Excel Advanced Tips: Sort and Filter

South Dakota

The 5th Annual Demography Conference

By

Shuang Li

Ph.D. Student, Teaching Assistant

Department of Sociology and Rural Studies

South Dakota State University

Introduction

Sort:

- The most important difference between Sort and Filter:
  - Sort does not remove entries from the data set.
  - It only puts the data into a new order.

Filter:

- Filter only keeps the items you select, and hide others.

Sort...

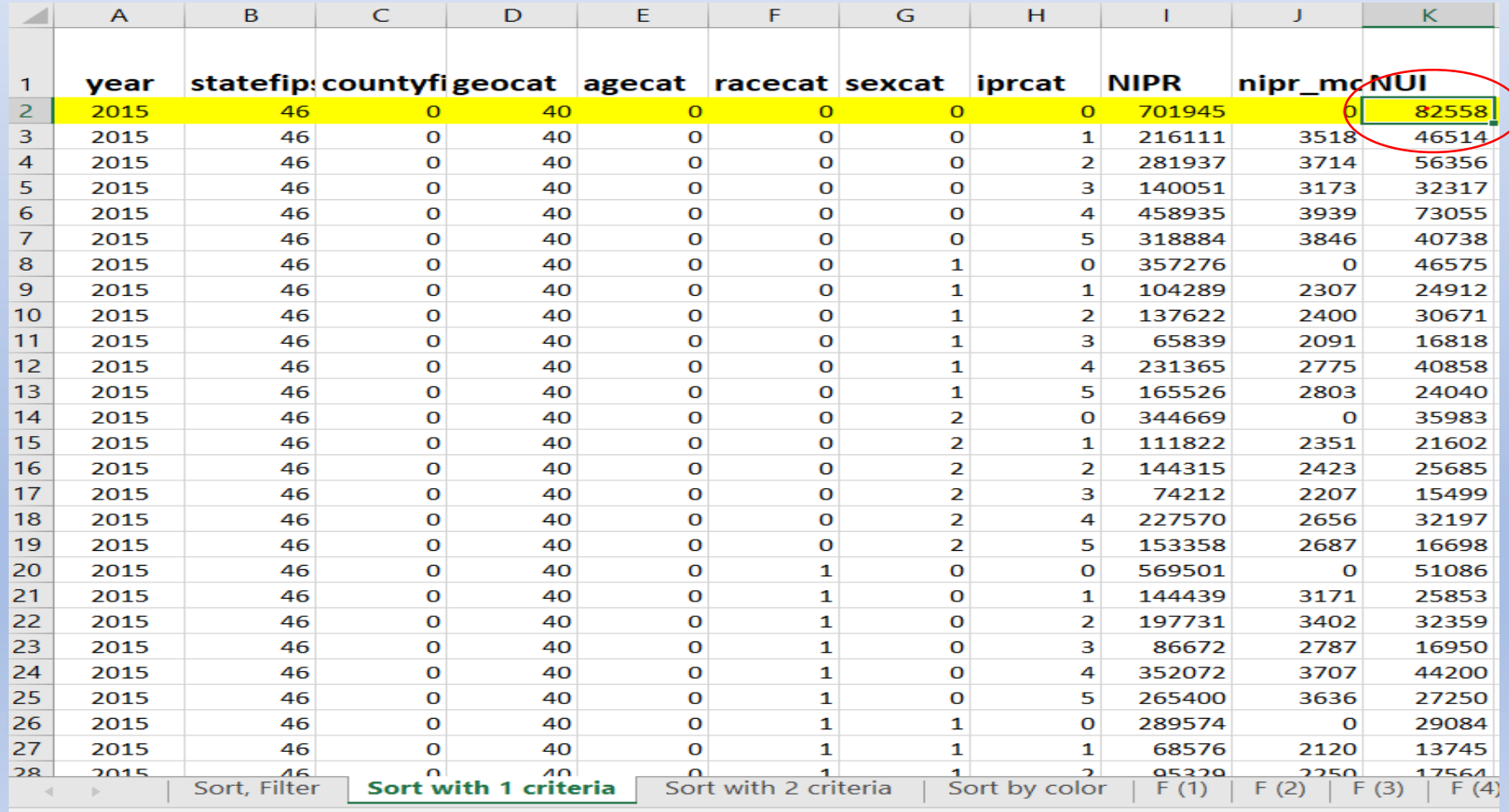
- Sort with 1 criteria: 1) Right-click, 2) Home Tab, 3) Data Tab
  - The whole table reminds intact, even though one single column was sorted.

- Sort with 2 criteria or more: when sorting with more than one column
  1) cell method, leave the Major sort for last
  2) Use Sort Dialog Box, the major sort is on top (preferred)

- Dataset downloaded from https://www.census.gov/did/www/sahie/

Step 1: Sort with 1 criteria

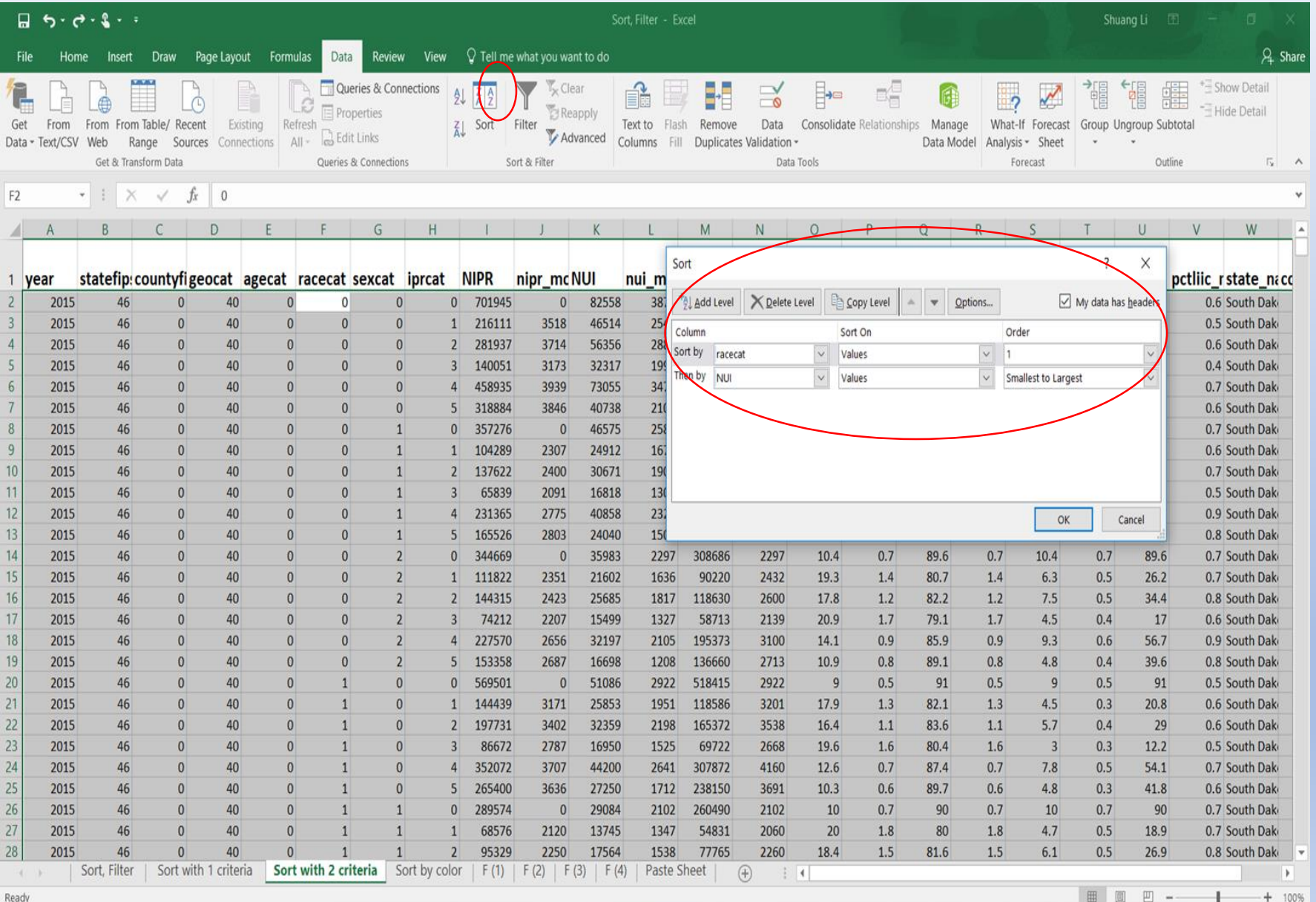
1. Click any cell in the column you want to sort.



