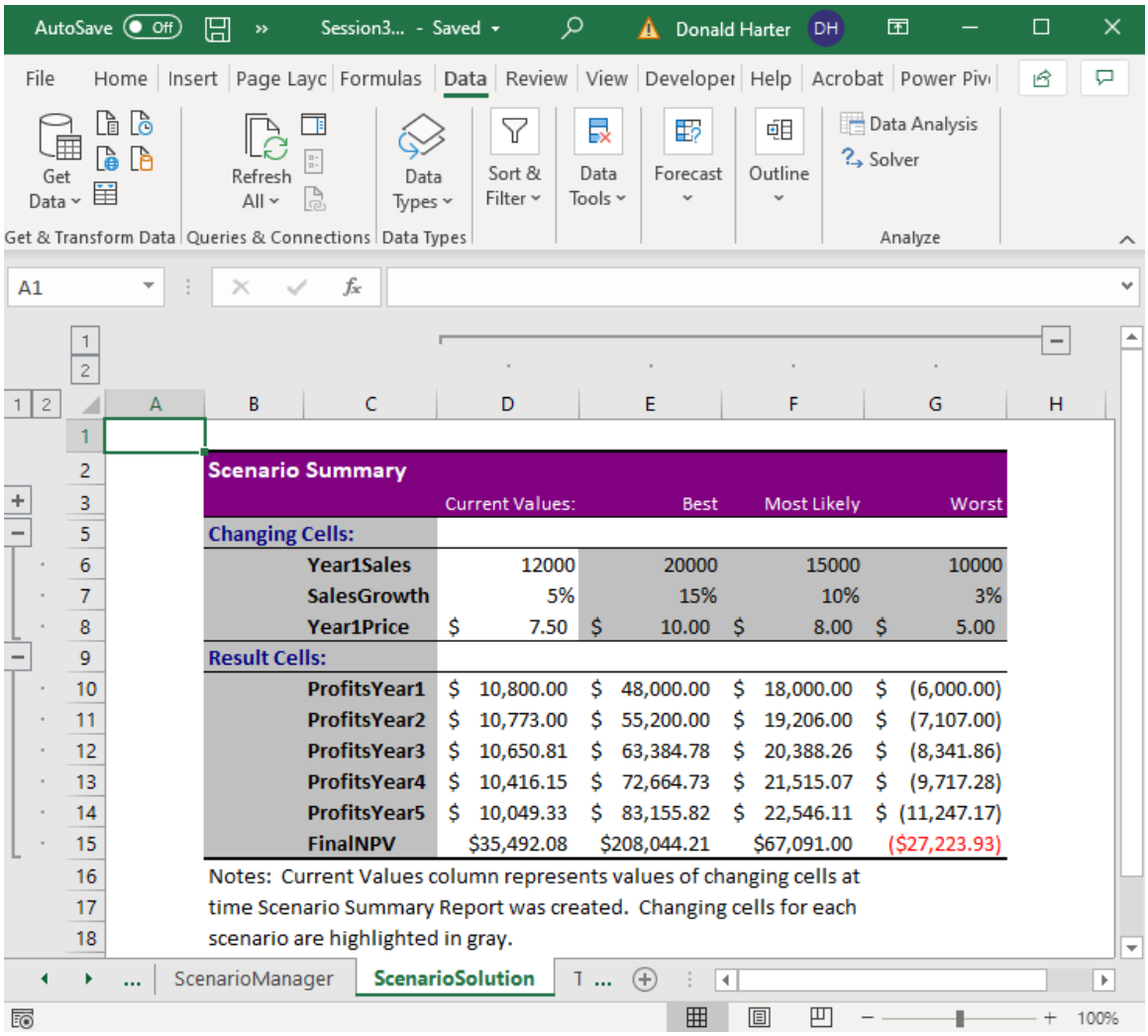
	A	B	C	D	E	F	G	H	I
1		taxrate	40%						
2		Year1sales	12000						
3		Sales growth	5%						
4		Year1price	$ 7.50						
5		Year1cost	$ 6.00						
6		intrate	15%						
7		costgrowth	5%						
8		pricegrowth	3%						
9	Year	1	2	3	4	5			
10	Unit Sales	12000	12600	13230	13891.5	14586.075			
11	unit price	$ 7.50	$ 7.73	$ 7.96	$ 8.20	$ 8.44			
12	unit cost	$ 6.00	$ 6.30	$ 6.62	$ 6.95	$ 7.29			
13	Revenues	$ 90,000.00	$97,335.00	$105,267.80	$113,847.13	$123,125.67			
14	Costs	$ 72,000.00	$79,380.00	$ 87,516.45	$ 96,486.89	$106,376.79			
15	Before Tax Profits	$ 18,000.00	$17,955.00	$ 17,751.35	$ 17,360.24	$ 16,748.88			
16	Tax	$ 7,200.00	$ 7,182.00	$ 7,100.54	$ 6,944.10	$ 6,699.55			
17	Aftertax Profits	$ 10,800.00	$10,773.00	$ 10,650.81	$ 10,416.15	$ 10,049.33			
18									
19	NPV	$35,492.08							

Let's vary Year1Sales, Sales Growth and Year1Price to determine the effect on Profits each year and Final NPV. First, assign names to the fields of interest. These steps have already been done.

1. Cell C2: Year1Sales
2. Cell C3: SalesGrowth
3. Cell C4: Year1Price
4. Cell B17: ProfitsYear1
5. Cell C17: ProfitsYear2
6. Cell D17: ProfitsYear3
7. Cell E17: ProfitsYear4
8. Cell F17: ProfitsYear5
9. Cell B19: FinalNPV

Follow these steps to create the three scenarios: Best, Most Likely, and Worst.

10. Click on the Data Tab.
11. In What-if Analysis, click on the Scenario Manager
12. When the Scenario Manager Screen pops up, there should be no scenarios. Click Add, then enter Best as the name and C2:C4 for the range. Click OK. On the next screen, enter 20000, .15, and 10, then OK.
13. Click Add, then enter a Most Likely scenario with values of 15000, .10, 8.
14. Click Add, then enter a Worst scenario with values of 10000, .03, 5.
15. When finished entering data, click on Summary to identify which fields to report. Let's use each year's after-tax profits (B17:F17) and NPV (B19). When the fields are non-contiguous, separate the groups with a comma, and click OK.
16. Excel will automatically create a scenario report in a new spreadsheet called Scenario Summary as shown below.



Scenario Summary				
	Current Values:	Best	Most Likely	Worst
Changing Cells:				
Year1Sales	12000	20000	15000	10000
SalesGrowth	5%	15%	10%	3%
Year1Price	$ 7.50	$ 10.00	$ 8.00	$ 5.00
Result Cells:				
ProfitsYear1	$ 10,800.00	$ 48,000.00	$ 18,000.00	$ (6,000.00)
ProfitsYear2	$ 10,773.00	$ 55,200.00	$ 19,206.00	$ (7,107.00)
ProfitsYear3	$ 10,650.81	$ 63,384.78	$ 20,388.26	$ (8,341.86)
ProfitsYear4	$ 10,416.15	$ 72,664.73	$ 21,515.07	$ (9,717.28)
ProfitsYear5	$ 10,049.33	$ 83,155.82	$ 22,546.11	$ (11,247.17)
FinalNPV	$35,492.08	$208,044.21	$67,091.00	($27,223.93)

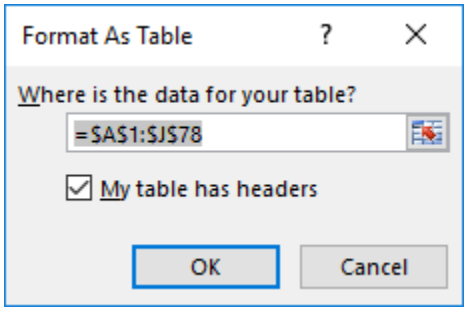
Notes: Current Values column represents values of changing cells at time Scenario Summary Report was created. Changing cells for each scenario are highlighted in gray.

