

Excel Data Analysis

DATA 301



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 an Excel spreadsheet with the following data:

	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 Sort dialog box is open, showing the following settings:

- My data has headers:
- Sort by: Month (Values, Jan, Feb, Mar, Apr, May, Jun, Jul)
- Then by: Category (Values, A to Z)

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 spreadsheet contains 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	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.

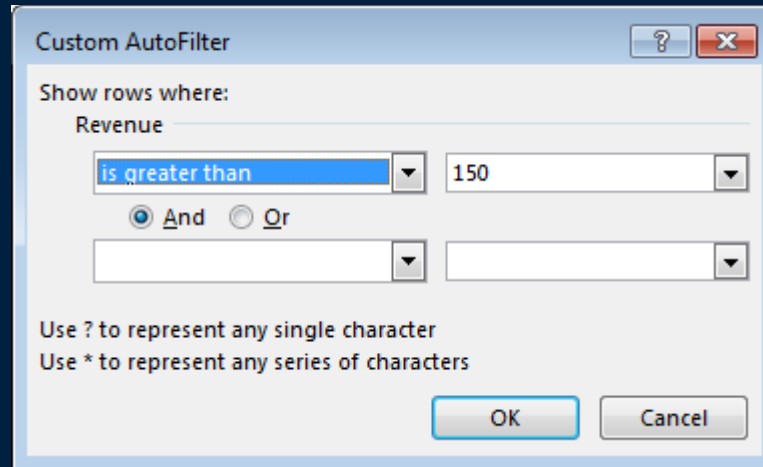
The screenshot shows an Excel spreadsheet with the following data:

1	Category	Product	Month	Volume	Price	Cost	Revenue
5	Food	Chocolate	Feb				
6	Clothing	Jacket	Feb				
12	Food	Apples	Mar				

The 'Revenue' column header is selected, and the 'Number Filters' menu is open. The 'Top 10...' option is highlighted. The 'Custom Filter...' option is also visible. The status bar at the bottom indicates 'Ready 3 of 12 records found'.

Custom Filter Example

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



The screenshot shows a 'Custom AutoFilter' dialog box. The title bar contains a question mark icon and a close button. The main area is titled 'Show rows where:' and contains the following elements:

- The column name 'Revenue' is displayed above a dropdown menu.
- The dropdown menu is open, showing 'is greater than' selected.
- To the right of the dropdown is a text input field containing the value '150'.
- Below these are two radio buttons: 'And' (which is selected) and 'Or'.
- Below the radio buttons are two empty dropdown menus.
- At the bottom, there are two buttons: 'OK' and 'Cancel'.
- Below the buttons, there are two lines of text: 'Use ? to represent any single character' and 'Use * to represent any series of characters'.

Custom Filter Result

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

	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					Total:		\$ 3,190.00				
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 spreadsheet displays a table with the following data:

	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 of the window shows '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.

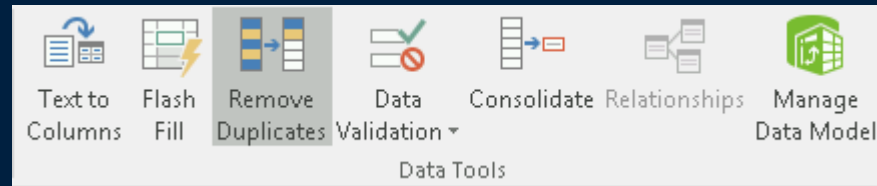
The screenshot shows an Excel spreadsheet with the following data table:

Category	Product	Month	Volume	Price	Cost	Revenue
Toys	Bat	Mar	10	\$ 75.00	\$ 50.00	\$ 750.00
Clothing	Jacket	Jan	15	\$ 50.00	\$ 35.00	\$ 750.00
Clothing	Jacket	Feb	10	\$ 50.00	\$ 35.00	\$ 500.00
Clothing	Jacket	Mar	8	\$ 50.00	\$ 30.00	\$ 400.00
Food	Chocolate	Mar	30	\$ 2.00	\$ 1.00	\$ 60.00
Toys	Ball	Jan	55	\$ 1.00	\$ 0.50	\$ 55.00
Food	Chocolate	Jan	20	\$ 2.00	\$ 1.00	\$ 40.00

The Excel interface shows the 'Data' tab with the 'Filter' button active. The filter criteria are: Volume < 20 and Revenue < 65. The status bar at the bottom indicates '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	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

