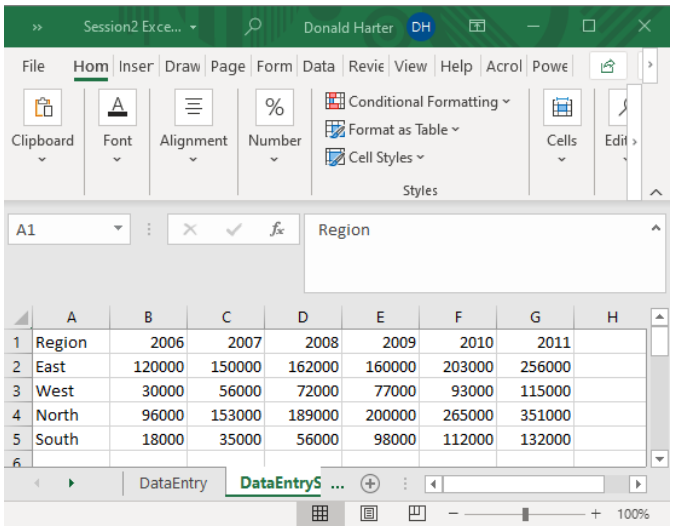
Microsoft Excel: Basics

Microsoft Excel

1. Go to the Course Website on BlackBoard
2. In Lectures, right click and save to your desktop “Session2 Excel 365.xlsx”

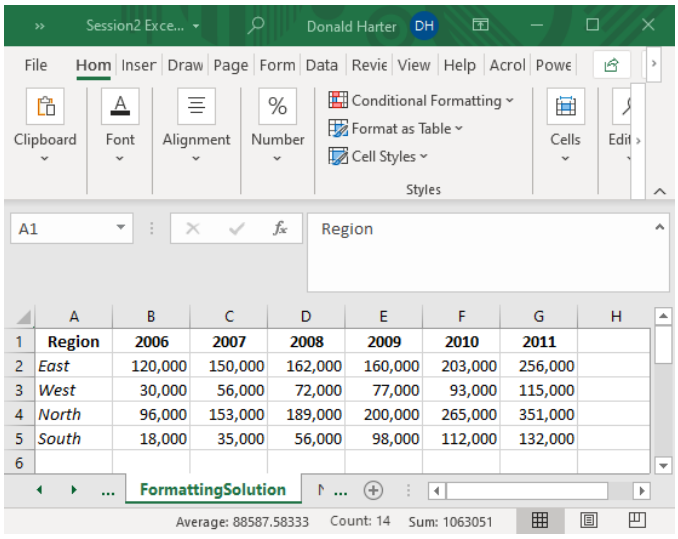
2.1 Data Entry

1. The spreadsheet is formatted as columns (A, B, C…) and rows (1, 2, 3 …). The column labels are across the top; the row labels are down the left side.
2. To type in data, click on the cell A1 and begin entering data. Use the tab key to move to the next cell. Use your mouse to click on any cell. Enter the data listed below.



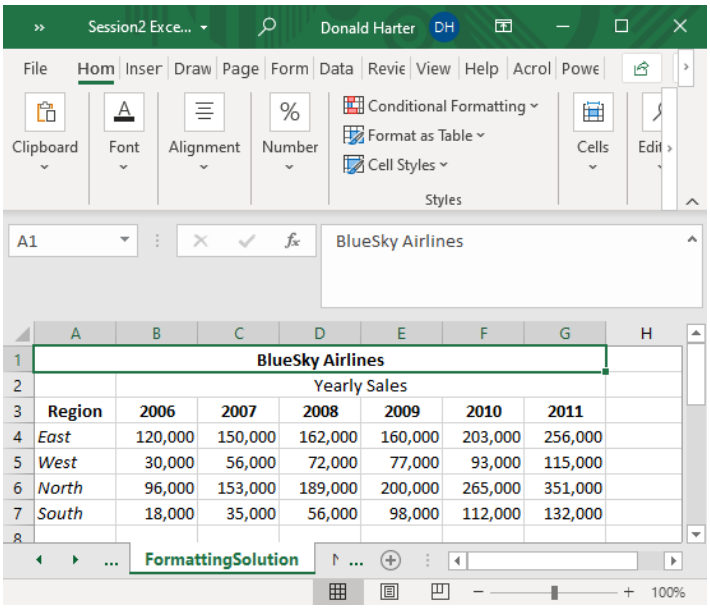
	A	B	C	D	E	F	G	H
1	Region	2006	2007	2008	2009	2010	2011	
2	East	120000	150000	162000	160000	203000	256000	
3	West	30000	56000	72000	77000	93000	115000	
4	North	96000	153000	189000	200000	265000	351000	
5	South	18000	35000	56000	98000	112000	132000	
6								

3. To change the format of a cell, click on the cell and use the buttons in the Font section.
    a. Change the first line to Bold and center
    b. Change the regions to Italics
    c. Use the comma (,) button to add commas to the sales, then reduce the number of zeros with the .00 -> .0 button

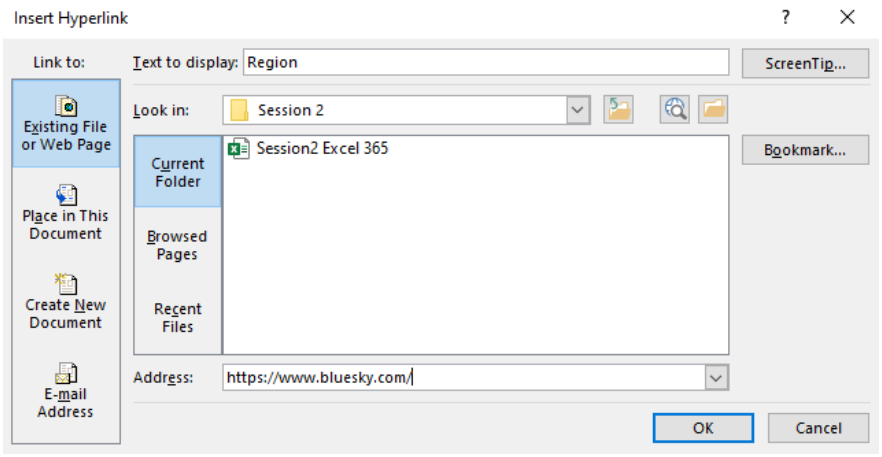


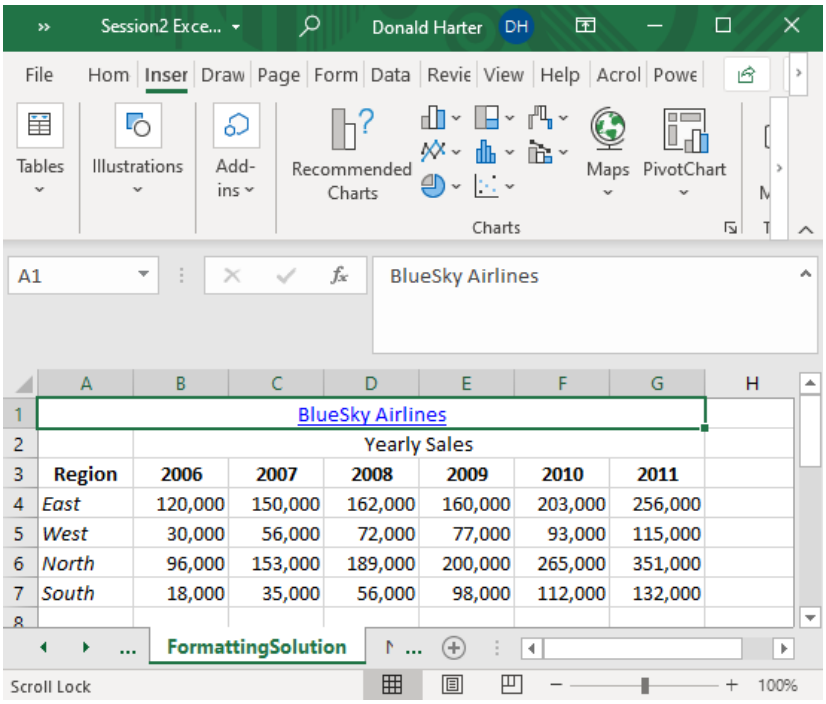
	A	B	C	D	E	F	G	H
1	**Region**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	
2	*East*	120,000	150,000	162,000	160,000	203,000	256,000	
3	*West*	30,000	56,000	72,000	77,000	93,000	115,000	
4	*North*	96,000	153,000	189,000	200,000	265,000	351,000	
5	*South*	18,000	35,000	56,000	98,000	112,000	132,000	
6								

4. To enter a new row before the first row, click on the 1 in the left column, which will highlight the first row
5. Use the drop-down arrow next to Insert and insert sheet rows. Insert two rows before the first row.
6. In the new cell A1, type in Blue Sky Airlines and change it to bold
7. In cell B2, enter Yearly Sales
8. To spread a label across cells, highlight the range of cells and click Merge and Center button in the alignment section
    a. Merge and center Blue Sky Airlines across cells A1 to G1, then change to Bold
    b. Merge and center Yearly Sales across cells B2 to G2

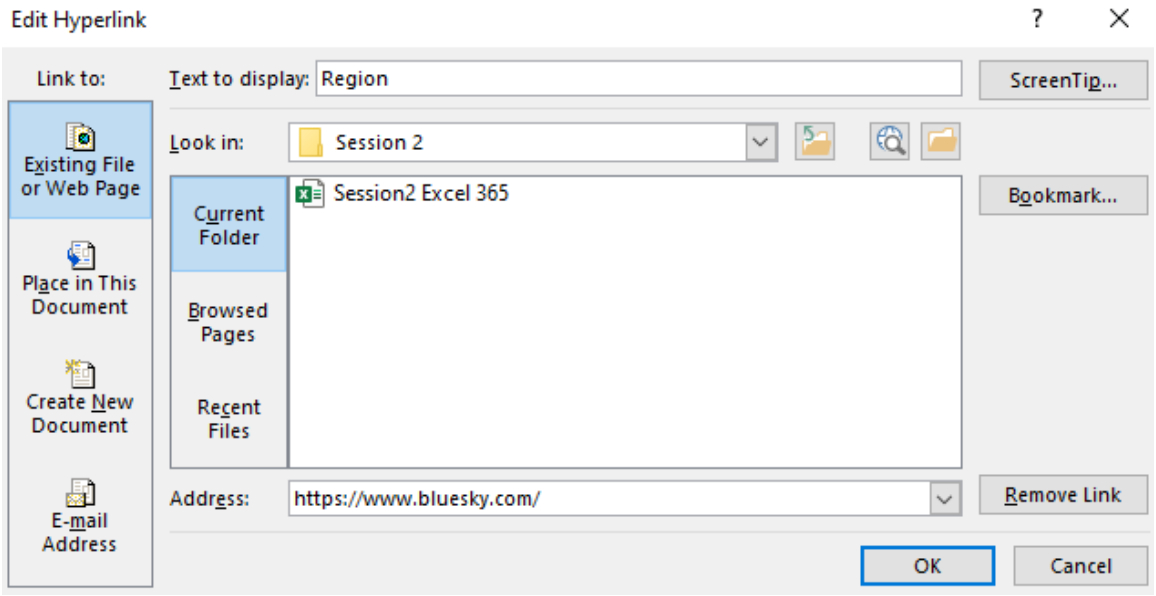


9. To add a hyperlink for BlueSky Airlines:
    a. Click on cell A1 (BlueSky Airlines)
    b. Click on the Insert tab at the top of the screen
    c. Click on Link, Insert Link
    d. In the Address box, type in https://www.bluesky.com/, and click OK





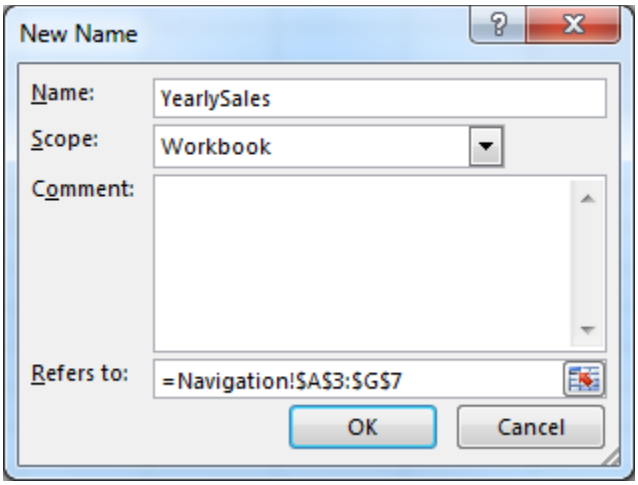
10. To remove a hyperlink
    e. Click on cell A1 (BlueSky Airlines) with the hyperlink
    f. Click on the Insert tab at the top of the screen
    g. Click on Link, Insert Link
    h. Next to the Address box, click Remove Link, then click OK



2.2 Navigate in worksheets and workbooks

Use the Navigation worksheet for this exercise

1. To name a cell range, such as A3:G7, as YearlySales
    a. Highlight cells A3 through G7
    b. Click on Formulas tab at the top of the screen
    c. Under Define Name, click on Define Name
    d. Type in YearlySales as the range name, then OK

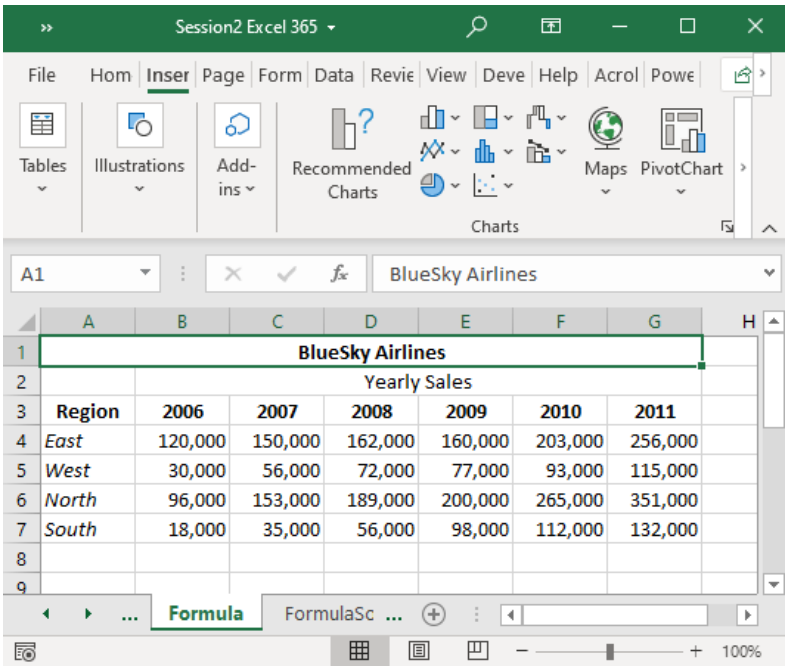


2. Alternatively, you can highlight the range, to the upper left corner of the screen, and type in YearlySales.

2.3 Calculations and Formulas

Use the Formula worksheet for this exercise.

1. Highlight the four sales values for 2006. Notice at the bottom of the page, it calculates the average, count and sum of the values.

