Intermediate GIS with QGIS and PostGIS

Richard Dunks, Co-Instructor

Eric Brelsford, Co-Instructor

Clayton Hunter, TA

Follow along: http://bit.ly/dot_intermediate_gis2

Goals for the class

- Provide a review of key spatial concepts and GIS functions in QGIS
- Demonstrate connecting to a spatial database in QGIS
- Introduce the structured query language (SQL) for querying relational databases
- Practice writing queries in SQL to accomplish key spatial analytical tasks with real-world data to answer meaningful analytical questions
- Practice best practices in communicating spatial analysis

Outcomes

- You will feel more comfortable with GIS concepts
- You will be more familiar with using QGIS
- You will understand the purpose of spatial databases
- You will have an understanding of the fundamentals of SQL
- You will be familiar with advanced styling techniques in QGIS

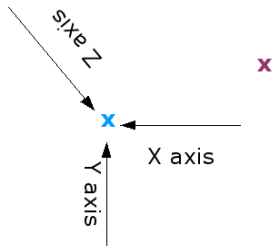
Goals for this morning

- Review basic geospatial principles
- Practice performing basic spatial operations in QGIS
- Discuss spatial databases
- Demonstrate and practice loading data from a spatial database in QGIS
- Practice answering an operational question with spatial data
- Demonstrate and practice more advanced styling techniques in QGIS

Basic Spatial Elements

Vector Point Feature

Point Geometry (indicates the x,y and z position of the feature)

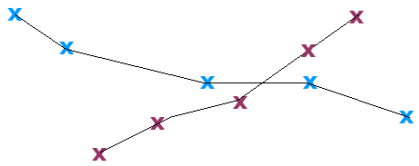


Point attributes (describe the feature)

Id, Name, Description
1, Tree, Outside our classroom 2, Light post, At the school entrance

Vector Polyline Feature

Polyline Geometry (a series of connected vertices that do not form an enclosed shape)

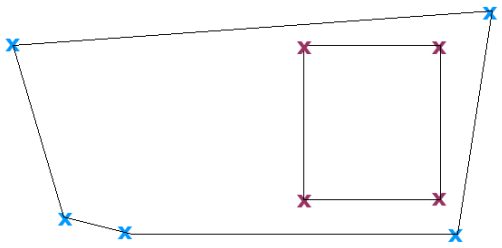


Polyline attributes (describe the feature)

Id, Name, Description
1, Footpath 1, From class to the playground 2, Footpath 2, From the school gate to the hall

Vector Polygon Feature

Polygon Geometry (a series of connected vertices that do form an enclosed shape)



Polygon attributes (describe the feature)

Id, Name, Description
1, School Boundary, Fenceline for the school 2, Sports Field, We play soccer here

Source: http://docs.qgis.org/2.8/en/docs/gentle_gis_introduction/vector_data.html#overview

Geographic Information System (GIS)

- Create interactive queries (user-created searches)
- Analyze spatial information
- Edit data in maps
- Present the results of all these operations

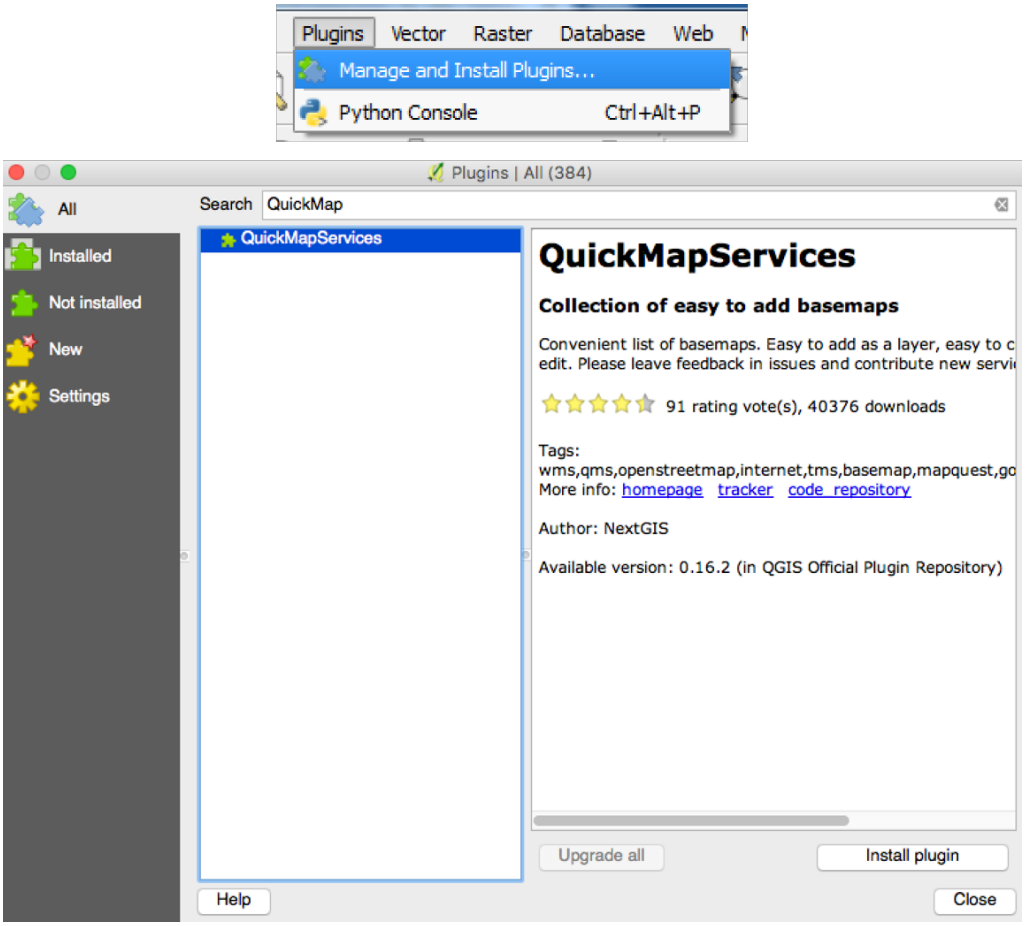
"Any system for capturing, storing, checking, and displaying data related to positions on the Earth's surface"

- National Geographic Education Encyclopedia

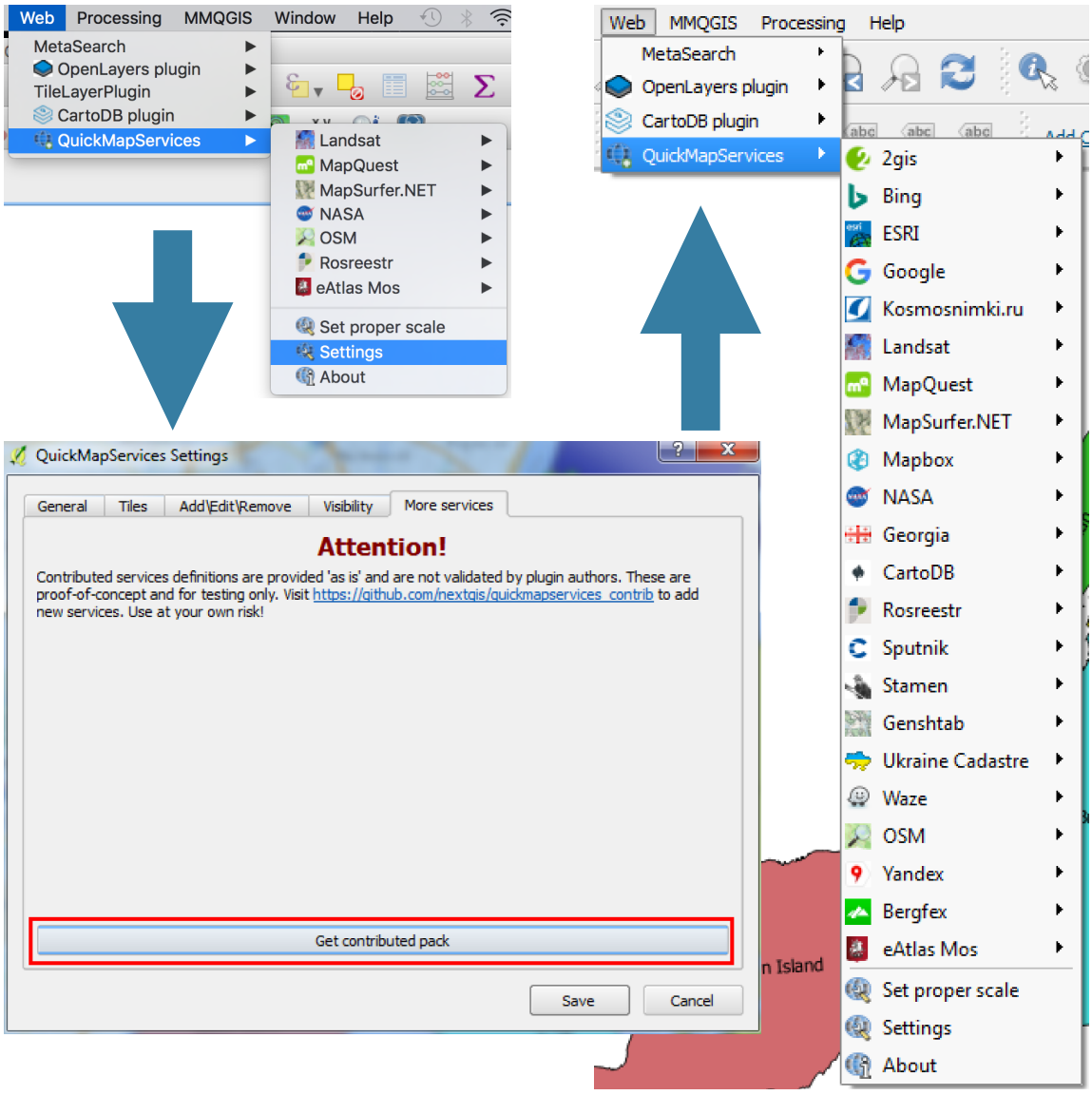
"In a GIS, you connect data with geography."

- GISgeography.com

Adding Basemaps (QuickMapServices Plugin)



Adding Base Maps (QuickMapServices Plugin)



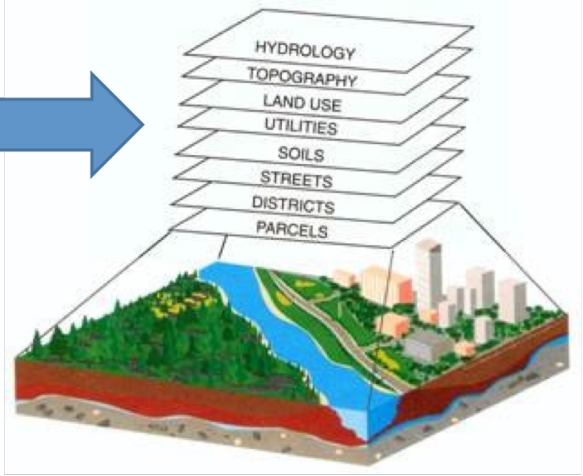
Choosing a Base Map

- Think about what someone reading your map needs to see for context
- Think about how the base map interacts with the data on your map

Adding Layers

Layers (Vector Data)

Basemap (Raster Data)



Shapefiles

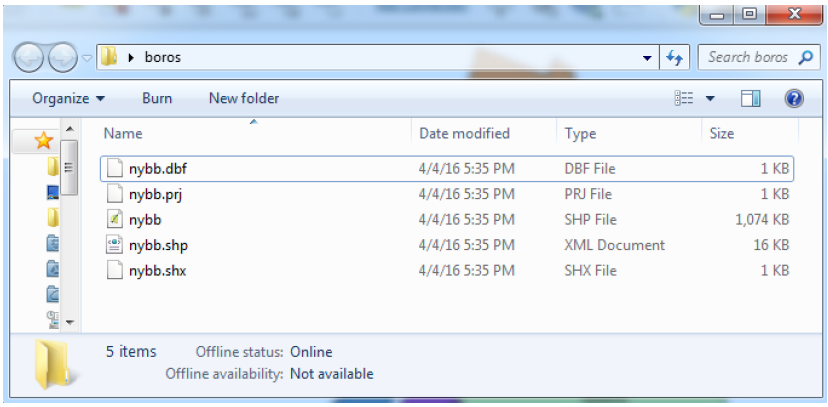
- Basic file for storing map elements
- Stores spatial data, like points, lines, and polygons
- Multiple files comprise a "shapefile"
- Column names can only be ten characters long

Shapefile file components:

- .shp — The main file that stores the feature geometry
- .dbf — The dBASE table storing attribute information of features
- .prj — The file that stores the coordinate system information
- .shx — The index file that stores the index of the feature geometry
- .cpg—Identifies the character set to be used
- .sbn and .sbx—The files storing spatial index of the features

