Assignment 3

PPHA 38520

**Buffering!**

*Goal of Lab: To learn more about spatial queries and what can be added through buffering and clipping  
Note: Before starting be sure to download the assignment files and unzip to a local folder*

* All the processes in this lab build on what we’ve done in the demos, assignment so far
* While you may consult with your classmates on individual steps or processes if you are stuck, you need to do your own work and submit via Canvas

**Background**

You need to conduct a study of households that are within three miles of the development around the Obama Presidential Library and Museum.

Due to the impacts of the development on the neighborhood and sensitive nature of the survey, the research organization wants to target households within low-income tracts (defined, for these purposes, as tracts with a median household income below $30,000). We will identify the households that meet these criteria, make a map, and answer some key questions.

**Data**

* Mappable file of IL tracts with ACS data
* Geocoded .csv file of all residential addresses within nearby ZIP codes

**New skills**

* Turning a ‘flat’ (.csv) file with latitude/longitude info into a ‘mappable’ file of point data
* Registering data to use a specific coordinate system
* Creating a new feature class
* Editing a layer
* Creating a buffer
* Clipping data

**Overview**

1. Plot geocoded .csv file using latitude/longitude data from the attribute table; register it with the same coordinate system as the IL tract file
2. Query tract file to show only those tracts with medhhinc < 30,000; add selection to layout
3. Create a new feature class ; edit new point file; add point at location of the Obama Library; export as a shapefile
4. Create 5-mile buffer around Obama Library point layer
5. Clip the tract select layer using the 5-mile buffer
6. Clip the household point data using the file from step 5
7. Make a map and answer questions

**Coordinate systems**

To start, unzip the data for assignment 3 into a new folder, and then open assignment3.qgz.

The goal of assignment 3 is to build a 3-mile buffer around the location of the Obama Presidential Library development to query a list of addresses. To do this, we need to ensure our files contain information projected in the same coordinate system as the data frame. When QGIS is started with a new, blank map, the default coordinate system (based on the application’s default settings) will use the coordinate system of the first layer added (note: in the absence of coordinate information, the default project settings will assume a layer is in World Geodetic System 1984, European Petroleum Survey Group [EPSG] code 4326). In our case, the first layer is in the Illinois tract file. Right-click on il\_tract and go to properties.

A screenshot of a computer

AI-generated content may be incorrect.  
  
You can confirm that the project coordinate reference system (CRS) matches EPSG 4269 – NAD 83 by looking to the bottom right of the window. A box shows the EPSG number for the QGIS project.

A screenshot of a computer

AI-generated content may be incorrect.

They are the same. As we create new mappable data, it may become necessary to give layers a specific projection. The software will automatically realign data layers if they have a different coordinate reference system, but it will generate warnings, slow performance, and – in some cases – result in bad calculations. You may want to be diligent about assigning projections to map layers.

**Plotting csv file**

A screenshot of a computer

AI-generated content may be incorrect.

The .csv file in the QGIS project named *obama\_fulladvo\_mappable\_t.csv* contains geocoded addresses of over 200,000 housing units in the ZIP codes surrounding the Obama Library. Geographic information is currently being stored in the file’s attribute table in longitude and latitude fields. In order to plot these addresses on our map, we’ll use a tool in the processing toolbox called “Create points layer from table.” Locate the gear icon  in the menu buttons near the top of the screen to open the processing toolbox. Search the text box for the create points layer tool, or find it using the drop-down arrows, under the “vector creation” section. Double-click to open it.

Set your X field to longitude, Y field to latitude, and keep your target CRS as WGS 84 (EPSG 4326). Use the ellipses to choose a folder location so that you do not create a scratch layer. Click to run.

**Note**: you can also generate a point layer at the time that you add a CSV file to your project, by indicating that there are geometry fields present.

## Query tract file

Open the Attribute Table for the ‘il\_tract’ layer. Click on the ‘Select Features Using Expression’ button (looks like a summation sign with a yellow square behind it) . You get a blank expression window to build a query, and a list of options in the middle panel. Click on the “Fields and Values” arrow to expand the available fields list. Double click items from the middle panel to add them to the query, and create a simple SQL expression to filter the database. In this case, we’ll use median income fields. Build the following query: "med\_hh\_inc" < 30000. Press Select Features to select matching features. To create a new layer from these selected tracts, right click on ‘il\_tract’ in the Layers panel and choose ‘Export’ > ‘Save Selected Features As...’ and save as a new shapefile (or GeoPackage, or .geojson format)

A screenshot of a computer

AI-generated content may be incorrect.

