

THE ADVANTAGES OF ACCESSING ACS SUMMARY DATA VIA FTP AND METHODS FOR CONNECTING ACROSS TIME

ACS DATA USERS CONF. 30 MAY 2014
JOHN P. DEWITT – U. OF MICHIGAN




BACKGROUND

- Univ. of Michigan Population Studies Center
- Principal Data Sources: ACS, Decennial Census, CPS
- ACS Needs:
 - Spatial: Full nation, multiple summary levels
 - Topics: migration, educ. attainment, income, many others
 - Schedule: often tight, consistency and speed valued
 - Table Types: basic/detailed table products
 - Temporal: 1/3/5-yr estimates, ability to connect to 1990, 2000

ACCESSING ACS SUMMARY DATA

- American FactFinder
- Third Party Portals
 - NHGIS (MPC)
 - Social Explorer
- **FTP**: Best option for specific set of users
 - High investment, high payoff
 - Access to all geos and summary levels
 - Easily connect tables
 - Filter based on criteria



The screenshot shows a web browser window displaying the FTP directory listing for the path `ftp://ftp.census.gov/acs2012_5yr/summaryfile/`. The title of the page is "Index of /acs2012_5yr/summaryfile/". The listing is organized into three columns: "Name", "Size", and "Date Modified".

Name	Size	Date Modified
[parent directory]		
2008-2012_ACSSF_All_In_2_Giant_Files(Experienced-Users-Only/		12/16/13 9:31:00 AM
2008-2012_ACSSF_By_State_All_Tables/		12/12/13 4:48:00 PM
2008-2012_ACSSF_By_State_By_Sequence_Table_Subset/		12/12/13 4:36:00 PM
ACS2012_5-Year_TableShells.xls	2.4 MB	12/17/13 1:17:00 PM
ACS_2008-2012_SF_Tech_Doc.pdf	3.2 MB	12/12/13 1:39:00 PM
SequenceNumberTableNumberLookup.sas7bdat	6.6 MB	11/7/13 9:55:00 AM
Sequence_Number_and_Table_Number_Lookup.txt	1.4 MB	11/7/13 9:55:00 AM
Sequence_Number_and_Table_Number_Lookup.xls	3.0 MB	11/7/13 10:00:00 AM
UserTools/		12/12/13 4:28:00 PM

ACS SUMMARY DATA FROM FTP

- Several Step Process for SAS
 1. Download
 2. Unpack/Unzip
 3. Modify code (it's not as scary as it sounds)
 4. Run code
- Once run, access to single measure to entire summary file
- Process time:
 - 5-yr (both): ~12 hours
 - 3-yr : 90 min
 - 1-yr: 30 min

PIECES TO DOWNLOAD

- Zipped data files
 - Geography files (text)
 - Estimates (text/csv)
 - Margins of Error (text/csv)
- Sequence/Table Lookup (SAS & Excel)
 - SAS dataset for processing
 - Excel file for human reading
- SAS macro code (text/SAS) – will require editing

HOW TO USE THE PIECES

- Data files – all are state-iterated

- Geographic files
- “Sequence” Files
 - Estimates
 - Margins of Error

Sequence File 00X					
Linking Data		B01001		B01002	
State	Record	Est 1	Est x	Est 1	Est x

- Interpretation file

- Sequence/Table Lookup
- Custom

File ID	Table ID	Sequence Number	Line Number	Start Position	Total Cells in Table	Total Cells in Sequence	Table Title
ACSSF	C08119	35		191	54 CELLS	238	MEANS OF TRANSPORTATION TO WORK BY WORKERS' EARNI
ACSSF	C08119	35					Universe: Workers 16 years and over with earnings
ACSSF	C08119	35	1				Total:
ACSSF	C08119	35	2				\$1 to \$9,999 or loss
ACSSF	C08119	35	3				\$10,000 to \$14,999
ACSSF	C08119	35	4				\$15,000 to \$24,999

- Processing

- SAS macro code: provides great starting point
- Custom

WITH EDITING

- Additions to SAS code provided by Census Bureau
- Other modifications
 - Check for redundancies
 - Macro variables
 - Increase reusability
- Result: All tables!

Name	Date modified	Type	Size
acs12_5_b00001.sas7bdat	1/17/2014 4:58 PM	SAS Data Set	70,504 KB
acs12_5_b00002.sas7bdat	1/17/2014 4:58 PM	SAS Data Set	70,440 KB
acs12_5_b01001.sas7bdat	1/17/2014 4:58 PM	SAS Data Set	227,816 KB
acs12_5_b01001a.sas7bdat	1/17/2014 4:58 PM	SAS Data Set	170,888 KB
acs12_5_b01001b.sas7bdat	1/17/2014 4:58 PM	SAS Data Set	138,608 KB
acs12_5_b01001c.sas7bdat	1/17/2014 4:58 PM	SAS Data Set	128,448 KB
acs12_5_b01001d.sas7bdat	1/17/2014 4:59 PM	SAS Data Set	132,248 KB
acs12_5_b01001e.sas7bdat	1/17/2014 4:59 PM	SAS Data Set	122,528 KB
acs12_5_b01001f.sas7bdat	1/17/2014 4:59 PM	SAS Data Set	132,408 KB
acs12_5_b01001g.sas7bdat	1/17/2014 4:59 PM	SAS Data Set	135,288 KB
acs12_5_b01001h.sas7bdat	1/17/2014 4:59 PM	SAS Data Set	168,768 KB
acs12_5_b01001i.sas7bdat	1/17/2014 4:59 PM	SAS Data Set	141,688 KB
acs12_5_b01002.sas7bdat	1/17/2014 5:00 PM	SAS Data Set	81,056 KB

BENEFITS OF ACCESS TO FULL SF

- Access to all tables → Can calculate/create custom measures
- Custom measures → connect more easily across time
 - Tables are *mostly* consistent across years for same release
 - More variation between releases from the same year
- Can calculate and connect measures directly within SAS code or...

