

DEM 5093/7093 – Spring 2023
Lab 3
Data Creation and Geocoding

1) Let's get some address data from the AtoZ databases.

Go to lib.utsa.edu → Databases (middle of screen under the search bar) → A → AtoZ Databases (you will need to sign in to your UTSA account)

2) Select the → Get Started option for 30 Million Businesses & Executives (bottom left)

3) We want to limit our search for the zip codes for the West side of San Antonio.

Check Search By → ZIP Code

Enter three zip codes: 78207, 78201, and 78204

We are going to search for WIC (Women, Infants, and Children) clinics.

Check Search By → NAICS Codes

Enter: 624190 (Dr. Sparks Googled this)

(<https://www.census.gov/naics/?99967>)

AtoZ databases FREE: Job Search, Business/People Profiles, Mailing Lists & Sales Leads, Email Service, Background/Criminal Search

Data Quality Intro Videos FAQ About ▾

Home 30 Million Businesses & Executives > Search

Search By —

Geography —

- ☐ State
- ☐ City
- ☐ Metro Area
- ☐ County
- ☒ ZIP Code
- ☐ Address
- ☐ One Per Address
- ☐ Radius Around an Address or ZIP Code
- ☐ Map Based Search

Keyword / Business Type —

- ☐ Keyword / Business Type
- ☐ Major Industry Group
- ☐ Home Based Business
- ☐ Business Type Code
- ☐ SIC Codes
- ☒ NAICS Codes

ZIP Code

ZIP Codes

Enter 3 or 5-digit ZIP Codes in boxes below.

78207 78201 78204

+ Paste up to 200 ZIP Codes below

NAICS Codes

NAICS Codes

Enter 4 or 6 digit [NAICS Codes](#) in boxes below.

624190

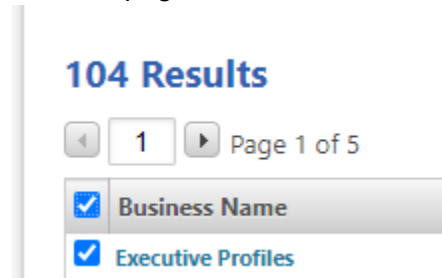
☒ All lines of business
☐ Primary line of business

+ Paste up to 200 NAICS Codes below

And click Search.

I got 104 results (in 2023 – you may get more/less another time). Now's the hard part, we need to select all of these.

Right under the results number there is a check box. Check it, then hit the arrow button to go to the next page and check the box for each of the 5 pages.



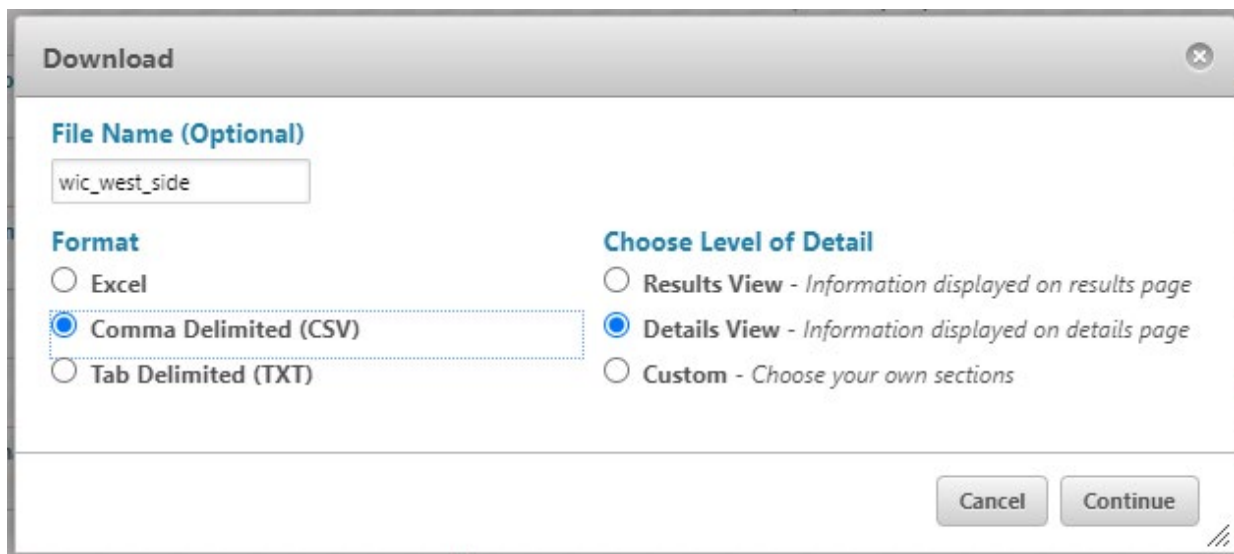
Do this for all 5 pages.

