SEPTEMBER 2022

TEXAS
The University of Texas at Austin

EXCEL WORKSHOP

Statistical Software in Public Policy Workshop Series

ALFONSO ROJAS-ÁLVAREZ, PH.D.
LBJ School of Public Affairs, The University of Texas at Austin

Welcome!

- Third of six statistical software workshops in the Fall:
  - September 9: Introduction to Stata in public policy
  - September 16: Stata lab
  - **September 23: Introduction to Excel in public policy**
  - September 30: Excel lab
  - October 14: Introduction to R in public policy
  - October 21: R lab

Fridays, from 12:15 to 1:45pm, in SRH 3.312/3.360

Contact: arojas@austin.utexas.edu | (512-552-9860)

Say the class/subject upfront!

About these workshops

- Workshops are *à la carte*; come to any you choose based on your needs.
- There is no evaluation, or required participation.
- The labs will usually build upon the content of the intro session.
- We assume no previous exposure to the software.
- We won't be going over installation of the software.
- 1 hour of lecture and 10-15 minutes of Q&A and troubleshooting
- Feel free to ask questions whenever!

Office Hours for technical, software or installation assistance (or anything else):

Monday 10-11:30am; Thursday 2-3:30pm, in **SRH 3.264.**

Introduction to Excel in Public Policy

Contents for today's workshop:

1. What is Excel
2. First steps on using Excel – sample data set
3. Using functions
4. Creating PivotTables
5. PowerQuery
6. Power Pivot

Contents at: https://github.com/LBJ-SoftwareWorkshops/excel_workshop
https://github.com/alf10087/excel_workshop

BASIC FUNCTIONS

SUM

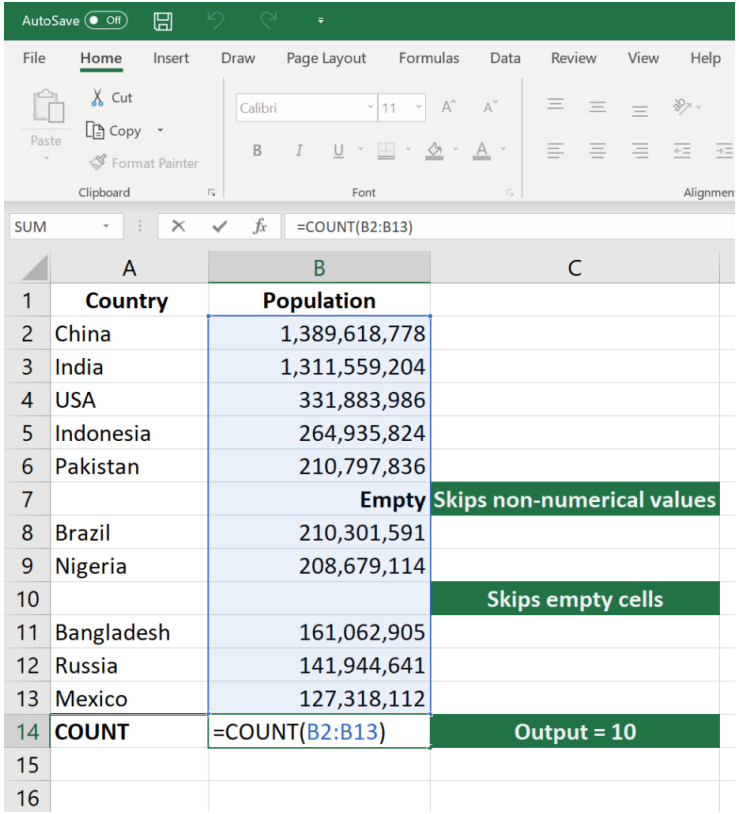
- This function will sum up a range of cells
- Example:
- **=SUM(B2:G2)** – A simple selection that sums the values of a row.
- **=SUM(A2:A8)** – A simple selection that sums the values of a column.
- **=SUM(A2:A7, A9, A12:A15)** – A sophisticated collection that sums values from range A2 to A7, skips A8, adds A9, jumps A10 and A11, then finally adds from A12 to A15.
- **=SUM(A2:A8)/20** – Shows you can also turn your function into a formula.

Average

- This functions tells you the average value of a range
- Example:
- **=AVERAGE(B2:B11)** – Shows a simple average, also similar to (SUM(B2:B11)/10)

COUNT

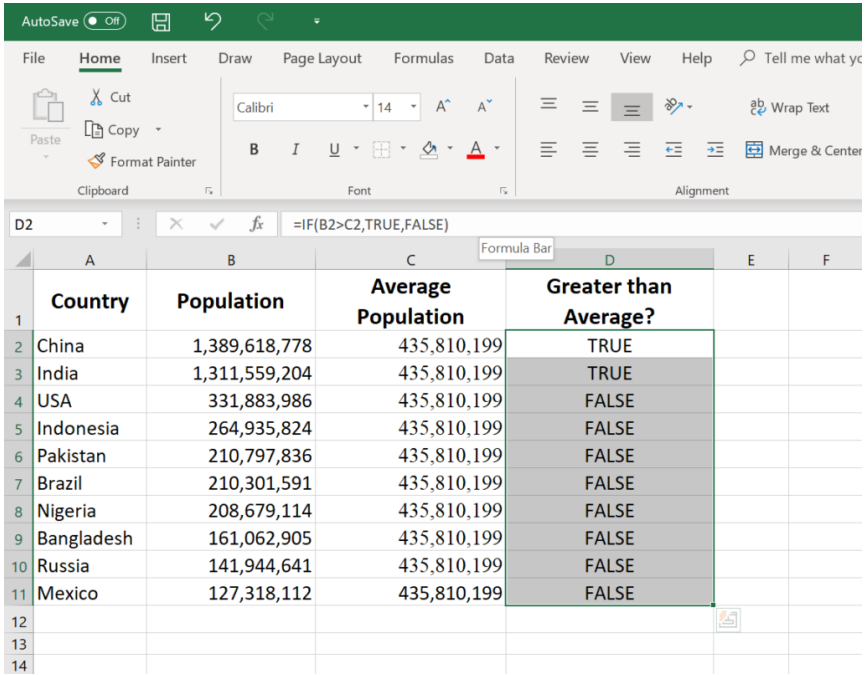
- The COUNT function counts all cells in a given range that contain only numeric values.
- Example:
- **=COUNT(A:A)** – Counts all values that are numerical in A column. However, you must adjust the range inside the formula to count rows.
- **=COUNT(A1:C1)** – Now it can count rows.