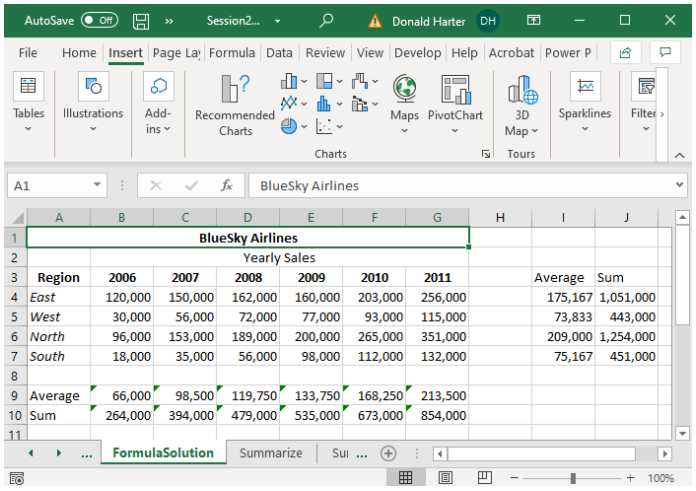


	A	B	C	D	E	F	G
1	BlueSky Airlines						
2		Yearly Sales					
3	**Region**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
4	*East*	120,000	150,000	162,000	160,000	203,000	256,000
5	*West*	30,000	56,000	72,000	77,000	93,000	115,000
6	*North*	96,000	153,000	189,000	200,000	265,000	351,000
7	*South*	18,000	35,000	56,000	98,000	112,000	132,000
8							
9							

2. Next, we'll enter a formula in the spreadsheet to automatically calculate averages and sums
3. In cell A9, type the word Average; also enter Average in I3
4. In cell A10, type the word Sum; also enter Sum in J3
5. To calculate a formula, use the equal sign, the name of the formula, and the data range
6. In cell B9, enter =average(b4:b7)
7. In cell B10, enter =sum(b4:b7)
8. In cell I4, enter =average(b4:g4)
9. In cell J4, enter =sum(b4:g4)
10. Notice that after you type the equals sign and part of the formula name, it gives you several options that start with the same spelling; you can click on the one you want, then highlight the range of data

2.4 Copy and Paste

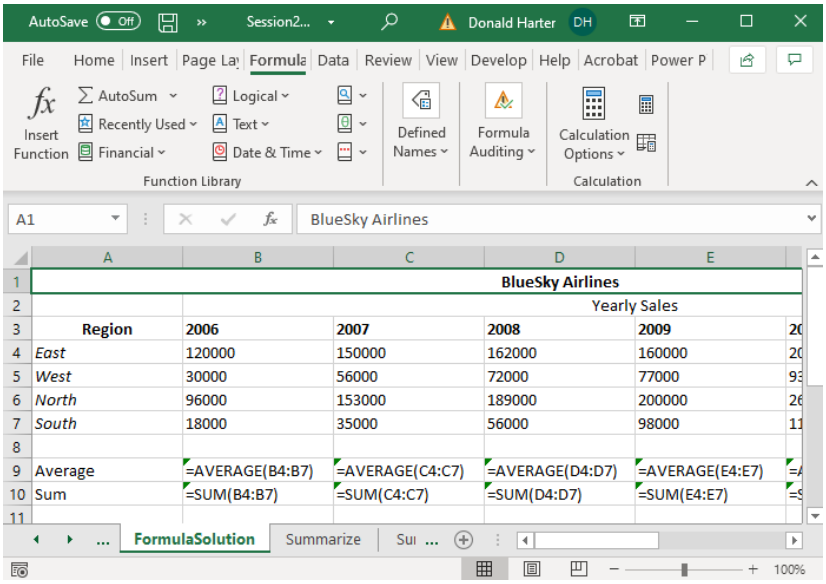
1. Copy and paste works the same in Excel as in Microsoft Word, but you can also copy and paste formulas and Excel will automatically update formulas
2. Copy the average and sum formulas across all columns and rows
    a. Click on cell B9, click on the copy button in the Clipboard section (or use control-C), highlight cells C9 to G9, and click the paste button in Clipboard
    b. Copy cell B10 across for the columns in the same way
    c. You can copy several formulas at once. Highlight cells I4 and J4. Click on copy, then highlight the range I5 to J7, and click paste.
    d. Look at the formulas in B9 and C9. How are they different?



	A	B	C	D	E	F	G	H	I	J
1	BlueSky Airlines									
2		Yearly Sales								
3	Region	2006	2007	2008	2009	2010	2011		Average	Sum
4	*East*	120,000	150,000	162,000	160,000	203,000	256,000		175,167	1,051,000
5	*West*	30,000	56,000	72,000	77,000	93,000	115,000		73,833	443,000
6	*North*	96,000	153,000	189,000	200,000	265,000	351,000		209,000	1,254,000
7	*South*	18,000	35,000	56,000	98,000	112,000	132,000		75,167	451,000
8										
9	Average	66,000	98,500	119,750	133,750	168,250	213,500			
10	Sum	264,000	394,000	479,000	535,000	673,000	854,000			
11										

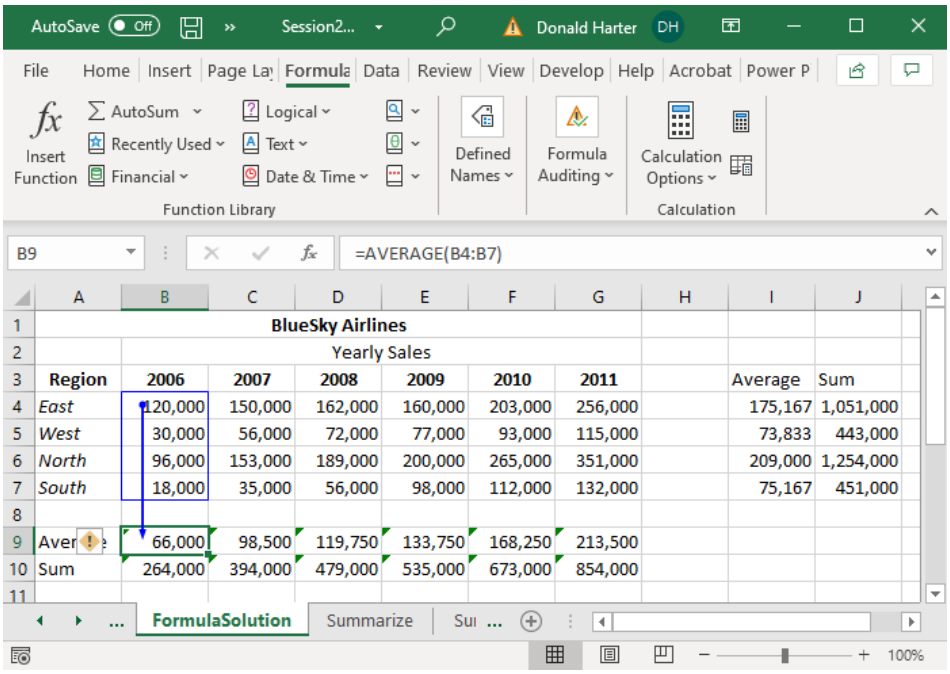
2.5 Checking Formulas

1. You can view formulas by clicking on the tab Formulas, the Show Formulas

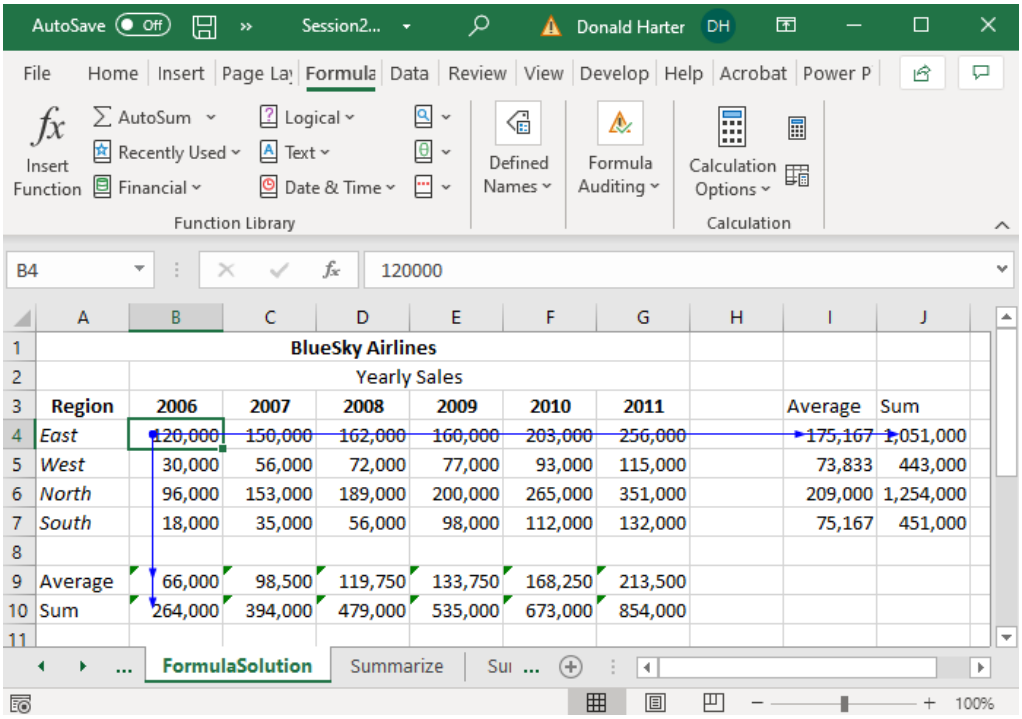


	A	B	C	D	E
1				BlueSky Airlines	
2				Yearly Sales	
3	Region	2006	2007	2008	2009
4	*East*	120000	150000	162000	160000
5	*West*	30000	56000	72000	77000
6	*North*	96000	153000	189000	200000
7	*South*	18000	35000	56000	98000
8					
9	Average	=AVERAGE(B4:B7)	=AVERAGE(C4:C7)	=AVERAGE(D4:D7)	=AVERAGE(E4:E7)
10	Sum	=SUM(B4:B7)	=SUM(C4:C7)	=SUM(D4:D7)	=SUM(E4:E7)
11					

2. Excel has an auditing feature to check formulas; click on the tab labeled Formulas
3. Click on B9, then click on Trace Precedents in Formula Auditing
4. The box represents the data which goes into the formula; the arrow points to the formula
5. To clear the arrows, click on Remove Arrows



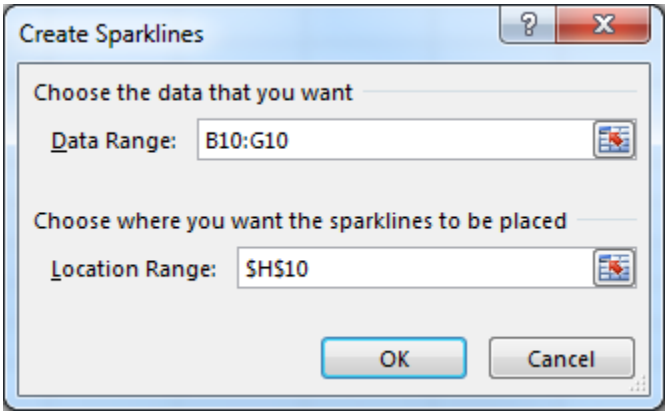
6. Click on B4, then click on Trace Dependents in Formula
7. The box represents the data which goes into the formula; the arrow points to the formulas
8. To clear the arrows, click on Remove Arrows



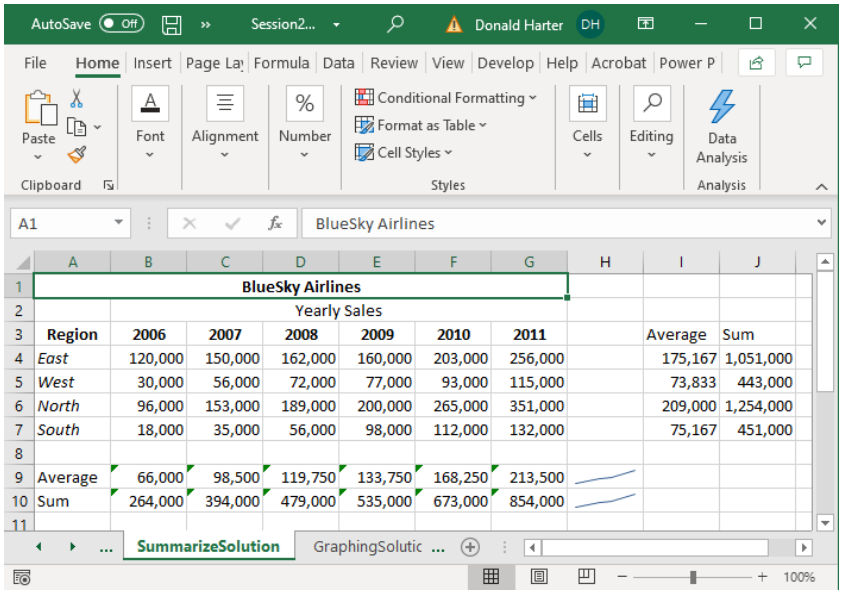
2.6 Summarize Data

Use the Summarize worksheet for this exercise.

1. A sparkline is miniature graph next to trend data
    a. To create a sparkline of the sum of yearly sales, highlight cells B10:G10
    b. On the Insert tab, click Sparklines, Line
    c. Enter H10 in Location Range, then OK

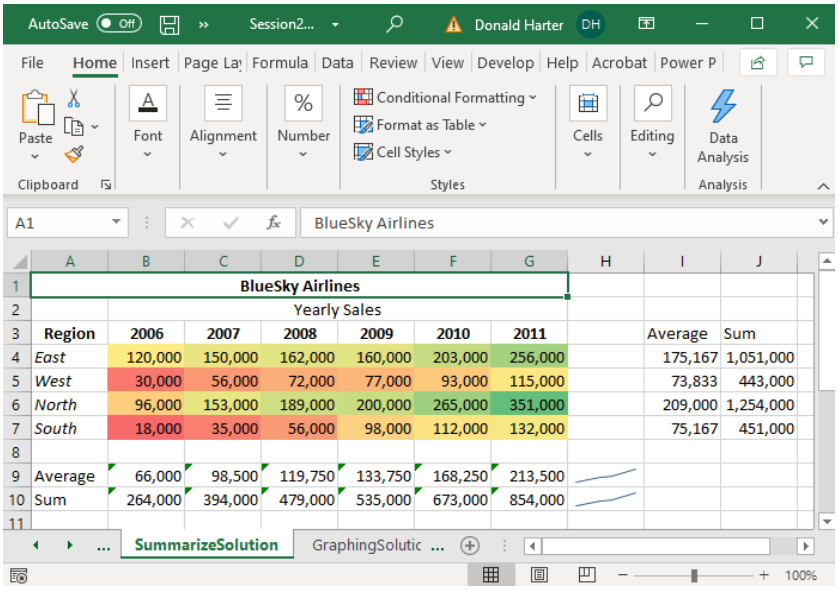


2. Similarly create sparklines for the average sales by year in cell H9



	A	B	C	D	E	F	G	H	I	J
1	BlueSky Airlines									
2		Yearly Sales								
3	**Region**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**		Average	Sum
4	*East*	120,000	150,000	162,000	160,000	203,000	256,000		175,167	1,051,000
5	*West*	30,000	56,000	72,000	77,000	93,000	115,000		73,833	443,000
6	*North*	96,000	153,000	189,000	200,000	265,000	351,000		209,000	1,254,000
7	*South*	18,000	35,000	56,000	98,000	112,000	132,000		75,167	451,000
8										
9	Average	66,000	98,500	119,750	133,750	168,250	213,500			
10	Sum	264,000	394,000	479,000	535,000	673,000	854,000			
11										