Then select Download.

Format: CSV

Choose Level of Detail: Details View

File Name: wic_west_side



And hit Continue. The CSV should download.

Move it to your GIS class folder!

You can open it and look at it.

We are going to use this for two different purposes. First, we will add the data to our Lab 2 project (did you save it?).

Open your Lab 2 project in QGIS.

Go to Layer → Add Layer → Add Delimited Text Layer → browse to your class folder and select the CSV file you downloaded.

Check Geometry Definition. The following should have been filled automatically:

X field: Longitude

Y field: Latitude

Geometry CRS: EPSG:4269 - NAD83

Here's what mine looked like:

Data Source Manager | Delimited Text

File name: C:\Users\xee291\Documents\UTSA 2022-2023\GIS Class\gis_class\Lab 3\wic_west_side.csv

Layer name: wic_west_side Encoding: UTF-8

File Format

- ☒ CSV (comma separated values)
- ☐ Regular expression delimiter
- ☐ Custom delimiters

Record and Fields Options

Geometry Definition

- ☒ Point coordinates X field: Longitude Z field:
- ☐ Well known text (WKT) Y field: Latitude M field:
- ☐ No geometry (attribute only table) ☐ DMS coordinates

Geometry CRS: EPSG:4269 - NAD83

Layer Settings

Sample Data

	Source	Date	Obsolescence Date	Business Name
	abc Text (string)	abc Text (string)	abc Text (string)	abc Text (string)
1	AtoZDatabases	02/22/2023	08/22/2023	Executive Profiles
2	AtoZDatabases	02/22/2023	08/22/2023	Methodist Hospital Texsan
3	AtoZDatabases	02/22/2023	08/22/2023	Central Cardiovascular Inst of San Antor
4	AtoZDatabases	02/22/2023	08/22/2023	Texas Diabetes Institute

Close Add Help

Hit Add.

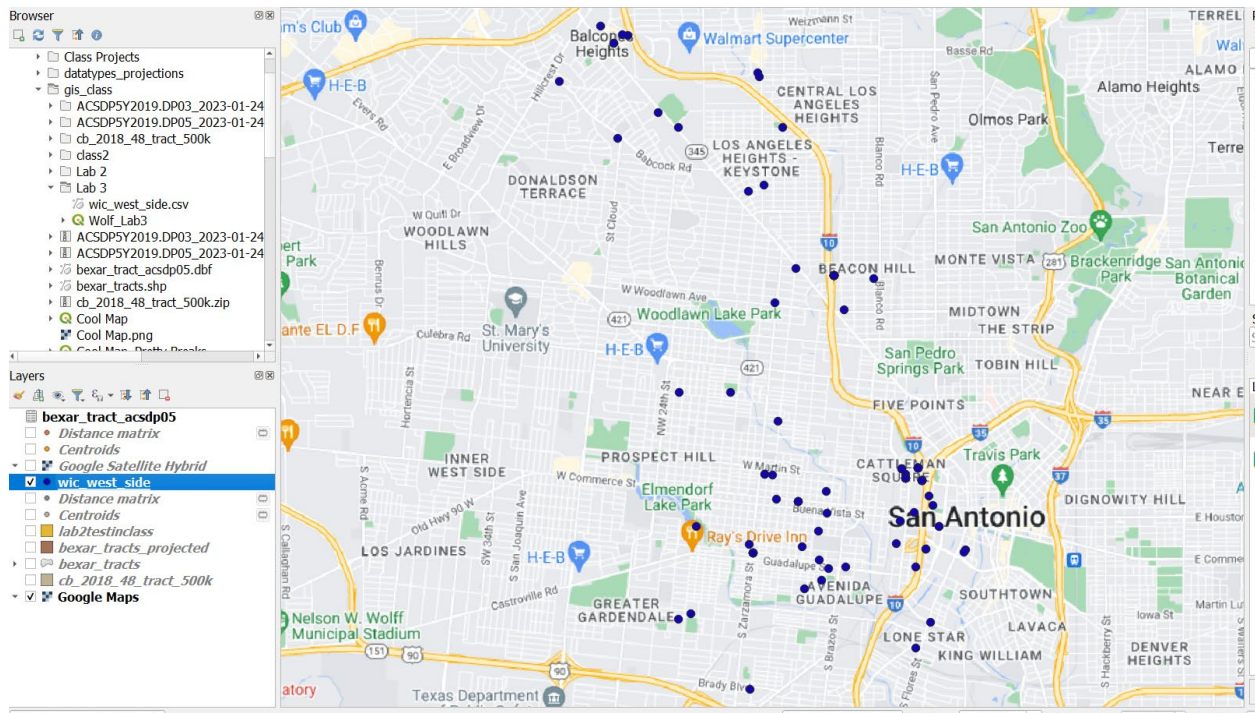
Hit Close.

