2009 ACS 1-Year Summary File: Technical Documentation

American Community Survey Office

U.S. Census Bureau

Version: 1.0
Date: September 27, 2010

Document History

Version Number	Date	Summary of Changes
1.0	September 27, 2010	Baseline Document

Table of Contents

Chapter 1 - Abstract

INTRODUCTION

The American Community Survey (ACS) is a relatively new survey conducted by the U.S. Census Bureau. It uses a series of monthly samples to produce annually updated data for the same small areas (census tracts and block groups) formerly surveyed via the Decennial Census long-form sample. Five years of data collection will be required to produce these small-area data. Once the Census Bureau has collected 5 years of data, new small-area data will be produced annually. The Census Bureau will produce 3-year and 1-year data products for larger geographic areas. The ACS includes people living in both housing units (HUs) and group quarters (GQs). The ACS is conducted throughout the United States and in Puerto Rico, where it is called the Puerto Rico Community Survey (PRCS). For ease of discussion, the term ACS is used here to represent both surveys.

The American Community Survey releases meaningful data products on the Census Bureau's American FactFinder website. To address other requests, the American Community Survey produces the Summary File product that contains the complete set of the detailed tables in a comma delimited file format to allow for importation of this data into databases, spreadsheets, SAS datasets, and other formats. The American Community Survey's Summary File follows the approach taken for the Census 2000 Summary File 3 format. The American Community Survey Summary File consists of approximately 1,300 detailed tables of demographic, social, economic, and housing characteristics for the 1-year and 3-year products. For information about Census geographies, please visit http://www.census.gov/geo/www/cob/metadata.html .

The American Community Survey detailed tables contain both a base and collapsed version. The base tables provide detailed information on a given topic, and the collapsed tables provide a consolidated version of a base table. The base table name begins with the letter "B" and the collapsed table name begins with the letter "C" to allow for easy identification.

The American Community Survey covers a broad spectrum of geographic areas within the United States and Puerto Rico. The American Community Survey divides characteristics into demographic, social, housing, and economic topics.

Demographic:

○ Sex	○ Age
○ Households by type	○ Race
○ Relationship	○ Hispanic origin

Social:

o School enrollment	o Educational attainment
o Fertility	o Residence one year ago
o Veteran status	o Disability status
o U.S. citizenship status	o Language spoken at home
o Marital status	o Place of birth
o Year of entry	o Ancestry
o Grandparents caring for children	o World region of birth of foreign born

Economic:

o Employment status	o Commuting to work
o Class of worker	o Income and benefits
o Industry	o Occupation
o Poverty status	o Health Insurance Coverage

Housing:

o Housing occupancy	o Housing tenure
o Units in structure	o Year structure built
o Number of rooms	o Number of bedrooms
o House heating fuel	o Housing value
o Occupants per room	o Vehicles available
o Mortgage status and costs	o Utility cost
o Year householder moved into unit	o Gross rent

GEOGRAPHIC CONTENT

The American Community Survey Summary Files include all detailed tables for all geographic areas published by the 2009 1-Year ACS. The main directory of the American Community Survey's Summary File product release, located within the file transfer protocol (FTP) site, contains documentation about the Summary Files and information pertaining to the United States, each of the individual states, the District of Columbia, and Puerto Rico.

The following list outlines the hierarchical geographic levels contained within the Summary File release.

State Files

The ACS 1-Year Summary File for states contains the following geographic areas:

- State
- County

- County Subdivision
- Place
- Metropolitan Statistical Area/Micropolitan Statistical Area-State-Principal City
- New England City and Town Area-State-Principal City
- State-School District (Elementary, Secondary, and Unified)
- State-Alaska Native Regional Corporation
- Congressional District (111th Congress)
- Public Use Microdata Sample Area (PUMA)

National Files

The ACS Summary File for the U.S. contains the following geographic areas:

- United States
- Region
- Division
- American Indian Area/Alaska Native Area/Hawaiian Home Land
- Metropolitan Statistical Area/Micropolitan Statistical Area
- Metropolitan Statistical Area-Metropolitan Division
- Combined Statistical Area
- Combined New England City and Town Area
- New England City and Town Area (NECTA)-NECTA Division
- New England City and Town Area
- Urban Area

CONTACT US

Please send any technical questions or comments you have via email to: nicholas.m.spanos@census.gov. If you have questions or comments about the American Community Survey, please contact us online at http://ask.census.gov.

Chapter 2 – How to Use the ACS Summary File

INTRODUCTION

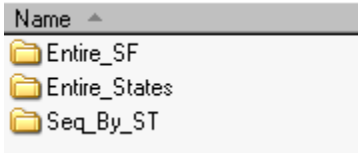
This chapter is intended to provide a guide for data users to the American Community Survey Summary Files. New users should review this chapter before attempting to use the Summary File product.

DATA FORMAT

The American Community Survey Summary Files (ACS-SF) are available in .zip file format. The Zip file format is a compressed format used to reduce the file size. Users can download the .zip files through the File Transfer Protocol (FTP) from the American Community Survey FTP site located at http://www2.census.gov/acs2009_1yr/summaryfile/. To read the Summary Files data, the compressed files must be uncompressed (unzipped) to allow access to the data.

NEW ORGANIZATION OF ACS SUMMARY FILES

Based on user feedback, the ACS Summary File is now organized into three categories



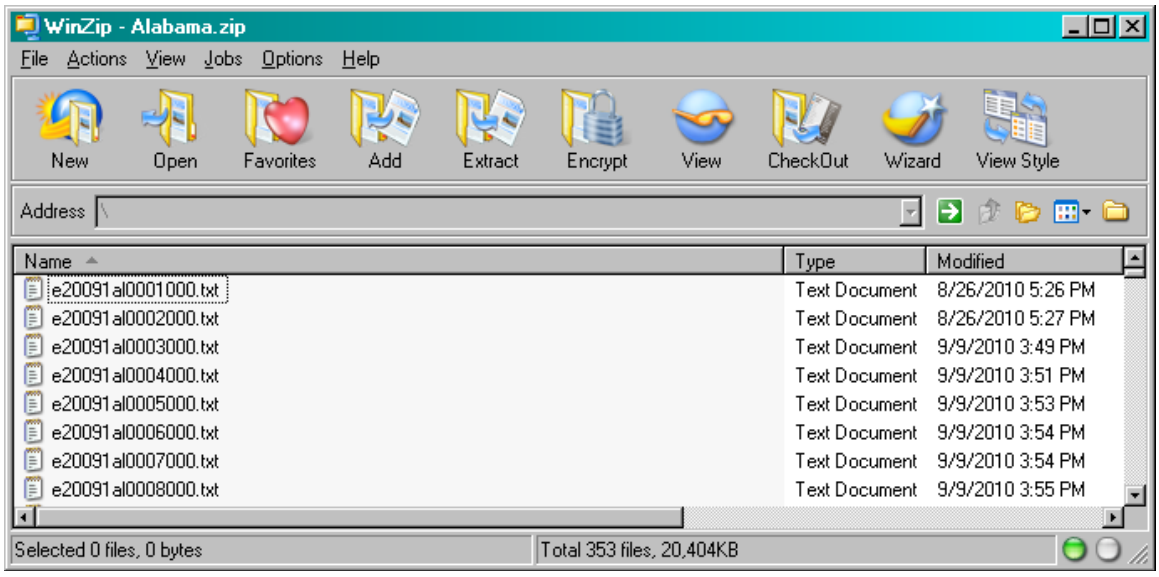
1. **Entire_SF:** This zip file contains all of the Summary File data, MOE's, and geography files - 1 File in Directory /acs2009_1yr/summaryfile/Entire_SF/*20091YRSF.zip*

For the user's convenience, this zip file contains all of the Summary Files for the U.S., Puerto Rico, states, and District of Columbia. For example, to obtain all Summary Files for nation (U.S.), states, District of Columbia, and Puerto Rico, just download and unzip the file *20091YRSF.zip*.

