# Appendix 2

# Calculating and Visualizing Four Indicators of Alcohol Outlet Density Using QGIS

This appendix demonstrates the initial steps of data preparation and the calculation and visualization of count-based and distance-based indicators for measuring alcohol outlet density using QGIS, described in the <u>Measuring Alcohol Outlet Density: A Toolkit for State and Local</u> <u>Surveillance</u> [PDF – 22 MB]. It was written by Elle Law, MPH, University of North Carolina Gillings School of Global Public Health and Mike Dolan Fliss, PhD, MPS, MSW, Research Scientist, University of North Carolina, Injury Prevention Research Center using QGIS version 3.

## Resources

- CDC Alcohol Research in Action
- <u>CDC Alcohol Measurement Guide</u> [PDF 32 pages]
- <u>CDC Alcohol Measurement Toolkit</u> 🚣 [PDF 52 pages]

# Contents

Part 1: Project Setup Tasks

Adding Shapefiles to QGIS

Adding CSV files as layers

<u>Geocoding</u>

Filtering to Active Licenses

Counting Outlets per County

Joining by Common Location Attribute (County)

Creating a Choropleth Map of Outlets per County

Part 2: Calculating Indicators of Alcohol Outlet Density

Calculating Count-Based Indicators

Counting Alcohol Outlets per Square Land Mile

Counting Alcohol Outlets per 10,000 people

Calculating Distance-Based Indicators

<u>Calculating the Average Distance from a Person to Their Nearest Alcohol Outlet</u> (Person-to-Outlet)

<u>Calculating the Average Distance from an Alcohol Outlet to Its Nearest Outlet (Outlet-to-Outlet)</u>

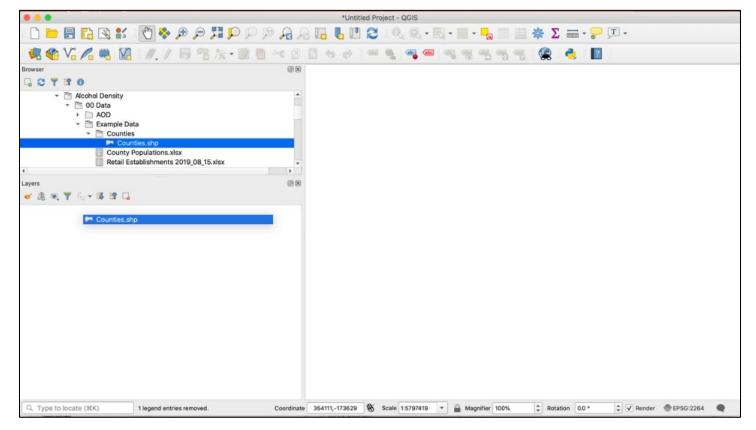
# Part 1: Project Setup Tasks

## **Adding Shapefiles to QGIS**

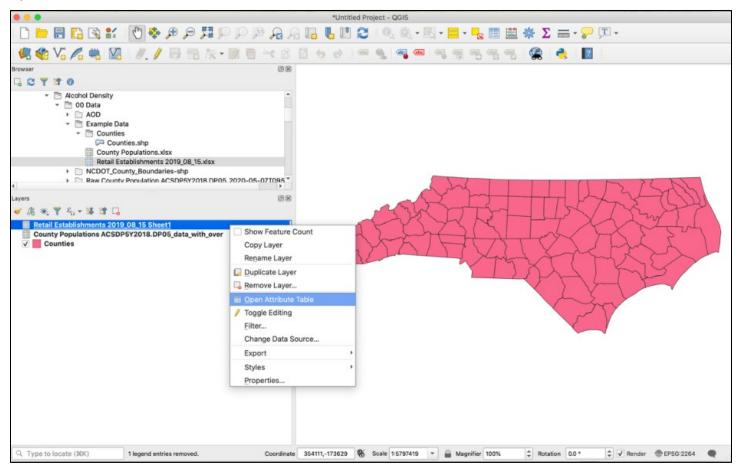
When adding data into QGIS, each file will appear as a layer. The two most common types will be shapefiles and tables. Store data files in one folder for easy reference and then add each dataset into QGIS, which will create a layer for each dataset:

- 1. Open a New Project in QGIS.
- 2. In the top left Browser pane, navigate to the folder containing the datasets.
- 3. Click and drag/drop each file into white space in the Layers window. Shapefiles will automatically appear in the map view once added.

Add layers using the drag/drop method into the Layers window.



Once layers are added, you can explore your layers by right-clicking on the layer to view the Attribute layer.



**Tip:** You can drag and rearrange shapefile layers so they appear in the order that you want. For example, dragging the layer containing alcohol outlets, as in the previous screenshot, the state shapefile layer will overlay outlet points on top of the state map.

**Troubleshooting:** If a shapefile is showing as a layer, but seems to not appear in the map view, click to View All to see if it is mapping in an unexpected location. If data does not overlay as expected, such as points in one area and the base map in another area, make sure shapefiles are on the same projection.

## Adding CSV files as layers

Any tables in .XLSX format must be converted into .CSV to be imported into QGIS.

Use the drag/drop method to add CSV files that do not contain spatial data. If a CSV file contains latitude and longitude decimal degree coordinates in separate columns, you can spatially plot that data.

	County	Status	Mailing Address	Mailing City	Mailing State	Mailing Zip	Phone	Fax	umber/Temp Permi	Latitude	Longitude
	Guilford	Pending-Tem	PO Box 853	Greensboro	NC	27402	3363838861	NULL	00280201MB	-79.7841	36.07426
	Gaston	Pending-Tem	2416 Arden G	Charlotte	NC	28262	NULL	NULL	T00276738	-80.75346	35.34643
	Nash	Active	NULL	NULL	NULL	NULL	2529773000	NULL	00255798AJ	-77.79734	35.94162
	Nash	Active	NULL	NULL	NULL	NULL	2529773000	NULL	00255798AL	-77.79734	35.94162
	Nash	Active	NULL	NULL	NULL	NULL	2529773000	NULL	00255798MB	-77.79734	35.94162
	Alamance	Active	NULL	NULL	NULL	NULL	3366038232	NULL	00236579AK	-79.5198039	36.102254
	Alamance	Active	NULL	NULL	NULL	NULL	3366038232	NULL	00236579AM	-79.5198039	36.102254
	Lincoln	Active	NULL	NULL	NULL	NULL	(704)735-7679	NULL	00229594AO	-81.2944601	35.4277222
	Lincoln	Active	NULL	NULL	NULL	NULL	(704)735-7679	NULL	00229594AK	-81.2944601	35.4277222
0	Lincoln	Active	NULL	NULL	NULL	NULL	(704)735-7679	NULL	00229594AM	-81.2944601	35.4277222
1	New Hanover	Active	7513 Old Oak	Wilmington	NC	28411	(910) 762-73	(919)762-7334	00129898AO	-77.932175	34.23303
2	New Hanover	Active	7513 Old Oak	Wilmington	NC	28411	(910) 762-73	(919)762-7334	00129898AM	-77.932175	34.23303
3	New Hanover	Active	7513 Old Oak	Wilmington	NC	28411	(910) 762-73	(919)762-7334	00129898AK	-77.932175	34.23303
4	Burke	Active	NULL	NULL	NULL	NULL	(828)368-00	NULL	00225462AJ	-81.5625027	35.7435004
5	Burke	Active	NULL	NULL	NULL	NULL	(828)368-00	NULL	00225462MB	-81.5625027	35.7435004
8	Burke	Active	NULL	NULL	NULL	NULL	(828)368-00	NULL	00225462AL	-81.5625027	35.7435004
7	Mecklenburg	Active	8101 Tifton R	Charlotte	NC	28226	(704)749-04	NULL	00214336AJ	-80.721196	35.1177459
3	Mecklenburg	Active	8101 Tifton R	Charlotte	NC	28226	(704)749-04	NULL	00214336AL	-80.721196	35.1177459
,	Randolph	Active	NULL	NULL	NULL	NULL	3364652751	NULL	00263313MB	-79.81345	35.70621
0	Randolph	Active	NIL II I	NIL II I	MITT	NUUL	3364652751	NII II I	0026331341	-79,81345	35 70621

## Example of CSV with Latitude and Longitude columns, in the attribute table view

In this case, each row in the CSV file represents an outlet, and each outlet has latitude and longitude (as in the previous screenshot). To add the CSV file spatially:

1. Go to Layer in Menu, Click Add Layer, Add Delimited Text Layer.

## Add Delimited Text Layer

GIS Project Edit View	Layer Settings Plugins Vector Raster Database Web Mesh Process	sing	
• • •	Data Source Manager		
	Create Layer          Add Layer          Embed Layers and Groups          Add Farm Layer Definition File	abr	
Browser	Add from Layer Definition File I Add Mesh Layer Copy Style 2, Add Delimited Text Layer 企業T		
	Paste Style   Read PostGIS Layers   合業D		
<ul> <li>Alcohol Density</li> <li>00 Data</li> <li>AOD</li> </ul>	<sup>1</sup> Copy Layer <sup>1</sup> Copy Layer <sup>1</sup> Add SpatiaLite Layer <sup>1</sup> Add MSSQL Spatial Layer <sup>1</sup> Paste Layer/Group <sup>1</sup> Add DB2 Spatial Layer <sup>1</sup> Ch # 2		
<ul> <li>Example Data</li> <li>Counties</li> <li>Counties.sh</li> <li>County Populat</li> <li>Retail Establish</li> <li>NCDOT_County_Br</li> </ul>	Save Layer Edits  Current Edits  Add ArcGIS MapServer Layer  Add WCS Layer	企業W	
<ul> <li>► NCDOT_County_Bit</li> <li>► Raw County Popul</li> <li>Layers</li> <li>✓ ▲ ● ▼ ▲ ▼ ↓ ↑ □</li> <li>County Populations ACSDP5Y20°</li> <li>✓ ■ Counties</li> </ul>		J X	

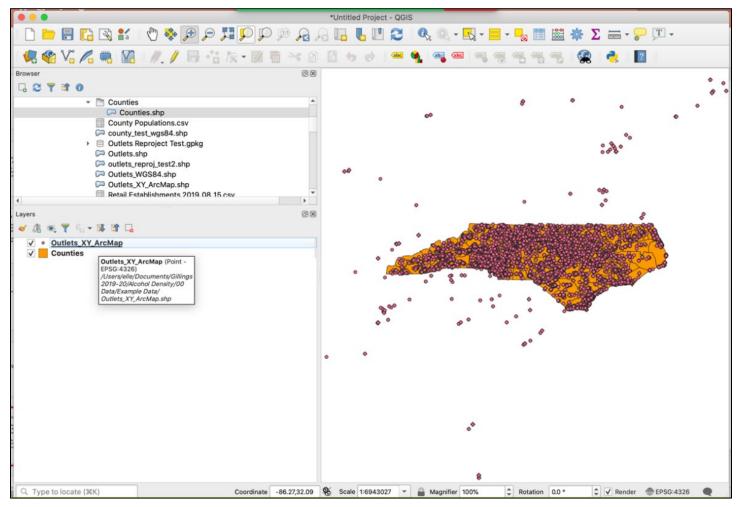
- 2. In the window pop-up, under File Name, browse to the CSV file you want to add by clicking on the "...", then click Open.
- 3. Under Geometry Definition, assign your longitude column to the X field and your latitude column to the Y field.
- 4. Next to "Geometry CRS," select the projection system for latitude-longitude data, "*EPSG:* 4326 WGS 84."

## Pop-up window for "Add Delimited Text Layer"

	Data Source Manager	Delimited Text									
Fowser	File name uments/Gillings 2019-20/Alco	hol Density/00 Data	/Example Data/Reta	il Establishments	2019_08_15.csv						
V <sub>+</sub> Vector	Layer name Retail Establishments 2019_08_15 Encoding UTF-8										
Raster	▼ File Format										
Mesh	CSV (comma separated values)										
Delimited Text	O Regular expression delimiter										
GeoPackage	O Custom delimiters										
🎢 SpatiaLite	Record and Fields Options										
PostgreSQL	▼ Geometry Definition										
MSSQL	Point coordinates	X field Longitu	ude 💌	Z field		•					
DB2 DB2	O Well known text (WKT)	Y field Latitud		M field		•					
Virtual Layer	O No geometry (attribute only table	-	S coordinates Project CRS: EPS0	3:4326 - WGS 84	4 👻						
💮 wms/wmts	Layer Settings										
and the second s	Sample Data										
💮 wfs	File Number / Temp Permit Number	Trade N		Corp N							
ArcGIS Map Server	1 00280201MB 2 T00276738	Club Eclipse Kings Food Mart 2		othy Watlington	Enterprises LLC	• •					
ArcGIS Feature Server											
GeoNode 🗸	Help				Add	Close					

5. Click Add, then Close. A new layer should appear, with outlets represented by points.

## Result of plotting XY coordinates of CSV file



6. Once you have imported the CSV with XY coordinates, it is a temporary file (you can prove this to yourself by checking the file pathway; it will be to the CSV file). You must save it as a new shapefile. To do so, right-click the newly created layer → Export → Save Features As... Click on "…" next to File name to choose your save location, choose a File Name, and provide a Layer Name (optional). Click Ok. Once the new permanent layer appears, you can delete the old temporary layer.

