

COSC 301

Introduction to Data Analytics

Spreadsheets: Microsoft Excel

Dr. Mostafa Mohamed
University of British Columbia Okanagan
Mostafa.Mohamed@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 Excel interface with the following data in the spreadsheet:

	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						

The formula bar shows: `=SUM(G2:G12)` (circled in blue, labeled "formula in cell").

Cell G13 is circled in blue and labeled "Cell G13".

A red bracket on the left indicates "rows".

A yellow bracket on the right indicates "columns".

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.

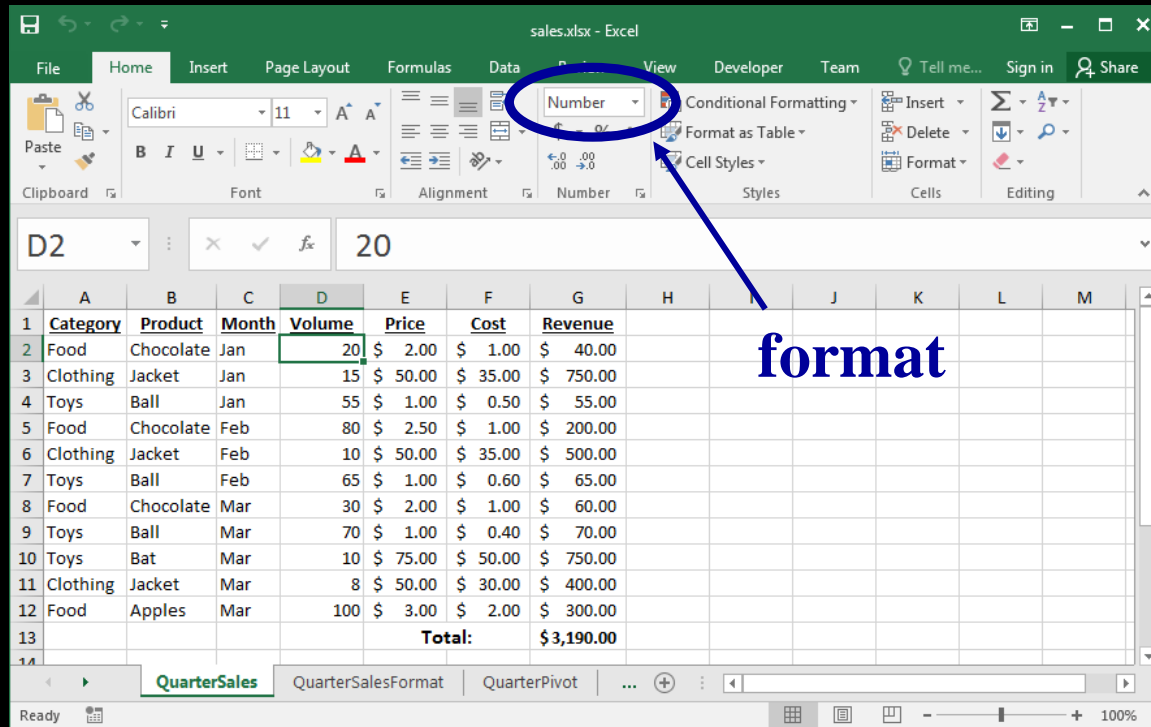
- e.g. B3 is the 2nd column and the 3rd row.

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 the 'sales.xlsx' file open. The 'Home' tab is active, and the 'Number' format is selected in the ribbon. The formula bar shows the value '20' entered in cell D2. The spreadsheet contains a table with columns: Category, Product, Month, Volume, Price, Cost, and Revenue. The data is organized by month (Jan, Feb, Mar) and product type (Food, Clothing, Toys). A blue arrow points from the word 'format' to the 'Number' dropdown in the ribbon.

	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. The 'Font' group on the ribbon is circled in blue, and an arrow points from the text 'format shortcuts' to it. The spreadsheet below shows a table with columns: Category, Product, Month, Volume, Price, Cost, Revenue. The 'Volume' column is highlighted in green, and the value '20' is entered in cell D2.

	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 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

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me... Sign in Share

Clipboard Font Alignment Number Styles Cells Editing

A3 Clothing

	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						

QuarterSales QuarterSalesFormat QuarterPivot ...

Ready Average: 115 Count: 28 Sum: 1840 100%

Selecting Individual Cells Example

Excel window showing a spreadsheet titled "sales.xlsx". The ribbon is set to "Home". The formula bar shows the formula `=SUM(G2,G5,G8,G10,G12` being entered into 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				Total:			G10,G12						

The status bar at the bottom shows "QuarterSales" is selected, with other tabs "QuarterSalesFormat" and "QuarterPivot" visible.

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

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me... Sign in Share

Clipboard Font Alignment Number Styles Cells Editing

D2 20

	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						

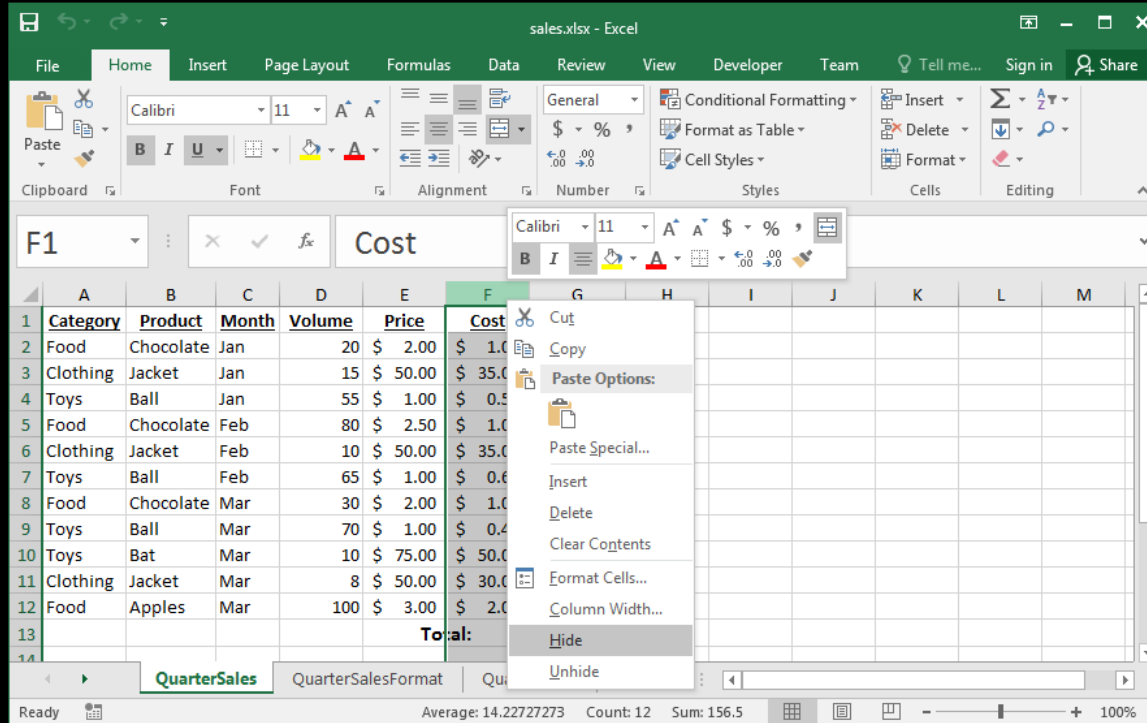
QuarterSales QuarterSalesFormat QuarterPivot ...

Ready 100%

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 the 'sales.xlsx' file open. The ribbon is set to 'Home'. The active cell is F1, containing the text 'Cost'. A right-click context menu is open over the column F header, displaying various options. The 'Hide' option is highlighted in the menu. 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	\$ 1.00							
3	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00							
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50							
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00							
6	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00							
7	Toys	Ball	Feb	65	\$ 1.00	\$ 0.60							
8	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00							
9	Toys	Ball	Mar	70	\$ 1.00	\$ 0.40							
10	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00							
11	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00							
12	Food	Apples	Mar	100	\$ 3.00	\$ 2.00							
13						Total:							

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 Microsoft Excel interface with the 'sales.xlsx' file open. The formula bar at the top displays the formula `=SUM(G2:G12)` in cell G13. A blue circle highlights the formula, and a blue arrow points to it with the label 'formula'. The worksheet contains a table with columns: Category, Product, Month, Volume, Price, Cost, and Revenue. The data rows show sales for various products over three months. The total revenue for March is calculated as \$3,190.00.