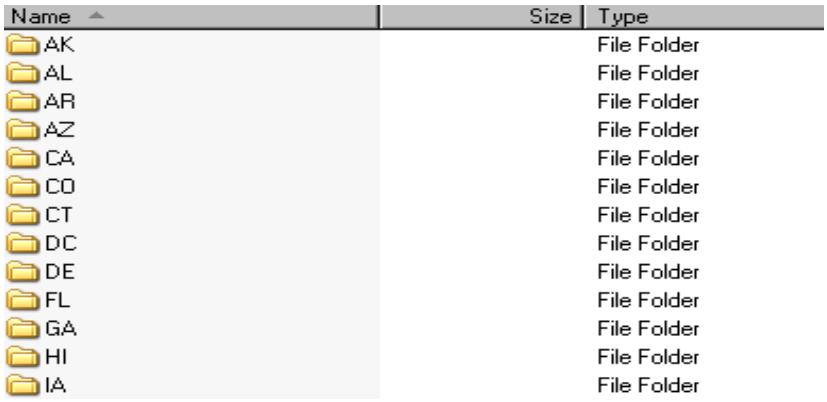
2. **Entire_States:** These files contain all of the summary file data, MOE's, and geography files per state - 53 Files in Directory (50 States, 1 US, DC, PR) /acs2009_1yr/summaryfile/Entire_States/<*FULL STATE NAME*>.*zip*

Name	Size	Type
Alabama.zip	5,865 KB	WinZip File
Alaska.zip	1,537 KB	WinZip File
Arizona.zip	7,840 KB	WinZip File
Arkansas.zip	3,805 KB	WinZip File
California.zip	47,385 KB	WinZip File
Colorado.zip	6,977 KB	WinZip File
Connecticut.zip	3,882 KB	WinZip File
Delaware.zip	1,458 KB	WinZip File
DistrictofColumbia.zip	475 KB	WinZip File
Florida.zip	19,386 KB	WinZip File
Georgia.zip	10,306 KB	WinZip File
Hawaii.zip	1,769 KB	WinZip File
Idaho.zip	2,561 KB	WinZip File
Illinois.zip	15,518 KB	WinZip File
Indiana.zip	8,688 KB	WinZip File

This directory contains 53 zip files that include the United States, each of the 50 states, District of Columbia, and Puerto Rico. Each zip file contains 176 estimate files, 176 Margin of Error files, and 1 geography file. See partial contents of the zip file for Alabama below.

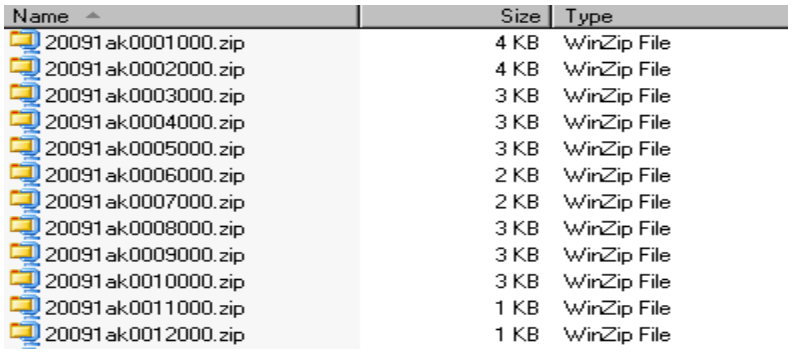


3. **Seq_By_ST:** 53 directories with 176 Compressed Sequence Files (both Estimates & MOE's) and 1 Geography file in each directory /acs2009_1yr/summaryfile/Seq_By_ST/<*ST ABBRV*>/<*Sequence ID*>.*zip*



Name	Size	Type
AK		File Folder
AL		File Folder
AR		File Folder
AZ		File Folder
CA		File Folder
CO		File Folder
CT		File Folder
DC		File Folder
DE		File Folder
FL		File Folder
GA		File Folder
HI		File Folder
IA		File Folder

Each directory contains 176 zip files and 1 geography file. Each zip file containing 1 estimate and 1 Margin of Error file. See a partial view of the AK directory below:

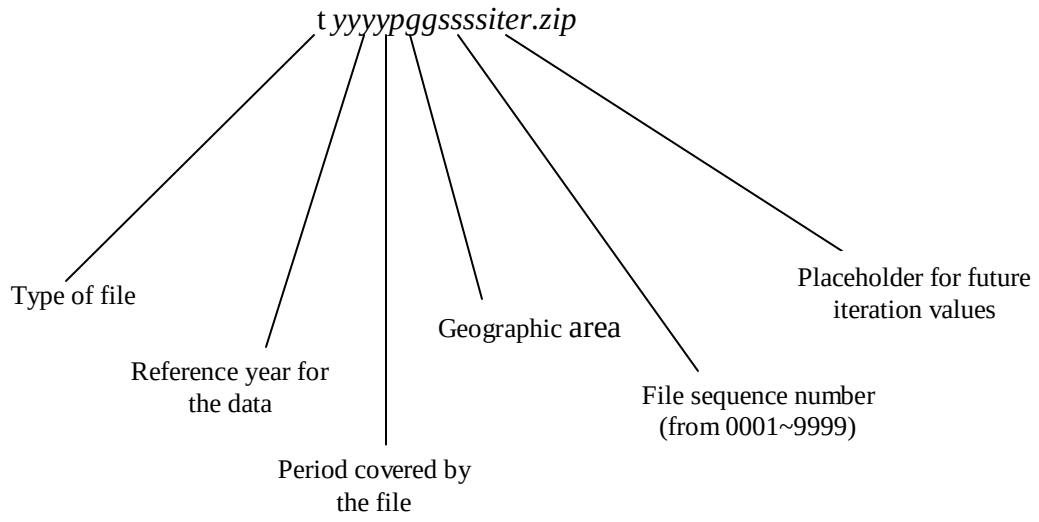


Name	Size	Type
20091ak0001000.zip	4 KB	WinZip File
20091ak0002000.zip	4 KB	WinZip File
20091ak0003000.zip	3 KB	WinZip File
20091ak0004000.zip	3 KB	WinZip File
20091ak0005000.zip	3 KB	WinZip File
20091ak0006000.zip	2 KB	WinZip File
20091ak0007000.zip	2 KB	WinZip File
20091ak0008000.zip	3 KB	WinZip File
20091ak0009000.zip	3 KB	WinZip File
20091ak0010000.zip	3 KB	WinZip File
20091ak0011000.zip	1 KB	WinZip File
20091ak0012000.zip	1 KB	WinZip File

FILE NAMING CONVENTIONS

The naming convention used for the Summary Files provides a hierarchy to allow for easier navigation to the desired information. Figures 2.1 and 2.2 provide an overview of the summary file naming convention.

Figure 2.1 Zip File Naming Convention Example

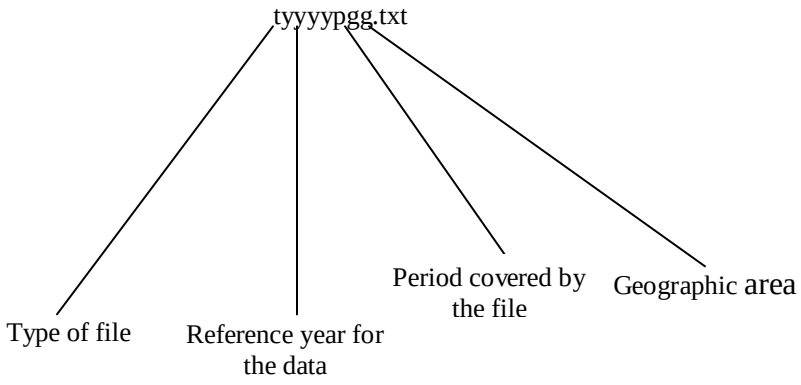


- *t* = type of data (e=estimate, m=margin of error)
- *yyyy* = reference year for the data, e.g., 2009
- *p* = period covered by the file (1=1-year, 3=3-year, 5=5-year)
- *gg* = geographic area (state or US) covered by the file; 'PR' for the Puerto Rico Community Survey
- *ssss* = file sequence number (valid range is 0001 through 9999)

Each sequence number will correspond to a series of detailed tables. Each record of this file will be for a unique geographic area published by the ACS. All the data cells contained in the detailed tables for that geographic area would be on this record. There will be a file for the estimates and a separate one for the margin of errors.

