Digging Deeper:
A Quick Tour Through Census FTP Downloads

Steve Barker
Program Manager
State Data Center

OKLAHOMA COMMERCE
900 N. STILES OKLAHOMA CITY | OKcommerce.gov

Building the Connection

"I have always imagined that Paradise will be a kind of library."

– Jorge Luis Borges

"In your thirst for knowledge, be sure not to drown in all the information."

– Anthony J. D'Angelo

"There are three kinds of lies…"

– Benjamin Disraeli

Today's Agenda

A Brief Overview of Census Data Sources

File Transfer Protocol – the FTP Site

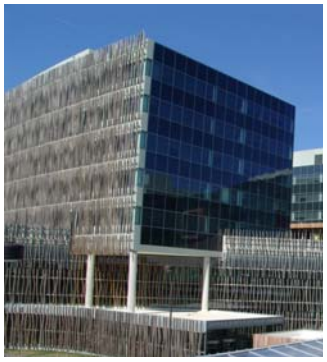
Finding Your Way

Example - American Community Survey

Example - 2010 Redistricting Data

Questions (time permitting)

A Brief Overview of Census Data Sources



Census Headquarters

You

Some Leading Census Data Sources

- Hard copy, CD/DVD
  - Census Bureau has made a conscious decision to slowly move away from these environments
- Census Bureau on the web
  - Main website www.census.gov
  - American FactFinder www.factfinder2.census.gov
  - Data Ferrett www.dataferrett.census.gov
  - FTP www2.census.gov

Assumptions About FTP

- Users have pre-existing expert knowledge of Census products
  - If not, you can be here but work with caution
- Users are comfortable working with large sets of data
  - Many sets have several thousand rows of data
- Users have the necessary tools
  - SAS or similar statistical software
  - Microsoft Office Suite 2007 or equivalent

FTP Limitations

- Text based instructions may be limited and technical
- There are very few tutorials
- Not everything is clearly labeled
  - lots of holdover files from era when names were limited to 8 characters
- Hard to walk you through the process in an 'over the phone' setting

So Why Use FTP?

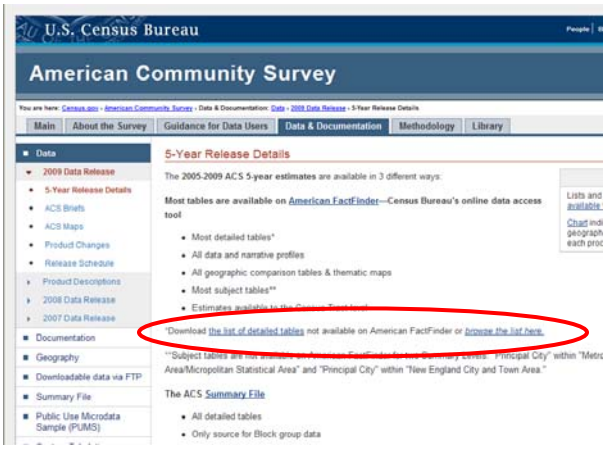
- Need an electronic file showing population by Census Block in
  - Tarrant County, TX
  - Orange County, CA
  - Oklahoma County, OK
  - "Your selection is too large to view this product, please remove some items from 'My Selections' and try again."
- Need electronic file with population characteristics from 1980 or 1990 Census
  - Can get *.pdf via archives, but exportable results not available via AFF2

Want Another Reason?

Not every ACS table is available on American FactFinder, but they are available in FTP

- 7 examples out of 295 such tables
  - Means of Transportation to Work by Age
  - Sex by Age by Citizenship Status (Hispanic or Latino)
  - Place of Birth By Marital Status in the United States
  - Grandchildren Under 18 Years Living With A Grandparent Householder By Age Of Child
  - Median Age at First Marriage by Race
  - Women 15 To 50 Years Who Had A Birth In The Past 12 Months By Marital Status And Poverty Status In The Past 12 Months
  - Mortgage Status by Age of Householder
  - Full list available at http://www.census.gov/acs/www/data_documentation/2009_5yr_data/

How to Find The List



- Using link on previous slide, you get this page
- Click 'list of detailed tables' to see what's not on ACS

The FTP Site

http://www2.census.gov/

What's Out There?

http://www2.census.gov/		Main Folder Options	
AHS – Requires SAS	acs2008_1yr	econ2003	outgoing
Econ2001_And Earlier	acs2008_3yr	econ2004	plmap
acs	acs2009_1yr	econ2005	pop
acs2002	acs2009_3yr	econ2006	prod2
acs2003	acs2009_5yr	econ2007	pub
acs2004	acs_special_tabs	econ2008	retail
acs2005	census_1980	econ2009	services
acs2005_2007_3yr	census_1990	econ2010	tms_data
acs2005-2009_5yr	census_2000	geo	video
acs2006	census_2010	govs	w3c
acs2006_2008_3yr	ces	hhes	wholesale
acsS2007_1yr	dadso	lehd	USEFUL
acs2007_3yr	econ	lost+found	NOT SO MUCH
acs2007_2009_3yr	econ2002	manufacturing	