## Tips:

- When telling QGIS which projection system to use when plotting coordinates for map viewing purposes, either the WGS84 system (EPSG: 4978) or the NA8 83 should work well in the United States. If your team has a local projection of choice, consider using that one.
- Make sure your latitude and longitude data are in decimal degrees. If they are in minutes and seconds, you will need to convert them first, before adding the CSV to QGIS.

**Troubleshooting**: If shapefiles are not mapping as expected, check to make sure projections match:

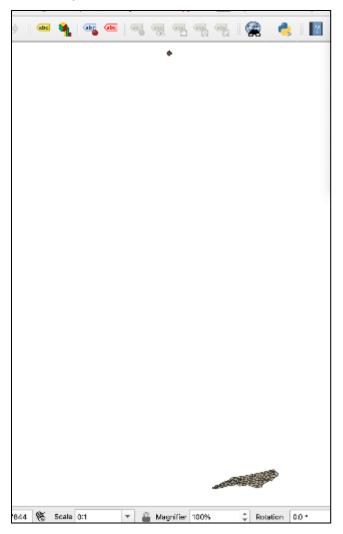
1. Right-click on each layer and select Layer Properties. The Coordinate Reference System (CRS) will specify what projection is being used. Here, it is NAD83 / North Carolina in feet, or EPSG code 2264.

Layer Properties showing the CRS: NAD83 / North Carolina in feet, or EPSG code 2264

		Layer Properties - Counties   Information	
Q	Information	from provider	<b>^</b>
information	Name Path	Counties /Users/elle/Documents/Gillings 2019-20/Alcohol Density/00 Data/Example Data/	
🍇 Source 🐳 Symbology	Storage Comment	Counties/Counties.shp ESRI Shapefile	
(abc) Labels	Encoding Geometry	UTF-8 Polygon (MultiPolygon) ERSC: 2364 NAD 82 / North Carolina (ft//S) - Projected	
M Diagrams	CRS Extent	EPSG:2264 – NAD83 / North Carolina (ftUS) – Projected 406832.6283580511808395,34987.9502319842576981 : 3052480.0220920592546463,1043624.4158946424722672	
Y 3D View Fields	Unit Feature count	feet 100	

2. To reproject a layer to match another: Go to Vector on the Menu, Data Management Tools → Reproject Layer...Select the layer to be reprojected and the target projection you want to transform it to. Under Reprojected, click on "..." and select where to save the output layer. Click Run, then close. Note: the default is to create a temporary layer unless you specify a new layer name and save location.

*Example of projection error: state shapefile is visible at bottom, and points are clustered far North instead of overlaid* 



## Geocoding

We will use QGIS to geocode the selection of ABC Stores within Greensboro, North Carolina, that is shown in the next screenshot.

## 15 addresses that need to be geocoded

		ABC Store A	ddres	ses		
Store	Address	City	State	ZIP	Hours	Phone
Greensboro ABC Board 1	3919 Burlington Rd	Greensboro	NC	27405	Mon-Sat 9am-9pm	336-375-2998
Greensboro ABC Board 2	2731 Ring Road	Greensboro	NC	27405	Mon-Sat 9am-9pm	NULL
Greensboro ABC Board 3	2309 Fleming Road	Greensboro	NC	27410	Mon-Sat 9am-9pm	336-668-2430
Greensboro ABC Board 4	2417 Lawndale Drive	Greensboro	NC	27408	Mon-Sat 9am-9pm	336-500-8806
Greensboro ABC Board 5	4411 West Wendover Ave.	Greensboro	NC	27407	9:00 am - 9:00 pm	336-292-1718
Greensboro ABC Board 6	4548 US Hwy 220 N	Summerfield	NC	27358	Mon-Sat 9am-9pm	336-643-3910
Greensboro ABC Board 7	500 Americhase	Greensboro	NC	27410	Mo-Thu10-7;Fr-Sa10-8	336-841-5510
Greensboro ABC Board 8	115-B North Cedar St	Greensboro	NC	27401	Mon-Fri 9am-5pm; Mixed Beverage	336-274-6304
Greensboro ABC Board 9	3716-A Battleground Plaza	Greensboro	NC	27410	Mon- Sat 9am-9pm	336-288-2383
Greensboro ABC Board 10	1101 Rotherwood Rd	Greensboro	NC	27401	Mon-Sat 9am-9pm	336-333-951
Greensboro ABC Board 11	3100 E Market St	Greensboro	NC	27405	Mon-Sat 9am-9pm	336-274-5929
Greensboro ABC Board 12	3923 Gate City Blvd.	Greensboro	NC	27407	Mon-Sat 9am-9pm	336-292-8199
Greensboro ABC Board 13	4633 West Market Street	Greensboro	NC	27410	Mon-Sat 9am-9pm	336-292-821
Greensboro ABC Board 14	403 Pisgah Church Rd	Greensboro	NC	27405	Mon-Sat 9am-9pm	336-282-1456
Greensboro ABC Board 15	115-C North Cedar St	Greensboro	NC	27401	Mon-Sat 9am-9pm	336-272-0875

You will need to download the MMQGIS plugin for geocoding in QGIS. MMQGIS is a collection of QGIS vector layer operation plugins. It can be used to animate maps, combine CSV files, create buffers around selected points, export attributes to a CSV file, and geocode a CSV file. We will use the MMQGIS plugin to obtain the coordinates for a selection of addresses stored in a CSV file.

## Adding MMQGIS plugin to QGIS

	Plugins   All	(541)
ii 👔	Q mmqgis	⊗
Installed	🏠 mmqgis	mmqgis
Not installed		A collection of QGIS vector layer operation plugins
New		MMQGIS is a set of Python plugins for manipulating vector map layers in Quantum GIS: CSV input/output/join,
🔆 Settings		geocoding, geometry conversion, buffering, hub analysis, simplification, column modification, and simple animation. MMQGIS provides an alternative to the Processing toolbox, with verbose progress reporting, an intuitive user interface, direct shapefile/CSV-file access, and some additional capabilities missing from other plugin sets.
		☆☆☆☆☆ 337 rating vote(s), 874164 downloads
		Tags vector, sort, merge, animate, delete
		More info homepage bug tracker code repository
		Author Michael Minn
		Available version 2020.1.16
		<u>د</u>
		Upgrade All Install Plugin
	Help	Close

After downloading the plugin, go to the new MMQGIS menu, hover over "Geocode," then select "Geocode CSV with Web Service."

## Geocoding using MMQGIS

Database Web Mesh Processing	MMQGIS Window	Help 93% 🗩 Fri Jul 10 1:13 PM Elle Law
Untitled Project - QGIS	Animate Combine Create Geocode	C     Geocode CSV with Web Service
	Import / Export Modify Search / Select	<ul> <li>Geocode from Street Layer</li> <li>Reverse Geocode</li> </ul>
lugins Repository Maintenance		

When "Geocode with CSV with Web Service" is selected, a window titled "Web Service Geocode" will pop up. For the "Input CSV File (UTF-8)" field, click on the "..." button and navigate to where your CSV file containing the addresses is saved.

Then, as shown in the next screenshot, ensure the following:

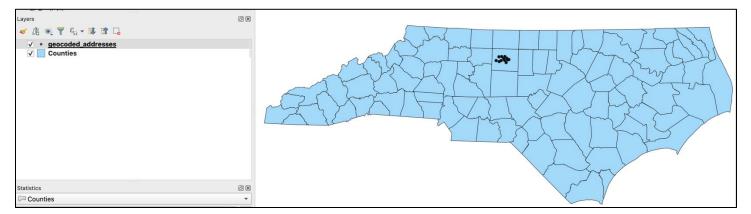
- For the Address field, select the correct column header from the CSV. Note that City, State, and Country fields are not required. When the window pops up, the column headers may already be automatically selected.
- For "Web Service" select "OpenStreetMap." You can also choose "Google," but if you lack a Google Maps API key, it may not work when searching for multiple addresses.
- For "Output Shapefile," select a valid location to save the resulting shapefile, which will contain the plotted coordinates that have been successfully geocoded from the addresses.
- For "Not Found Output List," select a valid location on your computer to save the resulting CSV. You will check this CSV to review what, if any, addresses were not able to be geocoded.

An example of settings for Web Service Geocode pop up. Note the bottom bar will display the number that geocoded successfully. In this case, you can see that 13 of the 15 addresses in the CSV geocoded successfully.

Web Service Geocode									
Input CSV File (UTF-8)									
0/Alcohol Density/00 Data/Example Data/Ge	ocoding/ABC Store Addresses.csv 🛛 📖								
Address	City								
Address	City -								
State	Country								
State -	(none) -								
Web Service									
OpenStreetMap / Nominatim	•								
API Key									
Duplicate Handling									
Use Only First Result	•								
Output File Name									
D/Alcohol Density/00 Data/Example Data/Ge	ocoding/geocoded_addresses.shp 🛛 📖								
Not Found Output List									
2019-20/Alcohol Density/00 Data/Example D	Data/Geocoding/not_geocoded.csv 🛚 🗔								
Geocode	d 13 of 15								
Apply	Close								

Click Apply to view the output shapefile in QGIS. Note that if you open the attribute table, you should see all the information present in the original CSV; QGIS will preserve any data attached to each address in the shapefile. It is helpful to have an existing map to ensure the coordinates that have plotted make sense, as shown.

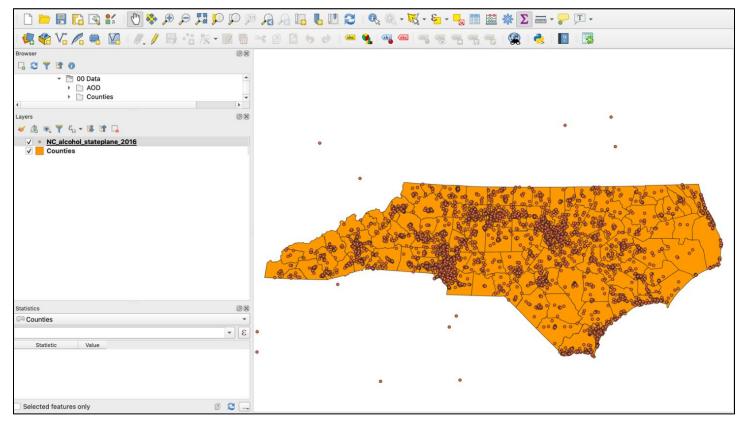
Overlaying the coordinates of the geocoded addresses with a map of North Carolina counties shows that they have plotted in as expected in Greensboro



## **Filtering to Active Licenses**

You can filter your data within QGIS, using the "Select by Attribute" feature. Here we have a map of North Carolina, along with a CSV of plotted alcohol outlets, each one represented by a dot.

## Map of North Carolina with Outlets Plotted from CSV File



When the attribute table of the alcohol outlets is opened (as in the next screenshot), notice it also contains outlets with non-active license types. Because we only want active licenses, we will need to subset the data to a new shapefile layer. To do this, you will write a simple expression to select for rows that only have "Active" for the license type column (the last column labeled "BBzStt\_D" in the next screenshot).

