

**DATA 301**

**Introduction to Data Analytics  
Spreadsheets: Microsoft Excel**

Dr. Ramon Lawrence

University of British Columbia Okanagan

ramon.lawrence@ubc.ca

# Why Spreadsheets and Microsoft Excel?

---

**Spreadsheets** are the most common, general-purpose software for data analysis and reporting.

Microsoft Excel is the most popular spreadsheet program with hundreds of millions of installations.

- The spreadsheet concepts translate to other products.

Excel and spreadsheets are not always the best tool for data analysis, but they are great for quick analysis, reporting, and sharing.

# Spreadsheet Overview

---

A **spreadsheet** organizes information into a two-dimensional array of cells (a **table**).

A **cell** has two components:

- an address - specified given a column letter and row number
- a location - that can store a number, text, or formula

The power of a spreadsheet is that we can write simple formulas (commands) to perform calculations and immediately see the results of those calculations.

Spreadsheets are very common in business and reporting applications.

# Spreadsheet Addressing

A **cell** is identified by a column letter and row number.

The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'sales.xlsx'. The formula bar at the top displays the formula `=SUM(G2:G12)` in cell G13. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13							\$ 3,190.00						
14													

ROWS

COLUMNS

Cell G13

# Spreadsheet Addressing

---

The rows in a spreadsheet are numbered starting from 1.

The columns are represented by letters.

- A is column 1, B is column 2, ..., Z is column 26, AA is column 27, ...

A cell is identified by putting the column letter first then the row number.

- e.g. B3 is the 2nd column and the 3rd row.

Question: What column number is AD? How about BAD?

# Spreadsheet Data Entry

An entry is added to a cell by clicking on it and typing in the data.

- The data may be a number, text, date, etc. Type and *format* are auto-detected.

The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'sales.xlsx'. The ribbon is set to 'Home', and the 'Number' format is selected in the ribbon. A blue circle highlights the 'Number' dropdown menu, and a blue arrow points from it to the word 'format' in the spreadsheet area. The spreadsheet contains a table with columns for Category, Product, Month, Volume, Price, Cost, and Revenue. The data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13					Total:		\$ 3,190.00						

# Spreadsheet Formatting

Formatting: bold, italics, underline, fonts, colors

The screenshot shows the Microsoft Excel interface with the 'Home' tab selected. A blue circle highlights the 'Font' group in the ribbon, which includes options for font face (Calibri), size (11), bold (B), italic (I), underline (U), and text color. A blue arrow points from the text 'format shortcuts' to the 'Font' group. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13						<b>Total:</b>	<b>\$ 3,190.00</b>						

# Spreadsheet Selecting Cells

---

Multiple ways of selecting cells:

- 1) With the mouse, (left) click and drag mouse to select a rectangle region of cells.
- 2) With keyboard, hold `SHIFT` key and use arrow keys to select a rectangle region of cells.
- 3) With mouse and keyboard, while holding `CTRL` key, (left) click on individual cells to select non-contiguous cells.
- 4) Click on a row number to select a whole row.
- 5) Click on a column header to select a whole column.



# Range Selecting Cells Example

The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'sales.xlsx'. The 'Home' tab is active, and the ribbon shows various options like Font, Alignment, Number, Styles, Cells, and Editing. The active cell is A3, and the formula bar displays 'Clothing'. A range of cells from A3 to G6 is selected, highlighted in green. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<b>Category</b>	<b>Product</b>	<b>Month</b>	<b>Volume</b>	<b>Price</b>	<b>Cost</b>	<b>Revenue</b>						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13						<b>Total:</b>	<b>\$ 3,190.00</b>						

The status bar at the bottom shows 'Ready', 'Average: 115', 'Count: 28', 'Sum: 1840', and '100%' zoom level.

# Selecting Individual Cells Example

The screenshot shows the Microsoft Excel interface with the following details:

- File Name:** sales.xlsx - Excel
- Formula Bar:** G13 :  $=\text{SUM}(G2,G5,G8,G10,G12$
- Spreadsheet Data:**

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<b>Category</b>	<b>Product</b>	<b>Month</b>	<b>Volume</b>	<b>Price</b>	<b>Cost</b>	<b>Revenue</b>						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13					<b>Total:</b>		<b>G10,G12</b>						

# Manipulating Cells

---

Once you have selected one or more cells, there are several common actions you can perform:

## 1) DELETE

- delete the contents of all cells by pressing delete key
- delete the contents and the cell locations (then shift remaining) by selecting `Edit` menu, `Delete...` or `Delete...` from pop-up menu (brought up by right click).

## 2) Cut, Copy, Paste

- cut - copies selected cells to clipboard and removes from document
- copy - copies selected cells to clipboard
- paste - copies cells in clipboard to sheet starting at currently selected cell

3) Add selected cells to a formula (requires that you were previously constructing a formula before selecting the cells).

# Manipulating Cells - Filling

---

**Filling** combines copy and paste.

There is a small box or tab beyond the cell's lower right corner (fill handle). Grab it with the cursor and pull to other cells.

# Cut, Copy, Paste

cut,  
copy,  
paste

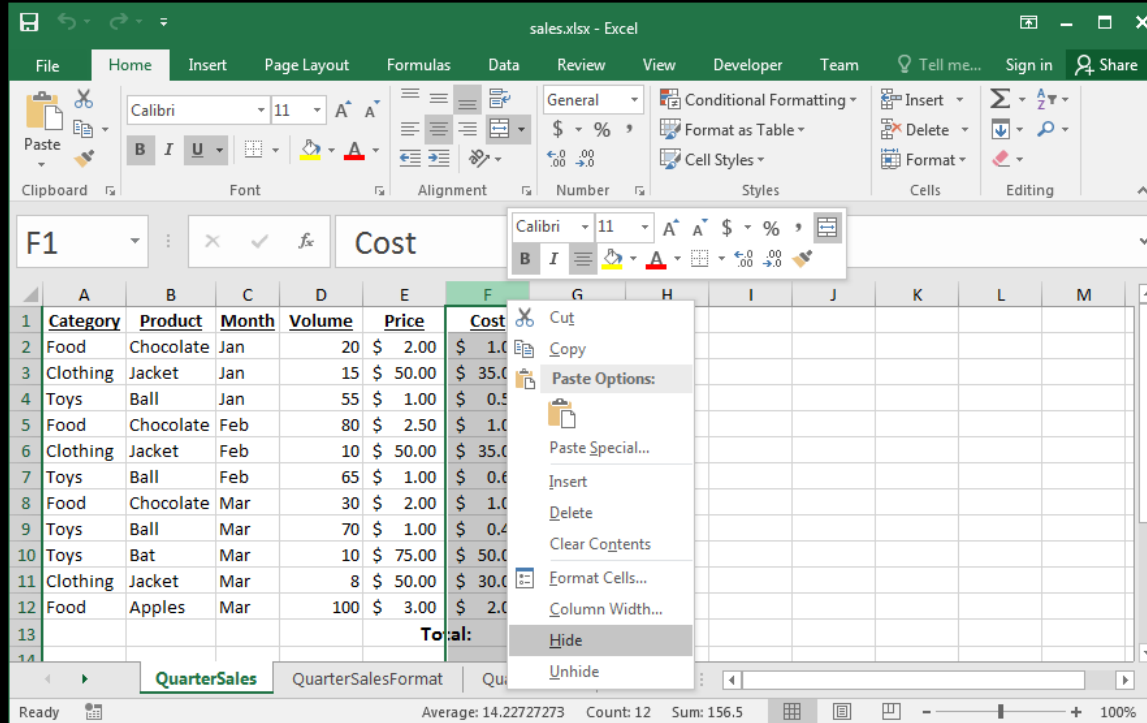
The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'sales.xlsx'. The ribbon is set to 'Home', and the 'Clipboard' group is visible. The active cell is D2, which contains the value 20. The spreadsheet contains a table with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13					<b>Total:</b>		<b>\$3,190.00</b>						

# Hiding Columns and Rows

Right-clicking on the column or row header and selecting **Hide**.

- The column/row still exists but will not be displayed or printed unless unhidden.



The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'sales.xlsx'. The 'Home' tab is active, and the 'Cost' column header (F1) is selected. A context menu is open over the 'Cost' column header, with the 'Hide' option highlighted. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost							
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.0							
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.0							
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.5							
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.0							
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.0							
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.6							
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.0							
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.4							
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.0							
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.0							
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.0							
13													
14													

The status bar at the bottom shows 'Ready', 'Average: 14.22727273', 'Count: 12', 'Sum: 156.5', and '100%' zoom level.

# Selecting Cells Question

---

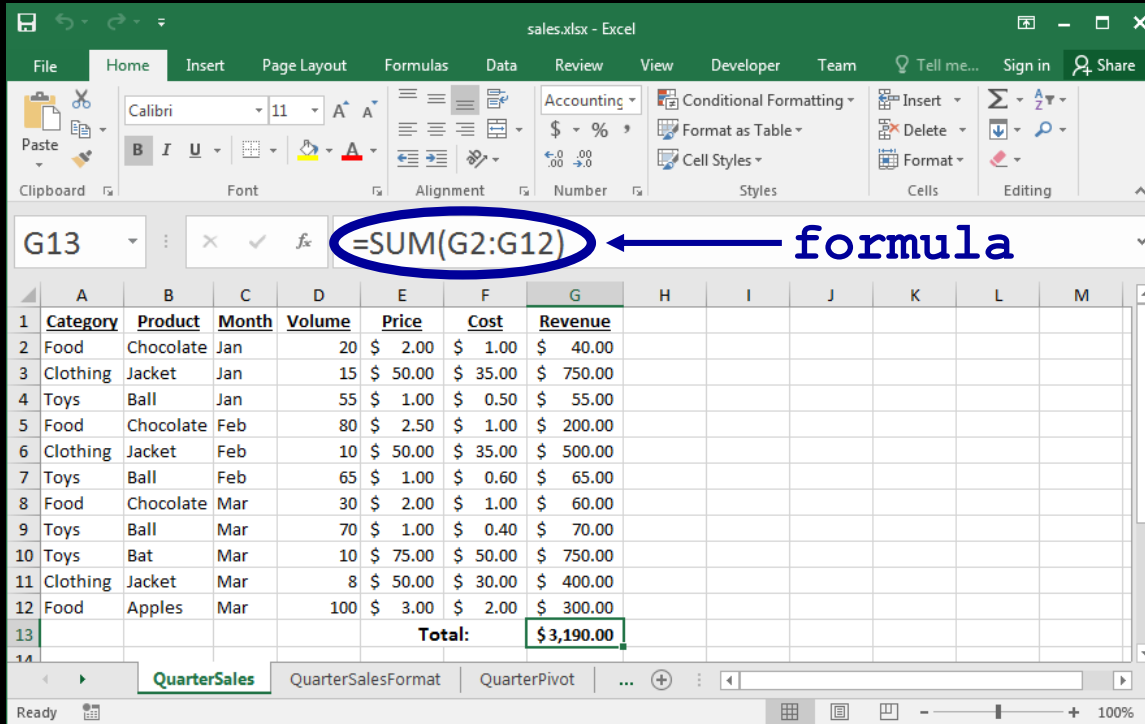
**Question:** Which method allows you to select non-contiguous cells in a spreadsheet?