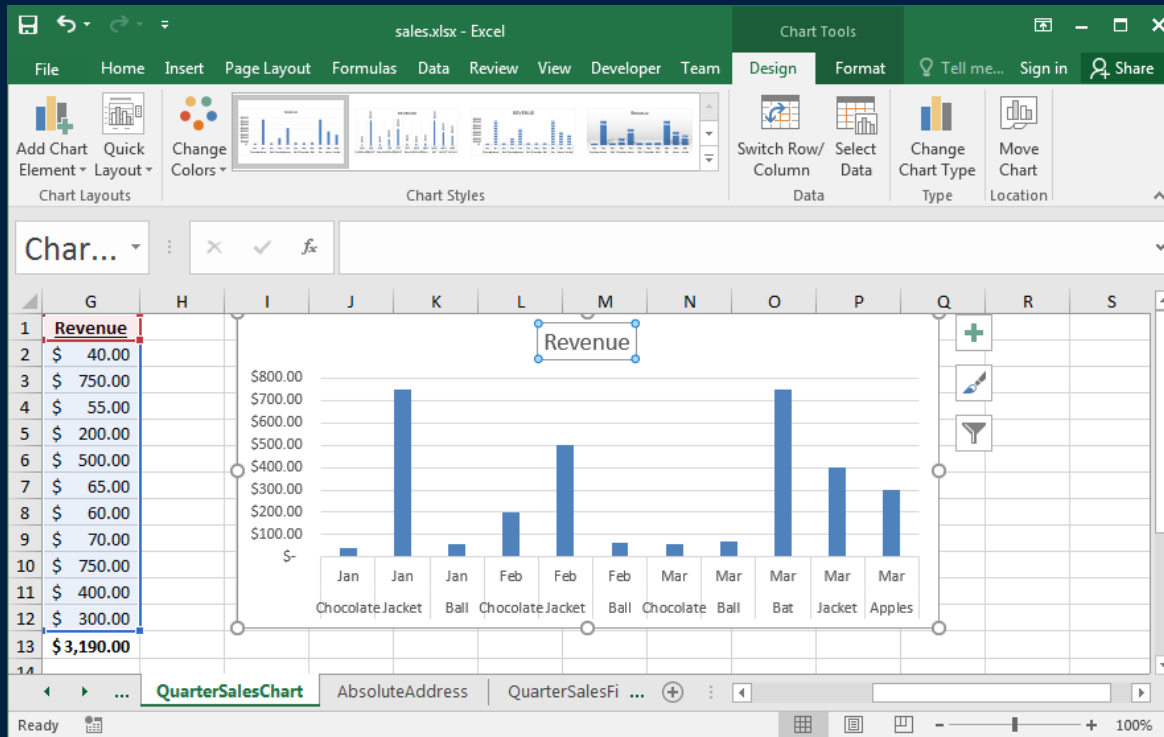
Select **Insert**, then click **Chart** Icon, and pick the chart type.

The screenshot displays the Microsoft Excel interface with the 'Insert Chart' dialog box open. The 'Column' chart type is selected, and a 'Clustered Column' chart is previewed. The background shows an Excel spreadsheet with sales data for various products over three months.

	A	B	C	D	E	F	G	H	I
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		

Chart Options

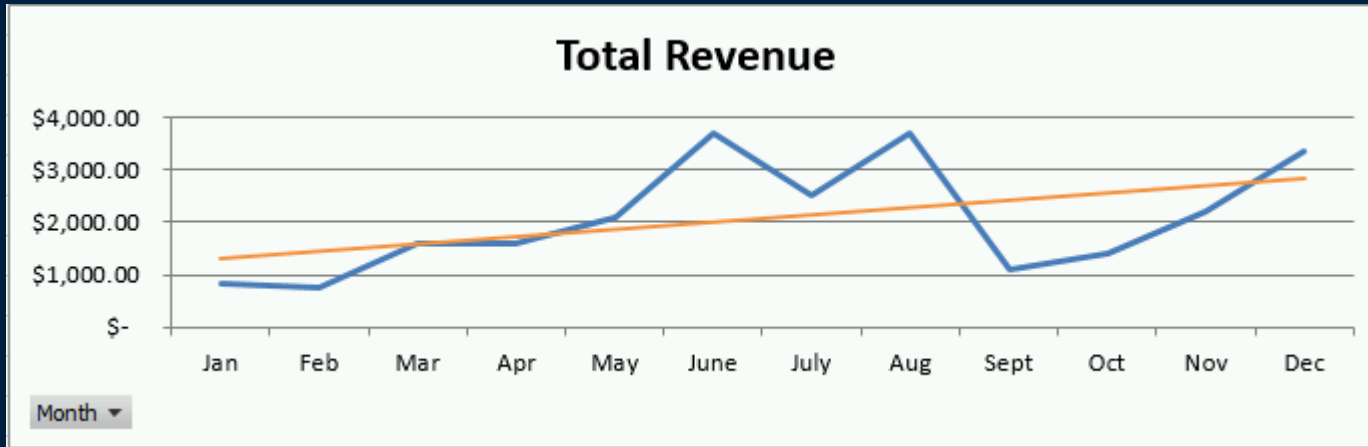
Chart Tools allows you to modify the data in the chart, change the chart type, and move the chart in the Worksheet.