	A	B	C	D	E	F	G	H	I	J	K
1	year	statefip:	countyfi	geocat	agecat	racecat	sexcat	iprcat	NIPR	nipr_mc	NUI
2	2015	46	0	40	0	0	0	0	701945	0	82558
3	2015	46	0	40	0	0	0	1	216111	3518	46514
4	2015	46	0	40	0	0	0	2	281937	3714	56356
5	2015	46	0	40	0	0	0	3	140051	3173	32317
6	2015	46	0	40	0	0	0	4	458935	3939	73055
7	2015	46	0	40	0	0	0	5	318884	3846	40738
8	2015	46	0	40	0	0	1	0	357276	0	46575
9	2015	46	0	40	0	0	1	1	104289	2307	24912
10	2015	46	0	40	0	0	1	2	137622	2400	30671
11	2015	46	0	40	0	0	1	3	65839	2091	16818
12	2015	46	0	40	0	0	1	4	231365	2775	40858
13	2015	46	0	40	0	0	1	5	165526	2803	24040
14	2015	46	0	40	0	0	2	0	344669	0	35983
15	2015	46	0	40	0	0	2	1	111822	2351	21602
16	2015	46	0	40	0	0	2	2	144315	2423	25685
17	2015	46	0	40	0	0	2	3	74212	2207	15499
18	2015	46	0	40	0	0	2	4	227570	2656	32197
19	2015	46	0	40	0	0	2	5	153358	2687	16698
20	2015	46	0	40	0	1	0	0	569501	0	51086
21	2015	46	0	40	0	1	0	1	144439	3171	25853
22	2015	46	0	40	0	1	0	2	197731	3402	32359
23	2015	46	0	40	0	1	0	3	86672	2787	16950
24	2015	46	0	40	0	1	0	4	352072	3707	44200
25	2015	46	0	40	0	1	0	5	265400	3636	27250
26	2015	46	0	40	0	1	1	0	289574	0	29084
27	2015	46	0	40	0	1	1	1	68576	2120	13745
28	2015	46	0	40	0	1	1	2	95329	2250	17564

Sort, Filter | Sort with 1 criteria | Sort with 2 criteria | Sort by color | F (1) | F (2) | F (3) | F (4)

Step 2: Sort with 2 criteria or more



1. Open the Sort Dialog– Add Level
2. The major sort (Racecat) is on top

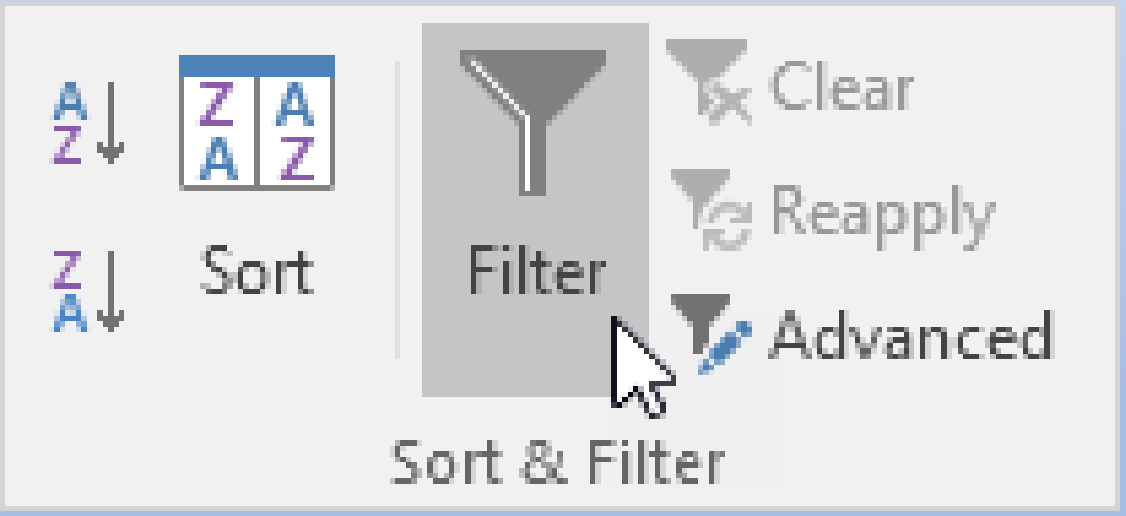
Filter…(Ctrl+Shift+L)

- Filter is similar to sort, but with one great advantage
- Filter only keeps the items you select, and hide others
  - Missing rows (row headers are blue)
  - Filter button is still on
- If copy the filtered items, the pasted will only show the filtered ones, not the whole dataset.
  - Cltrl+Shift+8 (number pad) to select the current