- *iter = '000'*, placeholder for future values for iteration
- *r* = data revision indicator, only present if needed. 'a', 'b', etc

Figure 2.2 Geoheader Summary File Naming Convention Example



- *t* = type of data (g= geographic header file)
- *yyyy* = reference year for the data, e.g., 2009
- *p* = period covered by the file (1=1-year, 3=3-year, 5=5-year)
- *gg* = geographic area is represented by 2 digit FIPS state code

READING THE GEOGRAPHIC FILE

The geographic file layout is below (Table 2.1). The geographic file provides the reference name, field description, field size, field starting position, and geographic summary levels that apply. The information pertaining to summary level identifies additional geographic information. Each record of the geographic file contains particular geographic area codes. For example, the reference name "PLACE" only contains information for summary levels 160 - State-Place, 312 - Metropolitan Statistical Area/Micropolitan Statistical Area-State-Principal City and 352 - New England City and Town Area-State-Principal City. For more detail on geographic definitions, see "Geographic Terms and Concepts"

Table 2.1 Geographic Header Record Layout for the ACS Summary File

Data Dictionary Reference Name	Description	Field Size	Starting Position	Geographic Summary Levels For Single-Year Tables
RECORD CODES				
FILEID	Always equal to ACS Summary File identification	6	1	All Summary Levels
STUSAB	State Postal Abbreviation	2	7	All Summary Levels
SUMLEVEL	Summary Level	3	9	All Summary Levels
COMPONENT	Geographic Component	2	12	All Summary Levels
LOGRECNO	Logical Record Number	7	14	All Summary Levels
GEOGRAPHIC AREA CODES				
US	US	1	21	010
REGION	Census Region	1	22	020
DIVISION	Census Division	1	23	030
STATECE	State (Census Code)	2	24	Blank for All Summary Levels
STATE	State (FIPS Code)	2	26	040, 050, 060,160, 230, 312, 352, 500, 795, 950, 960, 970, M01, M06, M07, M08, M11, M15, P01, P06, P07, P08, P11, P15
COUNTY	County of current residence	3	28	050, 060, M01, M08, P01, P08
COUSUB	County Subdivision (FIPS)	5	31	060, M08, P08
PLACE	Place (FIPS Code)	5	36	160, 312, 352, M07, M11, M15, P07, P11, P15
BLANK		6	41	**Reserved for future use**
BLANK		1	47	**Reserved for future use**
BLANK		5	48	**Reserved for future use**
AIANHH	American Indian Area/Alaska Native Area/ Hawaiian Home Land (Census)	4	53	250
AIANHHFP	American Indian Area/Alaska Native Area/ Hawaiian Home Land (FIPS)	5	57	250
BLANK		1	62	**Reserved for future use**
BLANK		3	63	**Reserved for future use**
BLANK		5	66	**Reserved for future use**
ANRC	Alaska Native Regional Corporation (FIPS)	5	71	230
CBSA	Metropolitan and Micropolitan Statistical Area	5	76	310, 312, 314, 332, M10, M11, M12, P10, P11, P12
CSA	Combined Statistical Area	3	81	330, M09, P09
METDIV	Metropolitan Statistical Area-Metropolitan Division	5	84	314, M12, P12
MACC	Metropolitan Area Central City	1	89	
MEMI	Metropolitan/Micropolitan Indicator Flag	1	90	010, 020, 030, 040, 314
NECTA	New England City and Town Area	5	91	350, 352, 355, M14, M15, M16, P14, P15, P16
CNECTA	New England City and Town Combined Statistical Area	3	96	335, M13, P13
NECTADIV	New England City and Town Area Division	5	99	355, M16, P16
UA	Urban Area	5	104	400
BLANK		5	109	**Reserved for future use**
CDCURR	Current Congressional District ***	2	114	500
BLANK		3	116	**Reserved for future use**
BLANK		3	119	**Reserved for future use**

BLANK		6	122	**Reserved for future use**
BLANK		3	128	**Reserved for future use**
BLANK		5	131	**Reserved for future use**
SUBMCD		5	136	**Reserved for future use**
SDELM	State-School District (Elementary)	5	141	950
SDSEC	State-School District (Secondary)	5	146	960
SDUNI	State-School District (Unified)	5	151	970
UR	Urban/Rural	1	156	010, 020, 030, 040
PCI	Principal City Indicator	1	157	010, 020, 030, 040, 312, 352
BLANK		6	158	**Reserved for future use**
BLANK		5	164	**Reserved for future use**
PUMA5	Public Use Microdata Area – 5% File	5	169	795
BLANK		5	174	**Reserved for future use**
GEOID	Geographic Identifier	40	179	All Summary Levels
NAME	Area Name	200	219	All Summary Levels

The Geographic Summary Levels labeled, as "Reserved for future use" will be used for the ACS period estimate data products.

IDENTIFICATION FIELDS WITHIN SUMMARY FILE

The first six columns of the estimates and margin of error files provide data and geographic information related to the tables contained within the same row.

- FILEID = File Identification - 6 characters
- FILETYPE= File Type - 6 characters
- STUSAB = State/U.S.-Abbreviation (USPS) - 2 characters
- CHARITER= Character Iteration - 3 characters
- SEQUENCE= Sequence Number - 4 characters
- LOGRECNO= Logical Record Number - 7 characters

A logical unique record number is assigned to all files for a specific geographic entity so that all records for a specific entity can be linked together across files. Besides the logical record number, other identifying fields are also carried over from the geographic header file to the table files. For more details on these fields, see the example below.

TABLE B08406:
SEX OF WORKERS BY MEANS OF TRANSPORTATION TO WORK FOR WORKPLACE GEOGRAPHY

FILEID	FILE TYPE	STUSAB	CHARITER	SEQUENCE	LOGRECNO	Total	Car, Truck, or Van	Drove Alone
ACSSF	2009e1	ak	000	0003	0000001	342551	275208	230130
ACSSF	2009e1	ak	000	0003	0000010	157650	141673	120945
ACSSF	2009e1	ak	000	0003	0000011	.	.	.
ACSSF	2009e1	ak	000	0003	0000012	.	.	.
ACSSF	2009e1	ak	000	0003	0000013	157650	141673	120945
ACSSF	2009e1	ak	000	0003	0000017	157650	141673	120945

SEQUENCE RULES AND TABLE ID

There are a few simple rules that are followed in the organization of tables into sequences in the ACS Summary File.

- There are never more than 245 cells in a sequence. Tables are grouped into sequences according to their subject area.
- Tables from different subject areas are not combined into the same sequence.
- Tables are grouped together when they share a collapsing arrangement. (Group B & C tables in same sequence.) For example, tables B02003 and C02003 should be in the same sequence. Place of Work (POW) tables are not grouped together with Place of Residence (POR) tables.

The data from the American Community Survey contain approximately 75 detailed tables that have different table layouts for the Puerto Rico Community Survey. The United States tables are limited to geographic areas within the United States that includes all 50 states plus the District of Columbia. The Puerto Rico tables are limited to areas within Puerto Rico. Separation of files between US and Puerto Rico was made because the comparable tables may have a different number of cells (such as C06001 with 45 cells and C06001PR with 37 cells). To identify what tables are available for Puerto Rico only look for the "PR" suffix in the table id (e.g. C06001PR). For more information on PR only sequences that are counterparts to states-only sequences, please reference the following table.