	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						

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

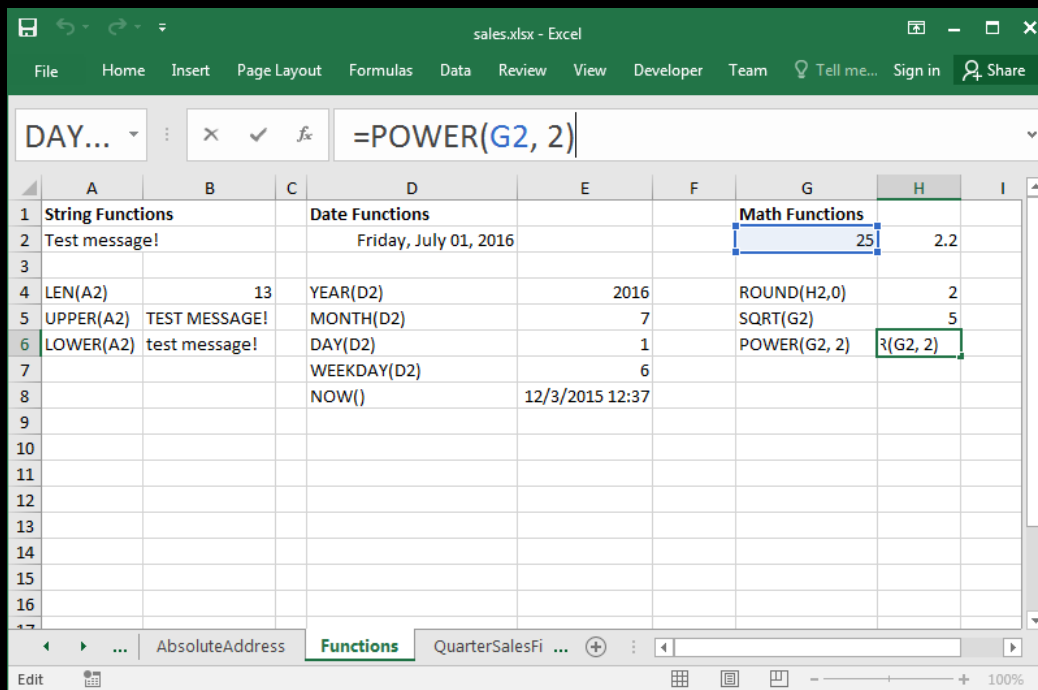
B) 18

C) $=2+4*3$

Using Excel Functions

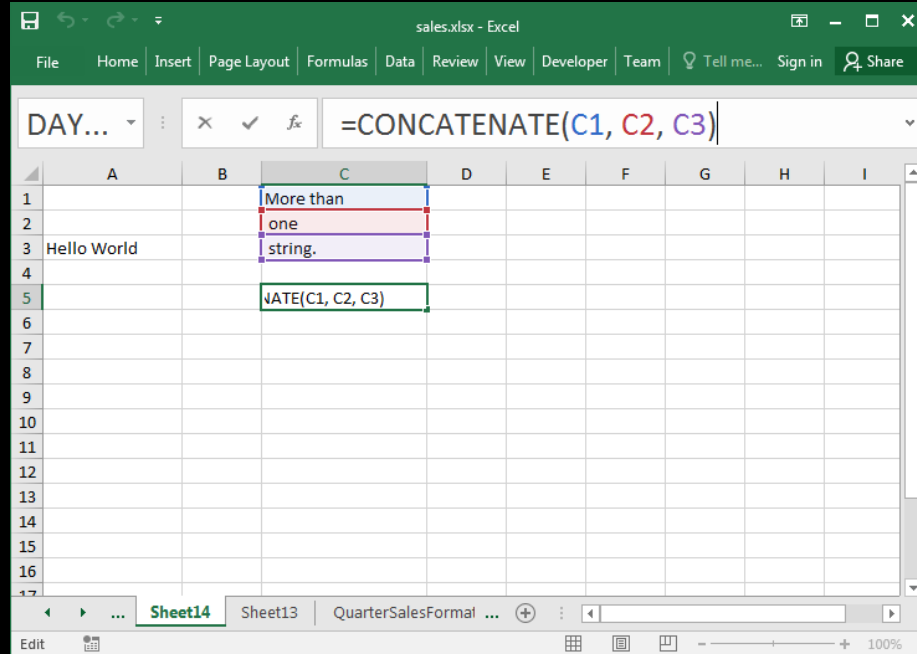
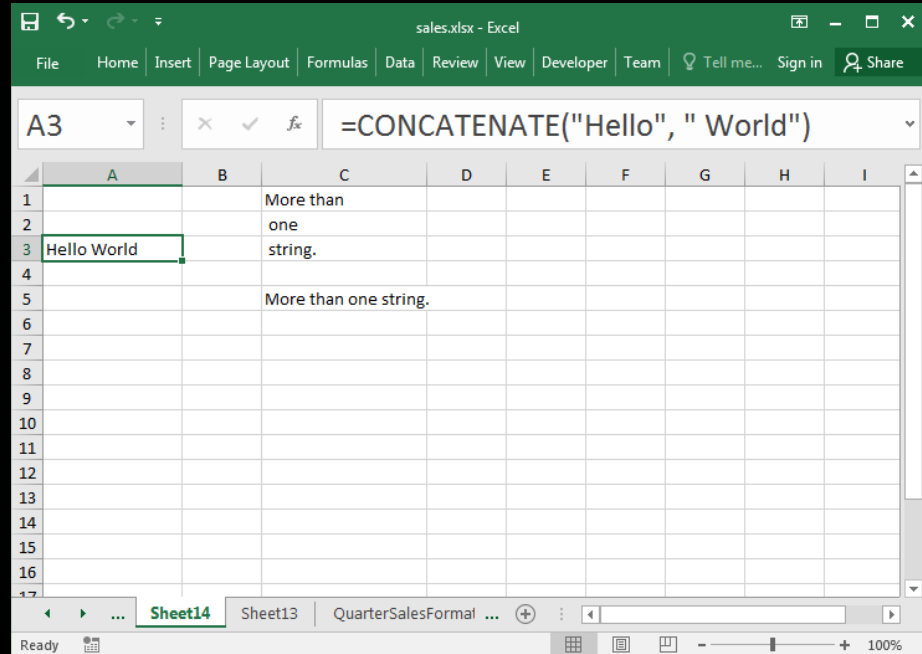
Excel has a large number of built-in functions to use.

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



Concatenation

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



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 Excel interface with the formula bar displaying `=LOOKUP(F2, A2:A6, C2:C6)`. A tooltip provides the function syntax: `LOOKUP(lookup_value, lookup_vector, [result_vector])`. The spreadsheet data is as follows:

	A	B	C
1	Product Id	Product Name	Product Price
2	1	Apple	\$ 3.99
3	2	Banana	\$ 2.99
4	3	Lettuce	\$ 1.99
5	4	Squash	\$ 6.99
6	5	Pumpkin	\$ 5.99

The formula bar also shows a partial formula `DAY...` on the left. The status bar at the bottom indicates the current cell is F2, and the formula bar shows the active formula `=LOOKUP(F2, A2:A6, C2:C6)`.

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:** Displays the formula `=INDEX(B2:B6,F2+1)`.
- Worksheet 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
- Formula Bar Tooltip:**

INDEX(array, row_num, [column_num])
 INDEX(reference, row_num, [column_num], [area_num])
- Cell F2:** Contains the value 2, highlighted with a red border.
- Cell B6:** Contains the value 5, highlighted with a green border.
- Formula Bar Tooltip (VLOOKUP):**

VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])

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 titled 'sales.xlsx - Excel'. The ribbon includes File, Home, Insert, Page Layout, Formulas, Data, Review, View, Developer, Team, Tell me..., Sign in, and Share. The Home ribbon is active, showing options for Clipboard, Font, Alignment, Number, Styles, Cells, and Editing. The formula bar shows 'H13'. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
	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						

The status bar at the bottom shows 'Ready' and 'QuarterSalesTryIt1'.