CONNECTING ACROSS TIME

- Maintain custom measures and their calculations in Excel sheet → Import and process with SAS
 - Separate and Organize the measures you want
 - Import spreadsheets into SAS
 - Process measures
- Code/custom measures can be reused

	A	B	C	D	E
1	measure	label	tableid	pos	
2	B0001	Total Population	B01003	1	
3	B0701	Total Pop for Mobility 1yr ago	C07401	1	
4	B0702	Total 25-34 1yr ago	C07401	5	
5	B0703	Total 55-64 1yr ago	C07401	8	
6	B0704	Total 65-74 1yr ago	C07401	9	
7	B0705	Total 75+ 1yr ago	C07401	10	

	A	B	C
1	measure	label	calculation
2	B00101	Net Migration All	(B1711 + B1712) - (B0711 + B0712)
3	B00102	Net Migration Age 25 to 34	(B1713+B1714) - (B0713+B0714)
4	B00103	Net Migration Age 55 and above	(B1715+B1716+B1717+B1718+B1719+B1720) - (B0715+B0716+B0717+B0718+B0719+B0720)
5	B00104	Net Migration College Grads (25+)	(B1733+B1734+B1735+B1736) - (B0733+B0734+B0735+B0736)
6	B00201	Net Migration Rate All	100* (B00101 / B1701)

RESULTS OF RUN

- Attach Two Columns: Year and Span (e.g. 2012 & 5)
- Highly customizable
 - Remove “garbage” columns
 - Labels/Variable names of own choosing

	FIPS	File Identification	State Postal Abbrevia	Summary Level	geographic Component	Logical Record Number	State (FIPS Code)	County	Place (FIPS Code)	Area Name	GEOID	TotalPopulation	TotalPopforMobility1yrago	Total25-341yrago	Total55-	
1	01	ACSSF	al	040	00	0000001	01			Alabama	0400000US01	4803488		4726550	610674	604212
2	02	ACSSF	ak	040	00	0000001	02			Alaska	0400000US02	723120		760924	116569	94168
3	04	ACSSF	az	040	00	0000001	04			Arizona	0400000US04	6477128		6324344	858604	736968
4	05	ACSSF	ar	040	00	0000001	05			Arkansas	0400000US05	2936822		2887255	374686	359412
5	06	ACSSF	ca	040	00	0000001	06			California	0400000US06	37686586		37064394	5352237	xxxxxx
6	08	ACSSF	co	040	00	0000001	08			Colorado	0400000US08	5117453		4976541	720252	618668
7	09	ACSSF	ct	040	00	0000001	09			Connecticut	0400000US09	3584561		3540260	425453	459938
8	10	ACSSF	de	040	00	0000001	10			Delaware	0400000US10	908351		887448	112747	114060

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	FIPS	FILEID	STUSAB	SUMLEVEL	STATE	NAME	GEOID	B0001	B0701	B0702	B0703	B0704	B0705	B0711	B0712	B0713	B0714	B0715	B0716	B0717	B0718	B0719	B0720	B0731	B0732
2	1	ACSSF	al	40	1	Alabama	0400000US01	4803488	4726550	610674	604212	383413	292595	158147	105044	35870	24081	8646	6569	4344	3032	3630	2658	451098	265312
3	2	ACSSF	ak	40	2	Alaska	0400000US02	723120	760924	116569	94168	38743	21934	19233	88797	4024	20177	1621	5624	312	1551	228	1280	83660	47678
4	4	ACSSF	az	40	4	Arizona	0400000US04	6477128	6324344	858604	736968	511965	395265	120998	198861	25274	43666	8413	14928	4628	7396	2892	7703	701901	400896
5	5	ACSSF	ar	40	5	Arkansas	0400000US05	2936822	2887255	374686	359412	241812	188437	105535	68335	22546	13707	5816	4717	2586	2995	3132	2708	156825	132363
6	6	ACSSF	ca	40	6	California	0400000US06	37686586	37064394	5352237	4187238	2395204	2014265	1065938	569939	259710	137270	64227	40666	30121	19345	25816	12805	4683937	2685545

SUMMARY

- Import and Process Summary File
- Create Custom Measures tables
- Toolbox
 - Code to calculate measures of change, range, etc
 - Sequence/Table Lookup comparison
 - Identify New tables, Lost tables, changed tables
 - Compare only tables used in custom measures
 - Code to process all tables for tracts and block groups for spatial comparability over time
 - Calculate margins of error for custom measures
- Future Directions
 - How we can share data
 - How we can share supplemental tools

TRY IT!

- Download a 1Yr Release and Try it!
- Code is easily adjusted for new releases
- Similar process for Censuses 1990, 2000, 2010
- Don't be scared of trying!

Contact:

John Paul DeWitt

jpdewitt@umich.edu