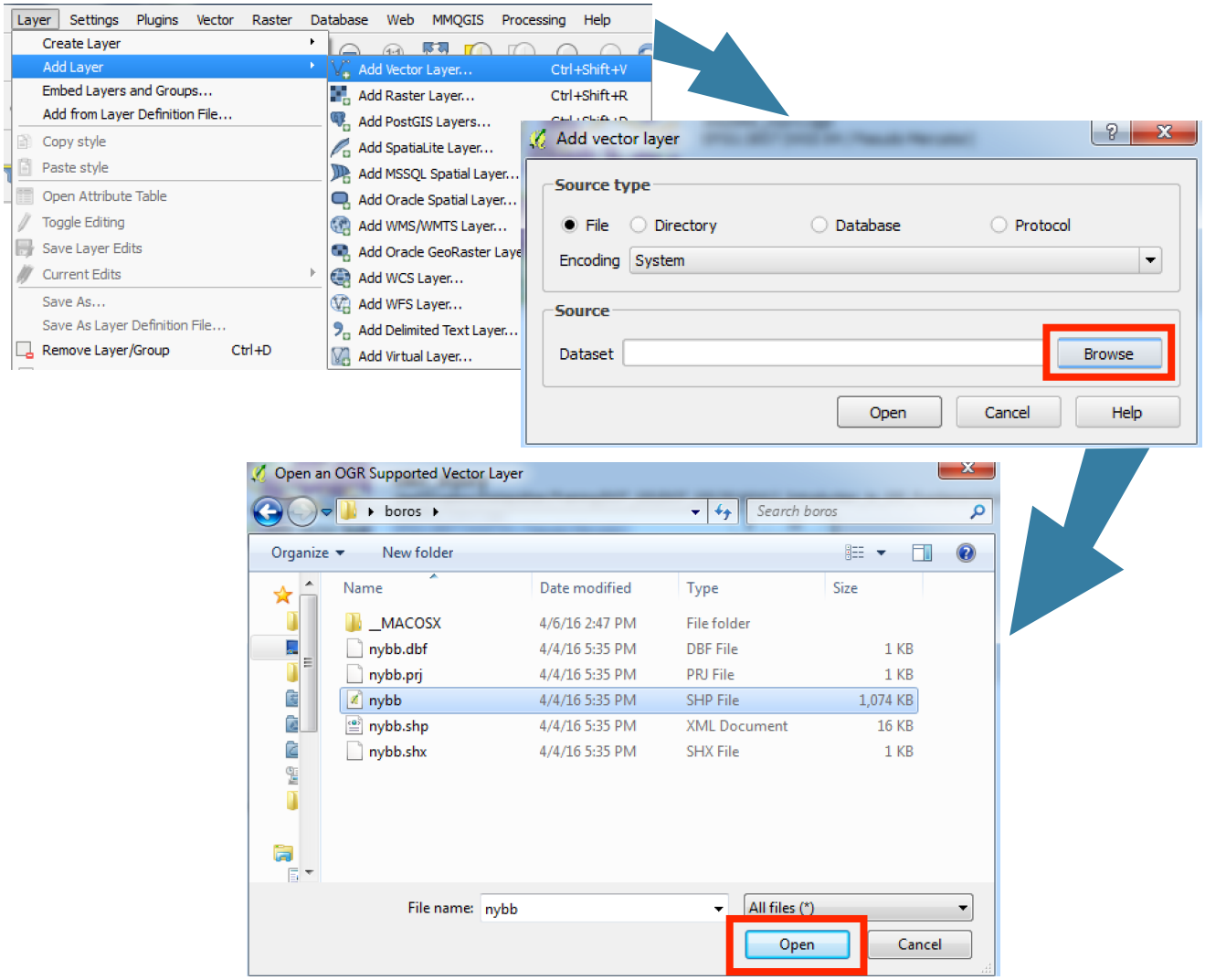
Layer Ordering

- Layers on top are drawn on top
- Just drag and drop within the Layers Panel to change order



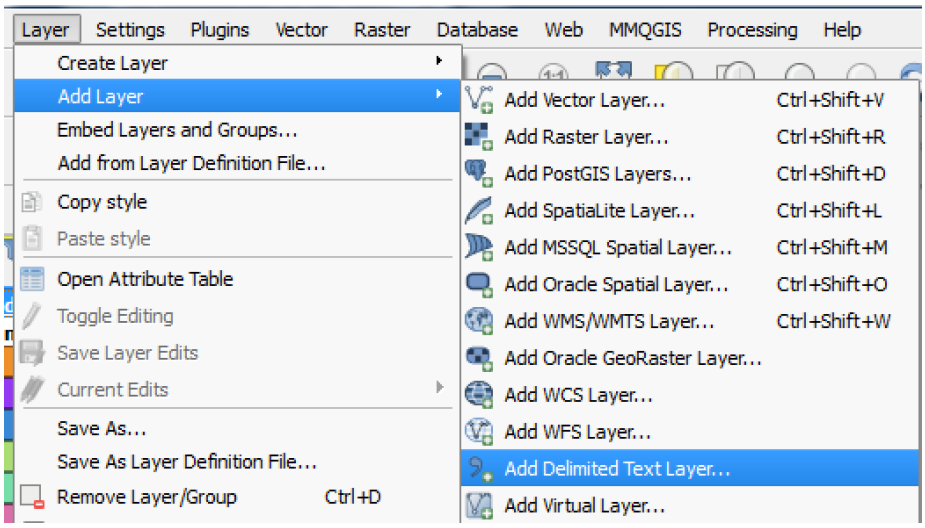
Download and Import Shapefile in QGIS

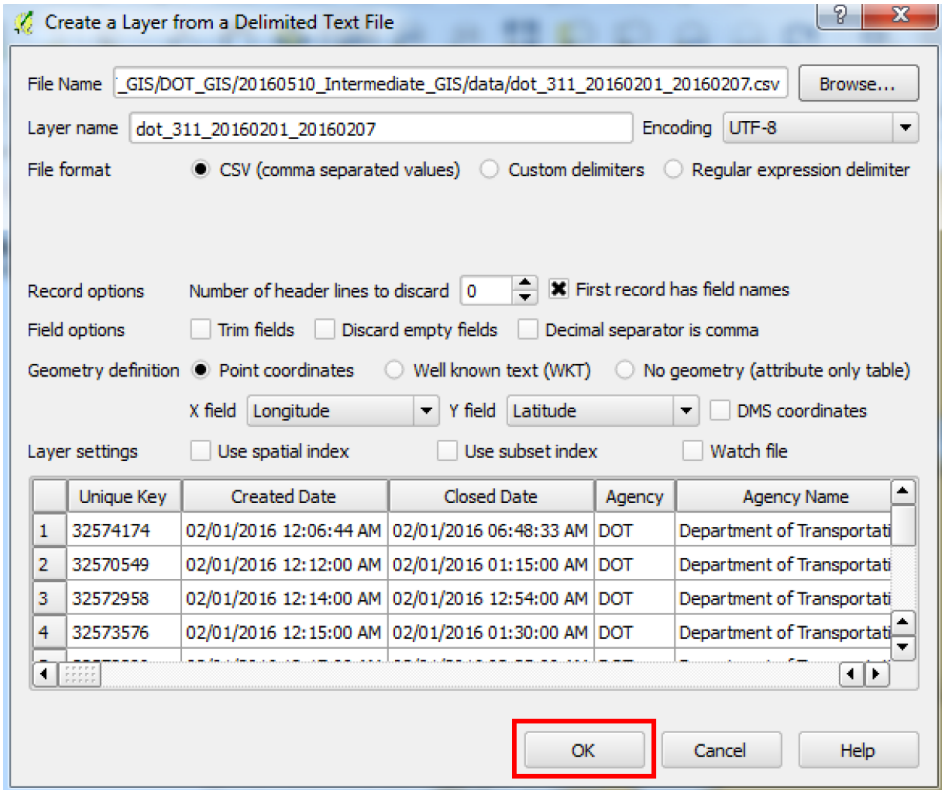
1. Click this link (http://www.datapolitan.com/DOT_GIS/20160510_Intermediate_GIS/data/boros/boros.zip) and download the file to your desktop
2. Unzip the file
3. Open in QGIS following steps below



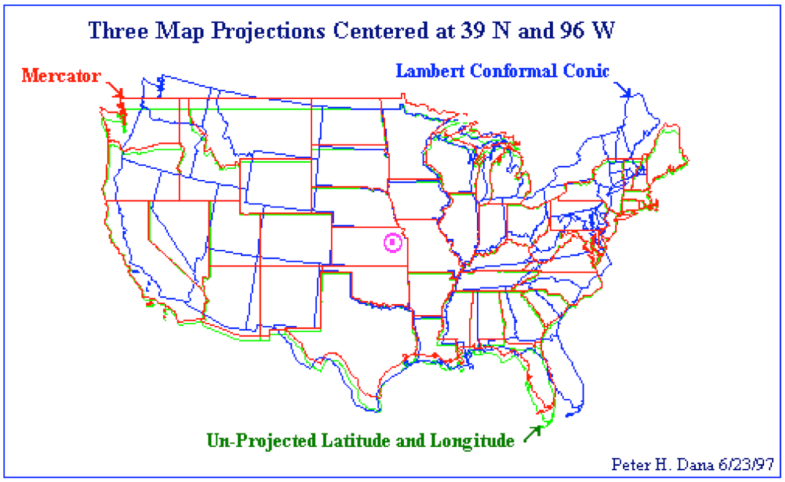
Download and Import a CSV in QGIS

1. Right click this link: http://www.datapolitan.com/DOT_GIS/20160510_Intermediate_GIS/data/dot_311_20160201_20160207.csv
2. Select "Save link as"
3. Save the file to your desktop





Projections



- No one's favorite part of GIS
- But a necessary part of it nonetheless
- Convert points on the 3-dimensional Earth (latitude and longitude) to x and y coordinates on a 2-dimensional map
- Every projection distorts some part of your map
- For the most part we will work in WGS 84 (latitude and longitude)
- In NYC, we use a more accurate projection NY State Plane/Long Island Zone
- Identified by unique IDs that make it easier to talk about them
- WGS 84 is referred to as EPSG:4326
- State Plane Long Island is referred to as EPSG:2263

Remember these two and you should be set

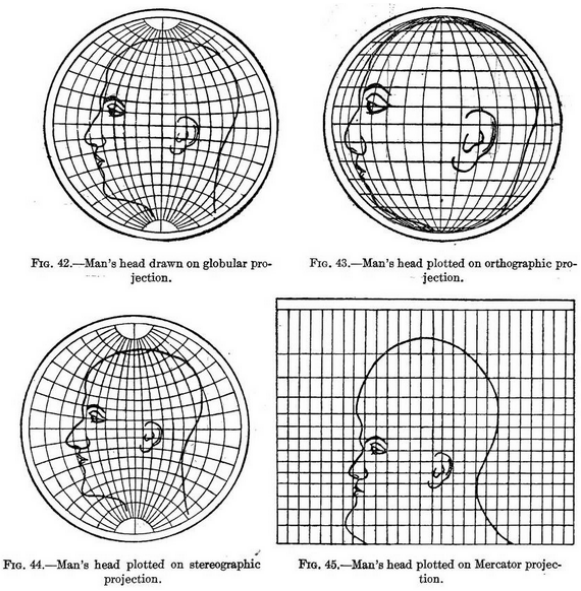


FIG. 42.—Man's head drawn on globular projection.

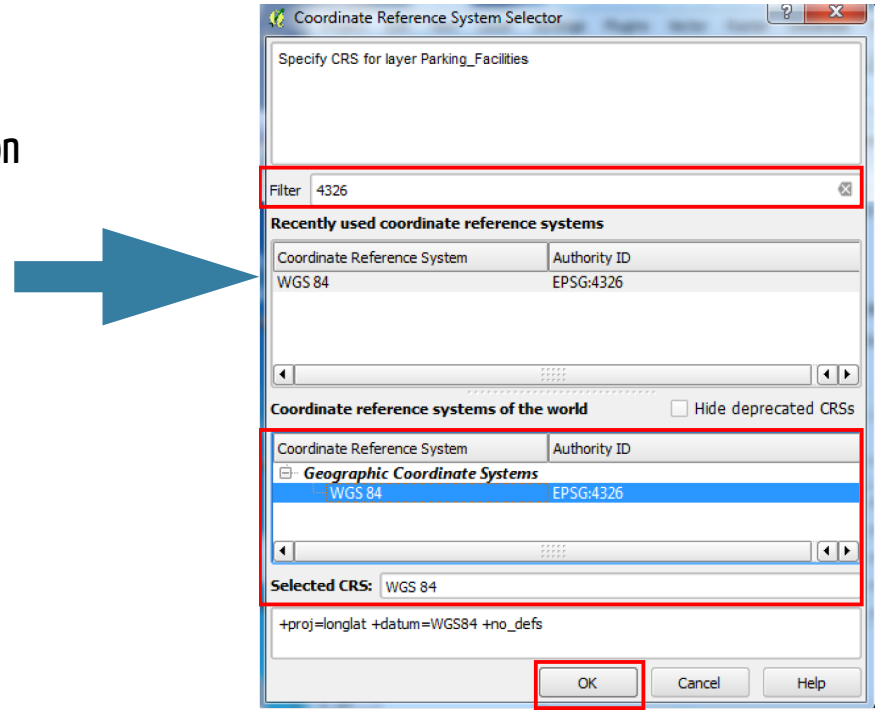
FIG. 43.—Man's head plotted on orthographic projection.

