



User Manual



Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
info@azavea.com | T 215.925.2600 | F 215.925.2663
Web azavea.com | Twitter [@azavea](https://twitter.com/azavea)

Table of Contents

Table of Contents 2

Background 3

Installation Instructions..... 3

 Installation from binaries 3

 Building from Source..... 3

Windows Form Application 4

Command Line Application 5

 Getting Help 5

 Getting the Data You Want 5

 Optional Parameters 7

Troubleshooting 9

Appendix A: Command Line Argument Reference..... 10

Background

The ACS Alchemist is a software tool that can help you extract specific portions of the American Community Survey (ACS). This version currently supports:

- ACS 2005-2009 (released in December 2010)
- ACS 2006-2010 (released in December 2011)

This tool was developed by Azavea in collaboration with Jerry Ratcliffe and Ralph Taylor of Temple University and partially funded by a Predictive Policing grant from the National Institute of Justice (Award # 2010-DE-BX-K004). The source code is released under a GPLv3 license and is available at:

<http://github.com/azavea/ACS-Alchemist/>

Installation Instructions

Currently the ACS Alchemist only runs on Windows. In order to view the output from this program, you'll need a program that can read Esri shapefiles.

Installation from binaries

1. Run ACSInstaller.exe
2. Click Next and check the prompts at each step, until the installation is complete

You can launch the application from the Start Menu as follows:

- Click on [Start]->[Programs]->[Azavea]->[ACS Alchemist]

The Alchemist will download files to an output directory, so you will need to have write permissions for this directory. If you receive permission errors when attempting to run the application in Windows Vista or Windows 7, please try the following steps:

On Windows 7 or Windows Vista:

1. Right->Click on [Start]->[Programs]->[Azavea]->[ACS Alchemist]
2. Select Properties -> Advanced -> Run As Administrator

Building from Source

1. Follow instructions on cloning the git repo for this project at <http://github.com/azavea/ACS-Alchemist/>
2. (Optional) Build the installer by downloading NSIS utility, and compiling the install script in /scripts/installer/setup.nsi

Windows Form Application

You can start the GUI version of the application by going to the Start menu and selecting the “ACS Alchemist GUI” application. This version of the application has essentially the same options as the command line application, but provides them in a single form. An example appears below.

The screenshot shows the 'ACS Data Ermine' application window. The title bar reads 'ACS Data Ermine'. The main window has a 'File' menu bar. Below the menu bar, the title 'ACS Data Ermine' is displayed in large blue text. To the right of the title, the 'Current Job File' is set to 'C:\sandbox\ACSDDataErmine\PAJob.bt'. The main area contains several input fields and checkboxes:

- 1. ACS Year: 2009 (dropdown)
- 2. State: Pennsylvania (dropdown)
- 3. Summary Level: Census Tract (140) (dropdown)
- 4. Variable File: C:\sandbox\ACSDDataErmine\PAvars.bt (text field with a 'Browse' button)
- 5. Output Directory: C:\sandbox\ACSDDataErmine (text field with a 'Browse' button)
- Optional Parameters section:
 - ☐ Preserve ACS Jam Values
 - ☒ Strip Extra GEOID text
 - Job Name: PAJob01 (text field) with a ☒ Replace Job checkbox
 - ☒ Export Shapefile
 - ☐ Export Fishnet (text field) Fishnet Cell Size (text field)
 - Fishnet Envelope (text field) with a 'Browse' button
 - Output SRID: 2272 (text field)
 - Pj File (WKT) (text field) with a 'Browse' button
 - Boundary Filter (text field) with a 'Browse' button

On the right side of the window, there is a 'Log Messages' section with a large empty text area. At the bottom right, there is a 'Save Log File as' section with the path 'C:\sandbox\ACSDDataErmine\PAlog01.bt' and a 'Browse' button.

Command Line Application

Getting Help

If you need a list of all available command line option, just enter the program name without any arguments.

```
C:\...> AcsAlchemist.exe
```

Getting the Data You Want

Specify State and Year

The American Community Survey has more than 20,000 variables for 50 states, District of Columbia and Puerto Rico and for multiple geographic levels. This tool is designed to enable you to extract up to 100 variables for a single geographic level and a single state in each run. The first step is to decide what data you want to extract. And the first decision is to pick a state and an ACS year.

This version of the application supports two years of ACS data:

- * ACS 2005-2009 (released in December 2010)
- * ACS 2006-2010 (released in December 2011).