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

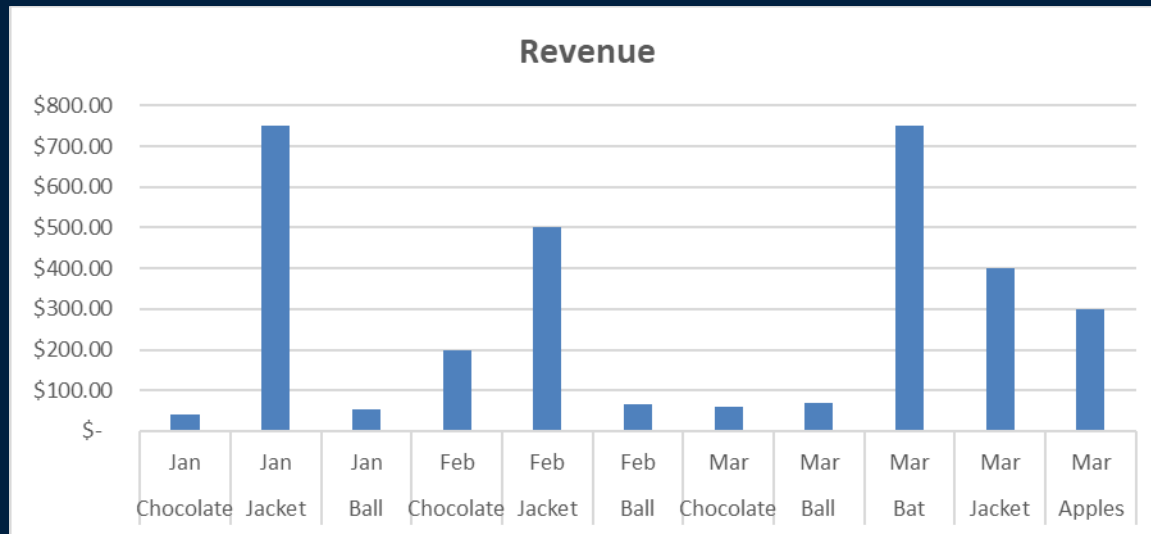
Set Intercept

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 per month.



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 active. The ribbon includes options for **Edit Data**, **Type** (Line, Column, Win/Loss), and **Show** (High Point, Low Point, Negative Points, First Point, Last Point, Markers). It also features **Style** options for Sparkline Color and Marker Color, and **Group** options (Group, Ungroup, Clear).

The worksheet displays a table of monthly revenue data. A sparkline chart is embedded in cell O5, showing a line chart of the monthly revenue data. The chart is labeled "Monthly Sales".