FIG. 44.—Man's head plotted on stereographic projection.

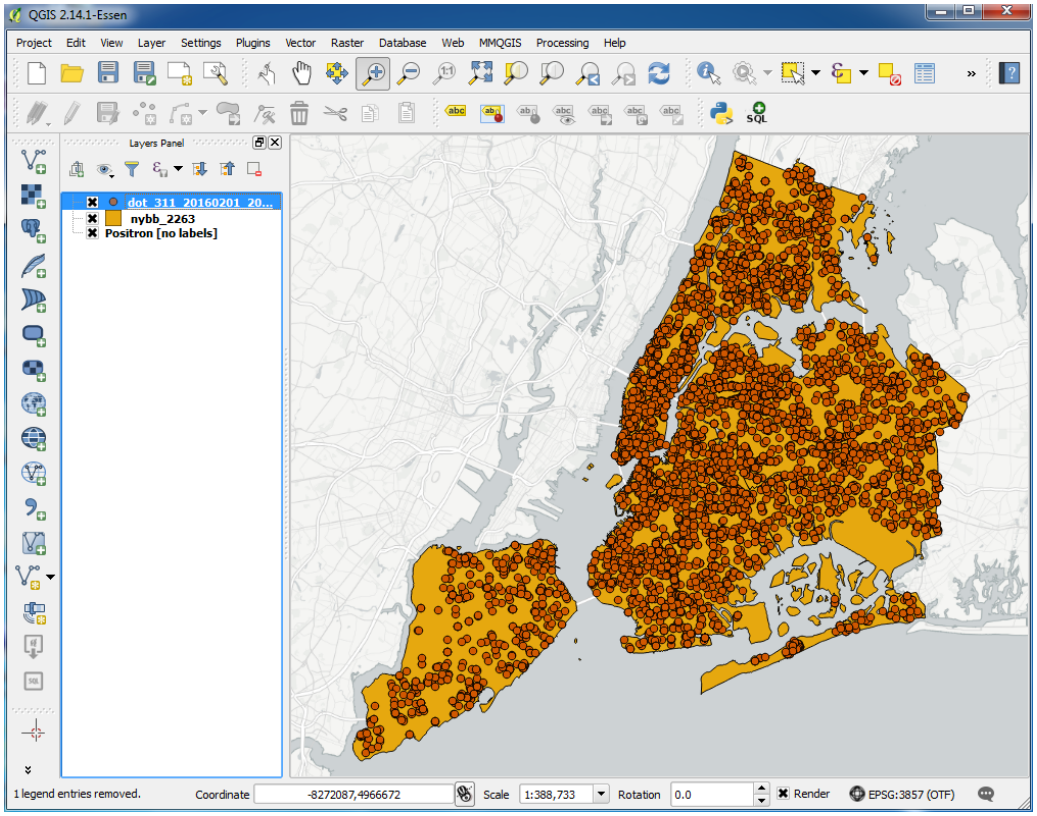
FIG. 45.—Man's head plotted on Mercator projection.

Setting the Projection

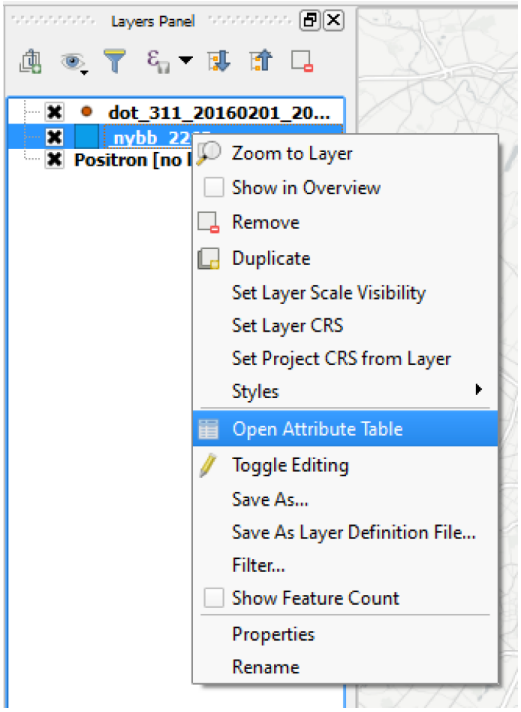
- Filter by “4326”
- Select “WGS 84”
- Click “OK”

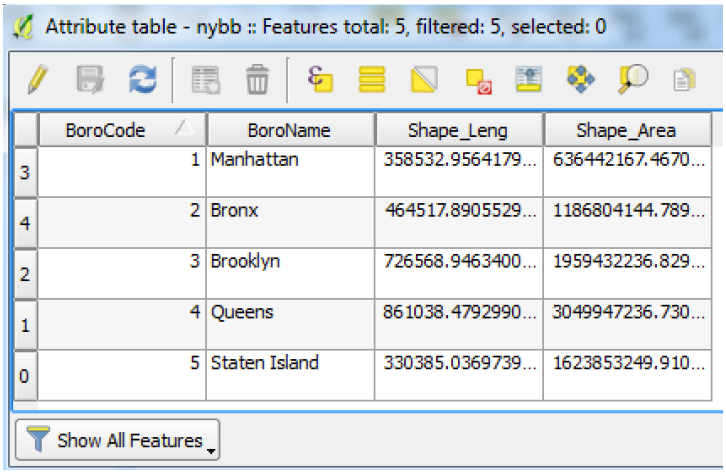


The Result



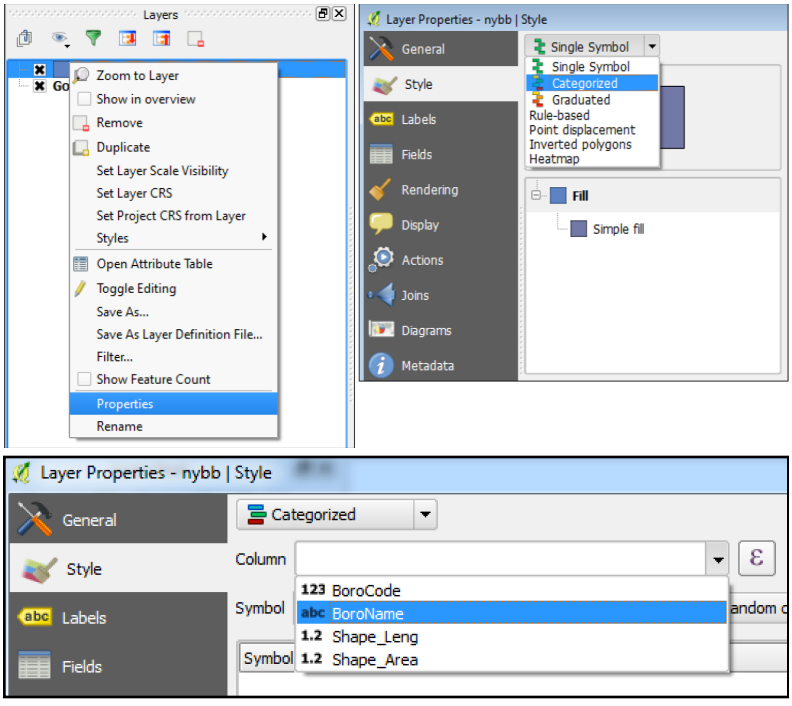
Attribute Table (Right click on layer)

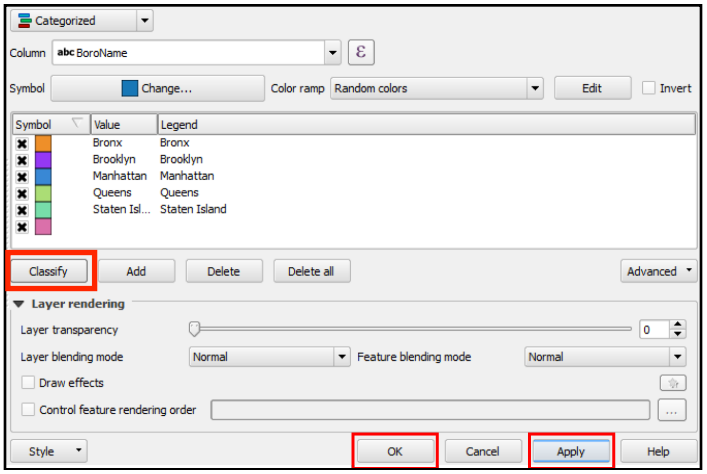


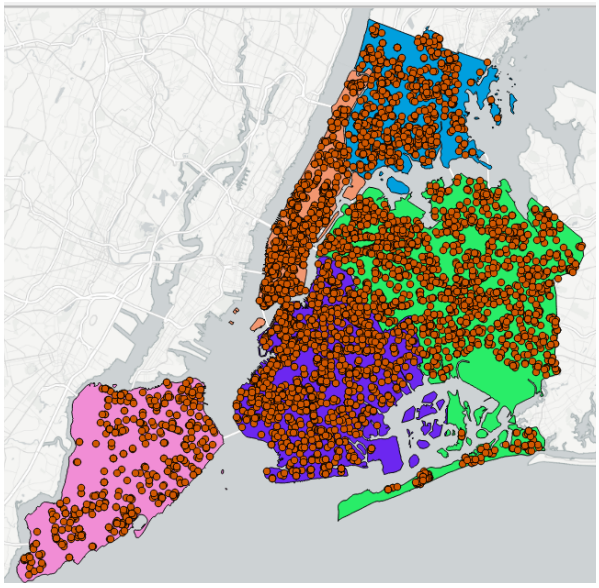


Styling Features

1. Right-click the layer and select the Properties option
2. Select "Style," and finally, choose "Categorized"
3. Select the column that has the data you want to style (from attribute table)
4. Select "Classify"
5. Click "Apply" and "Ok"







Example map after styling boroughs with random colors

Your turn

- Style the points and polygons however you'd like
- Try adding labels to the data
- Get familiar with the interface

Spatial Databases

Database

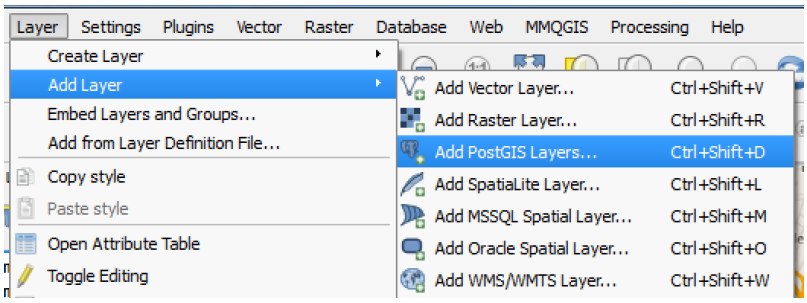
- "An organized collection of data" - Wikipedia
- Another source of data
- A program that often runs on a computer other than the one you're using (a server)
- You can connect to databases through many types of software, including QGIS

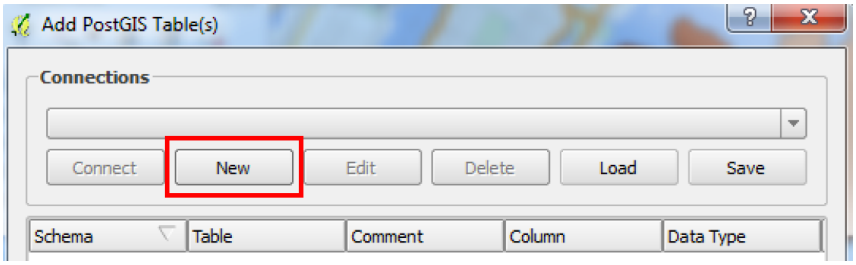
Spatial Database

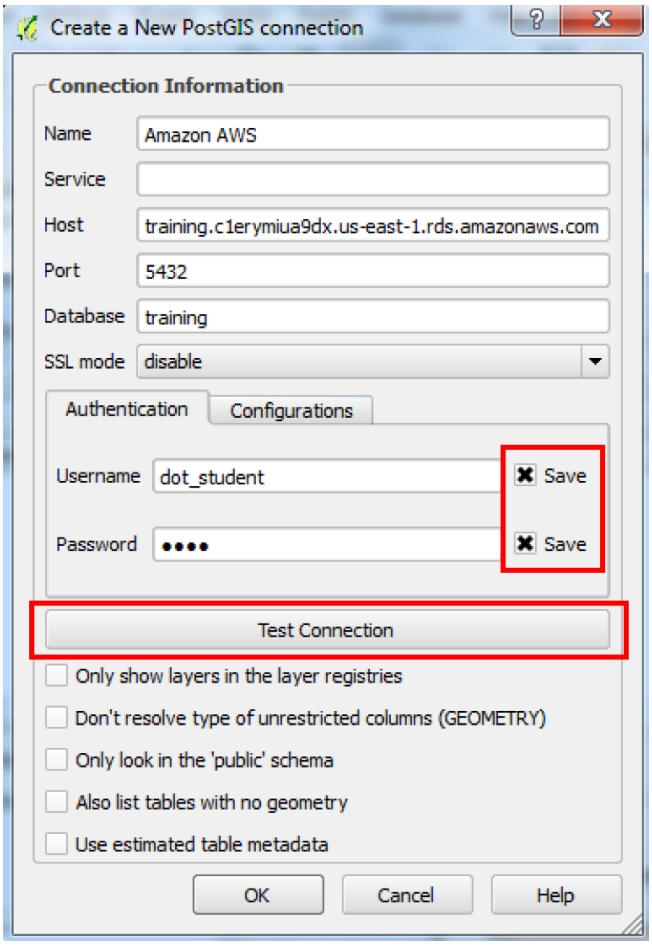
- A database that stores spatial features
- Optimized for processing spatial data, especially location-aware queries
- Both a storage and analysis tool