1. To see what states are available, use the following command.

```
C:\...> AcsAlchemist.exe -listStateCodes
```

2. To see what years are available, use the following command.

```
C:\...> AcsAlchemist.exe -listYears
```

3. When you know which state and year you'd like to use, you can pre-download the unfiltered state data files by using:

```
C:\...> AcsAlchemist.exe -s Wyoming
```

Specify Summary Level

Variable summaries are grouped into multiple geographic levels aggregation - tracts, blockgroups, etc. The ACS Alchemist requires a summary level specified by code. These can be listed by typing:

```
C:\...> AcsAlchemist.exe -listSummaryLevels
```

Specify Variables

Each data point for the ACS is represented by a unique variable ID. These are available at:

- http://www2.census.gov/acs2009_5yr/summaryfile/ACS2009_5-Year_TableShells.xls
- http://www2.census.gov/acs2010_5yr/summaryfile/ACS2010_5-Year_TableShells.xls

These files are also copied to the \docs\ directory with in the ACS Alchemist installation. From this file, select up to 100 Unique ID's per job, and save these variable names to a "variable file" in text format. This file is used to determine which variables should be included in your output. Each line corresponds to a variable and consists of two parts separated by a comma. The first is an ID from the "Table ID" column in the ACS Excel file, the second is an optional name for the column in the output shapefile (if no name is specified, the ID will be used instead). Note that column names are truncated to 10 characters due to limitations of the DBF file used in the shapefile format. In addition, you should avoid using spaces or special characters (other than ~, -, and _) If you specify column names that will result in duplicates, the Alchemist will throw an error. An example of the variables file would be:

```
B01001001,TOTALPOP  
B01001002,TOTALMALE  
B01001026,TOTALFEMALE
```

Assuming this is saved in myVariables.txt in your working directory, you can specify this variable file with the following command:

```
C:\...> AcsAlchemist.exe -s Wyoming -e 150 -y 2009 -v myVariablesFile.txt
```

Specify an Output Directory

By default, the ACS Alchemist will use the "AppData" directory in your Windows user directory as the output directory. For example, on Windows 7, this might be something like:

C:\Users\YourUserID\AppData\Local\ACSImporter\Data. If you want to change this default, you can specify an output directory using the "-outputFolder" switch as follows:

- -outputFolder c:\Sandbox\ACS

Specify a Working Directory

ACS Alchemist downloads several files from the Census Bureau and imports those files into various working databases. By default, the ACS Alchemist will use the “AppData” directory in your Windows user directory. For example, on Windows 7, this might be something like:

C:\Users\YourUserID\AppData\Local\ACSImporter\Data\Working. If you want to change this default, you can specify a working directory using the “-workingFolder” switch as follows:

- -workingFolder c:\Sandbox\ACS\Working

Naming your Job

Each run of the ACS Alchemist is called a “job”. The software will create a job name based on the state, the year and the date, unless you specify a job name.

- -jobName DesiredNameHere

Putting it all together

Here is an example import using what was discussed so far:

```
C:\...> AcsAlchemist.exe -s Wyoming -e 150 -y 2009 -v myVariablesFile.txt
-jobName Test01 -outputFolder c:\Sandbox\ACS -workingFolder
c:\Sandbox\ACS\Working
```

This command will make sure all necessary files are downloaded for Wyoming and import the selected variables for 2009 into an internal database (using SQLite) named "Test01". The variable values are at the blockgroup summary level (150). From here you can also generate your results as a shapefile using either the summary geography or using a 'fishnet' grid of polygon cells.

Optional Parameters

Exporting to a Shapefile

When you're ready to run your full export, you can convert the data to a Shapefile. "-exportToShape" will cause your data to be saved to a shapefile, and, optionally, re-project the data to the specified output projection (see below).

```
C:\...> AcsAlchemist.exe -s Wyoming -e 150 -y 2009 -v myVariablesFile.txt
-jobName Test01 -exportToShape
```

It is important to note that a shapefile is not a single file, but, rather, several files that together form a shapefile. The minimum files are .SHP, .DBF, and .SHX. The ACS Alchemist will also generate a PRJ file.

Setting an Output Projection

The geographic data associated with the ACS variables (tract or blockgroup boundaries) are downloaded from the Census Bureau in GRS80NAD83 decimal degrees (EPSG:4269). If you wish to specify a different projection, you can either add an EPSG spatial reference ID (SRID) or specify a filename containing the WKT of the desired projection. When a projection is listed, the ACS Alchemist will re-project when generating the shapefile. The ACS Alchemist uses the PROJ.4 and ProjNET libraries and the supported SRIDs are listed in the "SRID.csv" file in the installation directory. Here's an example that uses the PRJ file approach:

```
C:\...> AcsAlchemist.exe -s Wyoming -e 150 -y 2009 -v myVariablesFile.txt
-jobName Test01 -exportToShape -outputProjection myproj.prj
```

Saving Job Files

The ACS Alchemist has a lot of options (there are more listed in the Appendix of this document), and it is easy to make mistakes with the command line flags. We have therefore provided a mechanism for saving the variables as a "job file". A job file contains all the arguments you would normally specify on the command line in a text file. The syntax is simple:

- Blank lines and lines that start with "#", are ignored
- All other lines should be a command line flag starting with -, followed by an argument (if required for the particular flag)