US/State-only Sequences – Missing PR Data	**PR Only Sequences – Missing US/State Data**
1 (US)	167 (PR)
2 (US)	168 (PR)
23 (US)	169(PR)
24 (US)	170 (PR)
28 (US)	171 (PR)
29 (US)	172 (PR)
30 (US)	173 (PR)
31 (US)	174 (PR)
32 (US)	175 (PR)
165 (US)	176 (PR)

Sequences 126-143 contain 6 large median tables with 499 cells. Due to the size restriction on the sequences (no more than 245 cells), each of the tables is broken down into 3 sequences. These tables are only produced for the U.S. Please reference the following table for more details.

US Only Sequences	**US Only Tables**
126	B24121
127	B24121
128	B24121
129	B24122
130	B24122
131	B24122
132	B24123
133	B24123
134	B24123
135	B24124
136	B24124
137	B24124
138	B24125
139	B24125
140	B24125
141	B24126
142	B24126
143	B24126

Chapter 3 – Table Shells

A file containing all of the table shells is available at the top level of the Summary File directory along with the technical documentation, and it is called: ACS2009TableShells.xls. Here is an example of the contents of the file.

TABLE SHELLS

Table ID	Line	Unique ID	Stub
B00001			UNWEIGHTED SAMPLE COUNT OF THE POPULATION
B00001			*Universe: Total population*
B00001	1	B00001001	Total
B00002			UNWEIGHTED SAMPLE HOUSING UNITS
B00002			*Universe: Housing units*
B00002	1	B00002001	Total
B01001			SEX BY AGE
B01001			*Universe: Total population*
B01001	1	B01001001	Total:
B01001	2	B01001002	Male:
B01001	3	B01001003	Under 5 years
B01001	4	B01001004	5 to 9 years
B01001	5	B01001005	10 to 14 years
B01001	6	B01001006	15 to 17 years
B01001	7	B01001007	18 and 19 years
B01001	8	B01001008	20 years
B01001	9	B01001009	21 years
B01001	10	B01001010	22 to 24 years
B01001	11	B01001011	25 to 29 years
B01001	12	B01001012	30 to 34 years
B01001	13	B01001013	35 to 39 years
B01001	14	B01001014	40 to 44 years
B01001	15	B01001015	45 to 49 years

Chapter 4 – Summary Level Sequence Chart

2009 American Community Survey Data Products SUMMARY LEVEL SEQUENCE

Geographic Component	Summary Level	
00, 01, 43, A0, C0, C1, C2, E0, E1, E2, G0, H0, 89, 91-94	010	United States[1]
00, 01, 43, A0, C0, C1, C2, E0, E1, E2, G0, H0	020	Region[1]
00, 01, 43, A0, C0, C1, C2, E0, E1, E2, G0, H0	030	Division[1]
00, 01, 43, A0, C0, C1, C2, E0, E1, E2, G0, H0	040	State[2]
00	050	State-County[3]
00	060	State-County-County Subdivision
00	160	State-Place
00	500	State-Congressional District (111th)
00	795	State-Public Use Microdata Sample Area (5%)
00	950	State-School District (Elementary)/Remainder[4]
00	960	State-School District (Secondary) /Remainder[4]
00	970	State-School District (Unified)/Remainder[4]
00	230	State-Alaska Native Regional Corporation
00	250	American Indian Area/Alaska Native Area/Hawaiian Home Land
00	310	Metropolitan Statistical Area/Micropolitan Statistical Area
00	312	Metropolitan Statistical Area/Micropolitan Statistical Area-State-Principal City
00	314	Metropolitan Statistical Area-Metropolitan Division
00	330	Combined Statistical Area
00	335	Combined New England City and Town Area
00	350	New England City and Town Area
00	352	New England City and Town Area-State-Principal City
00	355	New England City and Town Area (NECTA)-NECTA Division
00	400	Urban Area

[1]Land area, water area, population counts, and housing unit counts for the United States, Regions, and Divisions do not include Puerto Rico.

[2]State, District of Columbia, or Puerto Rico.

[3]Parish in Louisiana, Borough or Census Area in Alaska, and Municipio in Puerto Rico; in Maryland, Missouri, Nevada, and Virginia, one or more cities are independent of counties and are treated as statistical equivalents of counties; the entire District of Columbia, which has no counties, is treated as a county equivalent.

Geographic Components

Geographic Components	*Geographic Component Description*
00	Not in any geographic component
01	Urban
43	Rural
A0	In metropolitan or micropolitan statistical area
G0	Not in metropolitan or micropolitan statistical area
C0	In metropolitan statistical area
C1	In metropolitan statistical area—in principal city
C2	In metropolitan statistical area—not in principal city
H0	Not in metropolitan statistical area
E0	In micropolitan statistical area
E1	In micropolitan statistical area—in principal city
E2	In micropolitan statistical area—not in principal city
89	American Indian Reservation and Trust Land – Federal
91	Oklahoma Tribal Statistical Area
92	Tribal Designated Statistical Area
93	Alaska Native Village Statistical Area
94	State Designated Tribal Statistical Area (Formerly SDAISA)

[4]Remainder of school districts are published for ACS 5-year data.

Chapter 5 – Data Dictionary - List of Detailed Tables and Sequence Numbers

The Summary Files are created in the American Standard Code for Information Interchange (ASCII) format, and they contain estimates, an ASCII header, file label, and data elements identified by its file position.

The files provided within the ACS Summary File product package contain one file providing geographic information that contains one record for each geographic entity, estimate data file and margin of error file.

The geographic information file includes logical record numbers assigned to geographies sequentially. The same geography sequence numbers match related data records creating a relationship between the data and geographic information.

The documentation on tables and associated sequence numbers can be found in these 2 files: merge_5_6.xls and merge_5_6.sas7bdat, which are available at the top level of the Summary File directory along with the technical documentation. Here is an example of the contents of the merge_5_6.xls file:

Table ID	Sequence Number	Line Number	Start Position	Total Cells in Table	Table Title	Subject Area
B08406	0003		7	51 CELLS	SEX OF WORKERS BY MEANS OF TRANSPORTATION TO WORK FOR WORKPLACE GEOGRAPHY	Journey to Work
B08406	0003				Universe: Workers 16 years and over	
B08406	0003	1			Total:	
B08406	0003	2			Car, truck, or van:	
B08406	0003	3			Drove alone	
B08406	0003	4			Carpooled:	
B08406	0003	5			In 2-person carpool	
B08406	0003	6			In 3-person carpool	
B08406	0003	7			In 4-or-more-person carpool	
B08406	0003	8			Public transportation (excluding taxicab):	
B08406	0003	9			Bus or trolley bus	
B08406	0003	10			Streetcar or trolley car (carro publico in Puerto Rico)	
B08406	0003	11			Subway or elevated	
B08406	0003	12			Railroad	
B08406	0003	13			Ferryboat	
B08406	0003	14			Bicycle	