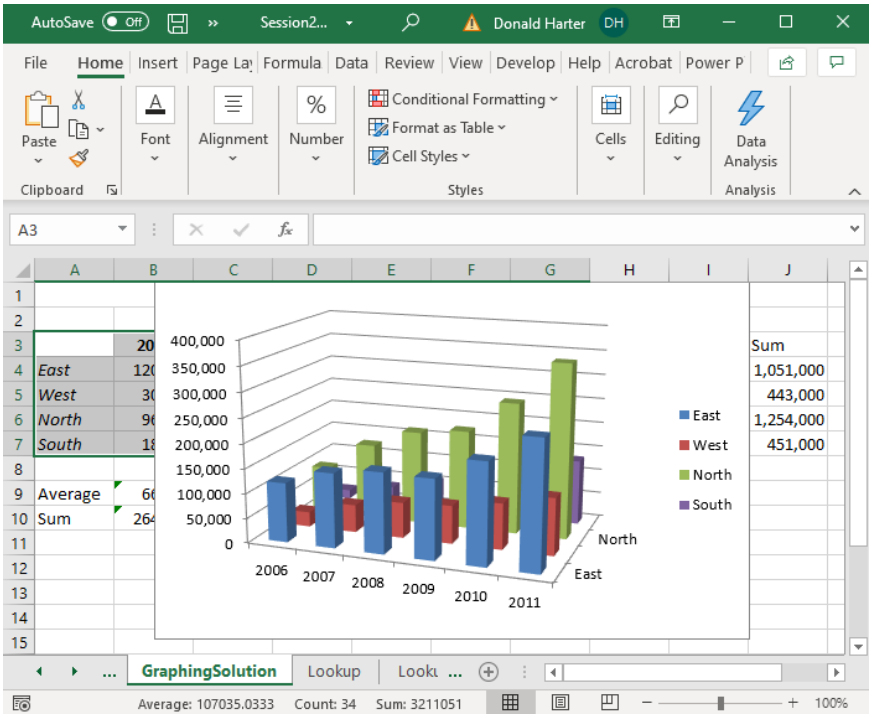
3. Conditional Formatting allows data to be highlighted by colors or icons
    a. Again, using the Summarize data for BlueSky Airlines, highlight sales in cells B4:G7
    b. Click on the Home tab
    c. Click on Conditional Formatting, Color Scales
    d. Alternatively, try Data Bars and Icon Sets



	A	B	C	D	E	F	G	H	I	J
1	BlueSky Airlines									
2		Yearly Sales								
3	**Region**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**		Average	Sum
4	*East*	120,000	150,000	162,000	160,000	203,000	256,000		175,167	1,051,000
5	*West*	30,000	56,000	72,000	77,000	93,000	115,000		73,833	443,000
6	*North*	96,000	153,000	189,000	200,000	265,000	351,000		209,000	1,254,000
7	*South*	18,000	35,000	56,000	98,000	112,000	132,000		75,167	451,000
8										
9	Average	66,000	98,500	119,750	133,750	168,250	213,500			
10	Sum	264,000	394,000	479,000	535,000	673,000	854,000			
11										

2.7 Graphing Data

1. Use the SummarizeSolution spreadsheet tab for this exercise.
2. Highlight the data by clicking on the cell labeled A3 through G7.
3. Next click on the Insert tab at the top of the screen.
4. In the section labeled Charts, click on the tiny icon in the lower right corner of charts to bring up the possible chart options.
5. Move the cursor over each option and select the 3-D column chart, click OK



6. Let's move the chart to another sheet. Click on the upper right corner of Excel, select Move Chart Location, click on New Sheet, then OK.
7. At the top of the screen click on Type, Change Chart Type to change the type of chart
8. In the Chart Tools: Design: Data, click on Switch Row/Column. What happens?
9. In the Chart Layouts: Quick Layout, there are several options (scroll down). How are they different?
10. Next, try Chart Tools: Design: Add Chart Element. Experiment with Labels and Axes. Under 3-D Rotation, change the X and Y rotation.

2.8 Saving your Spreadsheet

To save a spreadsheet, click on the Microsoft Office icon in the upper left corner of the screen, move the cursor over Save As. You will notice that you can save your Spreadsheet in a variety of formats.

Microsoft Excel: Lookup, Sorts, Filters, Pivot Tables

2.8 VLOOKUP

Refer to the Lookup worksheet for this exercise

Purpose: A lookup function scans on column of data to find a match, then returns a value from a second column. This technique is call vertical lookup (VLOOKUP), but there is also a horizontal lookup (HLOOKUP). The format is:

VLOOKUP(lookup value, table range, column index, range lookup)

Where: Lookup value in the value you search on
Table range is the group of rows and columns searched
Column index in the column number with the table range for the return value
Range lookup is a flag indicating whether non-exact matches are allowed; if FALSE, then only exact matches will return a value; if TRUE or omitted, then it assumes the data is sorted and uses the first row where the value is the table is less than or equal to the lookup value

Let's first identify tax rates by income level. The steps are:

1. We can create a new named range called TaxTable by highlighting the data under income, US Tax and NY Tax, then naming it TaxTable.
2. In A7, enter an income.
3. In B7, enter =vlookup(A7,TaxTable,2) to find the US Tax.
4. In C7, enter =vlookup(A7,TaxTable,3) to find the NY Tax.

Next, use VLOOKUP where the data is not sorted

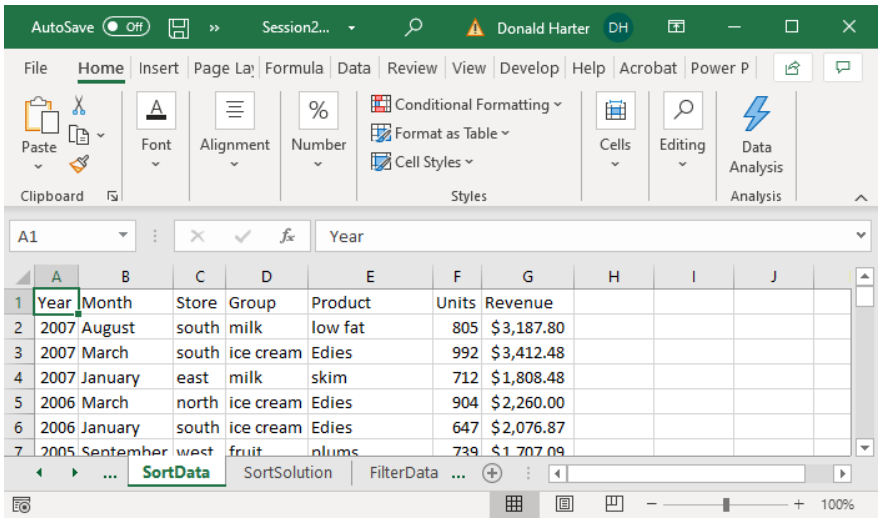
1. Create a range called ProductPrice by highlighting the data from A11 through B15
2. In A17, enter a product from the list (or not, to test the VLOOKUP)
3. In B17, enter =vlookup(A17,ProductPrice,2,FALSE)

Finally, build a VLOOKUP for sales by state.

1. In H1, type State
2. In I1, type Sales
3. In H2, type in a 2-letter code for a state
4. In I2, enter =vlookup(H2,E2:F51,2)

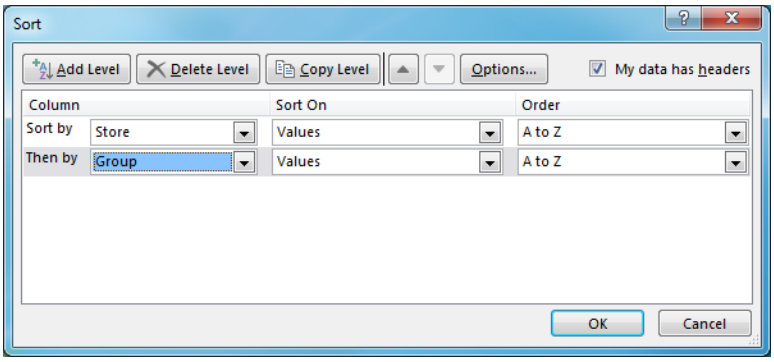
2.9 Sorting Data

1. Use the SortData spreadsheet for this exercise.



	A	B	C	D	E	F	G
1	Year	Month	Store	Group	Product	Units	Revenue
2	2007	August	south	milk	low fat	805	$3,187.80
3	2007	March	south	ice cream	Edies	992	$3,412.48
4	2007	January	east	milk	skim	712	$1,808.48
5	2006	March	north	ice cream	Edies	904	$2,260.00
6	2006	January	south	ice cream	Edies	647	$2,076.87
7	2005	September	west	fruit	plums	739	$1,707.09

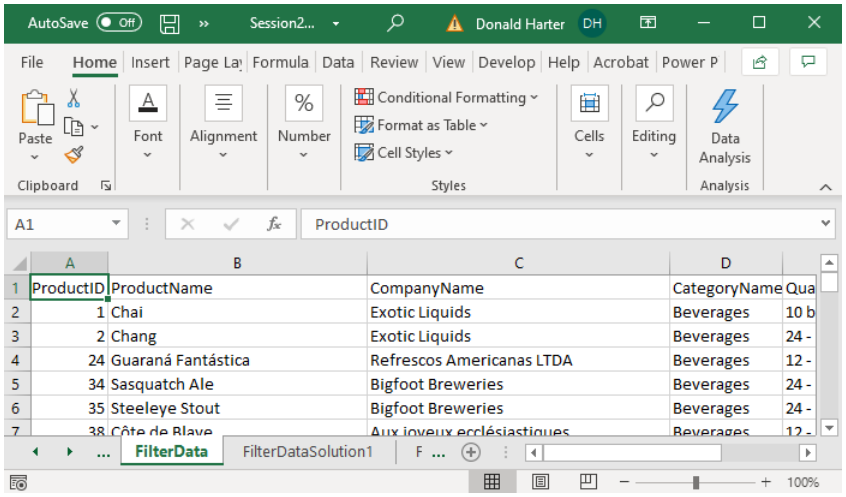
2. To sort data, highlight the top of the columns (A through G). The entire columns should be highlighted.
3. Click on the Data tab at the top of the spreadsheet.
4. Click on Sort; note that “My data has headers” is checked. Why is it checked?
5. Select the first sort criteria: Store
6. Click on Add Level, then add the second sort criteria: Group



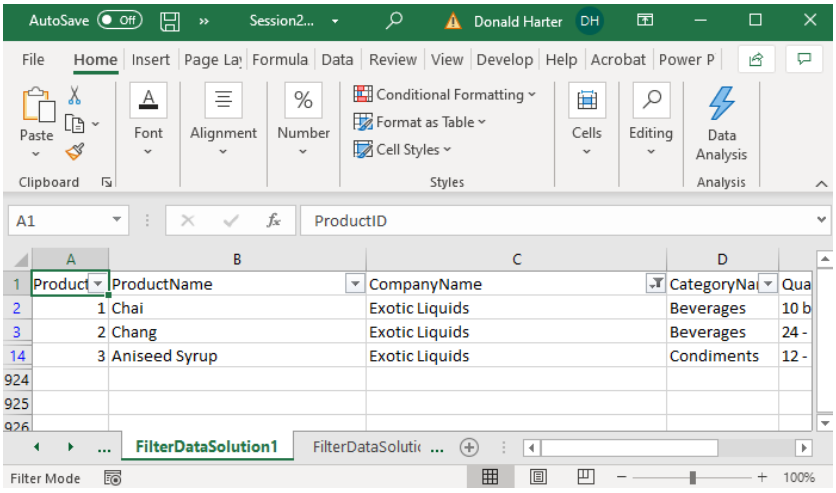
7. Click OK, and the data is sorted
8. How would you sort from Z to A (reverse alphabetical order)? Why would you want to sort in a different order?

2.10 Filters

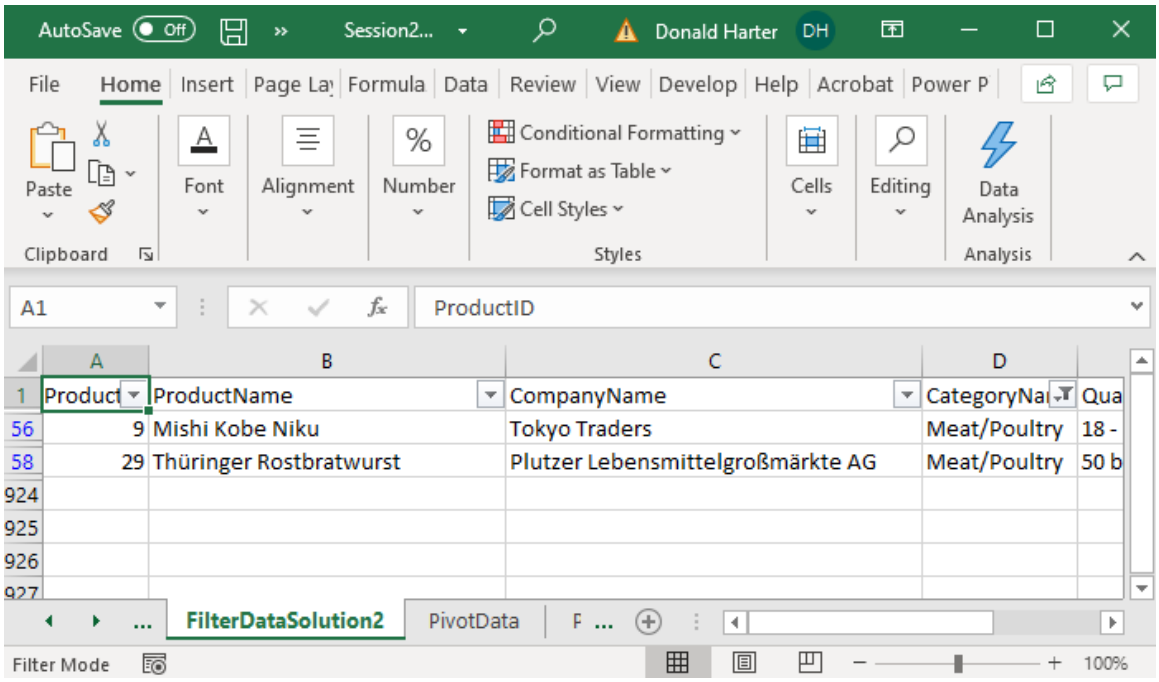
1. Use the FilterData spreadsheet for this exercise
2. A filter allows you to select specific data out of a long list
3. First, click on cell A1, then click on the Data tab, then Filter (looks like a funnel)



4. Notice that there are drop down arrows for each column. Let's find all products which Exotic Liquids produces.
5. Click on the drop-down arrow next to company name; uncheck "(Select All)", then check Exotic Liquids and OK. What happens?



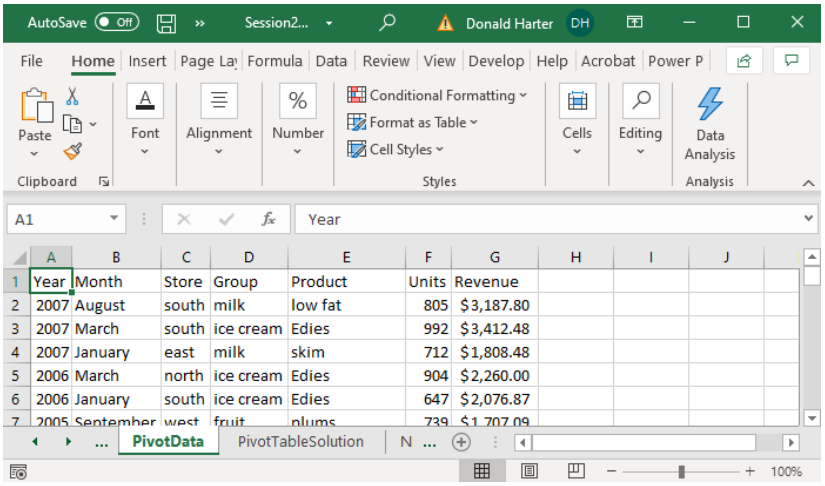
6. Go back and click Select All for Company Name
7. To find which products have a unit price greater than 50, click on the down arrow next to unit price, Number Filters, Greater Than, enter 50 in the field, and click OK
8. You can turn on multiple filters. Add a filter for Category Name equal to Meat/Poultry to see which Meat/Poultry products cost more than 50



9. To turn off the filter, click on the Filter icon and the drop-down arrows should disappear.

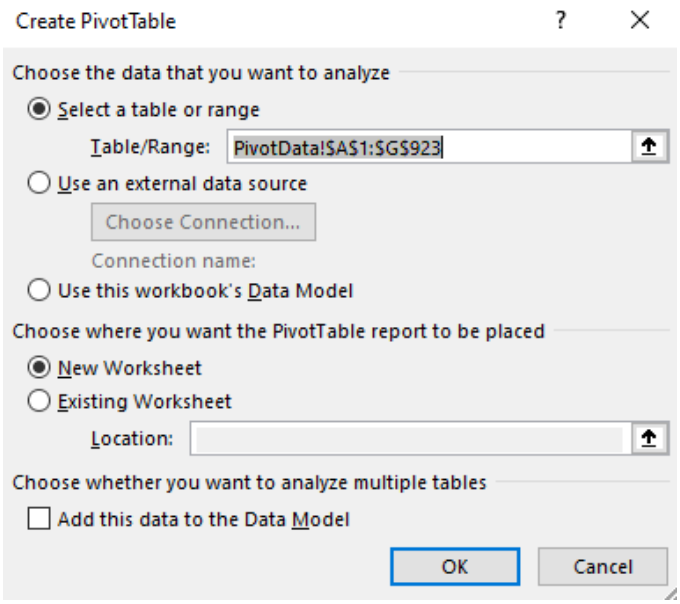
2.11 Pivot Tables and Charts

Pivot tables and charts are powerful techniques to quickly summarize and display large amounts of data. For this example, use the PivotData spreadsheet.