Now, Learn Your Way Around That In 45 Minutes or Less

Finding Your Way

An Example Using American Community Survey

AMERICAN COMMUNITY SURVEY
U.S. CENSUS BUREAU

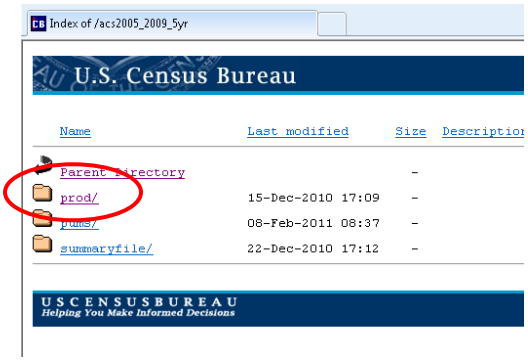
Start With The Documentation

- Look for key words in file names
- Here are some hints (not exhaustive list):
  - Methodology
  - README
  - Note
  - Datadict (data dictionary)
  - Help_guide
  - TableShells
  - TechDoc

Pick a Data Set From FTP Site

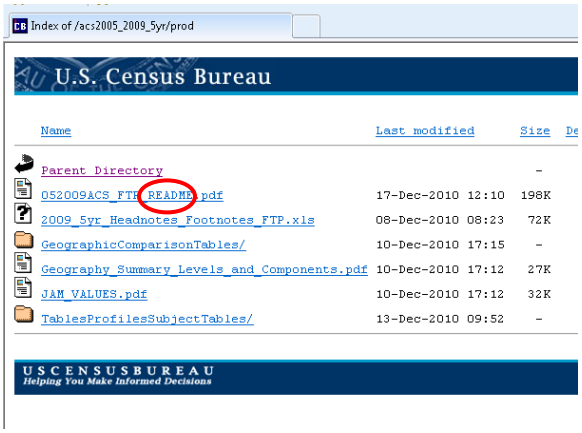
- Let's look at the American Community Survey 2005-2009 5 year dataset

Let the Digging Commence



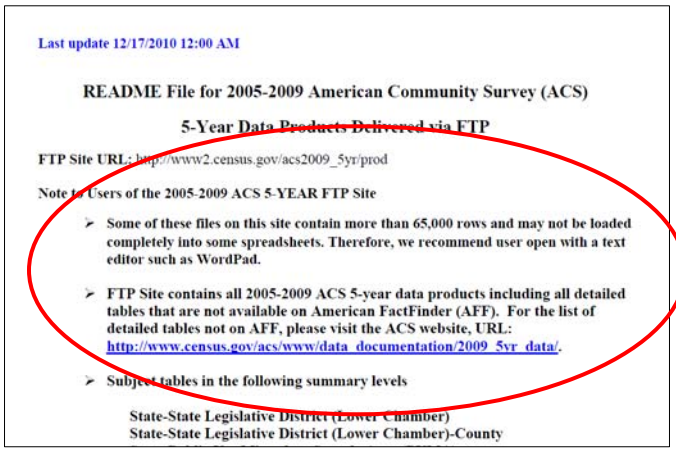
- Click on prod/

Our First Nuggets



- You see several options for guidance
- We start here with the README document

What The File Says



Last update 12/17/2010 12:00 AM

README File for 2005-2009 American Community Survey (ACS)

5-Year Data Products Delivered via FTP

FTP Site URL: http://www2.census.gov/acs2009_5yr/prod

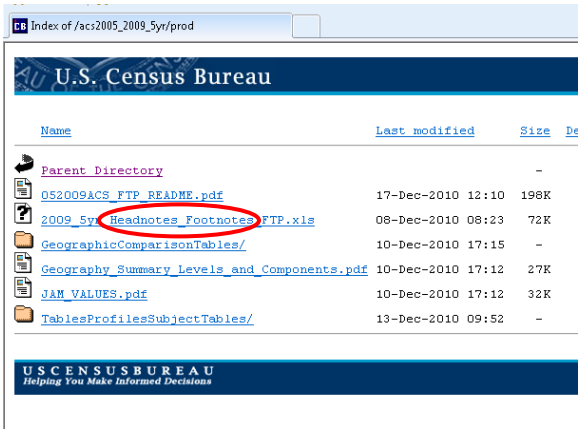
Note to Users of the 2005-2009 ACS 5-YEAR FTP Site

- Some of these files on this site contain more than 65,000 rows and may not be loaded completely into some spreadsheets. Therefore, we recommend user open with a text editor such as WordPad.
- FTP Site contains all 2005-2009 ACS 5-year data products including all detailed tables that are not available on American FactFinder (AFF). For the list of detailed tables not on AFF, please visit the ACS website, URL: http://www.census.gov/acs/www/data_documentation/2009_5yr_data/.
- Subject tables in the following summary levels

State-State Legislative District (Lower Chamber)
State-State Legislative District (Lower Chamber)-County