**Create a new layer**

There are several ways we could go about creating a buffer around the Obama Library. Perhaps the most intuitive would be to select the tract containing the Obama Library, create it as a new layer, and draw a buffer. But here, we decided to include a demo on how to create new mappable data in QGIS as it is a useful skill.

We know the precise center or location of the Obama Library within its census tract. Google Maps plots the Obama Library development at (41.784304, -87.585469). To create three-mile buffer around this point, we need to first create a layer that is this point. To do so, navigate to menu option Layer > Create Layer > New Shapefile Layer…

A screenshot of a computer

AI-generated content may be incorrect.

Click OK to create the layer.

**Edit new feature class**

The file temp\_point will now appear in the Layers list, but it currently holds no data. To edit this new layer, select it in the layer list by clicking on it, then click on the yellow pencil icon in the menu options (technically, this is the editing toolbar) to begin editing.

A screenshot of a computer

AI-generated content may be incorrect.

The icon will show a selection box and the red pencils icon next to the editing icon will illuminate to indicate the layer editing process is active, as shown above. Next, you will click the “Add Point Feature” icon to place a point – this is also in the editing toolbar, just to the right of the editing icon:

A screenshot of a computer

AI-generated content may be incorrect.

With the point feature option activated, you will get a target icon that allows you to place points directly into the layer. Use the coordinate box at the bottom of your window to hover over coordinate as close as possible to **41.7843, -87.585** (zoom in if needed) and click once to enter the point. You are required to assign an ID number to this point, just enter “1” for now. Click OK to add.

A screenshot of a computer

AI-generated content may be incorrect.

Click the yellow pencil icon once again to stop editing – click SAVE to save your addition. Note: if you are dissatisfied with point placement, you can reopen editing, then click the vertex editor tool (looks like a hammer and screwdriver with a node) from the same toolbar and right click to bring up the X/Y options and move the item.

**Reprojecting and Creating a buffer**

We now want to create a buffer extending 3 miles from our new point. However, the point we entered is in a different map projection that is based on distances stored in decimal degrees, which is a *geographic* rather than *projected* coordinate system (you may recall that linear distance measured in degrees is not uniform). So, for QGIS calculations like this, generally we will need to reproject our dataset to use a projection with linear distances that are stored in meters or feet. Use the processing toolbar to search for “reproject layer.” Set your new point as the input layer, make sure your target coordinate reference system (CRS) uses any projected coordinate system that does not store units in decimal degrees. Then, reproject any other layers if desired/needed.  
  
To create a buffer, return to the processing toolbox search for Buffer (or go to the Vector Geometry >> Buffer). As the Input Features select ‘temp\_point’ and save the output as obama\_buffer3. As the linear unit, select ‘US Survey Miles’ from the drop down and put ‘3’. This creates a 3 mile buffer overlay that is saved as a new polygon layer.

**Clipping**

Clipping is an intersection tool that changes the mappable features of an input layer by “cutting out” parts that do not align with another layer. We will now clip il\_tract selection layer using the buffer. Go to ‘Processing Toolbox’ > ‘Vector Overlay’ > ‘Clip’. Use the selected il\_tract layer as Input Layer, and obama\_buffer3 layer as Overlay Layer. Save the result as il\_tract\_clip. Use this il\_tract\_clip layer to clip the household addresses by running the ‘Clip’ tool again (with household points as Input and il\_tract\_clip as Overlay).an

**Questions**

1. Make a map to show the potentially-impacted households based on the distance criterion in the last step. Make your map professional-looking with appropriate scale, titles, and design elements as we discussed in class and illustrated during the last demo. (40 points).
2. Experiment with buffers of different sizes, e.g. 1 or 2 (or more) miles; how does that change the number of households who would be eligible for the survey? Would it include different places, neighborhoods? (10 points)
3. What other socioeconomic factors do you thinkthe developer should consider for the analysis, if any? How might those parameters influence analysis and results? (10 points)
4. Admittedly, the work done in this assignment is functional and not analytical. The end result is a list of addresses that could be used in a survey to understand the concerns of specific residents and not necessarily an idea to be communicated using a map. Nevertheless, imagine a situation in which you need to convey to someone what our process looks like. Create a map and show the results from one or more queries. If there are other variables and data that you consider important (parks, streets, ACS data, etc.), include it in your map. Make your map professional-looking with appropriate scale, titles, and design elements. (40 points)