Introduction to **Geographic Information Systems** (GIS) for Real Estate in NYC

Columbia GSAPP Real Estate Program
Raphael Laude
July 11 2023

Today

- 🗺️ Introduce **GIS** and mapping software
- 🗄️ Highlight **data sources** in New York City
- 🧰 Discuss **common GIS tasks**
- 🏗️ Introduce **soft site analysis**
- 📚 Provide **reference document** for future GIS adventures

🚨

This is an **interactive slide**.

When a green slide is up, we'll pause for you work through the next step.

Today (detailed)

🗺️ Intro to **GIS** 30m

- GIS basics: geometry, projections, file types
- Interface fly-through

🗄️ NYC GIS **data sources** 20m

- Land use and dev.: PLUTO (parcels), ACRIS, Zoning, permit data (?)
- Misc geographic
- Downloading spatial data

🧰 **Common GIS tasks** 30m

- Selecting and inspecting spatial data
- Normalization
- Adding a basemap to your project

🏗️ Introduce **soft site analysis** 30m

Today

- Introduce **GIS** and mapping software
- Highlight **data sources** in New York City
- Discuss **common GIS tasks**
- Introduce **site analysis**

Early spatial analysis
1854 Broad Street cholera outbreak, John Snow

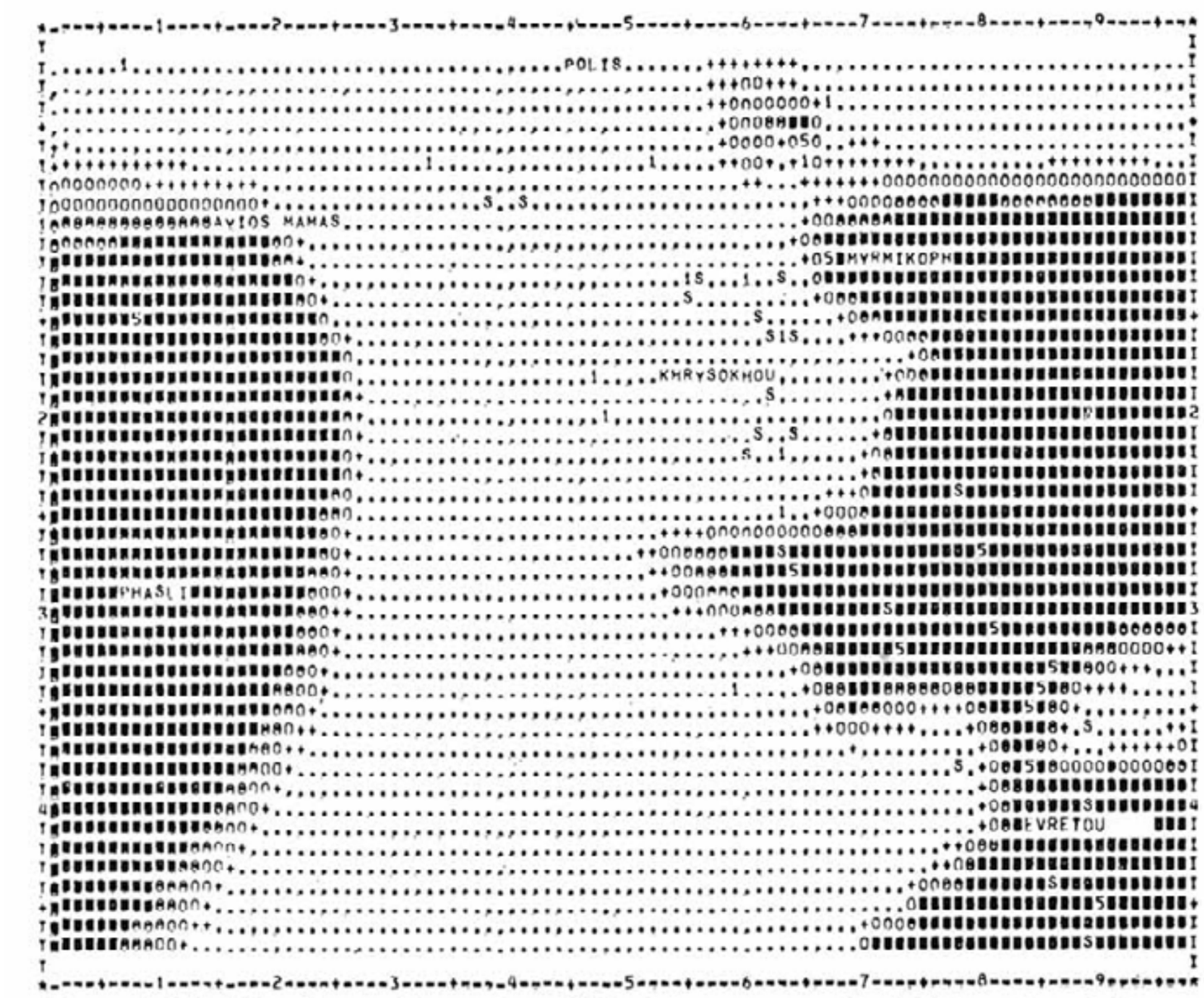


Figure 18 SYMAP of Cypro-Geometric to Hellenistic components in the Khrysokhou drainage.

Key: · Zone I
+, O, θ Intermediate Zone I/II
▮ Zone II
1, 5 Single component on single site locus
S Multiple components on single site locus

The start of modern GIS
Early GIS maps from
SYMAP, William Caraher

Modern, ubiquitous GIS
The Earth Wind Map,
Cameron Beccario

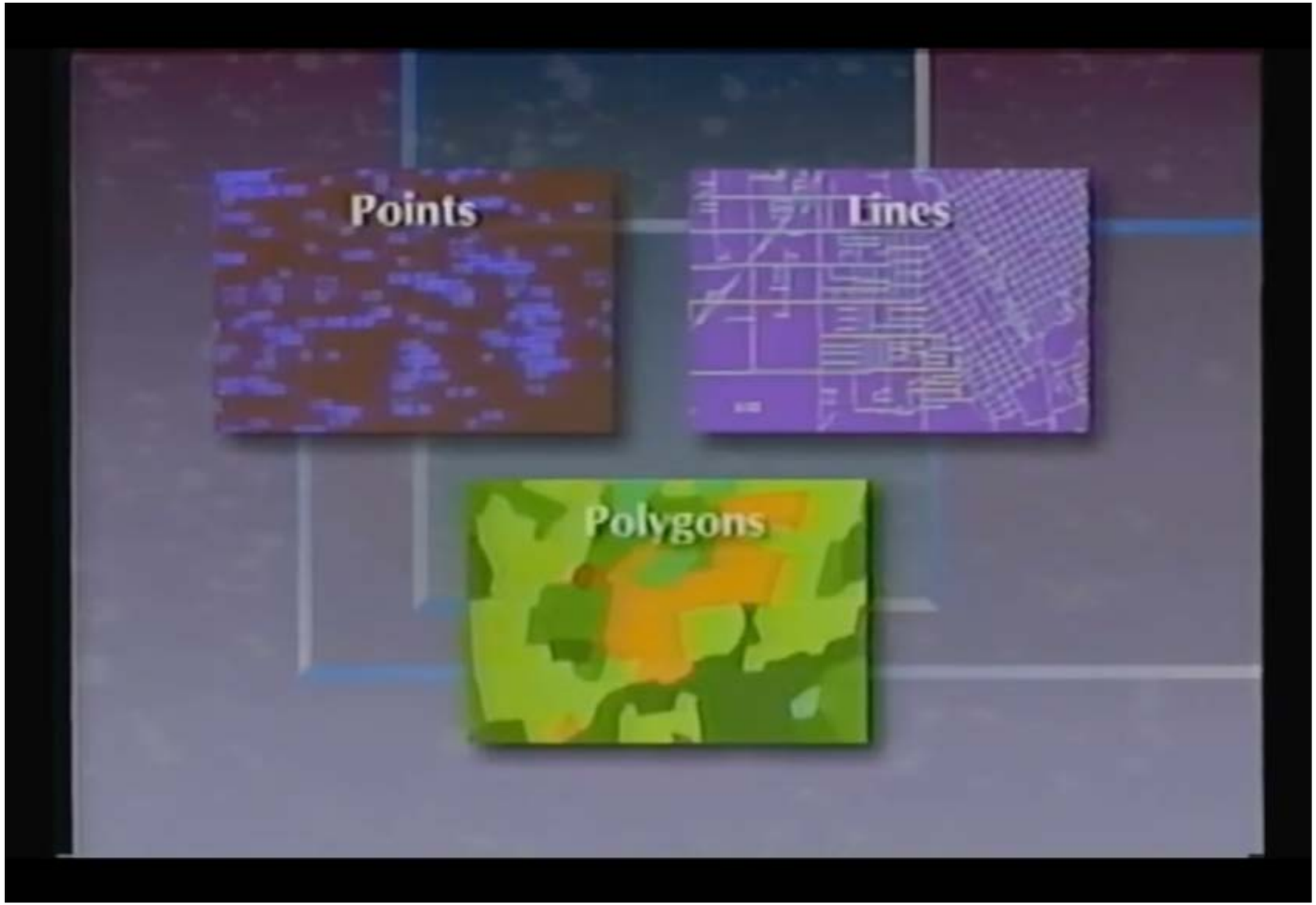
Computational (invisible) GIS
Computers read maps more than we do.