## Active / Closed Status of Plotted Outlets

			i 🗞 📕 🔼 🖣	5 Y 🎟 🌳 🕨							
_	Bsnss_I	trad_nm	Corp_Nm	address	City	State	zip	County	Latitud 🔻	Longitd	BzStt_D
	180571.0000	Hyatt Summe	VIII HII South	4920 South T	Charlotte	NC	28217	Mecklenburg	39.94760999	-75.1503100	Active
	213546.0000	Courtyard Ro	Hotel Rocky	250 Gateway	Rocky Mount	NC	27804	Nash	38.27105999	-85.4949799	Active
	142752.0000	Smokey Bone	Barbeque Int	3302 High Po	Greensboro	NC	27407	Guilford	37.71880999	-97.2622000	Active
	200699.0000	Fox and Hou	Last Call Ope	920 Town Ce	Wilmington	NC	28405	New Hanover	37.68531000	-97.2529499	Active
	166352.0000	Tazza Kitche	C C Stone Cr	600 Ledgest	Cary	NC	27519	Wake	37.64499000	-77.5759300	Active
	133998.0000	Pinehurst 7	Pinehurst LLC	Hwy 15-501	Pinehurst	NC	28374	Moore	37.52045590	-78.4948614	Active
	86200.00000	Cave Gentle	NULL	531 Richland	Jacksonville	NC	28540	Onslow	37.17693429	-77.4969002	Active
	105860.0000	Laurel Gray V	Laurel Gray V	5726 Old Hw	Hamptonville	NC	27020	Yadkin	37.16221670	-83.3733323.	Active
	110118.0000	Carolina Spor	NULL	12365 Hwy 4	Goldston	NC	27252	Chatham	36.63005619	-82.5478205	Revoked
	175214.0000	Buckaroos Gr	Buckaroo Gril	9531 NC Hwy	Piney Creek	NC	28663	Alleghany	36.56867600	-81.3064609	Cancelled
	135911.0000	Southland Re	Southland Tr	141 Caratoke	Moyock	NC	27958	Currituck	36.5486240	-76.1905617	Active
	75614.00000	Luckys	NULL	268 Gatewoo	Providence	NC	27315	Caswell	36.54074390	-79.3990579	Active
,	357.0000000	Warehouse	Three J's War	408 Greentre	Pelham	NC	27311	Caswell	36.54007579	-79.4701342	Active
	100642.0000	Aunt Millies	NULL	249 Broad St	Milton	NC	27305	Caswell	36.53866289	-79.2043778	Active
	198525.0000	Milton Tap an	NULL	116 Broad Str	Milton	NC	27305	Caswell	36.53863799	-79.2068289	Active
)	175600.0000	Cross Creek	Woltz Invest	1129 Greenhil	Mount Airy	NC	27030	Surry	36.53843500	-80.6196649	Active
	95273.00000	Food Lion 25	Food Lion LLC	50 Elam Road	Littleton	NC	27850	Warren	36.53780750	-77.9402908	Active
	28264.0000	New River Co	New River Co	611 Golf Cour	Sparta	NC	28675	Alleghany	36.5364569	-81.17221299	Active
;	132920.0000	Fresh Faces	Fresh Faces I	1196 River Ro	Henrico	NC	27842	Northampton	36.53417499	-77.8303379	Active
	168171 0000	Wildwood Fo	Wildwood Fo	139 Stanley R	Henrico	NC	27842	Warren	36 53344719	-77 9043575	Active

1. In the attribute table view, click on the leftmost yellow button, "Select Features Using Expression." The "Select by Expression" window will pop up.

## Select features tool example in QGIS

			$\bigcirc$		NC_alcol	hol_stateplane_2	016 :: Features T	otal: 9703, Filter	ed: 9703, Selecte	d: 1
1			E ) 🛛 🤻	, 🝸 🖀 🌺 🔎		= I 🗐 🍳				
	Bsnss_I	trad_n	ect feature	s using an expressi	on City	State	zip	County	Latitud	Longitd
1	147767.0000	Target Store	Target Corp	6350 Weddin	Wesley Chapel	NC	28104	Union	44.98743000	-93.271270
2	165400.0000	Stadium Foo	Stadium Foo	800 South Mi	Charlotte	NC	28202	Mecklenburg	42.8898300	-78.877120
3	206327.0000	Professional	Professional	408 Bellemea	Greensboro	NC	27401	Guilford	41.90059999	-87.624290
4	226057.0000	New Belgium	New Belgium	21 Craven Str	Asheville	NC	28806	Buncombe	40.5949800	-105.06802
5	180571.0000	Hyatt Summe	VIII HII South	4920 South T	Charlotte	NC	28217	Mecklenburg	39.94760999	-75.150310
6	213546.0000	Courtyard Ro	Hotel Rocky	250 Gateway	Rocky Mount	NC	27804	Nash	38.27105999	-85.494979
7	142752.0000	Smokey Bone	Barbeque Int	3302 High Po	Greensboro	NC	27407	Guilford	37.71880999	-97.262200
8	200699.0000	Fox and Hou	Last Call Ope	920 Town Ce	Wilmington	NC	28405	New Hanover	37.68531000	-97.252949
9	166352.0000	Tazza Kitche	C C Stone Cr	600 Ledgest	Cary	NC	27519	Wake	37.64499000	-77.575930
10	133998.0000	Pinehurst 7	Pinehurst LLC	Hwy 15-501	Pinehurst	NC	28374	Moore	37.52045590	-78.494861

- 2. In the top right pane, click on the drop-down for Fields and Values to display the license type column name.
- 3. Right-click on the field of interest and click All Unique to generate a list of all the distinct field values for that column in the right bottom panel.

● ○ ● Select b	y Expression - NC_alcohol_stateplane_2016	3
Expression Function Editor		
= + - / * ^    ( ) '\n'	Q Search     Show Value       > Aggregates     Image: Comparison of the second sec	group field
"BzStt_D" = 'Active'	<ul> <li>Aggregates</li> <li>Arrays</li> <li>Color</li> <li>Conditionals</li> <li>Conversions</li> <li>Date and Time</li> <li>Fields and Values</li> </ul>	Double-click to add field name to expression string. Right-Click on field name to open context menu sample value loading options.
	NULL	Notes
	1.2 Bsnss_I abc trad_nm abc Corp_Nm abc address abc City abc State abc Zip abc County 1.2 Latitud 1.2 Longitd	Loading field values from WFS layers isn't supported, before Values Search All Unique 10 Samples Active Cancelled Revoked Summary Suspension
Output preview: 1	abc BzStt_D abc min_dat abc mx_cnc_ abc mx_ssp_ abc Prm_T_D abc max_dat	Withdrawn
Help		Gelect Features ▼ Close

Selecting by expression window in QGIS

- 4. To write the expression, double click on the license type column name, "BBzStt\_D", type "=", then double-click on the column value to keep in the subset: "Active."
- 5. Click on Select Features. You may notice some rows become highlighted in the attribute table behind the pop-up box. In the attribute header, you should see the number after "Selected:" change to reflect the number of rows that qualify for your selection expression. Note that you can click on the drop-down button to the right of "Select Features," which allows you to further refine your selection.

*In the window header, note that of the total number of outlets in the attribute table (Features Total: 9703), the expression has selected 8490 "Active" outlets* 

	0 0					NC_alcoh	ol_statepl	ane_2016 :: Features Tota	al: 9703, Filtered	: 9703, Sele	ected: 8490			
/		0 - 0 0	8		. 🔻 🏦 🚳 🕻		-							
	Bsnss_I	trad_nm	(	000			Select	by Expression - NC_alcol	hol_stateplane_20	016			min_dat	mx_cnc_
1	147767.0000	Target Store	Targ	Expressi	on Function E	ditor							2008-12-15 1	NULL
2	165400.0000	Stadium Foo	Stad	= +	- / *	^    (	) '\n'	Q Search	Show Va	lues	a	oup field 📤	2010-09-16 1	NULL
3	206327.0000	Professional	Profe	"RzStt [	0" = 'Active'			Aggregates		-			2016-07-19 1	NULL
4	226057.0000	New Belgium	New					<ul> <li>Arrays</li> <li>Color</li> <li>Conditionals</li> </ul>		to e	uble-click to add 1 expression string. ht-Click on field r		2016-10-21 0	NULL
5	180571.0000	Hyatt Summe	VIII F					Conversions     Date and Time		ope	en context menu s ue loading options	ample	2012-08-21 1	
6	213546.0000	Courtyard Ro	Hote					<ul> <li>Fields and Values</li> </ul>				Notes	2016-09-21	NULL
7	142752.0000	Smokey Bone	Barb					NULL 1.2 Bsnss_I			ding field values f	from WFS	2008-03-19	NULL
8	200699.0000	Fox and Hou	Last					abc trad_nm abc Corp_Nm		lay	ers isn't supported	l, before 👻	2014-09-17 1	NULL
9	166352.0000	Tazza Kitche	сс					abc address abc City		Valu	es Q Search		2010-11-15 1	NULL
10	133998.0000	Pinehurst 7	Pine					abc State			All Unique	10 Samples	2007-05-06	NULL
11	86200.00000	Cave Gentle	NUL					abc zip abc County			celled		2001-01-02 1	NULL
12	105860.0000	Laurel Gray V	Laur					1.2 Latitud 1.2 Longitd		Sun	oked nmary Suspension		2003-10-27 1	NULL
13	110118.0000	Carolina Spor	NUL					abc BzStt_D		Wit	ndrawn		2006-04-30	NULL
14	175214.0000	Buckaroos Gr	Buck					abc min_dat					2011-08-29 1	NULL
15	135911.0000	Southland Re	Sout					abc mx_ssp_					2008-05-01	NULL
16	75614.00000	Luckys	NUL					abc Prm_T_D abc max_dat					2000-03-06	2016-04-30
17	357.0000000	Warehouse	Thre	Output p	review: 1			1.2 max_yer		¥			1998-03-09	2016-05-10
18	100642.0000	Aunt Millies	NUL	Help							Select Features	▼ <u>C</u> lose	2002-10-21 1	NULL
19	198525.0000	Milton Tap an	NULL	L	116 Broad Str	Milton	NC	27305	Caswell	36.53863	79 ⊱ Select Feat	ures	1	NULL
20	175600.0000	Cross Creek	Woltz	7 Invest	1129 Greenhil	Mount Airy	NC	27030	Surry	36 53843	Add to Curr	ent Selection	1	NIIII
4	how All Features											m Current Select	.ion	
	now An realtires		_	_							📮 🔁 Filter Curre	nt Selection		
					S 1									

- 7. Click Close. Notice the selected points also appear in the map view, as different colored points from the original color. To save the selection as a new file, right-click on the layer, go to Export, then select "Save Selected Features As...".
- 8. Click on the "..." button to the right of File name, navigate to the appropriate save folder, and enter the name of the new shapefile. Click Ok, and QGIS will automatically add the new layer for display.

**Tip:** Although it is possible to perform QGIS calculations with only selected features, it's good practice to save subsets as new layers to avoid confusion during analysis.

## **Counting Outlets per County**

Our goal is to add outlet count and population data to each county represented in the state shapefile. Once that data is present in the attribute table of the county shapefile, it can be used to calculate alcohol outlet density count indicators for each county.

First, we need to calculate outlet count per county. This can be done in two ways:

- Spatial Join. If you have two shapefiles, you can count the number of points from one shapefile within each county as represented by another shapefile.
- Tabular analysis. If you have categorical data for each outlet, you can count the number of outlets with the same category, such as the same county or census block. This method is demonstrated in the next paragraph and image.

Bring up the Toolbox window by clicking on the button with the image of a cog in the upper right-hand corner of the QGIS window. In the search bar at the top of the pane that appears, type in: "Statistics by Categories." In the resulting pop-up window, select the layer that contains a list of outlets with categorical data for each. Select the column name that contains the category of interest. Click Run.

#### Statistics by Categories Parameters Loa Statistics by categories Input vector layer This algorithm calculates statistics of - .... 🦻 fields depending on a parent class. " Outlets\_XY\_ArcMap [EPSG:4326] Field to calculate statistics on (if empty, only count is calculated) [optional] abc County Field(s) with categories 1 elements selected ... Statistics by category ... [Create temporary layer] ✓ Open output file after running algorithm

Pop-up window when using "Statistics by categories"

0%		Cancel
Help Run as Batch Process	Close	Run

For Output of Statistics by categories, open the Attribute Table of the output layer: number of outlets per county will be listed under "count"

	County	count	unique	empty	filled	min	max	min_length	max_length	mean_length
v	Vilkes	42	1	0		Wilkes	Wilkes	6	max_rengtri 6	filear_length
-										
	lertford	17	1	0		Hertford	Hertford	8	8	8
3 ⊢	lyde	20	1	0	20	Hyde	Hyde	4	4	4
4 B	runswick	176	1	0	176	Brunswick	Brunswick	9	9	9
5 S	tanly	36	1	0	36	Stanly	Stanly	6	6	6
6 C	leveland	49	1	0	49	Cleveland	Cleveland	9	9	9
7 V	ance	26	1	0	26	Vance	Vance	5	5	5
B V	Varren	21	1	0	21	Warren	Warren	6	6	6
	olk	35	1	0		Polk	Polk	4	4	
1										4
	erson	26	1	0		Person	Person	6	6	6
11 T	yrrell	5	1	0	5	Tyrrell	Tyrrell	7	7	7
12 0	Franville	29	1	0	29	Granville	Granville	9	9	9
13 R	lobeson	35	1	0	35	Robeson	Robeson	7	7	7
14 N	lorthampton	10	1	0	10	Northampton	Northampton	11	11	11
15 G	ates	3	1	0	3	Gates	Gates	5	5	5
16 A	inson	7	1	0	7	Anson	Anson	5	5	5
	erquimans	8	1	0		Perquimans	Perquimans	10	10	10
										10
18 A	she	21	1	0	21	Ashe	Ashe	4	4	4
19 S	wain	16	1	0	16	Swain	Swain	5	5	5
20 P	amlico	17	1	0	17	Pamlico	Pamlico	7	7	7

## Joining by Common Location Attribute (County)

Before joining data, review the tables to be joined. Check that the columns to be joined are recoded properly, named consistently, and expected to match across tables.