Session 3.10: Tables

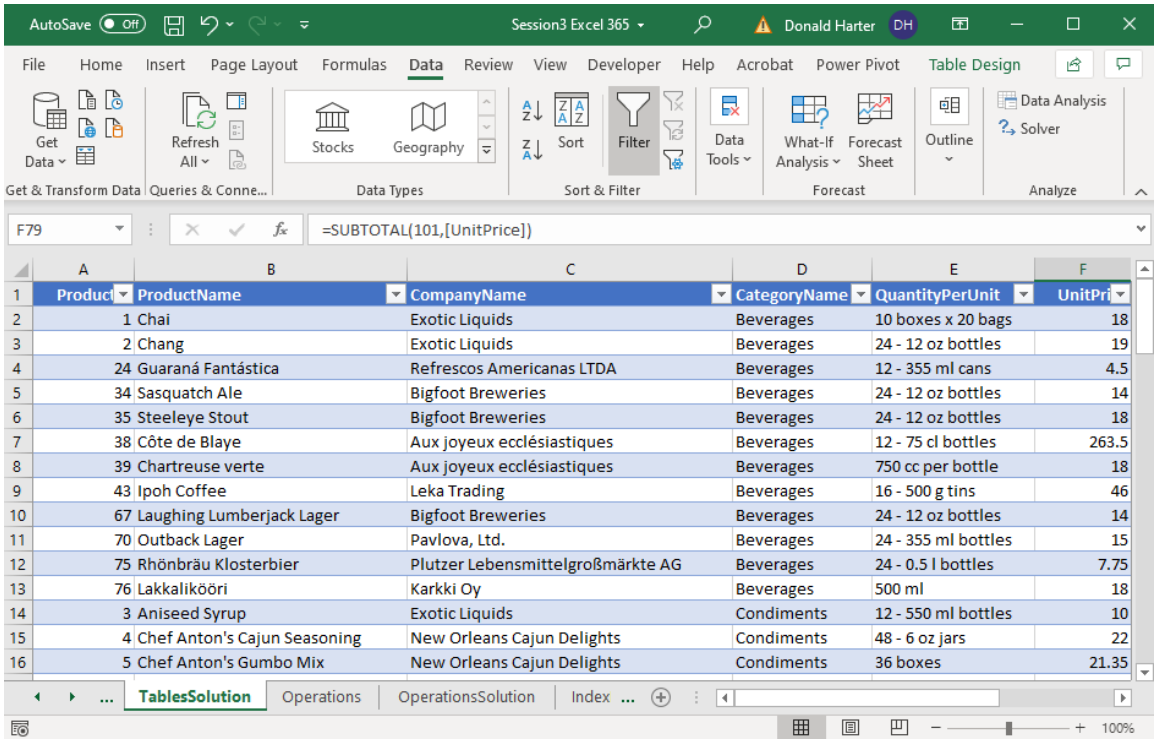
Tables allow you to format and organize data. Use the Tables worksheet for this exercise

Create and Manage Tables

1. To create an Excel Table from a set of data, highlight the data by clicking anywhere in the data, then press Control A
2. For Excel to correctly format your data as a table, the columns and rows must be contiguous. That is, there cannot be any blank columns or rows in the table area.
3. Click on the Home tab, and under Styles, click on Format as Table, select your color scheme, then confirm the data range. Check that your data has labels at the top and that the box for My table has headers is checked. Click OK.



4. The table will have your color selection and will also have drop down arrows for filters



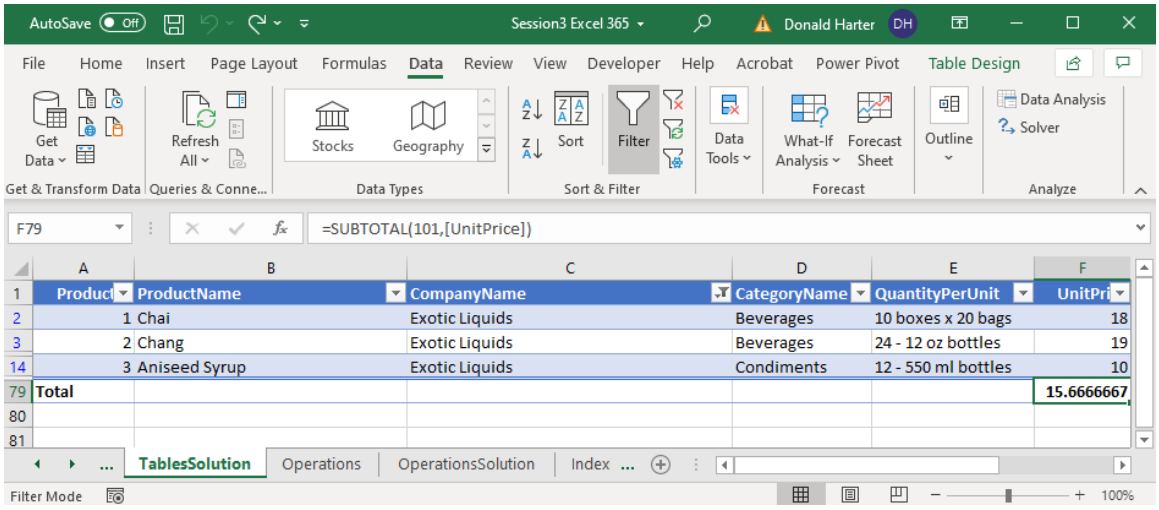
	A	B	C	D	E	F
1	Product	ProductName	CompanyName	CategoryName	QuantityPerUnit	UnitPri
2	1	Chai	Exotic Liquids	Beverages	10 boxes x 20 bags	18
3	2	Chang	Exotic Liquids	Beverages	24 - 12 oz bottles	19
4	24	Guaraná Fantástica	Refrescos Americanas LTDA	Beverages	12 - 355 ml cans	4.5
5	34	Sasquatch Ale	Bigfoot Breweries	Beverages	24 - 12 oz bottles	14
6	35	Steeleye Stout	Bigfoot Breweries	Beverages	24 - 12 oz bottles	18
7	38	Côte de Blaye	Aux joyeux ecclésiastiques	Beverages	12 - 75 cl bottles	263.5
8	39	Chartreuse verte	Aux joyeux ecclésiastiques	Beverages	750 cc per bottle	18
9	43	Ipoh Coffee	Leka Trading	Beverages	16 - 500 g tins	46
10	67	Laughing Lumberjack Lager	Bigfoot Breweries	Beverages	24 - 12 oz bottles	14
11	70	Outback Lager	Pavlova, Ltd.	Beverages	24 - 355 ml bottles	15
12	75	Rhönbräu Klosterbier	Plutzer Lebensmittelgroßmärkte AG	Beverages	24 - 0.5 l bottles	7.75
13	76	Lakkalikööri	Karkki Oy	Beverages	500 ml	18
14	3	Aniseed Syrup	Exotic Liquids	Condiments	12 - 550 ml bottles	10
15	4	Chef Anton's Cajun Seasoning	New Orleans Cajun Delights	Condiments	48 - 6 oz jars	22
16	5	Chef Anton's Gumbo Mix	New Orleans Cajun Delights	Condiments	36 boxes	21.35

Manage table styles and options

1. Use the Table Tools – Design tab at the top of the screen to change the style of the table
2. In the Design tab, Table Style Options, click on First Column, Last Column, or Banded Columns to change the formatting
3. To add a Total Row:
    a. Click on the Total Row box
    b. Scroll down to the bottom of the data
    c. Notice that the Total label is listed but not the numbers
    d. Click on cell F79 across from Total and under UnitPrice
    e. A drop-down arrow will appear; for this one, select Average, just as an example
    f. Excel will identify how to calculate the total, when possible, but when it is not sure, it will require you to make the selection

Filter and sort a table

1. Filter records by using the drop down arrows on the header row. For example, select Exotic Liquids from the Company Name column.