🗺️ GIS

data + geometry

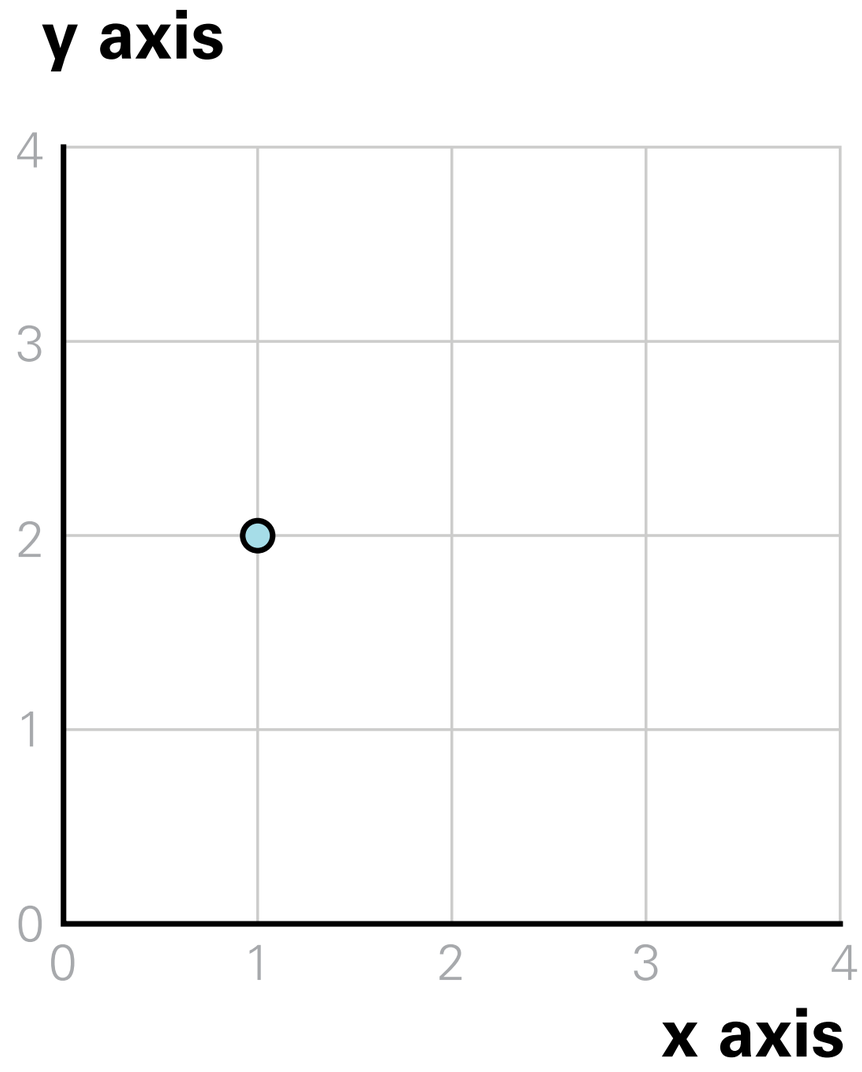
🗺️ GIS

data + geometry
& things you can do with geometry

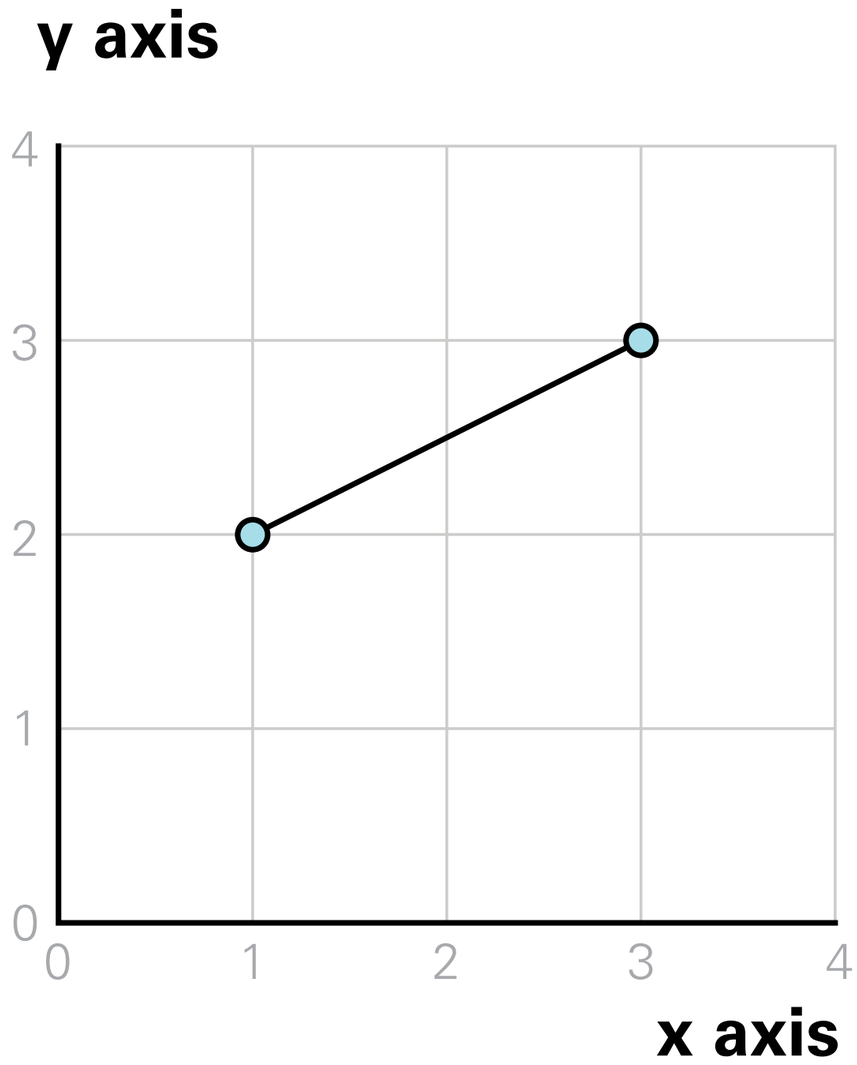