- To open the attribute table for a vector layer, activate the layer by clicking on it in the Layers Panel. Right-click on the layers you will be joining and select "Open Attribute Table." Make note of the column names that contains the geographic information you will be joining data by (e.g., "County Name").
- 2. To join two layers, right-click on the layer that you want data to be added to, select Properties, then navigate to the Joins page.
- 3. Click on the green plus (+) button on the bottom left. The pop-up window for "Add Vector Join" will appear.
- 4. Select the name of the layer to be added for "Join Layer."
- 5. Select the names of the columns with overlapping data for "Join Field" and "Target Field."
- 6. Select Ok to close the pop-up window. The new join will appear in the Layer Properties window. Click Ok to apply the join.
- 7. Verify the join performed as intended by opening the attribute table. The data should appear in new columns on the far right.
- 8. The join is temporary, so make sure to re-save as a new layer, or your work will be lost if you close and re-open QGIS.

Joining CSV (number of outlets per county) to shapefile of counties, using common location attribute of county



## Creating a Choropleth Map of Outlets per County

Now that a measure (outlets per county) is joined to a shapefile, you can create your first visual representation of outlet density: a choropleth map. To do this, use the Symbology pane of the shapefile:

- Right-click on the county shapefile and navigate to the Symbology pane.
- Click on the field at the very top of the window, which will have the default "Single Symbol" selected, meaning all the features have the same symbol applied to them. The three most used options are:
- Categorized: Allows you to style the layer using a categorical attribute field.
- Graduated: Allows you to classify the data by a numeric field attribute into discrete categories.
- Rule-based: Creates styling based on custom rules, based on SQL expressions.

- Select "Graduated," then for "Value," select the column name containing outlet counts per county.
- Change the color ramp if needed to make the map easier to read and understand.
- To populate a range of values for your selected attribute field, click on Classify. Notice how it populates the Classes pane. In this case, each color represents a different range of values. QGIS will apply a different color to each unique value in the field, based on the Mode selected. For more nuance, you can also change the number of Classes in the bottom right field.
- Click Ok. The window will close, and your map will display the updated symbol choices.

•	0. 🔴		Layer Properties	- Counties   Symbology		
9		불 Graduated				•
i	Information	Value	123 Outlets by County_	count		3 -
<b>3</b>	Source	Symbol				
*	Symbology	Legend format Color ramp	%1 - %2		Precision 4	↓ Trim
abc	Labels		istogram			
۹.	Diagrams	Symbol 🔻 Value	25	Legend		
<b>?</b>	3D View		00000 - 15.800000 00000 - 27.600000	3 - 16 16 - 28		
	Fields	✓ 50.4	00000 - 50.400000 00000 - 118.600000	28 - 50 50 - 119		ī
00	Attributes Form	✓ 118.	600000 - 1277.000000	119 - 1277		7
•	Joins					
	Auxiliary Storage					
<b>O</b>	Actions					
9	Display					
*	Rendering					
3	Variables					
2	Metadata		al Count (Quantile) 👻		Classes	
	Dependencies	Classify <ul> <li>Link class box</li> </ul>	Delete All			Advanced -
<b>8</b> -	Legend	Layer Rend				
	QGIS Server	Help	Style - Apply		Cancel	ОК

### *Classifying map symbol cut points and colors*

**Tip**: If you choose to create a new map, remember to click Classify again, or you will be using the value ranges from the previous map.

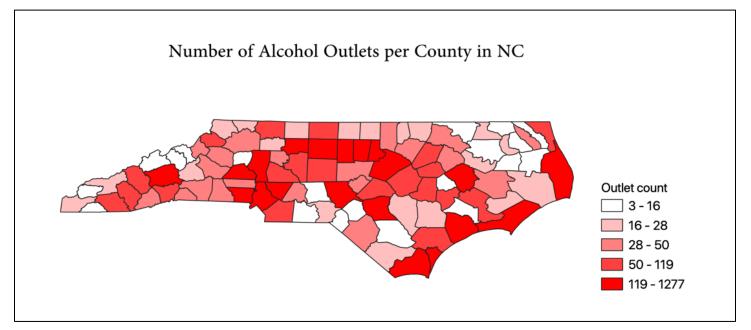
The map view should now display the counties, with a color gradient from white to red, darkening for counties with higher outlet counts. To export this map, copy it to your clipboard, save the image to your computer, or use the Print Layout mode for additional styling:

- To use the Print Layout mode, click on Project in the QGIS Menu bar, then click Layout Manager.
- Click the Add button to create and name a new layout. Click Ok, then Show.
- Click on the "Add New Map" button. Click and drag a box within the blank window, which will cause the map to appear. Position the map by clicking and dragging, then resize as needed using the boxes in the corners.
- To add a legend, click the "Add Legend" button, then click and drag a box over the location where it should appear.

**Tip:** When zooming in or out, click the refresh button to ensure the map is at the correct resolution when you export.

For additional guidance on how to create and export stylized QGIS maps with titles and legends, there are tutorials online at docs.qgis.org.

Map of number of alcohol outlets per county in North Carolina



# Part 2: Calculating Indicators of Alcohol Outlet Density

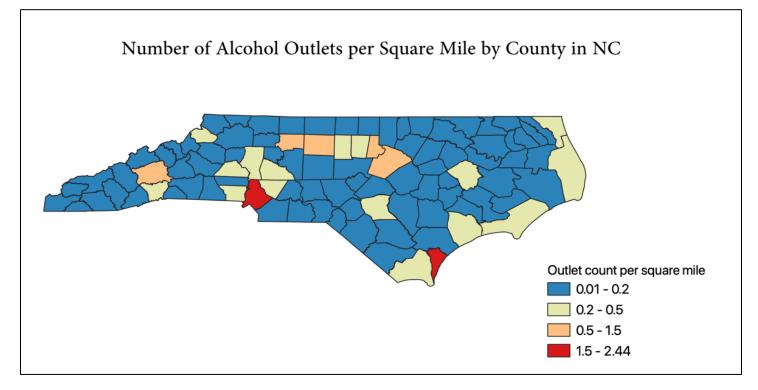
## **Calculating Count-Based Indicators**

Now that outlet count and population have been added for each county, you can begin to calculate your first set of indicators: calculate outlets per square mile and per 10,000 people. This will be done using the Field Calculator in the attribute table. See Step 6 in the *Measuring Alcohol Outlet Density Toolkit for more details on communicating count-based results.* 

## **Counting Alcohol Outlets per Square Land Mile**

- Open the Attribute table and click on the Field Calculator icon, represented by an abacus. The Field Calculator window will pop up.
- Make sure "Create a new field" is checked. Type the new column name into the "Output field name." Select decimal as your Output field type. Make sure the Output field length is long enough for the number of digits you want in your output numbers.
- In the Expression field, type in the formula for outlets per square mile in each county: "Outlets by County\_count"/"County Area." Click Ok. The new column will appear to the far right.
- Create a map by right-clicking on the layer, selecting Properties, and navigating to the Symbology pane. Follow the same steps as for the previous map, making sure to click the Classify button to reset the value ranges for the new indicator being mapped.

## Map of Number of Alcohol Outlets per Square Mile by County in North Carolina



**Tip:** If the Field Calculator returns a new column with blank or NULL values, check that you are not missing data in the fields used for calculation and that those field types are not a character field.

**Troubleshooting:** If a field does not appear for selection within the Symbology pane, make sure you have saved changes from the Field Calculator calculations in the attribute table.

## Counting Alcohol Outlets per 10,000 people

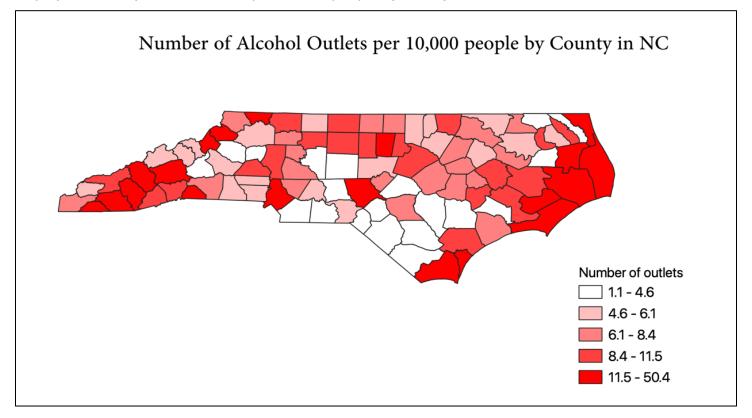
- 1. Open the Attribute table and click on the Field Calculator icon, represented by an abacus. The Field Calculator window will pop up.
- 2. Make sure "Create a new field" is checked. Type the new column name into the "Output field name." Select the appropriate Output field type, which will likely be decimal. Make sure the Output field length is long enough for the number of digits in your output numbers.
- 3. In the Expression field, type in the formula for your desired output. In this case, we want population per 10,000 in each county. The formula is: "Outlets by County\_count"\*10000 / "County Population". Click Ok. The new column will appear to the far right.

Using the Field Calculator to calculate outlets per 10,000 people in a county

					Counties :: Features Total: 100, Filtered: 100, Selected: 0			
/ 1	1 🗟 😂 🖷 🖷	× 6 🖸	ک 📒 ک	-	■ ▲ 「 ) ■ ■ ■ □ ▲ Field Calculator			
	OBJECTID	FIPS	CountyName	U		NAME	ShapeSTAre	ShapeSTLen
2	89	189	Watauga	WAT	Only update 0 selected features	Natauga	8735203026	444608.4927
3	38	49	Craven	CRA	✓ Create a new field Update existing field	Craven	21545383712	968408.423
1	52	89	Henderson	HEN	Create virtual field	Henderson	1045732200	599031.9450
5	45	123	Montgomery	MON	Output field name O per 10k	Montgomery	13971606100	597819.7041
6	7	183	Wake	WAK	Output field type Decimal number (real)	Nake	2388831046	702026.2737
7	62	159	Rowan	ROV	Output field length 10 ¢ Precision 3 ¢	Rowan	14603197216	570352.1367
3	27	101 .	Johnston	JOH	Expression Function Editor	Johnston	22177432359	673516.0998
•	47	71 (	Gaston	GAS	= + - / • ^    ( ) \n' Q Search Show Values	Gaston	10136315501.	496634.403
10	41	119	Mecklenburg	MEC	"Outlets by County_count" * 10000 /	vlecklenburg	15202529778	635613.2520
11	72	115	Madison	MAE	Population" 123 OBJECTID	Madison	12584845172	660215.9126
12	35	165	Scotland	sco	123 FIPS abc CountyName	Scotland	8936543764.	510787.10957
13	16	125	Moore	мос	abc UpperCount abc SapCountyl	Moore	1966795636	762525.0379
4	6	25	Cabarrus	CAB	123 DOTDistric	Cabarrus	10143637583	438825.3883
15	51	43	Clay	CLA'	123 DOTDivisio 123 SAP_CNTY_N	Clay	6153362609	449623.0713
16	19	147		PITT	123 CNTY_NBR	Pitt	18271650860	750034.4982
17	20	79	Greene	GRE	123 DSTRCT_NBR           Output preview: 2.871637790753326           123 DIV_NBR	Greene	7448298694.	415266.6631
18	33		Carteret	CAR		Carteret	2896656649	1003010.366
19	82		Chowan	СНС	You are editing information on this layer but the layer is currently not in edit mode. If you click OK, edit mode	Chowan		438858.6511
20	59		Swain	SWA	will automatically be turned on.	Swain		851283.1180
20	31		Onelow	ONS		Doslow	22747152410	703791 9334
(	w All Features				Help Cancel OK	J	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	

Resulting output from the Field Calculator, added as the right-most column

1. To generate a map displaying the indicator results. If you wish to generate maps displaying the indicator results, right-click on the layer and select Properties. In the left side bar, select Symbology. Choose Graduated in the topmost field, then select the column name that contains the indicator you want visualized. Click on Ok to add the new Symbology setting to your map.



