Nevada and Ohio’s

ASU Creation and Mapping Script in R

**Authors**

Michael Frank, Ohio LAUS Analyst

([michael.frank@jfs.ohio.gov](mailto:michael.frank@jfs.ohio.gov))

David Schmidt, Nevada Chief Economist

([deschmidt@detr.nv.gov](mailto:deschmidt@detr.nv.gov))

Contents

[Conceptual Approach 1](#_Toc140832453)

[The Snake Method Visualization 2](file:///R:\private\OWD\LMI\LMIStaging\LaborForceStats\LAUS\ASU\ASU%20Project%20with%20Nevada\Nevada%20and%20Ohio’s%20ASU%20Creation%20and%20Mapping%20Script%20in%20R%20Manual.docx#_Toc140832454)

[Input 2](#_Toc140832455)

[So Why R? 2](#_Toc140832456)

[Download R and R Studio 2](#_Toc140832457)

[Getting Started 3](#_Toc140832458)

[Open RStudio and the Code 3](#_Toc140832459)

[Update Variables 3](#_Toc140832460)

[Run the Code 3](#_Toc140832461)

[Manual Workflow 6](#_Toc140832462)

## Conceptual Approach

The conceptual approach behind this R script comes from Amanda Rohrer, former LAUS Analyst in Minnesota. As her state’s unemployment rate dropped and averaged below 4.0%, manual designation of ASUs became so labor intensive that it justified automation. Her process used SQL Server and ArcGIS to create and map potential ASUs.

The “snake” approach she settled on prioritized the contiguity of census tracts. Her process starts with the census tract with the highest unemployment rate. It then looks at all the tracts that are contiguous and adds the next tract with highest unemployment rate to the first census tract. This repeats until the first area’s combined unemployment rate falls below the ETA’s 6.45% requirement. After it finds the first ASU, the program starts all over with the tracts that haven’t been designated yet. It selects the tract with the highest unemployment rate, looks to its neighbor tracts, adds the one with the highest rate, and repeats until we fail to meet the 6.45% unemployment rate. The process continuously loops until there are no more ASUs it can designate.

The designated ASUs are guaranteed to be contiguous and have an unemployment rate of at least 6.45%. **However, no population control is used so analysts will still need to manually review ASUs that do not meet the population threshold of 10,000 persons.** Sometimes these smaller ASUs can be combined with another ASU or need to be dropped entirely.

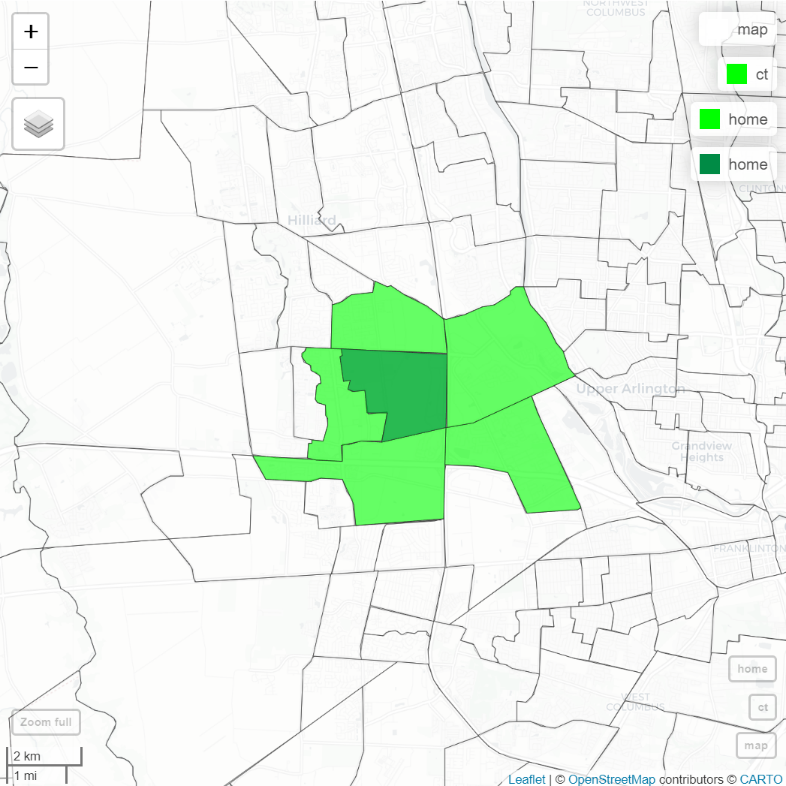
## The Snake Method Visualization

In Step 1, the tract with the 19.3% unemployment rate has been identified as our state’s highest unemployment rate tract. We then find the set of tracts that touch our highest tract.