B08406	0003	15			Walked	
B08406	0003	16			Taxicab, motorcycle, or other means	
B08406	0003	17			Worked at home	
B08406	0003	18			Male:	
B08406	0003	19			Car, truck, or van:	
B08406	0003	20			Drove alone	
B08406	0003	21			Carpooled:	
B08406	0003	22			In 2-person carpool	
B08406	0003	23			In 3-person carpool	
B08406	0003	24			In 4-or-more-person carpool	
B08406	0003	25			Public transportation (excluding taxicab):	
B08406	0003	26			Bus or trolley bus	
B08406	0003	27			Streetcar or trolley car (carro publico in Puerto Rico)	
B08406	0003	28			Subway or elevated	
B08406	0003	29			Railroad	
B08406	0003	30			Ferryboat	
B08406	0003	31			Bicycle	
B08406	0003	32			Walked	
B08406	0003	33			Taxicab, motorcycle, or other means	
B08406	0003	34			Worked at home	
B08406	0003	35			Female:	
B08406	0003	36			Car, truck, or van:	
B08406	0003	37			Drove alone	
B08406	0003	38			Carpooled:	
B08406	0003	39			In 2-person carpool	
B08406	0003	40			In 3-person carpool	
B08406	0003	41			In 4-or-more-person carpool	
B08406	0003	42			Public transportation (excluding taxicab):	
B08406	0003	43			Bus or trolley bus	
B08406	0003	44			Streetcar or trolley car (carro publico in Puerto Rico)	
B08406	0003	45			Subway or elevated	
B08406	0003	46			Railroad	
B08406	0003	47			Ferryboat	
B08406	0003	48			Bicycle	
B08406	0003	49			Walked	
B08406	0003	50			Taxicab, motorcycle, or other means	
B08406	0003	51			Worked at home	
B08412	0003		58	39 CELLS	SEX OF WORKERS BY TRAVEL TIME TO WORK FOR WORKPLACE GEOGRAPHY	Journey to Work

B08412	0003				Universe: Workers 16 years and over who did not work at home	
B08412	0003	1			Total:	
B08412	0003	2			Less than 5 minutes	
B08412	0003	3			5 to 9 minutes	
B08412	0003	4			10 to 14 minutes	
B08412	0003	5			15 to 19 minutes	
B08412	0003	6			20 to 24 minutes	
B08412	0003	7			25 to 29 minutes	
B08412	0003	8			30 to 34 minutes	
B08412	0003	9			35 to 39 minutes	
B08412	0003	10			40 to 44 minutes	
B08412	0003	11			45 to 59 minutes	
B08412	0003	12			60 to 89 minutes	
B08412	0003	13			90 or more minutes	
B08412	0003	14			Male:	
B08412	0003	15			Less than 5 minutes	
B08412	0003	16			5 to 9 minutes	
B08412	0003	17			10 to 14 minutes	
B08412	0003	18			15 to 19 minutes	
B08412	0003	19			20 to 24 minutes	
B08412	0003	20			25 to 29 minutes	
B08412	0003	21			30 to 34 minutes	
B08412	0003	22			35 to 39 minutes	
B08412	0003	23			40 to 44 minutes	
B08412	0003	24			45 to 59 minutes	
B08412	0003	25			60 to 89 minutes	
B08412	0003	26			90 or more minutes	
B08412	0003	27			Female:	
B08412	0003	28			Less than 5 minutes	
B08412	0003	29			5 to 9 minutes	
B08412	0003	30			10 to 14 minutes	
B08412	0003	31			15 to 19 minutes	
B08412	0003	32			20 to 24 minutes	
B08412	0003	33			25 to 29 minutes	
B08412	0003	34			30 to 34 minutes	
B08412	0003	35			35 to 39 minutes	
B08412	0003	36			40 to 44 minutes	
B08412	0003	37			45 to 59 minutes	
B08412	0003	38			60 to 89 minutes	
B08412	0003	39			90 or more minutes	

Appendix A – Supplemental Documentation

Supplemental documentation concerning the American Community Survey, to assist users using this technical document, is located on the ACS Website on: http://www.census.gov/acs/www/data_documentation/documentation_main/. Documents such as the Subject Definitions, Accuracy of the Data documentation, and Code Lists are available on the URL listed above.

Appendix B – Summary File User Tools

American Community Survey 1-Year Summary File Sample SAS ™ Programs and Excel Templates

The American Community Survey Office has provided sample SAS ™ programs and Excel Templates for each Summary File data file. SAS datasets or Excel files can be generated using the ASCII ACS Summary Files as input to these programs or using the Excel templates. All of these materials are located in a sub-directory named **/UserTools**.

Sample SAS Programs

a. Main Program to make all the 16,000+ SAS programs to read in each individual sequence - **SF_All_Macro.sas**

b. All 16,000+ SAS Programs: **SF20091YR_SAS.zip**

SF20091YR_SAS.zip - Contains 16,000+ sas programs, one for each summary file, which can be used to convert each estimate and margin of error summary file into a SAS Dataset, which include table stubs.

SF_All_Macro.sas – This is a detailed example SAS program containing SAS macros which accesses the three different type of Summary Files for one table for all geographies from the ACS summary file. It is recommended that you review this document before using this SAS program.

Basic Instructions for Using the Sample SAS Programs

Light modifications need to be made to the SAS programs in the zip file before using them. The SAS programs included in the sascode.zip have variable and value labels. Using SAS, a data user can open any of the programs, edit the libname and infile statements of the program, and modify the input, output file names and the directories being used to store data to reference the summary files after they have been copied onto the user's computer. Data users may also want to delete columns from output SAS data set not being used to reduce the data set size. After these modifications are made, these programs should be able to run using SAS. Users who do not have SAS are free to read these ACSII summary files into any program of their choice. All values in the ASCII summary files are in CSV-format (comma-separated values).

Sample SAS Code showing how to combine the contents of the data files with the geographic file:

```
/* Add a libname statement referencing the directory being used */

/* Sort all four summary file files by LOGRECNO */

PROC SORT data=census.g20091ak;
      by LOGRECNO;
RUN;

PROC SORT data=census.SFe0001ak;
      by LOGRECNO;
RUN;

PROC SORT data=census.SFm0001ak;
      by LOGRECNO;
RUN;

/* Merge all four files by LOGRECNO */

DATA census.coall;
MERGE census.g20091ak census.SFe0001ak census.SFm0001ak;
      by LOGRECNO;
RUN;
```

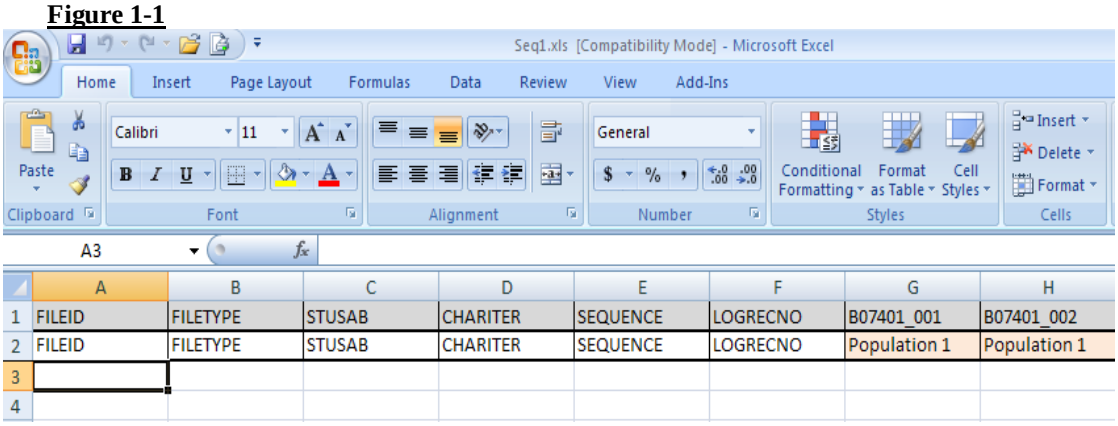
Sample Excel Templates

Excel Templates: **SummaryFileXLS.zip** - The Compressed file contains 176 individual XLS files, 1 per each Summary File Sequence. User can import the summary file (text file) into Excel to make the summary file more user friendly.