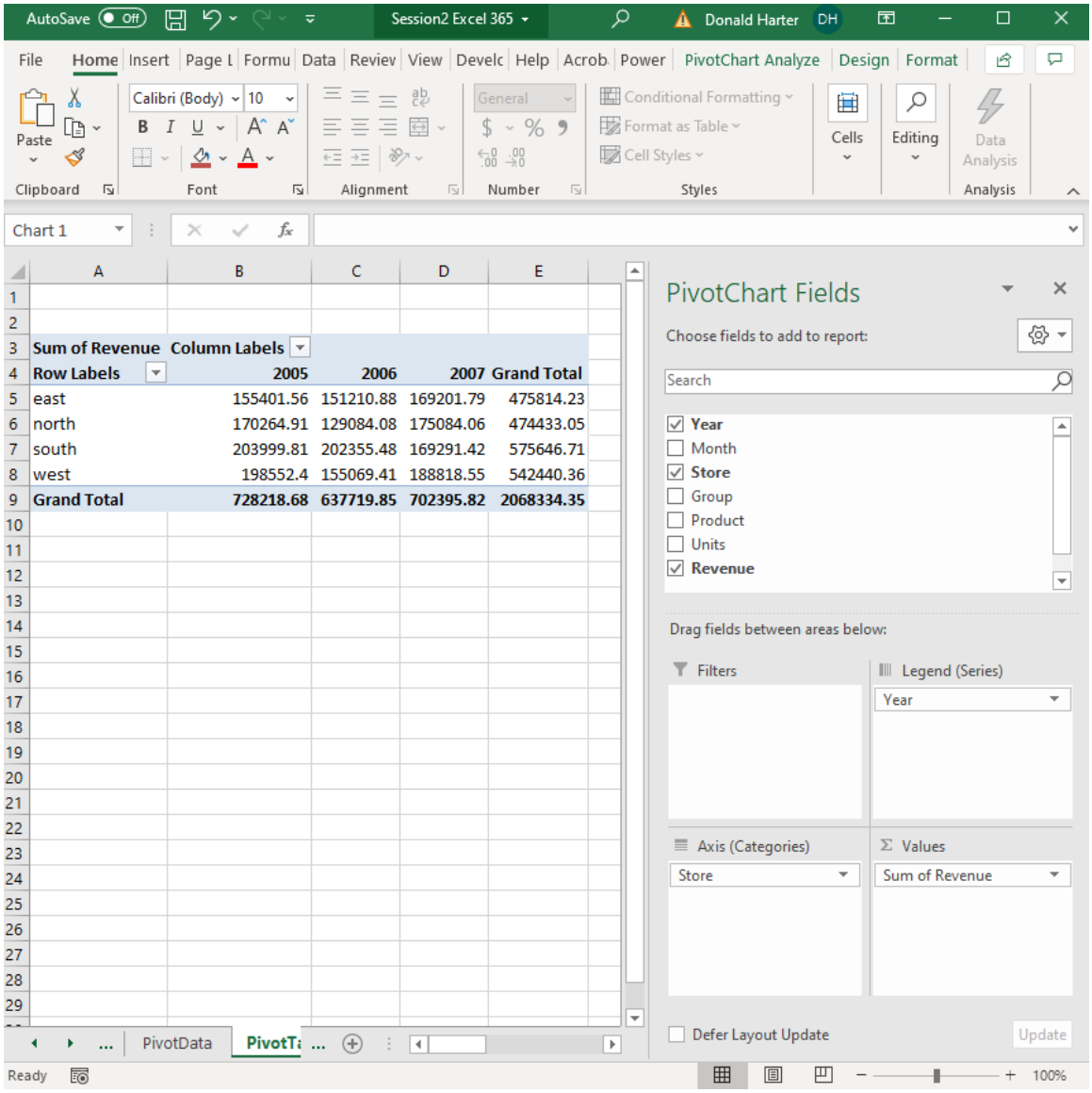
	A	B	C	D	E	F	G
1	Year	Month	Store	Group	Product	Units	Revenue
2	2007	August	south	milk	low fat	805	$3,187.80
3	2007	March	south	ice cream	Edies	992	$3,412.48
4	2007	January	east	milk	skim	712	$1,808.48
5	2006	March	north	ice cream	Edies	904	$2,260.00
6	2006	January	south	ice cream	Edies	647	$2,076.87
7	2005	September	west	fruit	plums	739	$1,707.09

1. The headings must be in the first row of your data.
2. Click anywhere in the data table.
3. Click on the Insert tab.
4. Click on PivotTable; it should automatically highlight the entire table.
5. Make sure New Worksheet is checked, then click OK.



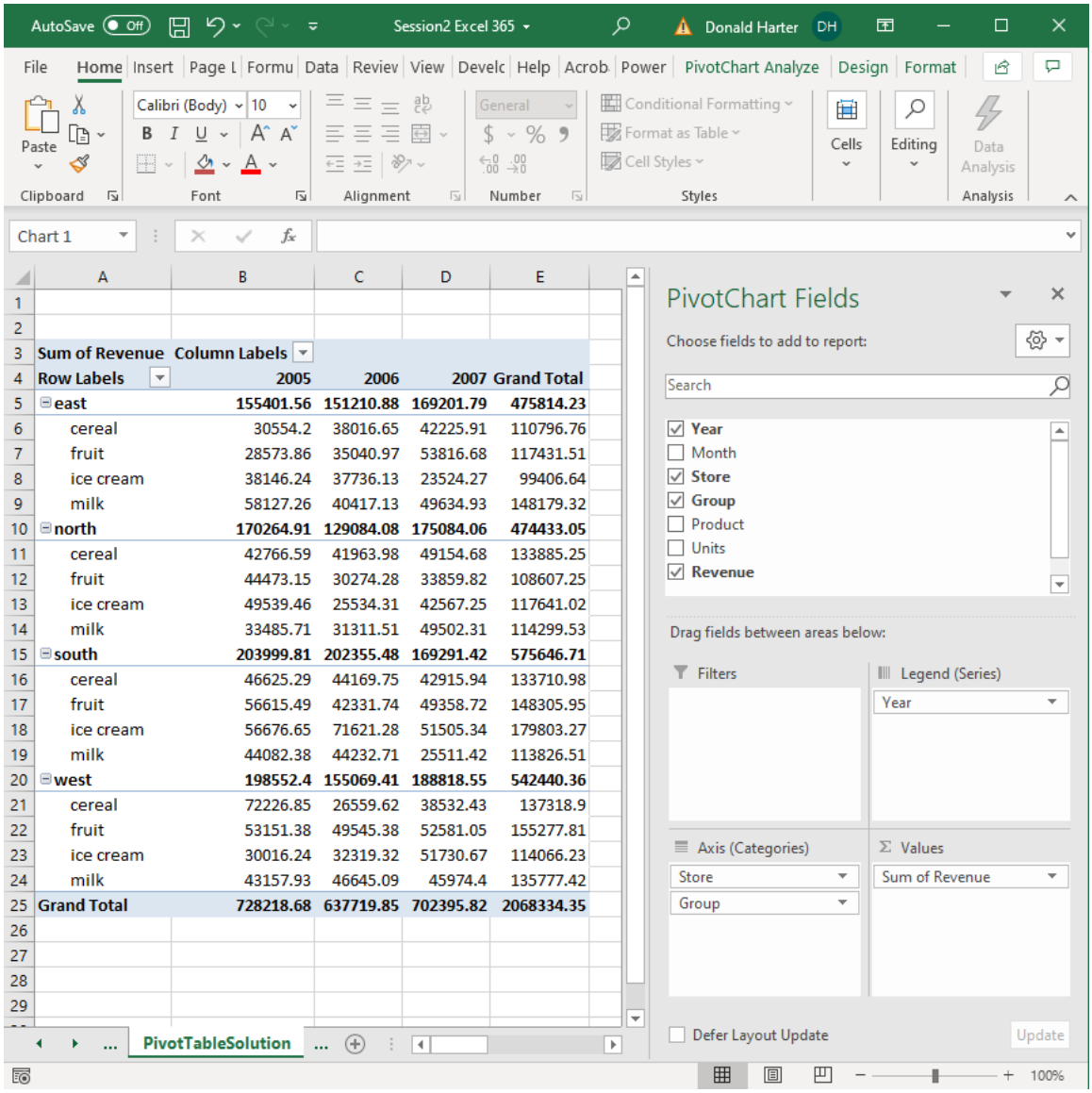
6. In the pivot table dialog box, we want to add row labels, column labels, and identify what goes into the table.
7. For this example, let's build a table with store revenue by year. First click on store and drag it to the row label. These will become the row labels on the left side of your table.
8. Next, click on year and drag it to the column label. These will become the column labels at the top of your table.

9. Finally, click on revenue and drag it to the values cell. These values will feed the interior of the table. The default function is Sum. If you want to change this function, click on the down arrow in the values box, then select the function (average, min, max, etc.)



Sum of Revenue	Column Labels			
Row Labels	2005	2006	2007	Grand Total
east	155401.56	151210.88	169201.79	475814.23
north	170264.91	129084.08	175084.06	474433.05
south	203999.81	202355.48	169291.42	575646.71
west	198552.4	155069.41	188818.55	542440.36
Grand Total	728218.68	637719.85	702395.82	2068334.35

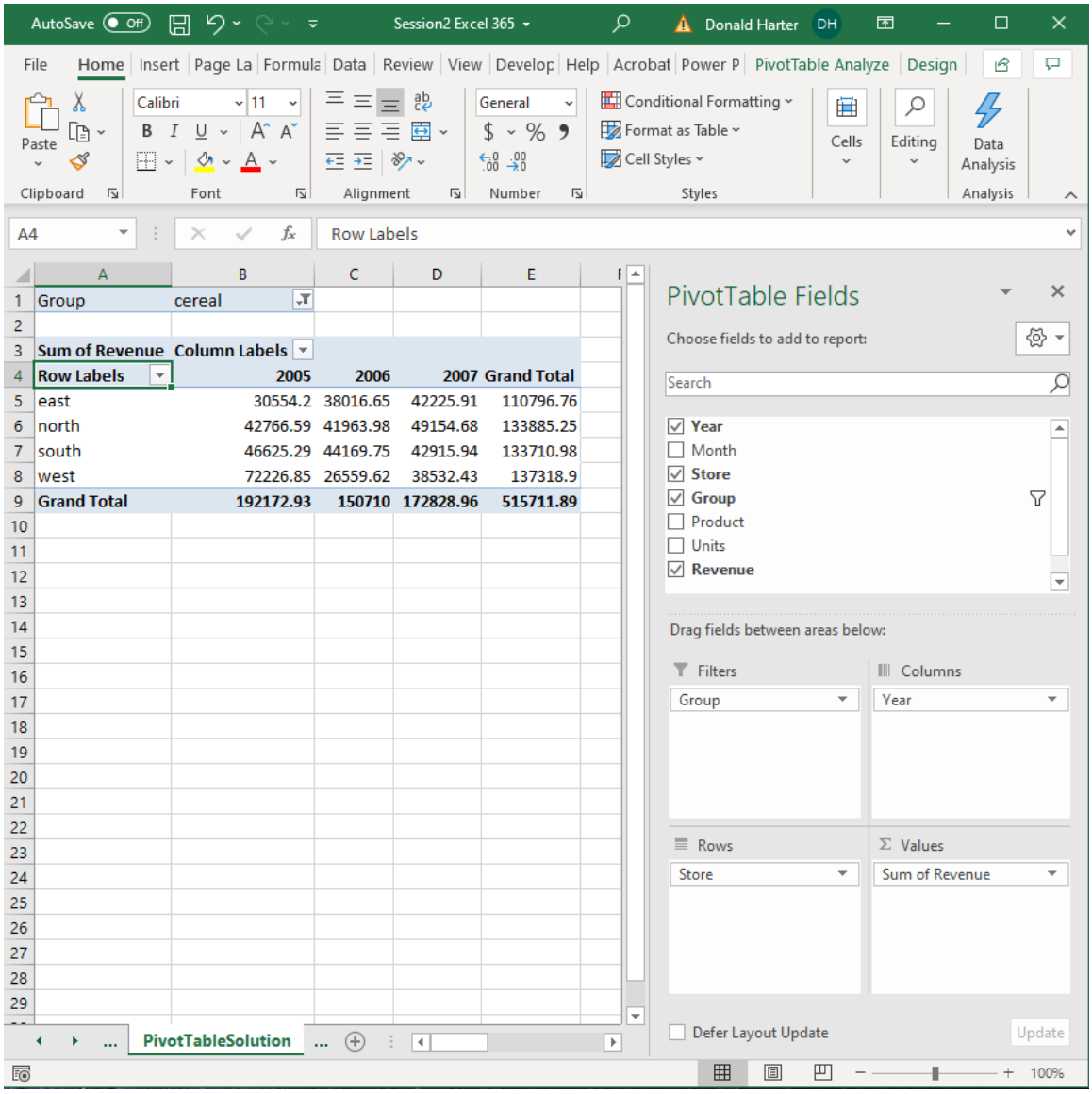
10. Add group to the row labels to see what happens.