	A	B	C
1	**Country**	**Population**	
2	China	1,389,618,778	
3	India	1,311,559,204	
4	USA	331,883,986	
5	Indonesia	264,935,824	
6	Pakistan	210,797,836	
7		**Empty**	**Skips non-numerical values**
8	Brazil	210,301,591	
9	Nigeria	208,679,114	
10			**Skips empty cells**
11	Bangladesh	161,062,905	
12	Russia	141,944,641	
13	Mexico	127,318,112	
14	**COUNT**	=COUNT(B2:B13)	**Output = 10**
15			
16			

IF

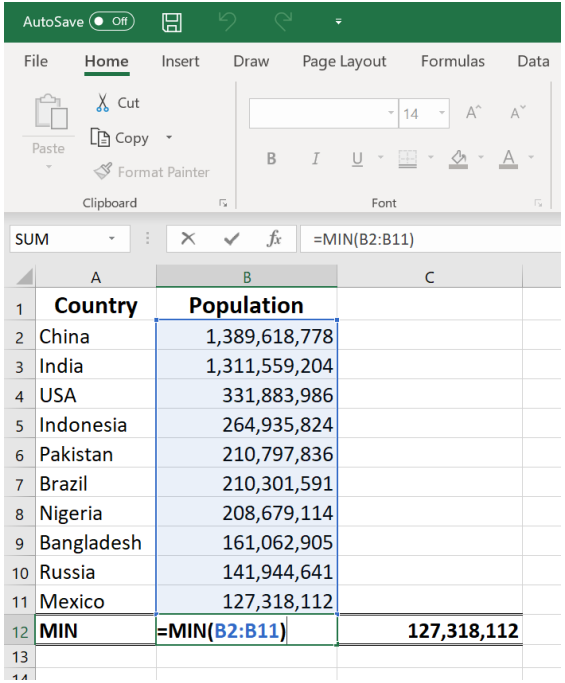
- The IF function is often used when you want to sort your data according to a given logic.
- Example:
- **=IF(C2<D3, 'TRUE,' 'FALSE')** – Checks if the value at C3 is less than the value at D3. If the logic is true, let the cell value be TRUE, else, FALSE
- **=IF(SUM(C1:C10) > SUM(D1:D10), SUM(C1:C10), SUM(D1:D10))** – An example of a complex IF logic. First, it sums *C1 to C10* and *D1 to D10*, then it compares the sum. If the sum of C1 to C10 is greater than the sum of *D1 to D10*, then it makes the value of a cell equal to the sum of *C1 to C10*. Otherwise, it makes it the SUM of *C1 to C10*.

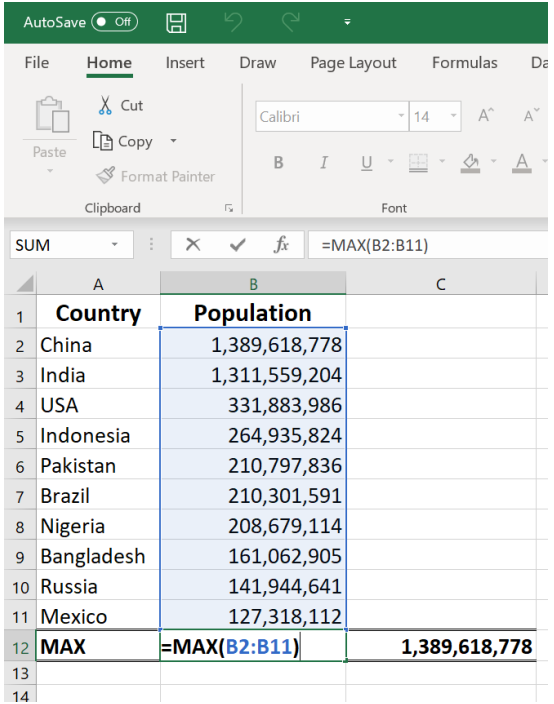


	A	B	C	D	E	F
1	Country	Population	Average Population	Greater than Average?		
2	China	1,389,618,778	435,810,199	TRUE		
3	India	1,311,559,204	435,810,199	TRUE		
4	USA	331,883,986	435,810,199	FALSE		
5	Indonesia	264,935,824	435,810,199	FALSE		
6	Pakistan	210,797,836	435,810,199	FALSE		
7	Brazil	210,301,591	435,810,199	FALSE		
8	Nigeria	208,679,114	435,810,199	FALSE		
9	Bangladesh	161,062,905	435,810,199	FALSE		
10	Russia	141,944,641	435,810,199	FALSE		
11	Mexico	127,318,112	435,810,199	FALSE		
12						
13						
14						

MAX & MIN

- The MAX and MIN functions tell you the maximum number and the minimum number in a range of values.
- Example:
- **=MIN(B2:C11)** – Finds the minimum number between column B from B2 and column C from C2 to row 11 in both columns B and C.
- **=MAX(B2:C11)** – Similarly, it finds the maximum number between column B from B2 and column C from C2 to row 11 in both columns B and C.



