2014 Planning Database (PDB)

November 19, 2014

Barbara O'Hare, Nancy Bates, Julia Coombs,
Travis Pape, Chandra Erdman

Office of Survey Analytics

Overview

- Tract and Block Group PDBs
- Useful for:
  - Identifying areas for outreach or recruitment
  - Stratifying small areas
  - Creating thematic maps
  - Enhancing reports with population metrics
- Will be updated every year with new ACS estimates

Background

- First PDB developed for 2000 census planning
  - Selected 1990 census tract data in easy-to-use format
  - Hard-to-Count Score
- ACS annual 5-year estimates for block groups resulted in revised PDB in 2012
- 2014 PDB
  - Latest 5-year ACS estimates
  - Percentages
  - Low Response Score

Contents of the 2014 PDB

- Both 2008-2012 5-year ACS estimates and 2010 census data
- Types of variables
  - Population: gender, age, education, poverty
  - Household: language, relationship, income
  - Housing unit: tenure, number of units
  - Census operational: mailout/mailback, bilingual

Access

- Available on the Census Bureau's Research @ Census page
- Link to the PDB: http://www.census.gov/research/data/planning_database/
  - Data files in zipped CSV format
  - Documentation describing the files in PDF format

Navigation to the PDB

- From the Census Bureau internet site (http://www.census.gov):
  1. Select “Our Research” from under the “About the Bureau” menu at the top of the page
  2. Select the “Data” tab
  3. Select the “Research Data Products” link
  4. Select “Planning Database” under the “Demographic – People and Households” heading
  5. Select the appropriate year under “Data and Documentation”

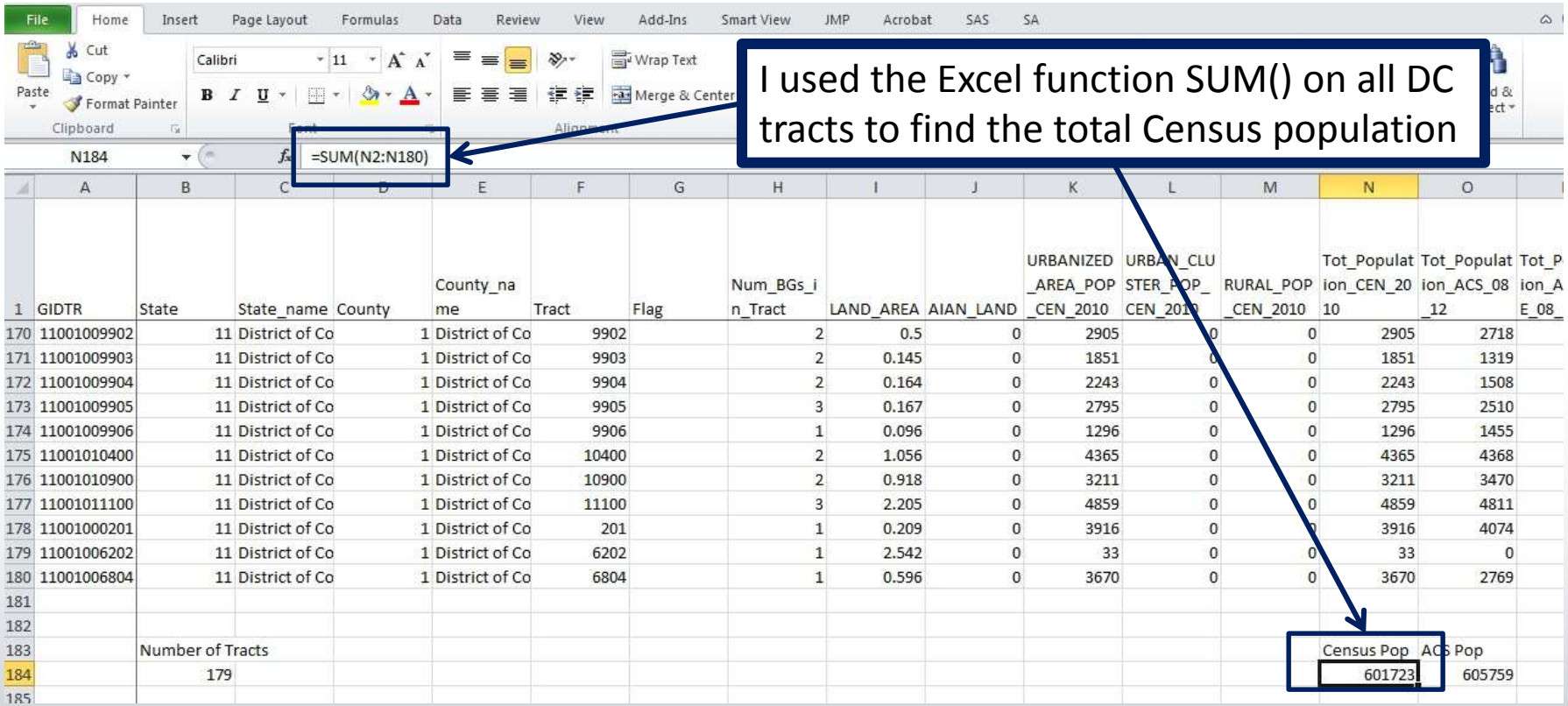
Examples Using the PDB

Area Demographics

- 601,723 people live in 179 tracts in DC

	DC	United States
Male to female ratio	0.90	0.97
Population under 5 years old	5.42%	6.54%
Population that identifies as Hispanic	9.10%	16.35%
Population that moved within the past year	19.45%	15.03%
Population that was not born in the US	13.54%	12.87%