Connecting to a spatial database

1. Layer -> Add Layer -> Add PostGIS Layers
2. Under Connections, select "New"
3. Add connection information
4. Click to save username and password
5. Click "Test Connection" to make sure everything is entered properly
6. Click "OK"
7. Click "Connect"
8. Click on the tables you want to load and select "Add"



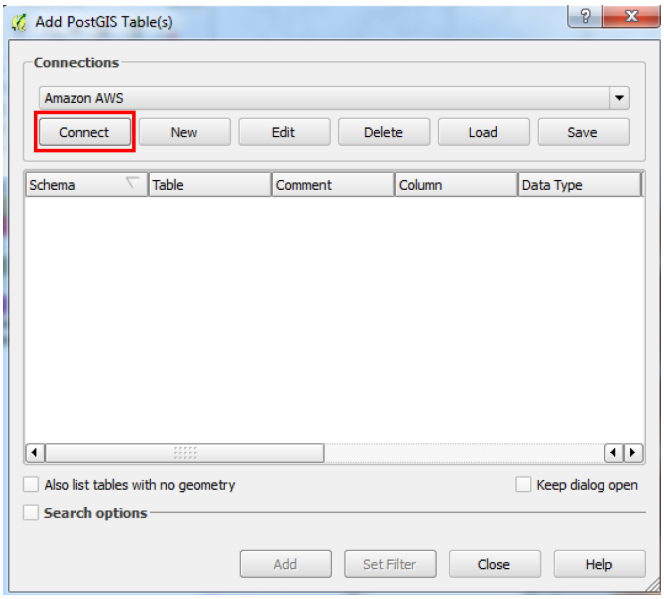


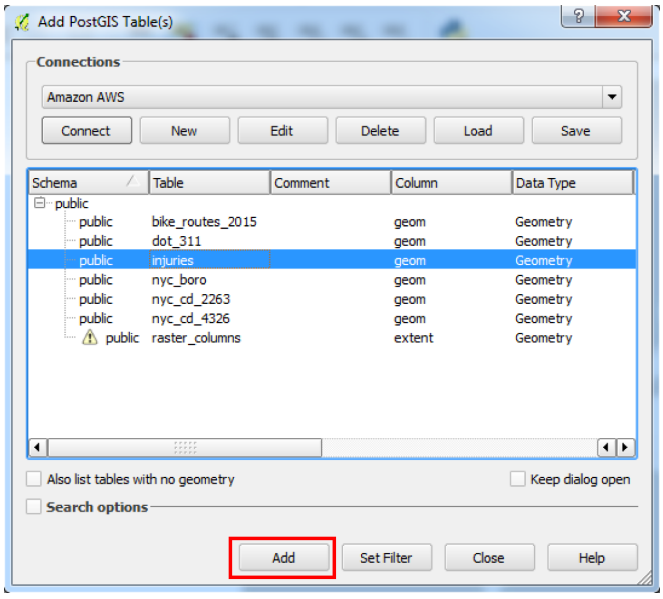


Now the data is just the same as any other data source (CSV, shapefile, etc)

Connection Information for class

- Name: (whatever you'd like)
- Host: training.c1erymiua9dx.us-east-1.rds.amazonaws.com
- Port: 5432
- Database: training
- Username: dot_student
- Password: qgis

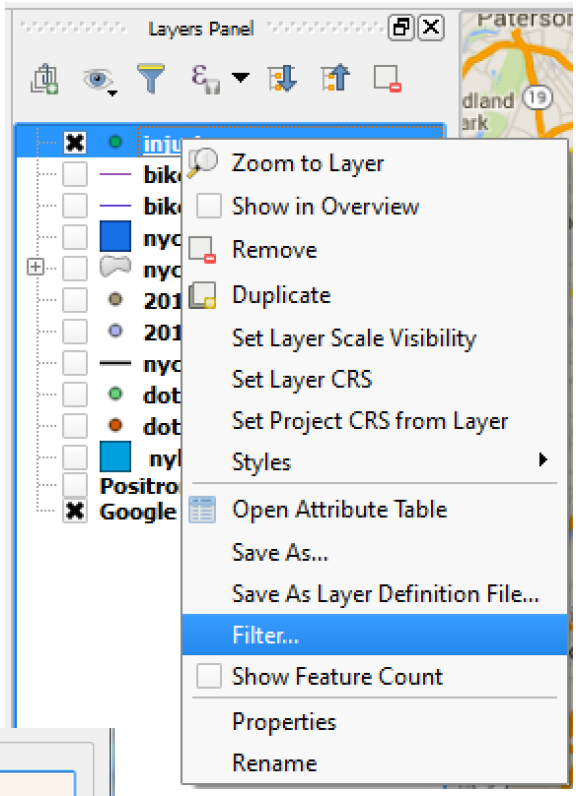


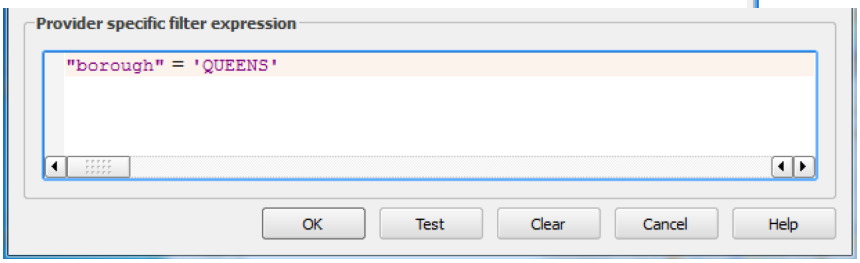


Exercise 1 - Selecting Injuries in a Borough

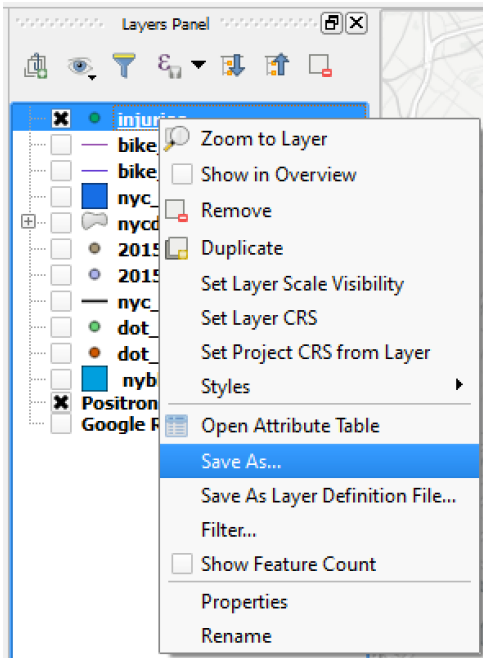
- Add the "injuries" table to your map
- Right click and select "Filter..."
- Select borough and query for your borough

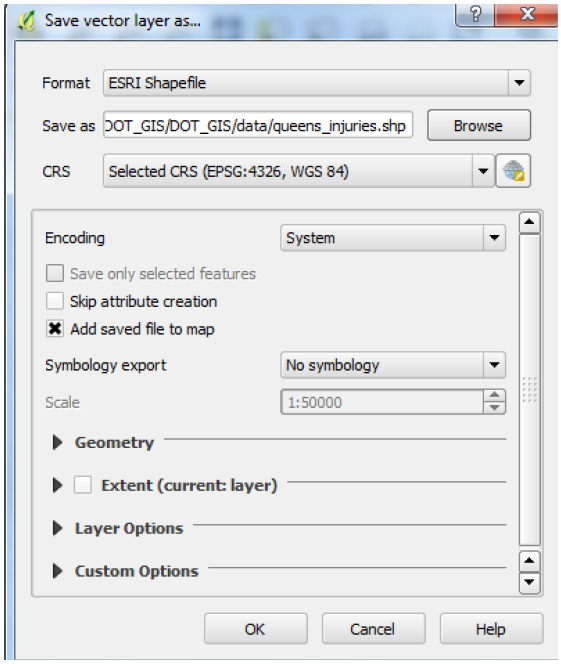
TIP: Always test your queries to make sure they return resulting rows





Save Selected Features as a Shapefile





Database datatypes

- Like attribute tables, database columns have types
- Numbers, text (strings), and dates are the main types
- Spatial databases add geometry columns
- Geometry columns let us map our data

Why use a spatial database?

- One authoritative copy of data
- Multiple people can use and edit it at once
- Generally is going to be faster than using a shapefile

Drawbacks of using a spatial database

- Not as simple as a CSV file or shapefile
- It takes some setup to connect to a database
- Need permission to connect to a database
- Need to know SQL to get the most out of it

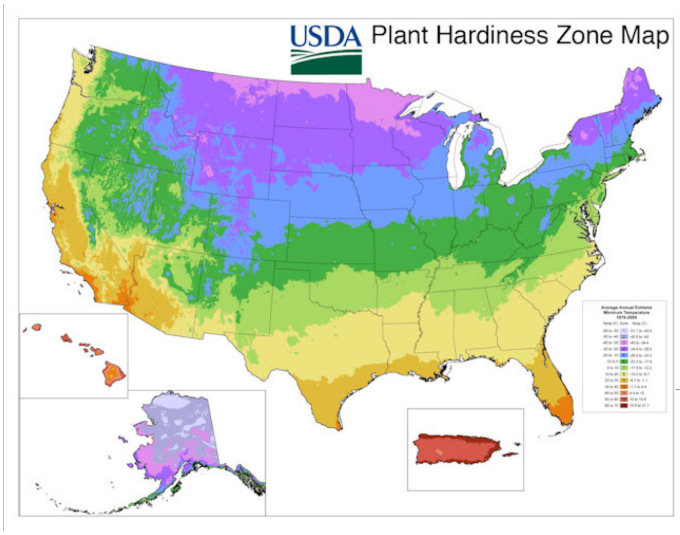
Types of Maps

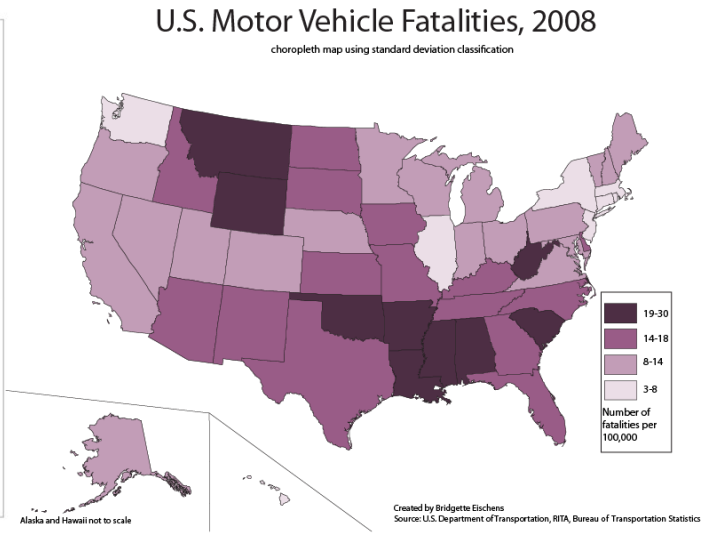
General Reference Maps

- Show important physical features of an area
- Include natural and man-made features
- Usually meant to help aid in the navigation or discovery of locations
- Usually fairly simple
- Can be stylized based on the intended audience (tourists vs locals)

Thematic Maps

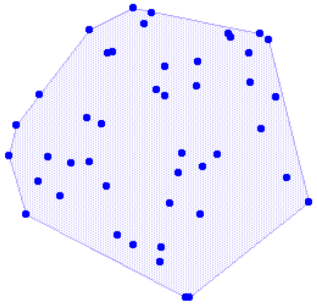
- Focuses on a specific theme or subject area
- Features on the map represent the phenomenon being mapped
- Spatial features used for reference