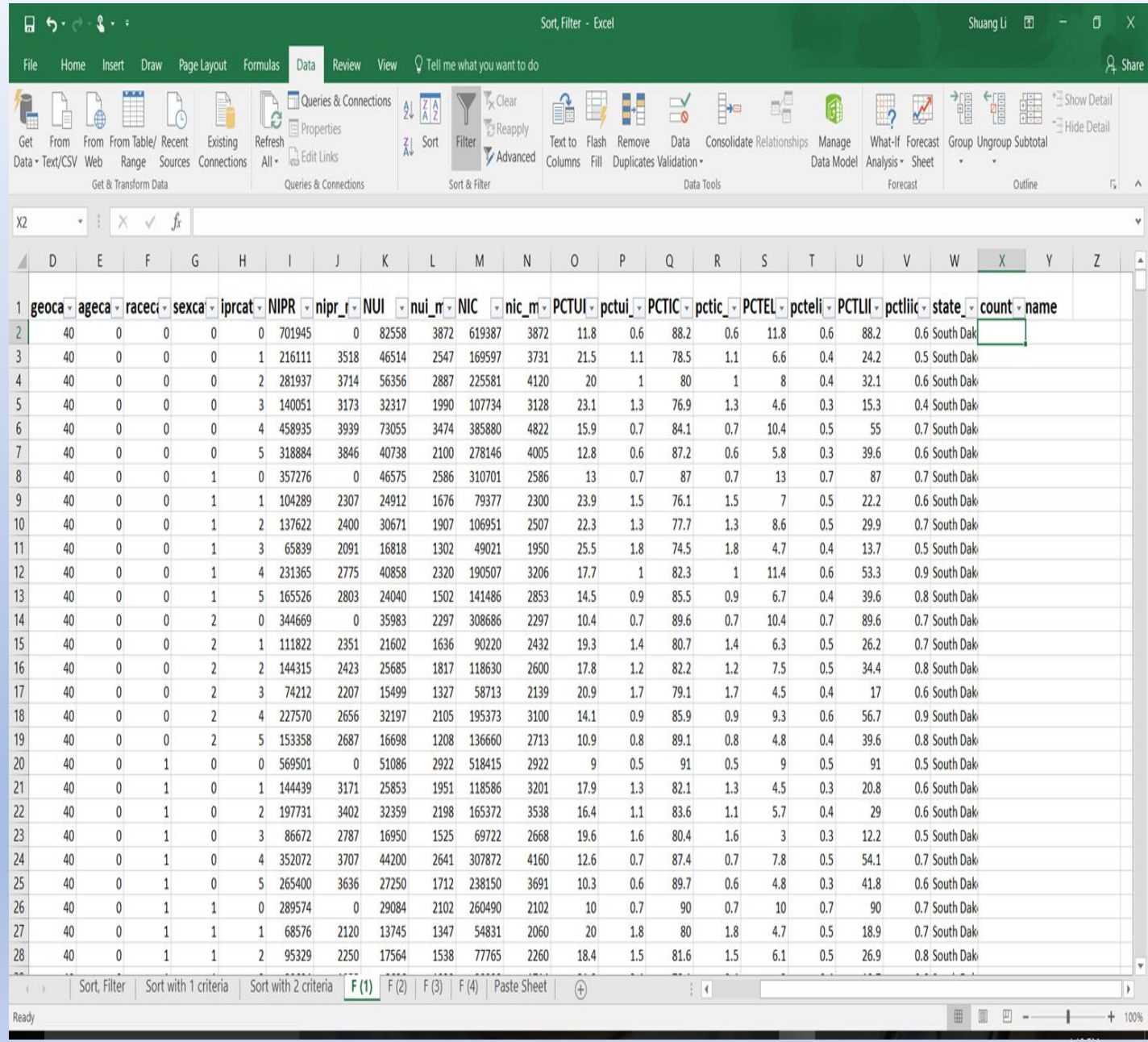
Steps in Examples

1. Click any single cell inside a data set.

2. On the Data tab, in the Sort & Filter group, click Filter.



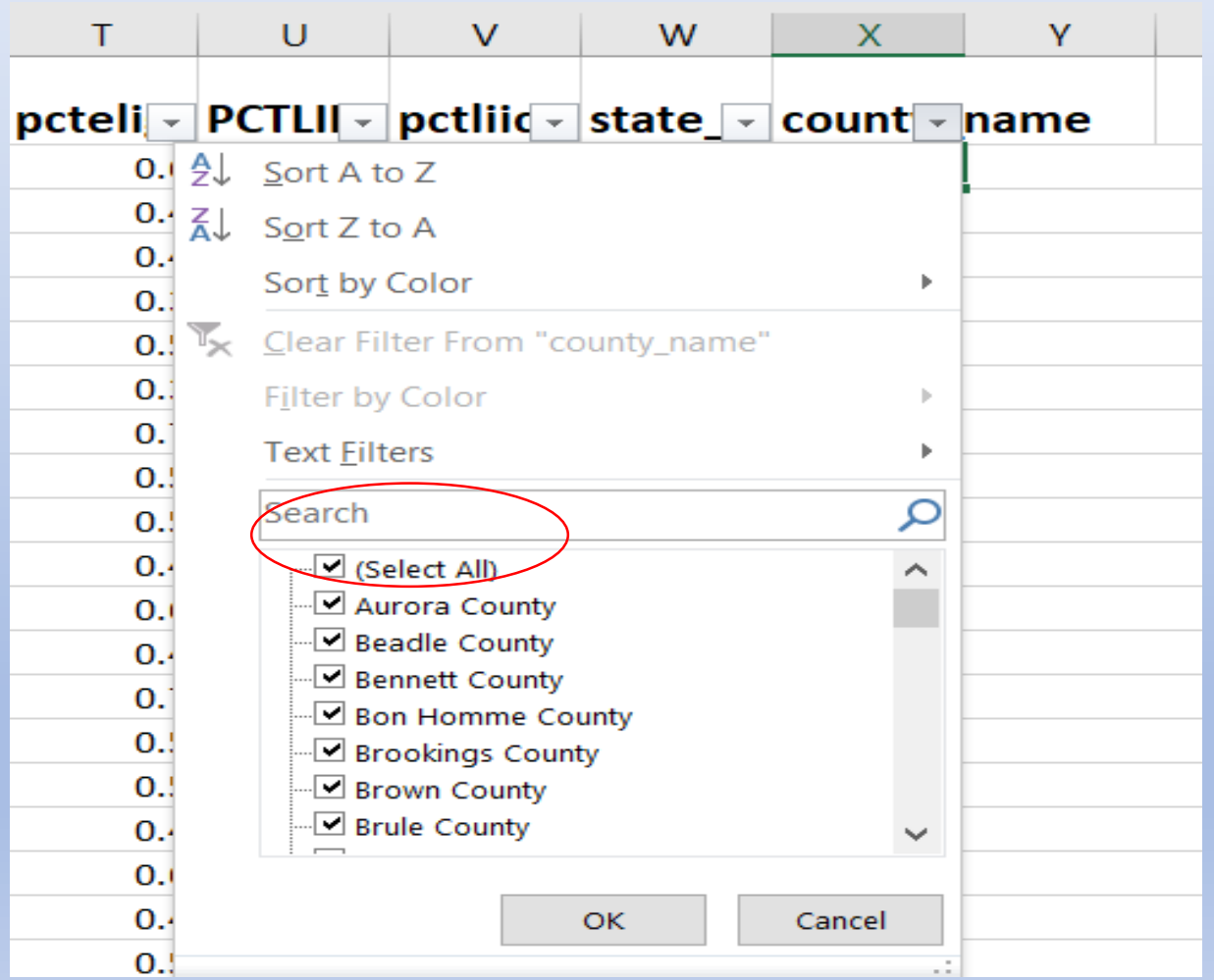
3. Arrows in the column headers appear.



Steps:

3. Click the arrow next to “county name”.

4. Click on Select All to clear all the check boxes, and click “Brookings County”.



Last Tip!

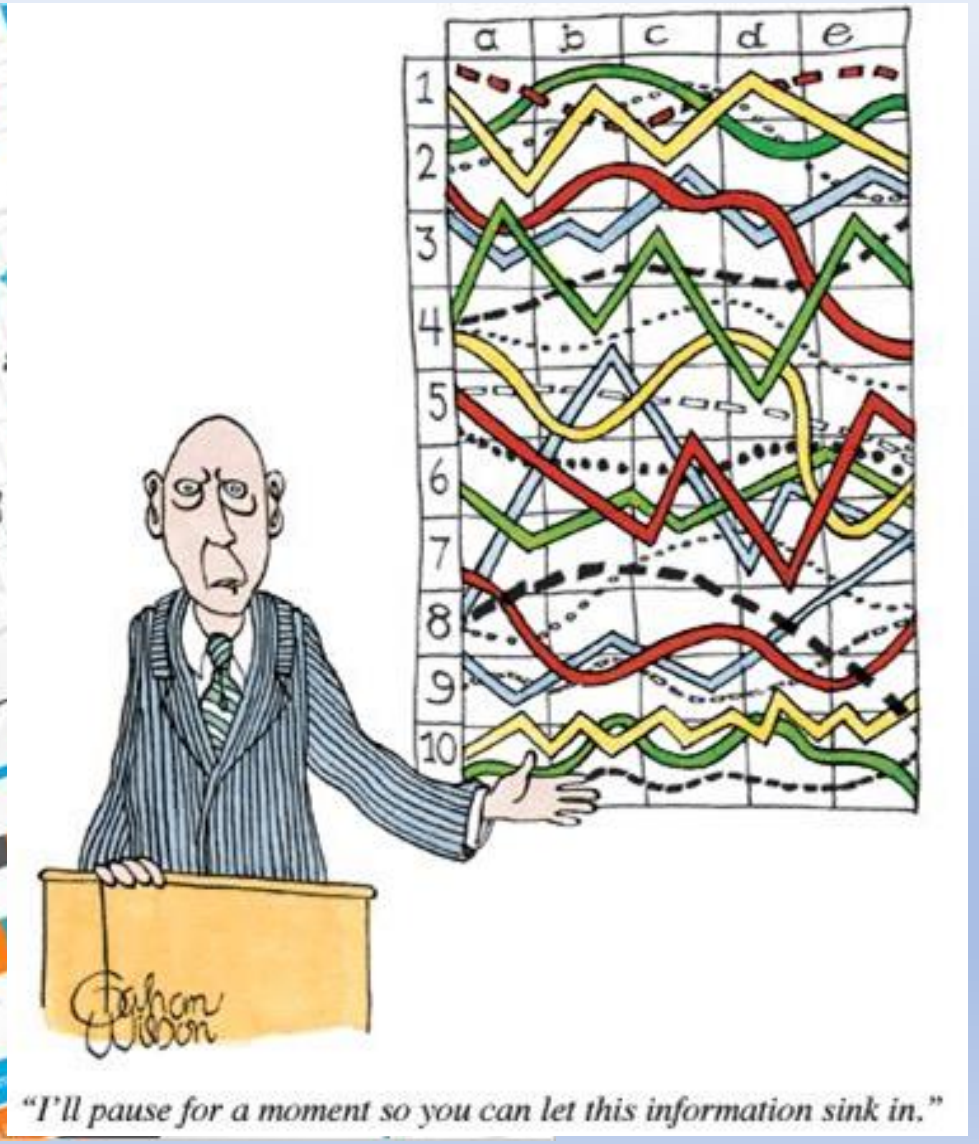
- One last tip: don’t forget you can filter AND sort at the same time.

- Questions? Comments?
- Contact Shuang Li via Shuang.li@sdstate.edu
- Thank you!

Data Visualizations: Lines and Pies

Virginia James
Ph.D. Student -
South Dakota State University

Data Visualization:

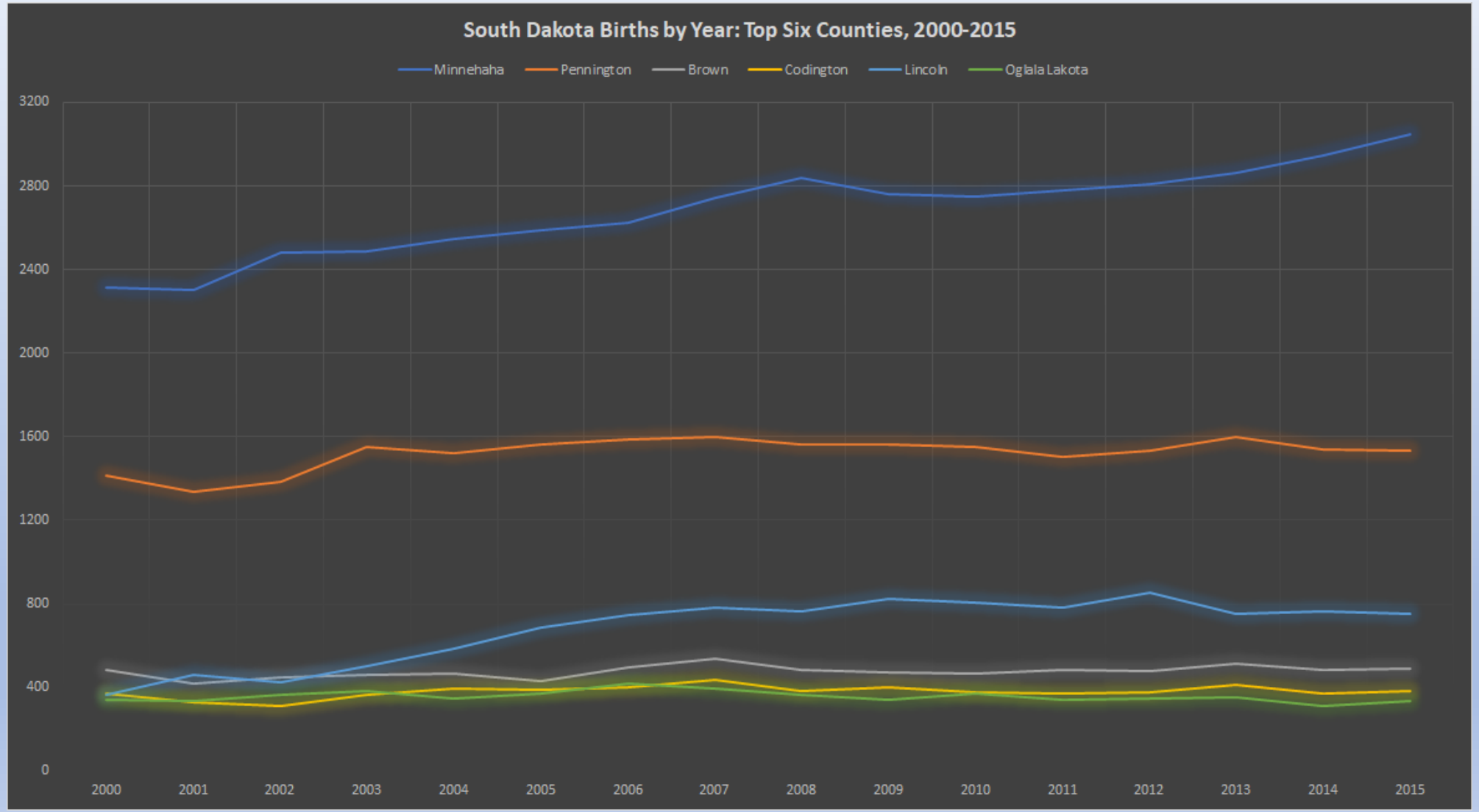


"I'll pause for a moment so you can let this information sink in."

Source: http://au.pcmag.com/tableau-desktop/43134/feature/10-free-data-visualization-tools

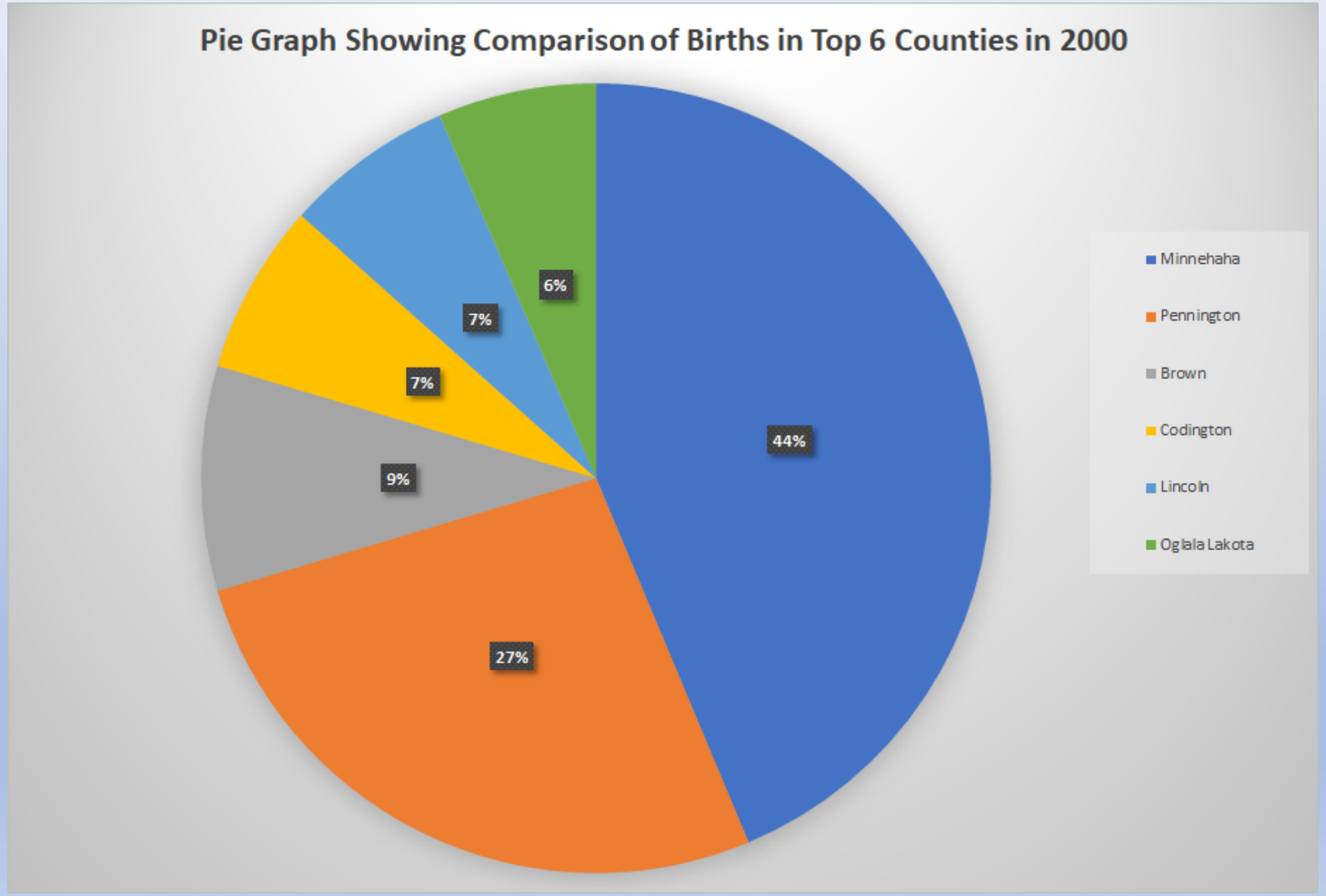
Source:https://lovestats.wordpress.com/dman/

Line Graph Building: Displaying Trends



Source of Data from Department of Health, South Dakota

Pie Graph Building: Displaying a Comparison



Source of Data from Department of Health, South Dakota

A Few Helpful Links:

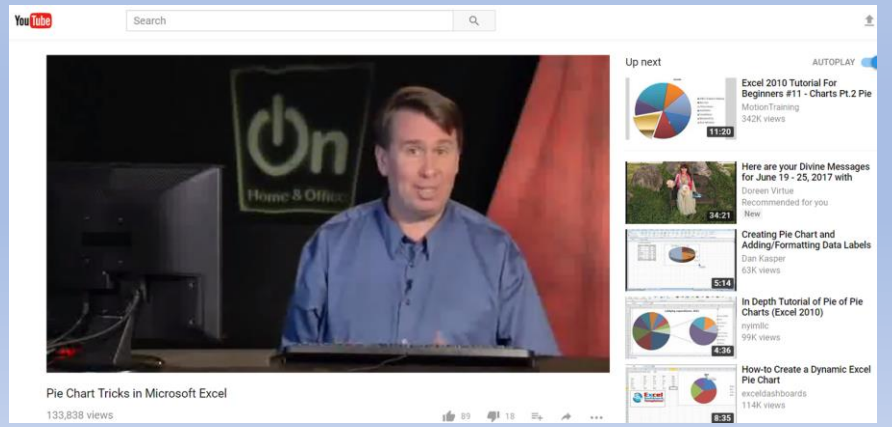
Line Graph Help:

Link to Microsoft Office Blogs for Chart Design: https://blogs.office.com/2012/05/30/our-eight-best-tutorials-on-excel-charts/
YouTube Video Tutorial for Line Graph Making: https://www.youtube.com/watch?v=9X0WTOqlmCI

Pie Graph Help:

YouTube video for Creating a Pie Chart: https://www.youtube.com/watch?v=BQqQmxwp5GU

Thank you for learning EXCEL with us!