Importing Summary Files into Excel Template (Excel 2007)

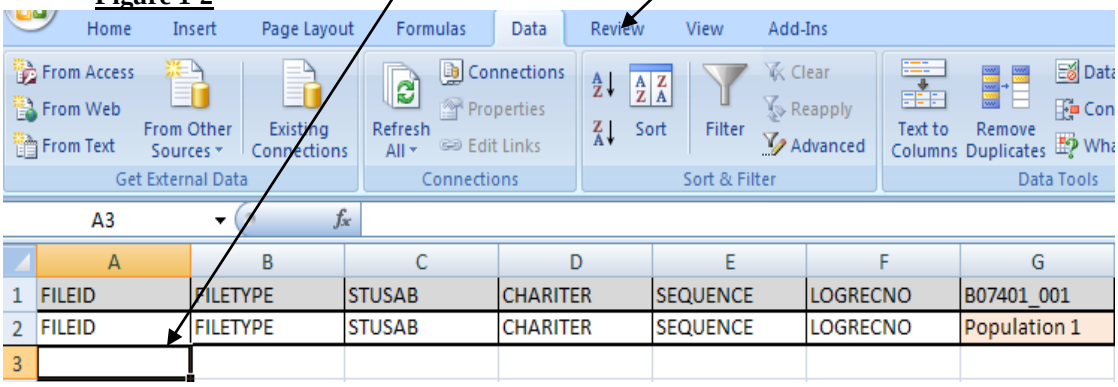
1) Open the appropriate Excel template for the Sequence ID you would like to import into Excel. The file will default to open on sheet "E". Both Sheet "E" and "M" have the same information; one sheet is for "Estimates" and one sheet for "Margin of Errors". When the file is open in Excel 2007 it should appear as Figure 1-1 below.

Figure 1-1



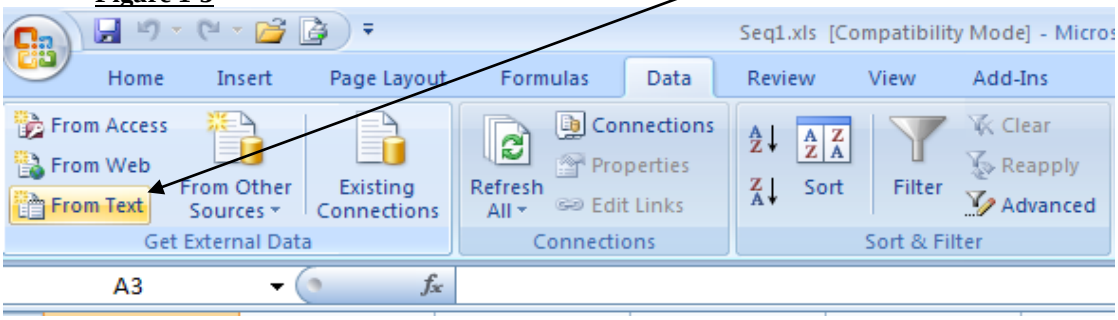
2) Place your cursor in cell "A3" and click on the "**Data**" tab in the Excel tool bar. As shown in Figure 1-2 below.

Figure 1-2



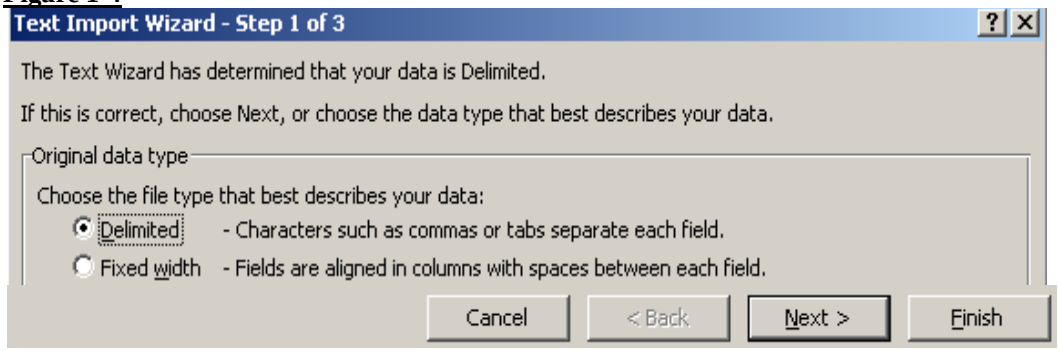
3) To import the Summary File text file into Excel click on "**From Text**" in the "Get External Data" section of the tool bar.

Figure 1-3



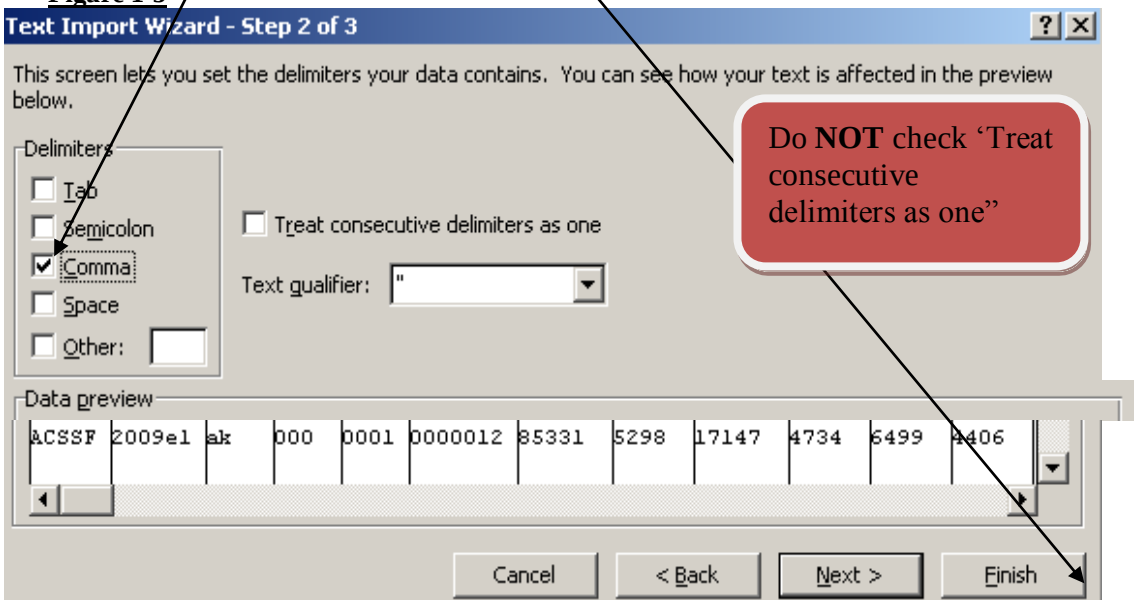
4) The Step 1 of the Excel "Text Import Wizard" will appear. Under "Original data type" choose "**Delimited**" , and then click "**Next**".

Figure 1-4



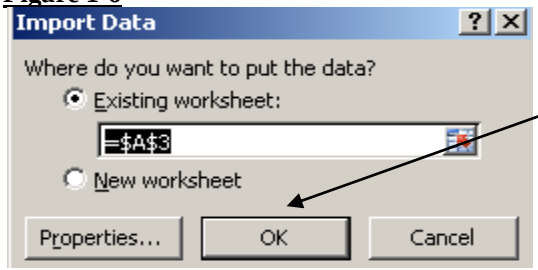
5) Step 2 of the Excel "Text Import Wizard" will appear. Under "Delimiters" choose "**Comma**", you may click "**Finish**" to import the file at this or choose "Next" to format the excel columns to your liking; Example purposes click "Next".

Figure 1-5



A Pop up window will appear for to confirm cell "A3" as the correct cell. Click "**OK**".