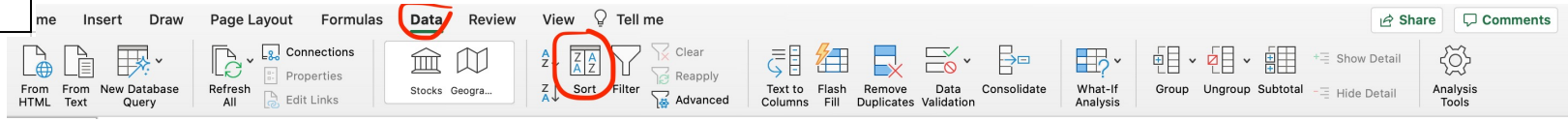


Sort Alphabetically

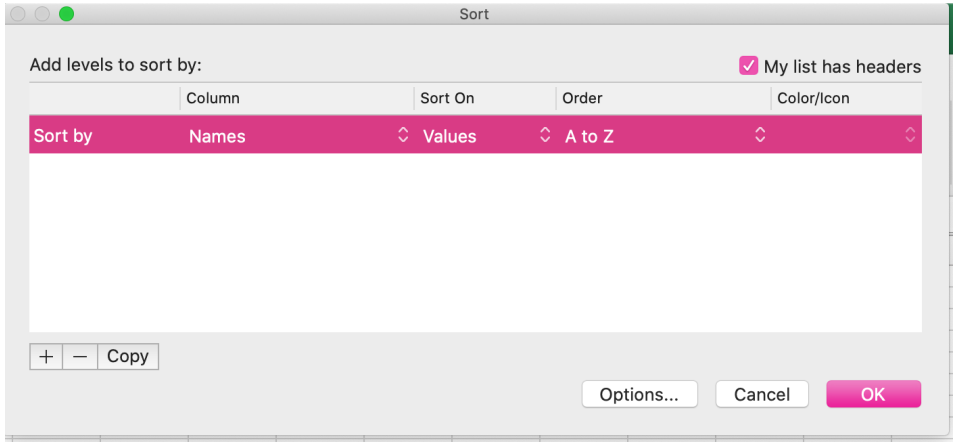
1

Names
Lolo
Bob
Zoya
Ruby
Sameer
Yessenia
Sonya
Noorah
Siraj
John
Jill
Jack

2



3



4

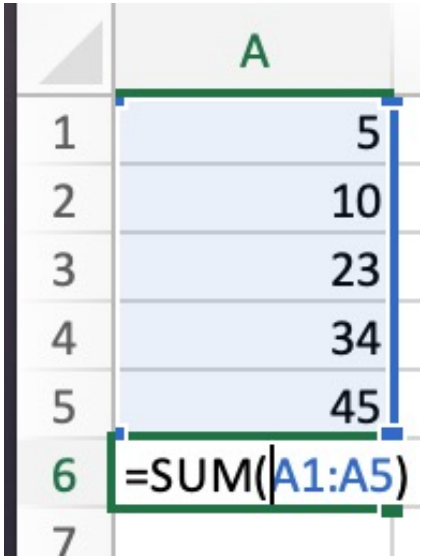
Names
Bob
Jack
Jill
John
Lolo
Noorah
Ruby
Sameer
Siraj
Sonya
Yessenia
Zoya

Select the cells you want to sort and click the sort button

HOW TO INSERT A FUNCTION

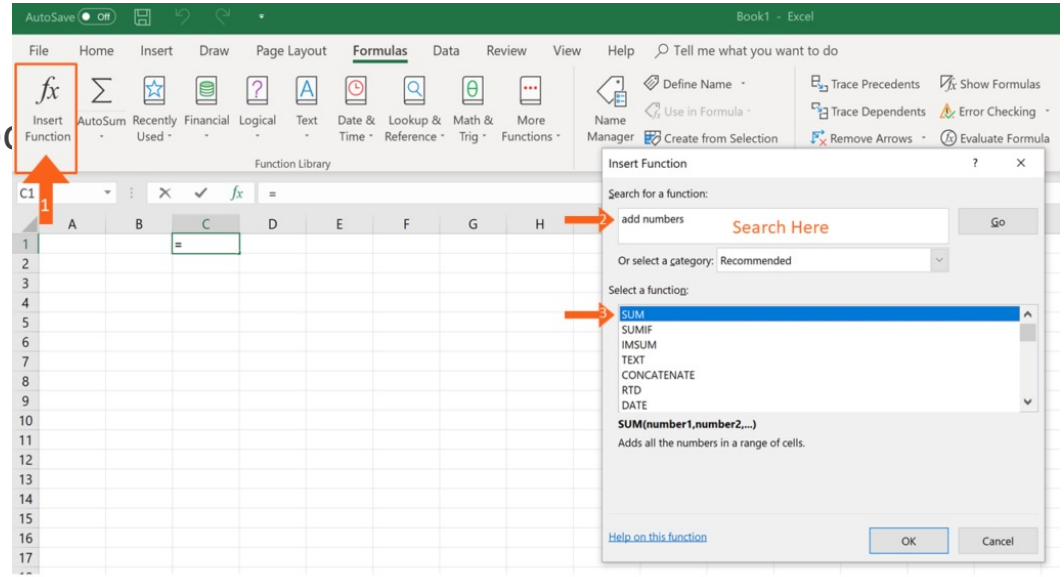
Option 1: Directly into a cell

- Always use the '=' **first**
- Name of function is **next**
- (starting cell : ending cell)
  - : means through i.e. range