- You find notes that help you understand what's included

Another Nugget



Index of /acs2005_2009_5yr/prod

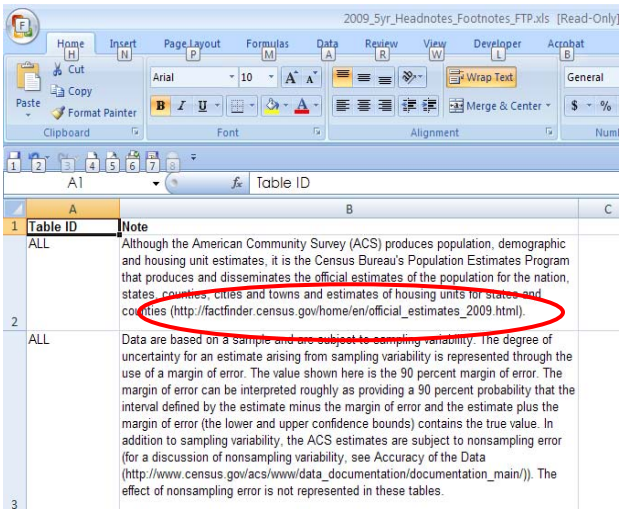
U.S. Census Bureau

Name	Last modified	Size	De
Parent Directory		-	
052009ACS_FTP_README.pdf	17-Dec-2010 12:10	198K	
2009_5yr_Headnotes_Footnotes_FTP.xls	08-Dec-2010 08:23	72K	
GeographicComparisonTables/	10-Dec-2010 17:15	-	
Geography_Summary_Levels_and_Components.pdf	10-Dec-2010 17:12	27K	
J&M_VALUES.pdf	10-Dec-2010 17:12	32K	
TablesProfilesSubjectTables/	13-Dec-2010 09:52	-	

USCENSUSBUREAU
Helping You Make Informed Decisions

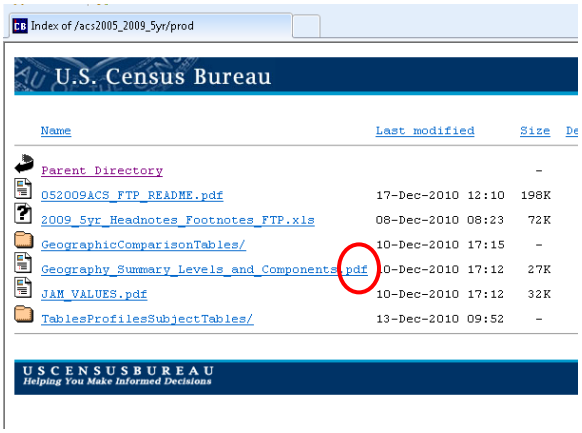
- Now we open the file marked with "notes" in the title

Tips For More Information



- Here you find web addresses for further explanation

And Another Nugget



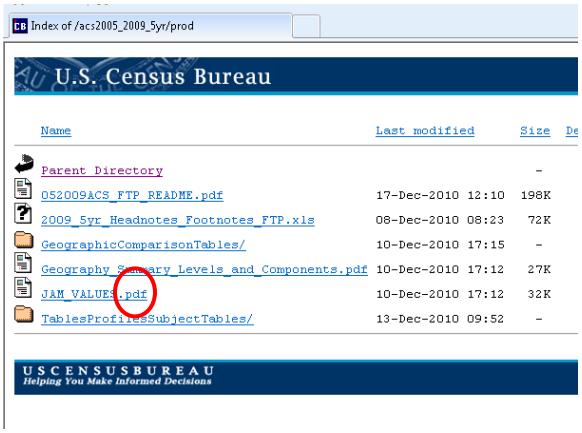
- Sometimes you need to key off of the *.pdf extension

A Key for Translating Data Tables

ACS 5-Year Summary Levels/Components for Detailed Tables

Summary Level	Component	Name
010	00	United States
010	01	United States; Urban
010	43	United States; Rural
010	89	United States; American Indian Reservation and Trust Land-Federal
010	90	United States; American Indian Reservation and Trust Land-State
010	91	United States; Oklahoma Tribal Statistical Area
010	92	United States; Tribal Designated Statistical Area
010	93	United States; Alaska Native Village Statistical Area
010	94	United States; State Designated Tribal Statistical Area (Formerly SDAISA)
010	95	United States; Hawaiian Home Land
010	A0	United States; In Metropolitan or Micropolitan Statistical Area
010	C0	United States; In Metropolitan Statistical Area
010	C1	United States; In Metropolitan Statistical Area-In Principal City
010	C2	United States; In Metropolitan Statistical Area-Not In Principal City
010	E0	United States; In Micropolitan Statistical Area
010	E1	United States; In Micropolitan Statistical Area-In Principal City
010	E2	United States; In Micropolitan Statistical Area-Not In Principal City
010	G0	United States; Not In Metropolitan or Micropolitan Statistical Area
010	H0	United States; Not In Metropolitan Statistical Area
020	00	Region

And Still More Nuggets To Review



- Again, look for clues

A Key to Translate Values

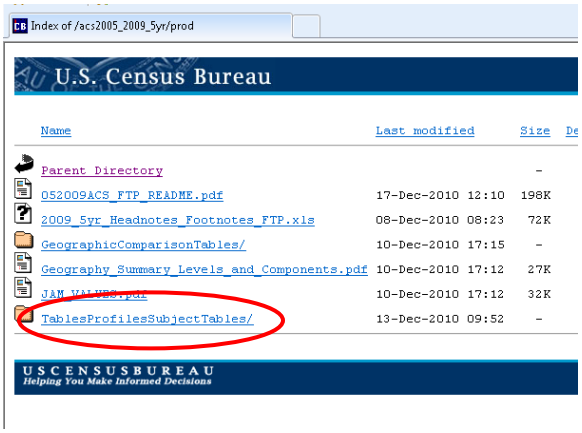
ACS 5-Year Jam Values

Some data values represent unique situations where either: a) 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 b) the information to be conveyed is an open-ended distribution, such as 115 or greater, represented by 116 or 115+.