Overall Revenue by Month	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

What-If

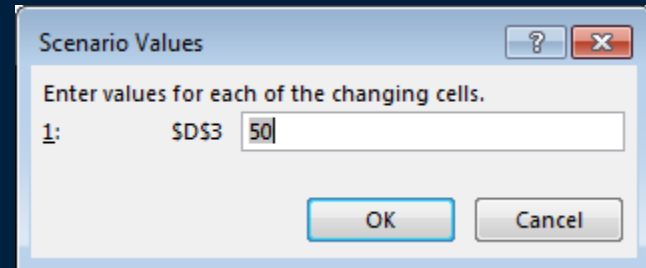
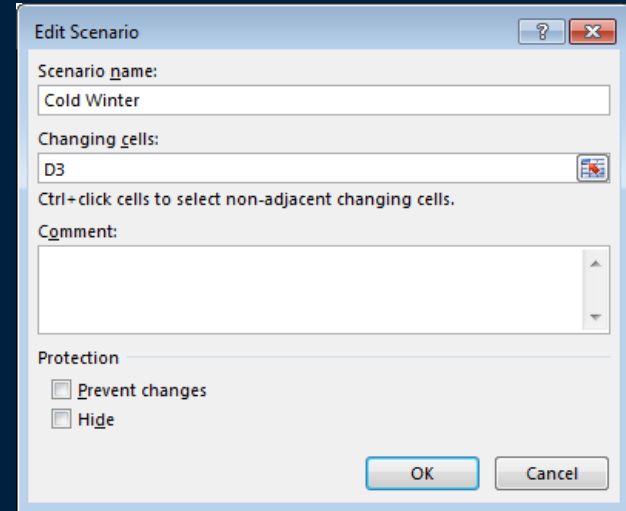
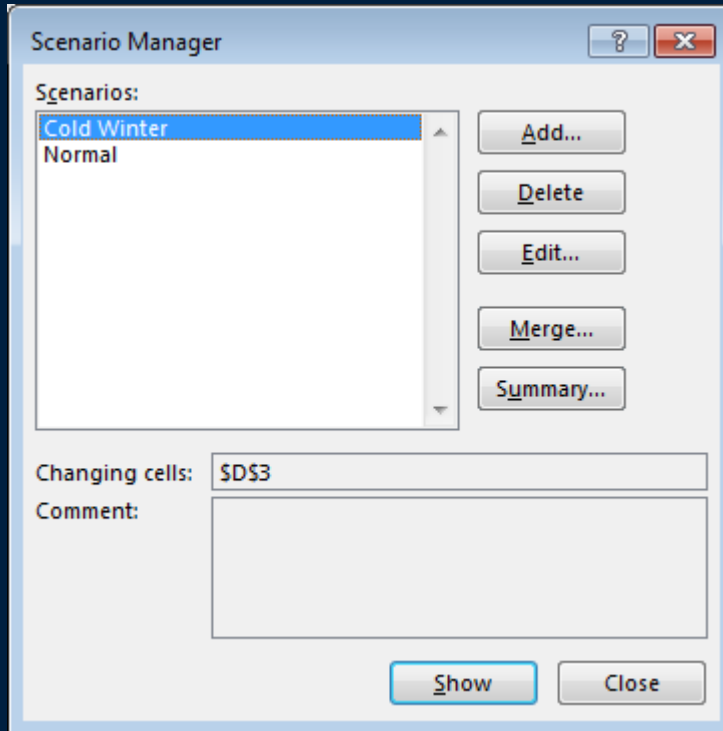
What-If scenarios help understand different possibilities.

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**.



What-If Scenarios Example

User can easily select scenario and see the result.

The screenshot displays the Microsoft Excel interface with the 'Data' tab selected. The main window shows a spreadsheet with columns for Category, Product, Month, Volume, Price, Cost, and Revenue. The 'Volume' cell in row 3 (D3) is highlighted with the value 50. The Scenario Manager dialog box is open, showing a list of scenarios: 'Cold Winter' (selected) and 'Normal'. The 'Changing cells' field is set to '\$D\$3' and the 'Comment' field is empty. The 'Show' button is highlighted.

	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			
14										

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 the Microsoft Excel interface with the 'Data' tab selected. The ribbon includes options for 'Sort & Filter', 'Data Tools', and 'Forecast'. The 'What-If Analysis' button is highlighted. Below the ribbon, the formula bar shows 'H13'. 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	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					

Try it: What-If Scenario Challenge

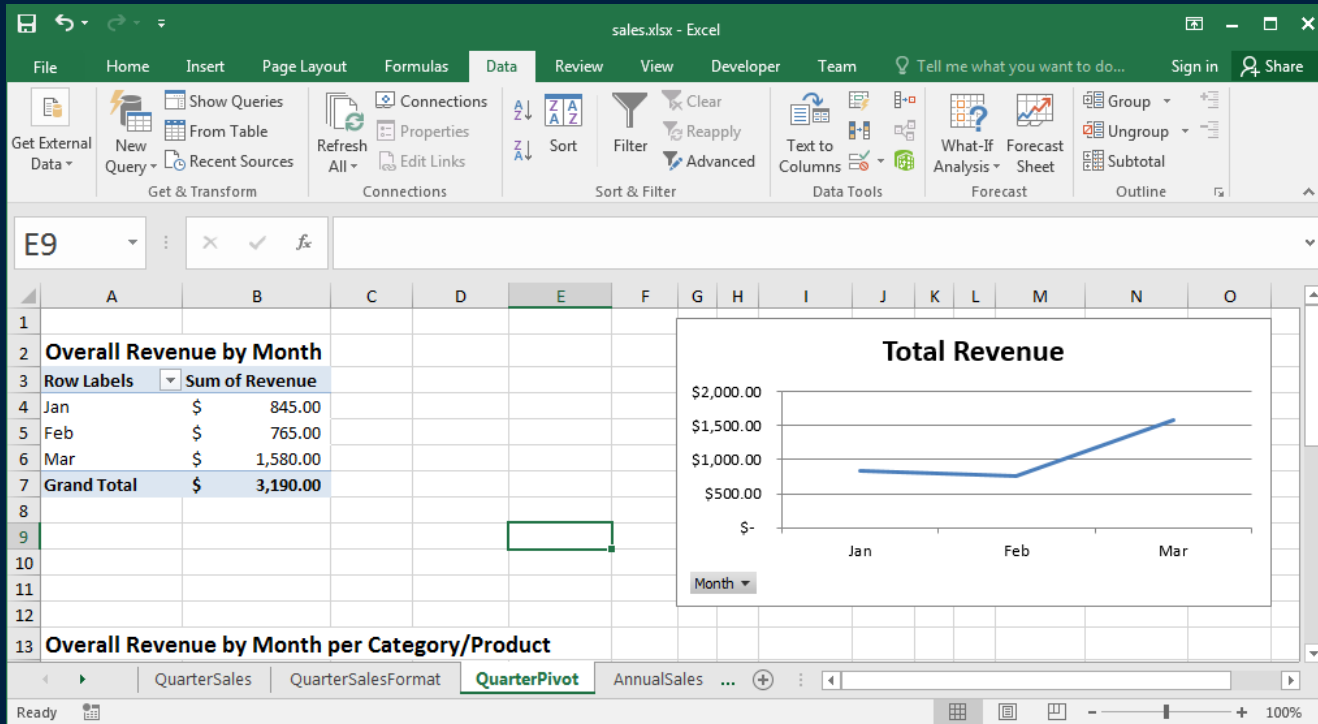
Question: Create a what-if scenario that all costs go up by 10% and volume down by 20%.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Category	Product	Month	Volume	Act. Vol.	Price	Cost	Act. Cost	Revenue				
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									Total:				\$2,552.00