In order to use a job file, simply specify its path as the only argument to the ACS Alchemist. Note that file paths specified in the job file should be relative to the directory the importer will be run in, not to the location of the job file. An example job file named "myJob.txt" for the previous command would look something like this:

```
# myJob.txt - gets some data about Wyoming and puts it in a shapefile
# specify Wyoming as the state
-s Wyoming
# specify year
-y 2009
# extract blockgroups
-e 150
# extract the variables in myVariablesFile.txt
-v myVariablesFile.txt
```



```
# save the data to a directory
-outputFolder C:\sandbox\ACS\
# use "Test01" as a job name
-jobName Test01
# export results to a shapefile
-exportToShape
# convert the output to projection specified in myproj.prj
-outputProjection myproj.prj
```

This Job file can then be run simply by using this command:

```
C:\...> AcsAlchemist.exe myJob.txt
```

Troubleshooting

This application carries out several tasks and there are a number of things that can go wrong. If you are having trouble, here are a few places to check:

- *Are you connected to the internet?* This application requires access to the internet in order to retrieve files from the U.S. Census Bureau. That means it won't generally run unless you are connected.
- *Do you have write access to folders?* ACS Alchemist needs to write files to a working directory and to an output directory. It can support remote file shares, but you will need to have both read and write access to these directories.
- *Check your variables and parameters?* There are a lot of options in the application. The variables need to be an exact copy of the ACS variables and the parameters must all be correct.
- *Still having trouble?* There is an error log generated in the installation directory: C:\Program Files (x86)\Azavea\ACS Alchemist\ This will sometimes have additional information that is not available in the user interface.

Appendix A: Command Line Argument Reference

- `-s` : State Code
 - **(Required)**
 - Should be the name of the state, see `-listStateCodes` for available states.
- `-y` : ACS Year
 - **(Required)**
 - Should be a numeric (integer) value corresponding to the year of the ACS survey (e.g. 2009, 2010, etc.)
- `-e` : Filter Spatially by Census Summary Level
 - **(Required)**
 - Should be a numeric summary code, see `-listSummaryLevels` for available codes
- `-v` : Export variables listed in file
 - **(Required)**
 - Should be the name of a file containing requested variable names in the **ACS2009_5-Year_TableShells.xls** file.
 - If you're exporting a shapefile, these aliases must conform to shapefile requirements (10 character max, unique). Note that the margin of error columns create 'm'+[Alias] columns, and may conflict with variables in the file. A maximum of 100 columns are allowed per shapefile export.
 - Format: UNIQUEID,ALIAS
- `-f` : Filter Spatially by optional shapefile
 - **(Optional)**
 - Should be the name of a shapefile whose geometries will be used to create a bounding box filter when importing the census data.
- `-j`
- `-jobName` : Specify a name for this job / shapefile
 - **(Optional)**
 - Enter any name to be used internally in the database, or as the shapefile name when exporting. Should conform to filename requirements on current Operating System.
- `-r`
- `-replaceJob` : Replace an existing job / shapefile
 - **(Optional)**
 - This parameter will cause the importer to replace its internal cache for the specified job.
- `-outputProjection` : Provide the .prj file of a desired projection to operate in
 - **(Optional)**
 - Should be a srid or a filename containing WKT of the desired output projection.

- o Default is to not reproject census data, the census data is in **NAD83WGS84**
- **-exportToShape** : Export results to shapefile
 - o **(Optional)**
 - o Creates a new shapefile in **-outputFolder**, with the name **-jobName**, using the census summary polygons specified by **-e**, and in the projection specified by **-outputProjection**.
- **-exportToGrid size** : Export results to fishnetted shapefile where value = # feet
 - o **(Optional)**
 - o Creates a new shapefile in **-outputFolder**, with the name **-jobName**. The geometries for this are vector cells of size **size**, generated in the **outputProjection**. The attributes for each cell are determined by whatever census summary geometry falls under the centroid of the cell.
- **-gridEnvelope** : Align the grid cells to an envelope in a file
 - o **(Optional)**
 - o This should be a filename of a file containing a reference envelope on which to align your grid.
- **-includeEmptyGeometries** : Keeps empty cells and shapes around during export
 - o **(Optional)**
 - o This parameter causes the exporter to include cells or geometries even though they do not overlap a summary polygon with data.
- **-outputFolder** : Specify where you'd like the results saved
 - o **(Optional)**
 - o Sets the location where exported shapefiles will be saved. Helpful when keeping output organized.
- **-workingFolder** : Specify where you'd like the working files to be saved
 - o **(Optional)**
 - o Sets the location where downloaded data will be stored while output is generated. Default is Windows user \AppData\ directory
- **-preserveJam** : Optional flag to preserve jam values
 - o **(Optional)**
 - o This parameter causes the exporter to not attempt to parse the error values as numbers, which will preserve non-numeric jam values like "."
- **-listStateCodes** : Displays a list of available state codes
 - o **(Optional)**
 - o This command will list the available states, and immediately exit.
- **-listSummaryLevels** : Displays a list of available census summary levels
 - o **(Optional)** This command will list available summary levels and immediately exit.

- `-stripGEOIDColumn`: Optional flag to create a stripped copy of the GEOID column
 - **(Optional)**
 - This parameter causes the exporter to add a copy of the GEOID column "GEOID_STRP" that does not have the leading "15000US"