Jam Value	Display Value	Description
-999999999	N	Indicates that an estimate or its margin of error cannot be provided because the number of sample cases is too small for the given geographic area.
-888888888	(X)	Indicates that the estimate is not applicable or not available.
-777777777	(Z)	Estimate is not available for an undefined reason.
-666666666	-	Indicates that no sample observations were available to compute an estimate, or a ratio of medians cannot be calculated because one or both of the median estimates falls in the lowest interval or upper interval of an open-ended distribution.

- Here you find value notes to explain certain table values

Finally, Time To Dig Deeper



- Now let's drill deeper

Look! More To Review

- And we have another *.pdf to examine

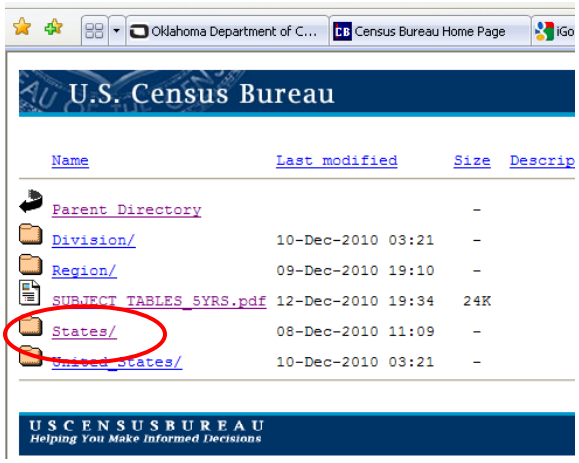
Another Key For Translating Tables

ACS 5-Year Subject Tables

Table ID	Table Title	Restricted Geographies
S0101	Age And Sex	
S0102	Population 60 Years And Over In The United States	US Geographies
S0102PR	Population 60 Years And Over In Puerto Rico	PR Geographies
S0103	Population 65 Years And Over In The United States	US Geographies
S0103PR	Population 65 Years And Over In Puerto Rico	PR Geographies
S0501	Selected Characteristics Of The Native And Foreign-Born Populations	
S0502	Selected Characteristics Of The Foreign-Born Population By Period Of Entry Into The United States	US Geographies
S0502PR	Selected Characteristics Of The Foreign-Born Population By Period Of Entry Into Puerto Rico	PR Geographies
S0503	Selected Characteristics Of The Foreign-Born Population By Region Of Birth: Europe	
S0504	Selected Characteristics Of The Foreign-Born Population By Region Of Birth: Africa, Northern America, And Oceania	
S0505	Selected Characteristics Of The Foreign-Born Population By Region Of Birth: Asia	
S0506	Selected Characteristics Of The Foreign-Born Population By Region Of Birth: Latin America	
S0601	Selected Characteristics Of The Total And Native Populations In The United States	US Geographies
S0601PR	Selected Characteristics Of The Total And Native Populations In Puerto Rico	PR Geographies
S0701	Geographic Mobility By Selected Characteristics In The United States	US Geographies
S0701PR	Geographic Mobility By Selected Characteristics In Puerto Rico	PR Geographies

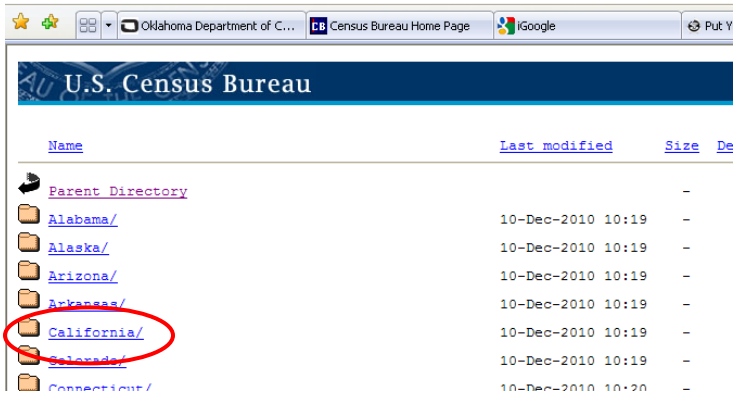
- Translating table ID numbers into specific table names

Digging Again



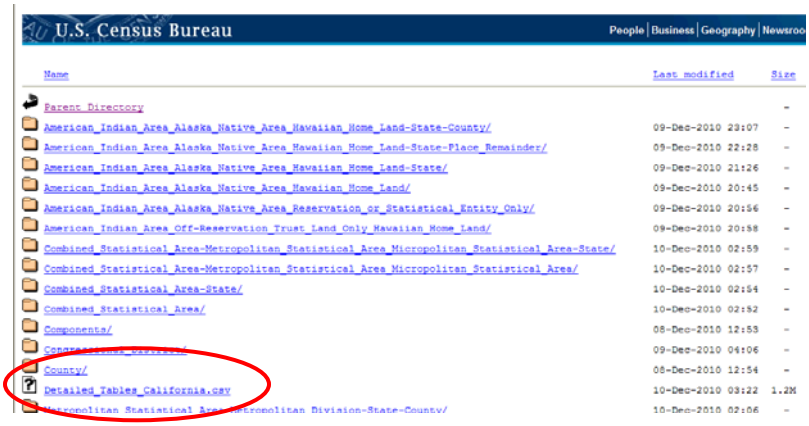
- Now we dig still deeper

Find Your State



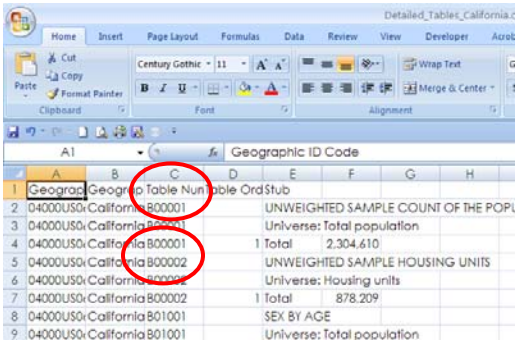
- And we dig deeper yet again

At Last – DATA!