Computer Graphics: GIS Geographic Information System (database, mapping, ARCinfo, ARCview) 1988

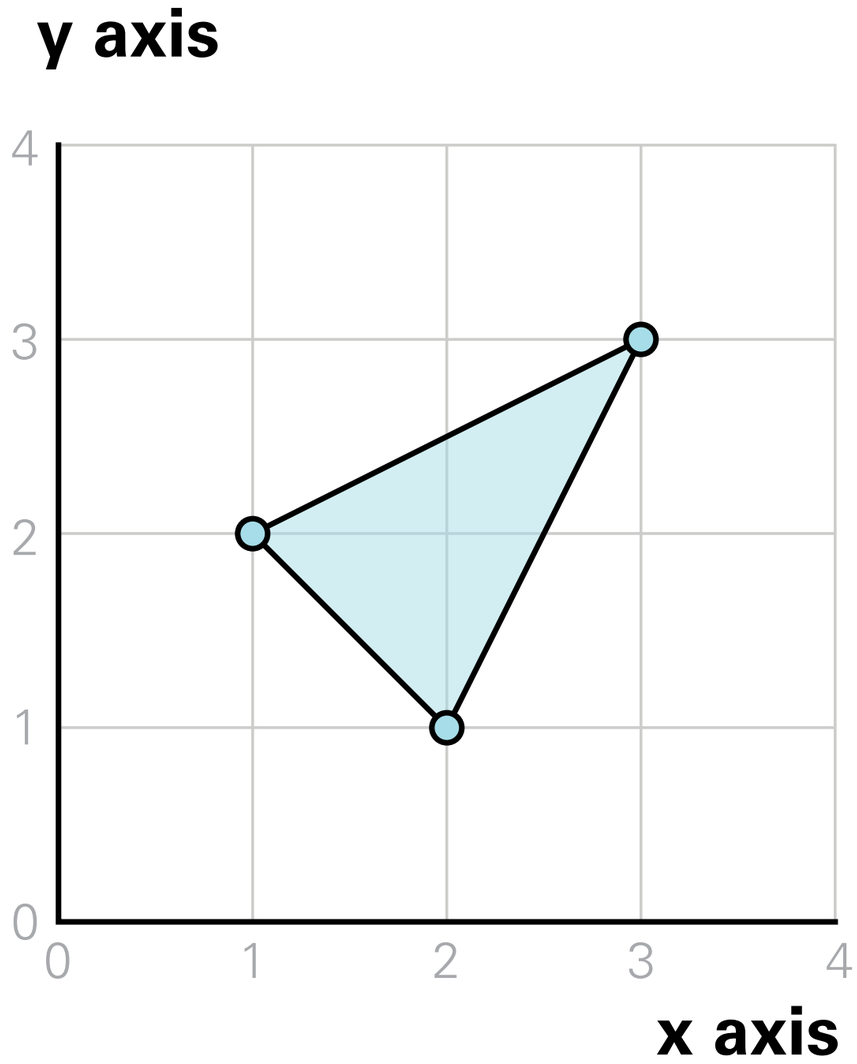
POINT
(1 2)