**Tip:** In the Search bar, you can type in the column name of interest. When it appears, doubleclick to autogenerate it within the Expression field.

## **Calculating Distance-Based Indicators**

Measuring distance-based indicators involves calculating a distance matrix, the results of which can then be manipulated to provide measurements at the desired area unit.

## Calculating the Average Distance from a Person to Their Nearest Alcohol Outlet (Personto-Outlet)

The average distance from a person to their nearest alcohol outlet will be measured using the data in the bulleted list. In this demonstration, the area unit is the county level, and the sub-area units are block groups:

- Block group shapefile for North Carolina.
- The population from 2018 ACS estimates for each block group in North Carolina.
- Total block group population for each block groups in each county.
- Centroid of each block group in North Carolina.
- Distance matrix of distance measured between centroids and nearest outlet.

To calculate the average distance from person to outlet, you will need centroids: a point layer where each point represents the center of a shape. From each block group shape's centroid, the distance to the nearest outlet (regardless of block group or county boundaries) will be measured. That distance will then be weighted by population. Once all the population-weighted block group distances are calculated, they will be averaged by county. The formula used is: (block group population / total county population) \* distance. The resulting calculation from each block group is then summed together with other block groups of that county to provide the average distance from person to outlet.

## **Gather Block Group and Population Data**

First, acquire a block groups shapefile with population data for each block group. If you cannot find a pre-linked shapefile or geodatabase on the US Census website, you may need to find and then join population data based on block group IDs. For this demonstration, there is a pre-linked geodatabase available with 5-year American Community Survey estimates for 2018–2019 at the block group level in North Carolina at the following link: www.census.gov/geographies/mapping-files/time-series/geo/tiger-data.html.

Navigate to the following:

- 1. Click on "Download these files from the FTP archive."
- 2. Select the desired year: 2018.
- 3. On the resulting page, scroll down until you see "ACS\_2018\_5YR\_BG\_37.gdb.zip." Click on it to begin the download, unzip it, and save it to your QGIS data folder. Note that "BG" stands for block group and "37" is the Federal Information Processing Standards (FIPS) code for North Carolina.
- You can drag and drop the saved geodatabase file to the layers panel in QGIS, which will open the box in the next screenshot. Select the layers you wish to import: "ACS\_2018\_5YR\_BG\_37\_NORTH\_CAROLINA," "BG\_METAFATA\_2018," and "X01\_AGE\_AND\_SEX." Then click Ok.

Choose which layers to add from the geodatabase

	Select Vector Layers to	Add							
Layer ID	Layer name		Number of features	Geometr 📤					
33	ACS_2018_5YR_BG_37_NORTH_CAROLINA		6155	MultiPoly					
32	BG_METADATA_2018		53992	None					
0	X00_COUNTS		6155	None					
10	X01_AGE_AND_SEX		6155	None					
1	X02_RACE		6155	None					
2	X03_HISPANIC_OR_LATINO_ORIGIN		6155	None					
11	X04_ANCESTRY		6155	None					
12	X05_FOREIGN_BORN_CITIZENSHIP		6155	None					
13	X06_PLACE_OF_BIRTH		6155	None					
14	X07 MIGRATION		6155	None 🔻					
Select A	14     X07 MIGRATION     6155     None       Image: Select All     Add layers to a group     Cancel     OK								

Once the layers have imported, view their attribute tables to familiarize yourself with their contents. The metadata layer will help you identify which column has your population, in this case, "B01001e1."

Using the metadata attribute table, identify which column name contains population data of interest

<b>—</b>		BG_METADATA_2018 :: Features Total: 53992, Filtered: 53992, Selected: T
🌌 📑 🔁 I 🖻 I	ه 🛛 🗧 ک	🝸 🍱 💠 🔎   🏭 🎆 📓   🚍   📾 🍭
OBJECTID	Short_Name	Full_Name
94368	B00001e1	UNWEIGHTED SAMPLE COUNT OF THE POPULATION: Total: Total Population (Estimate)
94369	B00001m1	UNWEIGHTED SAMPLE COUNT OF THE POPULATION: Total: Total Population (Margin of Error)
94370	B00002e1	UNWEIGHTED SAMPLE HOUSING UNITS: Total: Housing Units (Estimate)
94371	B00002m1	UNWEIGHTED SAMPLE HOUSING UNITS: Total: Housing Units (Margin of Error)
94372	B01001e1	SEX BY AGE: Total: Total Population (Estimate)
94373	B01001m1	SEX BY AGE: Total: Total Population (Margin of Error)
94374	B01001e2	SEX BY AGE: Male: Total Population (Estimate)

## **Create Block Group Population CSV Layer**

Right-click on the "X01\_AGE\_AND\_SEX" layer, then Export, and "Save Features As" to create a new CSV with only your target population data, as shown in the next screenshot. Click "Deselect all" then select "OBJECTID," "GEOID," and the population per block group: "B01001e1." Name the new layer "population table" and click Ok. Remove the "X01\_AGE\_AND\_SEX" layer from QGIS.

*Create a new CSV with only the target population data, saving only the first three columns from the "XO1\_AGE\_AND\_SEX" layer* 

0.		S	Save Vector Layer as	
Format	Comm	a Separated Val	lue [CSV]	Ŧ
File name	ensity/	00 Data/Exampl	le Data/block groups 2019/population table.csv	፼
Layer name				
CRS	invalid	projection		-
Save only	y selecte	ed features		
▼ Select fi	ields to	export and the	ir export options	
Nam	ne	Туре		
✓ OBJEC	CTID	Integer64		-
✓ GEOID	)	String		
✓ B0100	)1e1	Real		
B0100	)1m1	Real		
B0100	)1e2	Real		
B0100	)1m2	Real		-
	S	elect All	Deselect All	
▼ Geomet	ry			
Geometry	type		Automatic	•
Force n	nulti-typ	be		
🗌 Include	e z-dime	insion		
Hel	lp		✓ Add saved file to map Cancel	ОК

Check whether the "GEOID" column in your census sourced population file begins with "15000US," which refers to the summary level of the FIPS code (in this case "State-County-Census Tract-Block Group"). This addition will prevent joining this dataset to other datasets that contain only the FIPS code. Next, create a new column with only the 12-digit FIPS code. Use the following code to populate a new field: "fips\_BG", taking only the rightmost 12 digits from "GEOID": right("GEOID",12). This FIPS code will be used to join the population CSV to other datasets.

Creating a new field using the Field Calculator, extracting the 12-digit FIPS code from the GEOID

		Field Calculator
1 🖥 🗝 🖻 🖻 🗧 🔽 🗖 🕻	Only update 0 selected features	
3	✓ Create a new field	Update existing field
GEOID           15000US371830536101           15000US371830536091           15000US371830536061           15000US371830536062           15000US371830536062           15000US371830536061           15000US371830536061           15000US371830536061           15000US371830536021           15000US371830536031           15000US371830536031           15000US371830537151           15000US371830537151           15000US371830537161           15000US371830537161           15000US371830537152           15000US371830537132	Output field length     12     Precision       Expression     Function Editor       = + - / * ^    ( ) '\n'       right( "GEOID" ,12)	seo Show Values abc GEOID ecometry collect_geometries geom_from_gml geom_from_wkt geometry geometry geometry geometry geometry geometry num_geometries
15000US371830537131	Output preview: '370010201001'	
15000US371830537142		
15000US371830537141	Help	Cancel OK

## **Create Centroids**

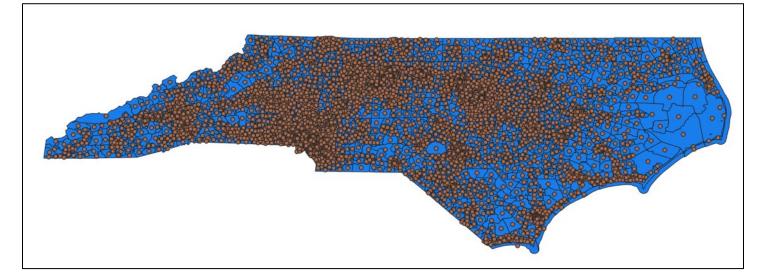
To create the centroid point layer, go to the Menu and select: Vector  $\rightarrow$  Geometry Tools  $\rightarrow$  Centroids. The window in the next screenshot should appear. Select the block groups shapefile as the Input layer and click Run.

## The pop-up window for the Centroids creation tool

	Centroids
Parameters       Log         Input layer	Centroids This algorithm creates a new point layer with points representing the centroid of the geometries in an input layer. The attributes associated to each point in the output layer are the same ones associated to the original features.
C	6 Cancel
Help Run as Batch Process	Close Run

Take a moment to visually explore the resulting centroids, overlaid with the block groups shapefile as shown.

Block groups of North Carolina, with points representing their centroids



**Troubleshooting:** If an attempt to calculate centroids returns errors, see if they can be resolved by applying a small buffer to the shapefile. This will smooth out any issues caused by how the boundaries are drawn, and it won't impact the accuracy of results due to the small scale. (Input Distance: enter 0.001 degrees if data has not yet been projected, otherwise use 1 foot/meter.)

**Alternate Analysis:** Click on the settings button (shaped like a cog) to bring up Processing Toolbox. Search for "Point on Surface." If you are working at the county level, using the "Centroid" tool for block groups is sufficient. However, if you are working at a small scale, consider whether it is worth using "Point on Surface" instead, which adds the additional constraint that the centroid point must be placed inside the boundaries.

**Tip:** After using the Centroids tool, you can check to see for points that do not spatially join to their block group (e.g., plotted out on water). This is more likely to occur with strangely shaped area units. Conduct a spatial join, then use "Select Features by Value" to select for any points that have a different location FIPS code from their own.

## **Build Distance Matrix**

Next, calculate the distance between each centroid and its closest outlet, using the **distance matrix** tool. In the menu bar, select Vector, then Analysis Tools, and click on distance matrix. Once selections are entered as shown in the bulleted list, click Run, then click Close.

- For the input point layer, select the newly created centroids point layer, and "GEOID" for the "Input unique ID field."
- For the "Target point layer," select the points layer containing the outlets. Make sure your outlets layer contains a unique ID field for each outlet, as that is the field you will select for "Target unique ID field."
- Make sure "Linear distance matrix" is selected for "Output matrix type" and that "1" is input for how many nearest target points to use, since you are calculating just the distance to the single closest point.

## Distance matrix analysis tool

Parameters Log	Distance matrix
Input point layer * * Centroids [EPSG:2264] Selected features only Input unique ID field ** GEOID * Target point layer * * Active_Outlets [EPSG:2264] Selected features only Target unique ID field 1:2 Bsnss_I Output matrix type Linear (N*k x 3) distance matrix Use only the nearest (k) target points 1 Create temporary layer] v Open output file after running algorithm	This algorithm creates a table containing a distance matrix, with distances between all the points in a points layer.
0% Help Run as Batch Process	Close Run

An output layer, "Distance Matrix" will appear. Open the attribute table and notice that the "InputID" contains the full 12-digit string FIPS code. "Distance" contains the distances from each block group centroid to nearest outlet, regardless of what block group or county the outlet is located in. You can examine this by visually exploring the layer.

## Distance matrix layer

•	Distance matrix :: Features Total: 6155, Filtered: 6155, Selected: 0									
Ø		🖬 🔀 🖻	ه 🛯 🔚 ک		<b>t</b>	<b>P</b>   <b>1</b>	×	>:		
	InputID	TargetID	Distance	_				•		
1	371190050002	110924.0000	338.8464478					_		
2	371190050001	110924.0000	601.0501372							
3	371190032011	214855.0000	791.8605507							
4	371190031082	220795.0000	675.7561199							
5	371190047001	209608.000	513.8875785							
6	371190046002	152319.0000	1194.325831							
7	371190031052	196776.0000	765.3523503							
8	371190030161	85894.0000	1037.5191008							
9	371190031091	185136.0000	728.6651609							
10	371190031061	112133.0000	312.4520627							
11	370939702022	121808.0000	7645.367785							
12	370710323022	205031.0000	382.1422675							
10	2700207040	121202 0000	10046 51402							

Each row will need the FIPS code for just the county, for when you later join total population data of each county to each distance. To create this, extract the county code from the 12-digit string in the "InputID" column, populating it into a new column to be named "fips\_county." Because the FIPS county code is the last 5 digits of the 12-digit string, we can use the following code (as shown in the next screenshot): left("InputID",5). Make sure the Output field type is "Text (string)" and the output field length is 5.