Sum of Revenue	Column Labels			
Row Labels	2005	2006	2007	Grand Total
east	155401.56	151210.88	169201.79	475814.23
cereal	30554.2	38016.65	42225.91	110796.76
fruit	28573.86	35040.97	53816.68	117431.51
ice cream	38146.24	37736.13	23524.27	99406.64
milk	58127.26	40417.13	49634.93	148179.32
north	170264.91	129084.08	175084.06	474433.05
cereal	42766.59	41963.98	49154.68	133885.25
fruit	44473.15	30274.28	33859.82	108607.25
ice cream	49539.46	25534.31	42567.25	117641.02
milk	33485.71	31311.51	49502.31	114299.53
south	203999.81	202355.48	169291.42	575646.71
cereal	46625.29	44169.75	42915.94	133710.98
fruit	56615.49	42331.74	49358.72	148305.95
ice cream	56676.65	71621.28	51505.34	179803.27
milk	44082.38	44232.71	25511.42	113826.51
west	198552.4	155069.41	188818.55	542440.36
cereal	72226.85	26559.62	38532.43	137318.9
fruit	53151.38	49545.38	52581.05	155277.81
ice cream	30016.24	32319.32	51730.67	114066.23
milk	43157.93	46645.09	45974.4	135777.42
Grand Total	728218.68	637719.85	702395.82	2068334.35

11. Uncheck Group to remove this level of detail.

12. To add a filter to the pivot table, drag Group to Filters.
13. In the upper left corner, click on cell B1 to use the drop-down arrow to select Cereal

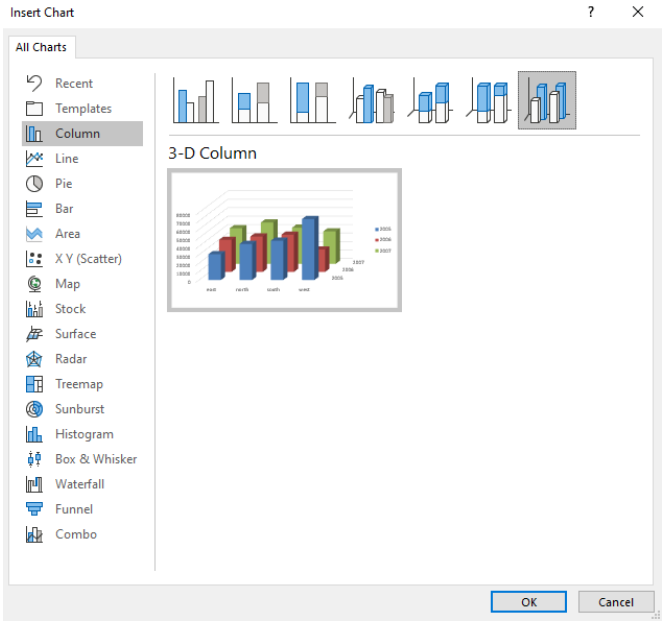


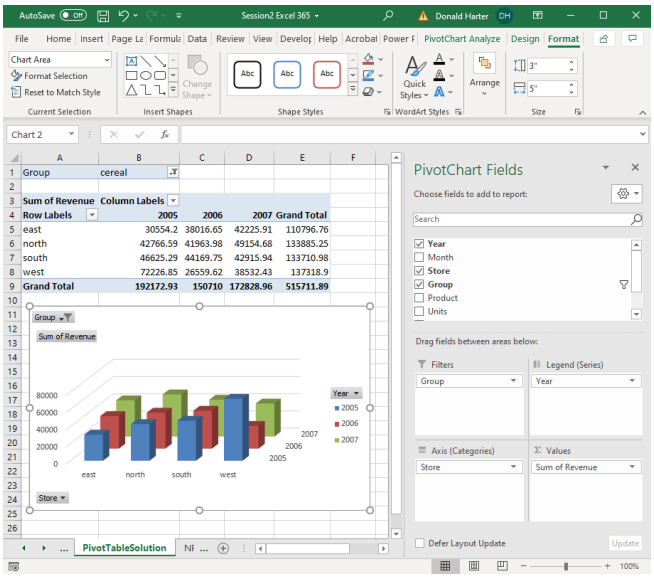
14. To select more than one Group, use the drop-down arrow in cell B1, check the box Select Multiple Items, then choose the Groups that you want to include in the filter
15. In the Pivot table fields box, uncheck Groups to remove the filter.

2.12 Pivot Chart

Let's now create a Pivot Chart. The steps are:

1. Uncheck Group to simplify the data.
2. Click anywhere inside your Pivot Table. An Analyze tab will appear. Click on the Analyze tab, then Pivot Chart.
3. A chart dialog box will appear. Click on your preferred type of chart.





To move the chart to its own page:

4. Right click on the picture
5. Click on Move Chart
6. Select New Sheet and name the new sheet

Microsoft Excel: Statistics

2.15 Data Analysis Add-in

The statistics options are available as an add-in to Excel. The steps to add it are:

1. In Excel, click on the File tab, then Options
2. Click on Add-Ins
3. Click Analysis ToolPak Add-in, then Go
4. Check the box for Analysis ToolPak, then OK

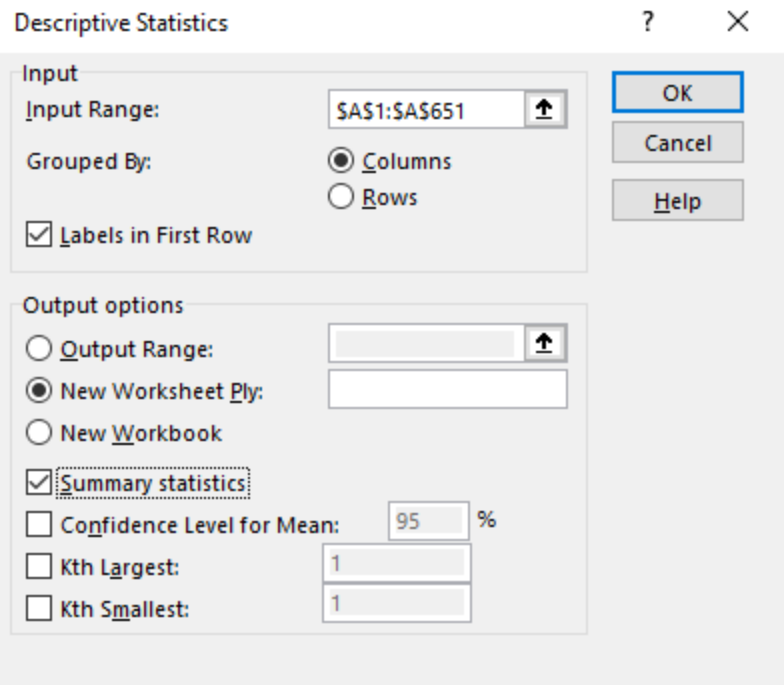
2.16 Data Analysis: Descriptive statistics

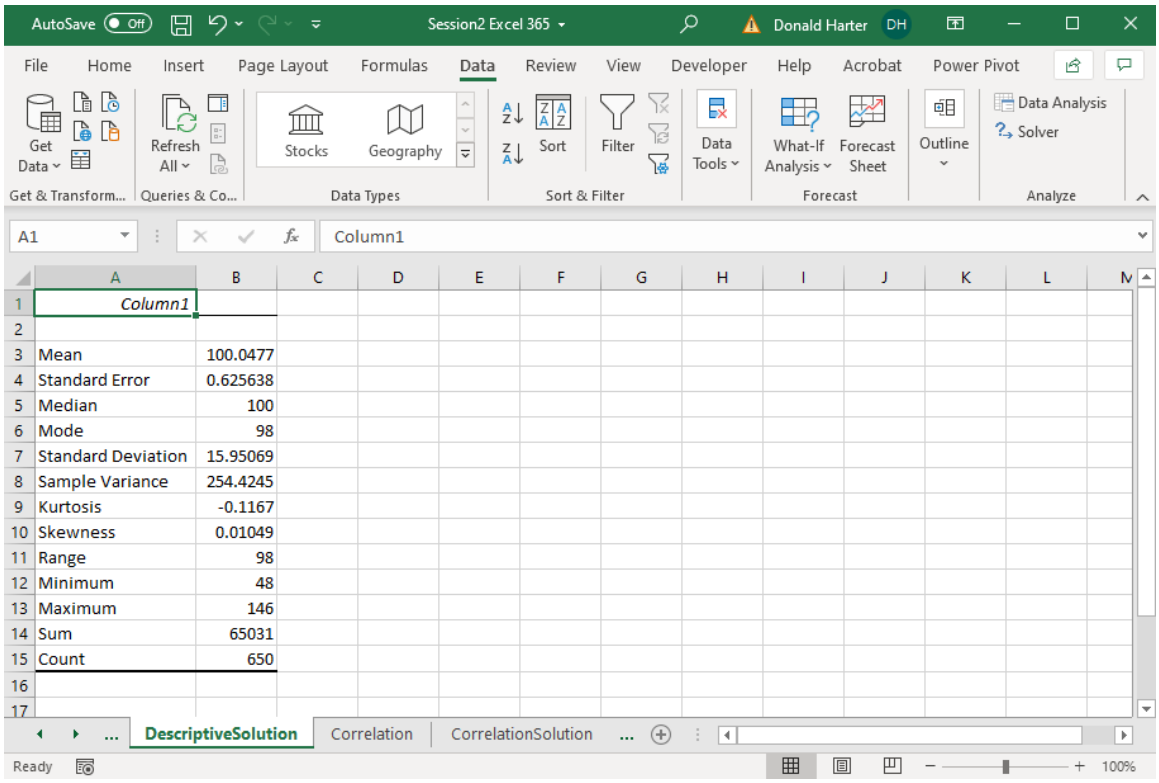
Use the DescriptiveStatistics spreadsheet tab for this exercise.
There are several descriptive statistics that can be automatically generated, including:

- Mean: arithmetic average
- Median: middle point in distribution
- Mode: most common value (highest frequency of occurrence)
- Kurtosis: is the data peaked higher or lower than normal?
- Skewness; is the peak shifted left or right?
- Standard deviation: measure of spread
- Range: highest value minus lowest value

To calculate the descriptive statistics:

1. Click on the data tab, then data analysis, descriptive statistics, and OK.
2. Enter the input range for the IQ data; if you include the header, click on Labels in first row.
3. Check Summary Statistics, then OK

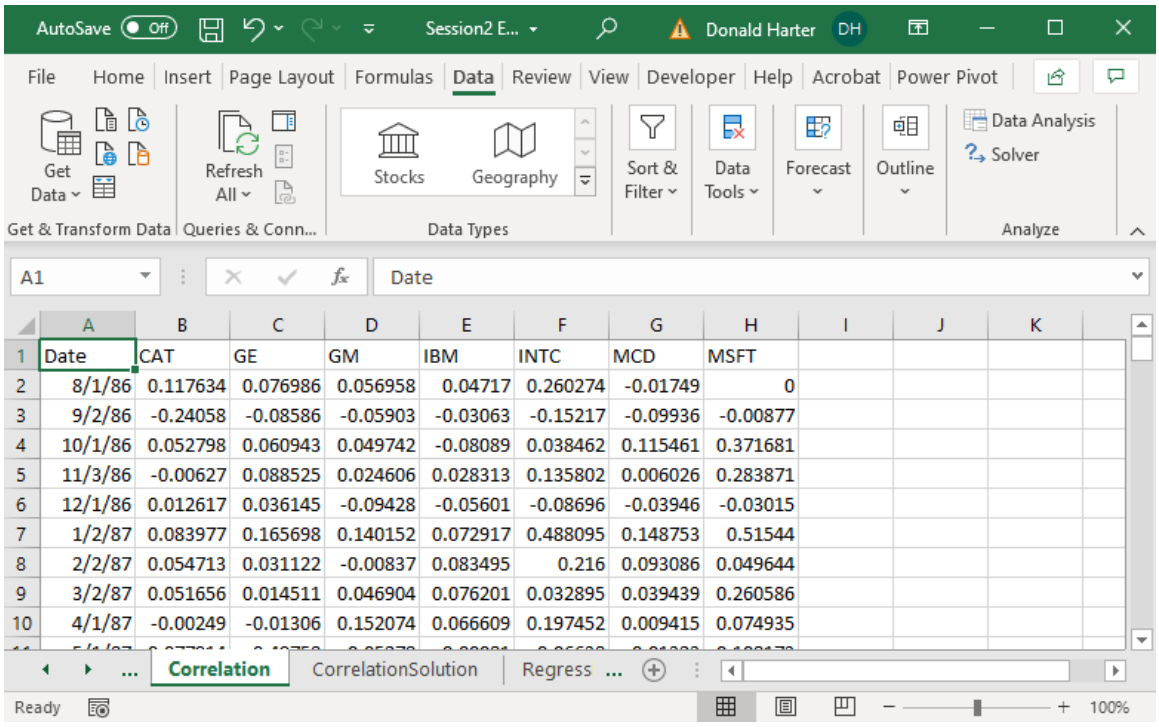




	A	B
1	Column1	
2		
3	Mean	100.0477
4	Standard Error	0.625638
5	Median	100
6	Mode	98
7	Standard Deviation	15.95069
8	Sample Variance	254.4245
9	Kurtosis	-0.1167
10	Skewness	0.01049
11	Range	98
12	Minimum	48
13	Maximum	146
14	Sum	65031
15	Count	650

2.17 Correlations

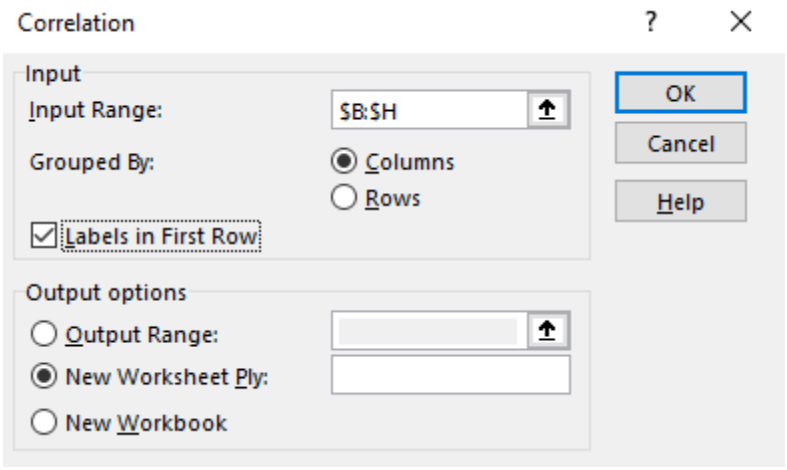
Correlation analysis identifies how two or more variables are related. For this exercise, use the Correlation spreadsheet. This spreadsheet records the upward or downward movement of stock by month.



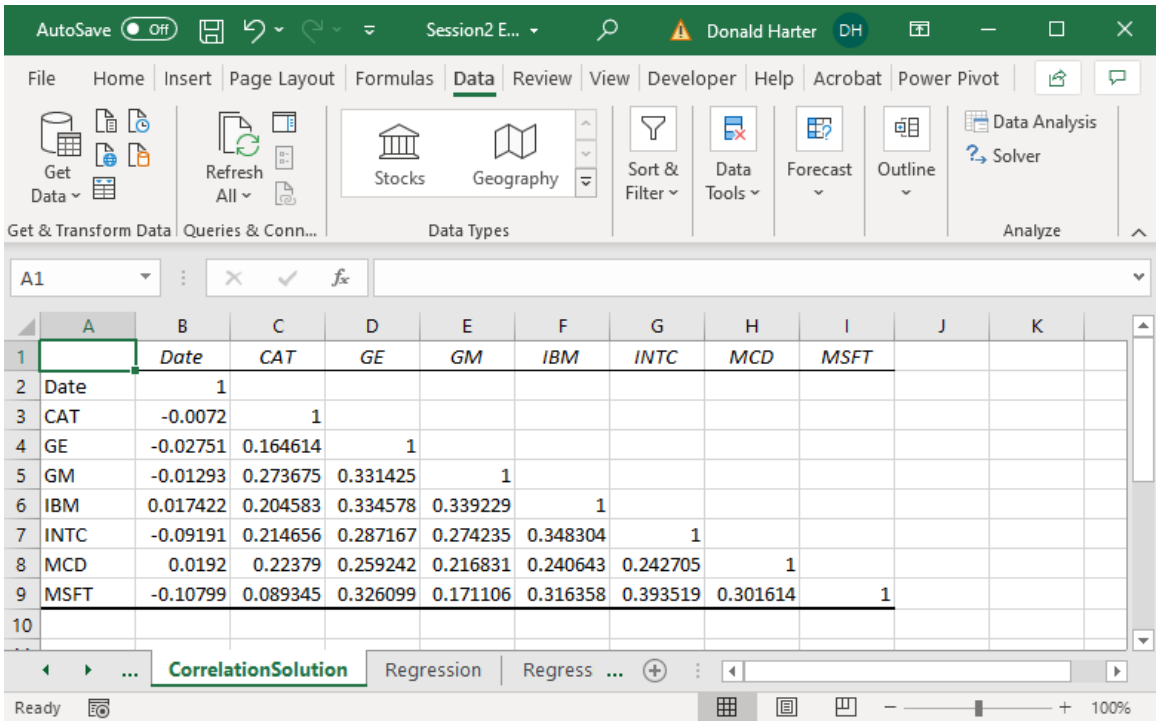
	Date	CAT	GE	GM	IBM	INTC	MCD	MSFT
2	8/1/86	0.117634	0.076986	0.056958	0.04717	0.260274	-0.01749	0
3	9/2/86	-0.24058	-0.08586	-0.05903	-0.03063	-0.15217	-0.09936	-0.00877
4	10/1/86	0.052798	0.060943	0.049742	-0.08089	0.038462	0.115461	0.371681
5	11/3/86	-0.00627	0.088525	0.024606	0.028313	0.135802	0.006026	0.283871
6	12/1/86	0.012617	0.036145	-0.09428	-0.05601	-0.08696	-0.03946	-0.03015
7	1/2/87	0.083977	0.165698	0.140152	0.072917	0.488095	0.148753	0.51544
8	2/2/87	0.054713	0.031122	-0.00837	0.083495	0.216	0.093086	0.049644
9	3/2/87	0.051656	0.014511	0.046904	0.076201	0.032895	0.039439	0.260586
10	4/1/87	-0.00249	-0.01306	0.152074	0.066609	0.197452	0.009415	0.074935

The stocks listed are Caterpillar, General Electric, General Motors, IBM, Intel, McDonalds and Microsoft.