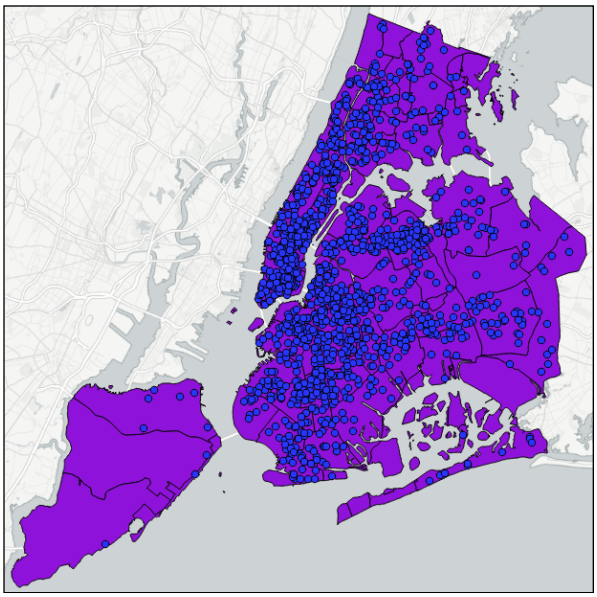


Basic Spatial Joins

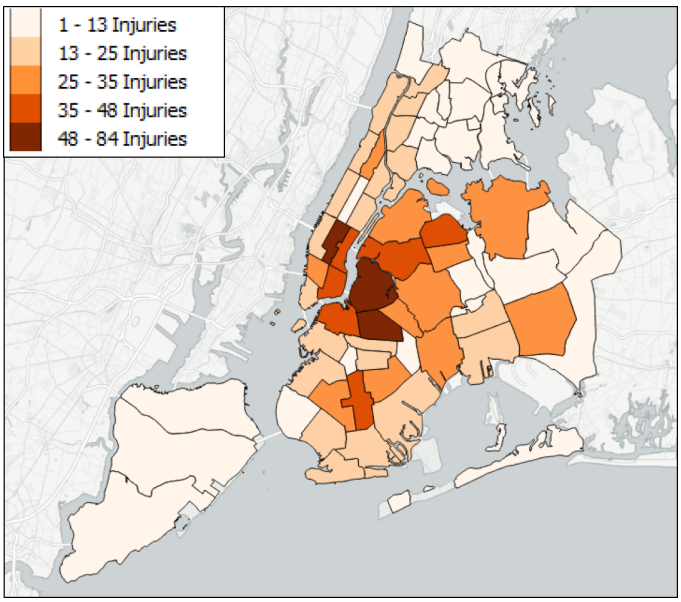
- Point to Polygon -> Relate points inside a polygon to that polygon (ex. count the number of points)
- Polygon to Point -> Points can take on value of enclosing polygon



Exercise 2 - Which Community Boards in NYC have the highest number of bicyclist injuries?



NYC Community Board boundaries and bicyclist injuries



Map of bicyclist injuries by NYC Community Board that we'll create

Gathering the data

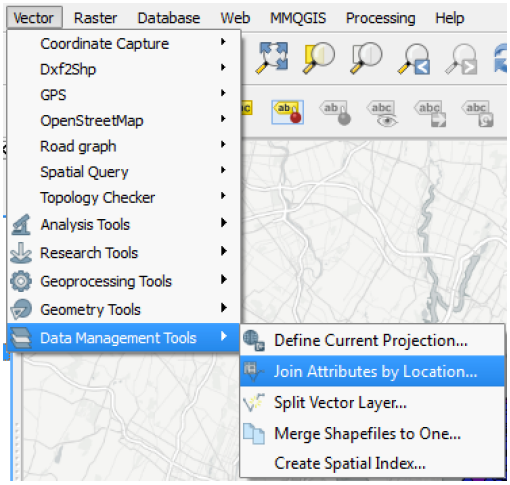
- Add the "nyc_cd_4326" table
- Add the "injuries" table

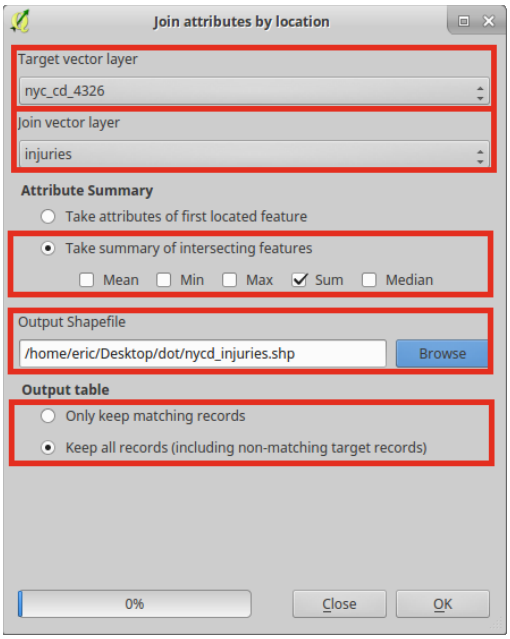
The Task

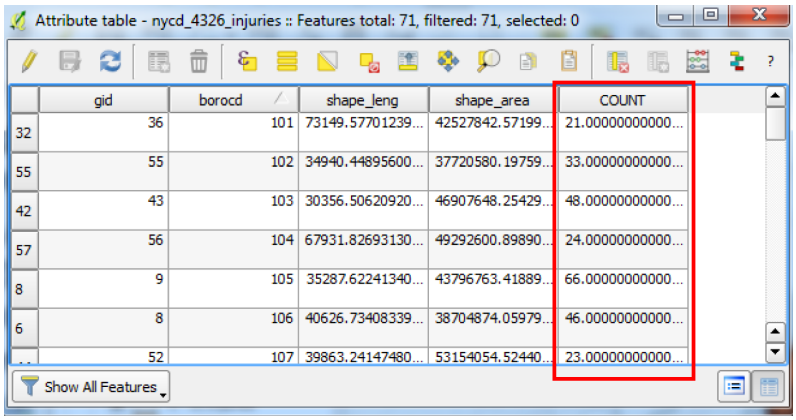
- Join injuries (points) to community districts (polygons)

Spatial Join steps

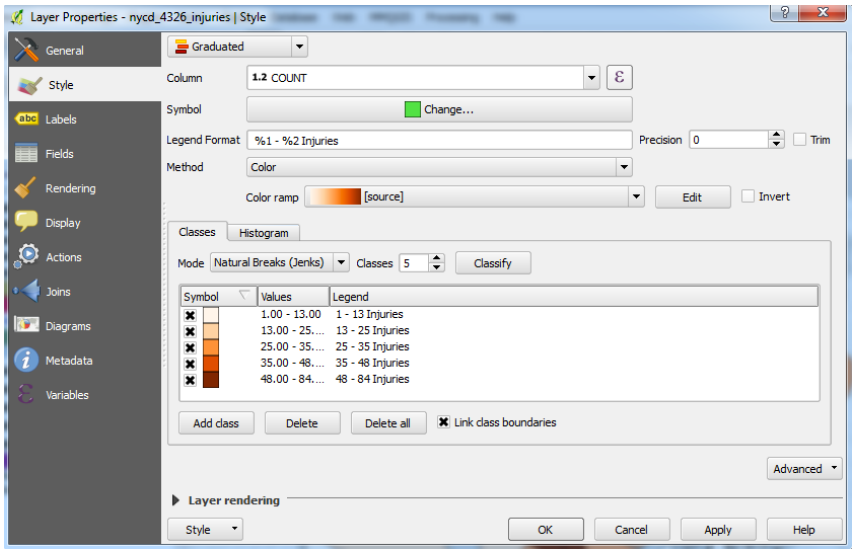
1. Vector->Data Management Tools->Join Attributes by Location
2. Select the target layer (the layer the attributes will go to)
3. Select the join layer (the layer attributes are taken from)
4. Select "Take summary of intersecting features", check "Sum"
5. Select the location for the shapefile
6. Select "Keep all records"
7. Click "OK"







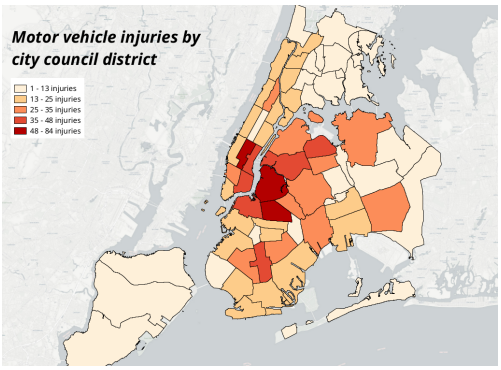
Attribute table with successfully joined features



Style settings for the choropleth map

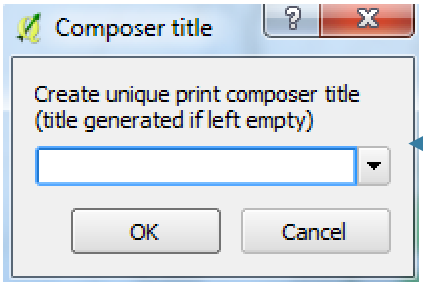
Print Composer

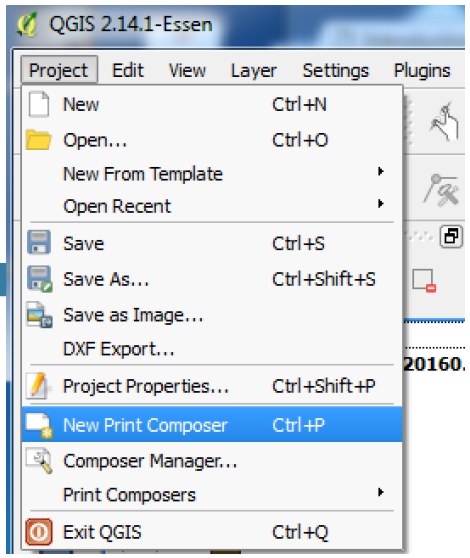
- How you make exportable and printable maps in QGIS
- Able to add map elements (legends, scales, text, etc)