Pivot Table

Wei (Vina) Gu

Research Assistant, Ph.D Student

State Data Center

Department of Sociology and Rural Studies, South Dakota State University

Email: *wei.gu@jacks.sdstate.edu*

Introduction

Pivot Table - a data summarization tool

Functions:

- Sort, Filter
- Count,
- Sum,
- Average,
- Cross Tabulation (two dimensional calculation)

Advantages

- Data safety:

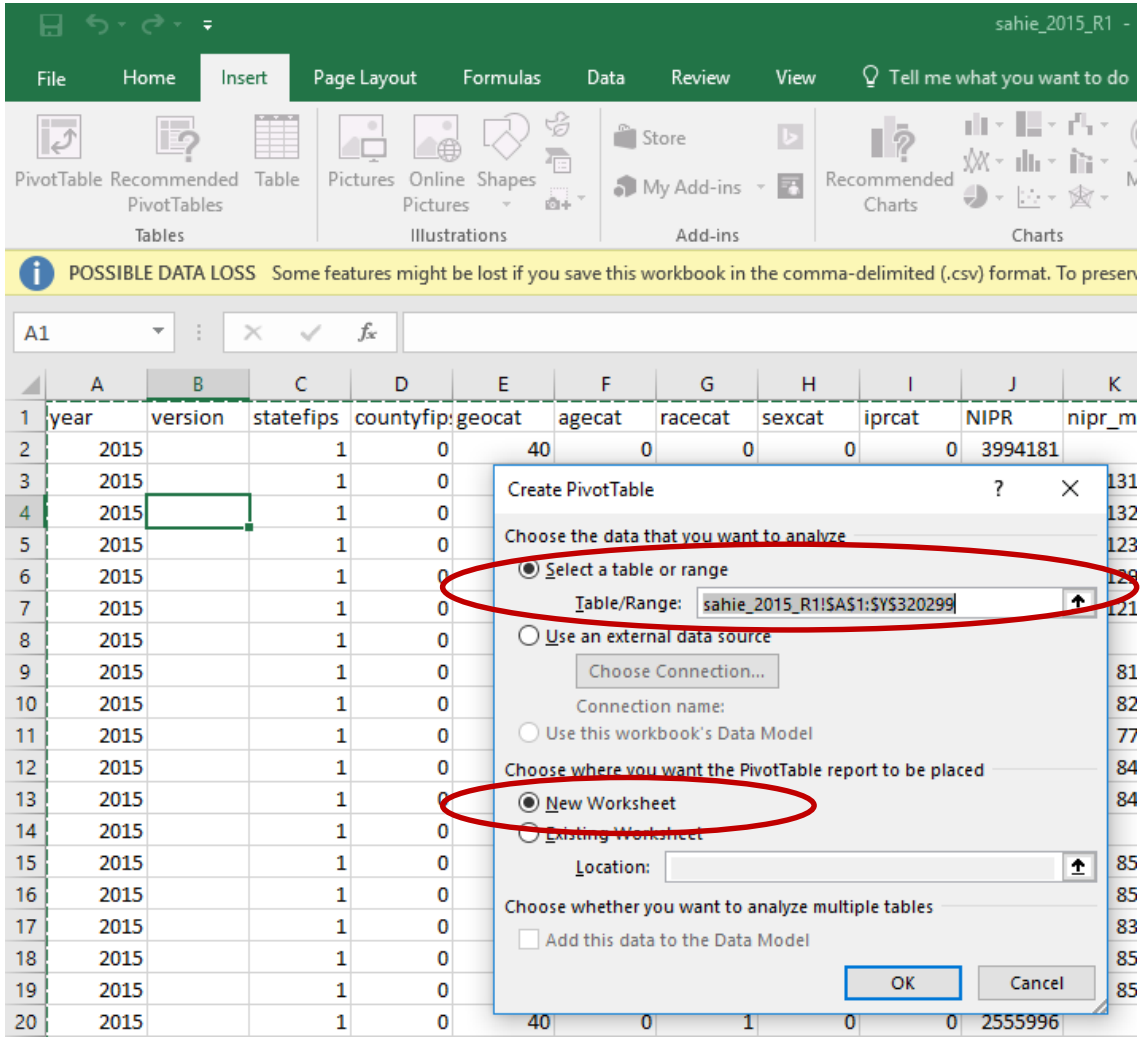
Keep original data by displaying the results in a second table

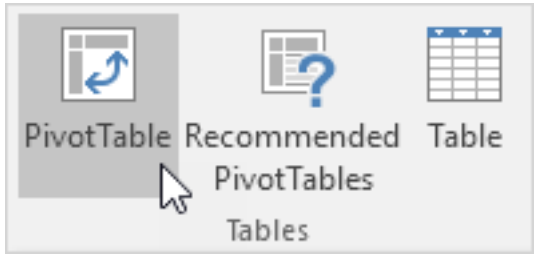
- Flexibility:

Sets up and changes the summary's structure by dragging and dropping fields graphically.

- Good tool for a large, detailed data set.

Step1: Insert Pivot Table and Select Table

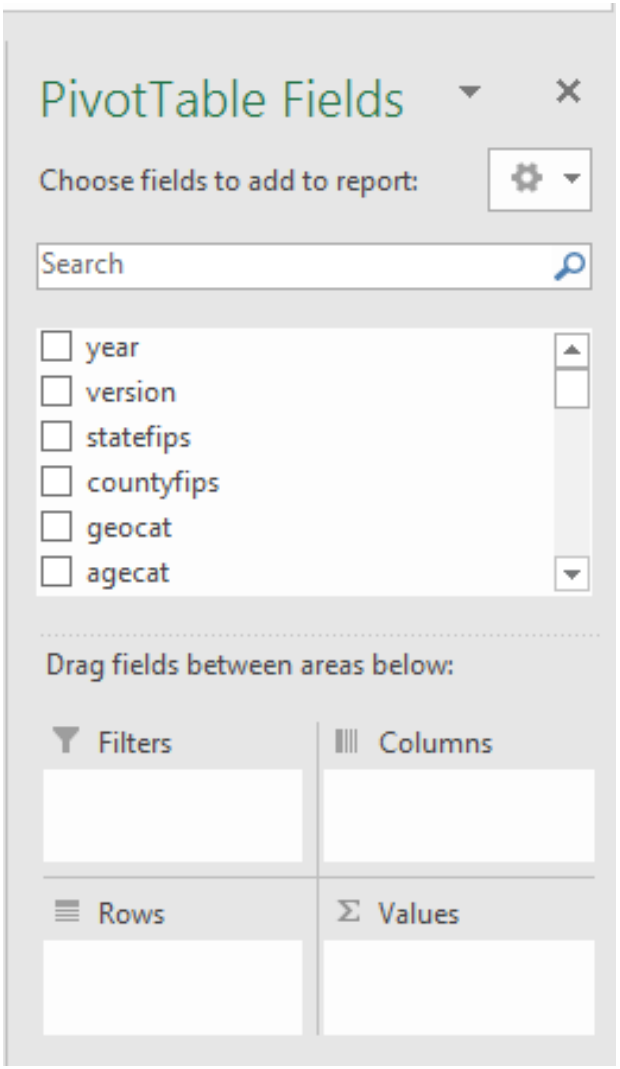




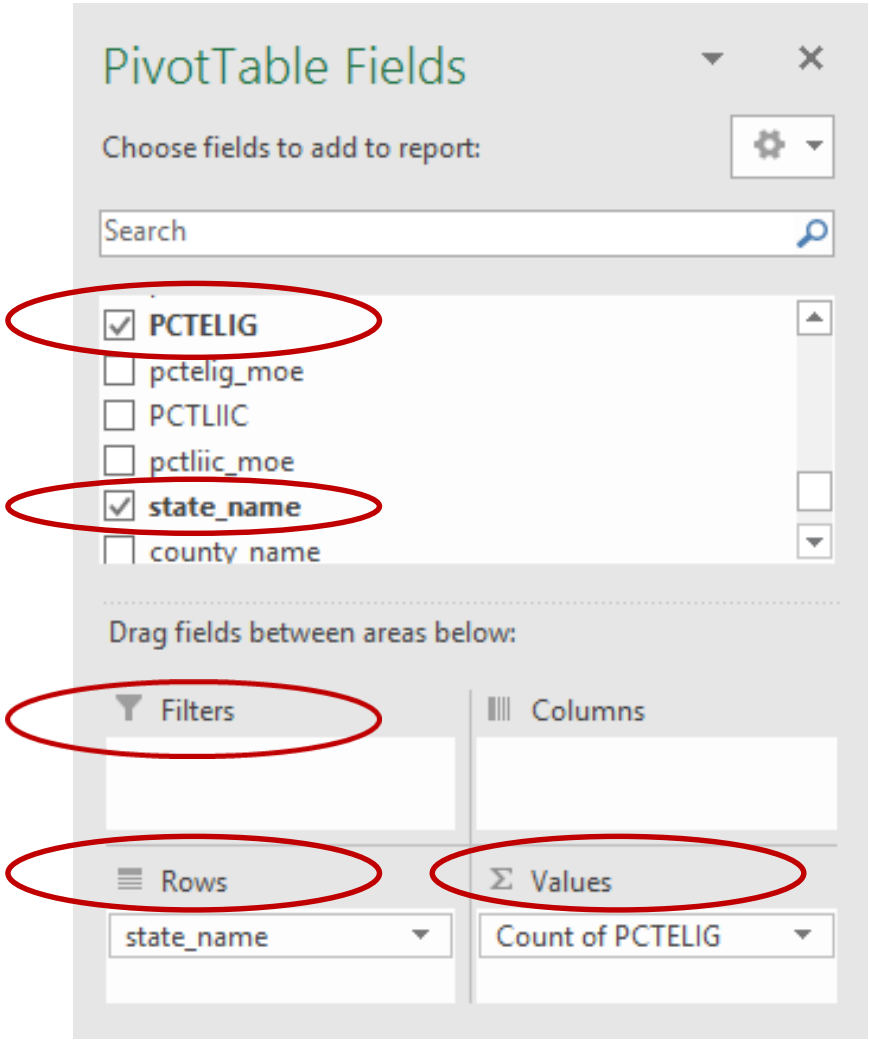
Tips: For large data set, before I insert a PivotTable, I will clean the dataset first- only keep titles of the variables and the data.