2. To sort the data, use the drop-down arrows to select the sorting option (A to Z, Z to A, smallest to largest, largest to smallest) to sort on Units in Stock.
3. To remove duplicate records, go to the Design Tab, Tools section, and click on Remove Duplicates

Session 3.11: Perform operations with formulas and functions

Use the Operations worksheet for these exercises

Summarize data by using functions

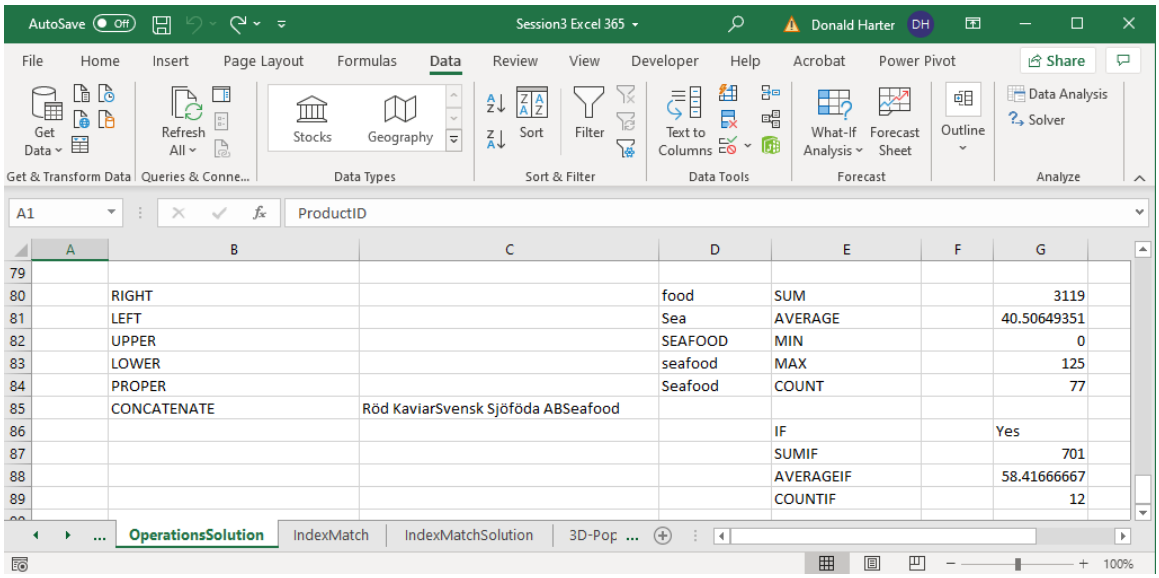
1. The functions SUM and AVERAGE were used earlier, but let's create a variety of calculations on the Operations data
2. In cell E80, type in the word SUM
3. In cell G80, enter the formula =sum(G2:G78) to calculate the sum of Units in Stock
4. In cell E81, type in the word AVERAGE
5. In cell G81, enter the formula =average(G2:G78)
6. In cell E82, type in the word MIN
7. In cell G82, enter the formula =min(G2:G78)
8. In cell E83, type in the word MAX
9. In cell G83, enter the formula =max(G2:G78)
10. In cell E84, type in the word COUNT
11. In cell G84, enter the formula =count(G2:G78)

Perform conditional operations by using functions

1. IF statement
    a. The IF statement can test whether a condition is true or false, and return an answer. The portions of an IF are the condition, the answer if the condition is true, and the answer if the condition is false. The structure is =IF(condition, result if true, result if false)
    b. In cell E86, type IF
    c. In cell G86, type =IF(G83>G82,"Yes", "No")
2. SUMIF
    a. SUMIF combines the SUM calculation with an IF condition
    b. In cell E87, type SUMIF
    c. In cell G87, type =SUMIF(D2:D78, "Seafood", G2:G78)
    d. SUMIF compares your criteria (Seafood) to the labels in D2:D78, and only Sums the G column where Seafood appears in the D column
3. AVERAGEIF – like SUMIF, but calculates the average
4. COUNTIF – like SUMIF (with the first two arguments), calculates the count
5. There are variations on each of these which allows multiple conditions (Expert level)
    a. SUMIFS – multiple criteria for a sum
    b. AVERAGEIFS – multiple criteria for an average
    c. COUNTIFS – multiple criteria for a count

Format and modify text

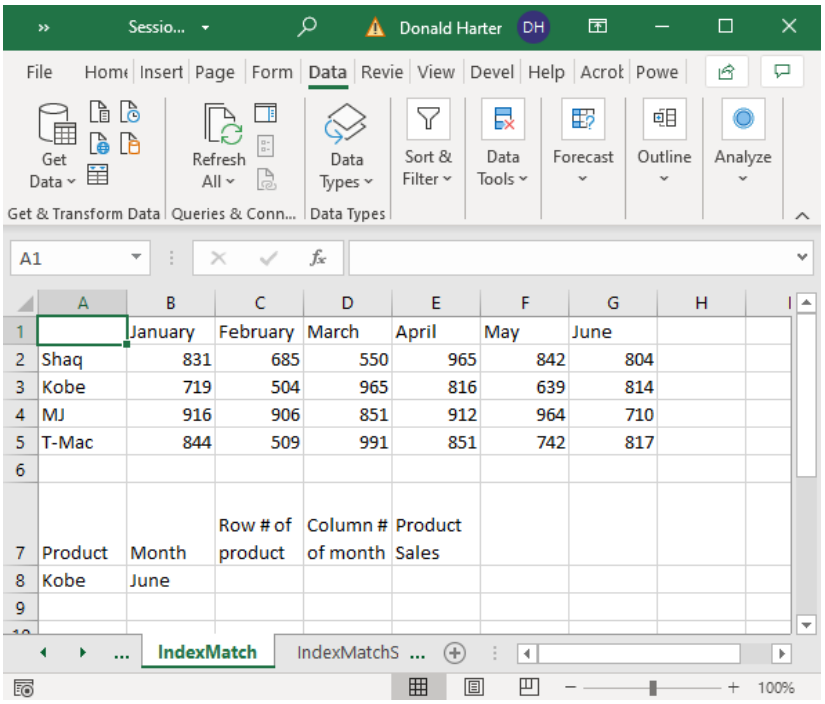
1. The RIGHT function strips off the right most characters of a text string
   b. In cell B80, type RIGHT
   c. In cell D80, type =right(D78,4) to get the 4 right most characters
2. The LEFT function strips off the left most characters of a text string
   a. In cell B81, type LEFT
   b. In cell D81, type =left(D78,3) to get the 3 left most characters
3. The UPPER function makes everything upper case
   a. In cell B82, type UPPER
   b. In cell D82, type =upper(D78)
4. The LOWER function makes everything lower case
   a. In cell B83, type LOWER
   b. In cell D83, type =lower(D78)
5. The PROPER function creates initial capitals
   a. In cell B84, type PROPER
   b. In cell D84, type =proper(D78)
6. The CONCATENATE function combines a group of text or number fields
   a. In cell B85, type CONCATENATE
   b. In cell C85, type =CONCATENATE(B78,C78,D78)
   c. Alternatively, you could type =B78&C78&D78
7. The TRIM function eliminates leading and trailing blanks
   a. In cell B86, type TRIM
   b. In cell D86, type =trim(B78)



	A	B	C	D	E	F	G
79							
80		RIGHT		food	SUM		3119
81		LEFT		Sea	AVERAGE		40.50649351
82		UPPER		SEAFOOD	MIN		0
83		LOWER		seafood	MAX		125
84		PROPER		Seafood	COUNT		77
85		CONCATENATE	Röd KaviarSvensk Sjöföda ABSeafood				
86					IF		Yes
87					SUMIF		701
88					AVERAGEIF		58.41666667
89					COUNTIF		12

Session 3.12: INDEX

Purpose: The index function, given and row and column, returns the value in a cell. This command is not very useful by itself but can be useful when combined with other commands. The following is the IndexMatch tab of your spreadsheet.