- A)** hold **SHIFT** key and use arrow keys
- B)** With the mouse left click on a cell and drag mouse
- C)** hold **CTRL** key and use arrow keys
- D)** hold **CTRL** key and left click on cells

# Entering Formulas

A **formula** is any expression that begins with an equal sign ("=").

- The equal sign means that a calculation must be done to compute the cell value.



The screenshot shows the Excel interface with the formula bar containing the formula `=SUM(G2:G12)`. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13							<b>Total:</b>	\$ 3,190.00					



# Formula Expressions

---

A **formula** expression can consist of literals (numbers, text strings), operators, functions, and cell references.

Simple mathematical expressions:

- = 1 + 5
- = 1.5 \* 3.14 + 42

Common functions:

- = ROUND (PI ( ) , 2) // Result is 3.14
- = CONCATENATE ("Hello", " World") // Hello World
- Other common functions for trigonometry, dates, and financial.

# Formula Expressions

---

The power of formulas comes from using cell references (similar to variable names in programming).

Cell reference examples:

- = A1 + A2
- = B1 + A3 - A4

# Formulas Question

---

**Question:** A cell contains the following:  $=2+4*3$  What is the value of the cell?

A) 14

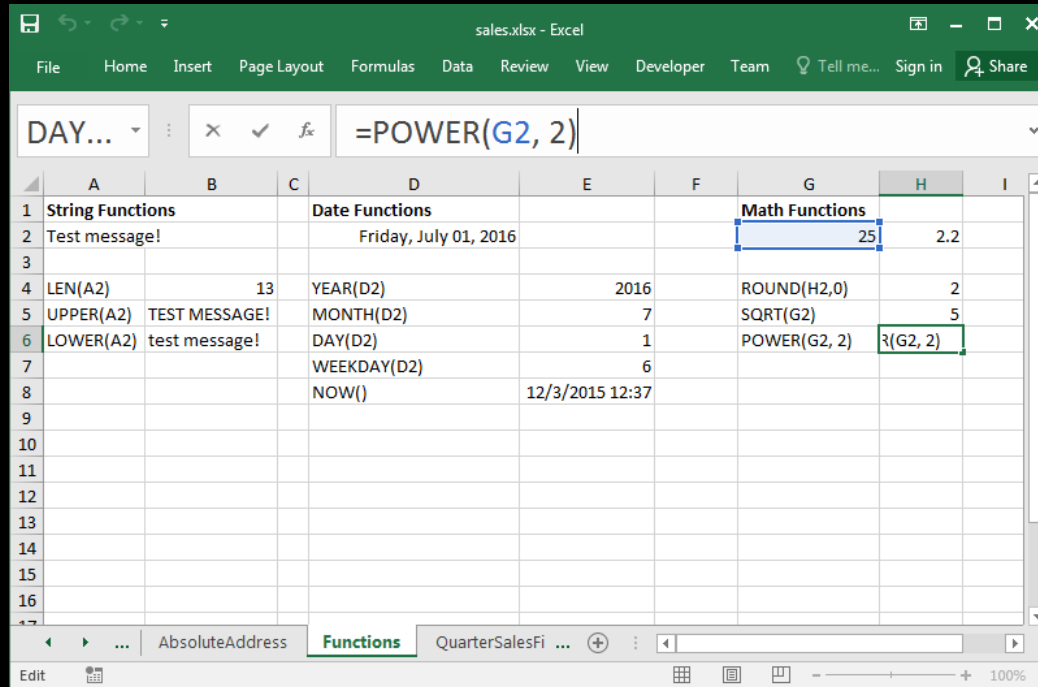
B) 18

C)  $=2+4*3$

# Using Excel Functions

Excel has a large number of built-in functions to use.

A **function** takes arguments as input and produces an output.



The screenshot shows the Microsoft Excel interface with a spreadsheet titled "sales.xlsx". The ribbon includes File, Home, Insert, Page Layout, Formulas, Data, Review, View, Developer, Team, Tell me..., Sign in, and Share. The formula bar displays the formula `=POWER(G2, 2)`. The spreadsheet contains the following data:

	A	B	C	D	E	F	G	H	I
1	String Functions			Date Functions			Math Functions		
2	Test message!			Friday, July 01, 2016			25	2.2	
3									
4	LEN(A2)	13	YEAR(D2)	2016			ROUND(H2,0)	2	
5	UPPER(A2)	TEST MESSAGE!	MONTH(D2)	7			SQRT(G2)	5	
6	LOWER(A2)	test message!	DAY(D2)	1			POWER(G2, 2)	<code>=POWER(G2, 2)</code>	
7			WEEKDAY(D2)	6					
8			NOW()	12/3/2015 12:37					
9									
10									
11									
12									
13									
14									
15									
16									
17									

The bottom status bar shows "Edit", "AbsoluteAddress", "Functions", "QuarterSalesFi ...", and "100%" zoom level.

# Concatenation

**String concatenation** is when two or more strings are combined by appending them in order. Function in Excel is `CONCATENATE ()` or `&`.

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I
1			More than						
2			one						
3	Hello World		string.						
4									
5			More than one string.						
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									

The formula bar shows the formula: `=CONCATENATE("Hello", " World")`. The result in cell A3 is "Hello World".

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I
1			More than						
2			one						
3	Hello World		string.						
4									
5			DATE(C1, C2, C3)						
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									

The formula bar shows the formula: `=CONCATENATE(C1, C2, C3)`. The result in cell C5 is "DATE(C1, C2, C3)".

# LOOKUP Function

The LOOKUP function searches for a value in a column.

- VLOOKUP searches a column in a table ; HLOOKUP searches a row in a table.

The screenshot shows the Microsoft Excel interface with the following data in the spreadsheet:

Product Id	Product Name	Product Price
1	Apple	\$ 3.99
2	Banana	\$ 2.99
3	Lettuce	\$ 1.99
4	Squash	\$ 6.99
5	Pumpkin	\$ 5.99

The formula bar displays: `=LOOKUP(F2, A2:A6, C2:C6)`

A tooltip for the LOOKUP function is visible, showing the syntax: `LOOKUP(lookup_value, lookup_vector, [result_vector])` and `LOOKUP(lookup_value, array)`. The spreadsheet also shows a table with columns for Product Id, Price (LOOKUP), and Price (VLOOKUP), with values corresponding to the data above.

# INDEX Function

INDEX () returns the value in the array of cells at the given index.

The screenshot shows the Microsoft Excel interface with the following details:

- Formula Bar:** Contains the formula `=INDEX(B2:B6,F2+1)`.
- Spreadsheet Data:**

Product Id	Product Name	Product Price
1	Apple	\$ 3.99
2	Banana	\$ 2.99
3	Lettuce	\$ 1.99
4	Squash	\$ 6.99
5	Pumpkin	\$ 5.99
- Lookup Results:**
  - Product Id: 2
  - Price (LOOKUP): \$ 2.99
  - Price (VLOOKUP): \$ 2.99
  - Product (INDEX): B6,F2+1
- Help Text:**
  - INDEX(array, row\_num, [column\_num])
  - INDEX(reference, row\_num, [column\_num], [area\_num])
- Task Pane:** The 'Lookup' task pane is active, showing the current formula and its components.

# Formulas Question

---

**Question:** A cell contains the following: **'ABC'+'DEF'**. What is the value of the cell?

**A)** error

**B)** ABCDEF

**C)** 'ABC'+'DEF'



# Formulas Question

---

**Question:** How many of the following statements are **TRUE**?

- 1) CONCATENATE function can take 3 arguments.
- 2) There is an Excel function that has 0 arguments.
- 3) =INDEX ( { 1 , 3 , 5 } , 2 ) returns 5.
- 4) =LOOKUP ( 5 , { 1 , 3 , 5 } , { "a" , "b" , "c" } ) returns "c".

**A) 0**                      **B) 1**                      **C) 2**                      **D) 3**                      **E) 4**

# Try it: Entering Formulas

**Question:** Add a column for expenses and profit as below:

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue	Expenses	Profit				
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00	\$ 20.00	\$ 20.00				
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00	\$ 525.00	\$ 225.00				
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00	\$ 27.50	\$ 27.50				
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00	\$ 80.00	\$ 120.00				
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00	\$ 350.00	\$ 150.00				
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00	\$ 39.00	\$ 26.00				
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00	\$ 30.00	\$ 30.00				
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00	\$ 28.00	\$ 42.00				
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00	\$ 500.00	\$ 250.00				
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00	\$ 240.00	\$ 160.00				
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00	\$ 200.00	\$ 100.00				
13							Total:	\$3,190.00					

# ★ Advanced Spreadsheet Addressing

The dollar sign "\$" is a symbol that indicates an **absolute address**.

- By default, addresses are "relative" in the sense that if they are in a formula that is copied to another cell, they will be changed relative to where they were copied from their origin.

Example:

- Cell A1 has the formula  $=A2+B1$
- Copy contents of cell A1 to cell C4.
- Formula changes to  $=C5+D4$  because moved down three rows and over two columns.
- If cell A1 had the formula  $=\$A\$2+\$B\$1$ , then the same formula would be in cell C4.
- Question: What if formula was  $=\$A2+B\$1$ ?

# Formulas and References Question

---

**Question:** Cell **A1** contains the following: **=\$B2+D\$4**. What is the formula if the cell is copied to cell **D3**?

**A)** error

**B)** =\$B2+D\$4

**C)** =\$B4+F\$4

**D)** =\$B4+G\$4



# Aggregate Functions

---

An **aggregate function** computes a summary function over a range of cells. The values can either be data values or cell locations.