★ 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.

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

Create PivotTable

Choose the data that you want to analyze

- S**elect a table or range

Table/Range:
- U**se an external data source

Connection name:
- U**se this workbook's Data Model

Choose where you want the PivotTable report to be placed

- N**ew Worksheet
- E**xisting Worksheet

Location:

Choose whether you want to analyze multiple tables

- A**dd this data to the Data Model

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 cells 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:**
 - 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

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 titled "Sum of Revenue" and is located in the range A3:E9. The PivotTable Fields task pane is open on the right side of the screen, showing 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...

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

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

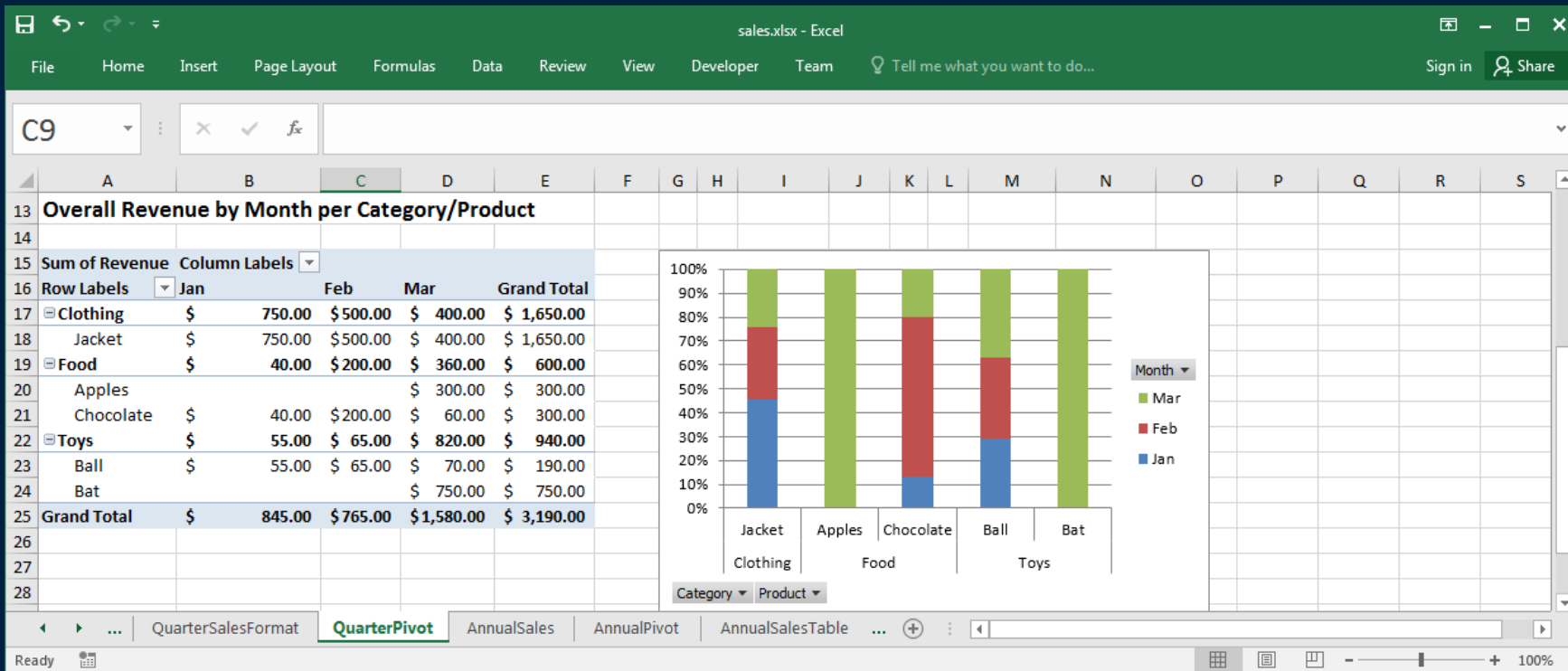
Ready TryItFilter2 TryItChart TryItWhatif TryItWhatif2 TryItPivot 100%