Using Excel to Analyze Demographics



1	GIDTR	State	State_name	County	County_name	Tract	Flag	Num_BGs_in_Tract	LAND_AREA	AIAN_LAND	URBANIZED_AREA_POP_CEN_2010	URBAN_CLUSTER_POP_CEN_2010	RURAL_POP_CEN_2010	Tot_Population_CEN_2010	Tot_Population_ACS_08_12	Tot_Pop_AE_08_
170	11001009902	11	District of Co	1	District of Co	9902		2	0.5	0	2905	0	0	2905	2718	
171	11001009903	11	District of Co	1	District of Co	9903		2	0.145	0	1851	0	0	1851	1319	
172	11001009904	11	District of Co	1	District of Co	9904		2	0.164	0	2243	0	0	2243	1508	
173	11001009905	11	District of Co	1	District of Co	9905		3	0.167	0	2795	0	0	2795	2510	
174	11001009906	11	District of Co	1	District of Co	9906		1	0.096	0	1296	0	0	1296	1455	
175	11001010400	11	District of Co	1	District of Co	10400		2	1.056	0	4365	0	0	4365	4368	
176	11001010900	11	District of Co	1	District of Co	10900		2	0.918	0	3211	0	0	3211	3470	
177	11001011100	11	District of Co	1	District of Co	11100		3	2.205	0	4859	0	0	4859	4811	
178	11001000201	11	District of Co	1	District of Co	201		1	0.209	0	3916	0	0	3916	4074	
179	11001006202	11	District of Co	1	District of Co	6202		1	2.542	0	33	0	0	33	0	
180	11001006804	11	District of Co	1	District of Co	6804		1	0.596	0	3670	0	0	3670	2769	

Linguistic Isolation

- What if you want to identify areas that may need support for a language other than English?
- Find tracts in that have a high percentage of housing units where no one over the age of 14 speaks English “very well”
- What language is spoken in these tracts?

Linguistically Isolated Tracts in DC

Rank	Tract	No one speaks English "very well"	Spanish	Indo-European	Asian/Pacific Islander	Other
1	001803	22.45 (6.83)	12.87 (5.25)	1.93 (1.80)	0.56 (0.93)	7.09 (3.91)
2	002802	16.01 (5.66)	13.49 (5.43)	1.01 (1.16)	0.81 (0.96)	0.70 (0.85)
3	002502	15.40 (6.35)	14.40 (6.11)	0.00 (0.90)	0.00 (0.90)	1.00 (1.60)
4	001804	11.32 (5.80)	4.74 (3.52)	0.16 (0.47)	0.00 (0.95)	6.42 (4.51)
5	002202	11.22 (5.42)	10.22 (5.12)	0.00 (1.01)	0.00 (1.01)	1.01 (1.55)

Low Response Score

Managing the PDB

It's a BIG dataset

Block Group Level

220,354 block groups X 329 variables =

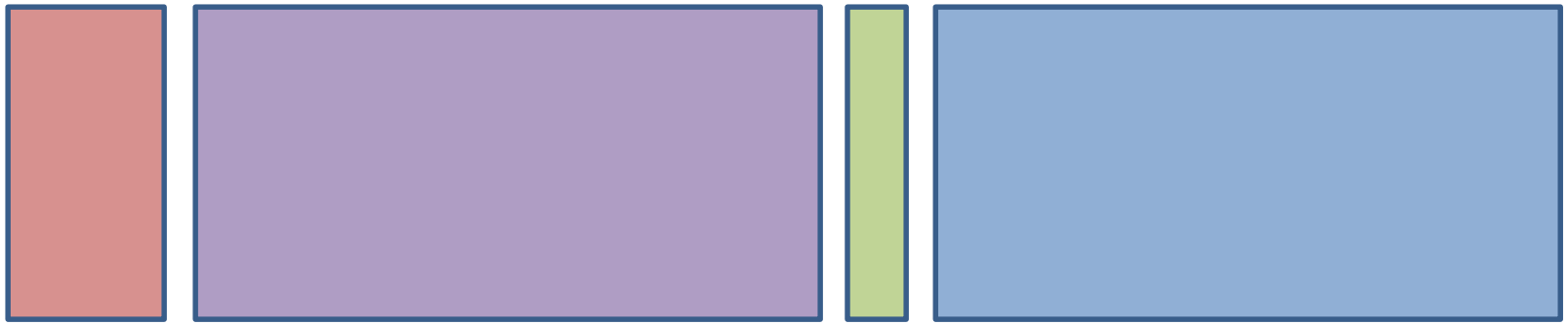
~72.4 **Million** cells

Tract Level

74,021 tracts X 550 variables =

~40.7 **Million** cells

The Structure



Geography Identifiers

- GIDBG (12 chars) = State (2 chars) + County (3 chars) + Tract (6 chars) + Block Group (1 char)
- GIDTR (11 chars) = State (2 chars) + County (3 chars) + Tract (6 chars)

Demographic, Socioeconomic, and Housing data.