Common functions are:

- `MIN(<value list>)` - returns minimum value in list
- `MAX(<value list>)` - returns maximum value in list
- `SUM(<value list>)` - returns sum of all values in list
- `AVERAGE(<value list>)` - returns average of values in list
- `COUNT(<value list>)` - returns count of values in list
- `MEDIAN(<value list>)` - returns median value of list

If specifying a cell rectangle, give the upper left and lower right corners, separated by a colon.

- e.g. `=AVERAGE(A3:E6)` - rectangle of 4 rows and 5 columns

# Aggregate Functions Example

The screenshot shows the Microsoft Excel interface with the following data table:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue	Expenses	Profit				
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00	\$ 20.00	\$ 20.00				
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00	\$ 525.00	\$ 225.00				
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00	\$ 27.50	\$ 27.50				
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00	\$ 80.00	\$ 120.00				
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00	\$ 350.00	\$ 150.00				
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00	\$ 39.00	\$ 26.00				
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00	\$ 30.00	\$ 30.00				
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00	\$ 28.00	\$ 42.00				
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00	\$ 500.00	\$ 250.00				
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00	\$ 240.00	\$ 160.00				
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00	\$ 200.00	\$ 100.00				
13				=max(D2:D12)	tal:		\$ 3,190.00						
14				MAX(number1, [number2], ...)									

The formula bar shows the formula: `=max(D2:D12)`

The status bar at the bottom shows: Point, QuarterSales, QuarterSalesFormat, QuarterPivot, 100%



# Aggregate Functions Question

---

**Question:** Assume the cells in the range A1 : C4 each contain a number that is equal to their row number (e.g. B3 contains 3). How many of the following statements are **TRUE**?

- 1) The number of cells in the range is 12.
- 2) The value of `SUM (A1 : C4)` is 20.
- 3) The value of `COUNTIF (A1 : B4, ">2")` is 4.
- 4) `AVERAGE (A1 : C4) > MAX (C2 : C3)`

**A) 0**                      **B) 1**                      **C) 2**                      **D) 3**                      **E) 4**



# Aggregate Functions Question

---

**Question:** Assume the three cells in the range A1 : C1 contain numbers. Which of these formula output results is **ALWAYS** the largest?

**A)** MAX (A1 : C1)

**B)** MIN (A1 : C1)

**C)** COUNT (A1 : C1)

**D)** SUM (A1 : C1)

**E)** none of the above are always guaranteed to be the largest

# Other Formatting: Column Width

The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'QuarterSales'. The spreadsheet contains the following data:

Category	Product	Month	Volume	Price	Cost	Revenue
Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00
Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00
Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00
Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00
Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00
Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00
Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00
Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00
Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00
Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00
Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00
<b>Total:</b>						<b>\$3,190.00</b>

The context menu is open, showing options for Cell Size, Visibility, and Protection. The 'Column Width...' option is highlighted.

Resizing columns/rows:  
Auto-resize by double clicking on border between columns or using the Format option.

Drag row/column border for manual resize.

# Conditional Formatting

**Conditional formatting** allows you to change the cell format based on data values. This is accessible under **Styles**.

- Other options: data bars, color scales

The screenshot displays the Microsoft Excel interface with a spreadsheet titled 'sales.xlsx'. The spreadsheet contains the following data:

	A	B	C	D	E	F	G
1	<b>Category</b>	<b>Product</b>	<b>Month</b>	<b>Volume</b>	<b>Price</b>	<b>Cost</b>	<b>Revenue</b>
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00
3	Clothing	Jacket	Jan	15	\$50.00	\$35.00	\$ 750.00
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00
6	Clothing	Jacket	Feb	10	\$50.00	\$35.00	\$ 500.00
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00
10	Toys	Bat	Mar	10	\$ 75.00	\$50.00	\$ 750.00
11	Clothing	Jacket	Mar	8	\$50.00	\$30.00	\$ 400.00
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00
13				<b>Total:</b>			<b>\$3,190.00</b>

The 'Conditional Formatting Rules Manager' dialog box is open, showing the following rules:

Rule (applied in order shown)	Format	Applies to	Stop If True
Cell Value >= 350	AaBbCcYvZ	=\$G\$2:\$G\$12	<input type="checkbox"/>
Cell Value < 75	AaBbCcYvZ	=\$G\$2:\$G\$12	<input type="checkbox"/>

# Conditional Formatting Result

The format painter button allows you to copy formatting to many cells. Select the cell, click paint button, then highlight cells to have identical formatting.

format  
painter  
button

The screenshot shows the Microsoft Excel interface with the ribbon set to the HOME tab. The Format Painter button, represented by a paintbrush icon, is circled in green in the Clipboard group. A tooltip for the Format Painter button is displayed, providing instructions on how to use it. The spreadsheet below shows a table with columns for Volume, Price, Cost, and Revenue. The Revenue column is highlighted in green, indicating that the formatting from the selected cell (G2) has been applied to the other cells in that column.

**Format Painter**  
Like the look of a particular selection? You can apply that look to other content in the document.

To get started:

1. Select content with the formatting you like
2. Click Format Painter
3. Select something else to automatically apply the formatting

FYI: To apply the formatting in multiple places, double-click Format Painter.

[Tell me more](#)

	D	E	F	G	H	I	J	K	L	M
	Volume	Price	Cost	Revenue						
1	20	\$ 2.00	\$ 1.00	\$ 40.00						
2	15	\$50.00	\$35.00	\$ 750.00						
3	55	\$ 1.00	\$ 0.50	\$ 55.00						
4	80	\$ 2.50	\$ 1.00	\$ 200.00						
5	10	\$50.00	\$35.00	\$ 500.00						
6	65	\$ 1.00	\$ 0.60	\$ 65.00						
7	30	\$ 2.00	\$ 1.00	\$ 60.00						
8	70	\$ 1.00	\$ 0.40	\$ 70.00						
9	10	\$75.00	\$50.00	\$ 750.00						
10	8	\$50.00	\$30.00	\$ 400.00						
11	100	\$ 3.00	\$ 2.00	\$ 300.00						
12	<b>Total:</b>			<b>\$3,190.00</b>						

# Try it: Conditional Formatting

**Question:** Format rows so: 1) bold/green if volume > 50, 2) italics/red if volume < 10, 3) yellow background otherwise as below:

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<b>Category</b>	<b>Product</b>	<b>Month</b>	<b>Volume</b>	<b>Price</b>	<b>Cost</b>	<b>Revenue</b>						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$50.00	\$35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$50.00	\$35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$75.00	\$50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$50.00	\$30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13					<b>Total:</b>		<b>\$3,190.00</b>						

# Try it: Conditional Formatting Challenge

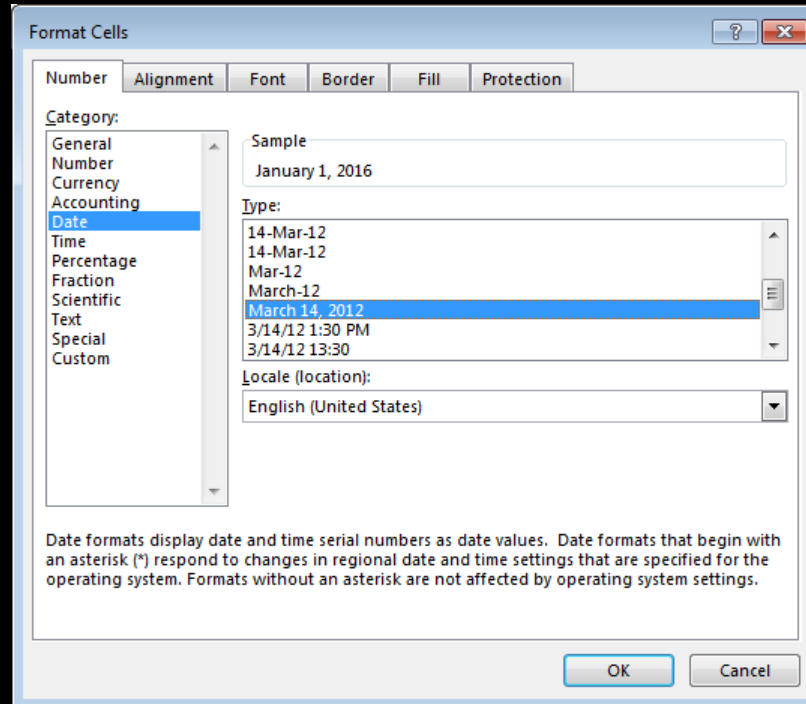
**Question:** Take the previous formatting and apply it to whole row:

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<b>Category</b>	<b>Product</b>	<b>Month</b>	<b>Volume</b>	<b>Price</b>	<b>Cost</b>	<b>Revenue</b>						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$50.00	\$35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$50.00	\$35.00	\$ 500.00						
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
10	Toys	Bat	Mar	10	\$75.00	\$50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$50.00	\$30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13					<b>Total:</b>		<b>\$3,190.00</b>						

# Date and Type Formats

Formatting data helps users read and understand data and is especially important for numbers and dates. Use built-in or custom formats.



# Spreadsheets for Data Management

---

A spreadsheet is often used as a "database". A database is an organized representation of information.

- Examples: schedules and calendars, timesheets, expenses and finances, records, notes, and recipes, data research/analysis

We can use a spreadsheet as a database by:

- Using a row to store all the information about something we want to represent.
- Giving each column a meaningful name. A column represents a property or feature of the object stored in the row.
- Using the formulas to calculate new facts from the data.
- Using sorting to organize the data by key features.
- Using simple filtering (querying) to only show the most important data or data of interest.



# Sorting Data

Data can be sorted by selecting the **Sort** option under the **Data** menu. Select the column(s) to sort on and order to sort by.

The screenshot shows the Microsoft Excel interface with the **Data** tab selected. The **Sort** dialog box is open, showing the following configuration:

- Column:** Month
- Sort On:** Values
- Order:** Jan, Feb, Mar, Apr, May, Jun, Jul
- Then by:** Category
- Sort On:** Values
- Order:** A to Z
- My data has headers

The spreadsheet data is as follows:

	A	B	C	D	E	F	G
1	Category	Product	Month	Volume	Price	Cost	Revenue
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00
13				<b>Total:</b>			<b>\$ 3,190.00</b>

# Try it: Sort

**Question:** Sort the data by revenue (desc) then product (asc).

The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The ribbon includes options for 'Sort', 'Filter', 'Advanced', 'Text to Columns', 'What-If Analysis', and 'Forecast Sheet'. The active cell is H13. The data table is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<b>Category</b>	<b>Product</b>	<b>Month</b>	<b>Volume</b>	<b>Price</b>	<b>Cost</b>	<b>Revenue</b>						
2	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
5	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
6	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
7	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
8	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00						
9	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00						
10	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
11	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
12	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
13							<b>Total:</b>	\$ 3,190.00					

# Filtering

---

A **filter** shows a subset of the rows in the spreadsheet that pass a given condition (test).