- And as we see lower geography options, we can get statewide data or dig still deeper

Sixteen Thousand Rows of Data

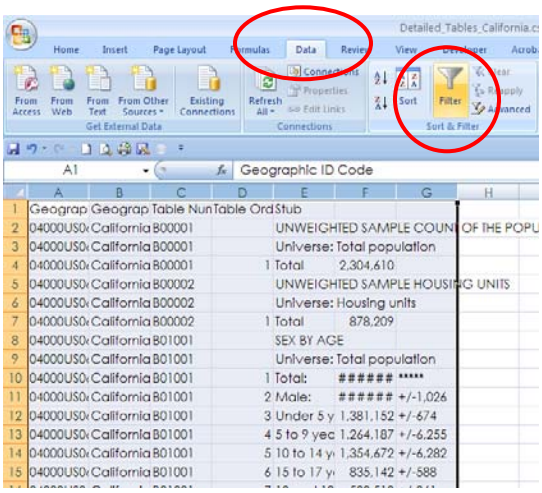


- Notice splits based on table numbers

- You can convert table numbers to table names by going to list at

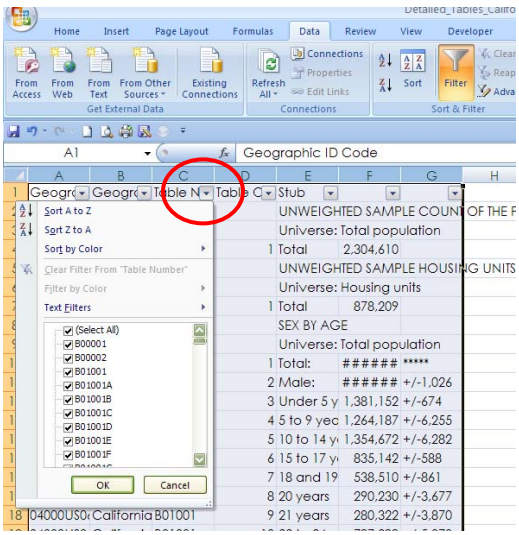
http://www.census.gov/acs/www/Downloads/data_documentation/2009_release/Tables2005_2009_5Year.xls

Filter to the Table You Want



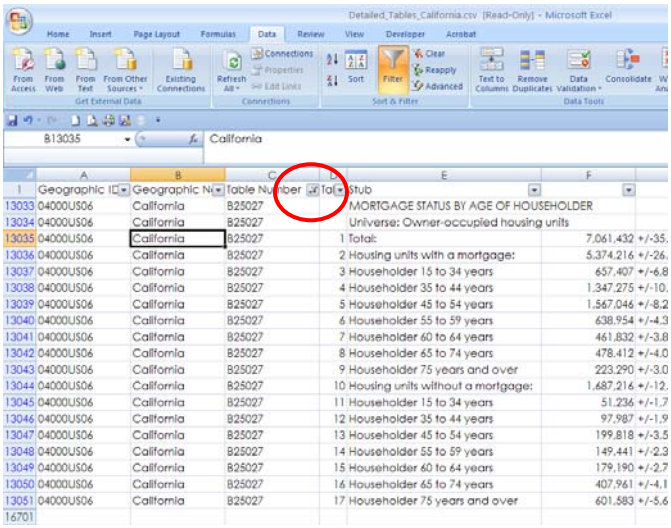
- Highlight filled columns on the page
- Click 'Data' at the top of the page
- Click 'Filter'

Filter Part Two



- Click on the down arrow at the top of the 'Table Num' column
- Use the filter options

How To See Just The Table You Want



	Geographic ID	Geographic Name	Table Number	Total	Stub	
13033	04000US06	California	B25027		MORTGAGE STATUS BY AGE OF HOUSEHOLDER	
13034	04000US06	California	B25027		Universe: Owner-occupied housing units	
13035	04000US06	California	B25027	1	Total:	7,061,432 +/-35,
13036	04000US06	California	B25027	2	Housing units with a mortgage:	5,374,216 +/-26,
13037	04000US06	California	B25027	3	Householder 15 to 34 years	657,407 +/-6,8
13038	04000US06	California	B25027	4	Householder 35 to 44 years	1,347,275 +/-10,
13039	04000US06	California	B25027	5	Householder 45 to 54 years	1,567,046 +/-8,2
13040	04000US06	California	B25027	6	Householder 55 to 59 years	638,954 +/-4,3
13041	04000US06	California	B25027	7	Householder 60 to 64 years	461,832 +/-3,8
13042	04000US06	California	B25027	8	Householder 65 to 74 years	478,412 +/-4,0
13043	04000US06	California	B25027	9	Householder 75 years and over	223,290 +/-3,0
13044	04000US06	California	B25027	10	Housing units without a mortgage:	1,687,216 +/-12,
13045	04000US06	California	B25027	11	Householder 15 to 34 years	51,236 +/-1,7
13046	04000US06	California	B25027	12	Householder 35 to 44 years	97,987 +/-1,9
13047	04000US06	California	B25027	13	Householder 45 to 54 years	199,818 +/-3,5
13048	04000US06	California	B25027	14	Householder 55 to 59 years	149,441 +/-2,3
13049	04000US06	California	B25027	15	Householder 60 to 64 years	179,190 +/-2,7
13050	04000US06	California	B25027	16	Householder 65 to 74 years	407,961 +/-4,1
13051	04000US06	California	B25027	17	Householder 75 years and over	601,583 +/-5,6
16701						