Step2: Drag Fields

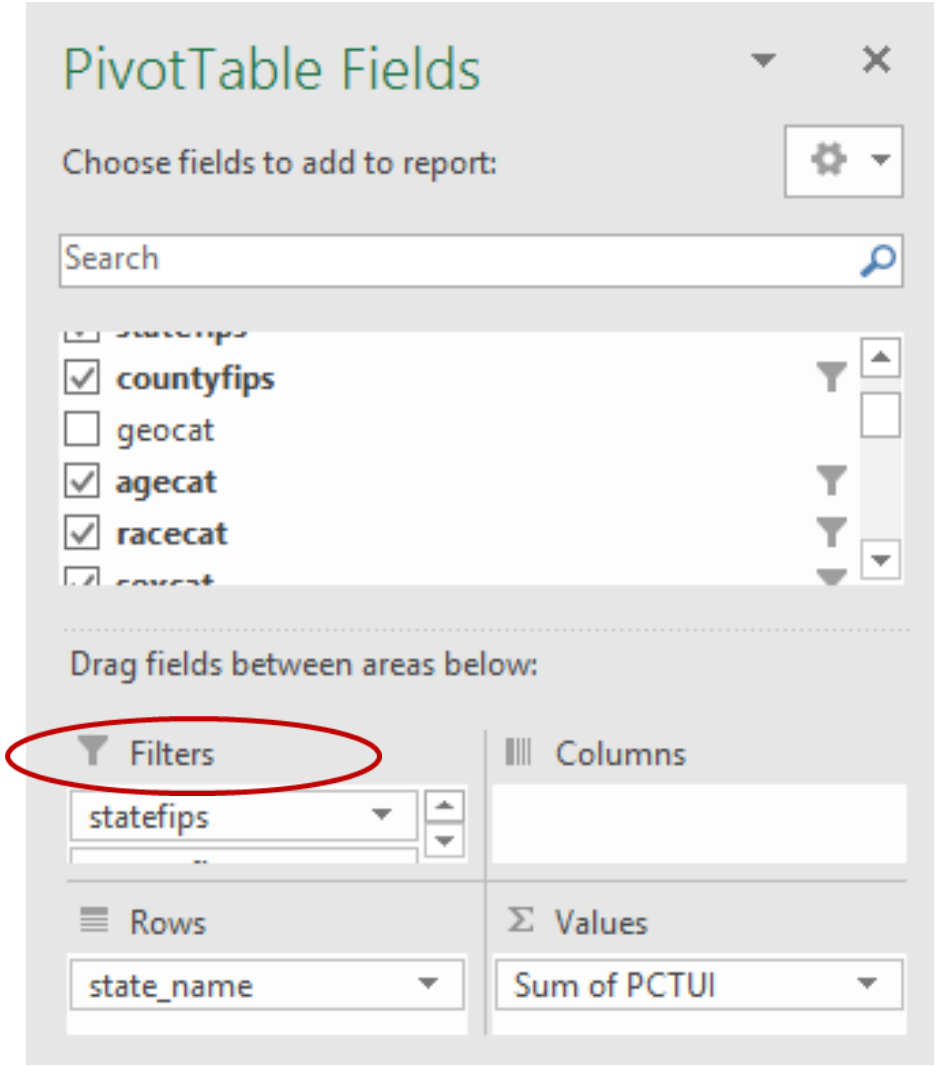
Pivot table fields pane



Select variables into areas



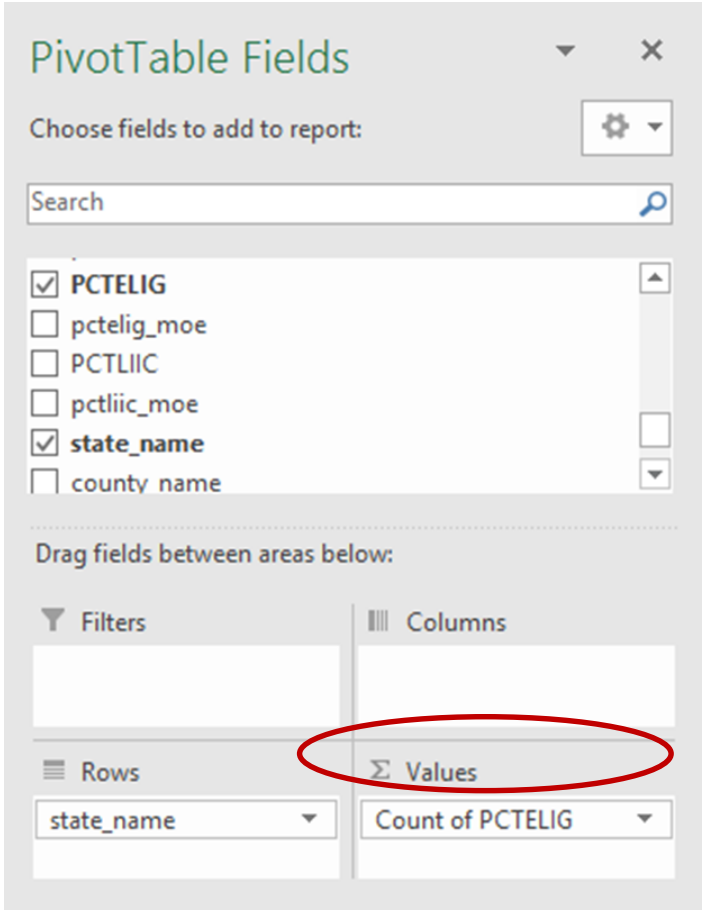
Step2: Drag Fields- Filters

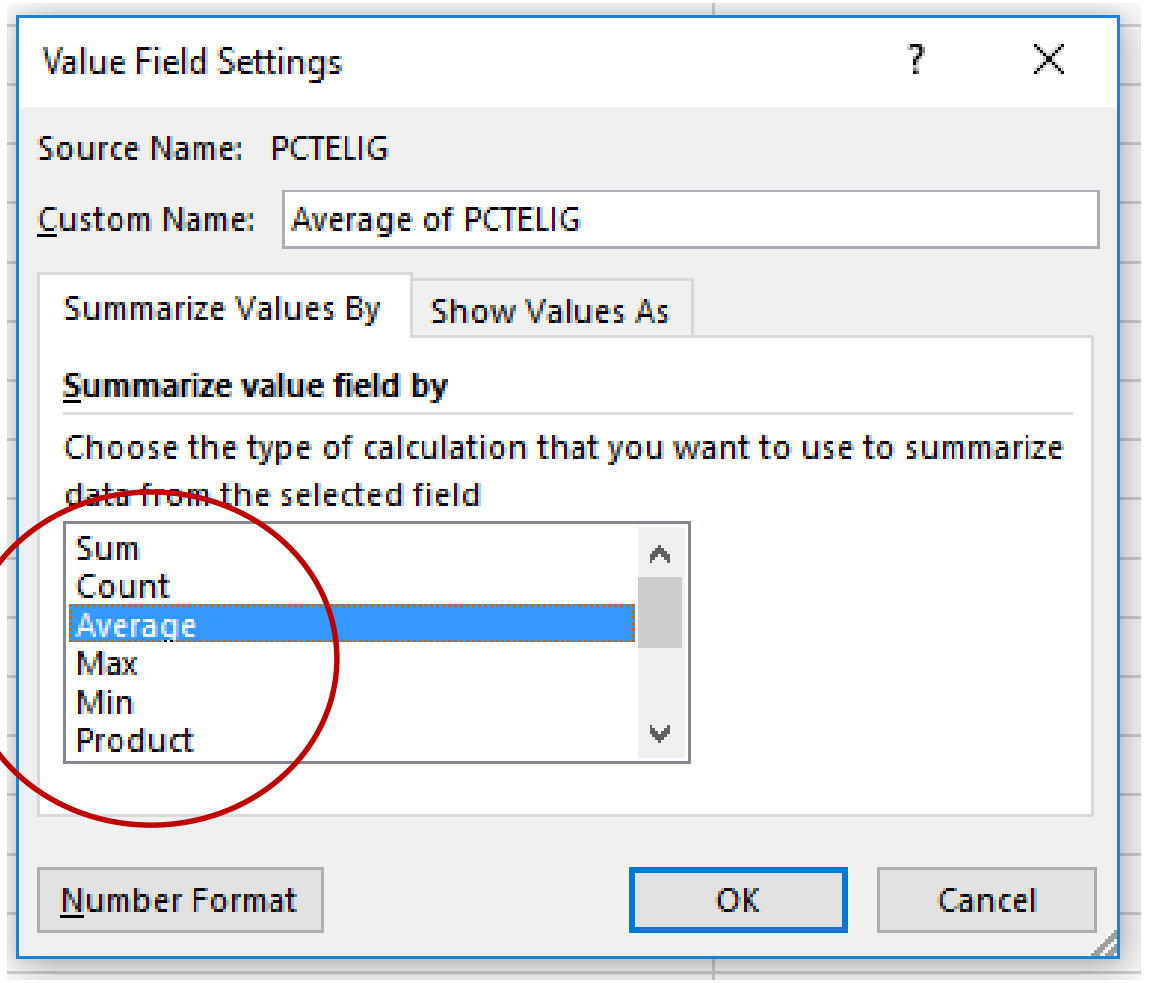


	A	B	C
1	statefips	(All)	
2	countyfips	0	
3	agecat	0	
4	racecat	0	
5	iprcat	0	
6	sexcat	0	
7			
8	**Row Labels**	**Sum of PCTUI**	
9	Alabama	11.9	
10	Alaska	16.3	
11	Arizona	12.8	
12	Arkansas	11.1	
13	California	9.7	
14	Colorado	9.2	
15	Connecticut	6.9	
16	Delaware	6.9	
17	District of Columbia	4.3	
18	Florida	16.3	
19	Georgia	15.8	

Step3: Set Summary Calculation

- Click the button under “Values” area;
- Then, click “Value Field Settings”;
- Choose the type of calculation. **How?**

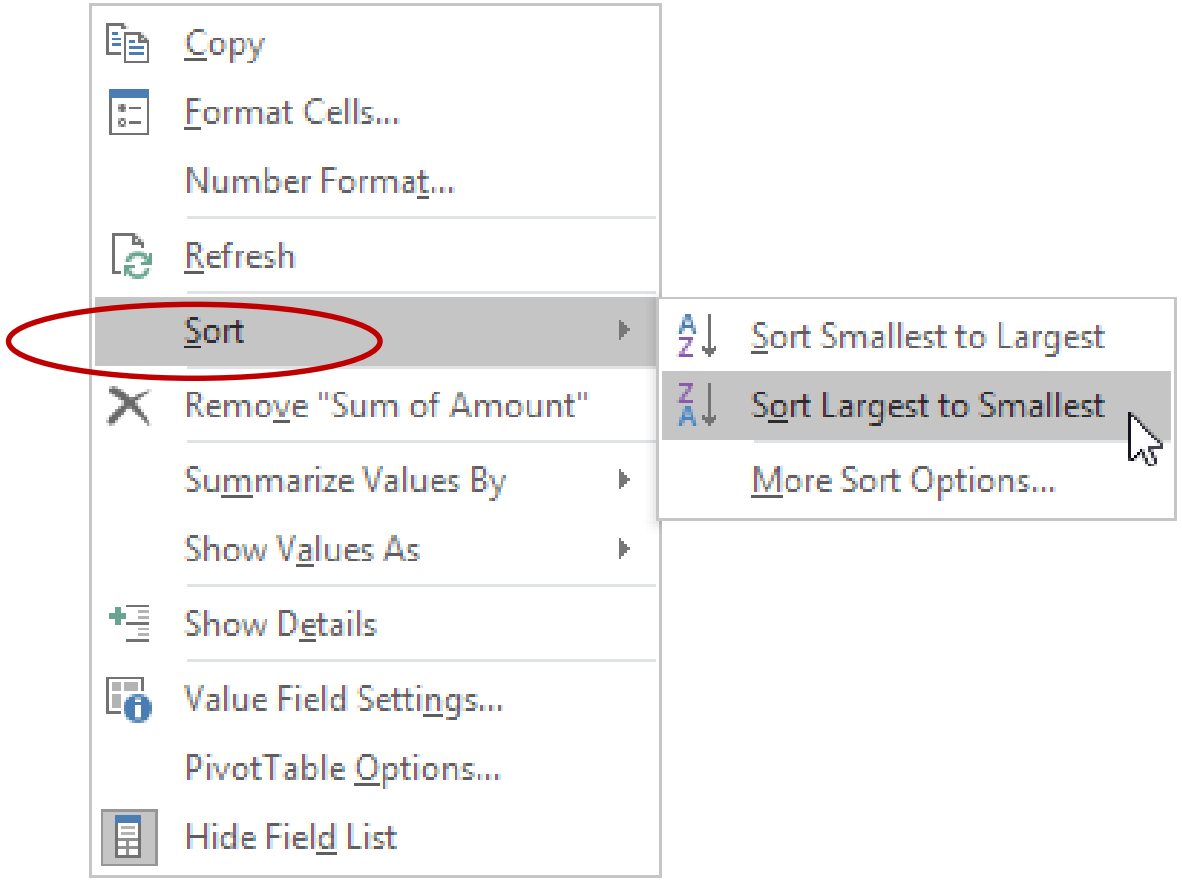