Select **Auto Filter** under the **Data** then **Filter** menu.

Once you select `Auto Filter`, each column heading has a drop-down list. By selecting a filtering criteria from the list, you can limit the rows that are displayed.

It is possible to filter on more than one column at the same time.

# Filter Example

Filter on Revenue column: Select value(s), Top 10, or custom filter.

The screenshot shows the Microsoft Excel interface with a data table. The table has columns: Category, Product, Month, Volume, Price, Cost, and Revenue. The 'Revenue' column is selected, and a filter menu is open. The menu shows options for sorting and filtering. The 'Number Filters' section is expanded, showing a list of values to filter by. The 'Top 10...' option is highlighted.

	A	B	C	D	E	F	G	H	I	J	K
1	Category	Product	Month	Volume	Price	Cost	Revenue				
5	Food	Chocolate	Feb								
6	Clothing	Jacket	Feb								
12	Food	Apples	Mar								
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											

Number Filters

Search

- (Select All)
- \$40.00
- \$55.00
- \$60.00
- \$65.00
- \$70.00
- \$200.00
- \$300.00
- \$400.00
- \$500.00
- \$750.00

OK Cancel

- Equals...
- Does Not Equal...
- Greater Than...
- Greater Than Or Equal To...
- Less Than...
- Less Than Or Equal To...
- Between...
- Top 10...**
- Above Average
- Below Average
- Custom Filter...

# Custom Filter Example

Filter on Revenue column: Custom filter with **Revenue > 150**

Custom AutoFilter

Show rows where:  
Revenue

is greater than 150

And  Or

Use ? to represent any single character  
Use \* to represent any series of characters

OK Cancel

# Custom Filter Result

Filter on Revenue: Custom filter result with **Revenue > 150**

The screenshot shows the Microsoft Excel interface with the 'Data' ribbon selected. The 'Filter' button is highlighted in the 'Sort & Filter' group. The data table below shows the results of a custom filter applied to the 'Revenue' column, showing only records where Revenue is greater than 150. The status bar at the bottom indicates 'Ready 7 of 12 records found'.

	A	B	C	D	E	F	G	H	I	J	K
1	Category	Product	Month	Volume	Price	Cost	Revenue				
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00				
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00				
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00				
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00				
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00				
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00				
13						<b>Total:</b>	<b>\$ 3,190.00</b>				
14											
15											
16											
17											
18											
19											

Ready 7 of 12 records found

# Try it: Filter

**Question:** Filter the data so only products with volume < 50 and revenue < \$100 are shown.

The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The 'Filter' button is highlighted in the ribbon. The data table is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Catego	Produc	Mon	Volum	Price	Cost	Revenu						
10	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
12	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													

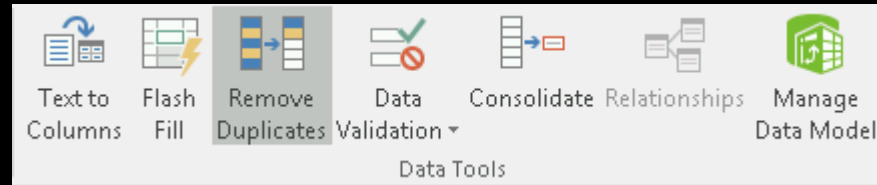
The status bar at the bottom indicates 'Ready 2 of 12 records found'.





# Removing Duplicates

To remove duplicates, select your Data then Remove Duplicates.



# Sorting Question

**Question:** Given this spreadsheet and sort order, what is the output?

Num	Char
1	A
1	a
1	B
1	b
2	A
2	b
3	a
3	B

Column	Sort On	Order
Sort by	Num	Largest to Smallest
Then by	Char	A to Z

**A)**

Num	Char
3	B
3	a
2	A
2	b
1	A
1	B
1	a
1	b

**B)**

Num	Char
3	a
3	B
2	A
2	b
1	a
1	A
1	b
1	B

**C)**

Num	Char
3	a
3	B
2	A
2	b
1	A
1	a
1	B
1	b

# Filtering Question

**Question:** Given this spreadsheet, how many of these statements are **TRUE**?

	A	B
1	Number	Letter
2	1	a
3	2	b
4	3	c
5	4	d
6	5	e
7		

- 1) The data is sorted ascending by Number.
- 2) Filter `Number > 3` shows 3 rows.
- 3) Filter `Letter >= "c"` shows 3 rows.
- 4) Filter `Number < 3 OR Letter > "b"` shows 5 rows.

**A) 0**                      **B) 1**                      **C) 2**                      **D) 3**                      **E) 4**

# Charts

---

A *chart* is a graphical representation of spreadsheet data.

A chart is of a particular type (line, bar, etc.) and requires the user to supply the data that will be displayed in the chart.

# Chart: Select Data and Type

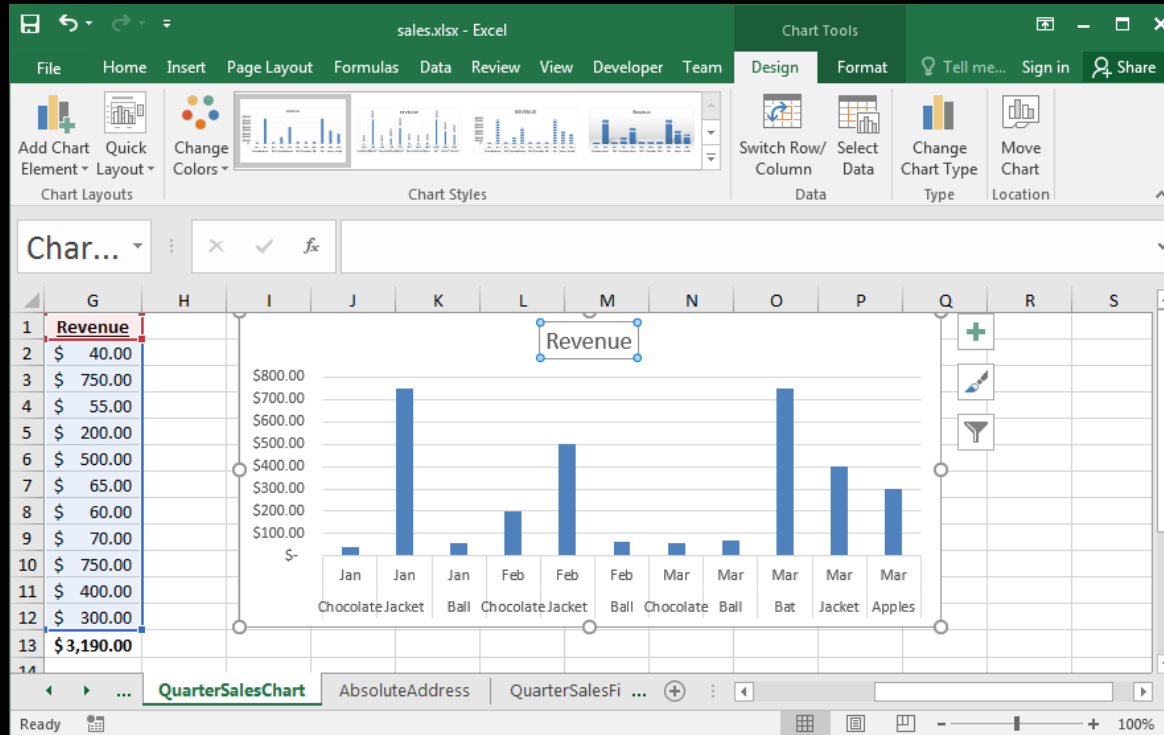
Select **Insert**, then click **Chart** Icon, and pick the chart type.

The screenshot shows the Microsoft Excel interface with the **Insert** tab selected. The **Charts** group is visible, and the **Insert Chart** dialog box is open. The dialog box shows the **Recommended Charts** tab, with **Clustered Column** selected. Two preview charts are shown: **Revenue** and **Revenue**.

	A	B	C	D	E	F	G	H	I
1	<b>Category</b>	<b>Product</b>	<b>Month</b>	<b>Volume</b>	<b>Price</b>	<b>Cost</b>	<b>Revenue</b>		
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00		
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00		
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00		
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00		
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00		
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00		
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00		
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00		
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00		
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00		
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00		
13						<b>Total:</b>	<b>\$ 3,190.00</b>		

# Chart Options

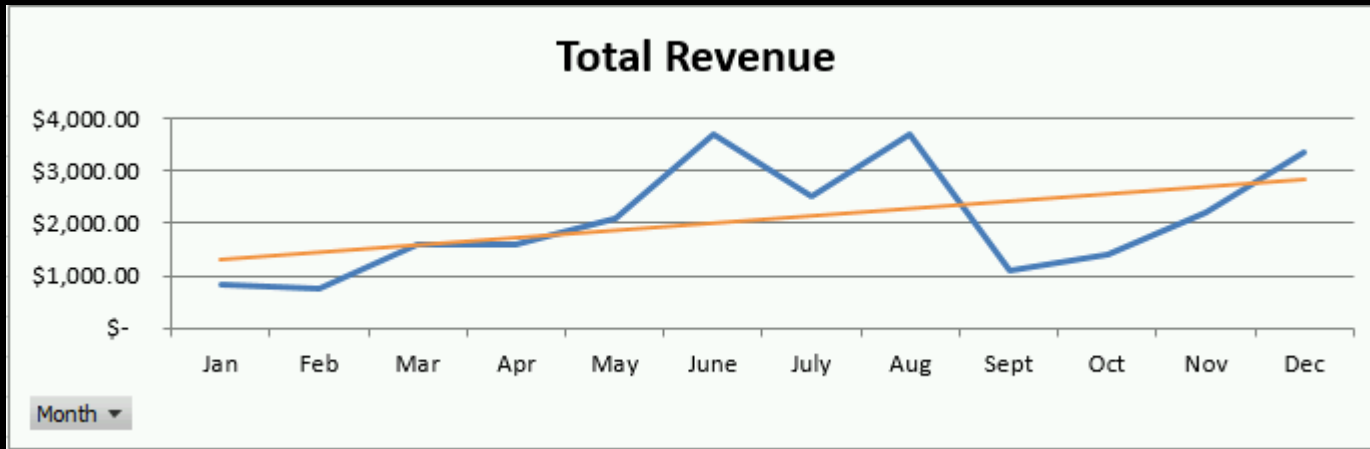
**Chart Tools** allows you to modify the data in the chart, change the chart type, and move the chart in the Worksheet.