In Step 2, we add the 7.3% tract because it is the tract with the next highest unemployment rate.

In Step 3, we find which tracts touch either of the two highest tracts. We then add the census tract with the next highest unemployment rate, which in this case would be the 9.5% tract to the northwest.

**Step 1**



**19.3**

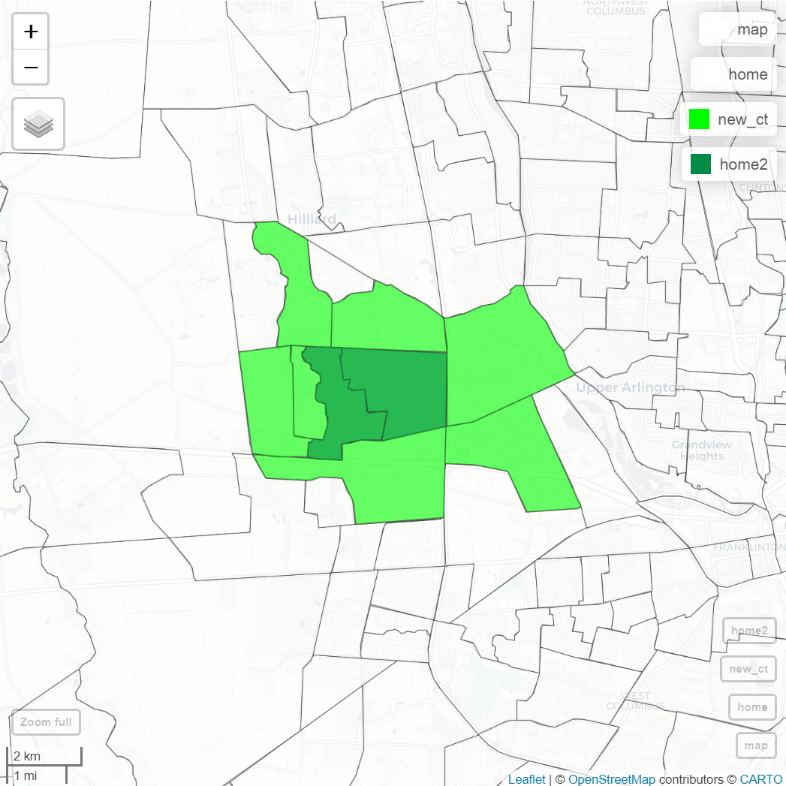
**6.4**

**3.2**

**4.7**

**1.2**

**7.3**



**4.7**

**6.4**

**7.3**

**19.3**

**3.2**

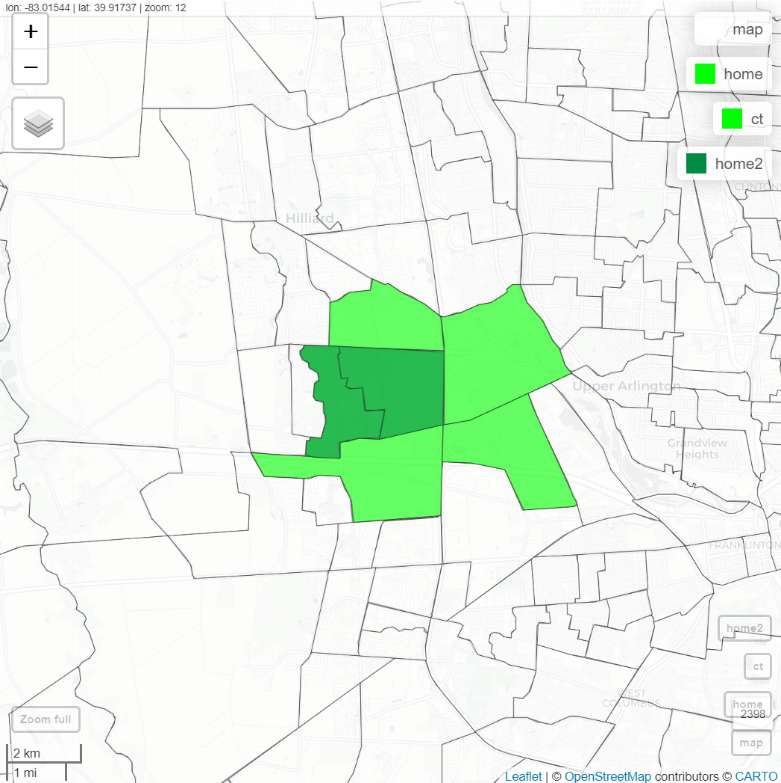
**1.2**

**9.5**

**5.6**

**8.8**

**Step 3**



**4.7**

**6.4**

**1.2**

**7.3**

**19.3**

**3.2**

**Step 2**

## Input

The only external input is the ***ST\_asuYY.xlsx*** file that the National Office distributes.

## So Why R?

The move from SQL Server and ArcGIS to R was motivated by accessibility and user experience. Truthfully, Minnesota’s approach is difficult to replicate. It requires analysts to have a SQL Server with the appropriate permissions as well as ArcGIS. On the other hand, R is a free, open-source software environment with powerful mapping capabilities. Only needing R is a huge plus and will hopefully mean that more states can have access to an automated ASU process.

## Download R and R Studio

The first step is to get R and RStudio installed. R is the software environment used for statistical computing and graphics. RStudio is an integrated development environment that providers a user-friendly interface for writing, running, and debugging R code. R is necessary for RStudio to run.

**Please check with your state’s own IT department about getting R and RStudio.**

For information about R, please visit [The R Project for Statistical Computing website](https://www.r-project.org/).

## Getting Started

* Create a folder where this year’s ASU files will be stored.
* Download the ***ST\_asuYY.xlsx*** file from EUSWeb. Save this file in the ASU folder.
* Save the ***ASU Creation and Mapping Script.R*** file in this folder as well.