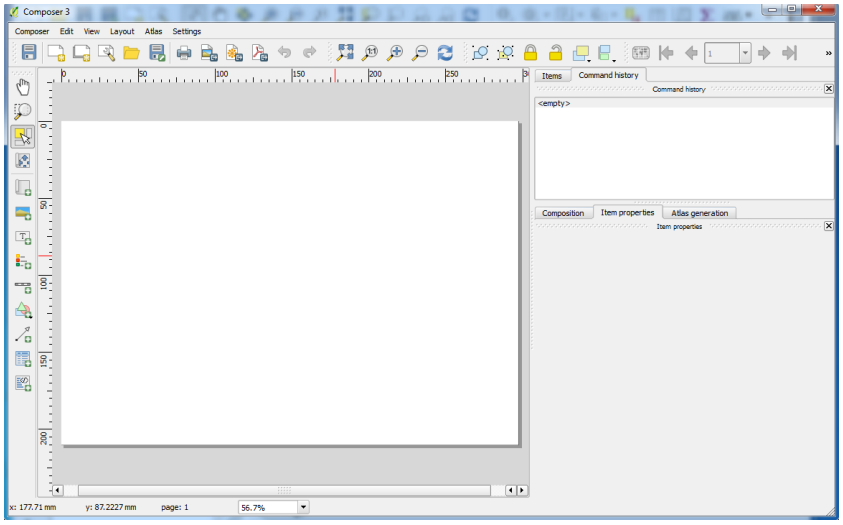


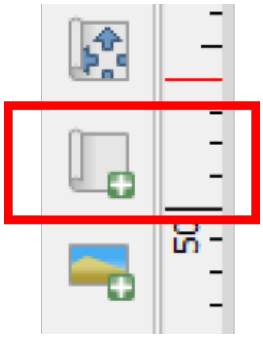
Creating a Print Composer

1. Select "New Print Composer"
2. Give it a title
3. Add elements to the blank canvas
4. Style elements

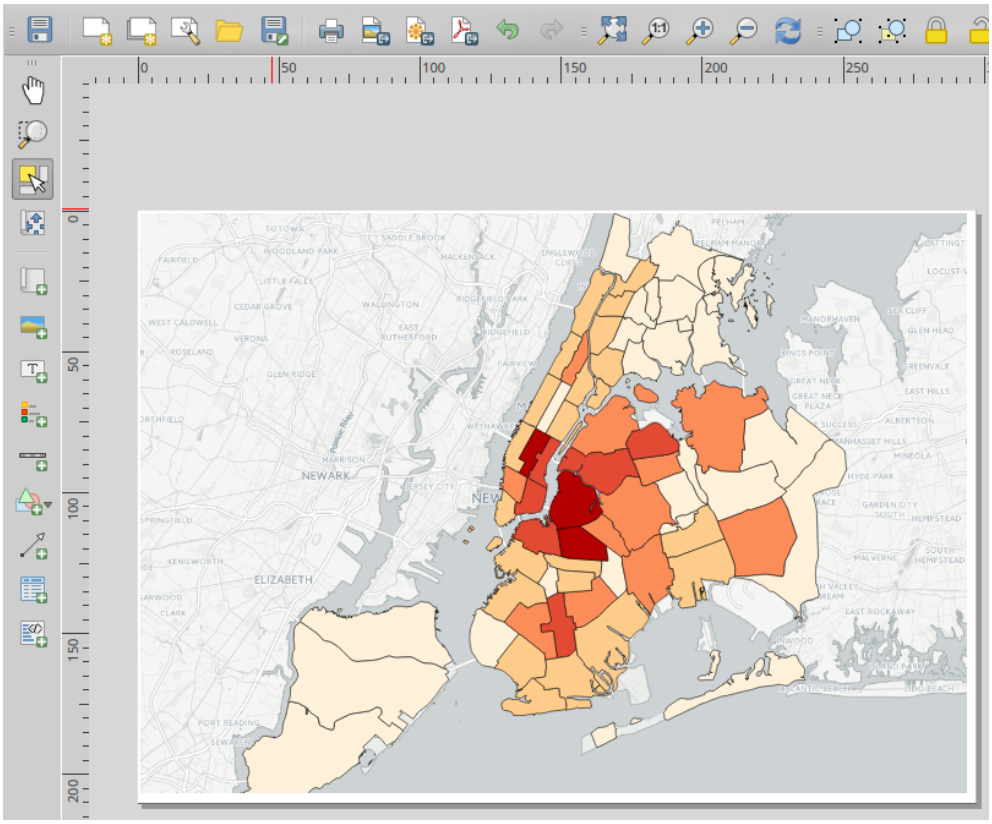




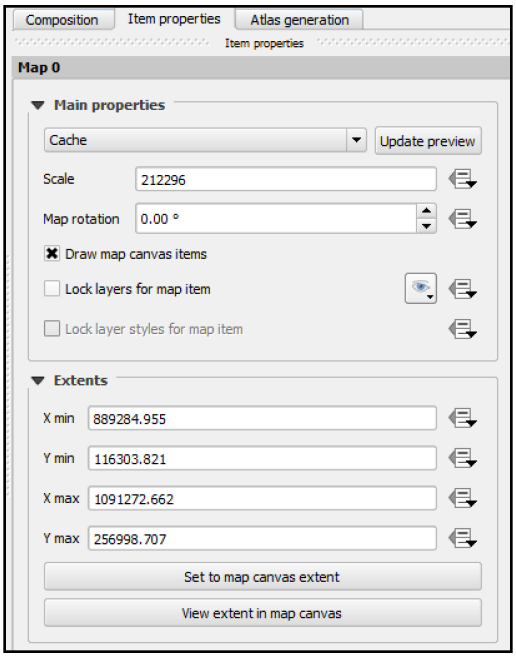




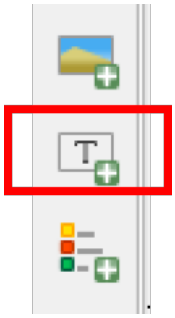
Draw Map Extent



Customize Item Properties



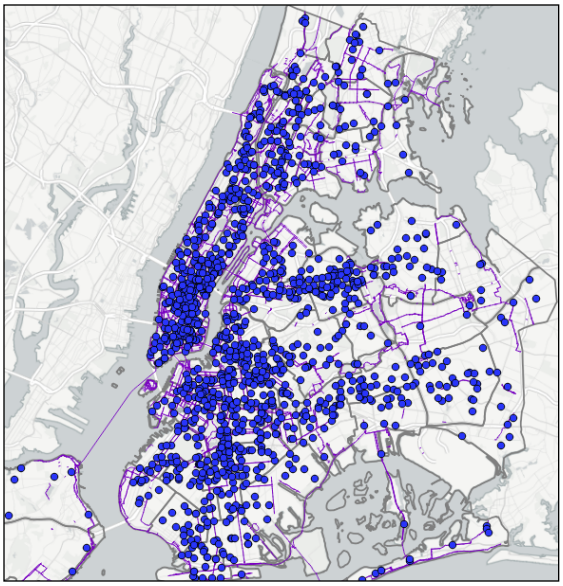
Add Text to Map



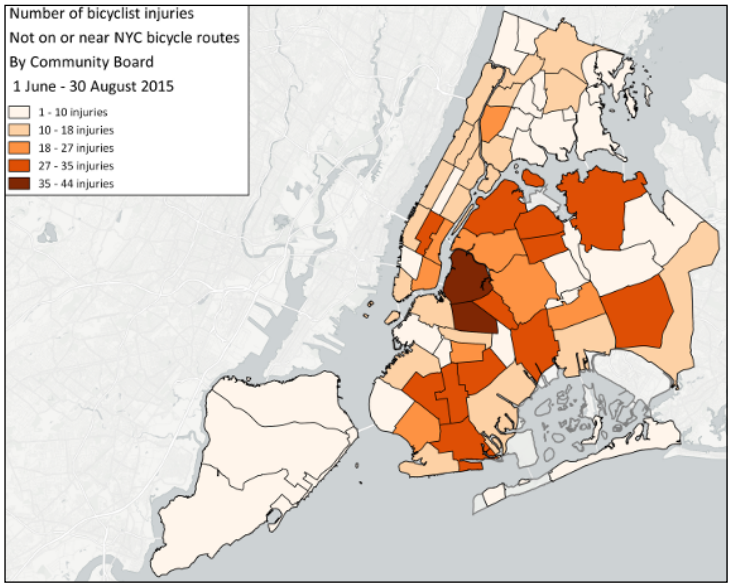
Print Composer Problems

- Map extent -> "Use current map extent"
- Moving map around -> Adjust with arrows
- "North" arrow -> need to manually align

Exercise 3 - Which community board has the highest number of bicyclist injuries off of bicycle paths?



Bicyclist injuries, NYC bike routes, and Community Board boundaries



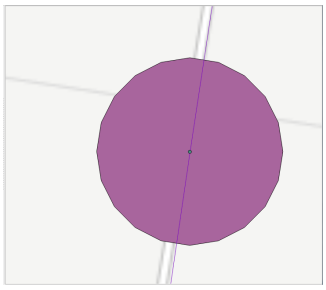
Choropleth of bicyclist injuries off DOT bike routes by Community Board

Steps to answer

1. Buffer bicyclist injuries (points)
2. Select injuries near bike lanes (intersection of buffers with bike lane)
3. Invert the selection to find injuries that happened off of bike lanes
4. Count the number of injuries off of bike lanes by community boards
5. Style the result

Buffering features

- Gives "size" to 2-D features
- Points don't have area
- Lines have length but don't have width
- Allows for analysis of overlap and intersection



Step 0: Load the tables we need

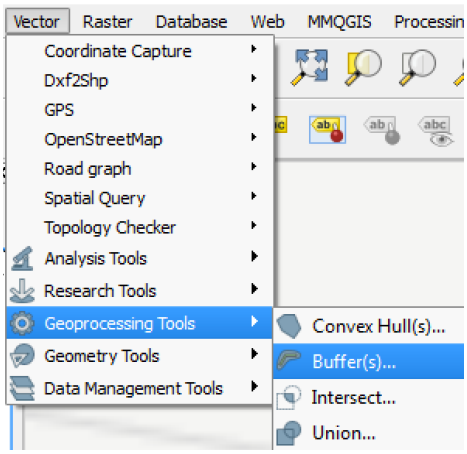
- "injuries"
- "bike_routes_2015"
- "nyc_cd_2263"

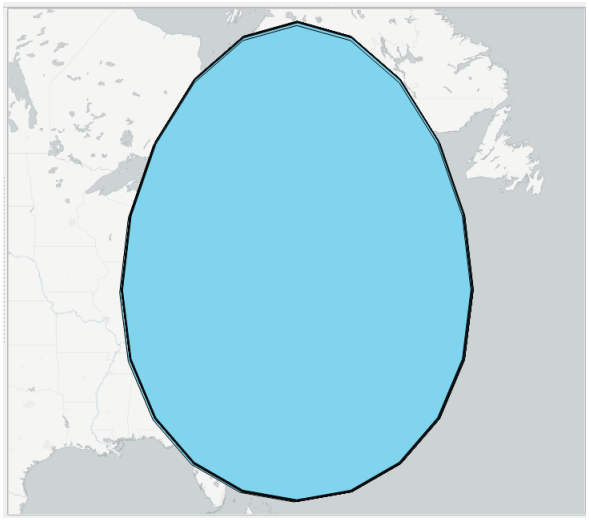
Step 1: Buffering Features in QGIS

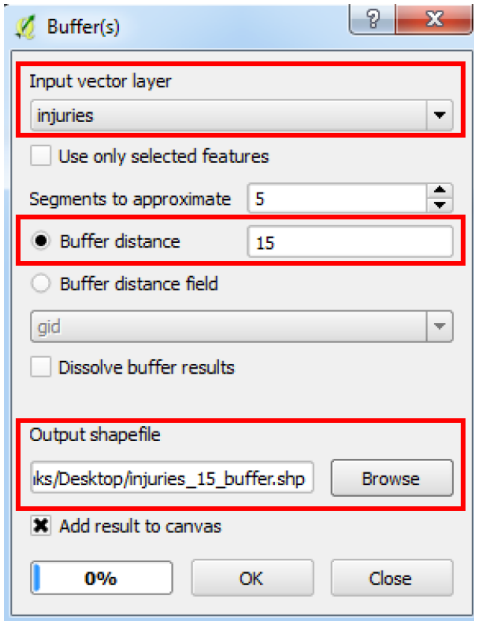
1. Vector -> Geoprocessing Tools -> Buffer(s)...
2. Select the layer you want to buffer on
3. Select the distance you want to buffer
4. Select the output shapefile
5. Click "OK"

SRID Map Units

- Each projection has a base unit of distance
- For WGS 84 it's decimal degrees
- Best to convert to a projection with simple map units
- For NYC, 2263 (NYS Plane Long Island Sound) is best (the map units are in feet)







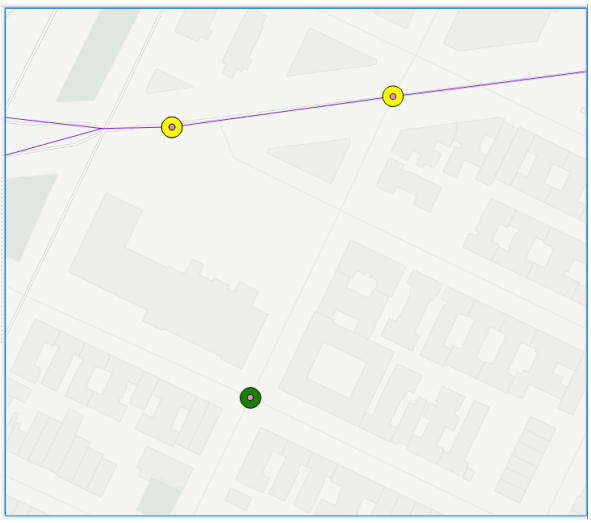
Reprojecting Data

- Right click the "injuries" layer and select "Save As..."
- Select the new projection (2263)
- Give the file a descriptive name
- Click OK

NOTE: we're saving data from the database locally on our computer when we reproject data this way

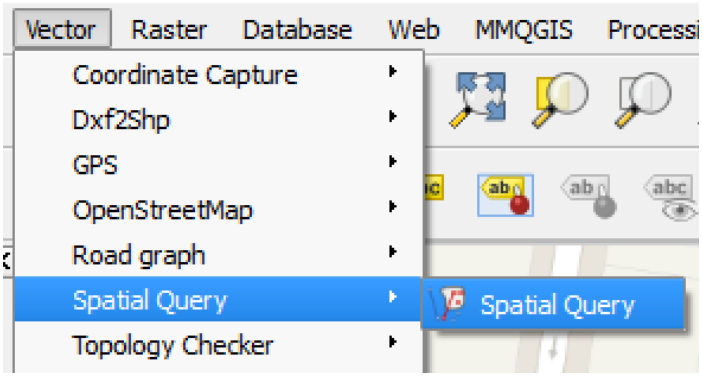
Buffer your features again, this time using the reprojected layer