- This is one of the tables not in ACS
- Note the ‘Filter’ symbol at top of column

Finding Your Way

An Example Using 2010 Census Redistricting Data

United States Census 2010

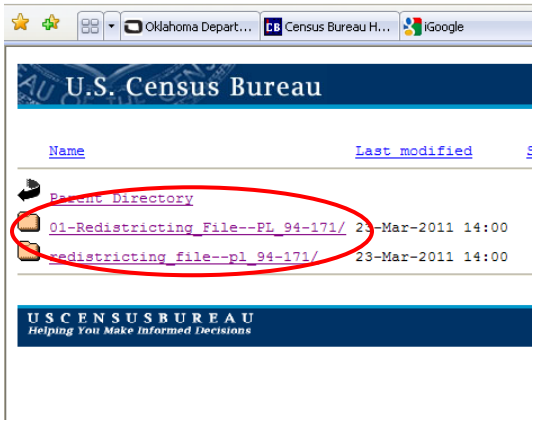
Notes About Redistricting FTP

- Census assumes you either have SAS or Access in order to view this data
- Its up to you to figure out how to view this data if you don't have those specific tools
- **Plan ahead!**
  - It took the better part of a morning to download necessary files, import into Access, and begin working with the data
  - That was using a high speed internet connection

Pick a Data Set From FTP Site

- Let's look at the Census 2010 dataset

Pick a Data Set From FTP Site

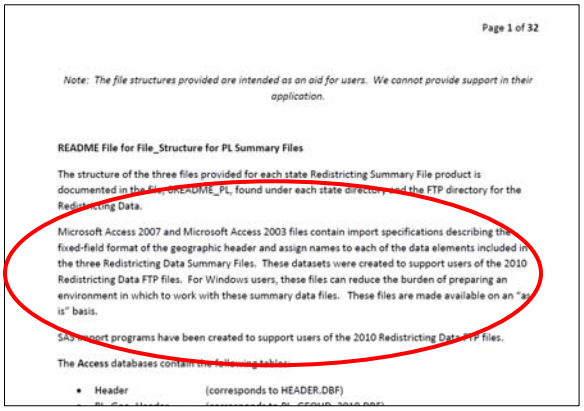


- Choose redistricting data set (top option)
- Not sure why they have two folders here

Finding the Support Documents

- Notice clues for helping navigate the tables
  - File_Structure
  - README
  - Matching

From the FILE_STRUCTURE file



Page 1 of 32

Note: The file structures provided are intended as an aid for users. We cannot provide support in their application.

README File for File_Structure for PL Summary Files

The structure of the three files provided for each state Redistricting Summary File product is documented in the file, README_PL, found under each state directory and the FTP directory for the Redistricting Data.

Microsoft Access 2007 and Microsoft Access 2003 files contain import specifications describing the fixed-field format of the geographic header and assign names to each of the data elements included in the three Redistricting Data Summary Files. These datasets were created to support users of the 2010 Redistricting Data FTP files. For Windows users, these files can reduce the burden of preparing an environment in which to work with these summary data files. These files are made available on an "as is" basis.

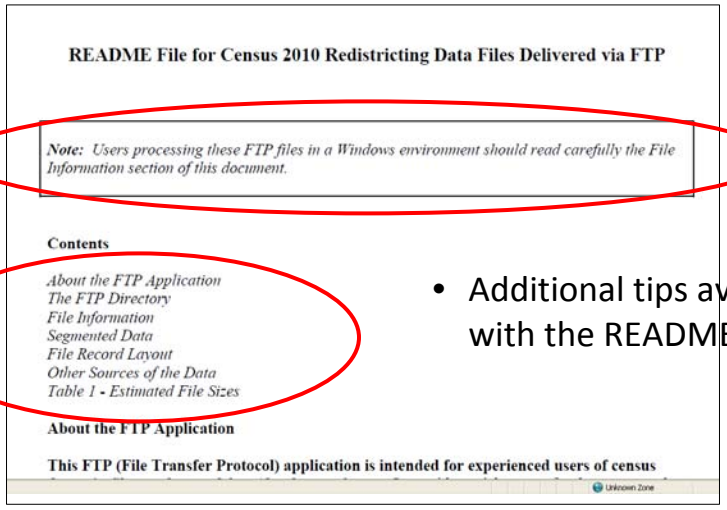
SAS import programs have been created to support users of the 2010 Redistricting Data FTP files.