Pivot Charts

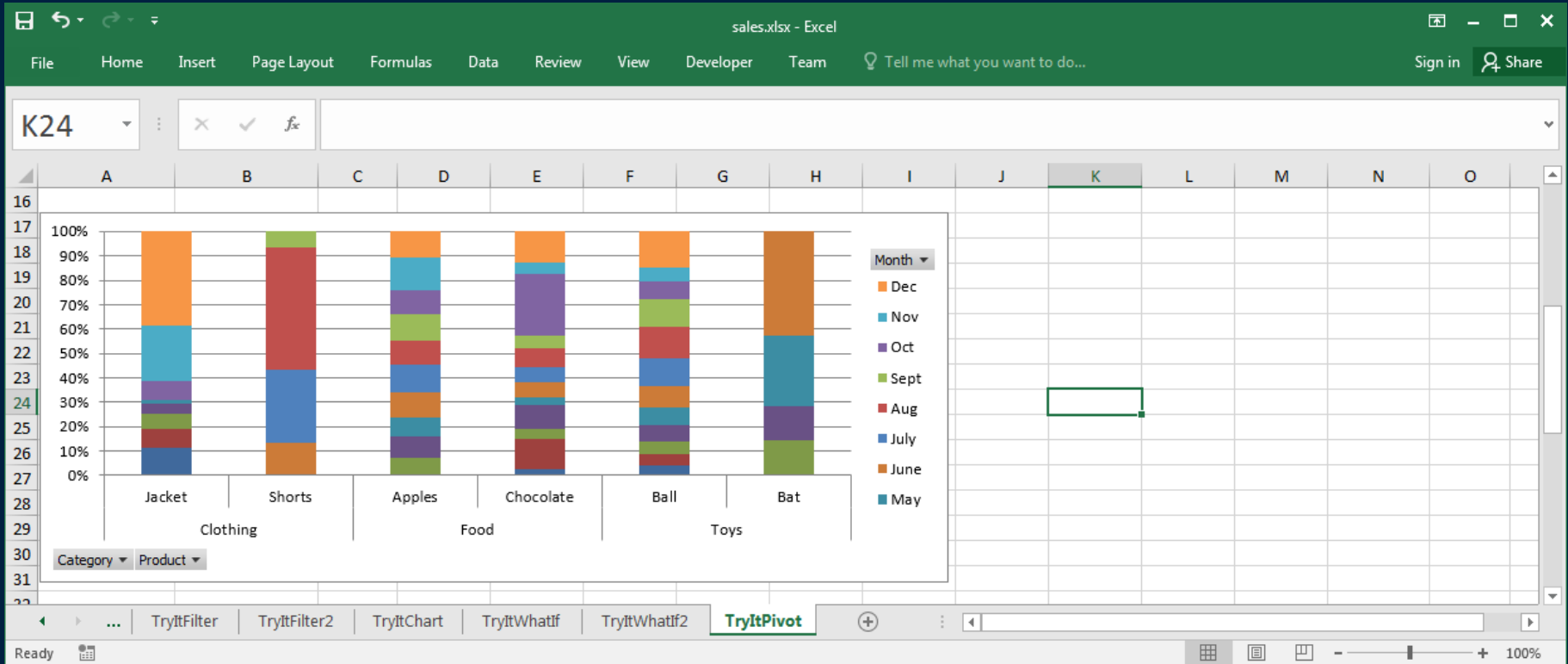


A **pivot chart** is a chart attached to a pivot table. Create it under Insert then Pivot Chart.



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 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**

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:

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
14							

The Goal Seek dialog box is open, showing:

- Set cell: G13
- To value: (empty)
- 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	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	Amoritization 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

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 (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.