Figure 1-6



6) The summary file will be imported in Excel, Figure 1-7

Figure 1-7

	G	H	I	J	K	L
1	B07401_001	B07401_002	B07401_003	B07401_004	B07401_005	B07401_006
2	Population 1 year and over in the United States	Population 1 year and over in the United States%1 to 4 years	Population 1 year and over in the United States%5 to 17 years	Population 1 year and over in the United States%18 and 19 years	Population 1 year and over in the United States%20 to 24 years	Population 1 year and over in the United States%25 to 29 years
3	711753	46387	132281	26759	58997	60325

Row 1 - Contains a unique identifier of Table ID and Line Number with a "_" between them
Row 2 - Contains the associated metadata for each unique Identifier
Row 3 – Is the first Row of the imported data

Appendix C – User Notes

Jam Values

Some data values represent unique situations where either the information to be conveyed is an explanation for the absence of data, represented by a symbol in the data display, such as "(X)", or the information to be conveyed is an open-ended distribution, such as 115 or greater, represented by 115+.

The following special data values (and their meaning), can appear in the ACS Summary File table as an explanation for the absence of data:

Missing Value = ".": A dot ".", indicates that the estimate is unavailable. All ACS tables contain these values when estimates are missing due to certain special conditions.

Filtered Value or Geographic Restriction = " ": A blank indicates that a value is filtered or is not shown due to a geographic restriction.

Margin of Error (MOE) = 0: Estimate is controlled, which indicates the estimates are controlled and a statistical test is not appropriate.

Rounding Rules, Margin of Errors (MOE)

A. B00001, B00002, B98001, and B98002 are sample counts, not estimates, and do not have margin of error (MOE) associated with them. Tables in series B99* imputation tables and B98* quality measure tables (except B98001 and B98002) do not provide margin of error calculations. The margin of error calculations are set to -1 for these tables.

B. There are a few special rules on how certain margin of error are determined for ACS estimates. The accuracy of the estimate (decimal place) within the detailed tables determines how many digits the margin of error is rounded.

The only exceptions to these rules are tables B19082 (Shares of aggregate household income by quintile) and B19083 (GINI index of income inequality). Estimates are rounded to 2 decimal places for all estimates and the margin of error to 3 decimal places.

Display of Estimates

Display of whole numbers in the summary files: The estimates in the summary files are stored using standard notation instead of in scientific notation. The largest estimate in a Summary File contains 14 digits and stored as whole numbers.

The following list shows median and ratio tables, their table types, and accuracy.

Table ID	Table Title	Table Type	Accuracy
B01002	MEDIAN AGE BY SEX	MEDIAN	Tenths
B01002A	MEDIAN AGE BY SEX (WHITE ALONE)	MEDIAN	Tenths
B01002B	MEDIAN AGE BY SEX (BLACK OR AFRICAN AMERICAN ALONE)	MEDIAN	Tenths
B01002C	MEDIAN AGE BY SEX (AMERICAN INDIAN AND ALASKA NATIVE)	MEDIAN	Tenths
B01002D	MEDIAN AGE BY SEX (ASIAN ALONE)	MEDIAN	Tenths
B01002E	MEDIAN AGE BY SEX (NATIVE HAWAIIAN AND OTHER PACIFIC ISLANDER ALONE)	MEDIAN	Tenths
B01002F	MEDIAN AGE BY SEX (SOME OTHER RACE ALONE)	MEDIAN	Tenths
B01002G	MEDIAN AGE BY SEX (TWO OR MORE RACES)	MEDIAN	Tenths
B01002H	MEDIAN AGE BY SEX (WHITE ALONE, NOT HISPANIC OR LATINO)	MEDIAN	Tenths
B01002I	MEDIAN AGE BY SEX (HISPANIC OR LATINO)	MEDIAN	Tenths
B05004	MEDIAN AGE BY CITIZENSHIP STATUS BY SEX	MEDIAN	Tenths
B06002	MEDIAN AGE BY PLACE OF BIRTH IN THE UNITED STATES	MEDIAN	Tenths
B06002PR	MEDIAN AGE BY PLACE OF BIRTH IN PUERTO RICO	MEDIAN	Tenths
B07002	MEDIAN AGE BY RESIDENCE 1 YEAR AGO IN THE UNITED STATES	MEDIAN	Tenths
B07002PR	MEDIAN AGE BY RESIDENCE 1 YEAR AGO IN PUERTO RICO	MEDIAN	Tenths
B08103	MEDIAN AGE BY MEANS OF TRANSPORTATION TO WORK	MEDIAN	Tenths
B08503	MEDIAN AGE BY MEANS OF TRANSPORTATION TO WORK FOR WORKPLACE GEOGRAPHY	MEDIAN	Tenths
B12008	MEDIAN AGE AT FIRST MARRIAGE	MEDIAN	Tenths
B12008A	MEDIAN AGE AT FIRST MARRIAGE (WHITE ALONE)	MEDIAN	Tenths
B12008B	MEDIAN AGE AT FIRST MARRIAGE (BLACK OR AFRICAN AMERICAN ALONE)	MEDIAN	Tenths
B12008C	MEDIAN AGE AT FIRST MARRIAGE (AMERICAN INDIAN AND ALASKA NATIVE ALONE)	MEDIAN	Tenths
B12008D	MEDIAN AGE AT FIRST MARRIAGE (ASIAN ALONE)	MEDIAN	Tenths
B12008E	MEDIAN AGE AT FIRST MARRIAGE (NATIVE HAWAIIAN AND OTHER PACIFIC ISLANDER ALONE)	MEDIAN	Tenths
B12008F	MEDIAN AGE AT FIRST MARRIAGE (SOME OTHER RACE ALONE)	MEDIAN	Tenths
B12008G	MEDIAN AGE AT FIRST MARRIAGE (TWO OR MORE RACES)	MEDIAN	Tenths
B12008H	MEDIAN AGE AT FIRST MARRIAGE (WHITE ALONE, NOT HISPANIC OR LATINO)	MEDIAN	Tenths
B12008I	MEDIAN AGE AT FIRST MARRIAGE (HISPANIC OR LATINO)	MEDIAN	Tenths
B19082	SHARES OF AGGREGATE HOUSEHOLD INCOME BY QUINTILE	GINI	Tenths
B19083	GINI INDEX OF INCOME INEQUALITY	GINI	Thousandths
B23013	MEDIAN AGE BY SEX FOR WORKERS 16 TO 64 YEARS	MEDIAN	Tenths
B23020	MEAN USUAL HOURS WORKED IN THE PAST 12 MONTHS FOR WORKERS 16 TO 64 YEARS	RATIO	Tenths

Table ID	Table Title	Table Type	Accuracy
B23021	MEAN WEEKS WORKED IN THE PAST 12 MONTHS FOR WORKERS 16 TO 64 YEARS	RATIO	Tenths
B25010	AVERAGE HOUSEHOLD SIZE OF OCCUPIED HOUSING UNITS BY TENURE	RATIO	Hundredths
B25018	MEDIAN NUMBER OF ROOMS	MEDIAN	Tenths
B25021	MEDIAN NUMBER OF ROOMS BY TENURE	MEDIAN	Tenths
B25071	MEDIAN GROSS RENT AS A PERCENTAGE OF HOUSEHOLD INCOME IN THE PAST 12 MONTHS (DOLLARS)	MEDIAN	Tenths
B25092	MEDIAN SELECTED MONTHLY OWNER COSTS AS A PERCENTAGE OF HOUSEHOLD INCOME IN THE PAST 12 MONTHS	MEDIAN	Tenths

Data Release Filtering Rules

Data release filtering rules. Filtering rules, based on statistical reliability of the survey estimates, are used because certain geographic areas contain detailed tables include estimates whose level of reliability is unacceptable. The data release rules for the American Community Survey data tables include the following.