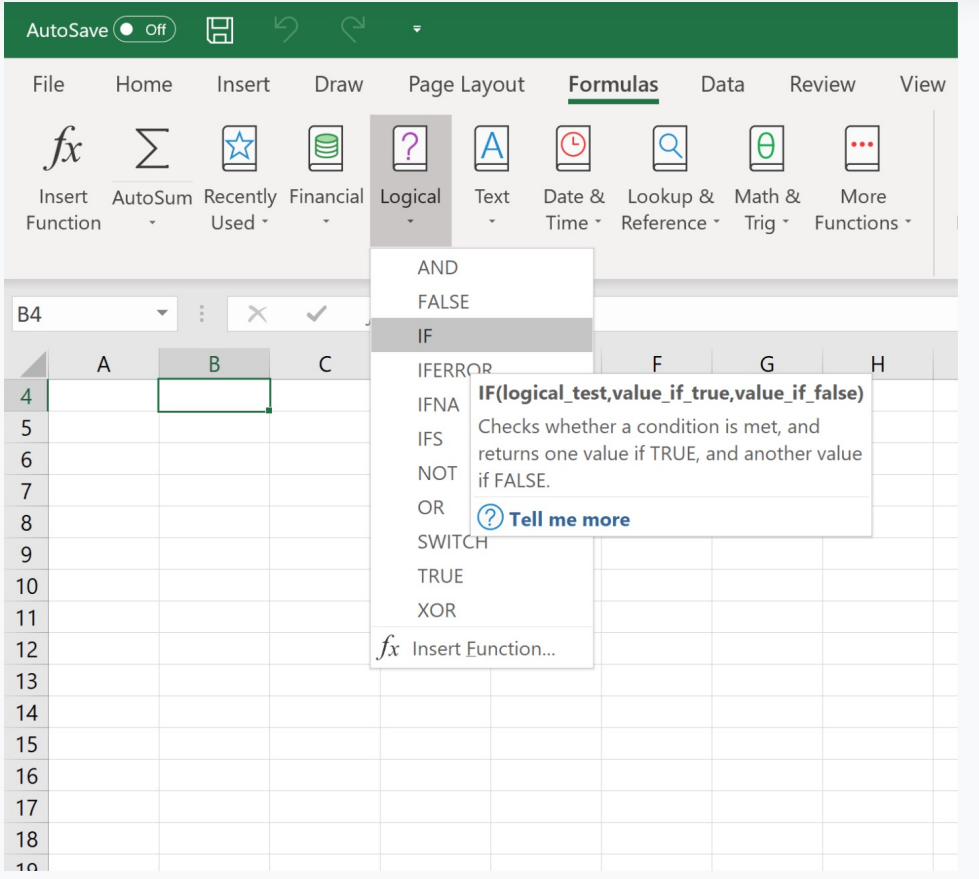


Option 2: Using the insert function

- Under formulas tab
- Click the 1st item on the ribbo
- Choose a function
- Click insert function
- Specify the range

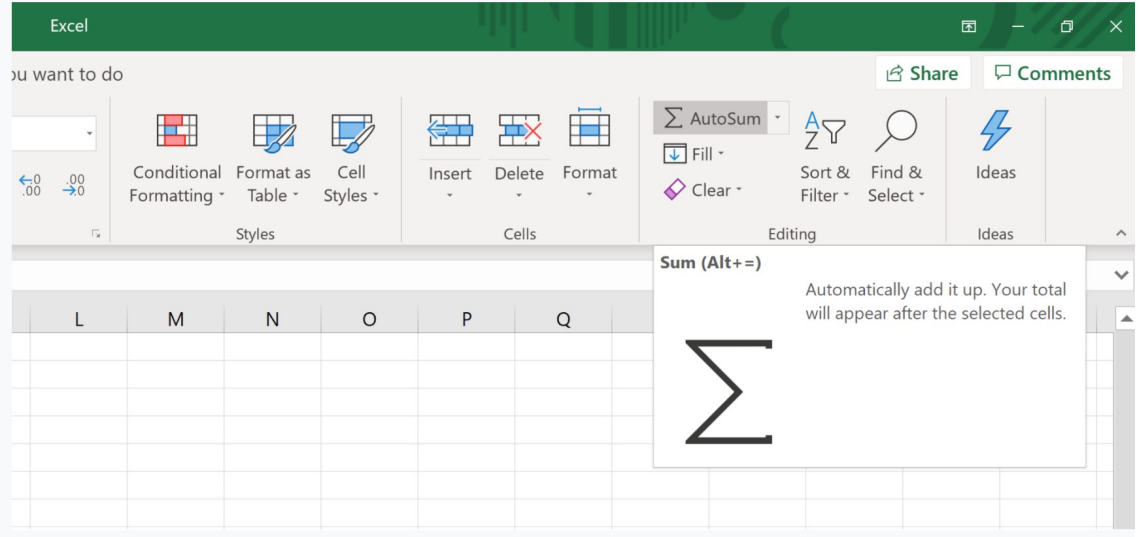


OPTION 3: Selecting a Formula from One of the Groups in Formula



Option 4: Using AutoSum Option

- Under the home tab



CONDITIONAL AGGREGATION

What does conditional Aggregation mean?

- **Data aggregation** represents a mathematical operation (summarization, counting, averaging) that we perform over a set of data.
- **Conditional aggregation** is performing data aggregation over a set of data that meets certain condition which is contained within a given data range.

EXAMPLE 1: SUMIF

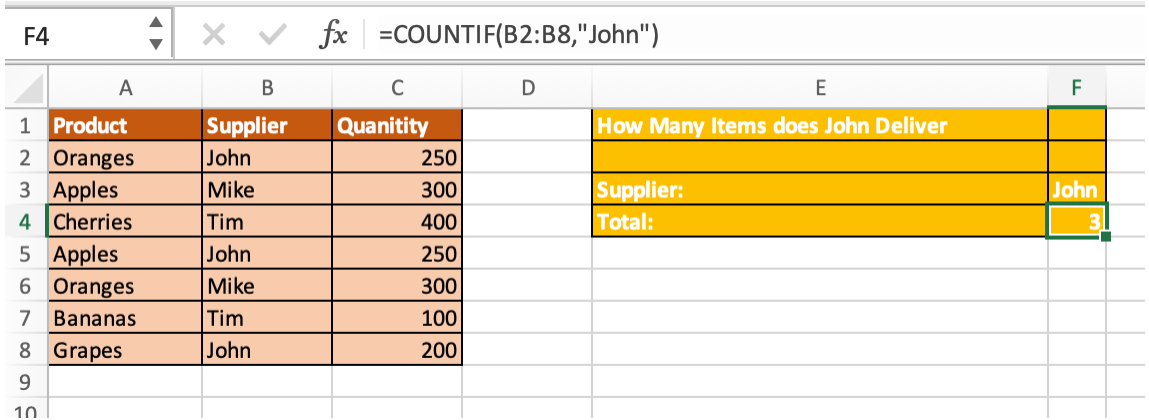
- To SUM up X value
- SUMIF (<range>, <criteria>,<sum range>)
  - Example below: To SUM up amount of products delivered by John

F4 | =SUMIF(B2:B8, "John",C2:C8)

	A	B	C	D	E	F
1	Product	Supplier	Quanitity		Total Amount of Products Delivered by John	
2	Oranges	John	250		Product:	Any
3	Apples	Mike	300		Supplier:	John
4	Cherries	Tim	400		Total:	700
5	Apples	John	250			
6	Oranges	Mike	300			
7	Bananas	Tim	100			
8	Apples	John	200			
9						
10						

EXAMPLE 2: COUNT IF

- To count X values
- COUNTIF (<range>, <criteria>)
  - To count how many items John delivers



	A	B	C	D	E	F
1	Product	Supplier	Quanitity		How Many Items does John Deliver	
2	Oranges	John	250			
3	Apples	Mike	300		Supplier:	John
4	Cherries	Tim	400		Total:	3
5	Apples	John	250			
6	Oranges	Mike	300			
7	Bananas	Tim	100			
8	Grapes	John	200			
9						
10						

PIVOT TABLES AND CHARTS

What is a pivot table?