## Open RStudio and the Code

Now that you have the setup done, we are ready to see the code.

* Find RStudio and open it.
* Press Ctrl+O. Browse for and then open the ***ASU Creation and Mapping Script.R*** file.

## Update Variables

For this script to work in your state, we need to update a few lines of code.

* Line 20 – The working directory is simply the folder which houses our files. Copy the file path from the folder you created above. Then, paste it in this format: setwd(“file path”). Change all backslashes ( \ ) to forward slashes ( / ).
* Lines 23 – Replace “39” with your state’s FIPS code.
* Line 24 – Update to your state’s BLS file name, ***ST\_asuYY.xlsx***.
* Line 25 – Update only if the unemployment rate threshold from ETA has changed.
* Line 28 – Update to the correct vintage year for the census tract shapefile.
* Lines 42, 144, 146, 151 – Update the year to match the population column in the BLS file.

## Run the Code

To run the code, highlight the lines you wish to run and then press Ctrl+Enter.

**Please run these sections one at a time**.

* Lines 6-7 We need to install “packages” in R to use some of the functions in this code.

Lines 6-7 will tell R to install the list of necessary packages.

The following warning message is okay and should be ignored:

WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the appropriate version of Rtools before proceeding: https://cran.rstudio.com/bin/windows/Rtools/

This warning message is not okay:

Warning in install.packages : unable to access index for repository https://cran.rstudio.com/src/contrib: cannot open URL ‘https://cran.rstudio.com/src/contrib/PACKAGES’

Potential Fixes:

* Go to Tools > Global Options > Packages > uncheck “Use secure download method for HTTP.
* For BLS users, packages may not install due to VPN and firewall settings. An application (CNTLM Proxy) running in the background for packages to be installed is the workaround. Contact LAN and request application.
* Lines 10-17 After installing our packages, we need to load them so they can be used.