Step4: Sort result

- Right click any cell inside the “Sum of PCTUI” column;
- Then, click sort.

	A	B	C
1	statefips	(All)	
2	countyfips	0	
3	agecat	0	
4	racecat	0	
5	iprcat	0	
6	sexcat	0	
7			
8	**Row Labels**	**Sum of PCTUI**	
9	Alabama	11.8	
10	Alaska	16.3	
11	Arizona	12.8	
12	Arkansas	11.1	
13	California	9.7	
14	Colorado	9.2	
15	Connecticut	6.9	
16	Delaware	6.9	
17	District of Columbia	4.3	
18	Florida	16.3	
19	Georgia	15.8	



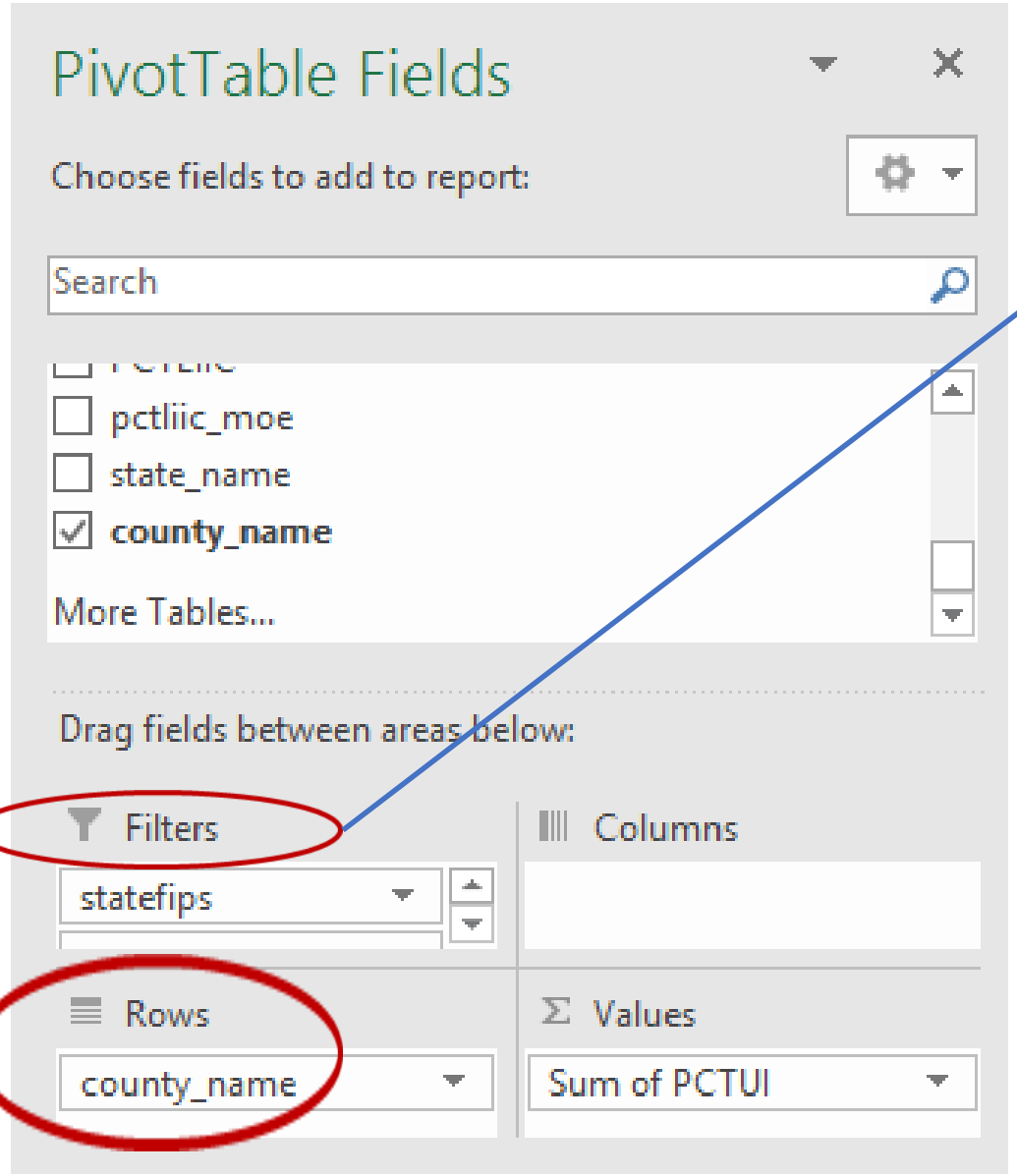
Two Examples

1. Percent uninsured in all States in the U.S.

2. Percent uninsured in all counties in South Dakota.

- Valuable name will use: “state_name” or “statefips”, “PCTUI”, “agecat”, “racecat”, “sexcat”, “iprcat”.
- Functions: Sort, Filter, Change Summary Calculation