	A	B	C	D	E	F	G
1		January	February	March	April	May	June
2	Shaq	831	685	550	965	842	804
3	Kobe	719	504	965	816	639	814
4	MJ	916	906	851	912	964	710
5	T-Mac	844	509	991	851	742	817
6							
7	Product	Month	Row # of product	Column # of month	Product Sales		
8	Kobe	June					
9							

The command for index is:

=index(array, row number, column number)

Where: array is the data only
Row number is the number within the array, not the spreadsheet
Column number is the number within the array, not the spreadsheet

To identify cell value of a specific column and row, use the following steps:

1. Identify the data range and name it Sales. In this case, the highlight B2 through G5, then enter a name of Sales.
2. In cell E8, under product sales, we will identify the row and column of the data (not the spreadsheet) for Kobe and June. Kobe is the second row; June is the sixth column. In E8, enter =index(sales, 2, 6)

MATCH

Purpose:

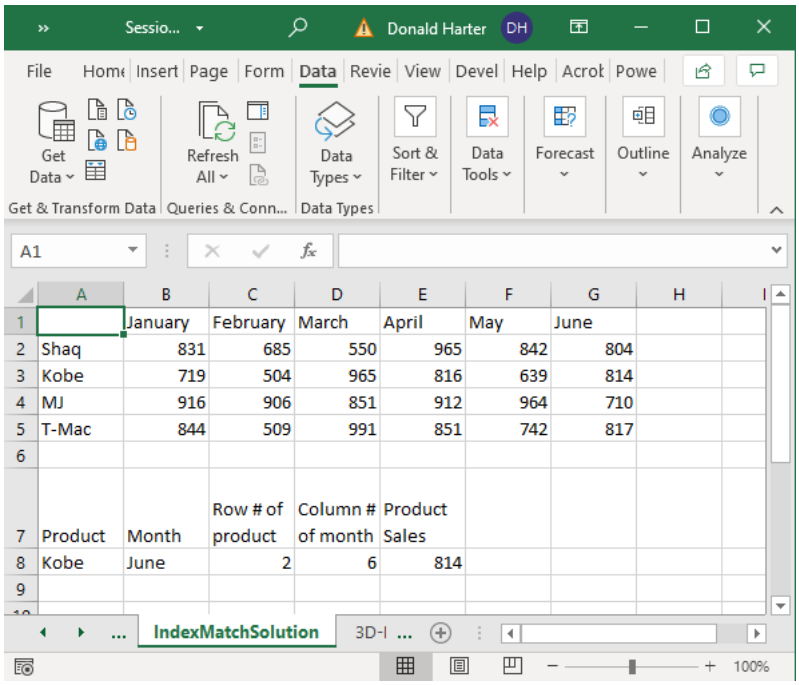
The Match function searches an array and finds a matching value. The match function takes the form:

=match(lookup value, lookup range, match type)

Where: lookup value is the value you are trying to match
Lookup range is the range to be searched
Match type can take on the values:
1: data in ascending order, returns the row location of the largest value in the range less than or equal to the lookup value
-1: descending order, returns the row location of the last value in the range that is greater or equal to the lookup value
0: returns the row location of the first exact match

In the pictured example above, let's find the row and column for Kobe's June sales automatically.

1. To make this a general procedure, note that Kobe in entered in A8 and June is entered in B8.
2. Let's assign a name of SportsStar to A2:A5. Also assign a name of SalesMonth to B1:G1
3. Now enter the formula for the row in C8. Since we are searching for the row number of Kobe, the formula would be =match(A8, SportsStar, 0). Why are we setting the last value to zero?
4. Enter the formula for the column in D8. The formula would be =match(B8, SalesMonth,0)
5. Now let's fix our index. The index created before had the row and column hard-coded. Set the row and column in the index cell E8 using the calculated row and column.
6. Finally, enter another star's name in A8. Enter a different month in B8.

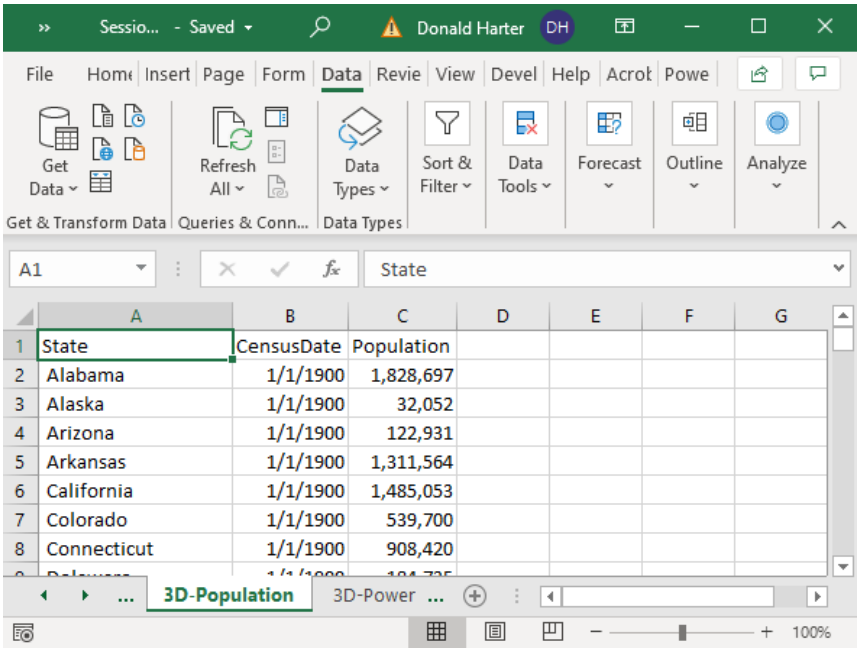


	A	B	C	D	E	F	G
1		January	February	March	April	May	June
2	Shaq	831	685	550	965	842	804
3	Kobe	719	504	965	816	639	814
4	MJ	916	906	851	912	964	710
5	T-Mac	844	509	991	851	742	817
6							
7	Product	Month	Row # of product	Column # of month	Product Sales		
8	Kobe	June	2	6	814		
9							

Session 3.13: 3D Maps

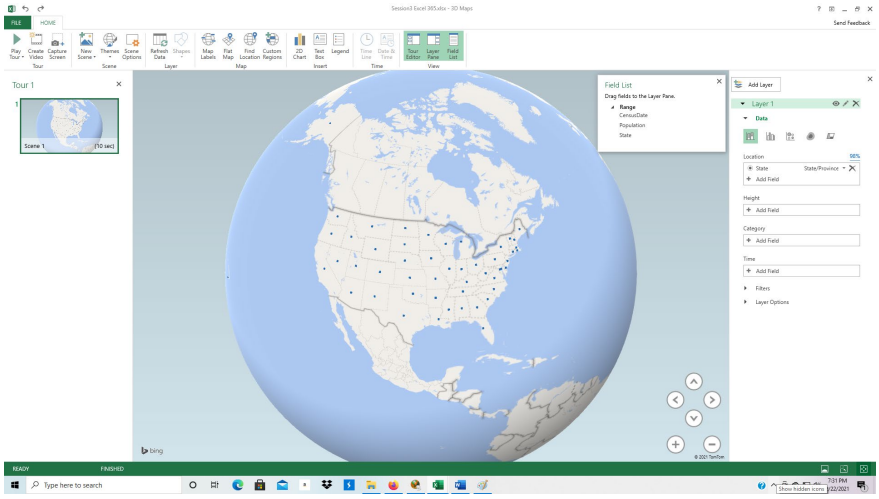
For this exercise, use the 3D Maps data on power plants in the United States and the population of each state over time. 3D Maps allows you to plot data on geographic maps and create videos displaying the data.

1. Click on the 3D Population tab at the bottom of the spreadsheet. This data is official census data of the population of each state in the U.S. from 1900 to 2010.