# Trendlines




Trendlines can be easily added to any chart.

- Linear trendline for monthly revenue. Good choice?



Format Trendline ✕

Trendline Options ▾

**Trendline Options**

Exponential

Linear

Logarithmic

Polynomial Order

Power

Moving Average Period

Trendline Name

Automatic Linear (Revenue)

Custom

Forecast

Forward  period

Backward  period

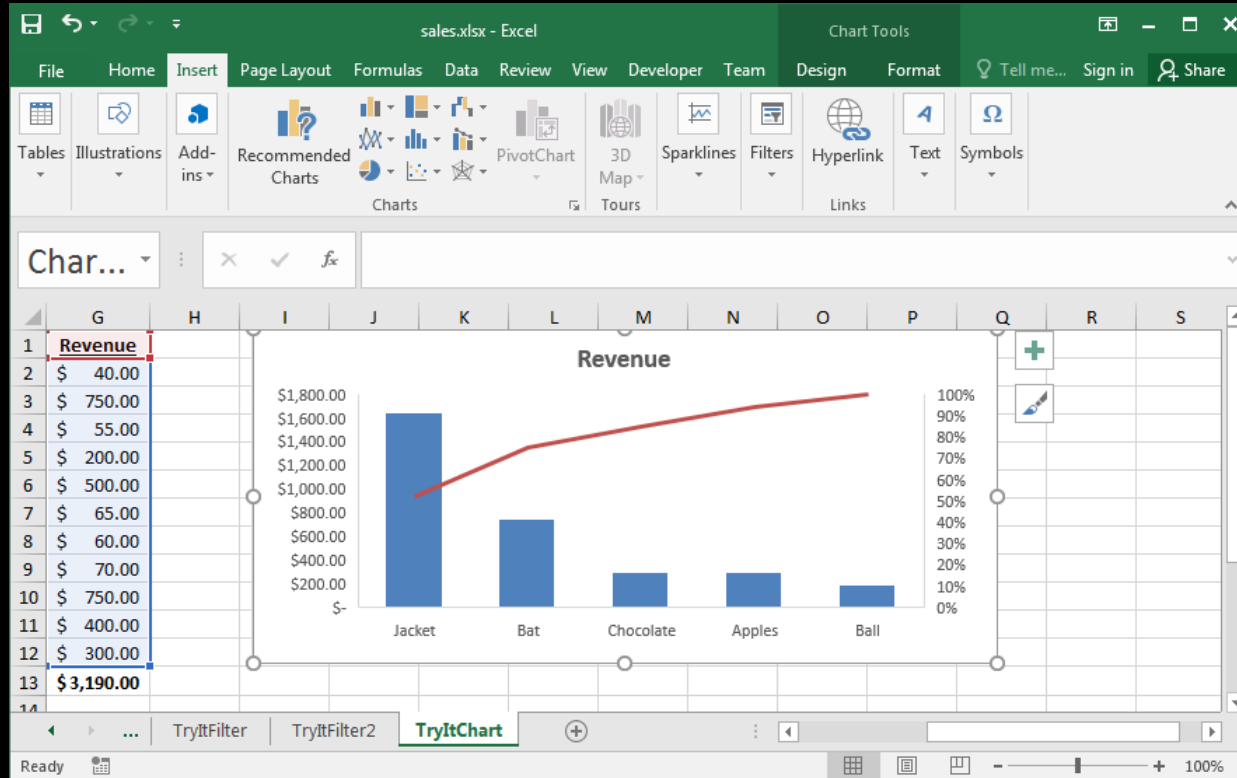
Set Intercept

Display Equation on chart

Display R-squared value on chart

# Try it: Chart

**Question:** Create a chart that makes it easy to see the best selling product.





# Sparklines

A **sparkline** is a tiny chart in a worksheet cell for a quick data overview.

- Insert then select a Sparkline (line, column, win/loss). May put text in sparkline cell.

The screenshot shows the Microsoft Excel interface with the Sparkline Tools ribbon selected. The ribbon includes options for Sparkline Type (Line, Column, Win/Loss), Show (High Point, Low Point, Negative Points, First Point, Last Point, Markers), Style (Line styles and colors), and Group (Axis, Group, Ungroup, Clear).

The worksheet displays a table titled "Overall Revenue by Month" with the following data:

	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Grand Total
Sum of Revenue	\$ 845.00	\$ 765.00	\$ 1,580.00	\$ 1,610.00	\$ 2,080.00	\$ 3,690.00	\$ 2,530.00	\$ 3,690.00	\$ 1,080.00	\$ 1,420.00	\$ 2,200.00	\$ 3,350.00	\$ 24,840.00

A sparkline chart is inserted in cell O5, labeled "Monthly Sales". The chart shows a line graph representing the monthly revenue data, with a dashed green line indicating a trend or target.

# What-If

---

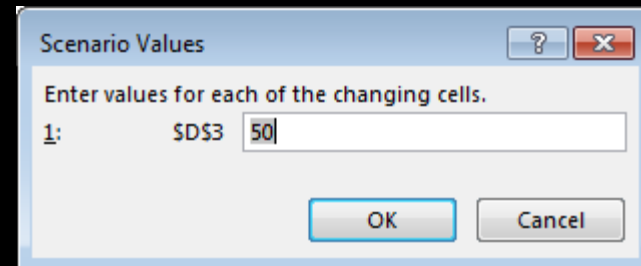
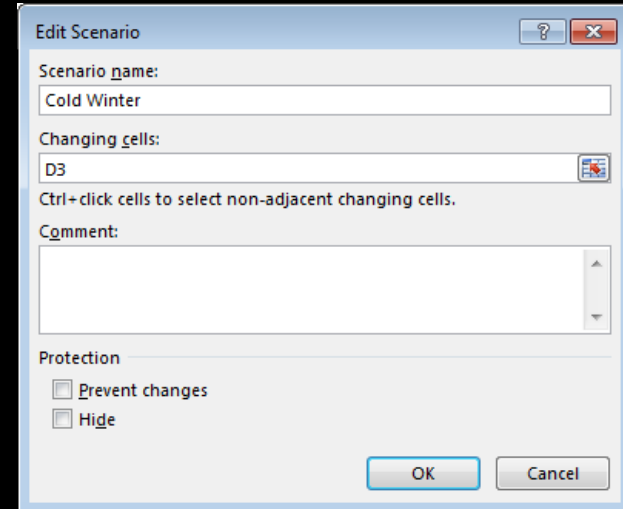
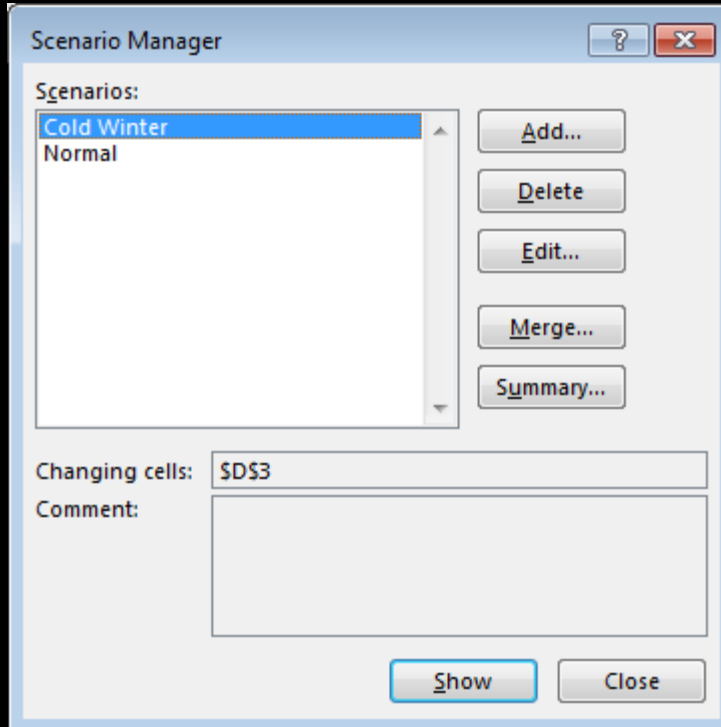
*What-If scenarios* help understand different possibilities.

A what-if scenario is created under Data then What-If Analysis then Scenario Manager.

To define a scenario, give it a name and list the cells that will change with this scenario.

# What-If Scenarios Example

Consider what happens with a cold winter and we predict to sell **50** jackets instead of the normal **15**.