Example 2



statefips	46	
racecat	0	
countyfips	(All)	
agecat	0	
iprcat	0	
sexcat	0	

Row Labels	Sum of PCTUI
	11.8
Aurora County	14.7
Beadle County	14.7
Bennett County	22.3
Bon Homme County	12.8
Brookings County	10.8
Brown County	10.2
Brule County	17.5
Buffalo County	20.7
Butte County	13.9

Tips

Before you extract a pivot table,

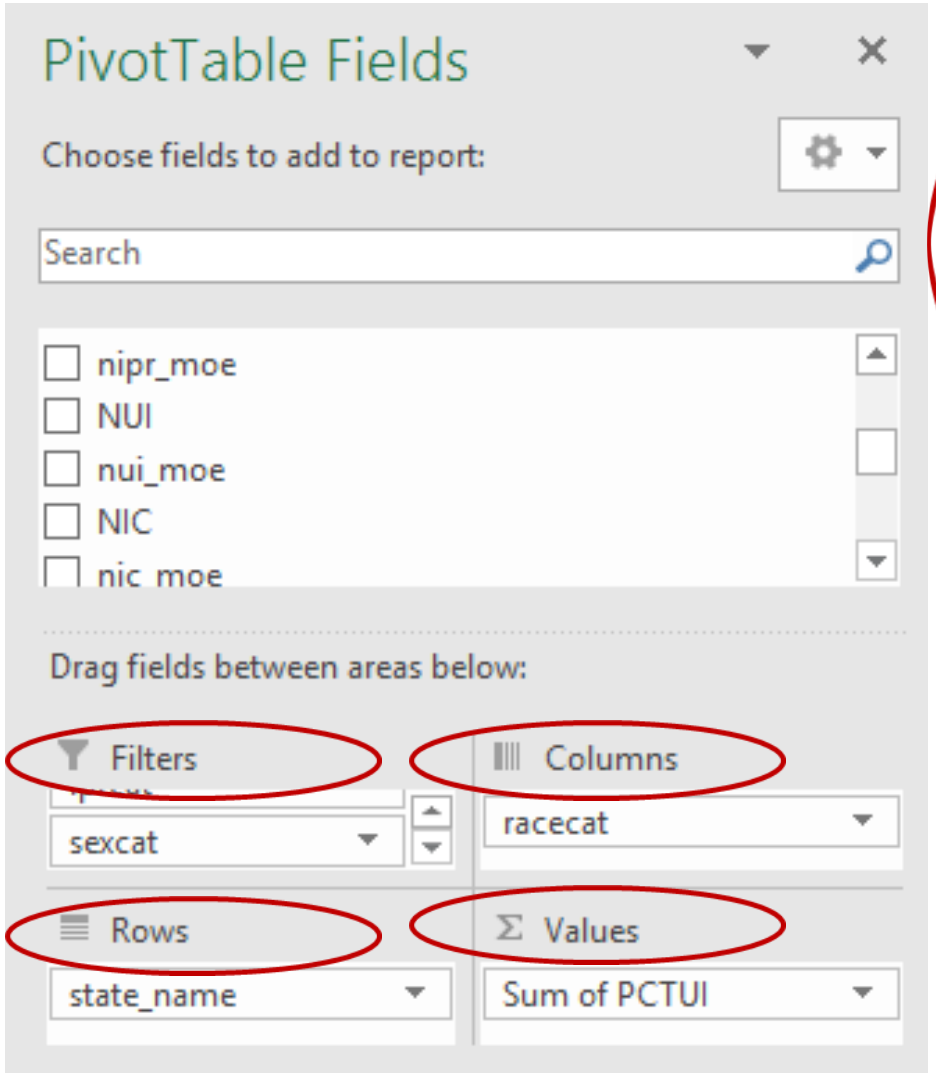
- Familiar with the data structure and arrangement,
- How to set the “Value Field Settings”.

Two-dimensional Pivot Table

- Rows & Columns

- Example: Comparing uninsured percentage by race in all States in the U.S.

Two-dimensional pivot table



statefips	(All)				
countyfips	0				
agecat	0				
iprcat	0				
sexcat	0				
Sum of PCTUI	**Column Labels**				
Row Labels	**0**	**1**	**2**	**3**	
Alabama	11.9	9.9	13.7	27.9	
Alaska	16.3	11.5	18.9	25.8	
Arizona	12.8	7.9	10.7	19.9	
Arkansas	11.1	9.3	11.3	25.2	
California	9.7	5.6	7.1	15	
Colorado	9.2	6.3	8.6	18.1	
Connecticut	6.9	4.2	8.1	16.1	
Delaware	6.9	5.3	7	16.5	
District of Columbia	4.3	2.4	4.3	10.4	

References:

- *Small Area Health Insurance Estimates Program, U.S. Census Bureau.*
- *Pivot Tables, http://www.excel-easy.com/data-analysis/pivot-tables.html. Excel Easy.*

Thank you!

Questions? Comments?

Email: *wei.gu@jacks.sdstate.edu*

SAHIE File Layout Overview

Product: SAHIE File Layout Overview
Filenames: SAHIE Text and SAHIE CSV files 2008 to 2015
Source: Small Area Health Insurance Estimates Program, U.S. Census Bureau.
Internet Release Date: March 2017
Description: Small Area Health Insurance Estimates (SAHIE)'s data file layout, variable names, and variable descriptions

Variable	Description
year	Year of Estimate
version	Release Version
	Blank: Year other than 2013, Only Version
	Original: 2013 only, Original Version
	Updated: 2013 only, Updated Version (May 2016)
statefips	Unique FIPS code for each state
countyfips	Unique FIPS code for each county within a state
geocat	Geography category:
	40 – State geographic identifier
	50 – County geographic identifier
agecat	Age category
	0 – Under 65 years
	1 – 18 to 64 years
	2 – 40 to 64 years
	3 – 50 to 64 years
	4 – Under 19 years
	5 – 21 to 64 years
racecat	Race category
	0 – All races
	Only state estimates have racecat=1,2,3 values
	1 – White alone, not Hispanic
	2 – Black alone, not Hispanic
	3 – Hispanic (any race)
sexcat	Sex category
	0 – Both sexes
	1 - Male
	2 - Female
iprcat	Income category
	0 – All income levels
	1 – At or below 200% of poverty
	2 – At or below 250% of poverty
	3 – At or below 138% of poverty
	4 – At or below 400% of poverty
	5 – Between 138% - 400% of poverty
NIPR	Number in demographic group for <income category>
nipr_moe	MOE for NIPR
NUI	Number uninsured
nui_moe	MOE for NUI
NIC	Number insured
nic_moe	MOE for NIC
PCTUI	Percent uninsured in demographic group for <income category>
pctui_moe	MOE for PCTUI
PCTIC	Percent insured in demographic group for <income category>
pctic_moe	MOE for PCTIC
PCTELIG	Percent uninsured in demographic group for all income levels
pctelig_moe	MOE for PCTELIG
PCTLIIC	Percent insured in demographic group for all income levels
pcteliic_moe	MOE for PCTLIIC
state_name	State name
county_name	County name

Note 1: A margin of error (MOE) is the difference between an estimate and its upper or lower confidence bounds. Confidence bounds can be created by adding the margin of error to the estimate (for an upper bound) and subtracting the margin of error from the estimate (for a lower bound). All published margins of error for the Small Area Health Insurance Estimates program are based on a 90 percent confidence level.

Note 2: The number in a demographic group is the number of people in the poverty universe in that age, sex, and race/Hispanic origin group.

Note 3: Values for Kalawao, HI (15-005)should be considered N/A or missing.

Note 4: MOEs of zero should be assumed to be <1 for counts and <0.1 for percentages.

General Note: Details may not sum to totals because of rounding.