- A pivot table allows you to extract the significance from a large, detailed data set.
- Example:
  - To summarize, sort, reorganize, group, count, total or average data
  - Filter data points i.e. who is assigned to do X task?
  - Sort Min to Max and vice versa

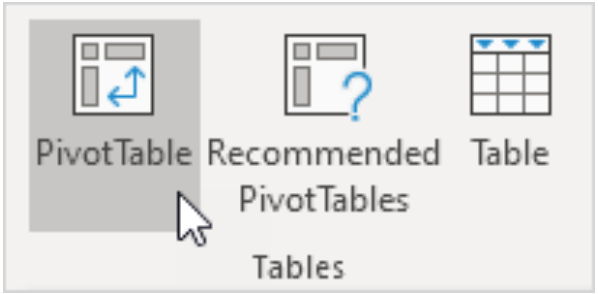
Example

OUR DATASET:

	A	B	C	D	E	F	G	H
1	**Order ID**	**Product**	**Category**	**Amount**	**Date**	**Country**		
2	1	Carrots	Vegetables	$4,270	1/6/2016	United States		
3	2	Broccoli	Vegetables	$8,239	1/7/2016	United Kingdom		
4	3	Banana	Fruit	$617	1/8/2016	United States		
5	4	Banana	Fruit	$8,384	1/10/2016	Canada		
6	5	Beans	Vegetables	$2,626	1/10/2016	Germany		
7	6	Orange	Fruit	$3,610	1/11/2016	United States		
8	7	Broccoli	Vegetables	$9,062	1/11/2016	Australia		
9	8	Banana	Fruit	$6,906	1/16/2016	New Zealand		
10	9	Apple	Fruit	$2,417	1/16/2016	France		
11	10	Apple	Fruit	$7,421	1/16/2016	Canada		

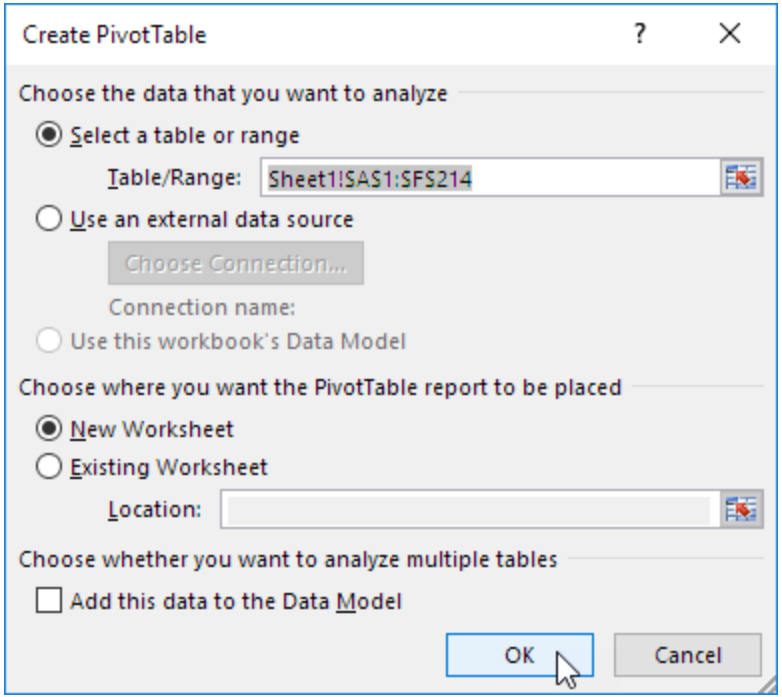
Step 1: Insert a pivot table

1. Click any single cell inside the data set.
2. Under the Insert tab, in the Tables group, click PivotTable.



Step 2: dialog box

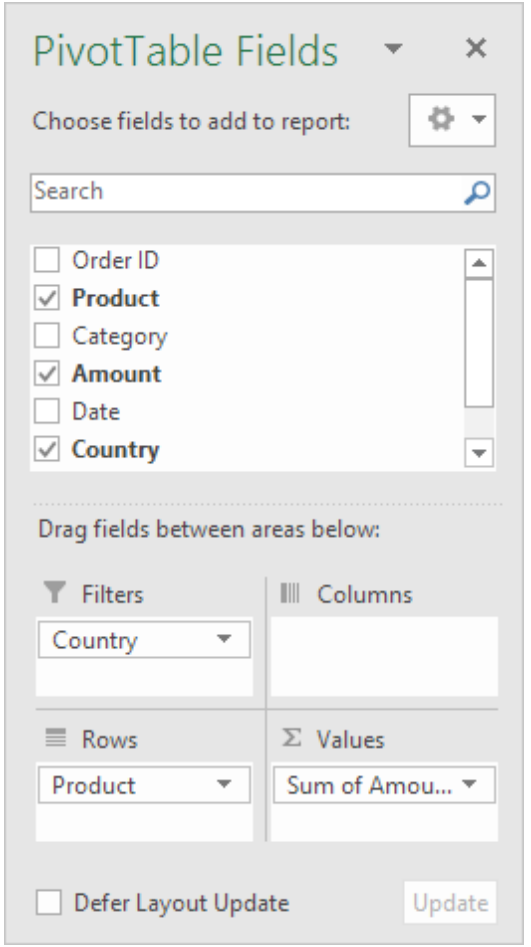
- Excel automatically selects the data for you. The default location for a new pivot table is New Worksheet.
- Click OK



STEP 3: Drag Fields

- The PivotTable Fields pane appears. To get the total amount exported of each product, drag the following fields to the different areas.