#### Extracting county FIPS code from detailed FIPS string

0	00	distance ma	atrix :: Features To	tal: 6157, Filtered: 6157, Selected: 1
	<b>7</b> 🗟 🕄 🗃	💼 🛰 🗈 🖪	ها 🛯 📄 کې	🍸 🖀 🗞 💭 I 🏪 🏙 I 🚍 I 📾 🍭
abc	nputID 👻 = 8			▼ Update AI Update Selected
	InputID	TargetID	Distance	Field Calculator
1	NULL	NULL	NULL	Only update 1 selected features
2	NULL	NULL	NULL	✓ Create a new field Update existing field
3	371619601004	125357.0000	6340.477405	Create virtual field
4	371619605006	179361.0000	1400.029582	Output field name fips_cnty
5	371619606002	130845.0000	2408.92666	Output field type Text (string)
6	371619610001	96201.00000	11091.20383	Output field length 5 ¢ Precision 3 ¢
7	371619610002	221613.0000	7168.406452	Expression Function Editor
8	371619607001	173884.0000	4340.762460	Show Values
9	371619601003	173884.0000	9853.682304	= + - / * ^ II ( ) \n group field
10	371619608001	152236.0000	796.4689861	left     "InputID" ,5     > Aggregates     Double-click to add field name       > Arrays     to expression string.
11	371619607003	173884.0000	1632.007911	Color     Right-Click on field name to     Open context menu sample
12	371619601001	181047.0000	8789.012055	Conversions     Date and Time
13	371619602004	46892.0000	9598.808703	✓ Fields and Values Values Q Search
14	371619602003	46892.0000	8029.674972	NULL     All Unique     10 Samples
15	371619604001	179361.0000	4488.950072	1.2 TargetID 1.2 Distance
16	371619605002	46892.0000	2623.804201	123 cnty_fips Files and Paths
17	371619607002	221613.0000	1293.489660	► Fuzzy Matching
18	371619605004	215457.0000	604.1437164	General     Geometry
19	371619605007	213424.0000	2011.355239	Output preview: NULL Map Layers
<b>T</b> s	how All Features			Help Cancel OK

**Alternate Analysis:** You can also find the nearest neighbor point on another layer, such as school locations. See <u>QGIS Nearest Neighbor Analysis</u> for more details.

#### Add Block Group Data

Now join block group population data to the distance matrix. Right-click on the distance matrix layer and navigate to Joins. Select "population table" as the Join layer and select the column containing the 12-digit FIPS code in each layer as the join and target fields, as shown in the next screenshot.

Favorites • Spatial Bookmarks ... Layer Properties - distance matrix | Joins A Home Q Setting Value 1 1 . Add Vector Join 👔 Information ayers Join layer population table \* 💉 🥼 🖲 🍸 🗞 🗕 🚺 🍃 🗞 Source Centroids Join field abc fips\_BG \* V 🥖 distance matrix 🧭 Symbology Target field abc InputID Ŧ Active\_Outlets population table abc Labels ✓ Cache join layer in virtual memory X01\_AGE\_AND\_SEX BG\_METADATA\_2018 🖌 Diagrams ACS\_2018\_5YR\_BG\_37\_NOR Dynamic form X00\_COUNTS 3D View V 2019\_bg\_EPSG\_2264 Editable join layer Fields Joined Fields Custom Field Name Prefix Attributes Form Joins auxiliary Storage Distance 🚺 Display Statistics Active\_Outlets Rendering Variables Statistic Value Cancel OK Metadata 1 🔁 Dependencies Selected features only

Join the population table to the distance matrix, using the 12-digit FIPS code columns

Open the attribute table to ensure the join occurs as expected. Both distance matrix and population table columns should appear together, as shown in the next screenshot. Notice QGIS added the layer name before the names of the added columns.

11	/ 🗃 😂 📆 🖷 🎽	< B 🖪 🗧 🗧			22 -	0		
						. Mark		
ibc	nputID = E				le vers av			✓ Update
1	InputID	TargetID	Distance	cnty_fips	fips_cnty	population table_OBJECTID	population table_GEOID	population table_B01001e1
	371999602004		2373.7111516		37199	6150	15000US371999602004	1643
-	371999604001	92617.00000	11302.93028	37199	37199	6153	15000US371999604001	1695
	371999602003	192590.0000	2623.230665	37199	37199	6149	15000US371999602003	1281
	371999603001	80467.00000	840.4929226	37199	37199	6151	15000US371999603001	1721
	371999604003	80467.00000	5436.166292	37199	37199	6155	15000US371999604003	1404
	371999601021	195281.0000	5734.50489	37199	37199	6145	15000US371999601021	858
	371999604002	92617.00000	6168.144599	37199	37199	6154	15000US371999604002	1126
	371999602001	192590.0000	9128.469953	37199	37199	6147	15000US371999602001	748
	371999601011	211902.0000	2880.795809	37199	37199	6143	15000US371999601011	1655
)	371999603002	80467.00000	7878.253315	37199	37199	6152	15000US371999603002	1459
	371999602002	192590.0000	4886.220751	37199	37199	6148	15000US371999602002	862
2	371999601012	211902.0000	6146.495186	37199	37199	6144	15000US371999601012	2185
3	371999601022	195281.0000	4371.253138	37199	37199	6146	15000US371999601022	1030
	371970504001	156081.0000	6686.959610	37197	37197	6130	15000US371970504001	1194
5	371970504004	156081.0000	4146.151939	37197	37197	6133	15000US371970504004	2048
5	371970504003	147090.0000	1072.960703	37197	37197	6132	15000US371970504003	1325
,	371970503002	202725.0000	2891.219411	37197	37197	6125	15000US371970503002	876
3	371970502002	179055.0000	4918.171365	37197	37197	6122	15000US371970502002	1794
,	371970501023	149085.0000	5281.1641431	37197	37197	6120	15000US371970501023	2048

#### *Convert text columns to numeric columns for calculations*

Verify that your population column is numeric rather than in text format. Text formats cannot be used in calculations. To check, click on any cell in the column, then hover over the column field name. A label will pop up, specifying the type. If it is text type, you will need to use the Field Calculator to generate a new column with output field type of "whole number," as shown in the next screenshot. Name the new block group population column "BG\_pop" and set it equal to the old column field name, making sure the output field type is whole number and the field length is appropriate for your population values. An example of this in the Field Calculator is demonstrated in the next screenshot.

• •	Field Calcu	llator			
<ul> <li>Only update 0 selected features</li> <li>✓ Create a new field</li> <li>Output field name BG_pop</li> <li>Output field type Whole number (integer)</li> <li>Output field length 10 ♀ Precision 3</li> <li>Expression Function Editor</li> <li>= + - / * ^    ( ) '\n'</li> <li>"population table_801001e1"</li> <li>"update 0 * 1928'</li> </ul>	<ul> <li>B01</li> <li>Fields and Val abc popula</li> <li>Recent (fieldo "populatio")</li> </ul>	ues ition table_B0	Double to exp Right- open o	gr e-click to add f ression string. Click on field r context menu s loading options Q Search All Unique	name to
Help				Cance	el OK

*Populating a new numeric column "BG\_pop" from the old text column "population table\_B01001e1"* 

Export the layer into a new shapefile to make the join permanent, naming it "distance matrix final." You will not be able to conduct calculations otherwise, as the added columns fields may not appear for selection when the Group Stats plugin is used. Right-click the layer, select Export, and go to "Save Vector Layer as." Use this opportunity to export only the fields needed. Two columns with the same FIPS code for county are not both needed, so one can be left unchecked.

			Save Vect	or Layer as			
Format	ESRI Shapefile						•
File name	20/Alcohol Densit	2019/distanc	e matrix final.sł	np 🛛 🗔			
Layer name							
CRS	EPSG:2264 - NA	D83 / North	Carolina (f	tUS)			•
Encoding				UTF-8			•
	y selected features ields to export an		ort options				
	Name	Туре	Replace wit	th displayed values	3		
✓ InputI	D	String					
✓ Target	tID	Real					
✓ Distan	ice	Real					
cnty_f	ips	Integer	🗌 Use Ra	nge			
✓ fips_ci	nty	String					
✔ BG_po	q	Integer	Use Ra	nge			•
	Select /	AII			Deselect	All	
Replac	e all selected raw f	ield values b	y displaye	d values			
Geomet	ry						
<ul> <li>Geomet</li> </ul>				Automatic			•
Geometry				Automatic			•
Geometry	type			Automatic			•
Geometry	<b>type</b> multi-type	1		Automatic			•
Geometry	<b>type</b> multi-type e z-dimension	_	030571.656				•
Geometry	type multi-type a z-dimension nt (current: layer)	_	)30571.656	69	ast 3049318	1.1556	<b>•</b>
Geometry Geometry Force r Ginclude	type multi-type a z-dimension nt (current: layer)	North 10	030571.656 9756.5396	69 E	ast 3049318	3.1556	
Geometry Geometry Force r Ginclude Circle Geometry Geomet	type multi-type a z-dimension nt (current: layer)	North 10 South 3		69 E		3.1556 vas Extent	

*Use "Save Vector Layer" as an opportunity to drop duplicate columns, such as "cnty\_fips"* 

#### **Calculate Total County Population**

As you can see in the attribute table in the next screenshot, we now have the FIPS code, distance, FIPS county code, and block group population. We still need the overall population of the county that each distance's centroid is located in. Use the block group populations to calculate overall county population, then return to the distance matrix and attach the overall county population to each distance, using matching county FIPS codes.

The block group populations "BG\_pop" will be summed by county to get overall county population

•	🔴 🔴 🔹 distance matrix final :								
/	1 B 2 B		ء 📃 들 4	7 🔳 💠 🖇	0   🖪 🖪 📓   🤅				
	InputID	TargetID	Distance	fips_cnty	BG_pop				
1	371190051001	74358.00000	706.8840392	37119	1063				
2	371190050002	110924.0000	338.8464478	37119	719				
3	371190046002	152319.0000	1194.325831	37119	1495				
4	371190032011	214855.0000	791.8605507	37119	1227				
5	371190048001	201839.0000	464.584484	37119	1385				
6	371190047001	209608.000	513.8875785	37119	1725				
7	371190031061	112133.0000	312.4520627	37119	3020				
8	371190031052	196776.0000	765.3523503	37119	2290				
9	371190031082	220795.0000	675.7561199	37119	2374				
10	371190031091	185136.0000	728.6651609	37119	2567				
11	370939703001	85333.00000	2169.069773	37093	653				
12	370939702022	121808.0000	7645.367785	37093	1496				
13	371470004004	205435.0000	487.7409838	37147	448				
14	3709397040	121808.0000	10046.51402	37093	2281				
15	370710323013	160226.0000	195.5926585	37071	951				
16	370710323012	210430.0000	1342.361949	37071	634				
17	370710323022	205031.0000	382.1422675	37071	937				
18	370710323021	130294.0000	1113.9747209	37071	1015				
19	370510032052	130299.0000	774.7203927	37051	2179				
20	370510032051	86052.0000	1223.613201	37051	2312				
<b>T</b> s	how All Features								

As mentioned in the *Measuring Alcohol Outlet Density Toolkit*, the Group Stats plugin can be used to summarize measurements of one column based on the category of another column. We can use this to average distances by their "category" of county: the county FIPS code column. If you have not already downloaded the plugin, do so now: In Menu view, select Plugins, then Manage and Install Plugins. In the search bar, type "Groups Stats," click on Install Plugin for the result, then click Close.

# Adding Group Stats plugin to QGIS

• • •	Plugins   All (	533)		
ii 🐴	् group s			۵
installed	🍅 Group Stats	Group St	ats	<b>7</b>
> Not installed		Stats and anal	ysis for vector layers	data
Install from ZIP			version by Faunalia. Sponsor RIKUS program in collaborati	
Settings		Regione Piemonte e Funzione Censiment Migrato alla versione Piemonte per il prog Regione Piemonte e Funzione Censiment	Dipartimento della Protezion o Danni e QGIS3 da Faunalia per Arpa ramma ERIKUS in collaboraz Dipartimento della Protezion	e Civile - a ione con e Civile -
		Tags	analysis, stats, pivot table	_
		More info	homepage bug tracker co repository	ode
		Author	Rajmund Szostok. Migrated	to QGIS 🛫
		Upgrade All	Install	Plugin
	Help			Close

Now when you select Vector on the QGIS menu bar, Groups Stats should appear as an option.

Open the Group Stats plugin. Select "distance matrix final" for "Layers" in the Control Panel. Drag "sum" and "BG\_pop" into the Value panel, and "fips\_cnty" into the Rows panel, then click Calculate. This will sum the block group populations by county. You may need to scroll down in the Fields pane if you don't see the field you want.

Group Stats 0 🗙 Control panel Layers fips\_cnty distance matrix final Ŧ Fields \* fips\_cnty InputID TargetID average count max median min stand.dev. sum unique variance Ŧ Filter Columns use NULL values Rows Value fips\_cnty sum BG\_pop Use only selected features Clear Calculate Calculate... 100% | generate view...100% | done.