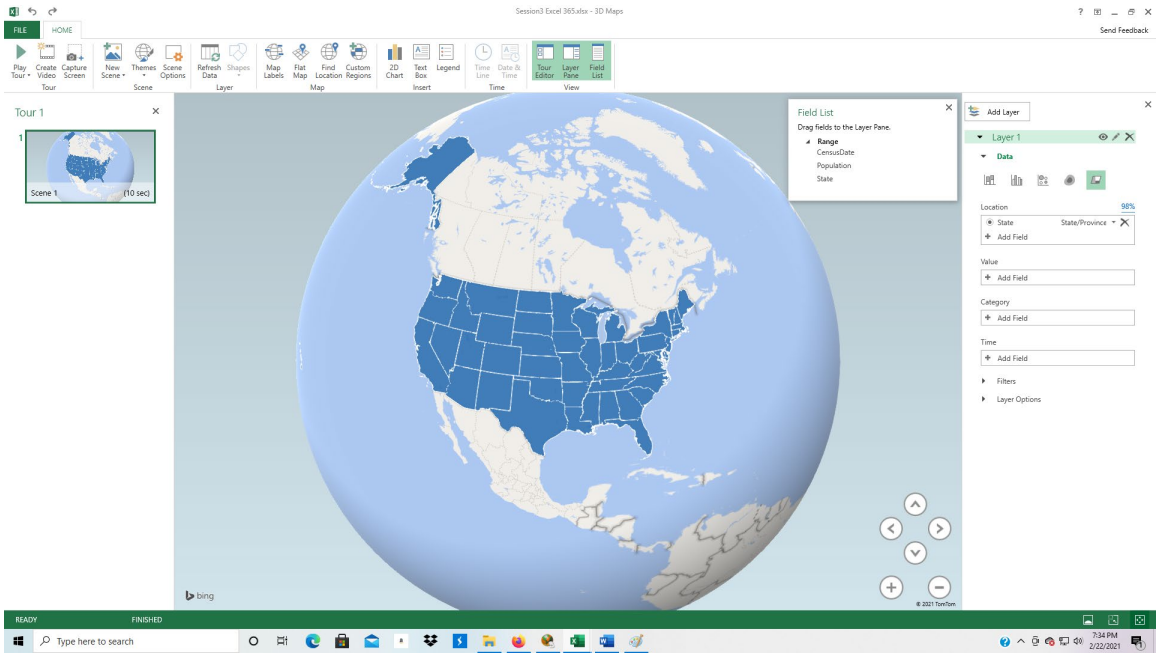


	A	B	C
1	State	CensusDate	Population
2	Alabama	1/1/1900	1,828,697
3	Alaska	1/1/1900	32,052
4	Arizona	1/1/1900	122,931
5	Arkansas	1/1/1900	1,311,564
6	California	1/1/1900	1,485,053
7	Colorado	1/1/1900	539,700
8	Connecticut	1/1/1900	908,420

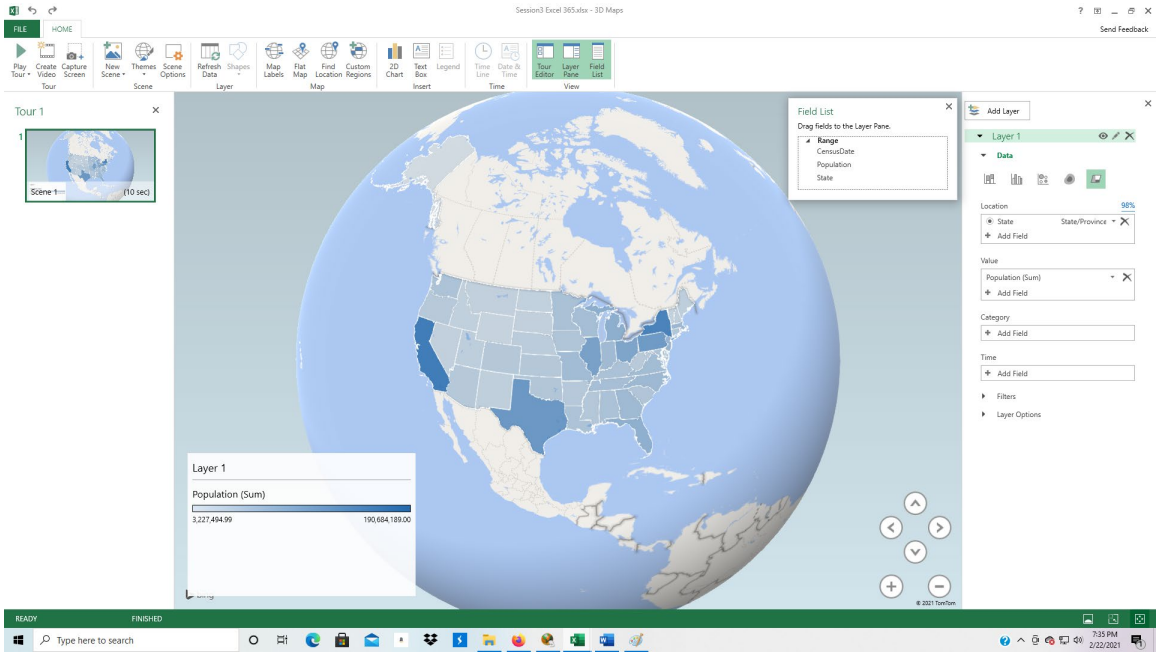
2. Click on Insert at the top of Excel, then 3D-Maps, then Open 3D Maps
3. Excel will populate the map. The map below is not very interesting.



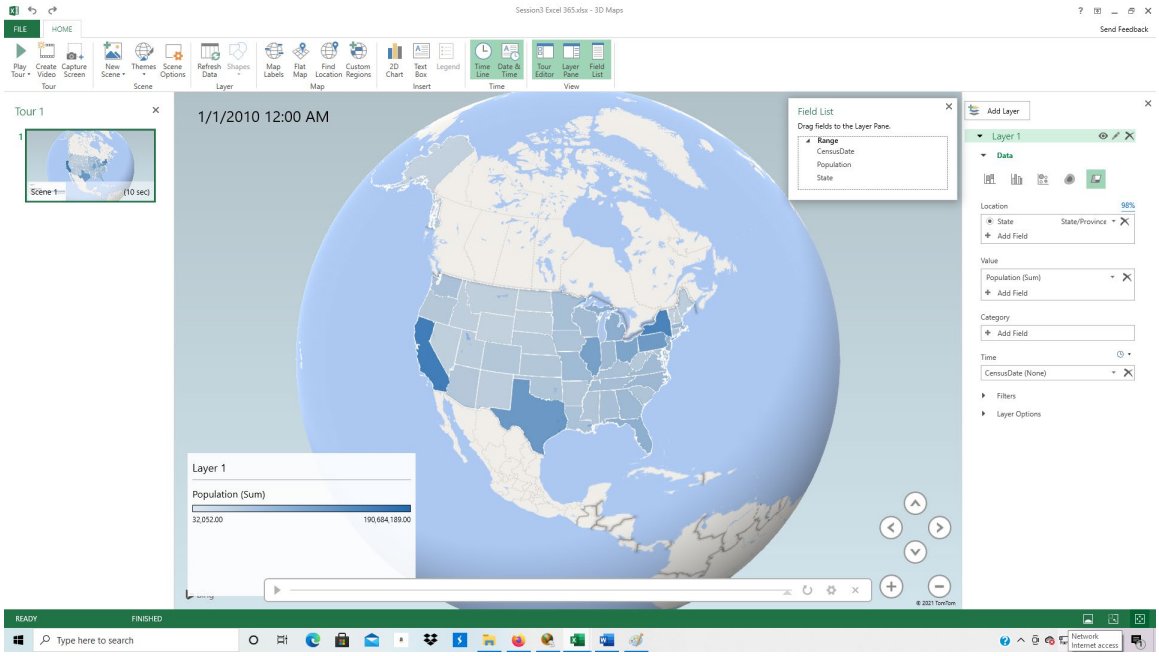
4. Under Layer 1, Data, click on Change Visualization to Region