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

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me... Sign in Share

Clipboard Font Alignment Number Styles Cells Editing

D2 \times \checkmark f_x `=max(D2:D12)`

	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				<code>=max(D2:D12)</code>	tal:		\$ 3,190.00						
14				<code>MAX(number1, [number2], ...)</code>									

QuarterSales QuarterSalesFormat QuarterPivot ... 100%

Try it: Aggregate Functions

Question: Create aggregate functions to match below:

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me... Sign in Share

Clipboard Font Alignment Number Styles Cells Editing

H13

	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				42	\$ 75.00	\$ 0.40	\$ 3,190.00						
14													

QuarterSalesTryIt1 QuarterSalesTryIt2

Ready 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)

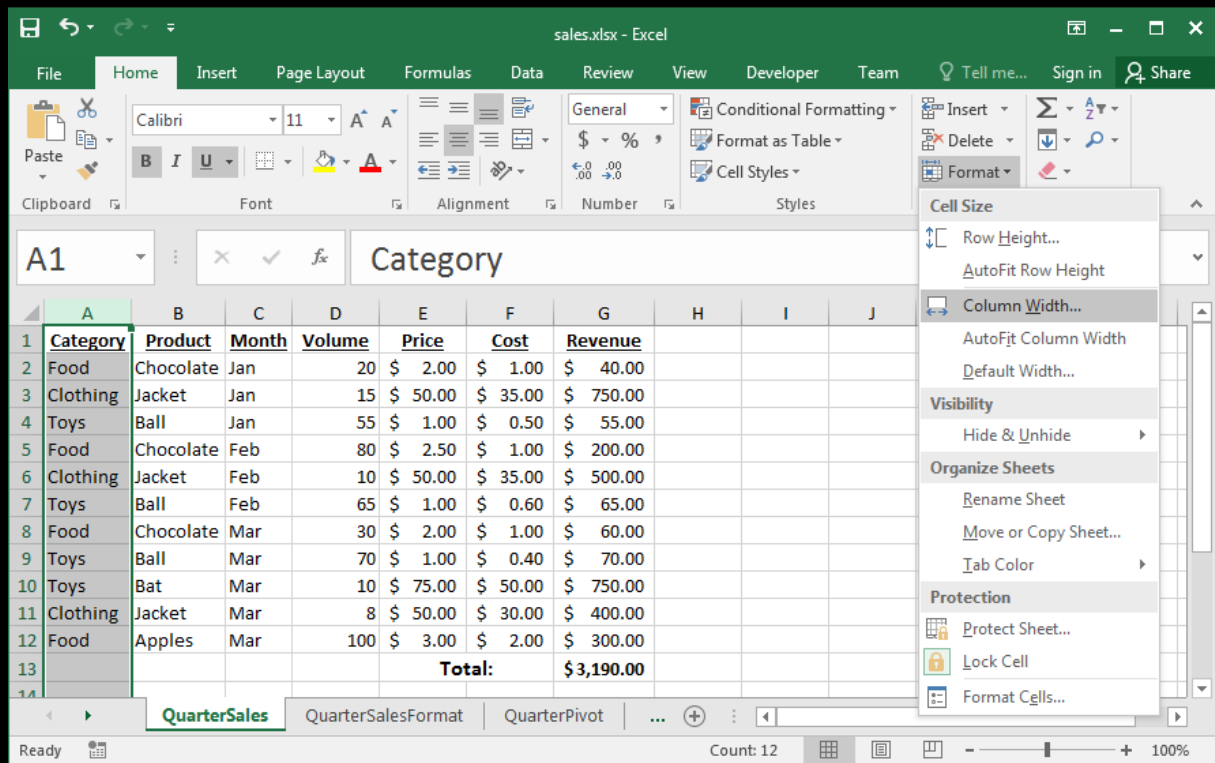
B) MIN (A1 : C1)

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 Excel interface with the 'Format' menu open. The 'Column Width...' option is highlighted. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J
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										
14										

The status bar at the bottom shows 'Count: 12' and '100%' zoom.

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 sales data table. The table has columns for Category, Product, Month, Volume, Price, Cost, and Revenue. The Revenue column (G) is highlighted with conditional formatting. The Conditional Formatting Rules Manager dialog is open, showing two 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"/>

The data table shows the following values for the Revenue column:

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
Total:						\$3,190.00

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

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	Total:			\$3,190.00						

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:

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me... Sign in Share

Clipboard Font Alignment Number Styles Cells Editing

D13

	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						

QuarterSalesTryIt2 QuarterSalesTryIt3

Ready 100%

Try it: Conditional Formatting Challenge

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

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me... Sign in Share

Clipboard Font Alignment Number Styles Cells Editing

Calibri 11 A A B I U \$ % 0 .00 0.00 General Conditional Formatting Insert Delete Format Format as Table Cell Styles

D13

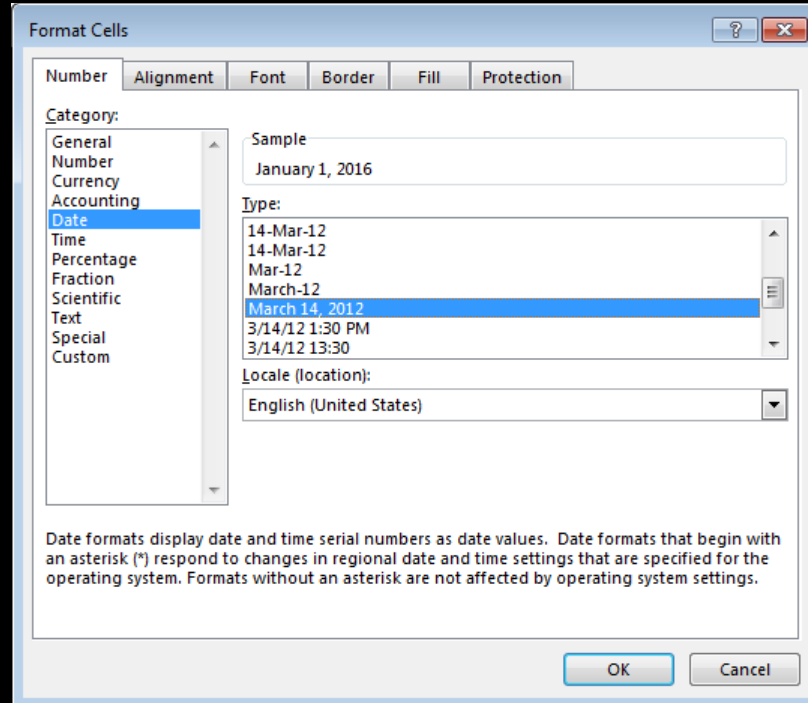
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue						
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
3	Clothing	Jacket	Jan	15	\$50.00	\$35.00	\$ 750.00						
4	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
5	Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00						
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						

QuarterSalesTryIt3 QuarterSalesTryIt4

Ready 100%

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, displaying the following settings:

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

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				Total:			\$ 3,190.00

The status bar at the bottom shows: Ready, Average: 91.97727273, Count: 77, Sum: 4047, 100% zoom.

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', 'Text to Columns', 'What-If Analysis', and 'Forecast Sheet'. The 'Sort' button is highlighted, indicating the sorting process. The data table below shows sales records with columns for Category, Product, Month, Volume, Price, Cost, and Revenue. The total revenue is \$3,190.00.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue						
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					Total:		\$ 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.

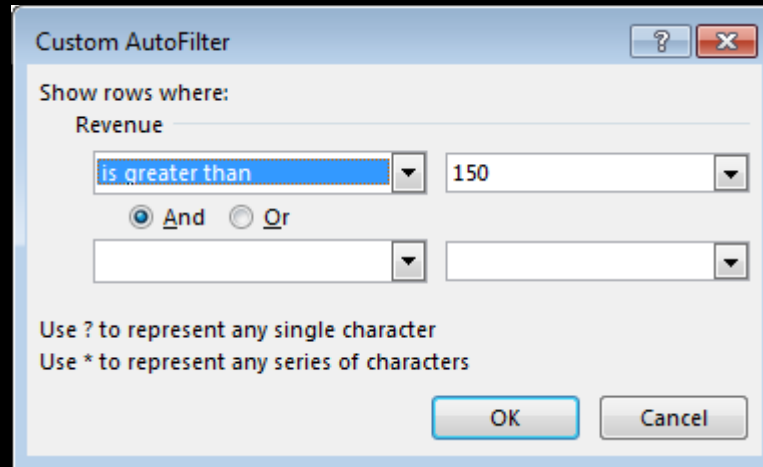
The screenshot shows the Microsoft Excel interface with a spreadsheet titled 'sales.xlsx'. The spreadsheet has columns A through K. Column A is 'Category', B is 'Product', C is 'Month', D is 'Volume', E is 'Price', F is 'Cost', and G is 'Revenue'. The 'Revenue' column is selected, and a filter menu is open. The filter menu shows options for sorting and filtering. The 'Filter by Color' option is selected, and a sub-menu is open showing a list of values with checkboxes. The 'Top 10...' option is highlighted.

Category	Product	Month	Volume	Price	Cost	Revenue
Food	Chocolate	Feb				
Clothing	Jacket	Feb				
Food	Apples	Mar				

QuarterSalesFilter

Custom Filter Example

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



Custom Filter Result

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

Excel screenshot showing a custom filter applied to the 'Revenue' column. The filter criteria are 'Revenue > 150', resulting in 7 of 12 records being displayed.

The data table is as follows:

	A	B	C	D	E	F	G	H	I	J	K
	Category	Product	Month	Volume	Price	Cost	Revenue				
1											
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					Total:		\$ 3,190.00				
14											
15											
16											
17											
18											
19											

Excel status bar: 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 ribbon includes options for 'Sort & Filter', 'Data Tools', and 'Forecast'. The 'Filter' button is highlighted. The data table is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Price	Cost	Revenue						
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'.

Try it: Filter Challenge

Question: Filter the data so only products with volume < 20 **or** revenue < \$65 are shown.