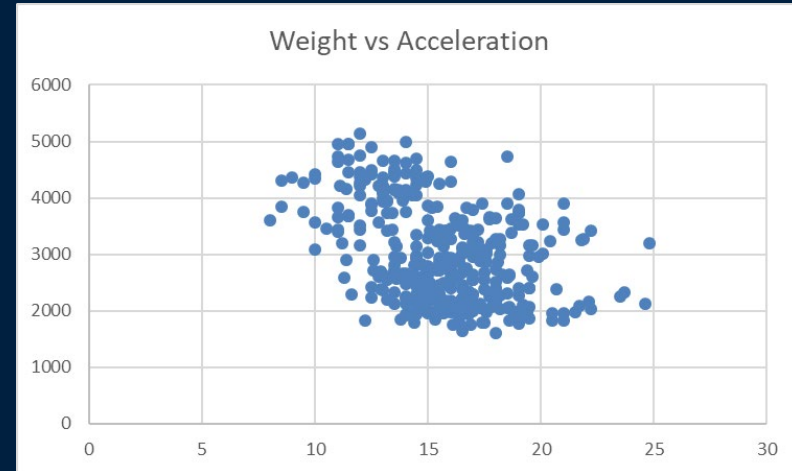
Regression in Excel

Excel provides a regression function that will calculate:

- R²
- 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.



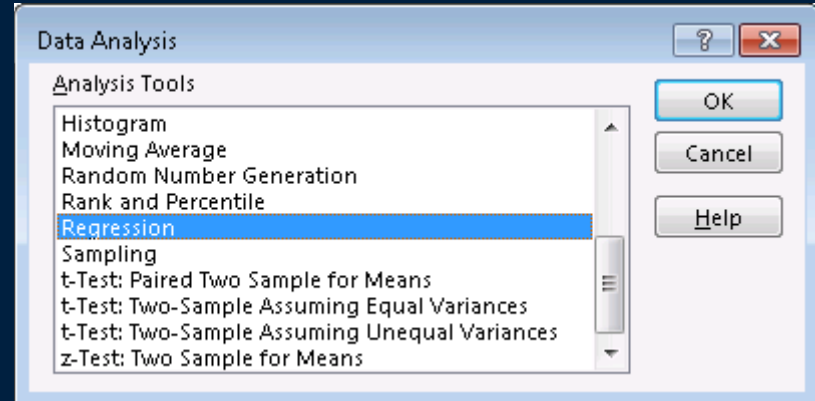
Regression Example

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.

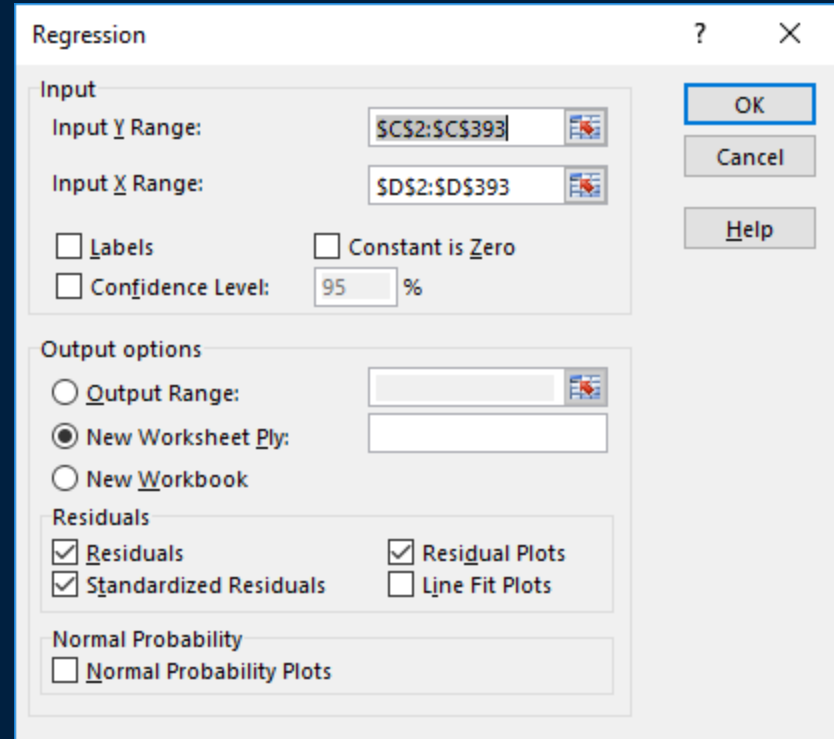
To start select, Data Analysis from the data tab and then select Regression and OK.