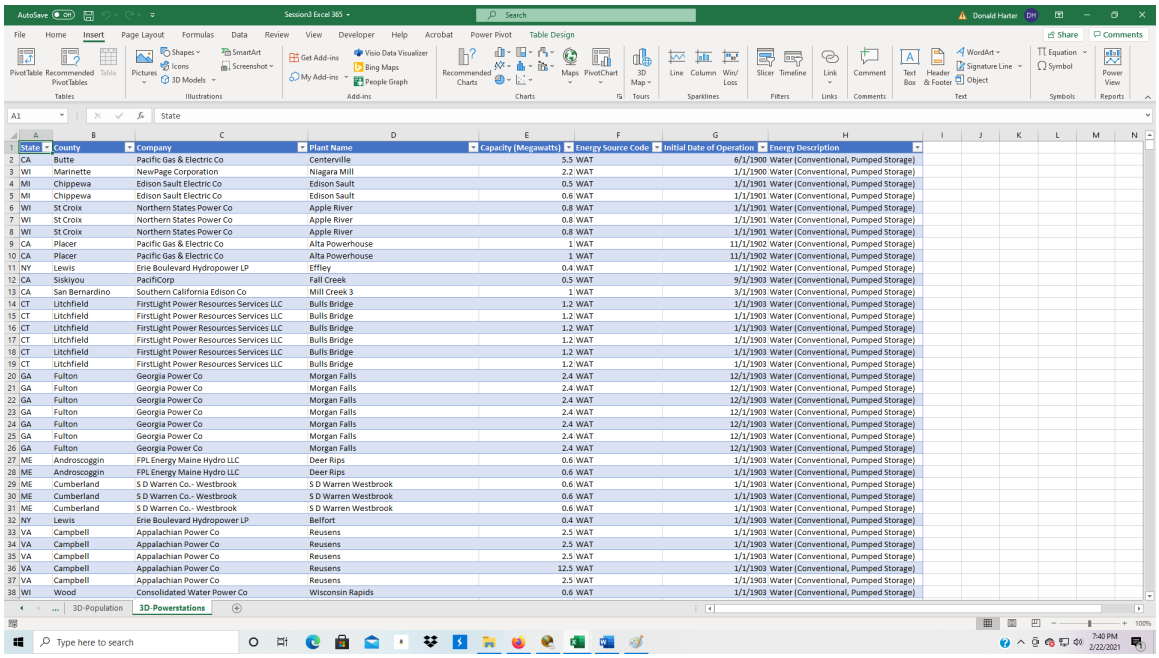
5. Still not very interesting; all the states are the same color.
6. In the Field list, drag the Population to Value in Layer 1



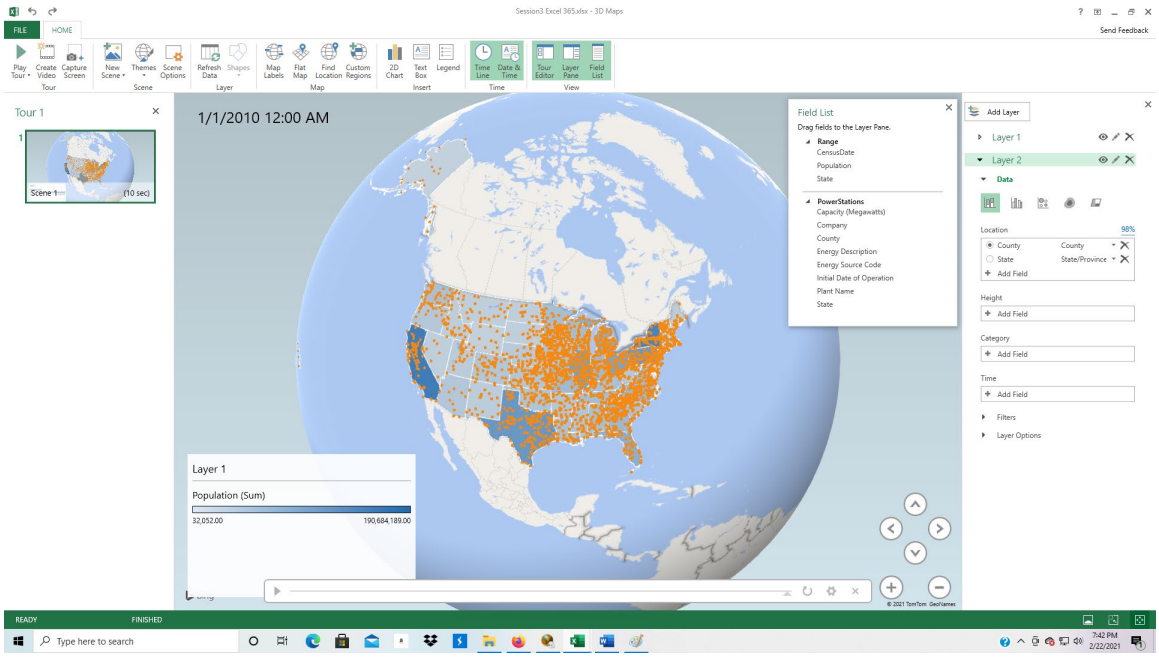
7. In the Field list, drag CensusDate to Time in Layer 1
8. At the bottom of the screen is now an arrow to play the video from 1900 to 2010
9. Click the arrow. When it is playing, you can stop the video by clicking on the timeline or move to any specific time.



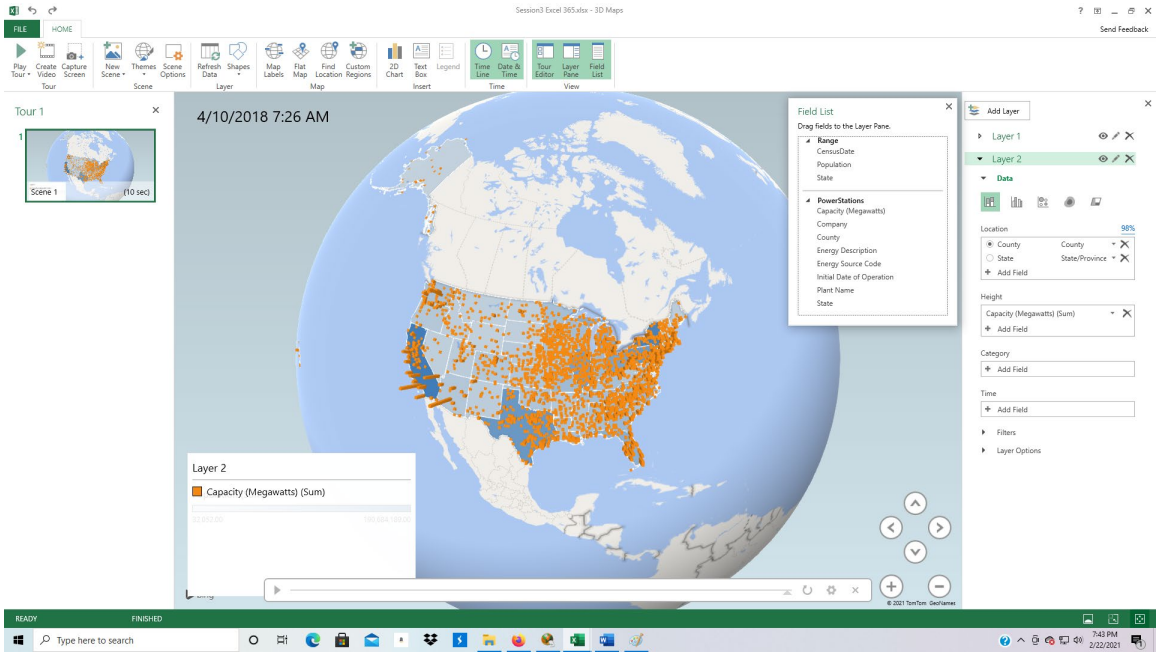
10. To add a second Layer, click back on the data in the spreadsheet.
11. Click on the 3D-PowerStations tab