1. Product field to the Rows area.
2. Amount field to the Values area.
3. Country field to the Filters area.



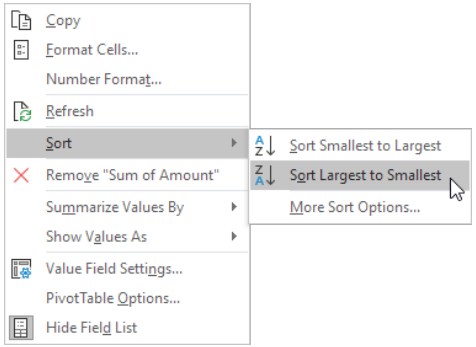
Result

What we can see: Bananas are our main export product because that is the highest number

	A	B	C
1	Country	(All)	
2			
3	**Row Labels**	**Sum of Amount**	
4	Apple	191257	
5	Banana	340295	
6	Beans	57281	
7	Broccoli	142439	
8	Carrots	136945	
9	Mango	57079	
10	Orange	104438	
11	**Grand Total**	**1029734**	
12			

What can we do with pivot tables

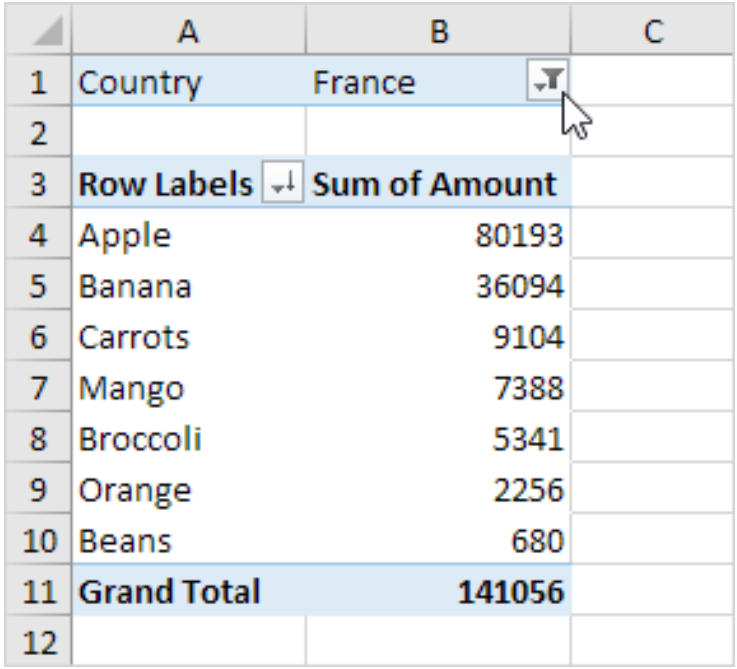
- **SORT**
- To get Banana at the top of the list, sort the pivot table.
- Click any cell inside the Sum of Amount column.
- Right click and click on Sort, Sort Largest to Smallest.



	A	B	C
1	Country	(All)	
2			
3	Row Labels	Sum of Amount	
4	Banana	340295	
5	Apple	191257	
6	Broccoli	142439	
7	Carrots	136945	
8	Orange	104438	
9	Beans	57281	
10	Mango	57079	
11	Grand Total	1029734	
12			

What can we do with pivot tables

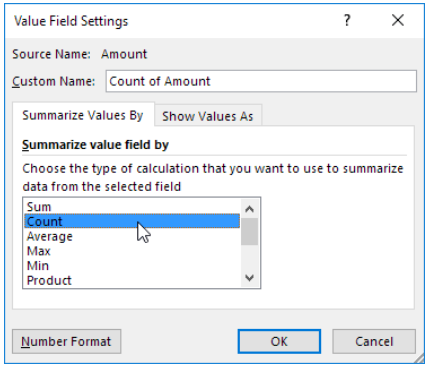
- **FILTER**
- Because we added the Country field to the Filters area, we can filter this pivot table by Country. For example, which products do we export the most to France?
- Click the filter drop-down and select France.
- Result. Apples are our main export product to France.

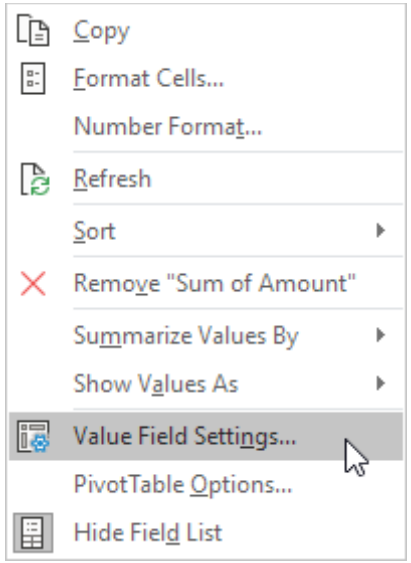


	A	B	C
1	Country	France	
2			
3	**Row Labels**	**Sum of Amount**	
4	Apple	80193	
5	Banana	36094	
6	Carrots	9104	
7	Mango	7388	
8	Broccoli	5341	
9	Orange	2256	
10	Beans	680	
11	**Grand Total**	**141056**	
12			

What can we do with pivot tables

- Change Summary Calculation
- By default, Excel summarizes your data by either summing or counting the items. To change the type of calculation
- Click any cell inside the Sum of Amount column.
- Right click and click on Value Field Settings.





	A	B	C
1	Country	France	
2			
3	Row Labels	Count of Amount	
4	Apple	16	
5	Banana	7	
6	Carrots	1	
7	Mango	1	
8	Orange	1	
9	Beans	1	
10	Broccoli	1	
11	Grand Total	28	
12			

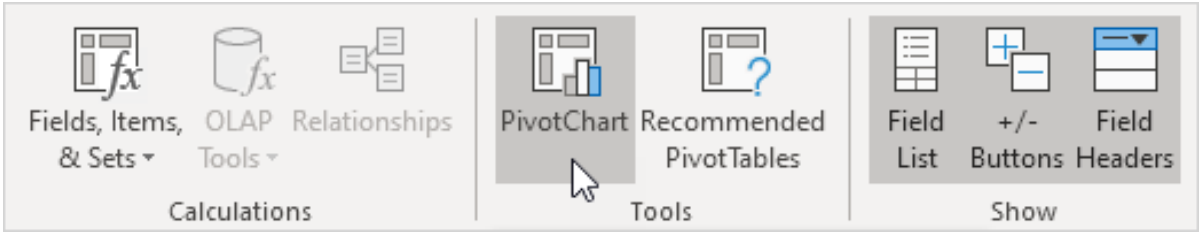
What is a Pivot chart?