You should now have lots of dots on the west side of San Antonio.

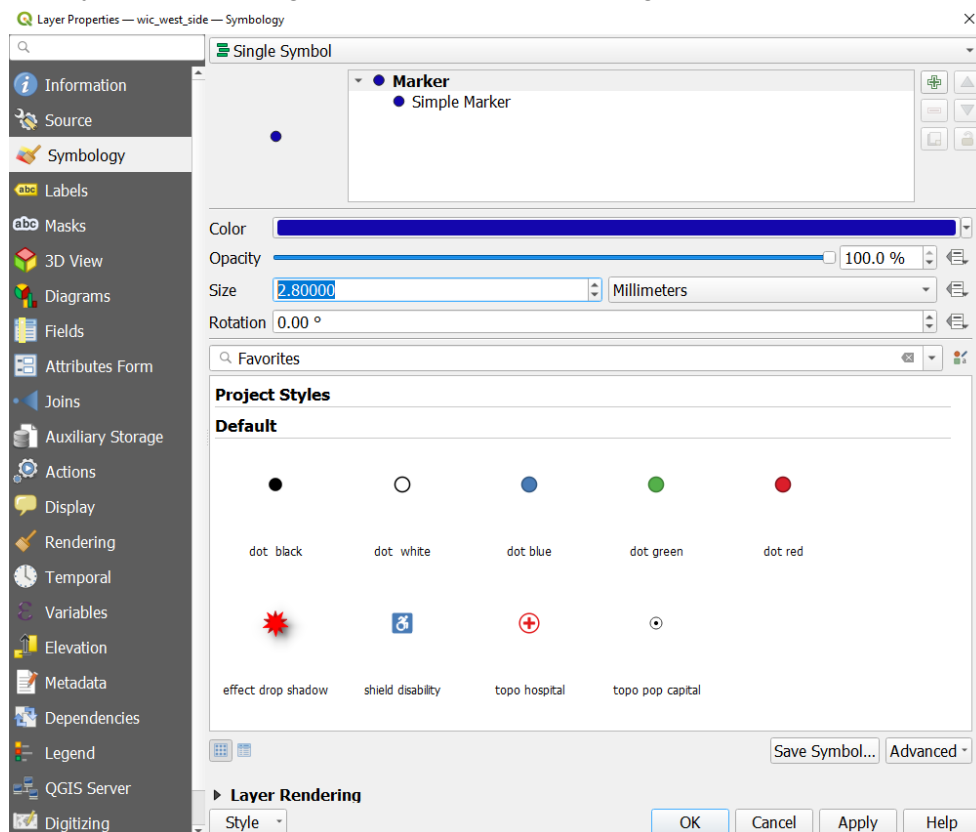
Right click on the layer and hit Zoom to Layer(s).

Great, you've got data now!

Here's mine:



If your dots are difficult to see, you can change the color and size of them by right clicking on the layer and selecting Properties. You can change the Color and Size.



Next, we will try to geocode these addresses ourselves to see how that is done.

If you haven't done so, go to Plugins → Manage and Install Plugins and search for MMQGIS and install it. Close this window when it's done.

Go to your class folder and open the CSV file you downloaded from AtoZ.

Sort the sheet by Physical Address (column F). There is one record without an address (it will be sorted to the last row whether you do Sort A to Z or Sort Z to A). Delete that row.

Save it again and close it.