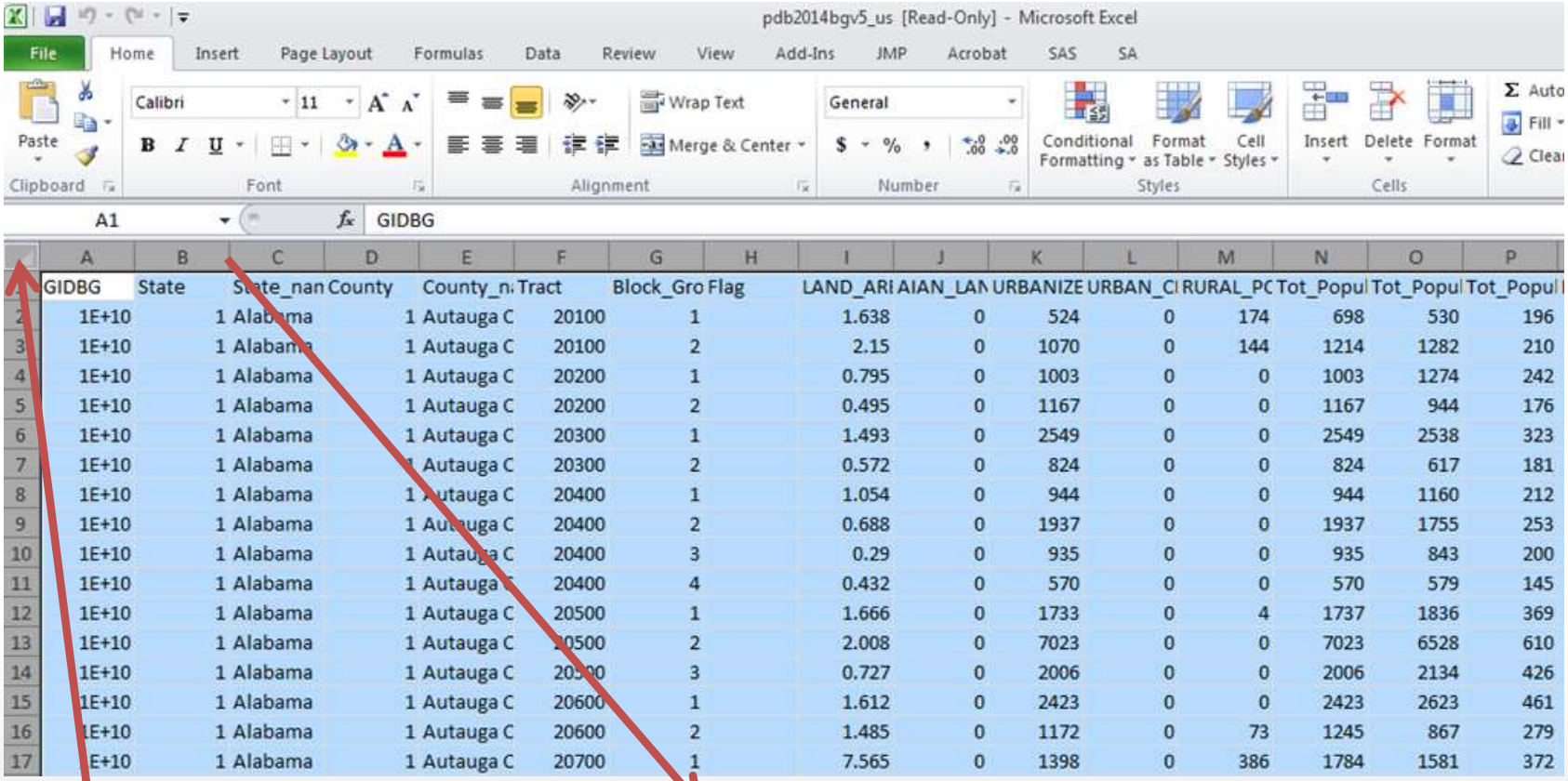
- Order of variables is consistent. Census data first, followed by ACS estimates and ACS MOEs.
- For example, Males_CEN_2010, Males_ACS_08_12, Males_ACSMOE_08_12

Census Operational data including Mail Return Rate and Low Response Score

Percentages and MOE Percentages. Listed in the same order as their respective estimate.

- Variables identified with 'pct_' added to their variable name.
- For example, pct_Males_CEN_2010, pct_Males_ACS_08_12, pct_Males_ACSMOE_08_12