Run Lines 10-17 to load our packages.

The following warning message is okay:

Warning message:

Package ‘XYZ’ was built under R version 4.2.3

* Line 20 Setting the working directory lets R know which folder our data file is in.

You updated this line previously, now it’s time to run it.

* Lines 23-25 Run these lines to set our State FIPS, BLS file name, and rate threshold.
* Line 28 This line uses the tigris package to download the census tract shapefile from

the US Census Bureau website. Run this line to save that geographic information

to the R workspace. It may take a few seconds to download***.***

***Note:*** If you close and reopen R Studio, you will need to run Lines 10-28.

* Line 31 Here we read in the data from BLS’s file. Run this line to load that data as well.
* Lines 34-47 Line 34 removes “14000US” from each entry in the geoid field so that it can be

joined with our shapefile. Lines 36-47 selects the ten most important fields.

* Line 50 Next, we join BLS’s data with the Census shapefile.
* Line 51 This line arranges our data by unemployment rate from highest to lowest, and

then by unemployment from highest to lowest. We do this to ensure our highest unemployment tracts are picked first.

* Lines 54-60 We set a few variables that are needed for the automatic designation.
* Lines 64-115 We have arrived at the meat and potatoes of this script – the creation loop.

These lines of code will act out the Snake Method we discussed above.

***Please note that this section of the code takes about 50 minutes to run for Ohio, which has 3,168 census tracts. Simply run the code and wait until you see “Finished!” print in the Console.***

***It may look like nothing is happening, but you can tell R is running by the red “STOP” icon on the right of the console window.***

* Lines 118-130 Now, we want to save this data to an Excel file so that we can manipulate the

census tract assignments. First, we remove the geographic information and join our ASU assignment to the BLS data. Then, we remove unneeded fields and assign “0” to any tracts that were not assigned to an ASU. Lastly, we create two Excel workbooks and save our ASU data on them. You will find these Excel files in the working directory folder we set above.

“ASU Map Assignment.xlsx” is for editing.

“Original ASU Map Assignment.xlsx” is the copy of our original assignment.

* Lines 133-134 Line 133 will read in the Excel file with our ASU designations and Line 134 joins

that data with the geographic shapefile.

* Lines 137-138 We filter the data to show only those tracts that are in an ASU.
* Lines 140-141 And then we filter the data to show those tracts that are not in an ASU.
* Lines 143-146 Here’s another great part! Lines 143-146 create an interactive map for you to

review and work with your ASUs. You can click on a census tract to open a popup, which displays all its relevant information (like geoid and rate).

* Lines 149-154 Run these lines to create a table to review the population and unemployment

rate of each ASU to ensure they meet ETA’s requirements.

* Line 157 View the criteria check table we just made.

***Note: asunum = 0 is the information for those tracts not assigned to an ASU.***

* Lines 160-173 Once you are finished with your ASU assignments, you can run

Lines 160-173 to create a batch file that can be uploaded to LSS. A file called “MyBatchFile.txt” will be saved in your working directory folder.

## Manual Workflow

Most of your work relating to this R script will be the manual changes you make after our script has given you an initial assignment. Here are some recommendations to get started:

1. Map your initial assignment and determine if any ASUs are touching. If not, on to Step 2.

If they are, you will want to open your Excel file and change the number in the asunum column to connect these areas. For example, Ohio’s 1st, 2nd, and 6th ASUs all touch. We then changed both “2” and “6” to “1” in the asunum column.

***Note: the asunum field should be no more than 4 digits long, e.g. “1”, “23”,”999”.***

1. Run Lines 149-157 and determine which ASUs are not meeting the population or rate requirements. There are sort buttons by the headers of the table so you can sort from lowest to highest.
2. Of the ASUs that do not meet the population requirement, see if you can merge them with nearby ASUs. This is the tedious part of turning tracts on/off to potentially join ASUs together.

***Note:*** Each time you make updates to the Excel file:

1. Save the workbook. You do not need to close the workbook between edits.
2. Rerun Lines 133-146. You can run all these lines at once, just expect it to take a few seconds.
3. Run Lines 149-157 to reassess the population and rate criteria of your ASUs.