Every base table consists of a series of estimates. If more than half the estimates are not statistically different from 0 (at a 90 percent confidence level), then the table fails. Each estimate is subject to sampling variability that is summarized by its standard error. Dividing the standard error by the estimate yields the coefficient of variation (CV) for each of the estimates. (If the estimate is 0, a CV of 100 percent is assigned.) To implement this requirement for each table at a given geographic area, CVs are calculated for each of the table's estimates, and the median CV value is determined. If the median CV value for the table is less than or equal to 61 percent, the table passes for that geographic area; if it is greater than 61 percent, the table fails. Tables that are too sparse will fail this test. In that case, the table will not be published for that geographic area. Whenever a table fails, a simpler table that collapses some of the detailed lines together can be substituted for the original, more detailed table. The rules are then applied to the simpler table. If it passes, the simpler table is released. If it fails, none of the estimates for that particular table is released for this geographic area. These rules are applied to single-year period estimates and multi-year period estimates based on three years of sample data.

Appendix D – Worked Examples

ACS Summary File Worked Example

Let's say that you want to create Table B08406, "Sex of workers by means of transportation to work for workplace geography," for the state of Alaska from the files on the ftp site http://www2.census.gov/acs2009_1yr/summaryfile/Seq_By_ST. Which files do you need? How do you read the files?

You will need three files:
1. The data dictionary (merge_5_6.sas7bdat)
2. The zip file/data file /acs2009_1yr/summaryfile/Seq_By_ST/*<ST ABBRV>/<Sequence ID>.zip* (20091ak0003000.zip - e2009ak0003000.txt)
3. The geography file (g20091ak.txt)

Start with the data dictionary, merge_5_6.sas7bdat. Under the "Tblid" column, look for the value "B08406". You will see that the "Sequence Number" is "0003." This means that the data you looking for are in the data file "e2008ak0003000.txt". How do you know this is the right file? You know this from the name of the file: the "e" stands for estimate, 2009 is the year, "ak" is the state (Alaska), and "0001" is the sequence number (which contains the data for Table B08406). See the "File Naming Conventions" section in Chapter 2.

Then use the geography file for Alaska to determine the location within the state to which the data refer. The appropriate file is g20091ak.txt, where "g" means "geography", 2009 is the year, 1 is the period estimate (in this case, 1-year estimate), and "ak" is the state. (For each state, the geography file contains the lower-case FIPS State Code.)

When you open the data file, e20091ak0003000.txt, you will see the following comma-delimited fields on the first line:

ACSSF,2009e1,ak,000,0003,0000001,342551,275208,230130,45078,35960,,…

The first six fields – from "ACSSF" to "0000001" – are identifiers.
A. The first field tells you that this is an ACS Summary File;
B. The second tells you that these data are one-year estimates for the year 2009 (notice the "e" before "2009" and the "1" at the end);
C. The third tells you the state ("ak" is Alaska);
D. The fourth is an iteration number;
E. The fifth is the sequence number,
F. The last is a logical record code (LOGRECNO). The LOGRECNO identifies the geographic area within a state.

The geography file, g20091ak.txt, defines the LOGRECNO. Each LOGRECNO specifies a geographic area pertaining to the state. For example, a LOGRECNO of

"0000001" means the state of Alaska; a LOGRECNO of "0000002" means just the urban areas in Alaska; a LOGRECNO of "0000003" refers to just rural areas in Alaska. Notice that each
state has its own geography file. For more information, see Chapter 2, Table 2.1.

The other fields in the data file, from the seventh on, are data values. Each field corresponds to the value of the "line number" variable in the data dictionary. So field number seven (the 342551 value, after the sixth comma) corresponds to line number one, which is "Total". Field number eight (the 275208 value, after the seventh comma) refers to line number two, which is "Car, Truck, or Van." Field number nine (the 230130 value) corresponds to line number three, which is "Drove alone." This continues all the way up to line number 51, at which point Table B08406 ends.

Were you to read this into a computer program using software such as SAS, you could translate the first nine fields of e20091ak0003000.txt as follows:

TABLE B08406:
SEX OF WORKERS BY MEANS OF TRANSPORTATION TO WORK FOR WORKPLACE GEOGRAPHY

FILEID	FILE TYPE	STUSAB	CHARITER	SEQUENCE	LOGRECNO	Total	Car, Truck, or Van	Drove Alone
ACSSF	2009e1	ak	000	0003	0000001	342551	275208	230130
ACSSF	2009e1	ak	000	0003	0000010	157650	141673	120945
ACSSF	2009e1	ak	000	0003	0000011	.	.	.
ACSSF	2009e1	ak	000	0003	0000012	.	.	.
ACSSF	2009e1	ak	000	0003	0000013	157650	141673	120945
ACSSF	2009e1	ak	000	0003	0000017	157650	141673	120945

See formatted example of Table B08406 below:

Table ID	Line Number	Sequence Number	Table Title	Estimates	Margin of Error
B08406		003	**SEX OF WORKERS BY MEANS OF TRANSPORTATION TO WORK FOR WORKPLACE GEOGRAPHY**		
B08406		003	Universe: Workers 16 years and over		
B08406	1	003	Total:	342551	+/-6,394
B08406	2	003	Car, truck, or van:	275208	+/-6,304
B08406	3	003	Drove alone	230130	+/-6,492
B08406	4	003	Carpooled:	45078	+/-3,523
B08406	5	003	In 2-person carpool	35960	+/-3,399
B08406	6	003	In 3-person carpool	5656	+/-1,243
B08406	7	003	In 4-or-more-person carpool	3462	+/-976

Advanced ACS Summary File Worked Example Using SAS

Here is an example of how to access the Summary Files for one table for all geographies from the ACS summary file.

Question: I am interested in downloading table B01001 for all published ACS geographies, how would I do this?

1. Go to Chapter 5 to locate sequence number for table B01001.

There are files that summarizes this information at:
http://www2.census.gov/acs2009_1yr/summaryfile/merge_5_6.xls
http://www2.census.gov/acs2009_1yr/summaryfile/merge_5_6.sas7bdat

2. Use the Summary File SAS Example Macros located at
http://www2.census.gov/acs2009_1yr/summaryfile/UserTools/summary_file_example_macros.sas

Run the macro for `%TableShell(B01001);`

This macro will provide metadata information on a given table, in this case B01001

3. The following SAS dataset will be created with information about table B01001

Table ID	Sequence Number	Line Number	Start Position	Total Cells in Table	Total Cells in Sequence	Table Title
B01001	0013		7	49 CELLS		SEX BY AGE
B01001	0013					Universe: Total population
B01001	0013	1				Total:
B01001	0013	2				Male:
B01001	0013	3				Under 5 years
B01001	0013	47				75 to 79 years
B01001	0013	48				80 to 84 years
B01001	0013	49				85 years and over

We can see that table B01001 is located in Sequence 0013; this applies to all published geographies.

4. We can read into SAS all tables in the 0013 sequence by running the SAS Example Macros.

`%CallSt;`
This macro will run a do loop creating State two-digit abbreviations, which will allow a simple way to read the Summary Files into SAS for all geographies. Each time a valid two digit state abbreviation is created, the macro ***%AllSeqs*** is ran with the two digit state abbreviation.

%AllSeqs Macro performs the following tasks

a) Read the Geographic header file - `%AnyGeo Macro.`

b) There is a do loop to allow you to choose which sequences you would like to read in for example if you wanted sequence 0010 set the loop to be `do x=13 %to 13;` - The 0 values will be filled in.

c) Within the do loop the following marcros will be executed:

I. `%TablesBySeq;` - This will give information about the whole 10 sequence, not just table B01001.

II. `%ReadDataFile` – Macro is called 2 times once for each type of estimate. This macro will generate and run SAS code for each sequence specified in the do loop in step 2 and for each geography specified in the `%CallSt` macro.

III. Lastly there is a merge statement that will merge together each of the three types of estimates and the geography header file by sequence number per geography.

5. You now will have all tables in the 0013 sequence read into SAS in the following dataset names (if the code is not modified), in the work directory Sf0013<st two-digit abbrev>.sas7bdat