Go to the MMQGIS menu, and select Geocode → Geocode CSV with Web Service

Select the CSV file you just created.

Address: Physical Address

City: Physical City

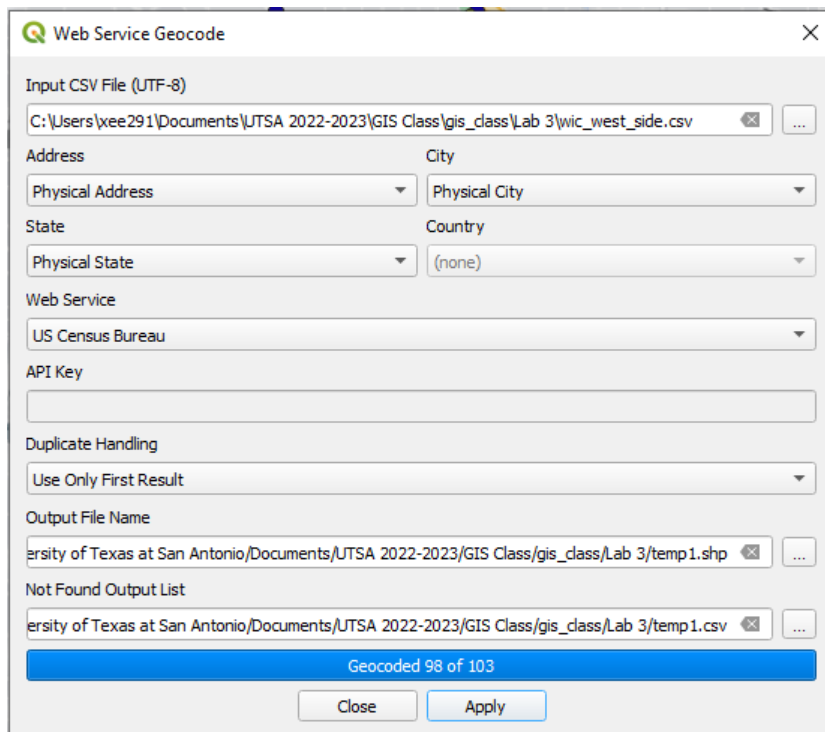
State: Physical State

Country: none

Web Service: US Census Bureau

Output File Name: Change temp1.shp to WIC1.shp (**note:** I didn't change mine below, but you should to help with Homework 4).

Hit Apply. It will take a few minutes to finish, as it works it will tell you how many it's found. Just let it go until it's finished.