# Try it: What-If Scenario

**Question:** Create a what-if scenario that wherever balls are sold, the volume is double than normal.

The screenshot shows an Excel spreadsheet titled 'sales.xlsx' with the 'Data' tab selected. The ribbon includes 'Sort & Filter', 'Data Tools', and 'Forecast'. The 'What-If Analysis' button is highlighted. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<b>Category</b>	<b>Product</b>	<b>Month</b>	<b>Volume</b>	<b>Price</b>	<b>Cost</b>	<b>Revenue</b>						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
4	Toys	Ball	Jan	110	\$ 1.00	\$ 0.50	\$ 110.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
7	Toys	Ball	Feb	130	\$ 1.00	\$ 0.60	\$ 130.00						
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
9	Toys	Ball	Mar	140	\$ 1.00	\$ 0.40	\$ 140.00						
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00						
13							<b>Total:</b>	<b>\$3,380.00</b>					

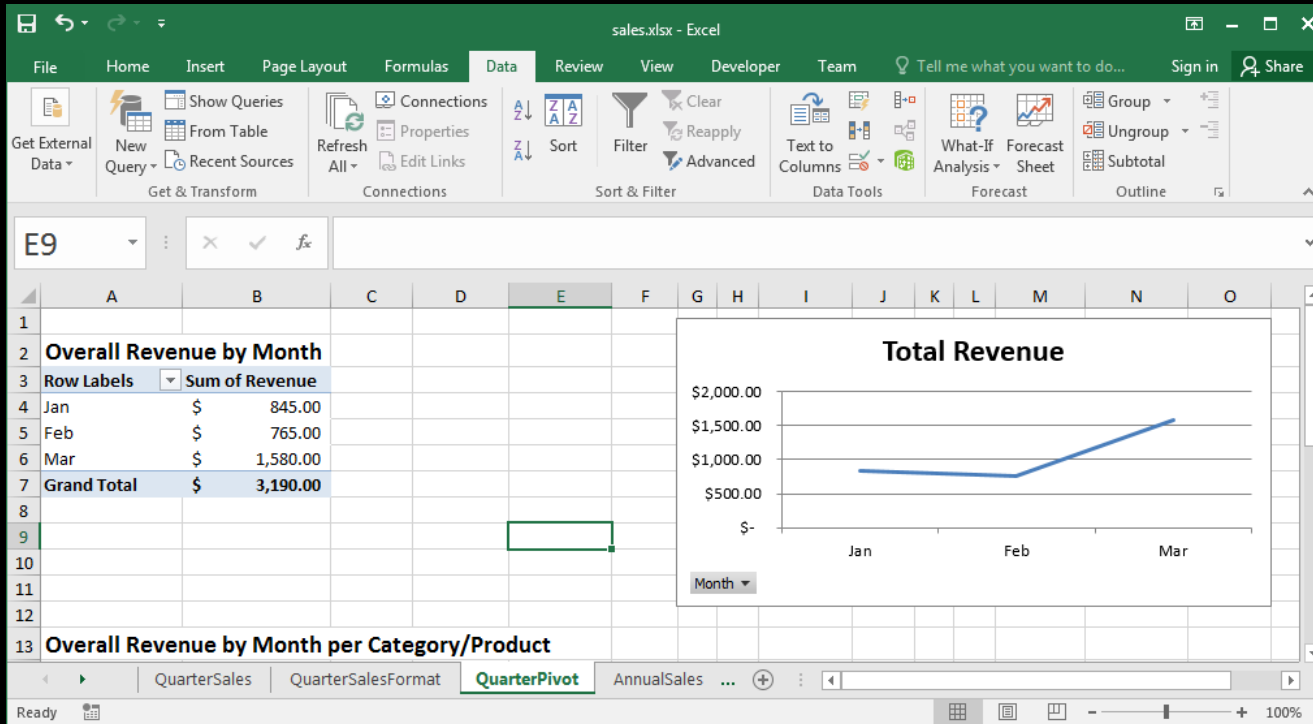
The 'Volume' column for 'Toys Ball' in row 4 is highlighted with a green border, indicating it is the target cell for the What-If Scenario. The status bar at the bottom shows 'Ready' and a zoom level of 100%.



# Pivot Tables

**Pivot tables** allow for easily aggregating and exploring large data sets.

- For example, our data set can be summarized by revenue by month.



# Creating a Pivot Table

To create, select the data and then Insert, Pivot Table.

The screenshot shows an Excel spreadsheet with the following data:

Category	Product	Month	Volume	Price	Cost	Revenue
Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00
Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00
Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00
Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00
Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00
Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00
Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00
Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00
Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00
Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00
Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00
<b>Total:</b>						<b>\$3,190.00</b>

The 'Create PivotTable' dialog box is open, showing the following options:

- Choose the data that you want to analyze:**
  - Select a table or range**: Table/Range: `QuarterSales!$A$1:$G$12`
  - Use an external data source: Choose Connection...
  - Use this workbook's Data Model
- Choose where you want the PivotTable report to be placed:**
  - New Worksheet**
  - Existing Worksheet: Location: [ ]
- Choose whether you want to analyze multiple tables:**
  - Add this data to the Data Model

Buttons: OK, Cancel



# Creating a Pivot Table

Add fields to pivot table.

Field may either be:

- Row value
- Column value
- Cell value (aggregated)
- Used in a filter

The screenshot displays the Microsoft Excel interface with a PivotTable and the PivotTable Fields task pane. The PivotTable is located in the range A3:D7 and shows the following data:

Row Labels	Sum of Revenue
Jan	845
Feb	765
Mar	1580
<b>Grand Total</b>	<b>3190</b>

The PivotTable Fields task pane on the right shows the following configuration:

- Choose fields to add to report:** Search,  Month,  Volume,  Price,  Cost,  Revenue
- Drag fields between areas below:**
- FILTERS:** (Empty)
- COLUMNS:** (Empty)
- ROWS:** Month
- VALUES:** Sum of Revenue
- Defer Layout Update
- UPDATE** button

The status bar at the bottom indicates "Ready" and "100%" zoom.

# Creating a Pivot Table Example

Products are rows.

Months are columns.

Each cell is a sum of revenue per product for that month.

Filter on product.

The screenshot shows an Excel spreadsheet with a PivotTable and the PivotTable Fields task pane. The PivotTable is titled "Sum of Revenue" and is located in the range A3:E9. The task pane shows the following configuration:

- PivotTable Fields:**
  - Choose fields to add to report:
    - Category
    - Product
    - Month
    - Volume
    - Price
- Drag fields between areas below:**
  - FILTERS:** (Empty)
  - COLUMNS:** Month
  - ROWS:** Product
  - VALUES:** Sum of Reven...
- Buttons:** Defer Layout Update (unchecked), UPDATE

The PivotTable data is as follows:

Row Labels	Jan	Feb	Mar	Grand Total
Ball	\$ 55.00	\$ 65.00	\$ 70.00	\$ 190.00
Bat			\$ 750.00	\$ 750.00
Chocolate	\$ 40.00	\$ 200.00	\$ 60.00	\$ 300.00
Jacket	\$ 750.00	\$ 500.00	\$ 400.00	\$ 1,650.00
<b>Grand Total</b>	<b>\$ 845.00</b>	<b>\$ 765.00</b>	<b>\$ 1,280.00</b>	<b>\$ 2,890.00</b>

# Try it: Pivot Table

**Question:** Create a pivot table using the annual sales data that shows revenue per month by category/product.

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me what you want to do... Sign in Share

12

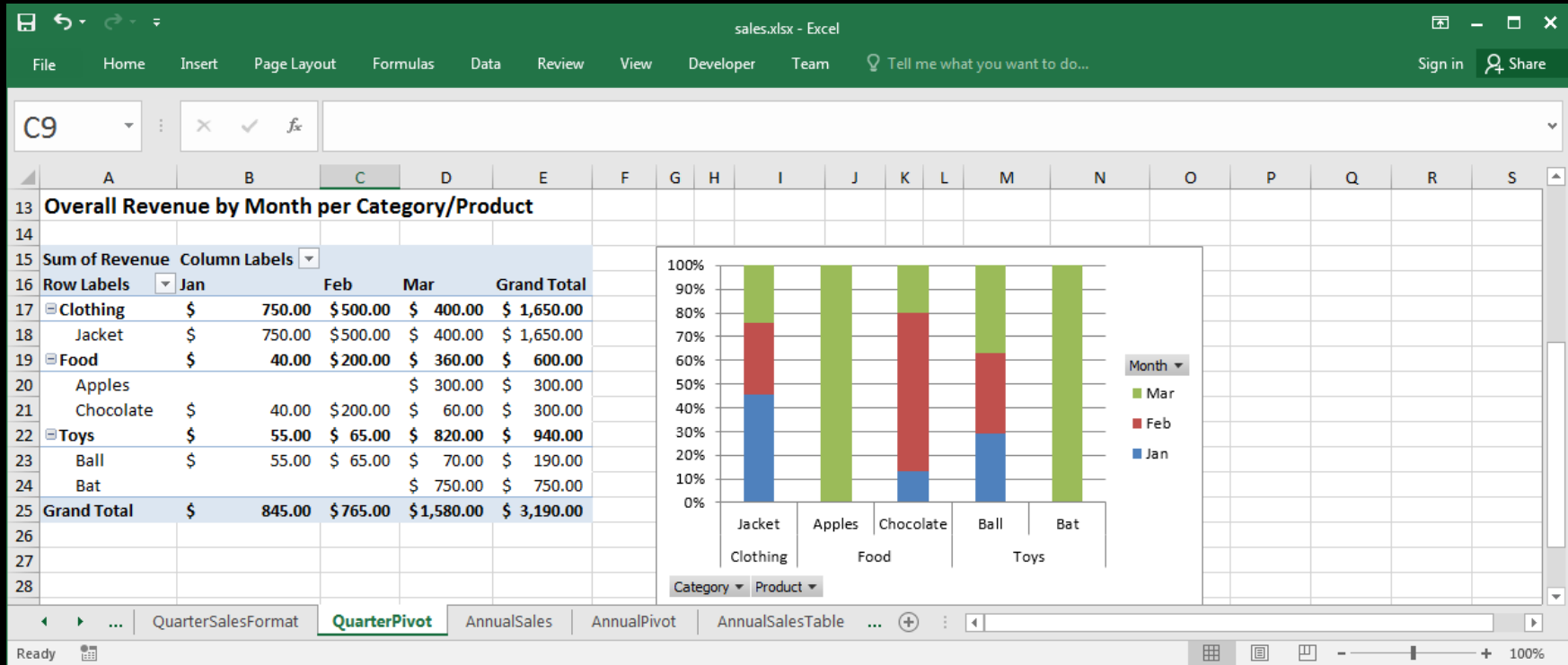
Overall Revenue by Month per Category/Product