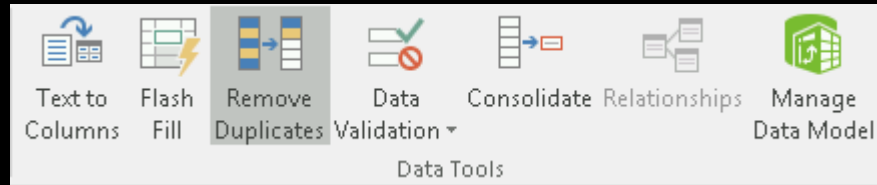
The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The 'Filter' button is active, and the 'TryItFilter2' filter is applied. The data table is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1				Volume			Revenue						
2				<20									
3							<65						
4													
5	Category	Product	Month	Volume	Price	Cost	Revenue						
6	Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00						
7	Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00						
8	Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00						
9	Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00						
14	Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00						
15	Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00						
16	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00						
18													
19													

The status bar at the bottom indicates 'Ready 7 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	Values
Then by	Char	Values
		Largest to Smallest
		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

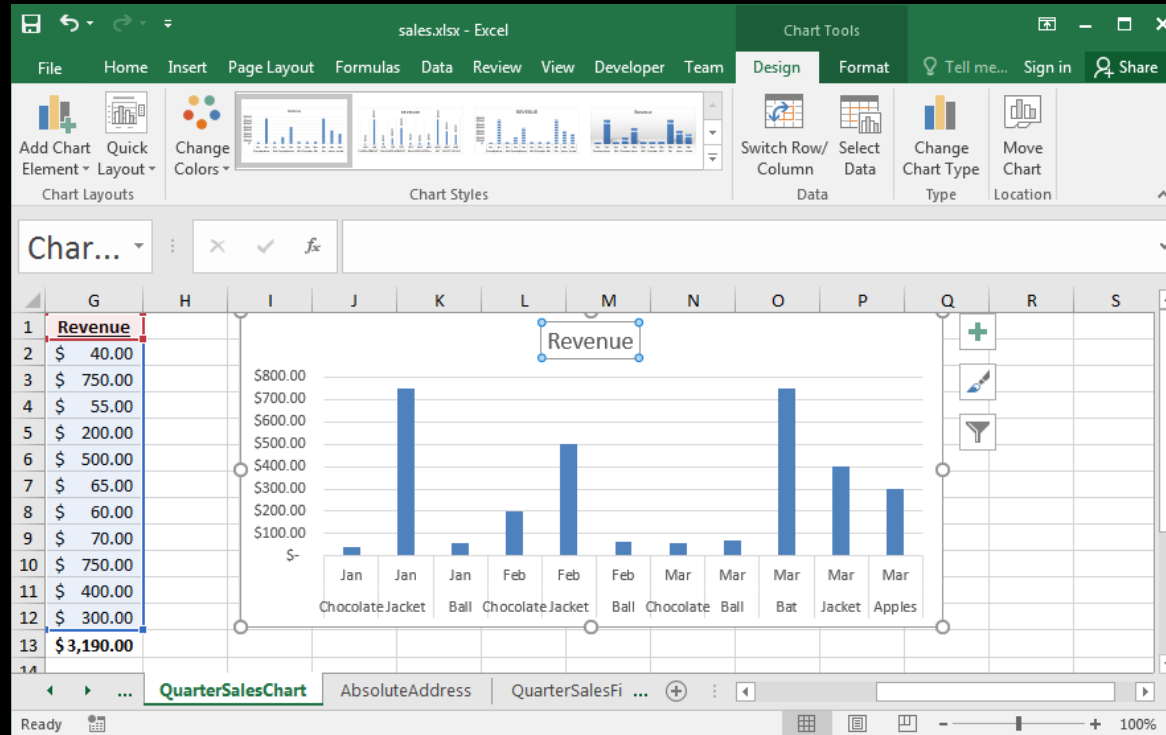
Charts

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 Options

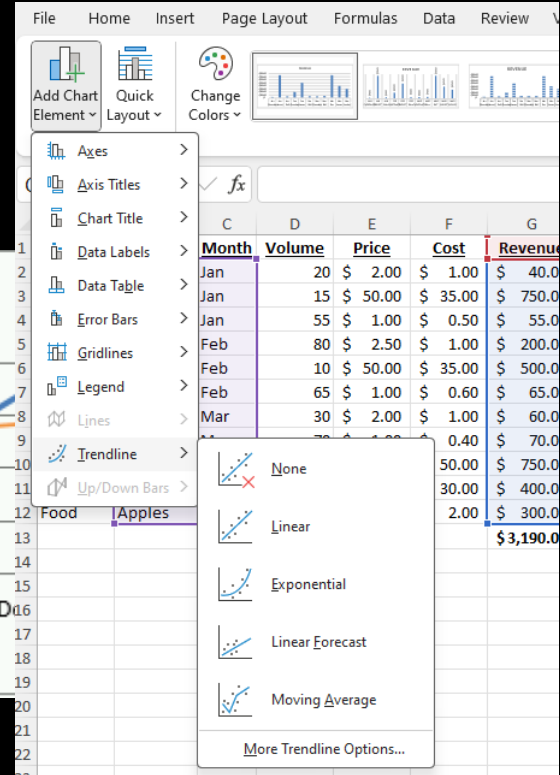
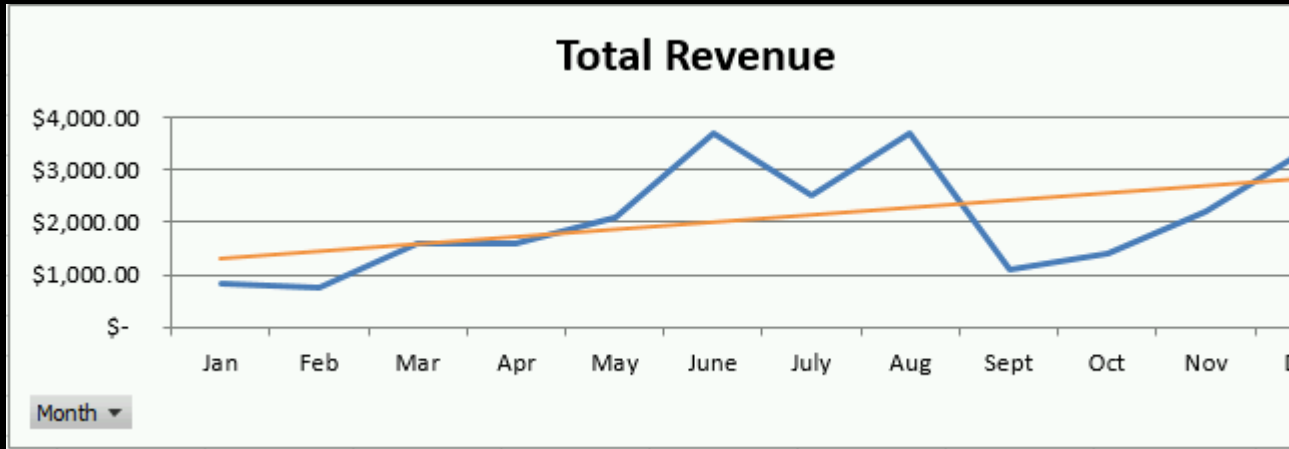
Chart Tools allows you to modify the data in the chart, change the chart type, and move the chart in the Worksheet.



Trendlines

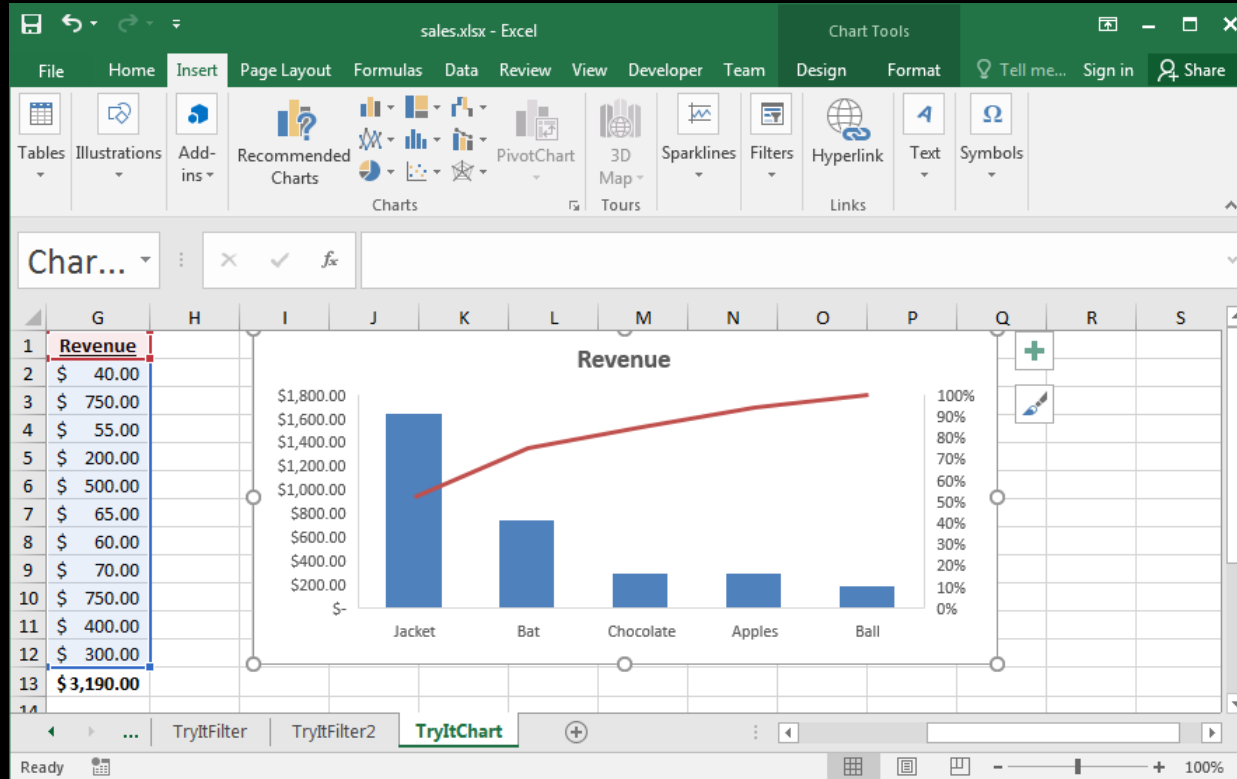
Trendlines can be easily added to any chart.