1. Click on the data tab, data analysis, correlation, then OK.
2. Highlight columns B through H (don't include the date column), group by columns, check Labels in First Row, then OK.



The result is shown below. A positive correlation means that when one variable increases, the other increases. A negative correlation means that when one increases, the other decreases.



	A	B	C	D	E	F	G	H	I
1		*Date*	*CAT*	*GE*	*GM*	*IBM*	*INTC*	*MCD*	*MSFT*
2	Date	1							
3	CAT	-0.0072	1						
4	GE	-0.02751	0.164614	1					
5	GM	-0.01293	0.273675	0.331425	1				
6	IBM	0.017422	0.204583	0.334578	0.339229	1			
7	INTC	-0.09191	0.214656	0.287167	0.274235	0.348304	1		
8	MCD	0.0192	0.22379	0.259242	0.216831	0.240643	0.242705	1	
9	MSFT	-0.10799	0.089345	0.326099	0.171106	0.316358	0.393519	0.301614	1
10									

2.18 Regression Assumptions

Regression is a technique that attempts to measure the relationship between and outcome variable (dependent) and explanatory variables (independent). To use linear regression, there are three key assumptions

1. relationship between x and y is linear
2. the x's are fixed numbers, not random variables (non-stochastic), not related to each other, i.e., independent: Corr(X_i,X_j)=0
3. the error terms:
    a. have zero mean and constant variance: $E(\varepsilon_i) = 0$, $V(\varepsilon_i) = \sigma^2$
    b. the error terms are independent: $Cov(\varepsilon_i,\varepsilon_j) = 0$
    c. the error terms are normally distributed $\sim N(0,\sigma^2)$

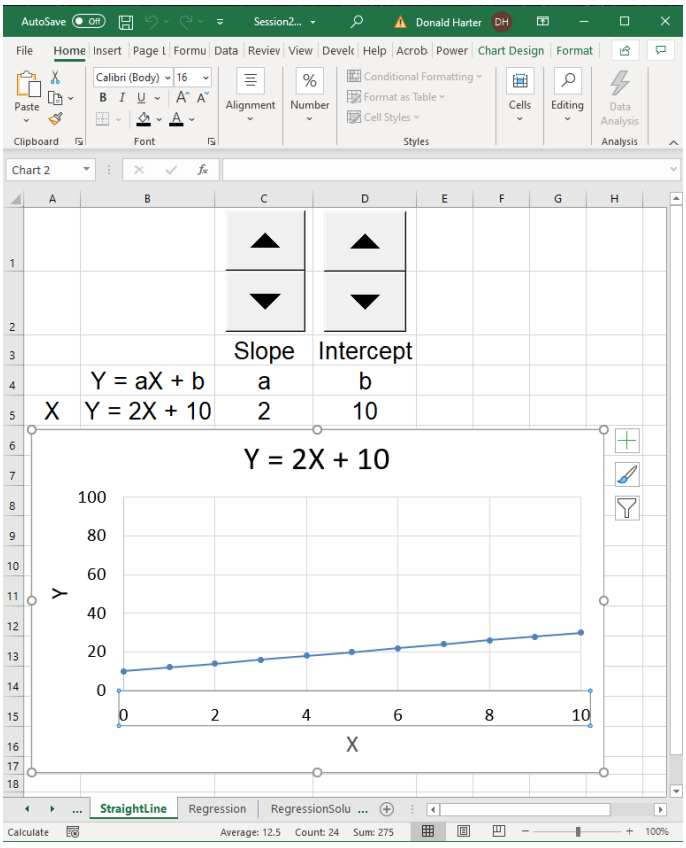
Violation of these assumptions requires the use of more sophisticated techniques.

2.19 Linear Regression: Straight line relationships

Before performing a linear regression, let us first review the equation of a straight line. A line is represented by the equation:

$$Y = aX + b$$

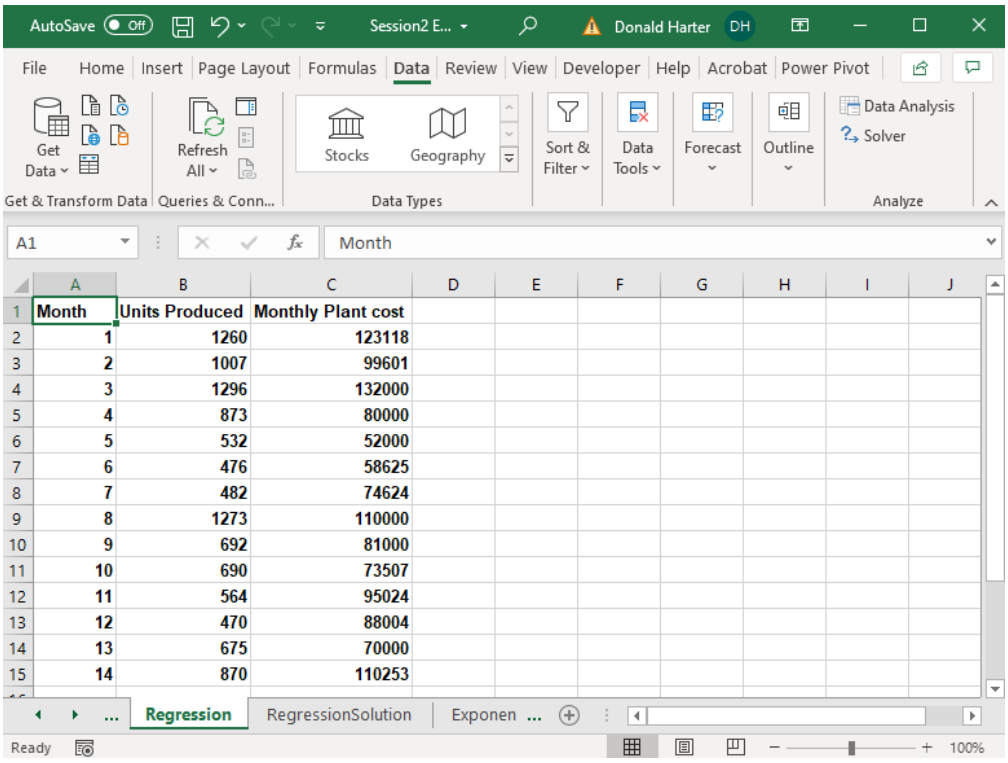
Where a is the slope of X and b is the intercept. We will use the example in the Straight Line spreadsheet tab to demonstrate what happens when the slope (coefficient of X) or the intercept change.



Next, we will perform a linear regression on factory costs (Y) explained by the number of units produced in a factory.

When you want to determine if there is a straight-line relationship in statistics, you can run a regression. Excel has the ability to perform regression analysis. For example, if you wanted to model the relationship between items produced and factory costs, you could estimate the linear relationship. Units produced would be called the independent variable; production costs would be the dependent variable. The output, costs, depends on the input, number of units produced.

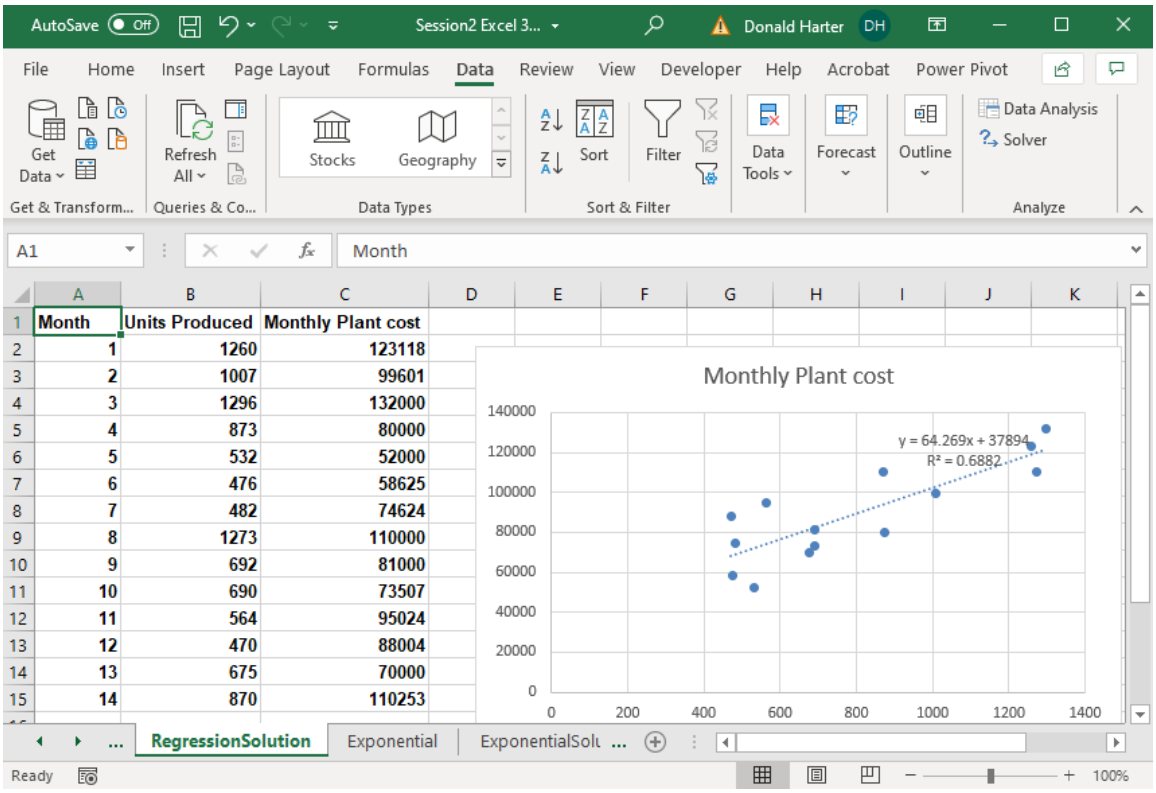
For this example, use the Regression spreadsheet.



	A	B	C
1	Month	Units Produced	Monthly Plant cost
2	1	1260	123118
3	2	1007	99601
4	3	1296	132000
5	4	873	80000
6	5	532	52000
7	6	476	58625
8	7	482	74624
9	8	1273	110000
10	9	692	81000
11	10	690	73507
12	11	564	95024
13	12	470	88004
14	13	675	70000
15	14	870	110253

Let's first draw a scatterplot to see what the data looks like.

1. Click on the Insert tab
2. Highlight the cells b1:c15
3. Click on Scatter in the charts group.



Month	Units Produced	Monthly Plant cost
1	1260	123118
2	1007	99601
3	1296	132000
4	873	80000
5	532	52000
6	476	58625
7	482	74624
8	1273	110000
9	692	81000
10	690	73507
11	564	95024
12	470	88004
13	675	70000
14	870	110253

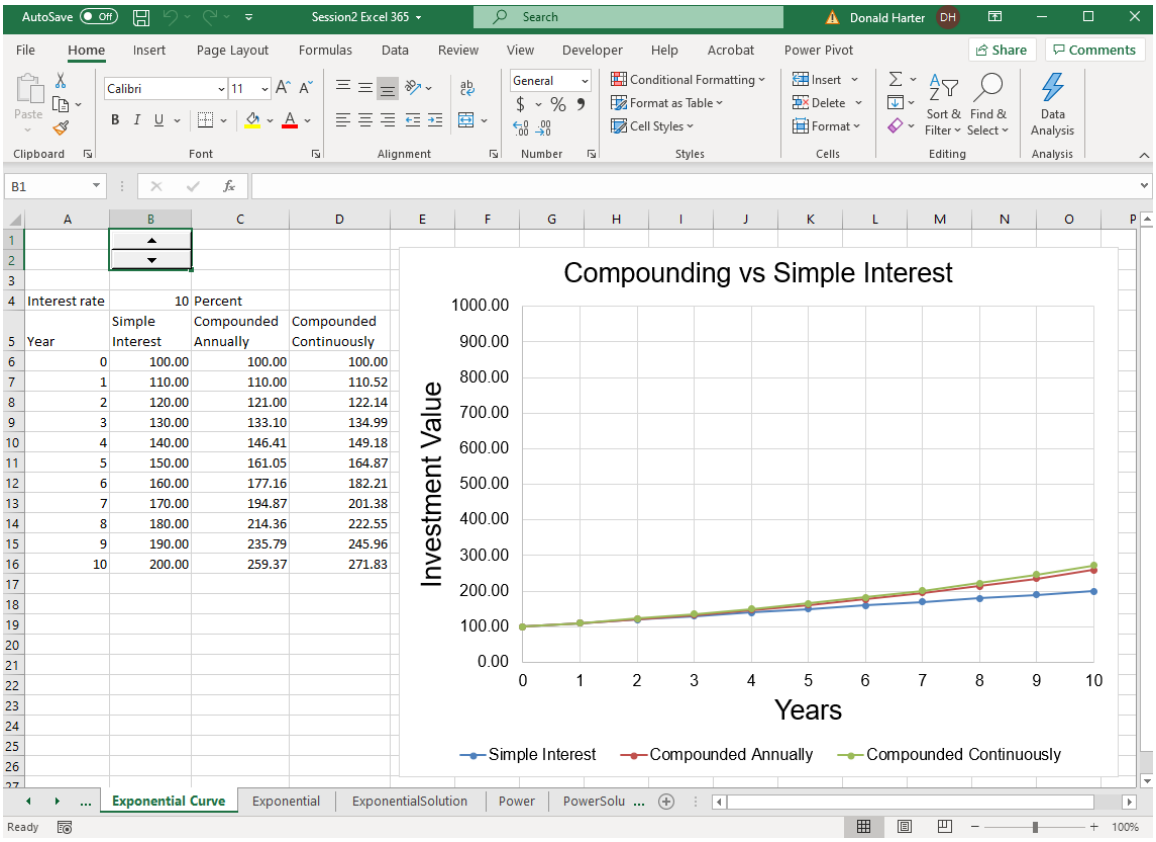
4. Reviewing the chart, it appears that there is a linear relationship. We will therefore perform a linear regression. Click on any data point, right click, then add trendline.

5. In the Format Trendline, Trendline Options, select Linear, then check the boxes for Display Equation and Display R-squared value.

6. In the picture above, the coefficient on x is approximately 64. This means that as unit production increases by one, costs increase by $64. What does the number 37,894 represent? What does the $R^2 = 0.6882$ mean?