Use Group Stats to sum populations of block groups by county to get total population of each county

Click on Data in the Menu bar and select "Save all to CSV file" to save the calculation to your computer as a CSV file: "Total County Pop." Open the CSV and name the column "county\_pop" before reimporting the CSV layer into QGIS.

*View the attribute table of "Total County Pop" to review column names and check the data is as expected.* 

	😑 🔵 Tot	al County Pop :: I	Features Total: 100, Filtered: 100, Selected: 0
	] 🗾 📑 😂 📑	6 × 6 C	) 🗧 📄 💊 🔽 🏘 🗭 🎼 🧱 🚍 🛛 x
abo	fips_cnty 🔹	= 8	▼ Update All Update Selected
	fips_cnty	County_pop	· · · · · · · · · · · · · · · · · · ·
1	37049	103082	
2	37029	10447	
3	37183	1046558	
4	37119	1054314	
5	37043	10813	
6	37005	10973	
7	37089	113625	
8	37073	11563	
9	37187	12156	
10	37191	124002	
11	37019	126860	
12	37137	12742	
13	37085	130361	
14	37155	133442	
15	37143	13459	
16	37159	139605	
17	37041	14205	
18	37173	14254	
19	37135	142938	<b>•</b>
7	Show All Features 🖵		

Now join the "Total County Pop" CSV layer to the "distance matrix final" layer, using the county FIPS column in both tables as the join and target fields.

	Layer Pr	operties - distance matrix fina	al   Joins	
Q Se	tting Value			
information		Add Vector Joi	in	
Source	Join layer	Total County	Рор	•
Y 💸 Symbology	Join field	abc fips_cnty		•
(abc) Labels	Target field	<sup>abc</sup> fips_cnty		<b>•</b>
Ciagrams		er in virtual memory		
3D View	Dynamic form			
Fields	Editable jo			
Attributes Form	Joined Fie     Custom F	elds Field <u>N</u> ame Prefix		
Joins				
Auxiliary Storage				
Actions				
🤛 Display				
Kendering				
🗧 Variables				
Metadata			Cancel	ок
Nependencies				
Egend				
QGIS Server	Help Style -	Apply		Cancel OK

Add total county population data to the "distance matrix final" layer

Now we can weight the measured distances of centroid to closest outlet by block group population, then sum them to receive an average county measurement. Using distance, county FIPS, block group population, and overall county population for each distance, calculate the population-weighted distance column "dist\_weight." Use the following formula, as shown in the next screenshot: ("BG\_pop"/"Total County\_pop")\*"Distance".

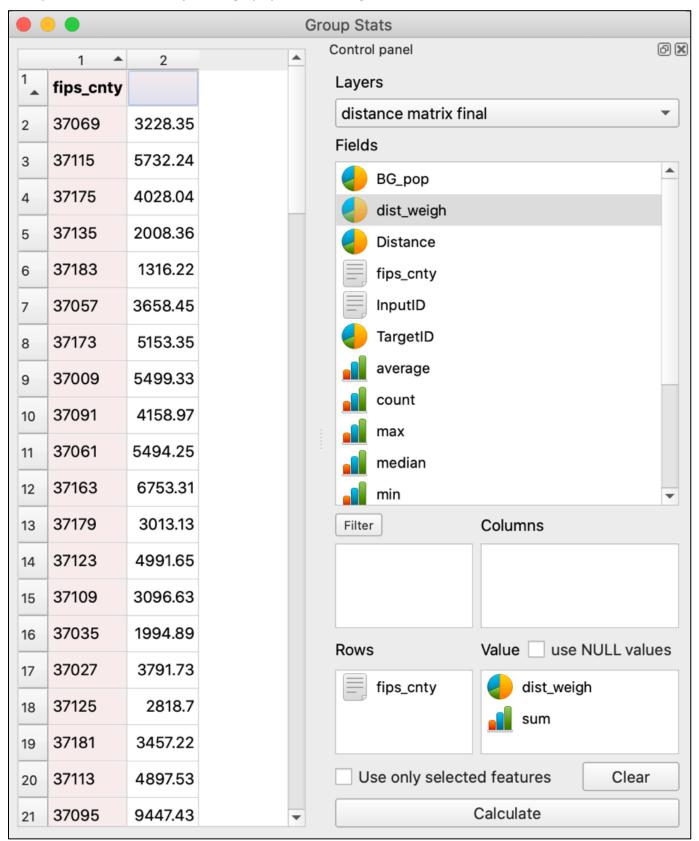
	•	3 =							▼ Update All	Jpdate Selecte
2	InputID 371779601	TargetID	Distance	fips_cnty	BG_pop	Total County	Pop_Cour	ity_pop		
	Contractor and the	• • •		F	ield Calculat	or				-
2	37177960'	Only update 0 se	lected features							
3	37177960	✓ Create a new fie	ld		Update	existing field				
•	37177960	Create virtual fie	eld							
5	371779601	Output field name	dist_weigh							
5	37095920	Output field type	Decimal numb	er (real) 🔹						
	37095920	Output field length	10 ‡ Pre	cision 3 🌲						
	37095990	Expression Fur	nction Editor							
	37095920			A Sea	arch	Show Values				
0	37095990	= + - /	* ^    (	) [./u.	number			g	roup field	
1	37095920	("BG_pop"/ "Tota Pop_County_pop"			regates			e-click to add ression string		
2	37075920:			Cold	or		Right-	Click on field	name to	
3	37075920:			▶ Con	ditionals versions		value l	oading option	s. 👻	
4	37075920				e and Time ds and Value	S	Values	Q Search		
5	37075920				s and Paths zy Matching			All Unique	10 Samples	
				▶ Gen	eral			, in onique		
6	37075920:			Map	metry Layers					
7	37075920:			▶ Map ▶ Mat						
8	37075920:				erators					
9	37075920				ord and Attri	butes				

# Generate the population-weighted distances using the Field Calculator

		-			inal Indicator - Q				
🗅 📁 🗟 🖾 🖄 💒 🛛 🕐 🖈 🔎	P	, Q Q .			C & &	- 🔣 - 🗧 -	🔓 🧱 🗱 Σ 💳	• 🖓 🖵 •	
🤽 🎕 Vi 🄏 🖏 🔀 🖊 🖊 🗒		•					ed: 6155, Selected: 0		
owser				8 = 6	L 🕈 🏛 🏶 🗜		i 🛱 🔍		
2 2 7 11 0	abc In	putID 👻 =				8.77	•	Update All Update Selected	
💕 tl_2019_37_bg.shp.ea.is		InputID	TargetID	Distance	fips_cnty	BG_pop	Total County Pop_County_pop	dist_weigh	
tl_2019_37_bg.shp.ea.is tl_2019_37_bg.shp.iso.x	(n	99603005 🖾		11371.406146		1503	103082	165.802	
tl_2019_37_bg.shp.iso.x	(n <sup>2</sup>	370499613013		153.6726790		712	103082	1.061	
Total County Pop.csv	3	3704996040	171970.0000	1418.757144	37049	1130	103082	15.553	
Geocoding	4	370499610022	103857.0000	428.0798494	37049	1018	103082	4.228	
Active License Only.shp	5	3704996040	75158.00000	1121.2132225	37049	1531	103082	16.653	
iyers	6	370499611003	20458.0000	1065.961606	37049	1761	103082	18.210	
🖌 🕼 🔍 🍸 🖏 + 🗰 🖬 🗔	7	370499607001	158803.0000	878.4974323	37049	1035	103082	8.821	
Centroids     Distance pop table for excel	8	3704996080	204849.000	1724.041430	37049	808	103082	13.514	
🕖 Total County Pop	9	3704996090	204849.000	50.15009818	37049	978	103082	0.476	
<ul> <li>✓ <u># distance matrix final</u></li> <li>✓ • distance matrix</li> </ul>	10	3704996080	140273.0000	714.6642920	37049	830	103082	5.754	the way
Active_Outlets	11	370499601022	213413.0000	6931.484125	37049	889	103082	59.779	
population table           X01_AGE_AND_SEX	-			10574.07433		2022	103082	207.415	
BG METADATA 2018	-	3704996050		405.5430219		1134	103082	4.461	States of P
* Active_Outlets				11981.701499		719	103082	83.573	
Active_Oddets	-								
Statistic Value		3704996060		463.8030284		784	103082	3.527	
	16	3704996030	191515.0000	4436.008101	37049	497	103082	21.388	
	17	370499610021	103857.0000	1545.541030	37049	2035	103082	30.511	
	18	3704996100	155869.0000	1481.5159701	37049	572	103082	8.221	
	19	370499610011	155869.0000	1214.264618	37049	654	103082	7.704	
	20	3704996060	158803.0000	516.81285111	37049	1186	103082	5.946	
	T Sho	ow All Features 🖕						8	
Selected features only		8 2							
C Type to locate (#K) Toggles the editing	etata cf.			Coordinate	97,32.400 <b>%</b> So	1:2790620	Magnifier 100%	Rotation 0.0 °	Render @EPSG:4269

*Final columns of "distance matrix final" including the newly generated "dist\_weigh"* 

Sum the dist\_weigh by county to get the total average distance for that county. To do this, open the Group Stats plugin, then drag "dist\_weigh" and "sum" to the Value panel, and "fips\_cnty" into the Rows column. Click Calculate, then click on Data in the menu bar and select "save all to CSV file." The results should show the final indicator measurement value for each county, as represented by the 5-digit county FIPS code.



Group Stats calculation of average population-weighted distance to nearest outlet

Check your units. Notice that the units in this example are in feet, which are not as easy to read as miles. To convert feet to miles, use Excel (or QGIS as shown in the next screenshot). Reimport the resulting CSV file into QGIS and navigate to the Field Calculator in the Attribute Table. Name the new output field name "avgdist\_mi," and select "Decimal number" for the output field type, with a field length of 10. In the Expression pane, multiply the final indicator measurement column field name by the conversion factor for feet to miles: "0.000189394."

	Field Calculator								
<ul> <li>Only update 0 set</li> <li>Create a new fie</li> </ul>		d							
Create virtual fiel Output field name Output field type Output field length Expression Fur = + - /	ield avgdist_mi Decimal number (real) h 10 Precision 3 unction Editor * ^    ( ) '\n' Fields and Values	- ▼ group field							
"Indicator per co _avg_distance" *0	ounty results	Double-click to add field name to expression string. Right-Click on field name to Values Search All Unique 10 Samples							
Output preview: 0.811178100486         Vou are editing information on this layer but the layer is currently not in edit mode. If you click OK, edit mode will automatically be turned on.									
Help		Cancel OK							

Convert indicator values from miles to feet using the Field Calculator

Exit editing mode to save the new column. Check your results one last time. The number of total features should match the number of counties, and the values should be within reason. For example, counties with the smallest average distance should have higher densities of outlets.

🔴 🕘 🔵 Indicator Results Final :: Features Total: 100,								
J	/ 🛛 🕃 🏹 🖷 🔫 🖻 🖹 🏷 🧏							
	fips_cnty	None						
1	37009	1.043						
2	37119	0.131000000						
3	37059	0.63400000						
4	37179	0.572000000						
5	37137	1.30900000						
6	37113	0.9300000						
7	37005	1.00300000						
8	37027	0.716000000						
9	37125	0.531000000						
10	37165	0.743000000						

#### Mapping Indicator Results

To create a map displaying the calculated indicator of outlet to person, join the CSV layer to a county shapefile by FIPS code.

You can create your own county shapefile, generate one from your block groups shapefile by creating a 5-digit county FIPS column and then using the dissolve tool to merge block groups into county boundaries. This is especially useful when you have attribute data at a subcounty level, but not at the county level yet. The dissolve tool will preserve the attribute data as you merge spatial boundaries into larger units. Alternately, county-level results can be joined to an existing county shapefile.

In this block group shapefile, as shown in the next screenshot, the state and county digits are separated in the attribute table. Create a new column that joins them to create a county FIPS code: "STATEFP" + "COUNTYFP." Make sure the output field type is "Text (string)." Remember to exit editing mode to save the new column by clicking on the upper left pencil icon in the attribute table.