- Linear trendline for monthly revenue. Good choice?



Try it: Chart

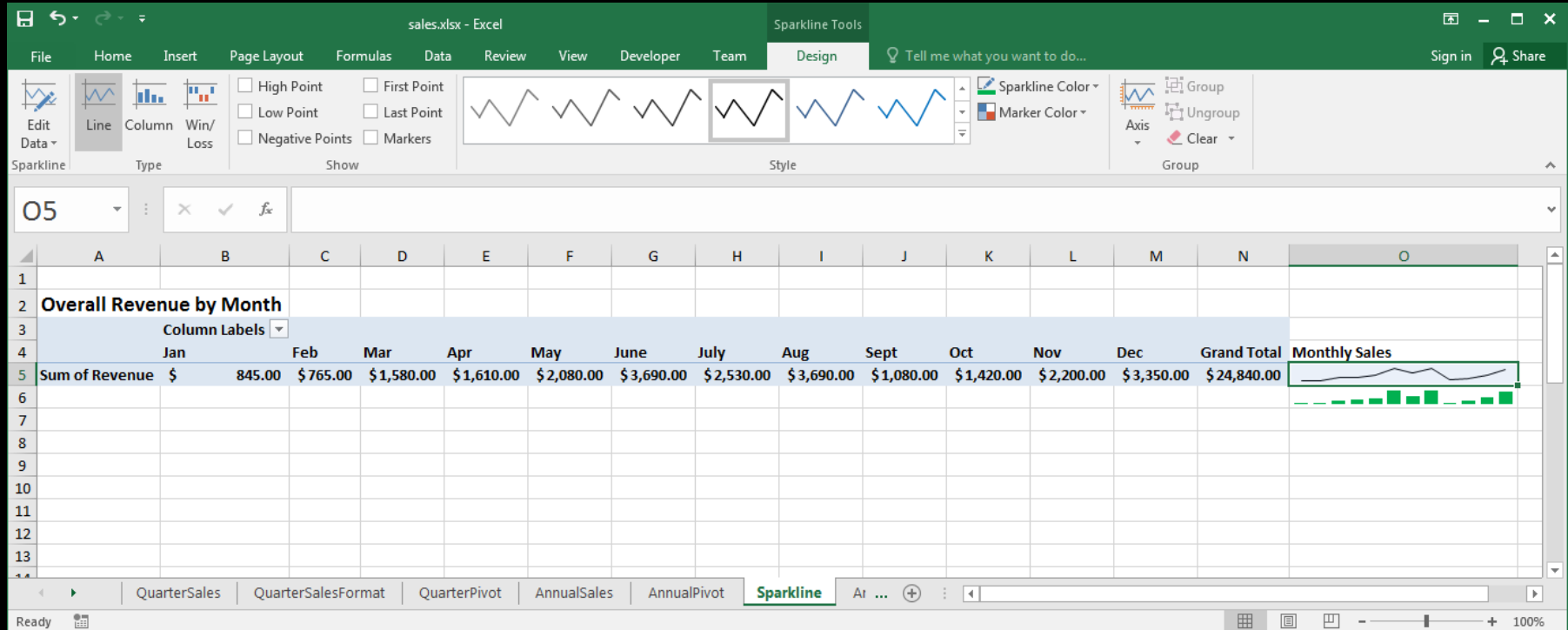
Question: Create a chart that makes it easy to see the best selling product.



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.



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**.

The Scenario Manager dialog box is shown. It has a title bar with a question mark and a close button. Inside, there is a list of scenarios: "Cold Winter" (selected) and "Normal". To the right of the list are buttons: "Add...", "Delete", "Edit...", "Merge...", and "Summary...". Below the list, there is a "Changing cells:" field containing "\$D\$3" and a "Comment:" text area. At the bottom are "Show" and "Close" buttons.

The Edit Scenario dialog box is shown. It has a title bar with a question mark and a close button. Inside, there is a "Scenario name:" field containing "Cold Winter". Below it is a "Changing cells:" field containing "D3" with a selection icon. A note says "Ctrl+click cells to select non-adjacent changing cells." Below that is a "Comment:" text area. At the bottom, there is a "Protection" section with two checkboxes: "Prevent changes" and "Hide", both of which are unchecked. At the bottom right are "OK" and "Cancel" buttons.

The Scenario Values dialog box is shown. It has a title bar with a question mark and a close button. Inside, there is a text prompt: "Enter values for each of the changing cells." Below it, there is a list of changing cells with their corresponding values. The first entry is "1: \$D\$3" followed by a text box containing the value "50". At the bottom right are "OK" and "Cancel" buttons.

What-If Scenarios Example

User can easily select scenario and see the result.

The screenshot shows the Microsoft Excel interface with the 'Scenario Manager' dialog box open. The dialog box has a title bar 'Scenario Manager' and a list of scenarios: 'Cold Winter' (selected) and 'Normal'. To the right of the list are buttons: 'Add...', 'Delete', 'Edit...', 'Merge...', and 'Summary...'. Below the list, the 'Changing cells:' field is set to '\$D\$3' and the 'Comment:' field is empty. At the bottom of the dialog are 'Show' and 'Close' buttons.

The background Excel spreadsheet shows a table with the following data:

	A	B	C	D	E	F	G	H	I	J
1	Category	Product	Month	Volume	Price	Cost	Revenue			
2	Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00			
3	Clothing	Jacket	Jan	50	\$ 50.00	\$ 35.00	\$2,500.00			
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:		\$4,940.00			

Try it: What-If Scenario

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

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me... Sign in Share

Get External Data New Query Refresh All Properties Edit Links Connections Sort & Filter Filter Clear Reapply Advanced Text to Columns Data Tools What-If Analysis Forecast Outline

H13

	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	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					Total:		\$ 3,380.00						

Ready

Try it: What-If Scenario Challenge

Question: Create a what-if scenario that all costs go up by 10% and volume down by 20%.

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Developer Team Tell me... Sign in Share

Get External Data New Query Refresh All Connections Properties Edit Links Get & Transform