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

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

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	Regression Statistics								
4	Multiple R	0.416839202							
5	R Square	0.17375492							
6	Adjusted R Square	0.171036343							
7	Standard Error	773.0793697							
8	Observations	392							
9									
10	ANOVA								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
12	Regression	1	49016353.59	49016354	82.01491	6.56562E-18			
13	Residual	390	233084167.6	597651.7					
14	Total	391	282100521.2						
15									
16		<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>
17	Intercept	4972.111073	223.6731365	22.22936	2.61E-71	4532.355074	5411.867073	4532.355074	5411.867073
18	X Variable 1	-128.3369786	14.17115976	-9.05621	6.57E-18	-156.1984045	-100.4755526	-156.1984045	-100.4755526

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

Coefficients for the regression equation

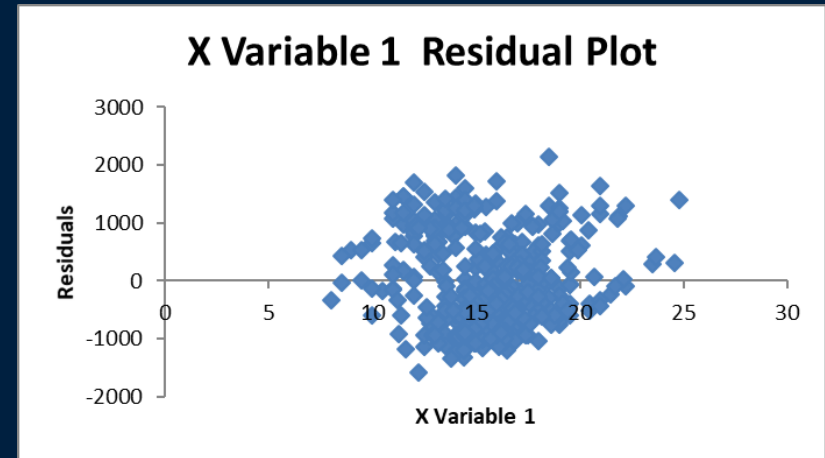
All of the output is put into a new sheet. Read the values off of the table and form the regression equation:

- weight = $-128.34 * \text{acceleration} + 4972.11$

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.

22	RESIDUAL OUTPUT			
23				
24	<i>Observation</i>	<i>Predicted Y</i>	<i>Residuals</i>	<i>Standard Residuals</i>
25	1	3432.067331	71.9326693	0.093166154
26	2	3496.23582	196.76418	0.254846123
27	3	3560.404309	-124.4043093	-0.161126664
28	4	3432.067331	0.932669301	0.00120798
29	5	3624.572799	-175.5727985	-0.227399352
30	6	3688.741288	652.2587122	0.84479606
31	7	3817.078266	536.9217336	0.695413272
32	8	3881.246756	430.7532443	0.557905379
33	9	3688.741288	736.2587122	0.953591647
34	10	3881.246756	-31.24675565	-0.040470346
35	11	3688.741288	-125.7412878	-0.162858299

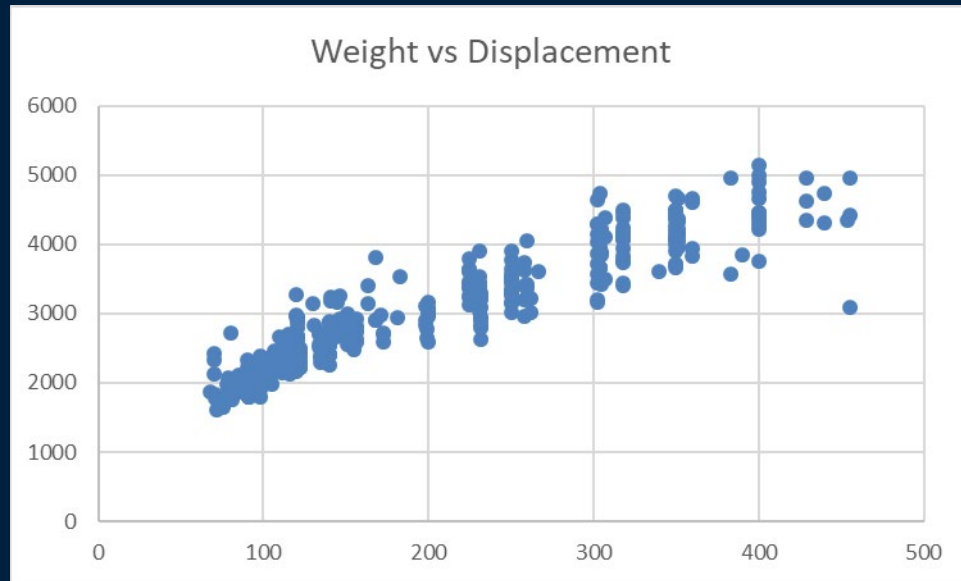


Expected a weak relationship and this is demonstrated by the R^2 value.

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

Pivot tables and charts are useful for data aggregation and filtering and displaying results.

Analysis ToolPak contains tools for what-if scenarios, goal seek, linear solvers, and statistical analysis tools.

Objectives

- 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.



THE UNIVERSITY OF BRITISH COLUMBIA