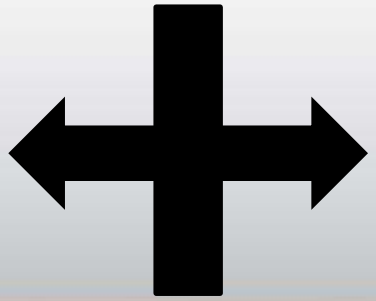
Managing Columns



GIDBG	State	State_nan	County	County_n	Tract	Block_Gro	Flag	LAND_ARI	AIAN_LAN	URBANIZE	URBAN_CI	RURAL_PC	Tot_Popul	Tot_Popul	Tot_Popul
1E+10	1	Alabama	1	Autauga C	20100	1		1.638	0	524	0	174	698	530	196
1E+10	1	Alabama	1	Autauga C	20100	2		2.15	0	1070	0	144	1214	1282	210
1E+10	1	Alabama	1	Autauga C	20200	1		0.795	0	1003	0	0	1003	1274	242
1E+10	1	Alabama	1	Autauga C	20200	2		0.495	0	1167	0	0	1167	944	176
1E+10	1	Alabama	1	Autauga C	20300	1		1.493	0	2549	0	0	2549	2538	323
1E+10	1	Alabama	1	Autauga C	20300	2		0.572	0	824	0	0	824	617	181
1E+10	1	Alabama	1	Autauga C	20400	1		1.054	0	944	0	0	944	1160	212
1E+10	1	Alabama	1	Autauga C	20400	2		0.688	0	1937	0	0	1937	1755	253
1E+10	1	Alabama	1	Autauga C	20400	3		0.29	0	935	0	0	935	843	200
1E+10	1	Alabama	1	Autauga C	20400	4		0.432	0	570	0	0	570	579	145
1E+10	1	Alabama	1	Autauga C	20500	1		1.666	0	1733	0	4	1737	1836	369
1E+10	1	Alabama	1	Autauga C	20500	2		2.008	0	7023	0	0	7023	6528	610
1E+10	1	Alabama	1	Autauga C	20500	3		0.727	0	2006	0	0	2006	2134	426
1E+10	1	Alabama	1	Autauga C	20600	1		1.612	0	2423	0	0	2423	2623	461
1E+10	1	Alabama	1	Autauga C	20600	2		1.485	0	1172	0	73	1245	867	279
1E+10	1	Alabama	1	Autauga C	20700	1		7.565	0	1398	0	386	1784	1581	372

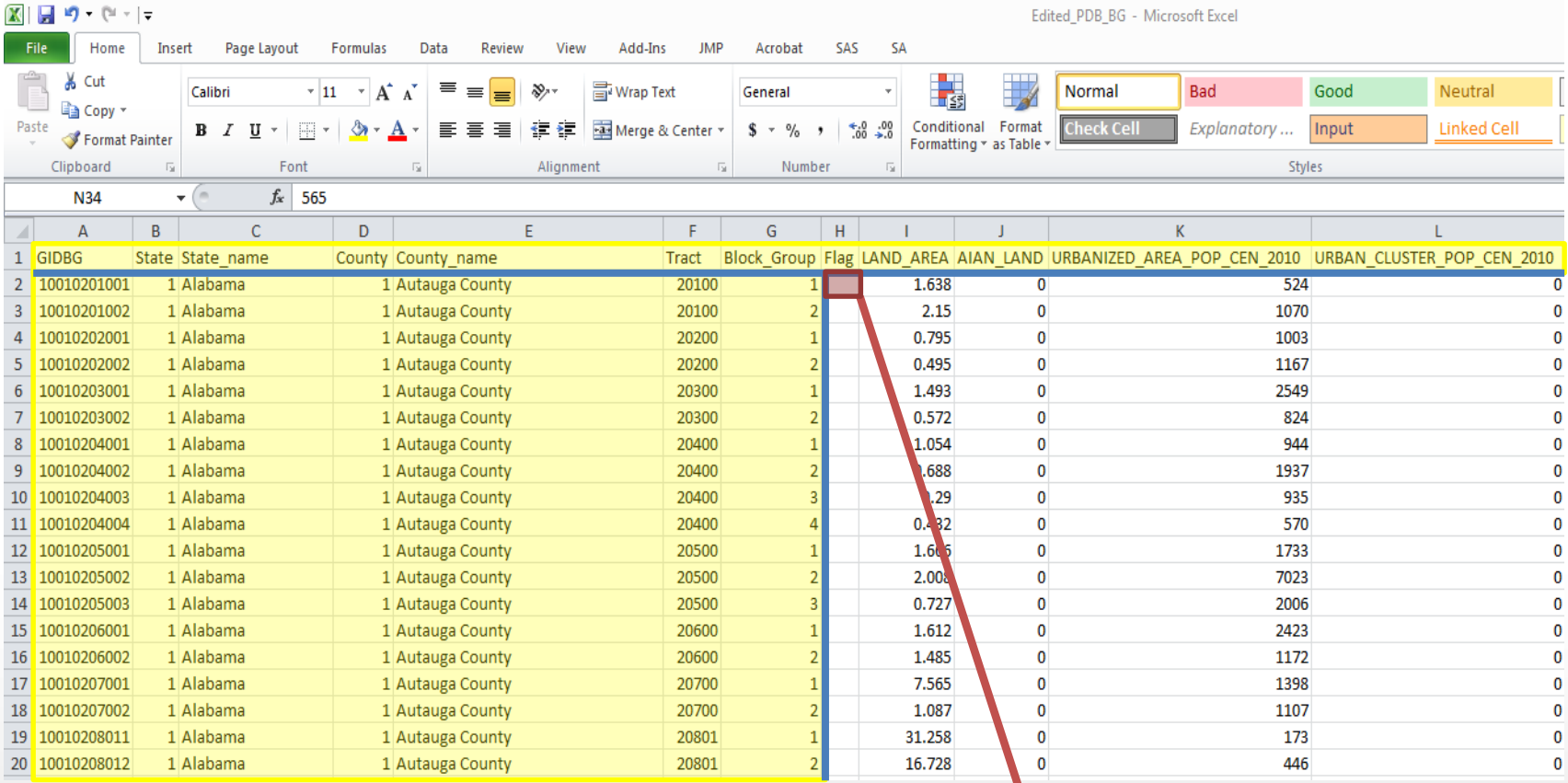
Click this box to highlight all cells in your active workbook

When curser is between columns, icon is…



Double click to expand columns to show the whole variable name.