Sort Filter Sort & Filter

Text to Columns Data Tools

What-If Analysis Forecast Sheet Forecast

K10

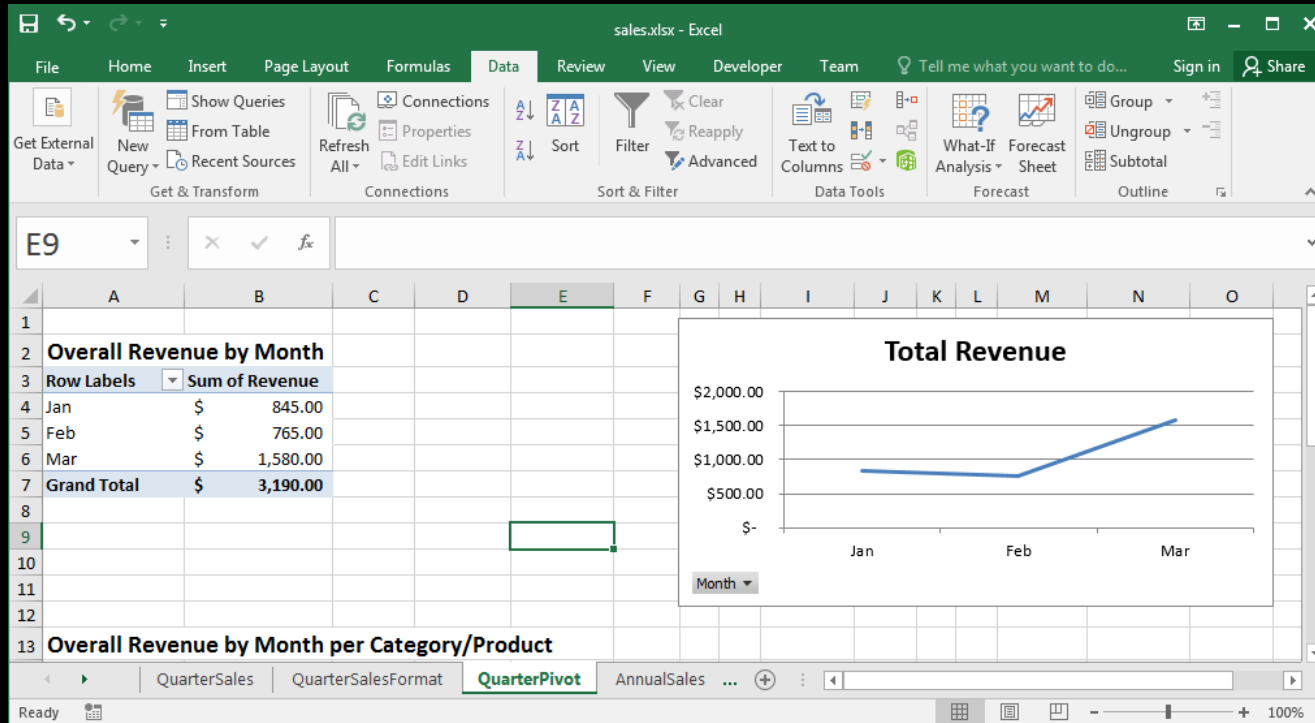
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<u>Category</u>	<u>Product</u>	<u>Month</u>	<u>Volume</u>	<u>Act. Vol.</u>	<u>Price</u>	<u>Cost</u>	<u>Act. Cost</u>	<u>Revenue</u>				
2	Food	Chocolate	Jan	20	16	\$ 2.00	\$ 1.00	\$ 1.10	\$ 32.00		Costs	10%	
3	Clothing	Jacket	Jan	15	12	\$ 50.00	\$ 35.00	\$ 38.50	\$ 600.00		Volume	-20%	
4	Toys	Ball	Jan	55	44	\$ 1.00	\$ 0.50	\$ 0.55	\$ 44.00				
5	Food	Chocolate	Feb	80	64	\$ 2.50	\$ 1.00	\$ 1.10	\$ 160.00				
6	Clothing	Jacket	Feb	10	8	\$ 50.00	\$ 35.00	\$ 38.50	\$ 400.00				
7	Toys	Ball	Feb	65	52	\$ 1.00	\$ 0.60	\$ 0.66	\$ 52.00				
8	Food	Chocolate	Mar	30	24	\$ 2.00	\$ 1.00	\$ 1.10	\$ 48.00				
9	Toys	Ball	Mar	70	56	\$ 1.00	\$ 0.40	\$ 0.44	\$ 56.00				
10	Toys	Bat	Mar	10	8	\$ 75.00	\$ 50.00	\$ 55.00	\$ 600.00				
11	Clothing	Jacket	Mar	8	6	\$ 50.00	\$ 30.00	\$ 33.00	\$ 320.00				
12	Food	Apples	Mar	100	80	\$ 3.00	\$ 2.00	\$ 2.20	\$ 240.00				
13									\$2,552.00				
14													

TryItChart TryItWhatIf TryItWhatIf2

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.

sales.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View

Clipboard Font Alignment Number

A1 Category

	A	B	C	D	E	F	G	H
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	

Sheet1 QuarterSales QuarterSalesFormat Q ...

Ready Average: 91.97727273 Count: 84 Sum: 40

Create PivotTable

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...

Connection name:

☐ 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

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 shows an Excel spreadsheet with a PivotTable and the PivotTable Fields task pane. The PivotTable is located in the range A3:D7 and has the following data:

Row Labels	Sum of Revenue
Jan	845
Feb	765
Mar	1580
Grand Total	3190

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

- Choose fields to add to report:** Search bar, checkboxes for Month, Volume, Price, Cost, and Revenue (checked).
- Drag fields between areas below:**
 - FILTERS:** (Empty)
 - COLUMNS:** (Empty)
 - ROWS:** Month
 - VALUES:** Sum of Revenue
- Defer Layout Update:** (Unchecked)
- UPDATE:** (Button)

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. The PivotTable is located in the range A3:E9. The PivotTable Fields task pane is open on the right side of the screen.

PivotTable Data:

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
Grand Total	\$ 845.00	\$ 765.00	\$ 1,280.00	\$ 2,890.00

PivotTable Fields Task Pane:

- 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...
- ☐ Defer Layout Update
- UPDATE**

Try it: Pivot Table

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

Excel window: sales.xlsx - Excel

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

Formula bar: 12

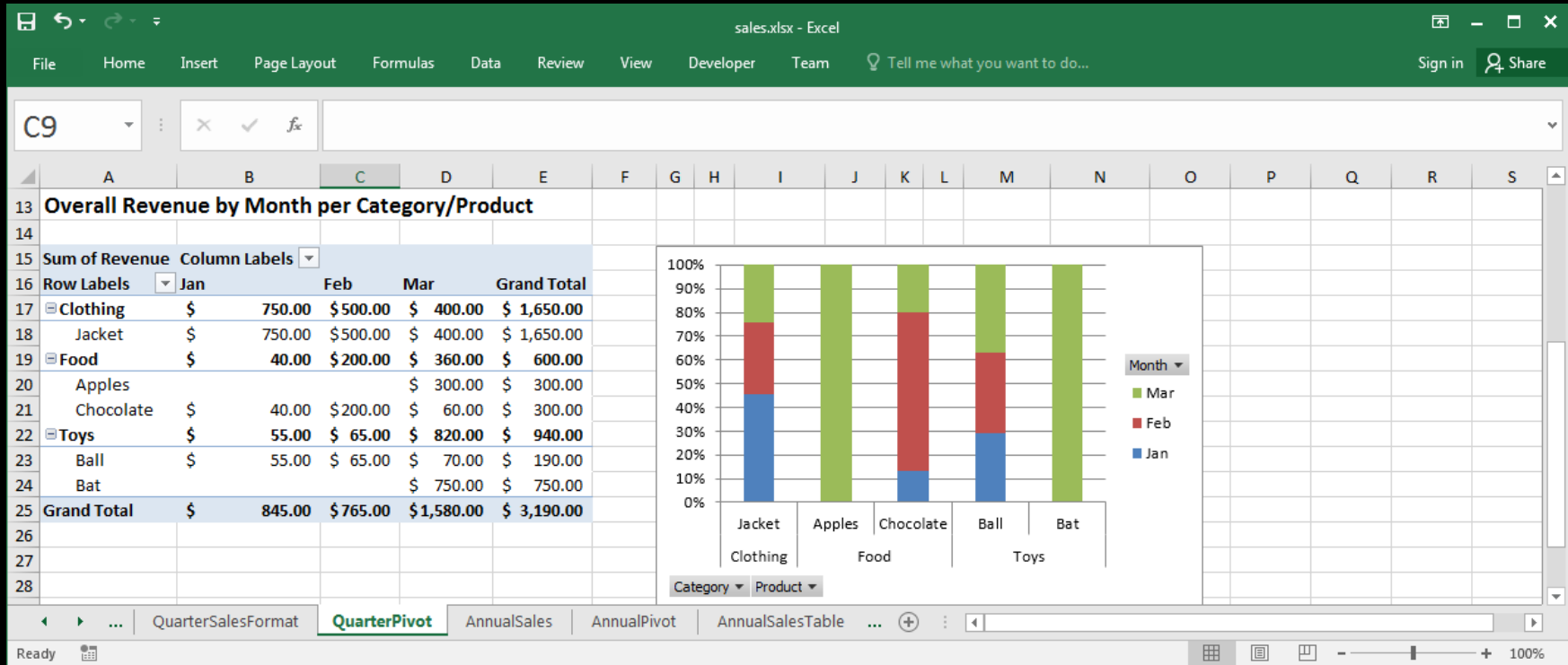
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Overall Revenue by Month per Category/Product														
2															
3	Sum of Revenue	Column Labels													
4	Row Labels	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Grand Total	
5	Clothing	\$ 750.00	\$ 500.00	\$ 400.00	\$ 250.00	\$ 100.00	\$ 800.00	\$ 1,800.00	\$ 3,000.00	\$ 400.00	\$ 500.00	\$ 1,500.00	\$ 2,500.00	\$ 12,500.00	
6	Jacket	\$ 750.00	\$ 500.00	\$ 400.00	\$ 250.00	\$ 100.00					\$ 500.00	\$ 1,500.00	\$ 2,500.00	\$ 6,500.00	
7	Shorts						\$ 800.00	\$ 1,800.00	\$ 3,000.00	\$ 400.00				\$ 6,000.00	
8	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	
9	Apples			\$ 300.00	\$ 360.00	\$ 330.00	\$ 420.00	\$ 480.00	\$ 390.00	\$ 450.00	\$ 420.00	\$ 540.00	\$ 450.00	\$ 4,140.00	
10	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	
11	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	
12	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	
13	Bat			\$ 750.00	\$ 750.00	\$ 1,500.00	\$ 2,250.00							\$ 5,250.00	
14	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	
15															