- A pivot chart is the **visual representation** of a pivot table in Excel.
- Our data set (same as before)

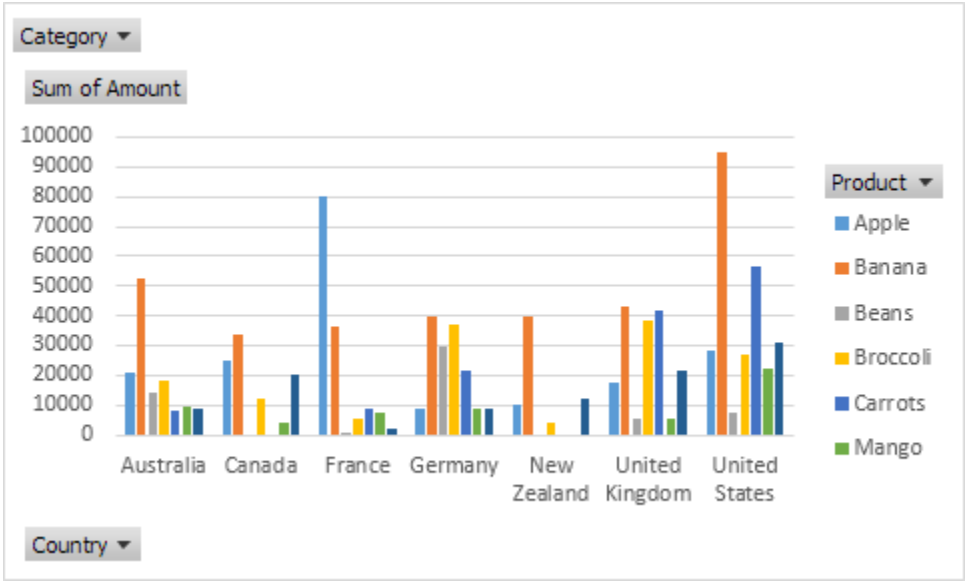
	A	B	C	D	E	F	G	H	I	J
1	Category	(All)								
2										
3	Sum of Amount	Column								
4	Row Labels	Apple	Banana	Beans	Broccoli	Carrots	Mango	Orange	Grand Total	
5	Australia	20634	52721	14433	17953	8106	9186	8680	131713	
6	Canada	24867	33775		12407		3767	19929	94745	
7	France	80193	36094	680	5341	9104	7388	2256	141056	
8	Germany	9082	39686	29905	37197	21636	8775	8887	155168	
9	New Zealand	10332	40050		4390			12010	66782	
10	United Kingdom	17534	42908	5100	38436	41815	5600	21744	173137	
11	United States	28615	95061	7163	26715	56284	22363	30932	267133	
12	**Grand Total**	**191257**	**340295**	**57281**	**142439**	**136945**	**57079**	**104438**	**1029734**	
13										

Step 1: Insert Pivot chart

1. Click any cell inside the **pivot table**.
2. On the Analyze tab, in the Tools group, click PivotChart.

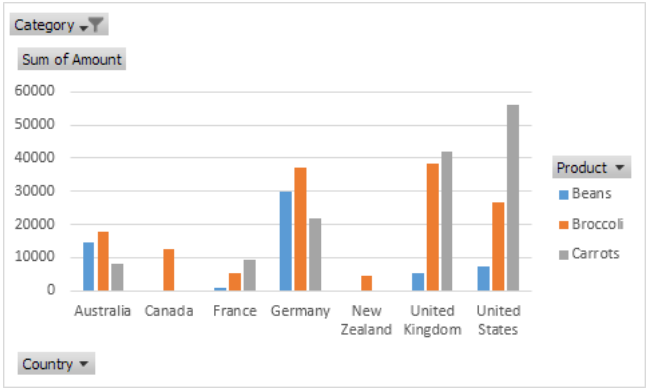


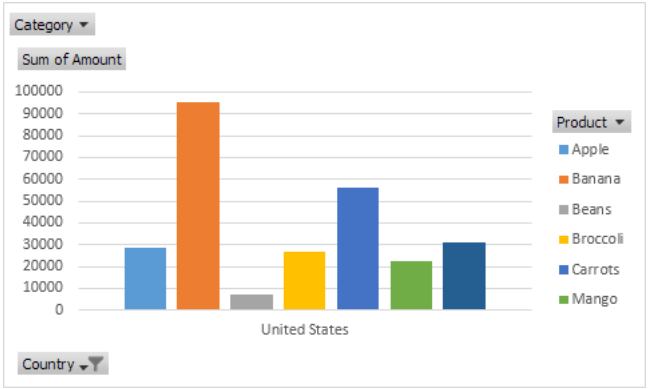
Result


Category
Sum of Amount
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
Australia
Canada
France
Germany
New Zealand
United Kingdom
United States
Product
Apple
Banana
Beans
Broccoli
Carrots
Mango
Country

What can we do with pivot charts

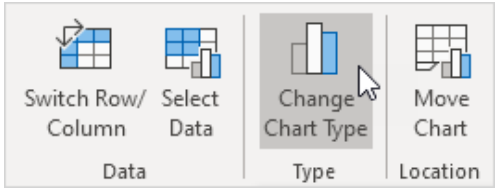
- Use the standard filters (triangles next to Product and Country). For example, use the Country filter to only show the total amount of each product exported to the United States.
- Explore: Select and deselect different fields to play with the data

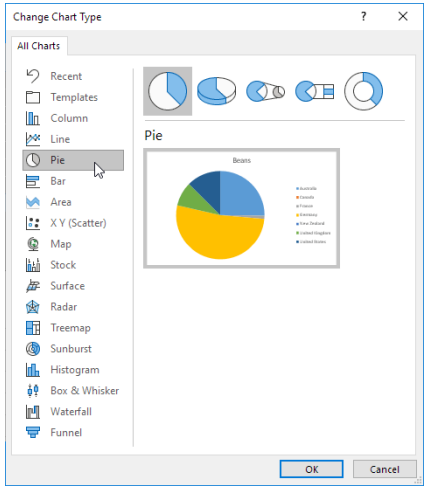


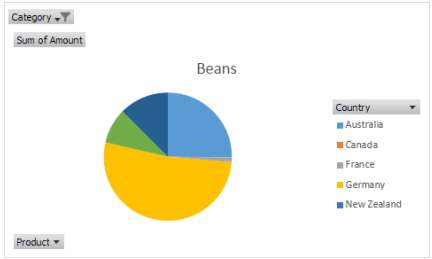


What can we do with pivot charts

- Change Pivot Chart Type
  1. Select the chart.
  2. On the Design tab, in the Type group, click Change Chart Type







NOTES:

- Any changes you make to the pivot chart are immediately reflected in the pivot table and vice versa.
- Pie charts always use one data series (in this case, Beans). To get a pivot chart of a country, swap the data over the axis.
  1. Select the chart.
  2. On the Design tab, in the Data group, click Switch Row/Column.