Buffered injuries

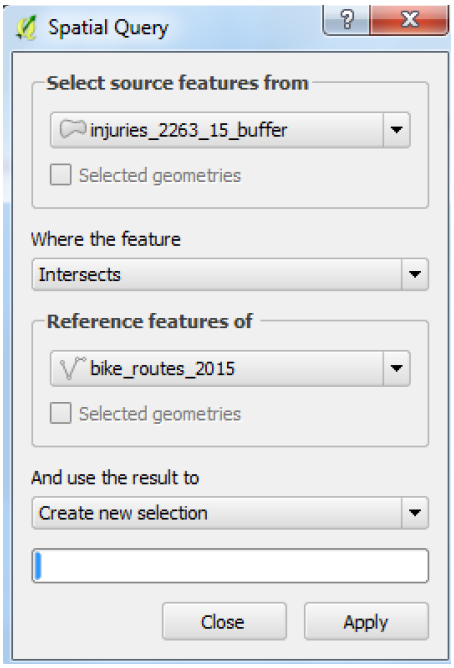


Selected buffers that intersect bike lanes

Step 2: Select injuries near bike lanes

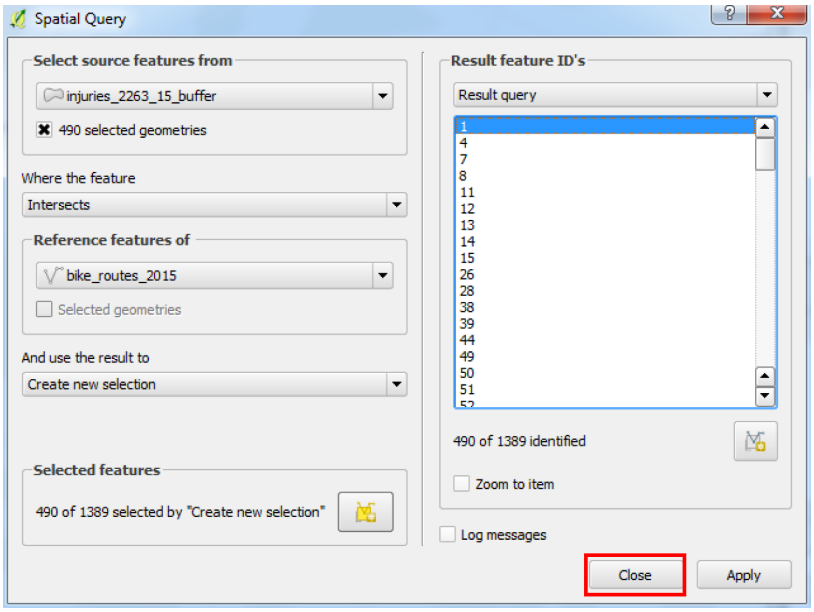


Select by location in QGIS



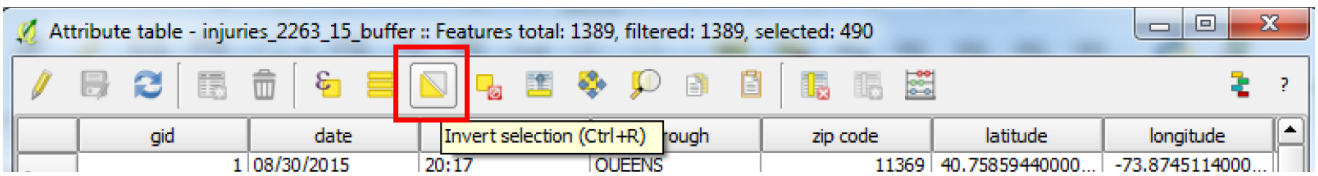
- Select the buffered injuries layer
- Select "Intersects"
- Select the bike routes layer
- Select "Create new selection"
- Click "Apply"

QGIS will now find all the injuries that happened near or on a bike route and select just those features

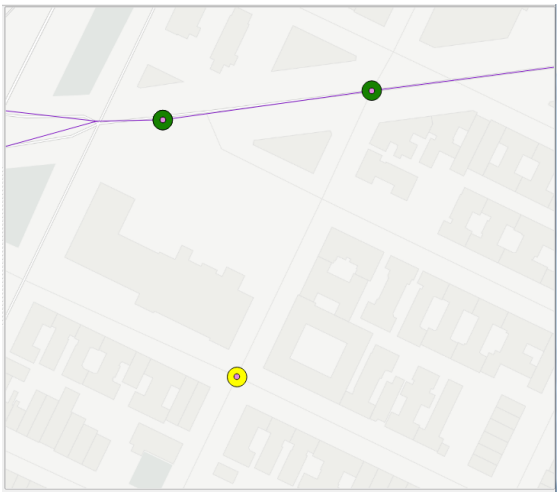


Step 3: Find injuries off of bike lanes

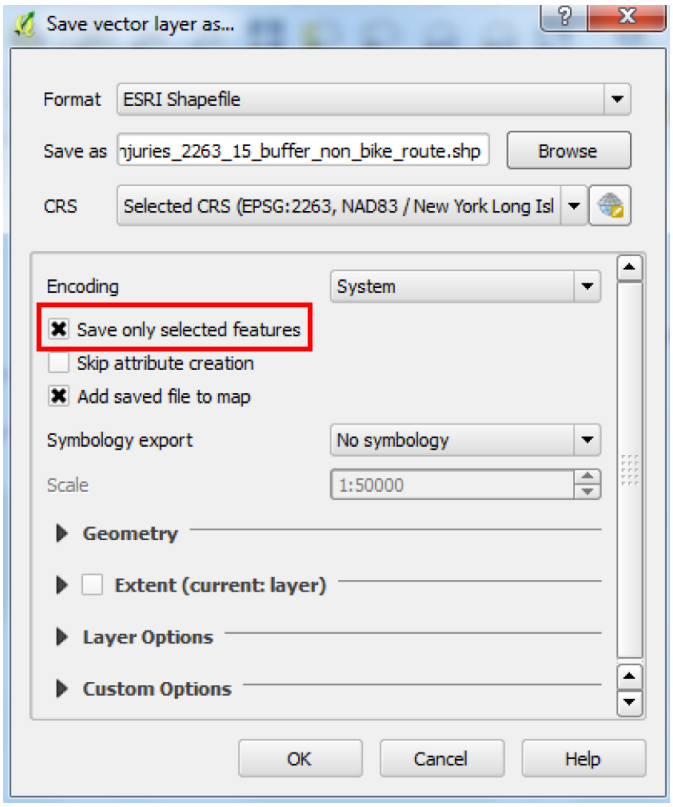
In the attribute table, select "Invert Selection"



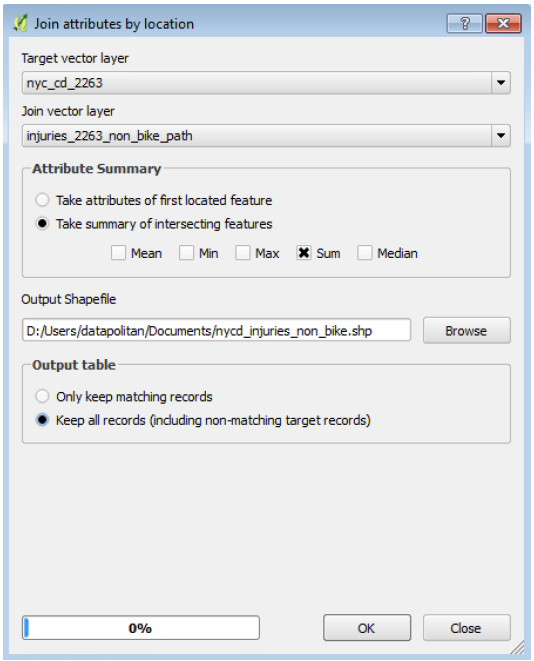
Save the inverted selection as it's own layer



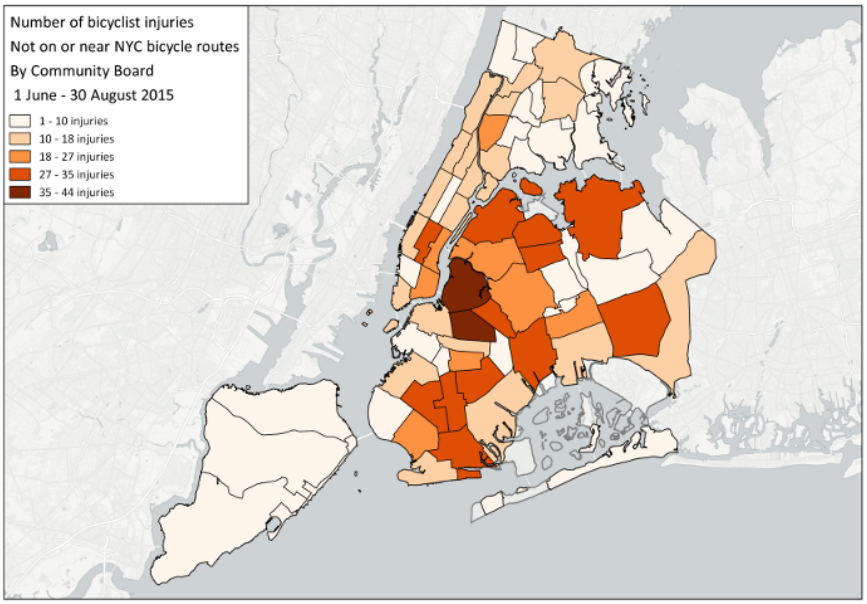
Result after inverting the selection



Step 4: Join the result to the Community Boards (another spatial join)



Step 5: Style the result



Your Turn

- Filter the result for your borough
- Style a map of just the community boards in your borough
- Develop a hypothesis as to why the injuries are happening in those areas

---

Goals for the Afternoon

- Discuss QGIS and other GIS tools
- Provide a basic understanding of open-source technology as it relates to GIS tools
- Introduce Structured Query Language for querying databases
- Practice using SQL for querying and processing spatial data
- Introduce heatmaps and practice creating them in QGIS

Proprietary vs Open Source Software

Open-Source software

- Free to download
- Free to use
- Able to change and customize
- Usually no enterprise support
- Supported by a community

Proprietary software

- Pay to download
- Pay to use (usually with a license)
- Not able to change or customize (legally)
- Usually supported by a team of paid developer

SQL

- Structured Query Language
- The language databases understand

SQL queries

- Can view data
- and change data

We're going to stick to viewing data

SQL Examples

```
SELECT *
FROM injuries
```

In english: "select all the injuries"

```
SELECT *
FROM injuries
WHERE BOROUGH = 'QUEENS'
```

In english: "select all the injuries in Queens"

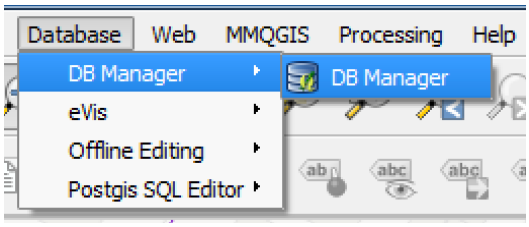
Combining conditions

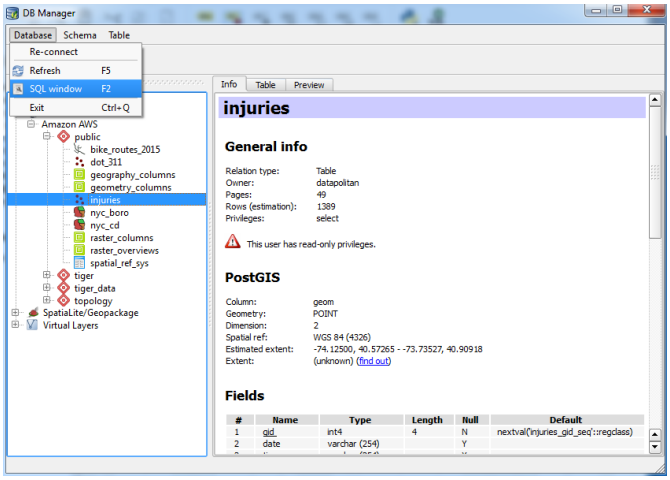
```
SELECT *
FROM injuries
WHERE borough = 'QUEENS' AND contributi = 'Fatigued/Drowsy'
```

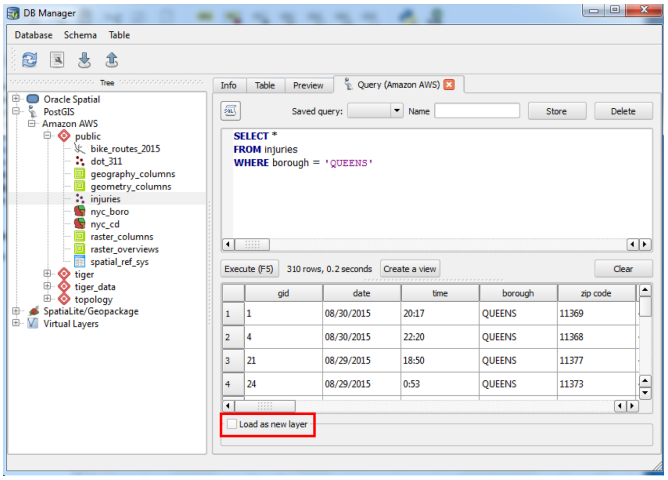
```
SELECT *
FROM injuries
WHERE contributi = 'Outside Car Distraction' OR contributi = 'Fatigued/Drowsy'
```

DB Manager

- One way to see your database connections and tables
- Also a way to run SQL queries on your tables

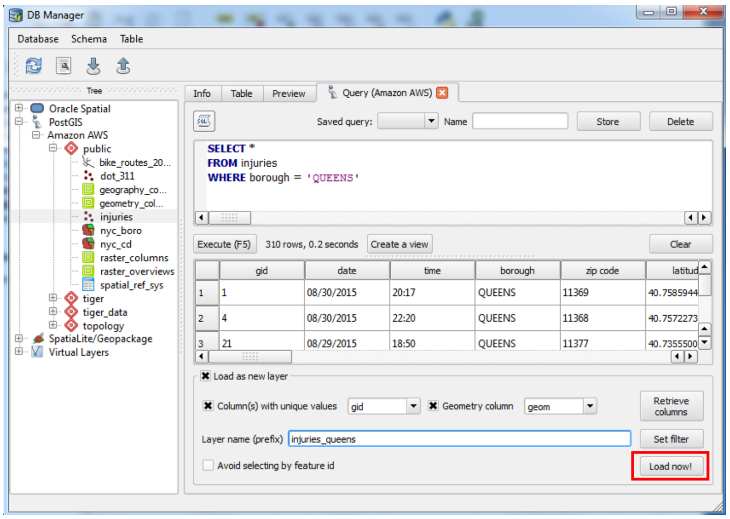






Load Query Result as Layer

- Run query
- Select "Load as new layer"
- Identify the geometry column (should be "geom")
- Give the layer a prefix
- Select "Load now!"