Sum of Revenue	Column Labels	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Grand Total
Row Labels														
Clothing		\$ 750.00	\$ 500.00	\$ 400.00	\$ 250.00	\$ 100.00	\$ 800.00	\$ 1,800.00	\$ 3,000.00	\$ 400.00	\$ 500.00	\$ 1,500.00	\$ 2,500.00	\$ 12,500.00
Jacket		\$ 750.00	\$ 500.00	\$ 400.00	\$ 250.00	\$ 100.00				\$ 500.00	\$ 1,500.00	\$ 2,500.00	\$ 6,500.00	
Shorts							\$ 800.00	\$ 1,800.00	\$ 3,000.00	\$ 400.00				\$ 6,000.00
Food		\$ 40.00	\$ 200.00	\$ 360.00	\$ 520.00	\$ 380.00	\$ 520.00	\$ 580.00	\$ 510.00	\$ 530.00	\$ 820.00	\$ 620.00	\$ 650.00	\$ 5,730.00
Apples				\$ 300.00	\$ 360.00	\$ 330.00	\$ 420.00	\$ 480.00	\$ 390.00	\$ 450.00	\$ 420.00	\$ 540.00	\$ 450.00	\$ 4,140.00
Chocolate		\$ 40.00	\$ 200.00	\$ 60.00	\$ 160.00	\$ 50.00	\$ 100.00	\$ 100.00	\$ 120.00	\$ 80.00	\$ 400.00	\$ 80.00	\$ 200.00	\$ 1,590.00
Toys		\$ 55.00	\$ 65.00	\$ 820.00	\$ 840.00	\$ 1,600.00	\$ 2,370.00	\$ 150.00	\$ 180.00	\$ 150.00	\$ 100.00	\$ 80.00	\$ 200.00	\$ 6,610.00
Ball		\$ 55.00	\$ 65.00	\$ 70.00	\$ 90.00	\$ 100.00	\$ 120.00	\$ 150.00	\$ 180.00	\$ 150.00	\$ 100.00	\$ 80.00	\$ 200.00	\$ 1,360.00
Bat				\$ 750.00	\$ 750.00	\$ 1,500.00	\$ 2,250.00							\$ 5,250.00
Grand Total		\$ 845.00	\$ 765.00	\$ 1,580.00	\$ 1,610.00	\$ 2,080.00	\$ 3,690.00	\$ 2,530.00	\$ 3,690.00	\$ 1,080.00	\$ 1,420.00	\$ 2,200.00	\$ 3,350.00	\$ 24,840.00

Ready TryItFilter2 TryItChart TryItWhatIf TryItWhatIf2 TryItPivot 100%

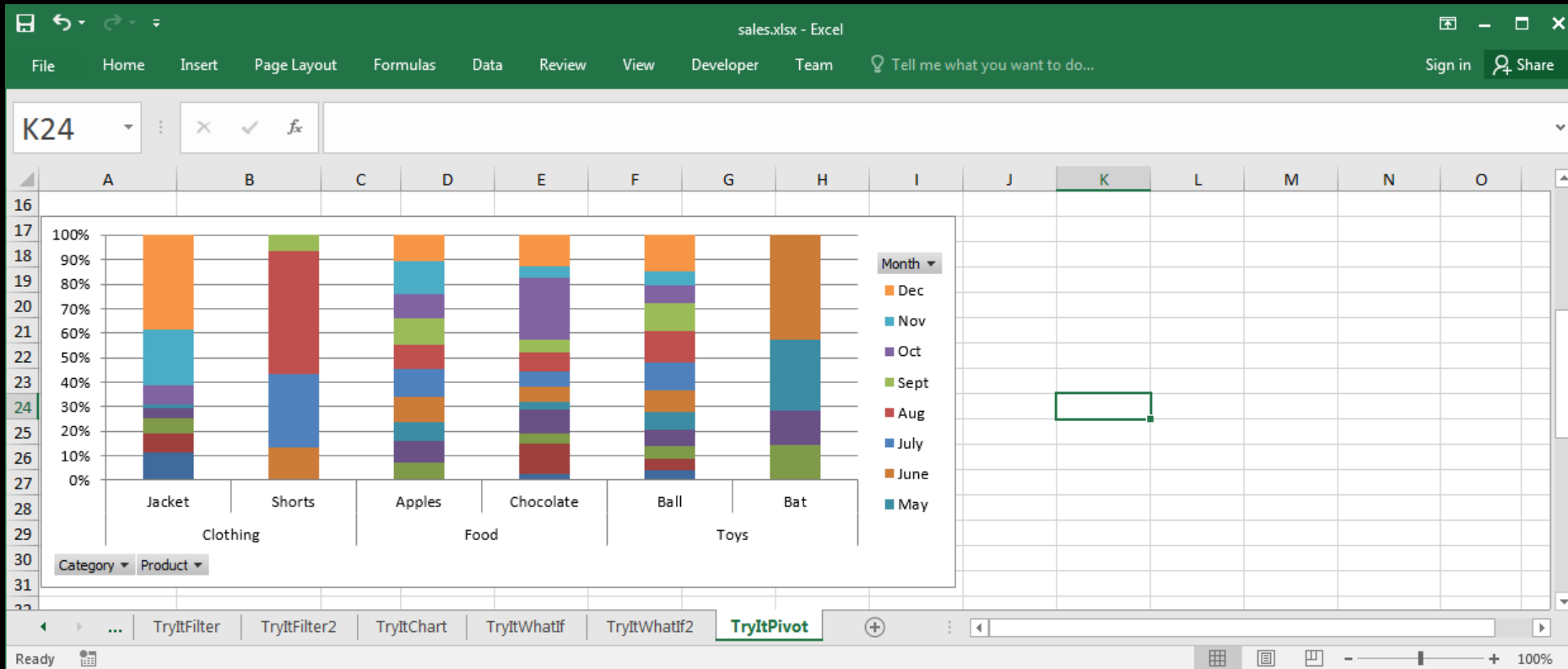
# Pivot Charts

A **pivot chart** is a chart attached to a pivot table. Create it under Insert then Pivot Chart.



# Try it: Pivot Chart

**Question:** Create a pivot chart for previous pivot table.



# What-if and Pivot Tables Question

---

**Question:** How many of the following statements are **TRUE**?

- 1) A what-if scenario can have multiple cells change not just one.
- 2) A pivot table field can be used in ROWS and COLUMNS at the same time.
- 3) A pivot table field can be used in VALUES more than once.
- 4) In our sales spreadsheet example, if Product and Category are both used in ROWS then the order they are list does not matter.
- 5) It is not possible for a field that is a string to be used in VALUES.

**A) 0**      **B) 1**      **C) 2**      **D) 3**      **E) 4**

# Conditions and Decisions

---

A **condition** is an expression that is either TRUE or FALSE.

Conditions are used to make decisions and perform different actions depending on the condition value.

Excel condition and decision functions:

- FALSE () – returns FALSE
- TRUE () – returns TRUE
- AND (cond1, cond2) – returns TRUE if both cond1 and cond2 are true
- OR (cond1, cond2) – returns TRUE if either or both of cond1 and cond2 are true
- NOT (cond) – returns TRUE if cond is FALSE

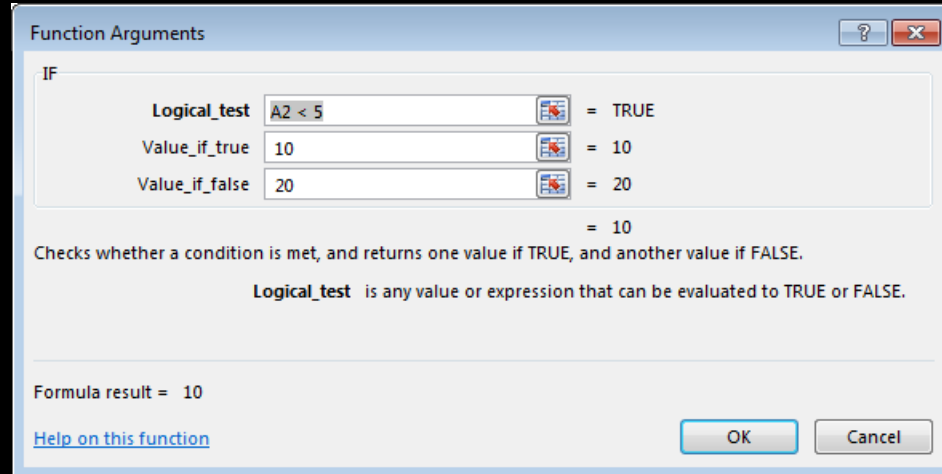
# Decisions using IF ( )

The IF ( ) function is used to make a decision based on a condition.

- $IF(\text{condition}, \text{value\_if\_true}, \text{value\_if\_false})$

Example: If cell A2 is less than 5, return 10 otherwise return 20.

= IF (A2 < 5, 10, 20)





# Try it: Conditions and IF ( )

**Question:** Create two conditions:

- 1) If cell B2  $\geq$  10, then show C2, otherwise D2.
- 2) If cell B2  $<$  15 and C2  $>$  20, return B2\*C2, otherwise if D2  $<$  10, return 1, else 4.

The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'sales.xlsx'. The active cell is D6. The spreadsheet contains the following data:

	A	B	C	D	E	F	G	H	I
1	What-If Conditions								
2	4	10	15	20					
3	Q1:		Q2:						
4	15		4						
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									

The formula bar shows the active cell D6 is empty. The status bar at the bottom indicates 'Ready' and '100%' zoom.

## Decisions using IF () Question

**Question:** How many of these statements are **TRUE**? **A1=40** , **A2=10**

1) =AND ( FALSE () , TRUE () )

2) =OR ( FALSE () , NOT ( TRUE () ) )

3) =IF ( A1=40 , 5 , 10 ) returns 10.

4) =IF ( OR ( A1=40 , A2>10 ) , 1 , 2 ) returns 2.

5) =IF ( A2=10 , IF ( A1=40 , FALSE () ) , TRUE () )

**A) 0**

**B) 1**

**C) 2**

**D) 3**

**E) 4**

# Goal Seek

**Goal seek** is used to have Excel solve for a variable given the target value of another cell.

- Example: How many balls would we have to sell in January to have total revenue for first 3 months of \$4000? Answer: 865

The screenshot shows an Excel spreadsheet with the following data:

Category	Product	Month	Volume	Price	Cost	Revenue
Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00
Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00
Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00
Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00
Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00
Toys	Ball	Feb	65	\$ 1.00	\$ 0.60	\$ 65.00
Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00
Toys	Ball	Mar	70	\$ 1.00	\$ 0.40	\$ 70.00
Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00
Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00
Food	Apples	Mar	100	\$ 3.00	\$ 2.00	\$ 300.00
<b>Total:</b>						<b>\$3,190.00</b>

The Goal Seek dialog box is configured as follows:


- Set cell: G13
- To value: 4000
- By changing cell: D54

# Linear Programming with Solver


Solver performs linear programming to maximize or minimize a given function by changing multiple variables subject to constraints.

	A	B
1	<b>Solver Example: Maximum House You Can Afford</b>	
2		
3	<b>House Value:</b>	<b>\$ 500,000.00</b>
4	<b>Down payment:</b>	\$ 20,000.00
5	<b>Mortgage amount:</b>	\$ 480,000.00
6	<b>Amoritization in years:</b>	25
7	<b>Monthly mortgage expense:</b>	<b>\$2,533.62</b>
8		
9	<b>Monthly income:</b>	\$ 5,000.00
10	<b>Mortgage rate:</b>	4%
11		
12	<b>Maximum mortgage expense percentage:</b>	30%
13	<b>Maximum mortgage expense amount:</b>	\$ 1,500.00
14		\$1,033.62

**Solver Parameters**

Set Objective:  

To:  Max  Min  Value Of:

By Changing Variable Cells:  

Subject to the Constraints:

Make Unconstrained Variables Non-Negative

Select a Solving Method:

Solving Method  
 Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.