Task pane: TryItFilter2 TryItChart TryItWhatIf TryItWhatIf2 TryItPivot

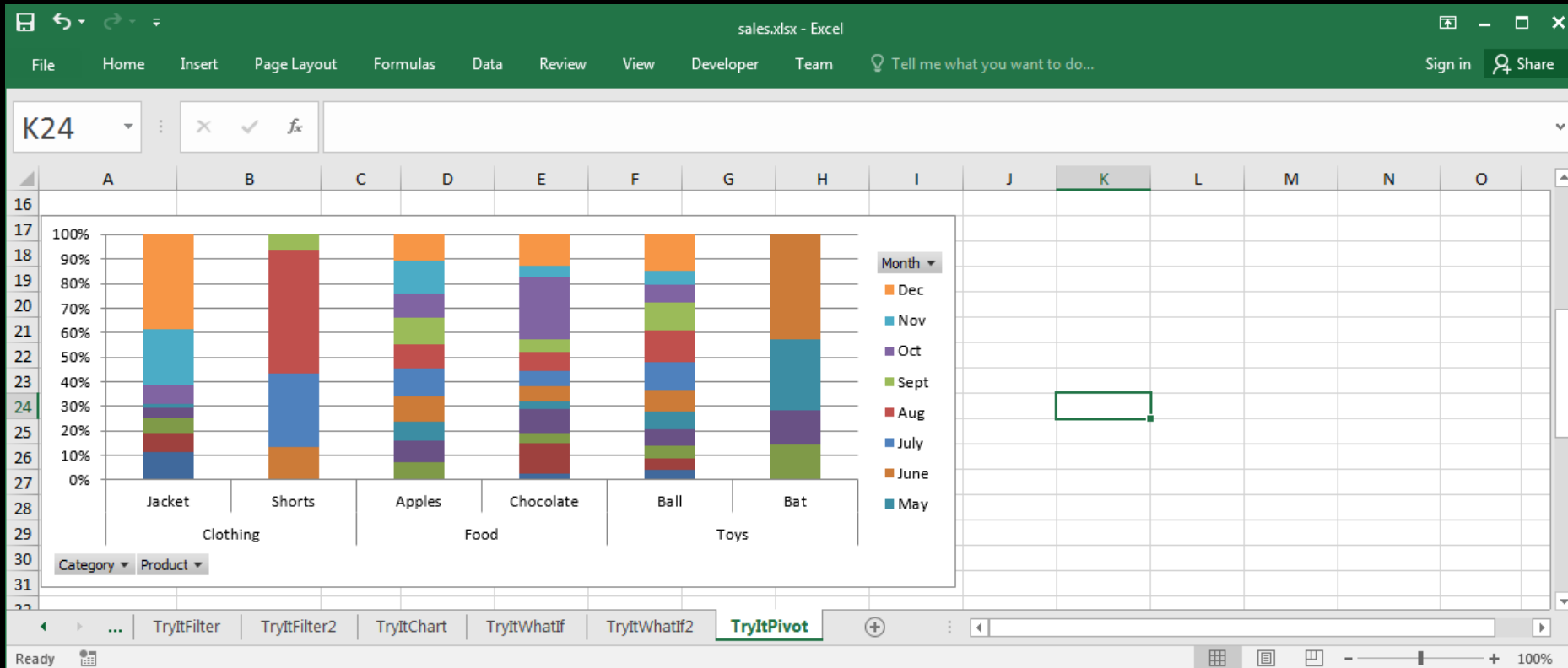
Status bar: Ready 100%

Pivot Charts

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



Question: Create a pivot chart for previous pivot table.



What-if and Pivot Tables Question

Question: choose the correct statements?

- 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 listed 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

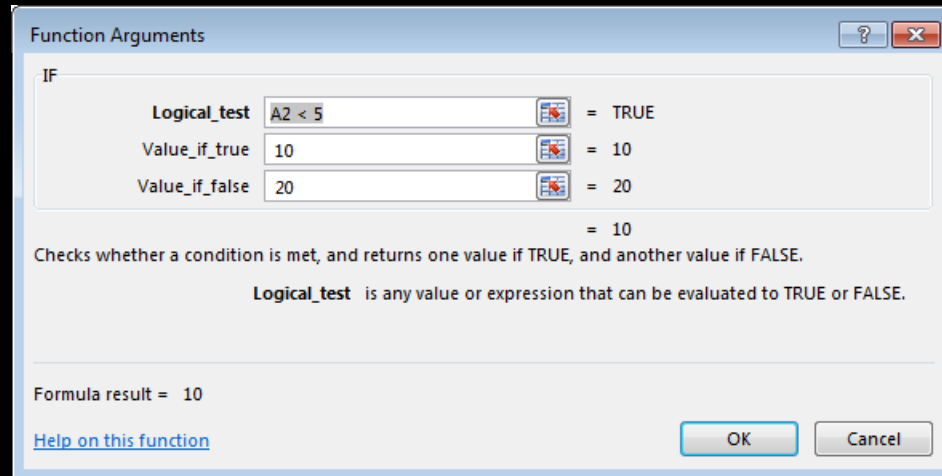
Decisions using IF ()

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

- `IF(condition, value_if_true, value_if_false)`

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

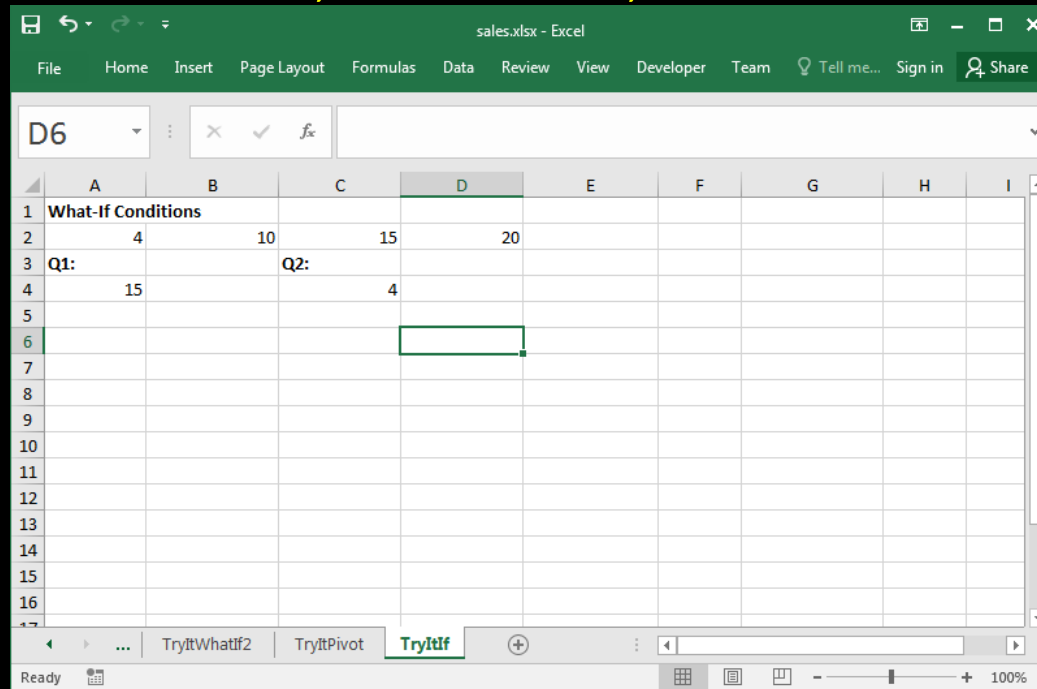
= IF(A2 < 5, 10, 20)



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.



Decisions using IF () Question

Question: How many True statements: **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

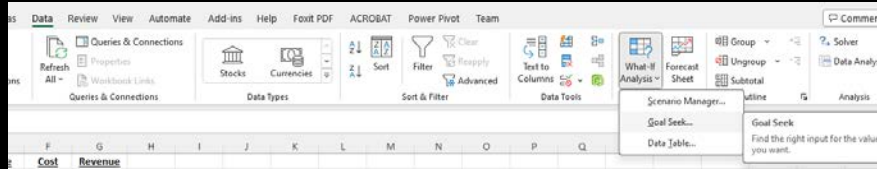
D) 3

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

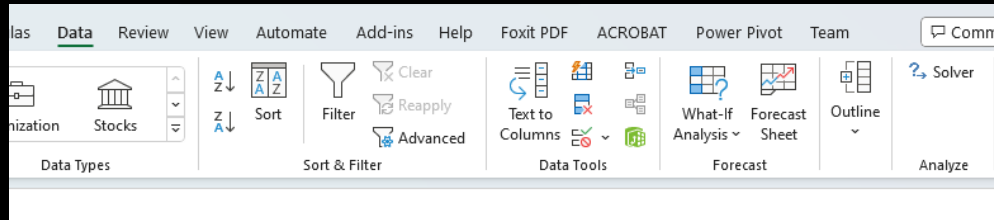


This screenshot shows the Excel spreadsheet with a table of product sales data. The 'Goal Seek' dialog box is open, showing the 'Set cell' as 'G13' and the 'To value' as '4000'. The 'By changing cell' is 'D4'.