Row/Column Lock (A.K.A. Freeze Panes)

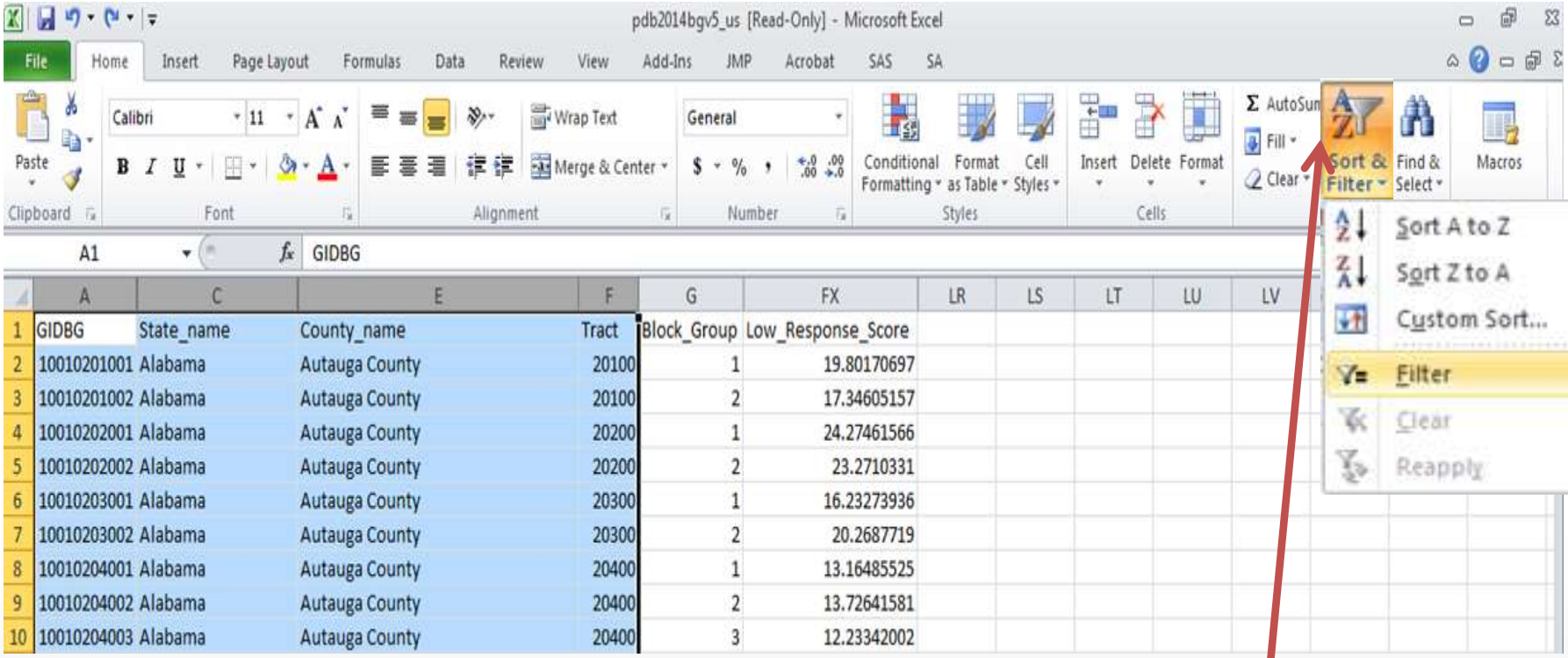


	A	B	C	D	E	F	G	H	I	J	K	L
1	GIDBG	State	State_name	County	County_name	Tract	Block_Group	Flag	LAND_AREA	AIAN_LAND	URBANIZED_AREA_POP_CEN_2010	URBAN_CLUSTER_POP_CEN_2010
2	10010201001	1	Alabama	1	Autauga County	20100	1		1.638	0	524	0
3	10010201002	1	Alabama	1	Autauga County	20100	2		2.15	0	1070	0
4	10010202001	1	Alabama	1	Autauga County	20200	1		0.795	0	1003	0
5	10010202002	1	Alabama	1	Autauga County	20200	2		0.495	0	1167	0
6	10010203001	1	Alabama	1	Autauga County	20300	1		1.493	0	2549	0
7	10010203002	1	Alabama	1	Autauga County	20300	2		0.572	0	824	0
8	10010204001	1	Alabama	1	Autauga County	20400	1		1.054	0	944	0
9	10010204002	1	Alabama	1	Autauga County	20400	2		1.688	0	1937	0
10	10010204003	1	Alabama	1	Autauga County	20400	3		1.29	0	935	0
11	10010204004	1	Alabama	1	Autauga County	20400	4		0.482	0	570	0
12	10010205001	1	Alabama	1	Autauga County	20500	1		1.606	0	1733	0
13	10010205002	1	Alabama	1	Autauga County	20500	2		2.008	0	7023	0
14	10010205003	1	Alabama	1	Autauga County	20500	3		0.727	0	2006	0
15	10010206001	1	Alabama	1	Autauga County	20600	1		1.612	0	2423	0
16	10010206002	1	Alabama	1	Autauga County	20600	2		1.485	0	1172	0
17	10010207001	1	Alabama	1	Autauga County	20700	1		7.565	0	1398	0
18	10010207002	1	Alabama	1	Autauga County	20700	2		1.087	0	1107	0
19	10010208011	1	Alabama	1	Autauga County	20801	1		31.258	0	173	0
20	10010208012	1	Alabama	1	Autauga County	20801	2		16.728	0	446	0

We want to be able to consistently see the geo identifiers and the column names as we view the data.

Cell H2 is the intersection cell we will use to lock on.