The Access databases contain the following tables:

- Header (corresponds to HEADER.DBF)

- The next slides are based on process for opening data in Access
- You can read more detail in the FILE_STRUCTURE document

From the README file



README File for Census 2010 Redistricting Data Files Delivered via FTP

Note: *Users processing these FTP files in a Windows environment should read carefully the File Information section of this document.*

Contents

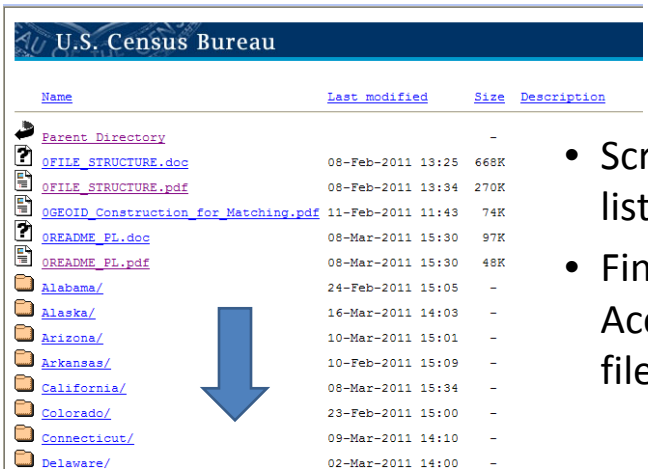
About the FTP Application
The FTP Directory
File Information
Segmented Data
File Record Layout
Other Sources of the Data
Table 1 - Estimated File Sizes

About the FTP Application

This FTP (File Transfer Protocol) application is intended for experienced users of census

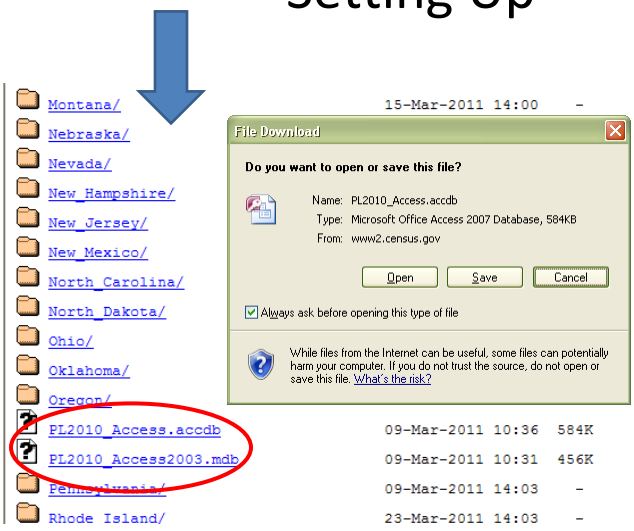
- Additional tips available with the README file