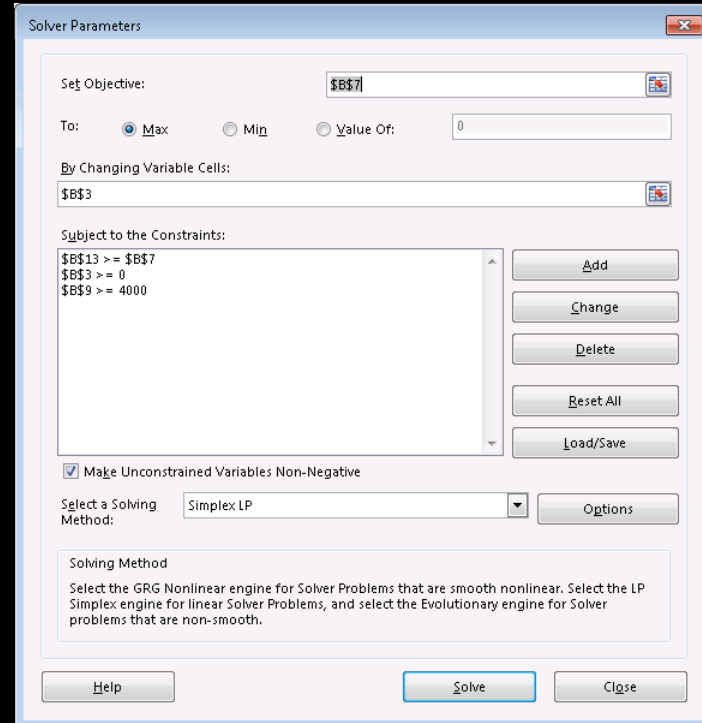
Category	Product	Month	Volume	Price	Cost	Revenue
Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00
Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00
Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00
Food	Chocolate	Feb	80	\$ 2.50	\$ 1.00	\$ 200.00
Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00
Toys	Ball	Feb	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
Total:						\$ 3,190.00

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



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 (Windows)
- To install: Tools → Excel Add-Ins (Mac)
- 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.

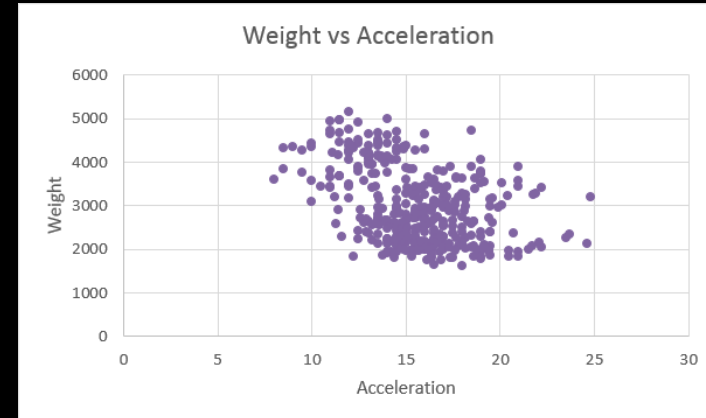
Regression in Excel

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.



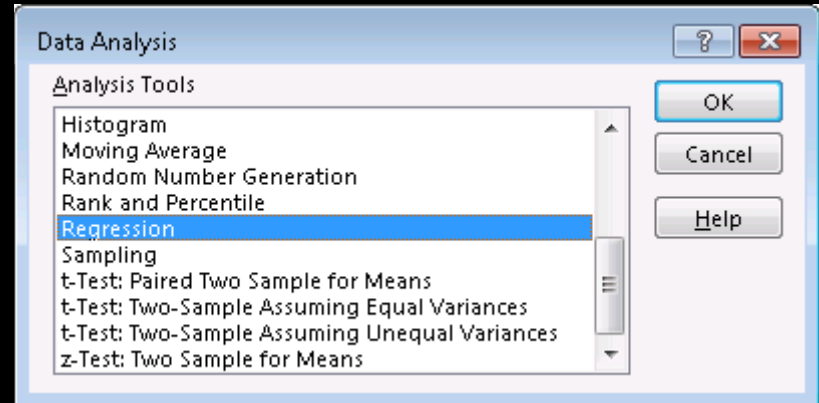
Regression Example

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.

Regression ? x

Input


Input Y Range: 

Input X Range: 

☐ Labels ☐ Constant is Zero

☐ Confidence Level: %

Output options

☐ Output Range: 

☒ New Worksheet Ply:

☐ New Workbook

Residuals

☒ Residuals ☒ Residual Plots

☒ Standardized Residuals ☐ Line Fit Plots

Normal Probability

☐ Normal Probability Plots

OK Cancel Help

Regression Example Results

1	SUMMARY OUTPUT								
2									
3	Regression Statistics								
4	Multiple R	0.416839282							
5	R Square	0.17375492							
6	Adjusted R Square	0.171636343							
7	Standard Error	2.510965983							
8	Observations	392							
9									
10	ANOVA								
11		df	SS	MS	F	Significance F			
12	Regression	1	517.0999442	517.0999442	82.01491373	6.56562E-18			
13	Residual	390	2458.930566	6.304950169					
14	Total	391	2976.03051						
15									
16		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
17	Intercept	19.57266581	0.462860061	42.28635703	9.5445E-148	18.66265269	20.48267893	18.66265269	20.48267893
18	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

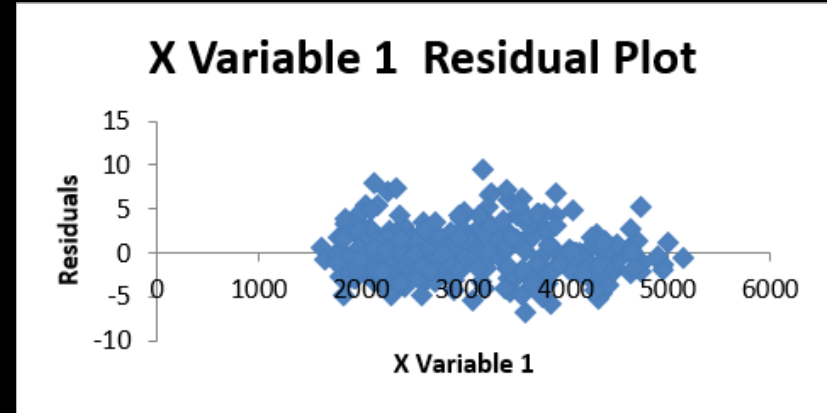
All of the output is put into a new sheet. Read the values off of the table and form the regression equation:

- 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

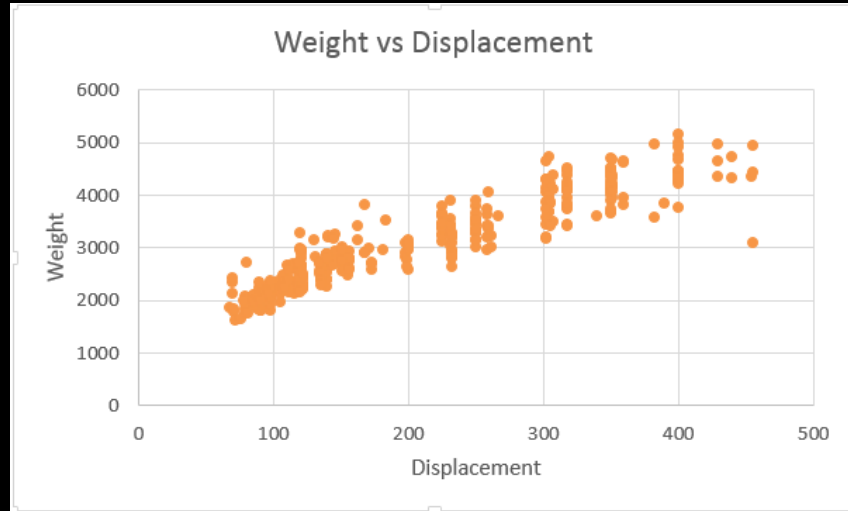


Expected a weak relationship and this is demonstrated by the R^2 value.

- 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.



$$R^2 = 0.870$$

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.
- Use the Analysis ToolPak including computing a regression.