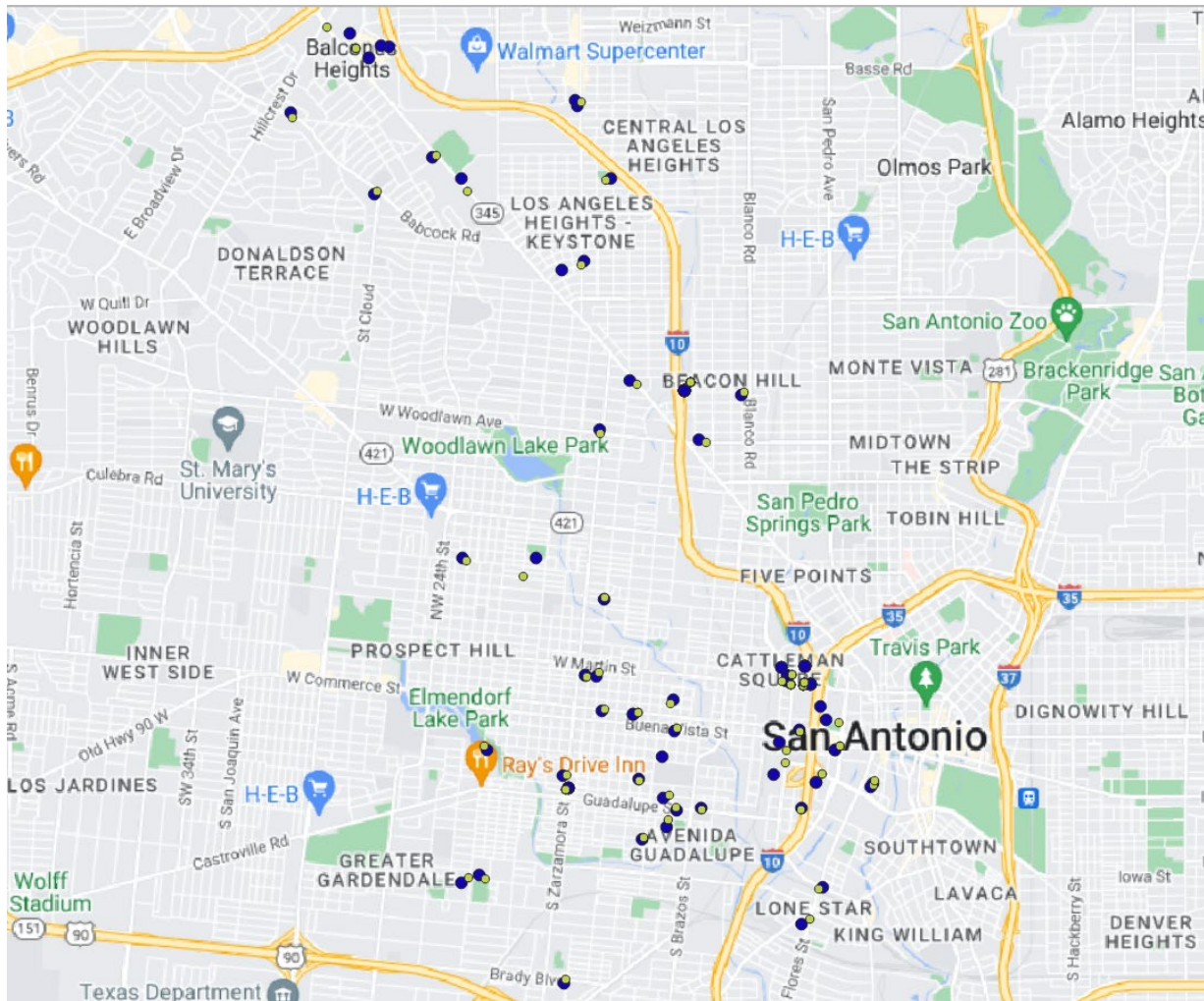


The screenshot shows the 'Web Service Geocode' dialog box. It has a title bar with a green icon and a close button. The 'Input CSV File (UTF-8)' field contains the path 'C:\Users\xee291\Documents\UTSA 2022-2023\GIS Class\gis_class\Lab 3\wic_west_side.csv'. Below this are four dropdown menus: 'Address' (Physical Address), 'City' (Physical City), 'State' (Physical State), and 'Country' ((none)). The 'Web Service' dropdown is set to 'US Census Bureau'. There is an empty 'API Key' field. The 'Duplicate Handling' dropdown is set to 'Use Only First Result'. The 'Output File Name' field contains 'ersity of Texas at San Antonio/Documents/UTSA 2022-2023/GIS Class/gis_class/Lab 3/temp1.shp'. Below that, the 'Not Found Output List' field contains 'ersity of Texas at San Antonio/Documents/UTSA 2022-2023/GIS Class/gis_class/Lab 3/temp1.csv'. A blue progress bar at the bottom indicates 'Geocoded 98 of 103'. At the very bottom are 'Close' and 'Apply' buttons.

Mine got 98 of 103.

The process created a file in your class folder with the mistakes (temp1.csv; you could have named it in the last step). You can open it and look at the ones it couldn't find. Most of mine were W Interstate 10 addresses.

I can also see that some of the geocoded addresses are not exactly the same as the Lat/Long points we initially loaded in. In the map below, the dark blue circles are the Lat/Long points and the light green circles are the geocoded points we just did with the addresses.



You can repeat this process with the OpenStreetMap/ Nominatim service (instead of the US Census Bureau). You may get a different number matched.

This is an imperfect process.

If you zoom in, you'll see the geocoded dots are in streets but slightly on the right or left side. The Lat/Long dots are exactly where the buildings are off the streets.

Homework 4. – Only in QGIS.

Repeat this process in the same map project, but use NAICS code 445110, for grocery stores. Use the same zip codes to search within. Generate an image of your map with a legend for both the WIC and grocery store layers (**Hint:** remember Lab 1?).

Report the geocoding results in terms of % correctly geocoded. Use the US Census Bureau's geocoding service.

Submit your map with the data displayed via Blackboard before **6pm on 3/1/2023**.

Hint for the Legend (This is only Lab 3 showing, I haven't added the homework part yet):