SLICERS AND TIME SLICERS

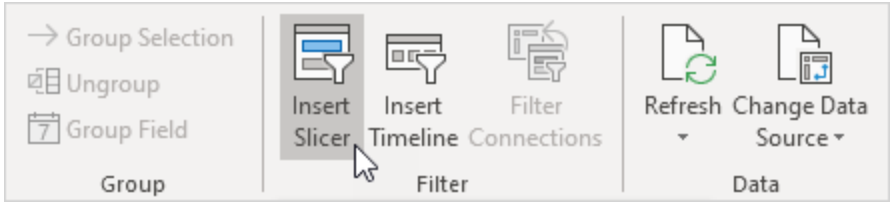
WHAT ARE SLICERS IN EXCEL

- Use slicers in Excel to quickly and easily filter pivot tables.
- OUR DATA:

	A	B	C
1	Country	(All)	
2			
3	**Row Labels**	**Sum of Amount**	
4	Apple	191257	
5	Banana	340295	
6	Beans	57281	
7	Broccoli	142439	
8	Carrots	136945	
9	Mango	57079	
10	Orange	104438	
11	**Grand Total**	**1029734**	
12			

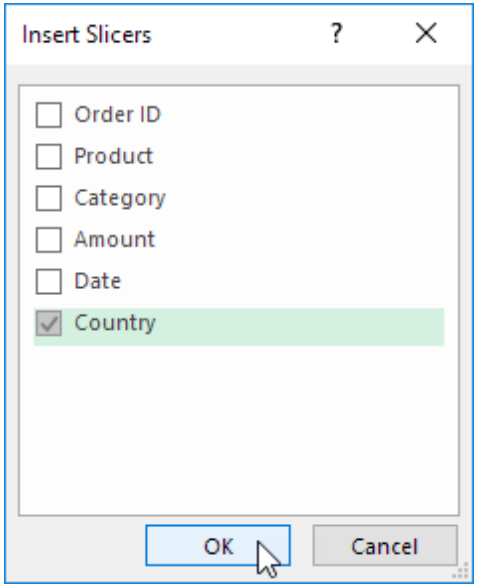
Step 1: Insert a slicer

- Click any cell inside the pivot table.
- On the Analyze tab, in the Filter group, click Insert Slicer.

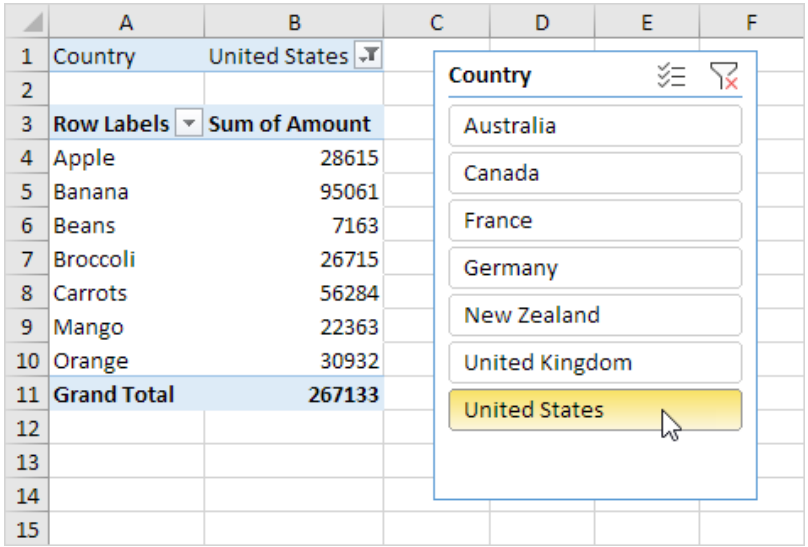


STEP 2: select relevant fields

Check Country and click OK.

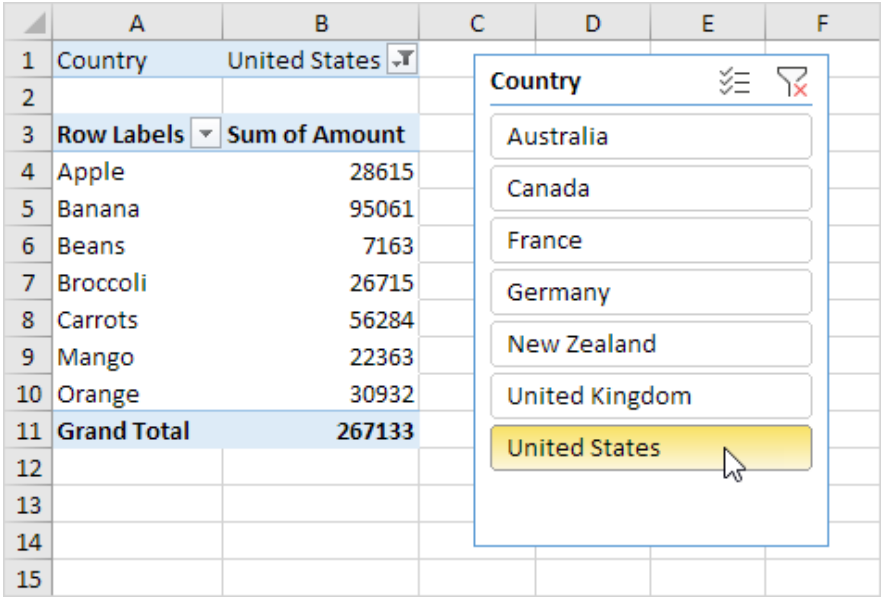


Click United States to find out which products we export the most to the United States.



RESULT:

- Bananas are our main export product to the United States because that is the max value



	A	B
1	Country	United States
2		
3	Row Labels	Sum of Amount
4	Apple	28615
5	Banana	95061
6	Beans	7163
7	Broccoli	26715
8	Carrots	56284
9	Mango	22363
10	Orange	30932
11	Grand Total	267133
12		
13		
14		
15		

Q&A

Office Hours for technical, software or installation assistance (or anything else):

Monday 10-11:30am; Thursday 2-3:30pm, in **SRH 3.264.**

Contact: arojas@austin.utexas.edu | (512-552-9860)

Say the class/subject upfront!