# Analysis ToolPak

---

The Analysis ToolPak is an Excel add-in that has a set of statistical and data analysis tools such as ANOVA, covariance, regression, and t-test.

Analysis ToolPak is not installed by default.

- To install: File → Options → Add-Ins
- Select Excel Add-ins in the Manage: box and select Go...
- Choose AnalysisToolPak and select OK

You should now see Data Analysis under the Data tab

# Regression

---

**Linear regression** models the relationship between a dependent variable  $y$  and explanatory variables  $X$ .

- Simple linear regression has one explanatory variable:  $y = Bx + \varepsilon$
- Used to fit a predictor model on observed data and also used to determine the strength of the relationship between  $y$  and  $X$  variables.

**Trend lines** are often calculated using linear regression.

The technique provides a way to determine patterns in the data set and model the data so that new values can be predicted.

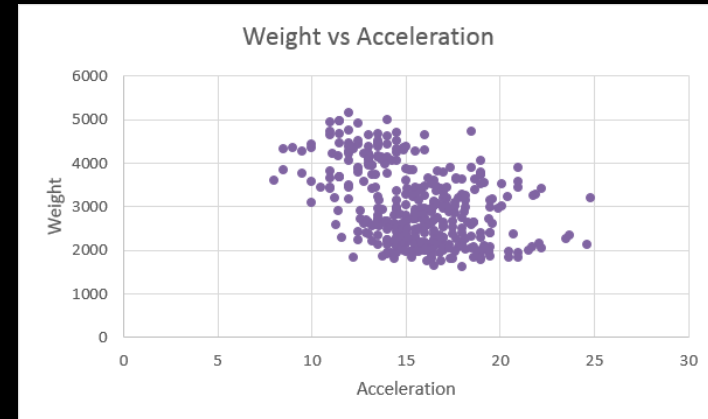
# Regression in Excel

Excel provides a regression function that will calculate:

- $R^2$
- ANOVA table
- regression equation coefficients
- standardized and unstandardized residuals

Example: Given a data set of car weight and acceleration, determine if there is any relationship between them.

Scatterplot shows weak relationship with no strong patterns, and we would expect to see this shown in the regression analysis.



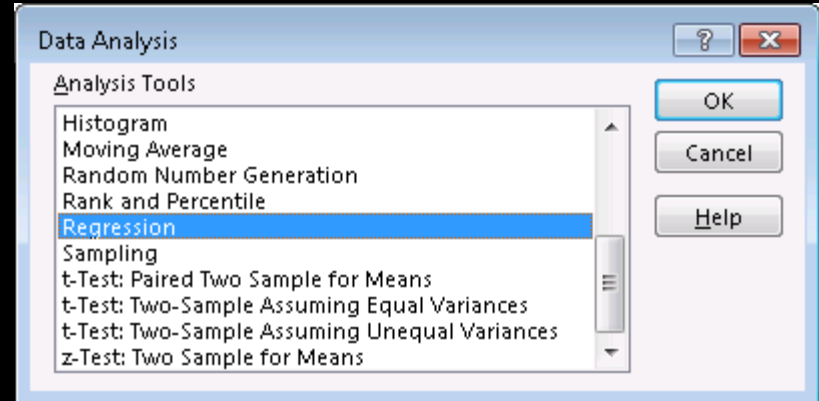
# Regression Example

Regression computes constants  $m$  and  $b$  in formula:

$$\text{weight} = m * \text{acceleration} + b$$

Weight is the dependent variable and acceleration is the independent variable.

To start select, Data Analysis from the data tab and then select Regression and OK.





# Regression Example Settings

## Settings:

- Response (dependent) data for the Input Y Range
- Columns for the explanatory (independent) data (X Range).
- For residual information select, Residuals, Standardized Residuals, and Residual Plots from the Residuals section.

The screenshot shows the 'Regression' dialog box with the following settings:

- Input**
  - Input Y Range: \$D\$2:\$D\$393
  - Input X Range: \$C\$2:\$C\$393
  - Labels
  - Constant is Zero
  - Confidence Level: 95 %
- Output options**
  - Output Range: [Empty]
  - New Worksheet Ply: [Empty]
  - New Workbook
- Residuals**
  - Residuals
  - Standardized Residuals
  - Residual Plots
  - Line Fit Plots
- Normal Probability**
  - Normal Probability Plots

Buttons: OK, Cancel, Help

# Regression Example Results

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.416839282							
R Square	0.17375492							
Adjusted R Square	0.171636343							
Standard Error	2.510965983							
Observations	392							
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	517.0999442	517.0999442	82.01491373	6.56562E-18			
Residual	390	2458.930566	6.304950169					
Total	391	2976.03051						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	19.57266581	0.462860061	42.28635703	9.5445E-148	18.66265269	20.48267893	18.66265269	20.48267893
X Variable 1	-0.001353896	0.000149499	-9.056208574	6.56562E-18	-0.001647821	-0.001059971	-0.001647821	-0.001059971

$R^2 * 100\%$  = percentage of variation in dependent variable explained by independent variable

Coefficients for the regression equation

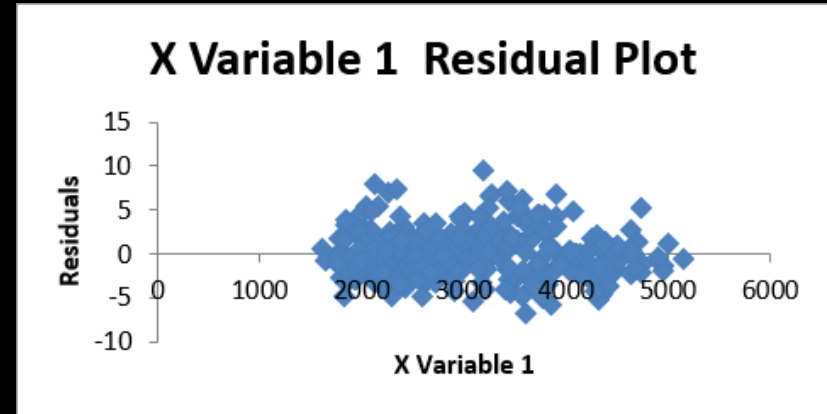
All of the output is put into a new sheet. Read the values off of the table and form the regression equation:

- weight =  $-0.001 * \text{acceleration} + 19.572$

## Regression Example Results (cont.)

Below the previous tables are the predicted y values (from the regression equation) as well as the residuals and standardized residuals. All plots are placed to the right of the charts.

Observation	Predicted Y	Residuals	Standard Residuals
1	14.82861427	-2.828614269	-1.127947728
2	14.57272793	-3.072727927	-1.22529131
3	14.9206792	-3.920679196	-1.563423207
4	14.92474088	-2.924740884	-1.1662795
5	14.90307855	-4.403078548	-1.755786393
6	13.69540333	-3.695403327	-1.473591445
7	13.67780268	-4.67780268	-1.865336311
8	13.73466631	-5.234666311	-2.087393123
9	13.58167606	-3.581676065	-1.428241179
10	14.36016626	-5.860166257	-2.336819583
11	14.74873441	-4.748734406	-1.893621284

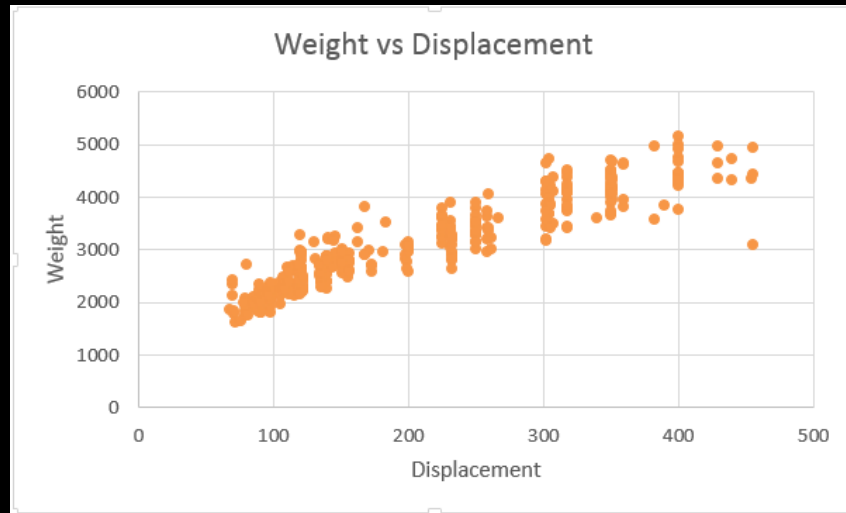


Expected a weak relationship and this is demonstrated by the  $R^2$  value.

- Only 17.4% of the variation in weight is explained by acceleration.

# Try It: Regression

**Question:** Perform a regression analysis between weight (dependent) and displacement (independent) variable.



# Conclusion

---

**Spreadsheets** are general purpose tools for data analysis that consist of a table of cells which contain data and formulas.

Formulas contain data values, cell references, and functions.

- Aggregate functions summarize multiple data values into a single value.
- Functions exist for statistics, string manipulation, lookup/indexing, and decisions.

Spreadsheets provide tools for data sorting, filtering, visualization using charts, and summarization (pivot tables).

- Also contain tools for what-if scenarios, goal seek, linear solvers, and statistical analysis tools.

# Objectives

---

- Explain what a spreadsheet is.
- Explain how cells are addressed in a spreadsheet.
- List some of the ways to select cells in a spreadsheet.
- Define and explain: formula, function, argument, concatenation
- Use these functions: concatenate, lookup, index
- Explain the difference between an absolute and relative address.
- Explain how an aggregate function works. List some examples.
- Explain how to use conditional formatting.
- Explain how spreadsheets can be used as a database. Use sorting and filtering.
- Be able to create and edit charts and use chart features: trendlines, sparklines
- Explain the usefulness of: what-if scenarios, goal seek, solver
- Use and create pivot tables and charts.
- Evaluate and create conditions. Use IF() to make decisions.