Your turn

- Use DB Manager and SQL to select the injuries for the Bronx
- Add the results as a layer in QGIS

Spatial joins in SQL

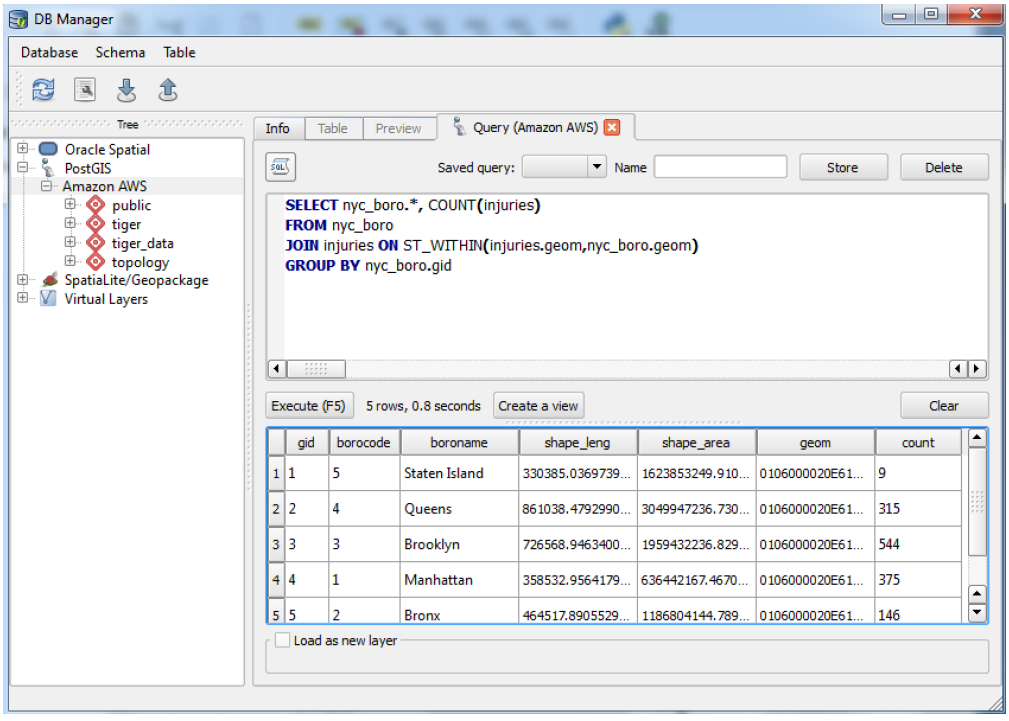
- You can do the same spatial joins you would do in QGIS using SQL
- They give you more control than QGIS does
- And you don't have to create new shapefiles

Join Points to Polygons

- "Select all of the nyc_boro columns and the count of injuries"
- "Use the nyc_boro table"
- "Join with injuries, associate injuries with the nyc_boro they are within"
- "Count by borough"

```
SELECT nyc_boro.*, COUNT(injuries)
FROM nyc_boro
JOIN injuries ON ST_Within(injuries.geom, nyc_boro.geom)
GROUP BY nyc_boro.gid
```

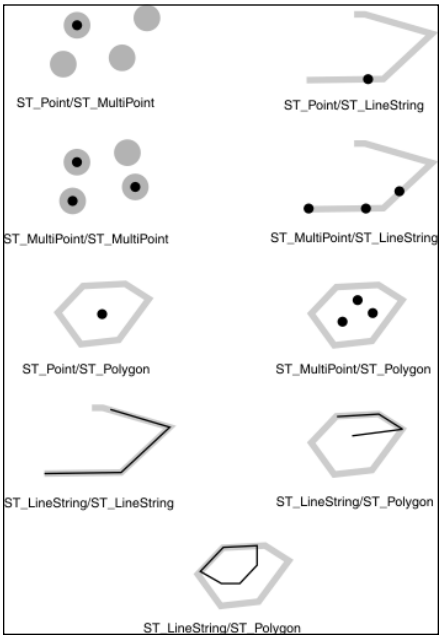
Joining Tables in SQL



Query to join bicyclist injuries to boroughs, similar to what we did with QGIS for Community Boards

ST_Within

- A spatial SQL function
- Similar to other SQL conditions (borough = 'QUEENS')
- We're checking whether the geometries of one table are within geometries from another



ST_Within for different spatial objects

Spatial SQL functions

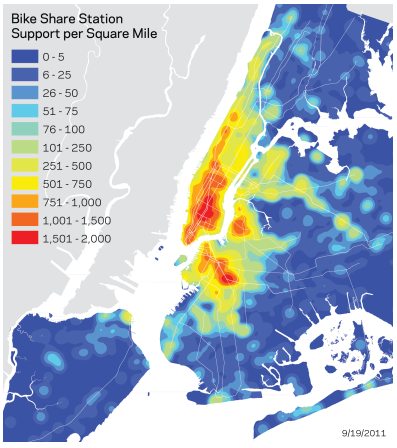
- There are many more you can use
- We're only covering ST_Within today
- But we will give you resources for finding other spatial SQL functions

Your turn

- Use DB Manager and SQL to do a points in polygon join to count injuries in community districts (table "nyc_cd")
- Add the results as a layer in QGIS
- Style the community districts by the number of injuries

Heatmaps

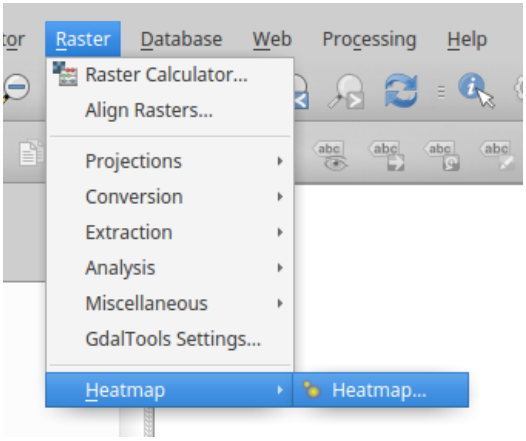
- One way of mapping density when you have many points
- An alternative to choropleth maps if you want to show a little more nuance

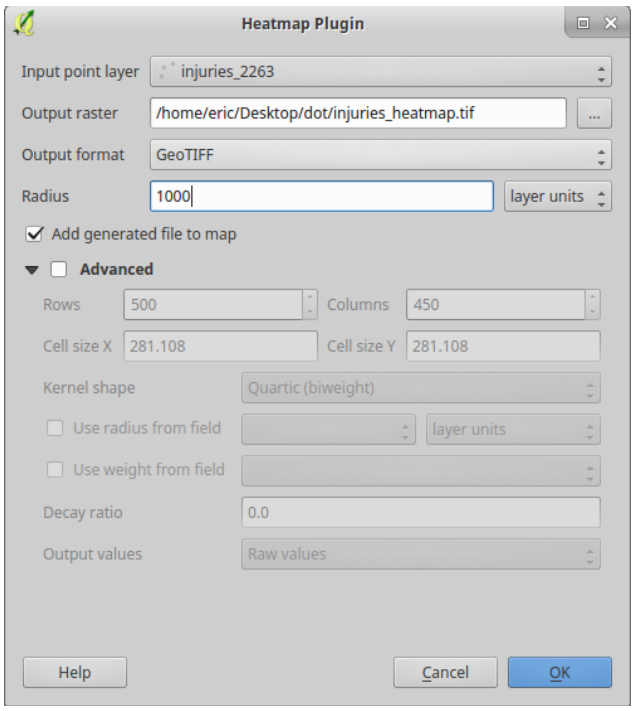


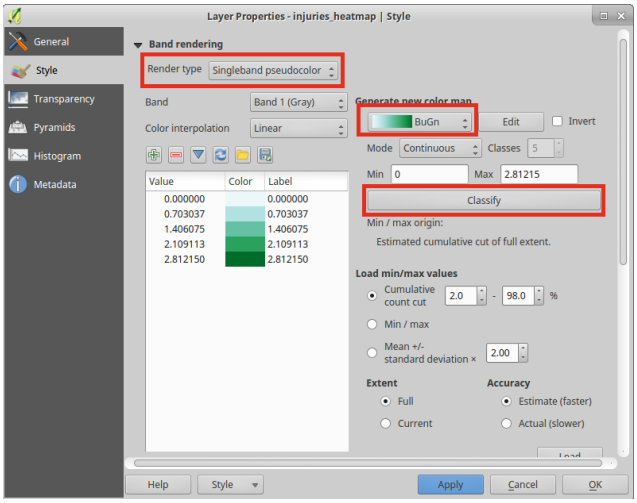
Source: http://a841-tfpweb.nyc.gov/bikeshare/2011/09/20/station-suggestion-recap/

Heatmaps in QGIS

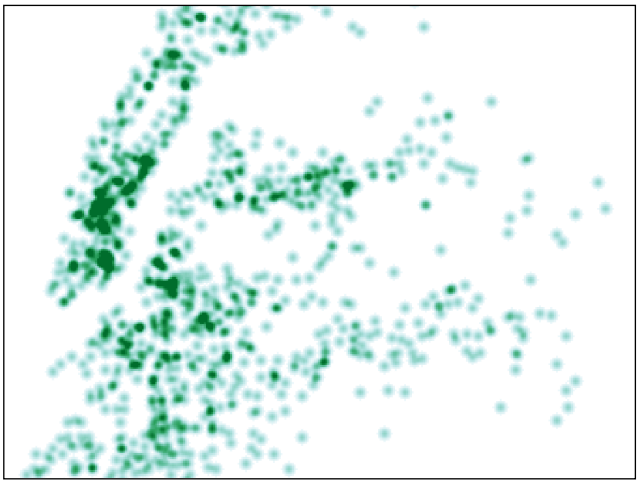
- Use the Heatmap plugin
- Create a raster (image) showing the density of the points
- Style the raster







Style dialogue for heatmaps



Style result without a basemap

Your turn - Putting it all together

- Filter DOT 311 requests (dot_311) by borough and by type
- Join filtered requests to community districts
- Create a choropleth of the number complaints in each community district
- Create a heatmap of complaints in your borough
- You can do this in SQL or QGIS (your choice)
- We'll be around to help
- Hint: Be thinking about the problems we've run into with joining layers

Spatial data formats we didn't cover

- JSON and GeoJSON -> used for digital data exchange and online mapping applications
- XML and KML -> used by Google Earth and other Google-related mapping products

Geocoding (convert addresses to spatial locations)

- You'll want a CSV with the street address, city, and state in separate columns
- Enable the MMQGIS plugin

Resources

Books

- QGIS Map Design
- Learning QGIS 2.0
- PostGIS in Action

Websites

- http://postgis.net/documentation - PostGIS documentation
- http://gis.stackexchange.com/ - Online Q&A site for all things GIS
- http://docs.qgis.org/2.8/en/docs/ - QGIS documentation
- http://www.qgistutorials.com/en/index.html - Great tutorial site for QGIS
- https://anitagraser.com/ - Great blog on QGIS features and news

Richard Dunks, Co-Instructor
- richard@datapolitan.com
- @rdunks1/@datapolitan

Eric Brelsford, Co-Instructor
- ebrelsford@gmail.com
- @ebrelsford

Clayton Hunter, TA

Reminders

- Know your layer projection
- Always test your queries when you filter
- Use descriptive filenames
- Keep your data organized
- Save your work often (especially your project file)
- When in doubt, save, close, and restart