Filters

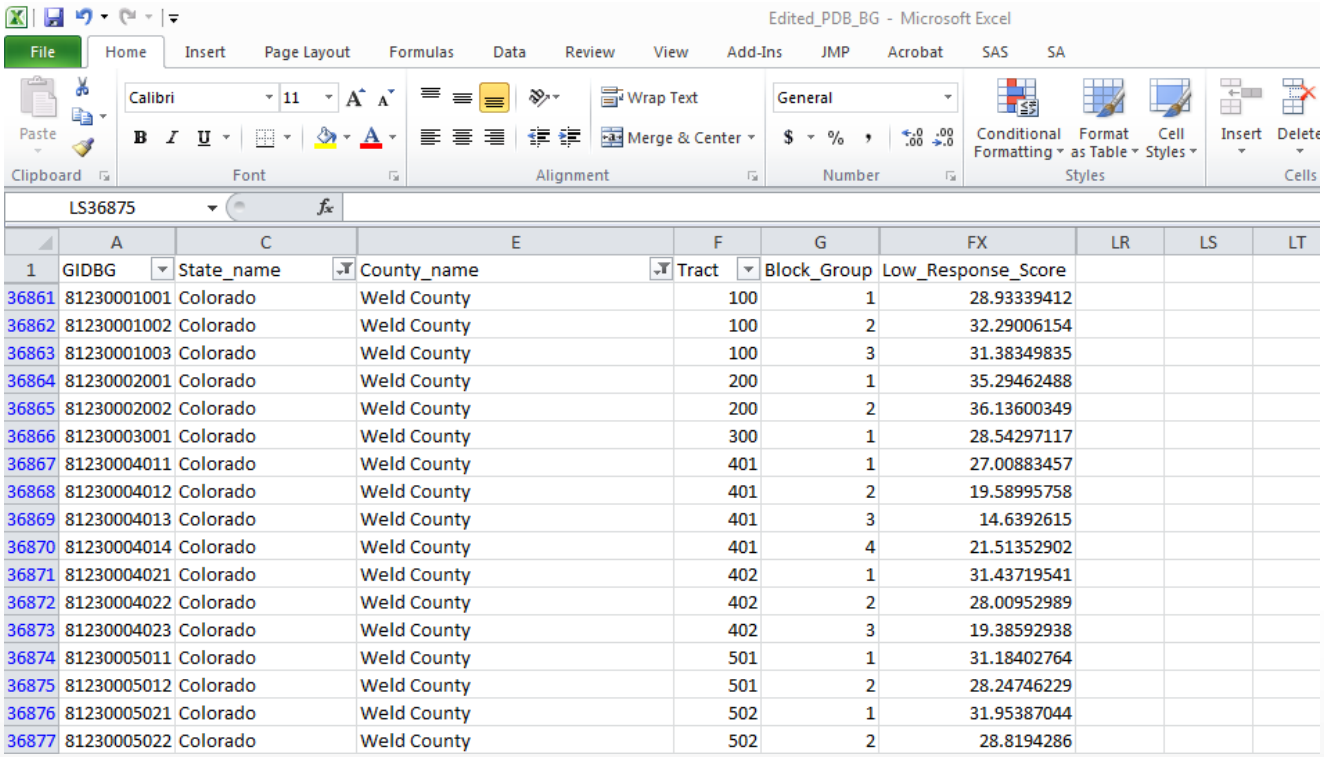


	A	C	E	F	G	FX	LR	LS	LT	LU	LV
1	GIDBG	State_name	County_name	Tract	Block_Group	Low_Response_Score					
2	10010201001	Alabama	Autauga County	20100	1	19.80170697					
3	10010201002	Alabama	Autauga County	20100	2	17.34605157					
4	10010202001	Alabama	Autauga County	20200	1	24.27461566					
5	10010202002	Alabama	Autauga County	20200	2	23.2710331					
6	10010203001	Alabama	Autauga County	20300	1	16.23273936					
7	10010203002	Alabama	Autauga County	20300	2	20.2687719					
8	10010204001	Alabama	Autauga County	20400	1	13.16485525					
9	10010204002	Alabama	Autauga County	20400	2	13.72641581					
10	10010204003	Alabama	Autauga County	20400	3	12.23342002					

Filters give you the ability to only view data which match specified conditioning.

- Highlight all geo identifiers you which to be able to able a filter on.
- Then select ‘Sort & Filter’ from the Home tab and then select Filter.