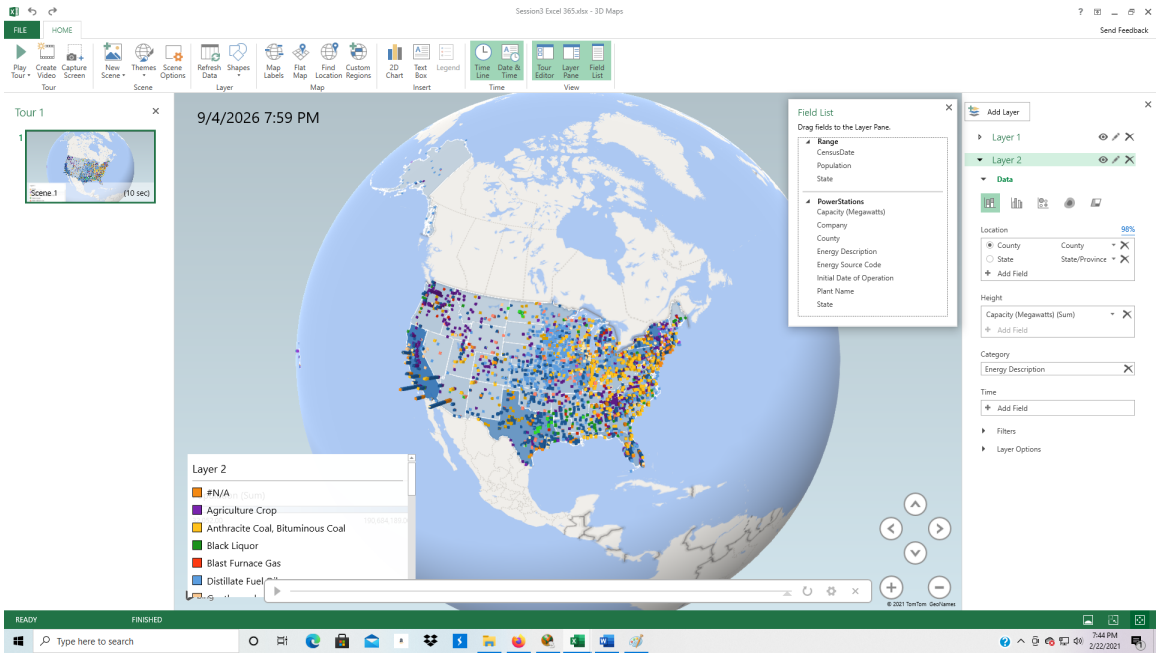
12. Click on the Insert tab, 3D Map, Add Selected Data to 3D Maps



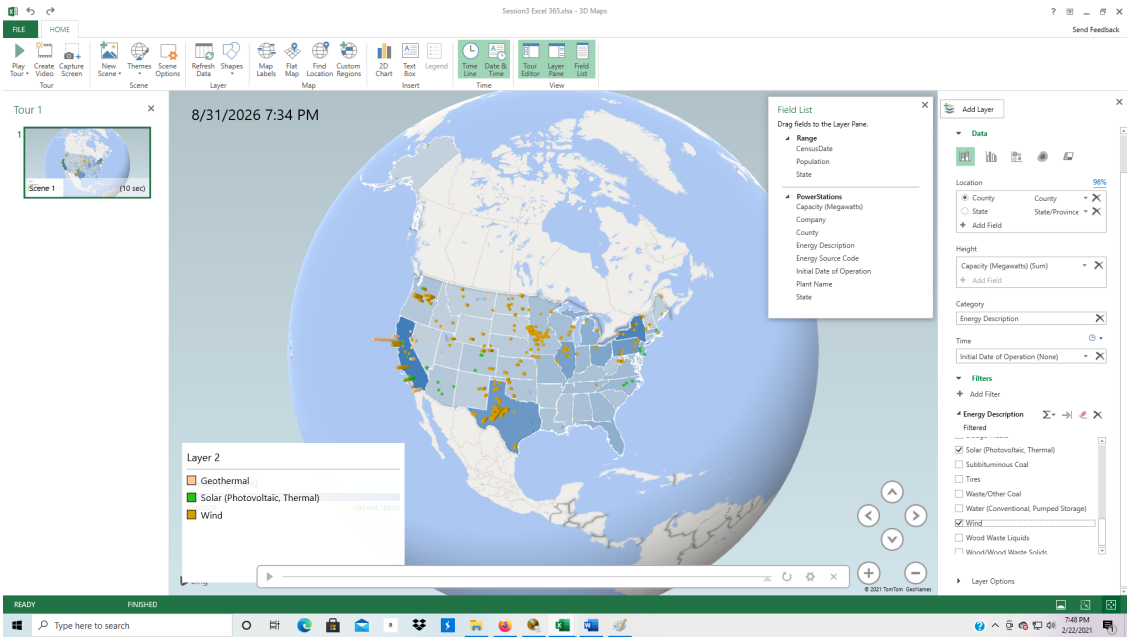
13. Next we are going to add a measurement scale to each data point. From the Field List, PowerStations, drag Capacity (Megawatts) to Height (Add Field)



14. Next, we can color code categories of Power Stations by dragging Energy Description from Field List, PowerStations to Category (Add Field) on the right side of the screen

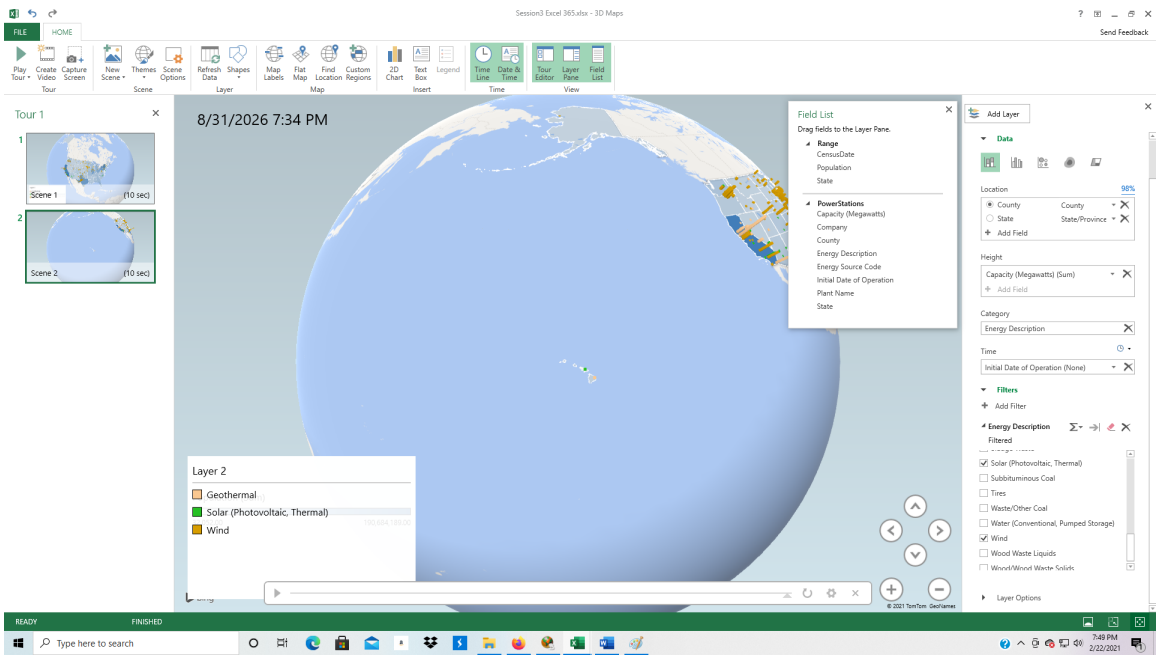


15. To add a timeline to the data, drag Initial Date of Operation from Field List, PowerStations to Time (Add Field) on the right side of the screen.
16. The video play bar appears at the bottom of the screen. You can click on the arrow at the left of the video play bar to play the animation over time.
17. The next feature allows us to add a filter to the display. Click on the arrow next to Filters on the right. Click on the Add Filter under Filters. Click on Energy Description to turn on that filter. You can now select Geothermal, Solar and Wind for renewable energy.