Finding the Access Tools



- Scroll down the list of states
- Find the Microsoft Access application files

Setting Up

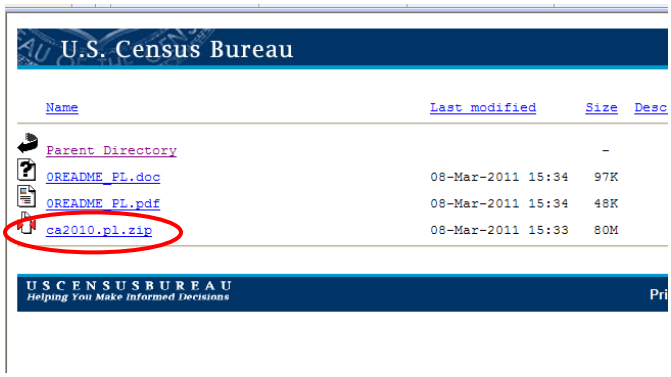


- Click on each Access database file and save to your computer

Searching for Data

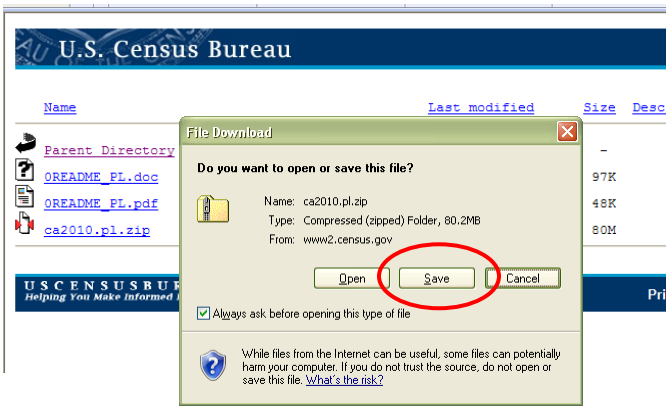
- Pick your state

Digging Deeper



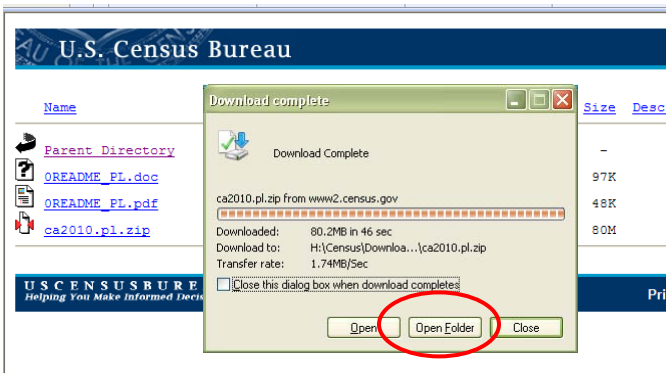
- Data is contained in the *.zip file

Opening the Data



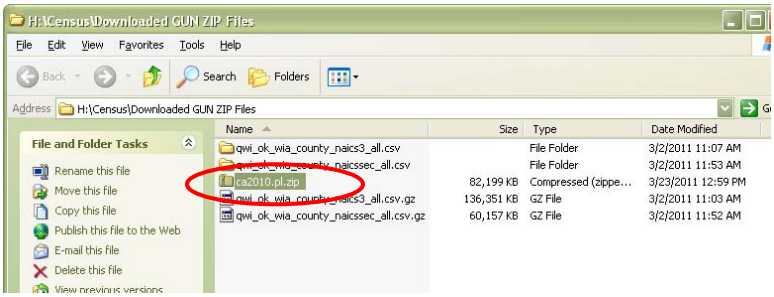
- Save the file so you can work with it

Opening the Data



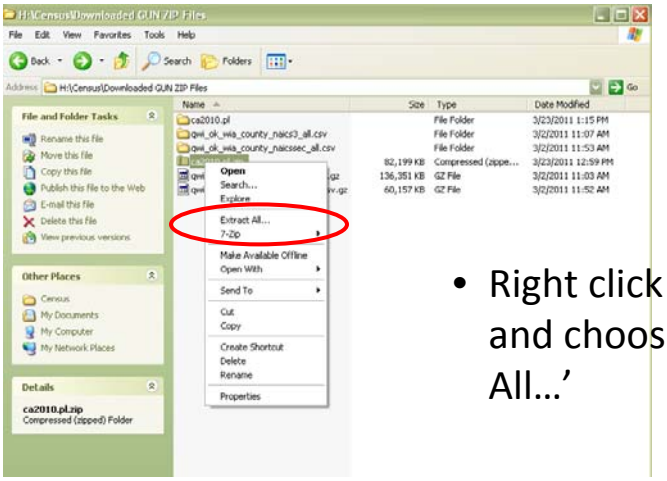
- Open Folder

Opening the Data



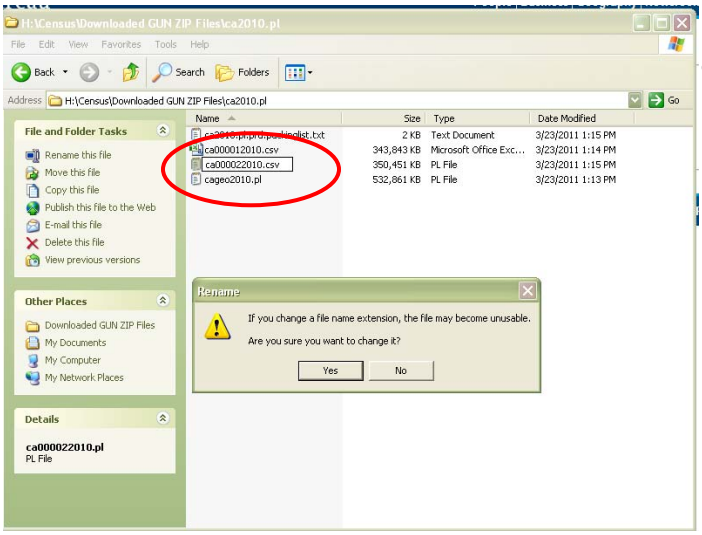
- This will appear different on your computer, depending on where you saved the file

Opening the Data



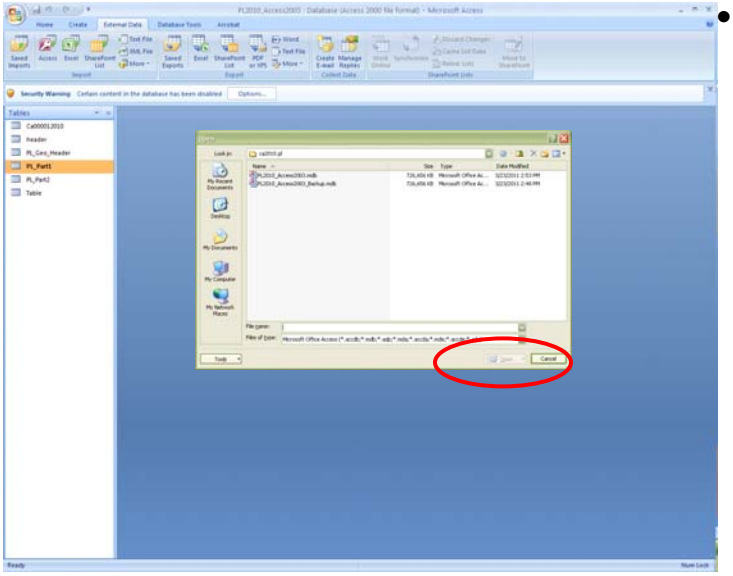
- Right click on the file and choose ‘Extract All…’

Opening the Data



- After extracting, change file extensions from *.pl to *.csv

Opening the Data



- Using Microsoft Access, open the *.mdb file that you downloaded from the FTP site

Steve Barker
State Data Center Program Manager
steven_barker@okcommerce.gov
(405) 815-5182

OKLAHOMA COMMERCE

900 N. STILES OKLAHOMA CITY | OKcommerce.gov