Identifying LRS Groups

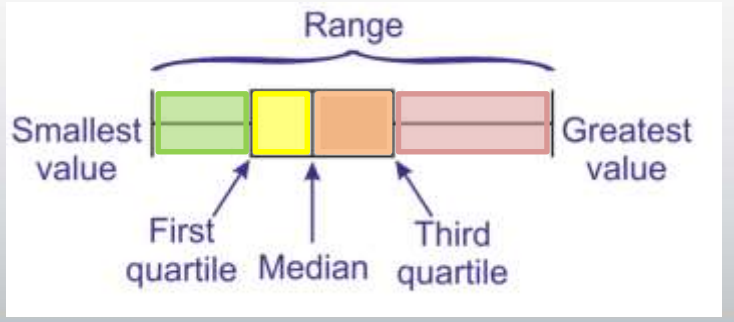


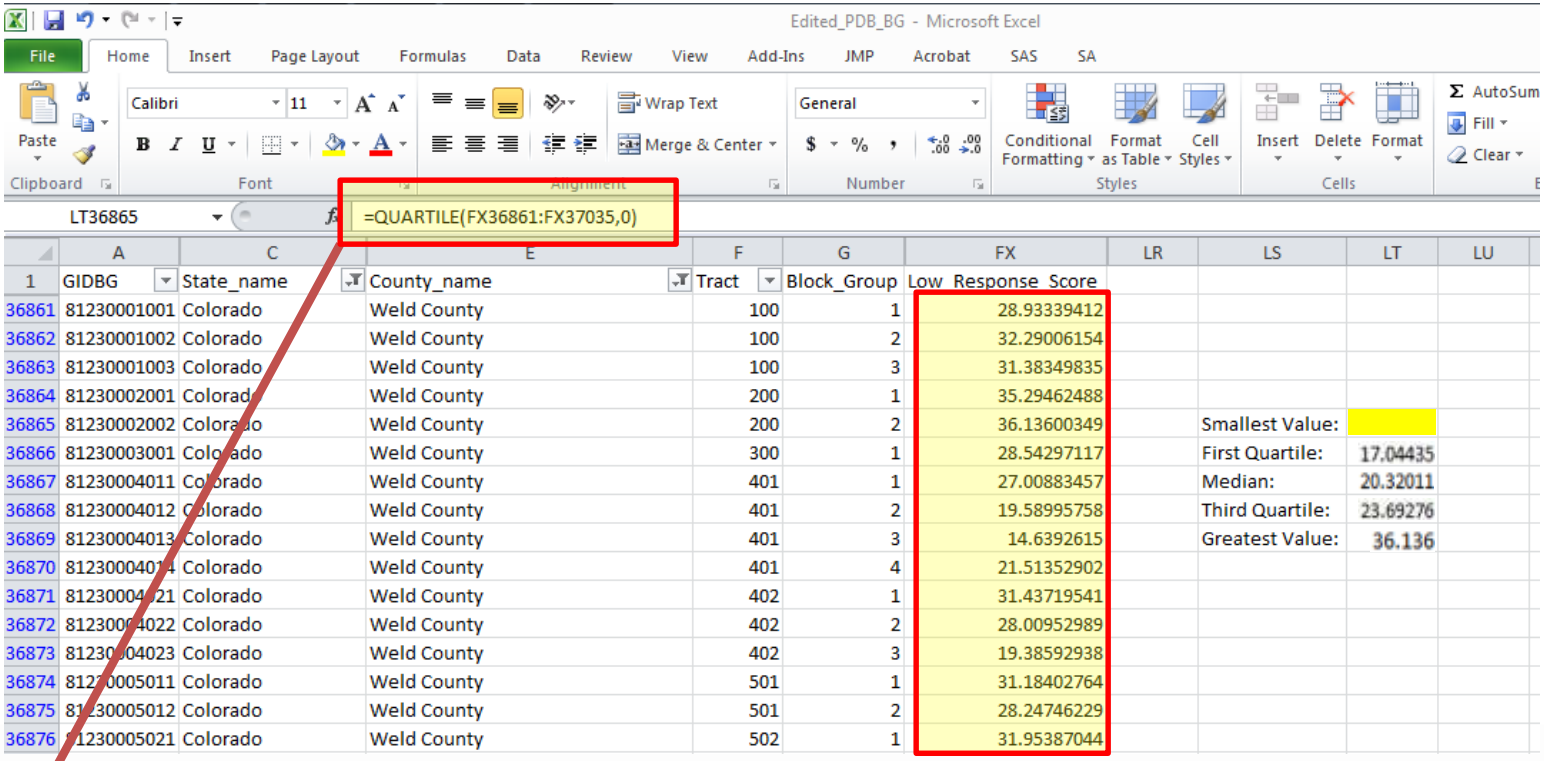
	A	C	E	F	G	FX	LR	LS	LT
1	GIDBG	State_name	County_name	Tract	Block_Group	Low_Response_Score			
36861	81230001001	Colorado	Weld County	100	1	28.93339412			
36862	81230001002	Colorado	Weld County	100	2	32.29006154			
36863	81230001003	Colorado	Weld County	100	3	31.38349835			
36864	81230002001	Colorado	Weld County	200	1	35.29462488			
36865	81230002002	Colorado	Weld County	200	2	36.13600349			
36866	81230003001	Colorado	Weld County	300	1	28.54297117			
36867	81230004011	Colorado	Weld County	401	1	27.00883457			
36868	81230004012	Colorado	Weld County	401	2	19.58995758			
36869	81230004013	Colorado	Weld County	401	3	14.6392615			
36870	81230004014	Colorado	Weld County	401	4	21.51352902			
36871	81230004021	Colorado	Weld County	402	1	31.43719541			
36872	81230004022	Colorado	Weld County	402	2	28.00952989			
36873	81230004023	Colorado	Weld County	402	3	19.38592938			
36874	81230005011	Colorado	Weld County	501	1	31.18402764			
36875	81230005012	Colorado	Weld County	501	2	28.24746229			
36876	81230005021	Colorado	Weld County	502	1	31.95387044			
36877	81230005022	Colorado	Weld County	502	2	28.8194286			

Here I have hidden all columns except for some geo locating variables and the Low Response Score.

I have also applied a filter to only show data from Weld County, CO.

Lets use Excel to help us locate and identify those block groups with a Low Response Scores that fall within the quartile groups.