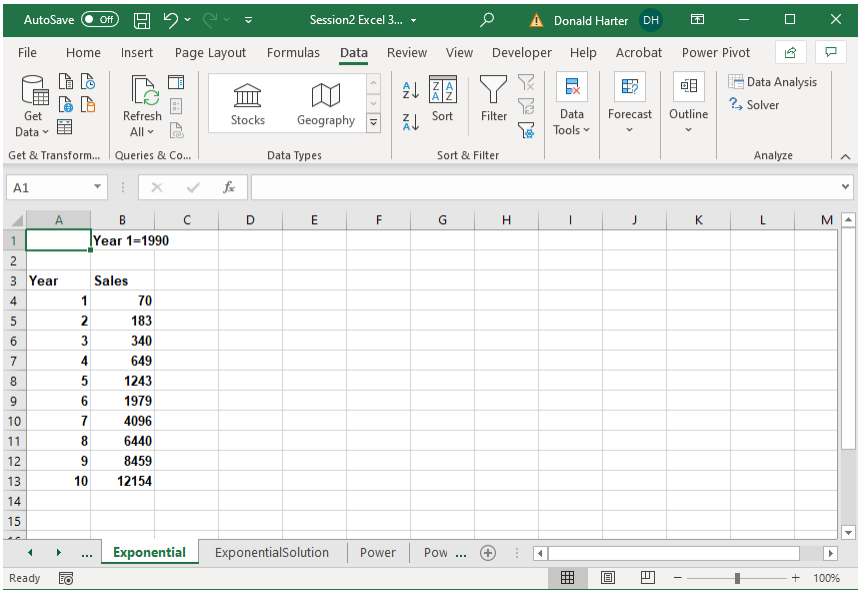
2.20 Exponential Regression: Exponential growth, compounded growth

Before performing an exponential regression, we will examine what an exponential curve looks like. For this demonstration, use the Exponential Curve spreadsheet tab.



	A	B	C	D
4	Interest rate	10	Percent	
5	Year	Simple Interest	Compounded Annually	Compounded Continuously
6	0	100.00	100.00	100.00
7	1	110.00	110.00	110.52
8	2	120.00	121.00	122.14
9	3	130.00	133.10	134.99
10	4	140.00	146.41	149.18
11	5	150.00	161.05	164.87
12	6	160.00	177.16	182.21
13	7	170.00	194.87	201.38
14	8	180.00	214.36	222.55
15	9	190.00	235.79	245.96
16	10	200.00	259.37	271.83

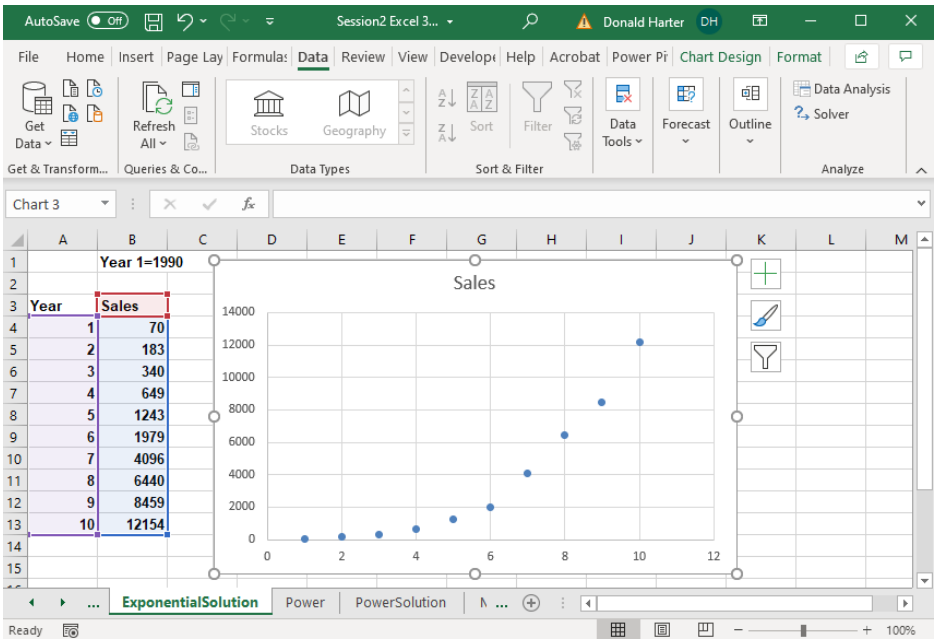
Some data relationships are not linear, but grow at an increasing rate. These curves often follow the exponential growth curve. An exponential growth curve will have the same percentage growth per period compounded over time. Use the Exponential spreadsheet.



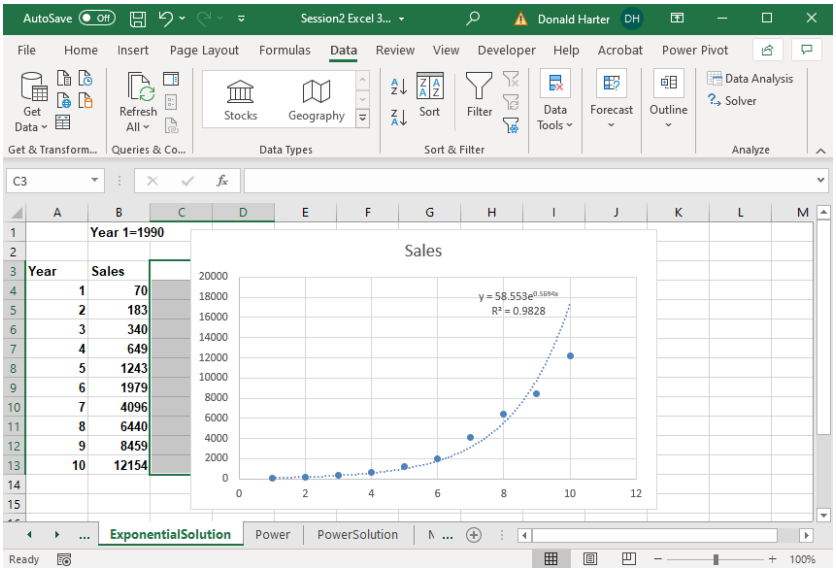
	A	B
1		Year 1=1990
3	Year	Sales
4	1	70
5	2	183
6	3	340
7	4	649
8	5	1243
9	6	1979
10	7	4096
11	8	6440
12	9	8459
13	10	12154

Let's first draw a scatterplot to see what the data looks like.

1. Click on the Insert tab
2. Highlight the cells a3:b13
3. Click on Scatter in the charts group.

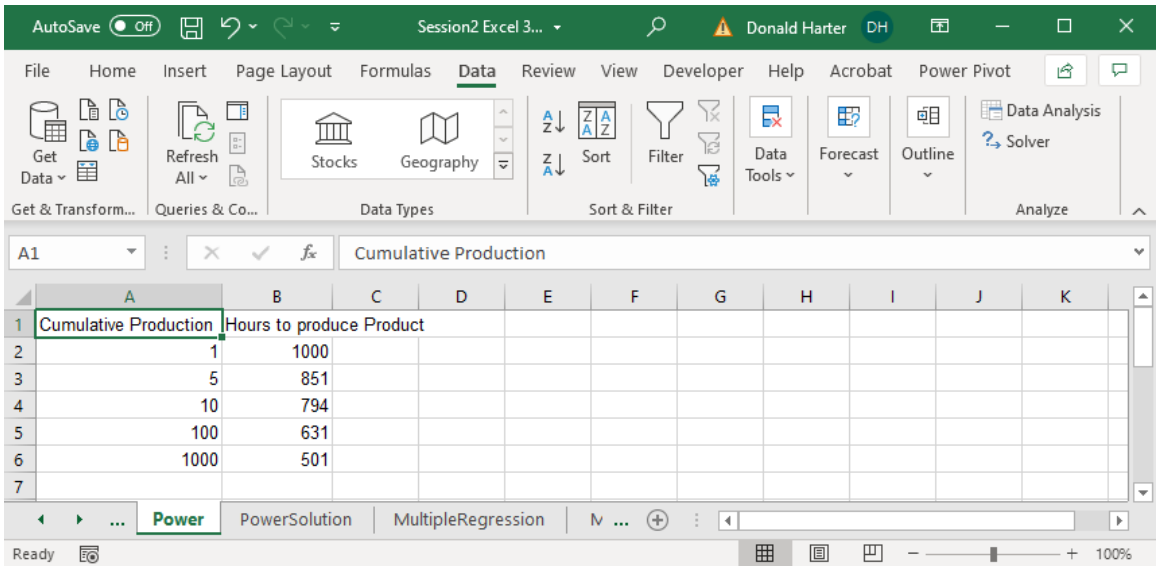


4. This data definitely does not look linear. So let's use the exponential curve. Click on any data point, right click, then add trendline. Select exponential, display equation and display R-squared, then Close.



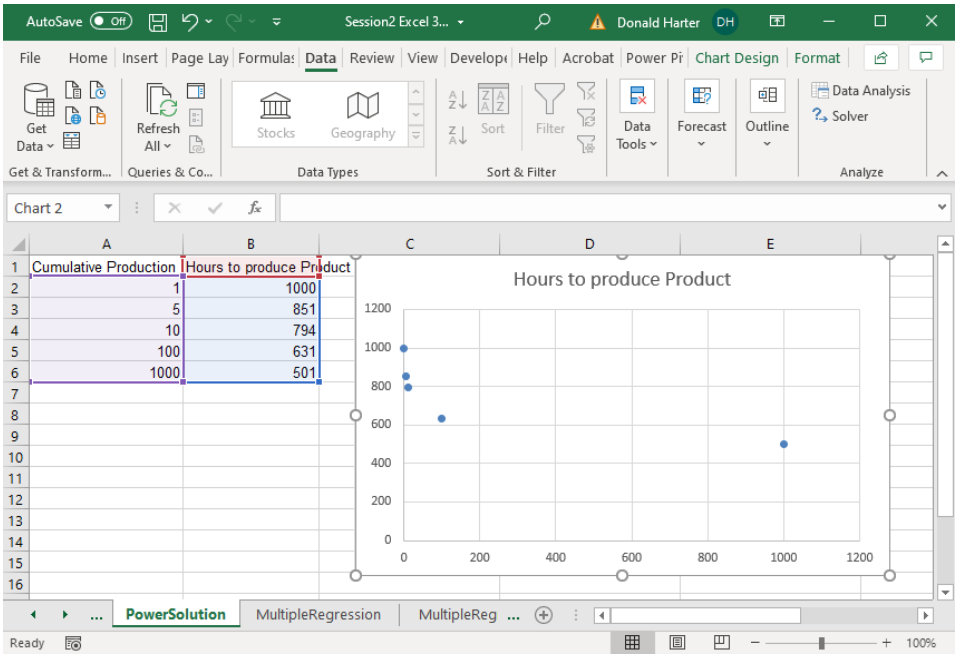
2.21 Power Regression: Power curve for learning curve, economies of scale

The power curve allows you to examine economies of scale and diseconomies of scale. Economies of scale means that you become more efficient as volume increases. Diseconomies of scale means that you become less efficient as volume increases. Use the Power spreadsheet.



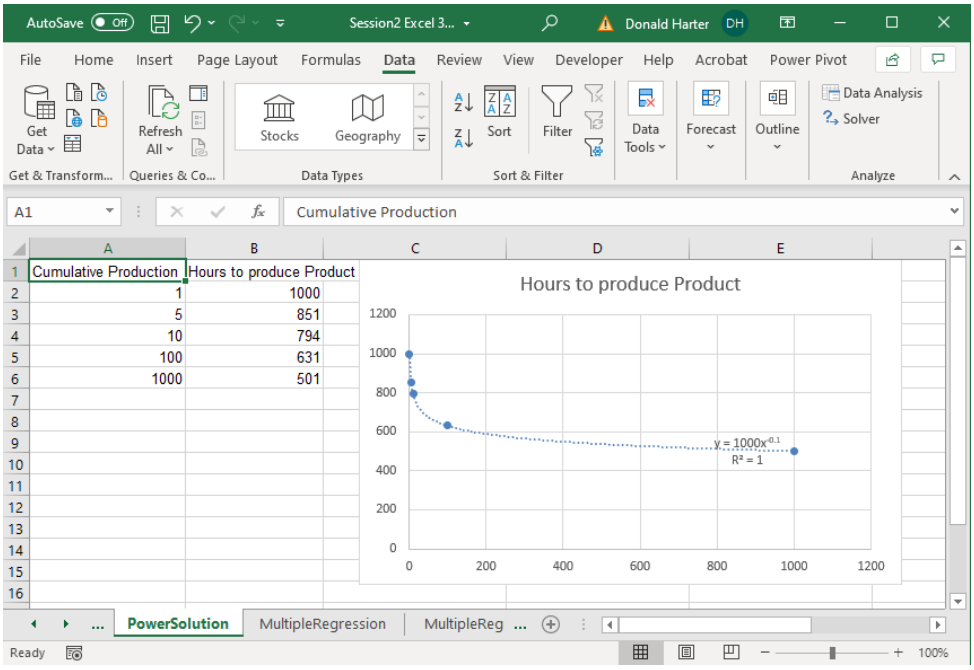
Let's graph as before.

1. Click on the Insert tab.
2. Highlight the cells a1:b6
3. Click on Scatter in the charts group



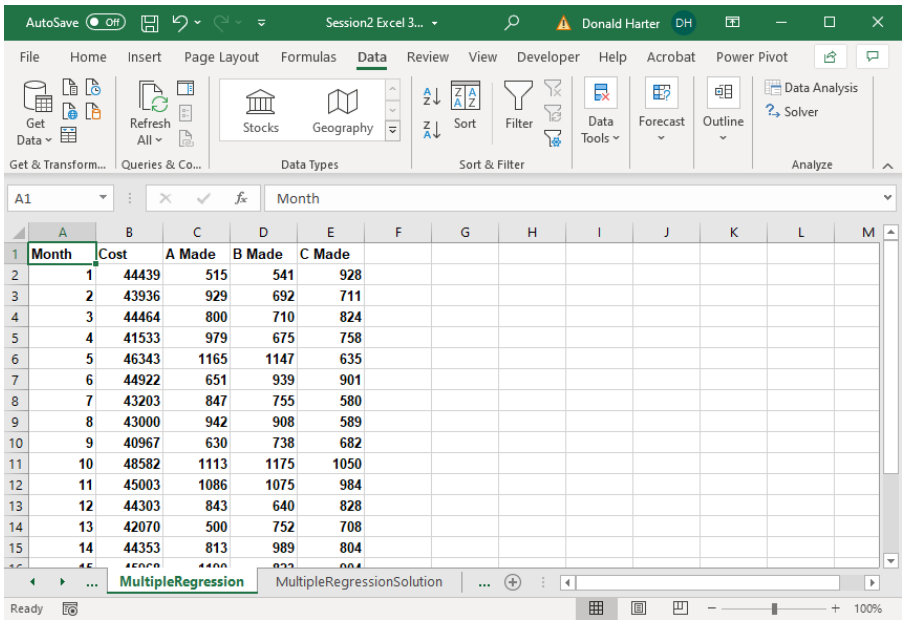
4. Now click on any data point in the graph, right click, add trendline.

5. Click on Power, display equation, display R-squared.



2.22 Multiple regression

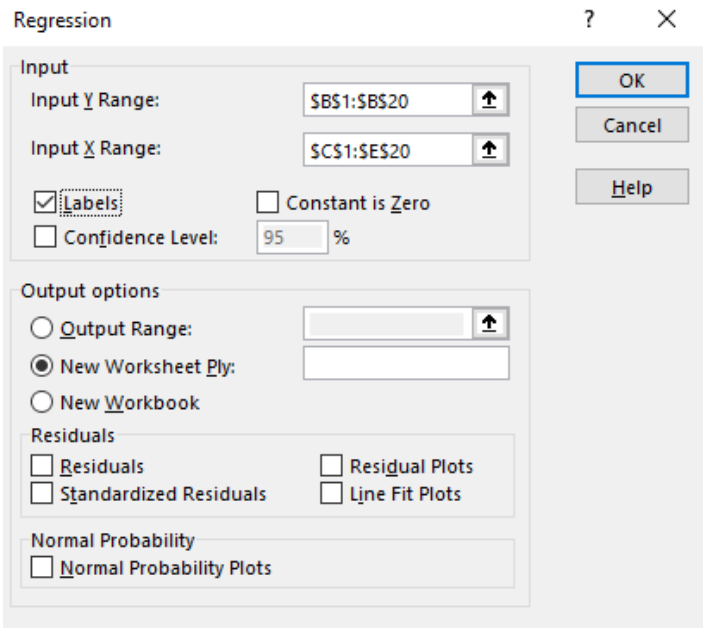
When we reviewed linear regression earlier, we only had one independent variable. Multiple regression includes several independent variables. Use the Multiple Regression spreadsheet.