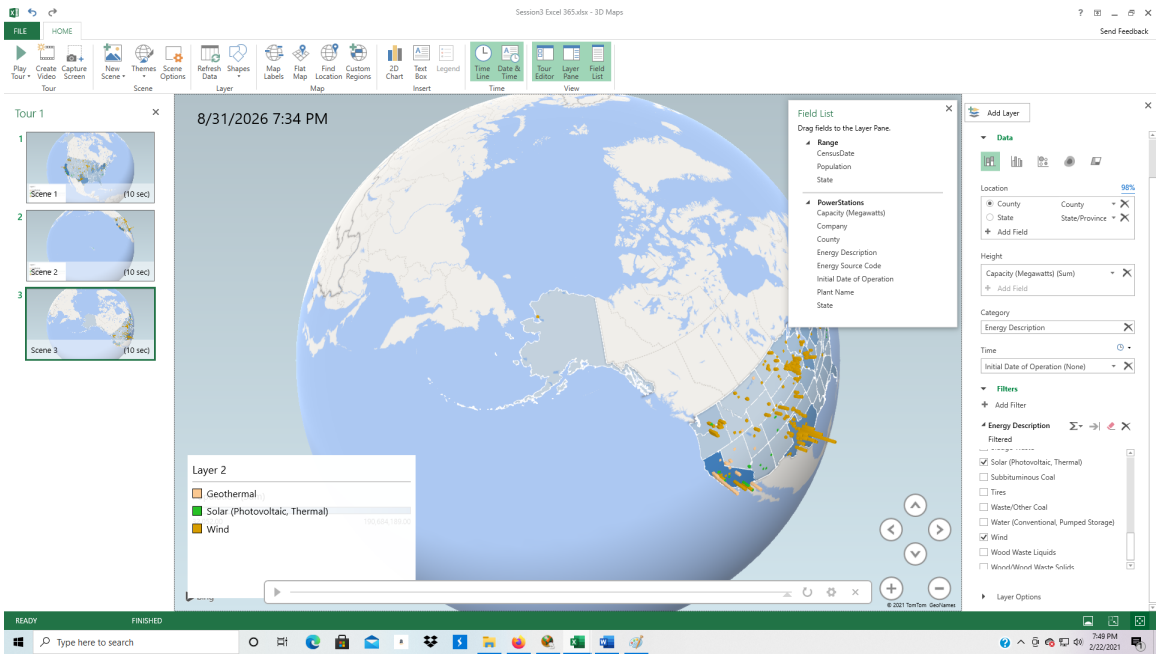


Creating a Video

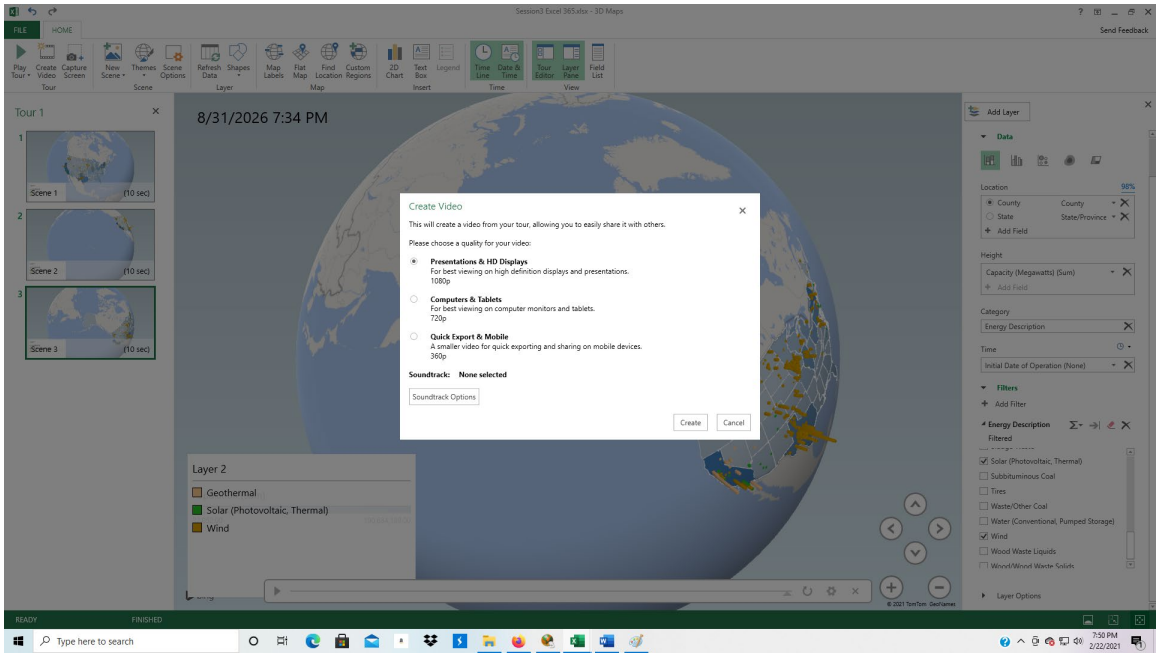
1. To create a video, we first need to create new scenes. Click on New Scene, then rotate the globe to show Hawaii. You can create as many scenes as you want



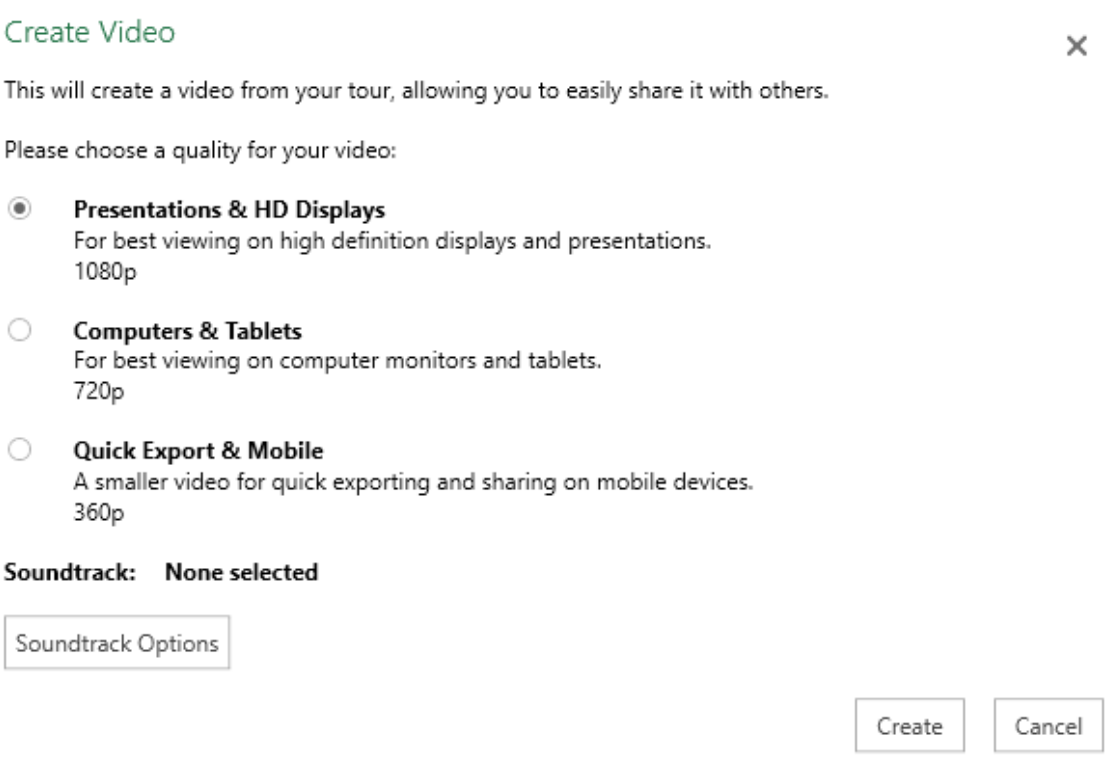
2. Click New Scene, then rotate the globe to show Alaska



3. To play the video, click Play Tour
4. If you want to save the video to your computer, click on Create Video



5. Select Presentations & HD Displays if you have a high definition screen. Otherwise select a low resolution option.
6. Click Create, and save the video to your computer



Earth from a different perspective

Most maps are displayed flat rather than round like a globe. If you believe that the earth is flat, click on the Flat Map button at the top of the screen.