	A	C	E	F	G	FX	LR	LS	LT	LU
1	GIDBG	State_name	County_name	Tract	Block_Group	Low_Response_Score				
36861	81230001001	Colorado	Weld County	100	1	28.93339412				
36862	81230001002	Colorado	Weld County	100	2	32.29006154				
36863	81230001003	Colorado	Weld County	100	3	31.38349835				
36864	81230002001	Colorado	Weld County	200	1	35.29462488				
36865	81230002002	Colorado	Weld County	200	2	36.13600349		Smallest Value:		
36866	81230003001	Colorado	Weld County	300	1	28.54297117		First Quartile:	17.04435	
36867	81230004011	Colorado	Weld County	401	1	27.00883457		Median:	20.32011	
36868	81230004012	Colorado	Weld County	401	2	19.58995758		Third Quartile:	23.69276	
36869	81230004013	Colorado	Weld County	401	3	14.6392615		Greatest Value:	36.136	
36870	81230004014	Colorado	Weld County	401	4	21.51352902				
36871	81230004021	Colorado	Weld County	402	1	31.43719541				
36872	81230004022	Colorado	Weld County	402	2	28.00952989				
36873	81230004023	Colorado	Weld County	402	3	19.38592938				
36874	81230005011	Colorado	Weld County	501	1	31.18402764				
36875	81230005012	Colorado	Weld County	501	2	28.24746229				
36876	81230005021	Colorado	Weld County	502	1	31.95387044				

=QUARTILE(Starting_Cell:Ending_Cell,0) ➔ Smallest Value

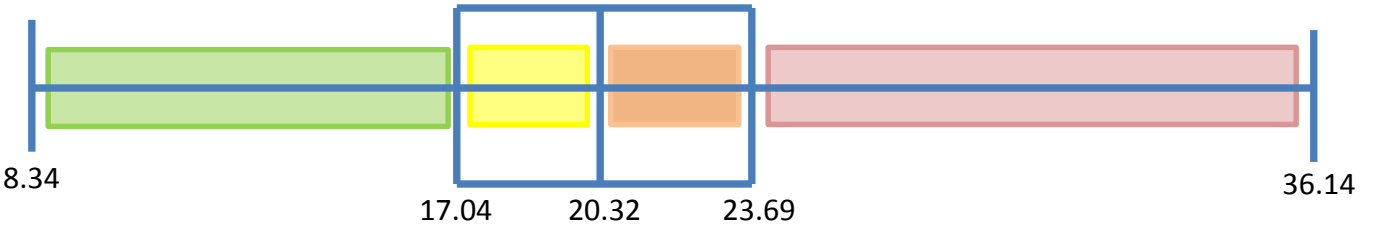
=QUARTILE(Starting_Cell:Ending_Cell,1) ➔ First Quartile

=QUARTILE(Starting_Cell:Ending_Cell,2) ➔ Median

=QUARTILE(Starting_Cell:Ending_Cell,3) ➔ Third Quartile

=QUARTILE(Starting_Cell:Ending_Cell,4) ➔ Greatest Value

Tip: When selecting your starting and ending cells, after selecting your starting cell hold shift and scroll to your ending cell and select. This will highlight all cells in between.



Edited_PDB_BG - Microsoft Excel

File | Home | Insert | Page Layout | Formulas | Data | Review | View | Add-Ins | JMP | Acrobat | SAS | SA

Calibri 11 | Wrap Text | General | Conditional Formatting | Format as Table | Cell Styles | Insert | Delete | Format

Clipboard | Font | Alignment | Number | Styles | Cells

LW36871

	A	C	E	F	G	FX	LR	LS	LT	LU
1	GIDBG	State_name	County_name	Tract	Block_Group	Low_Response_Score				
36861	81230001001	Colorado	Weld County	100	1	28.93339412				
36862	81230001002	Colorado	Weld County	100	2	32.29006154				
36863	81230001003	Colorado	Weld County	100	3	31.38349835				
36864	81230002001	Colorado	Weld County	200	1	35.29462488				
36865	81230002002	Colorado	Weld County	200	2	36.13600349		Smallest Value:	8.339976	
36866	81230003001	Colorado	Weld County	300	1	28.54297117		First Quartile:	17.04435	
36867	81230004011	Colorado	Weld County	401	1	27.00883457		Median:	20.32011	
36868	81230004012	Colorado	Weld County	401	2	19.58995758		Third Quartile:	23.69276	
36869	81230004013	Colorado	Weld County	401	3	14.6392615		Greatest Value:	36.136	
36870	81230004014	Colorado	Weld County	401	4	21.51352902				
36871	81230004021	Colorado	Weld County	402	1	31.43719541				
36872	81230004022	Colorado	Weld County	402	2	28.00952989				
36873	81230004023	Colorado	Weld County	402	3	19.38592938				
36874	81230005011	Colorado	Weld County	501	1	31.18402764				
36875	81230005012	Colorado	Weld County	501	2	28.24746229				
36876	81230005021	Colorado	Weld County	502	1	31.95387044				
36877	81230005022	Colorado	Weld County	502	2	28.8194286				
36878	81230006001	Colorado	Weld County	600	1	33.52910431				
36879	81230007011	Colorado	Weld County	701	1	30.9415182				
36880	81230007012	Colorado	Weld County	701	2	28.78310569				
36881	81230007031	Colorado	Weld County	703	1	26.13298805				
36882	81230007032	Colorado	Weld County	703	2	32.75918394				

Questions?

Census.PDB.questions@census.gov