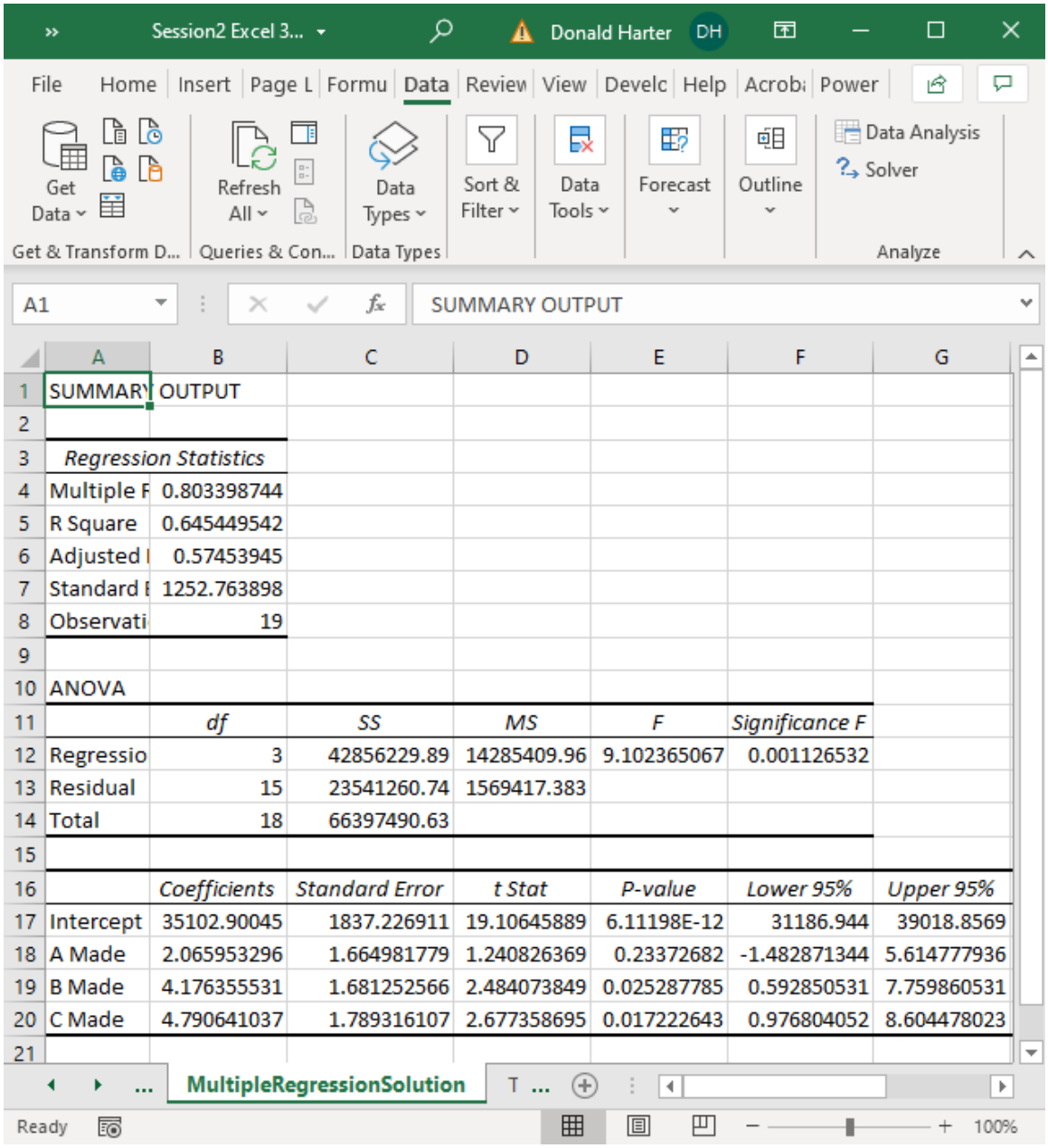
Month	Cost	A Made	B Made	C Made
1	44439	515	541	928
2	43936	929	692	711
3	44464	800	710	824
4	41533	979	675	758
5	46343	1165	1147	635
6	44922	651	939	901
7	43203	847	755	580
8	43000	942	908	589
9	40967	630	738	682
10	48582	1113	1175	1050
11	45003	1086	1075	984
12	44303	843	640	828
13	42070	500	752	708
14	44353	813	989	804

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.



6. The first test is whether the equation is good. If the Significance of F is < 0.05, then the equation is significant. If the equation is not significant, throw it out and do not proceed any further.
7. Second, the R-Square measures the percent change in Y explained by the change in X. In this example, 64.5% of the change in factory cost is explained by the change in the number of A, B, and C produced.
8. Third, the p-value of each coefficient determines if it is significant and can be interpreted. If the p-value is < 0.05, then it is significant. If the p-value is > 0.05, the coefficient is not significant and you cannot say anything about it.



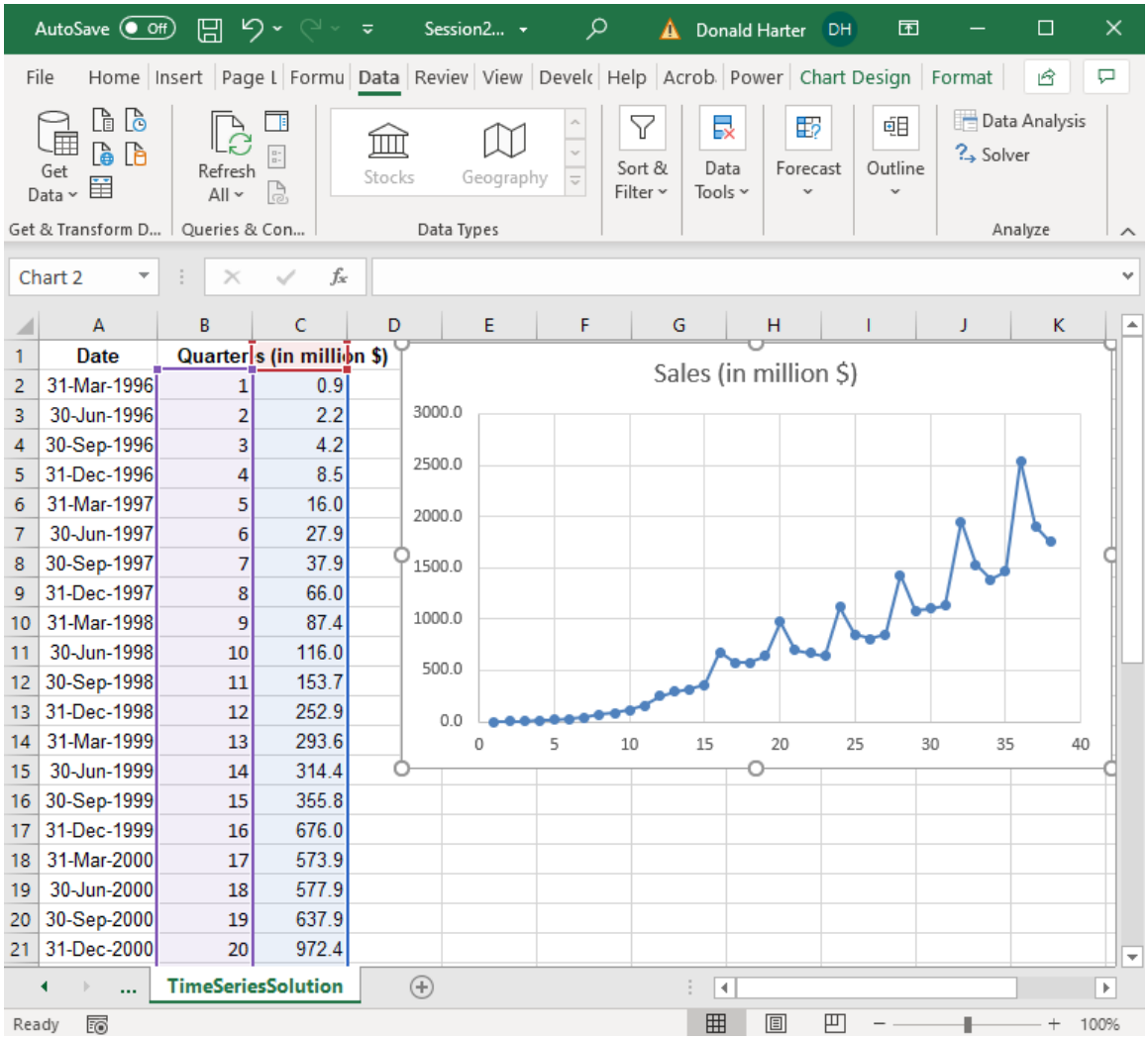
	A	B	C	D	E	F	G
1	SUMMARY	OUTPUT					
2							
3	*Regression Statistics*						
4	Multiple F	0.803398744					
5	R Square	0.645449542					
6	Adjusted I	0.57453945					
7	Standard E	1252.763898					
8	Observati	19					
9							
10	ANOVA						
11		*df*	*SS*	*MS*	*F*	*Significance F*	
12	Regressio	3	42856229.89	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							

2.23 Time Series and Moving Average

Time series problems have data where one data point is dependent on the previous data point. For example, the closing price of Microsoft stock can be tracked day by day. Today's price is dependent on yesterday's price. This dependency from one day to the next, or one time period to the next, is a characteristic of time series data.

Often there is so much variation in time series data that it's hard to see trends. Seasonality also masks a trend. Seasonality is variations in data due to high or low points that occur at regular intervals. Create a scatter plot for the Amazon data using the Time Series spreadsheet. For this scatter plot, select the option to connect the dots.

1. Click on the Insert tab
2. Highlight b1:c39
3. Click on Scatter in the charts group
4. Select the option to connect the dots.

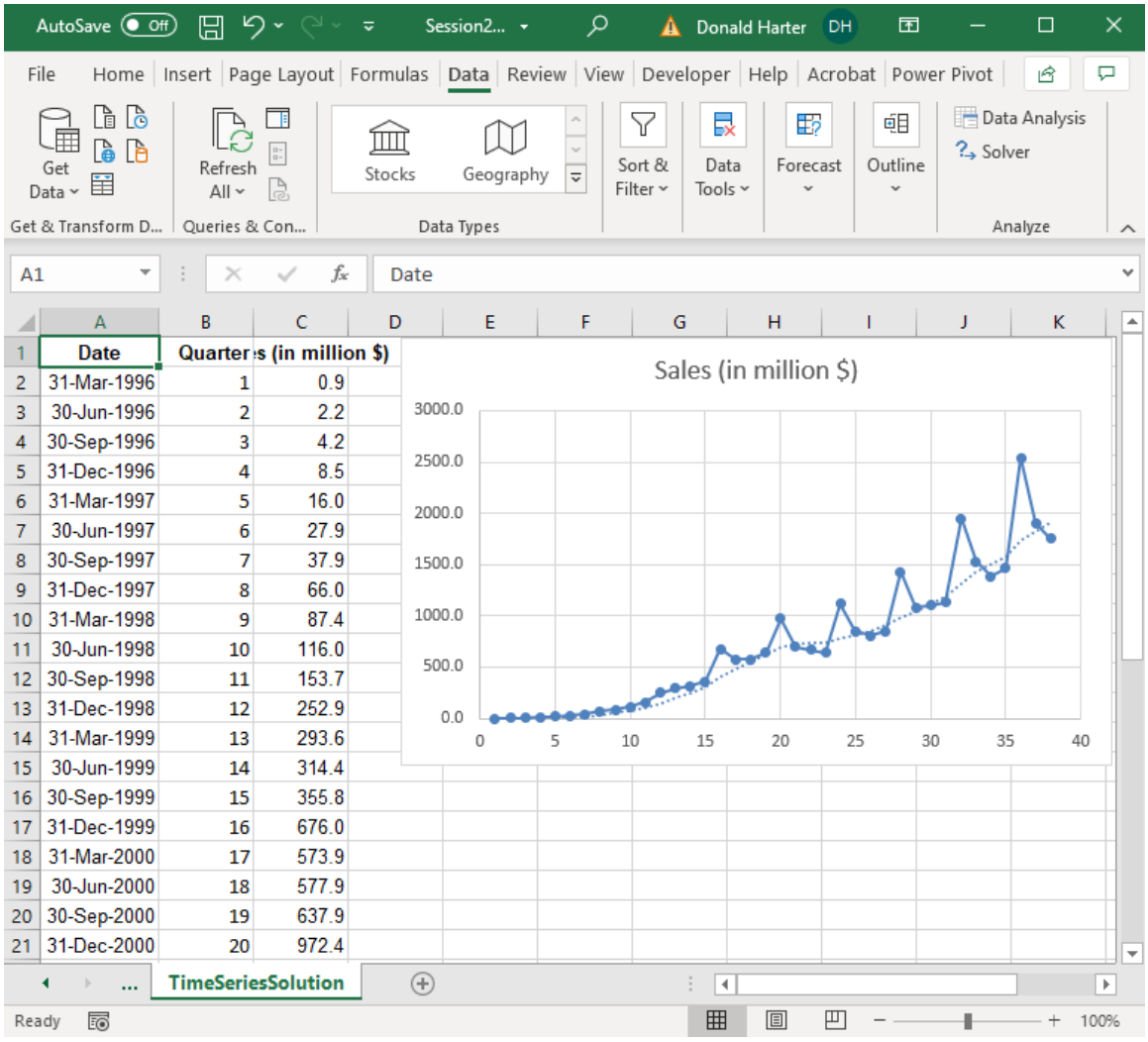


	A	B	C
1	Date	Quarters	Sales (in million $)
2	31-Mar-1996	1	0.9
3	30-Jun-1996	2	2.2
4	30-Sep-1996	3	4.2
5	31-Dec-1996	4	8.5
6	31-Mar-1997	5	16.0
7	30-Jun-1997	6	27.9
8	30-Sep-1997	7	37.9
9	31-Dec-1997	8	66.0
10	31-Mar-1998	9	87.4
11	30-Jun-1998	10	116.0
12	30-Sep-1998	11	153.7
13	31-Dec-1998	12	252.9
14	31-Mar-1999	13	293.6
15	30-Jun-1999	14	314.4
16	30-Sep-1999	15	355.8
17	31-Dec-1999	16	676.0
18	31-Mar-2000	17	573.9
19	30-Jun-2000	18	577.9
20	30-Sep-2000	19	637.9
21	31-Dec-2000	20	972.4

Notice that there is seasonality in the data. Amazon sales tend to peak during the fourth quarter of each year due to holiday sales. However, this seasonality masks the true trend. A moving average helps to see the trend.

To add a moving average line, follow these steps:

1. Right click on a data point.
2. Click on Add Trendline
3. Click the checkbox for moving average. Since we have quarterly data, let's identify the number of periods as four.



	A	B	C
1	Date	Quarters (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
12	30-Sep-1998	11	153.7
13	31-Dec-1998	12	252.9
14	31-Mar-1999	13	293.6
15	30-Jun-1999	14	314.4
16	30-Sep-1999	15	355.8
17	31-Dec-1999	16	676.0
18	31-Mar-2000	17	573.9
19	30-Jun-2000	18	577.9
20	30-Sep-2000	19	637.9
21	31-Dec-2000	20	972.4

The moving average line is superimposed on the graph. It's now very clear what the trend looks like when a moving average accounts for seasonality.