*Combine the 2-digit state codes and the 3-digit county codes into the 5-digit FIPS code for each block group* 

0	00		2019_bg_	PSG_2264 :: Features Total: 61	55, Filtered: 6155, Se	lected: 0				?
/	🗾 🖬 😂 📑	6 ~ 6 6	8	1 🛃 🍳 🏶 🖺 🕈 🖉	<b>i i</b> i i i i i i i i i i i i i i i i i					
abc STATEFP $\mathbf{v} = \mathbf{E}$						*	Update All	Update \$	Selected	
-	STATEFP	COUNTYFP	TRAC			Field Calculator	FUNCS	ТАТ		
1	37	119	005100			Tield Calculator				
2	37	119	005000	Only update 0 selected feat	ures	_				
3	37	119	004600	✓ Create a new field		Update ex	kisting field			
4	37	119	003201	Create virtual field	221	_				
5	37	119	004800	Output field name CNTY_FIP						
6	37	119	004700	Output field type Text (strin Output field length 5	ng) Precision 3					
7	37	119	003106			•				
8	37	119	003105	Expression Function Edito	or					
9	37	119	003108	= + - / * ^	( ) (\n'		ow Values		g	roup field 📤
10	37	119	003109	"STATEFP" + "COUNTYFP"		Fields and Values abc COUNTYFP			e-click to add	
11	37	093	970300		•	Map Layers Total County Po	D		ression string. Click on field i	
12	37	093	970202		-	Total County Po Recent (fieldcalc)			ontext menu s oading option	
13	37	147	000400	'37' + "SapCountyl"		Q Search				
14	37	093	970402			"Outlets by Cou	inty_cou			
15	37	071	032301			"CountyArea" / "CountyArea" /	/ 25900		All Unique	10 Samples
16	37	071	032301			"CountyArea" * "Outlets by Cou				
17	37	071	032302							
18	37	071	032302							
19	37	051	003205	Output preview: '37161'						
-	77	051	000005							
<b>V</b>	Show All Features 🚽		8	Help					Canc	el OK

The FIPS column will be the basis on which block groups are grouped before dissolving them into their respective counties. Go to Vector in the menu bar, select "Geoprocessing Tools...," and select "Dissolve...". The Dissolve Tool window will pop up. Select the block group shapefile as the input layer and click on the "..." button for "Dissolve field(s)" to checkmark the FIPS county field. Click Ok, then Run, and close.

Input layer	<b>Dissolve</b> This algorithm takes a vector layer and combines their features into new
AWAIER AWAIER INTPTLAT INTPTLON CNTY_FIPS Select All Clear Selection Toggle Selection	features. One or more attributes can be specified to dissolve features belonging to the same class (having the same value for the specified attributes), alternatively all features can be dissolved in a single one. All output geometries will be converted to multi geometries. In case the input is a polygon layer, common boundaries of adjacent polygons being dissolved will get erased.
0% Help Run as Batch Process	Close Run

### Dissolve the block groups into larger units based on shared county FIPS codes

This will create a temporary layer, "Dissolved," so make sure to re-save it as a new layer to make it permanent: "County final." Then right-click on the "County final" shapefile and navigate to the "Joins" pane. Click on the plus (+) button to open the new join pop-up and use the county FIPS columns of each layer as the join and target fields. Click Ok.

	Edit Vector Join					
Join layer	Indicator per county results	•				
Join field abc fips_cnty						
Target field	abc CNTY_FIPS					
✓ Cache join layer in virtual memory						
Create attribute index on	join field					
Dynamic form						
Editable join layer						
▶						
Custom Field Name Prefix						
	Cancel	ОК				

*Join the indicator value CSV layer to the shapefile output of the dissolve tool* 

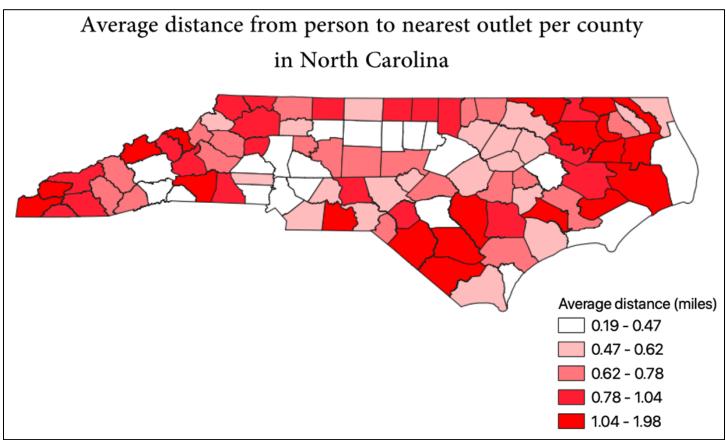
After verifying the join is successful, right-click again on the "County final" shapefile and navigate to the "Symbology" pane. Select "Graduated" for the top field and select the indicator column for the value: "avgdist\_mi." Click "Classify" to generate the color-coded ranges, and then click Ok. The shapefile should now demonstrate different fill colors for each county based on the value.

	0		Layer Pro	operties - Co	ounty final   Syr	nbology			
Q		불 Graduated							•
i	Information	Value	1.2 avgdist_m	i				•	3
2	Source	Symbol							-
~~		Legend format	%1 - %2				Precision 2	-	Trim
~~	Symbology	Color ramp							-
abc	Labels	Classes H	listogram						
۹.	Diagrams	Symbol 🔻 Value	es Le	egend					
Ŷ	3D View		360 - 0.4658 0 658 - 0.6190 0						
	Fields	✓ 0.61	190 - 0.7814 0 314 - 1.0434 0	.62 - 0.78					
-8	Attributes Form		134 - 1.9790 1.						
•	Joins								
	Auxiliary Storage								
<b>Q</b> _	Actions								
9	Display								
*	Rendering								
3	Variables								
	Metadata	Mode 🕕 Equa	al Count (Quan	tile) 🔻			Classes	5	\$
	Dependencies			elete All				Advand	ed -
-	Legend	✓ Link class bo	oundaries						
-		Layer Rend	lering						
	QGIS Server	Help	Style 🔹	Apply			Cancel	Oł	<

#### Adjust layer properties, creating a graduated map of the average distance in miles

Use "New Print Layout" under "Project" in the menu bar to add a title, legend, or other features before exporting your map. A sample map output is shown in the next graphic.

Map of average distance from person to nearest alcohol outlet by county in North Carolina



# Calculating the Average Distance from an Alcohol Outlet to Its Nearest Outlet (Outlet-to-Outlet)

Average distance between outlets by county can be calculated in QGIS using the **distance matrix** tool together with the **Group Stats** plugin.

If you use Nearest Neighbors from the Processing Toolbox, it will contain distances within the origin points. Instead, use the **Distance Matrix tool** from the menu bar: Vector  $\rightarrow$  Analysis Tools  $\rightarrow$  Distance Matrix. This version ignores calculating the distance of the origin points with itself. By selecting the option to calculate only the nearest k = 1 points, the distance matrix will have rows representing all outlets and a single column representing the distance to its nearest one outlet.

# Calculate Each Outlet's Distance to its Next Nearest Outlet:

- 1. In the menu bar, select Vector, then Analysis Tools, and click on **distance matrix**.
- 2. Make sure your outlets layer has a unique ID for each row/outlet. If there is not one in the existing attribute table, create a new column called "Temp\_ID," based on the row number.

- 3. In the pop-up window, select the appropriate layer and fields:
  - 1. *Input point layer:* Select the layer containing all the plotted outlet points.
  - 2. Input unique ID field: Choose the field name containing a unique ID for each point.
  - 3. *Target point layer:* Since you are measuring points within the same layer, select the same layer as you did for Input point layer.
  - 4. *Output matrix*: Linear (N\*k x 3) distance matrix.
  - 5. Use only the nearest (k) target points: 1.

Using Distance Matrix to calculate distance to each point's nearest neighbor.

• • •	Distance Matrix		
Parameters Log		•	Distance matrix
Input point layer		-	This algorithm creates a table containing
,* NC_alcohol_stateplane_2016 []			a distance matrix, with distances
Selected features only			between all the points in a points layer.
Input unique ID field			
122 Temp_ID	*		
Target point layer			
* NC_alcohol_stateplane_2016 []			
Selected features only			
Target unique ID field			
122 Temp_ID	•		
Output matrix type			
Linear (N*k x 3) distance matrix	•		
Use only the nearest (k) target points			
1	€ 2		
Distance matrix		Ŧ	
	0%		Cancel
Help Run as Batch Process			Close Run

- 4. Click Run, then close. A new temporary layer called "Distance Matrix" will have appeared.
- 5. What unit of measurement are your distances calculated in?
- 6. Double check distance units by going to the QGIS Project Properties, General, Measurements, Units for distance measurement. If the units are incorrect, change the unit and then rerun your outputs.

# Join the Distance Matrix Results to the Original Outlets Layer

- 1. Right-click on the outlets layer to which you will be adding the distance matrix results. Select Properties, then navigate to the "Joins" page.
- 2. Click on the green plus (+) button on the bottom left. The pop-up window for "Add Vector Join" will appear.
- 3. Select the appropriate layer and fields:
  - 1. Join Layer: Distance matrix.
  - 2. Join Field: Input ID.
  - 3. Target Field: The name of the unique ID field, in this case, Temp\_ID.
- 4. Select Ok to close the pop-up window. The new join will appear in the Layer Properties window. Click Ok to apply the join.
- 5. Open the attribute table of the outlets layer. The data should appear in new columns on the far right. Distance matrix will create a table with InputID, TargetID, and Distance. When you do a join, the columns may appear as your unique ID column, "Distance\_m" and "Distance\_l," with TargetID renamed as Distance\_m. The actual distance you want to use in the next step is Distance\_l. If you are unsure which column is correct, select a row in the attribute table and navigate to the map view. The selected point will be highlighted, and you can use the ruler measurement tool to manually measure the closest neighbor, then see which column contains that value.
- 6. Export and save the outlets layer, otherwise the distance matrix data and join are temporary.

#### Join the distance matrix results to the original outlets layer

😑 🔿 🗧 Edit Vec	ctor Join
Join layer	" Distance matrix
Join field	123 InputID 👻
Target field	123 Temp_ID 👻
✓ Cache join layer in virtual memory	
Create attribute index on join field	
Dynamic form	
Editable join layer	
Joined Fields	
Custom Field Name Prefix	
	Cancel OK

#### Summarize average outlet to nearest outlets distance by county

We will use Group Stats to average distances by their "category" of county, listed in another column. (Note that you can also use "Statistics by Category" to calculate this in QGIS 3.0.)

- 1. Go to the Menu bar, select Vectors, mouse over Group Stats, and select "Group Stats."
- 2. In the pop-up window, a control panel has selectable options for Layers and Fields:
  - 1. Layers: Select the layer that contains the list of outlets and each distance to the nearest outlet
  - 2. Fields: Drag and drop "County" into the Rows box, then drag and drop the field name for distances, "distance\_l" and "average" into the Value box. You may need to scroll to the end of Fields to see "average."
- 3. Select "Calculate." A table should appear in the left portion of the pop-up window.

Group Stats PX Control panel 1 2 . . 1 Layers County Ŧ Outlet distances Ŧ Alamance 506.677 2 Fields Alexander 3897.19 3 . address Alleghany 1452.45 4 Bsnss\_I 5 Anson 265.629 BzStt\_D Ashe 2101.18 6 City Avery 695.056 Corp\_Nm 7 County Beaufort 1279.51 8 Distance m Bertie 4656.15 9 Distance\_1 Bladen 3882.23 10 Latitud Brunswick 822.603 11 Longitd Buncombe 1864.84 12 max\_dat Burke 1051.27 13 max\_yer Cabarrus 638.958 14 Filter Columns Caldwell 636.396 15 1535.06 Camden 16 17 Carteret 733.037 Value 📃 use NULL values Rows 2949.97 Caswell 18 County Distance\_1 Catawba 772.999 19 average Chatham 3463.35 20 Use only selected features Clear Cherokee 1642.45 21 Calculate 22 Chowan 1039.68 Calculate... 100% | generate view...100% | done.

The right of the window is where you can drag and drop fields; the results will appear on the left after clicking Calculate.

- 4. To save the results, click on Data on the QGIS menu, and select "Save all to CSV file."
- 5. Navigate to your computer's folder to verify and view the CSV file saved. It should consist of two columns: the county name and the average distance in meters between outlets for each.
- 6. To use a distance measure that is more easily understood, use Excel to convert meters to miles by dividing the meters by 1609.

#### Visualize results

To visualize these results, add the CSV file as a layer and join it to the county shapefile using instructions from previous steps. As before, use Symbology to create an output as shown in the next image. Select Graduated for the top field and select the indicator column for the value: "avgdist\_outlets." Click Classify to generate the color-coded ranges, and then click Ok. As before, you can use the Print Layouts and Layout Manager for more detailed annotation. Detailed instructions for using Print Layouts in QGIS are available online at docs.qgis.org.

Map of average distance from alcohol outlet to next nearest outlet by county in North Carolina