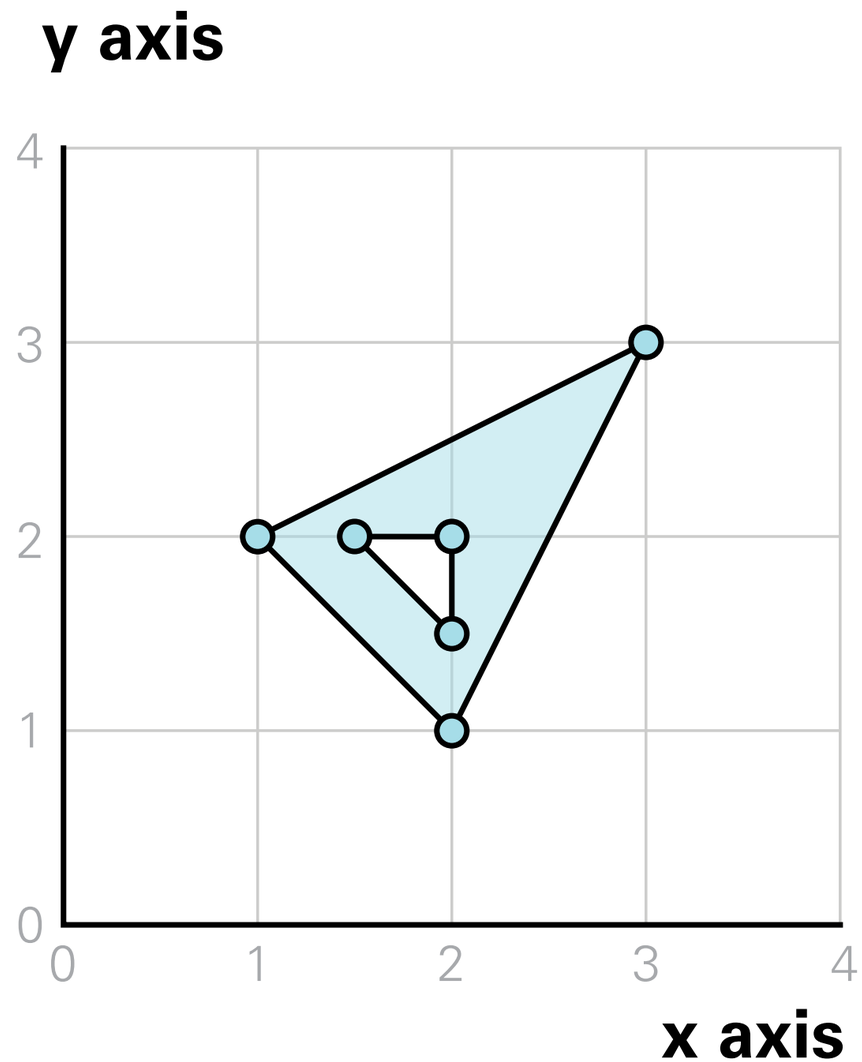
LINESTRING
((1 2), (3 3))



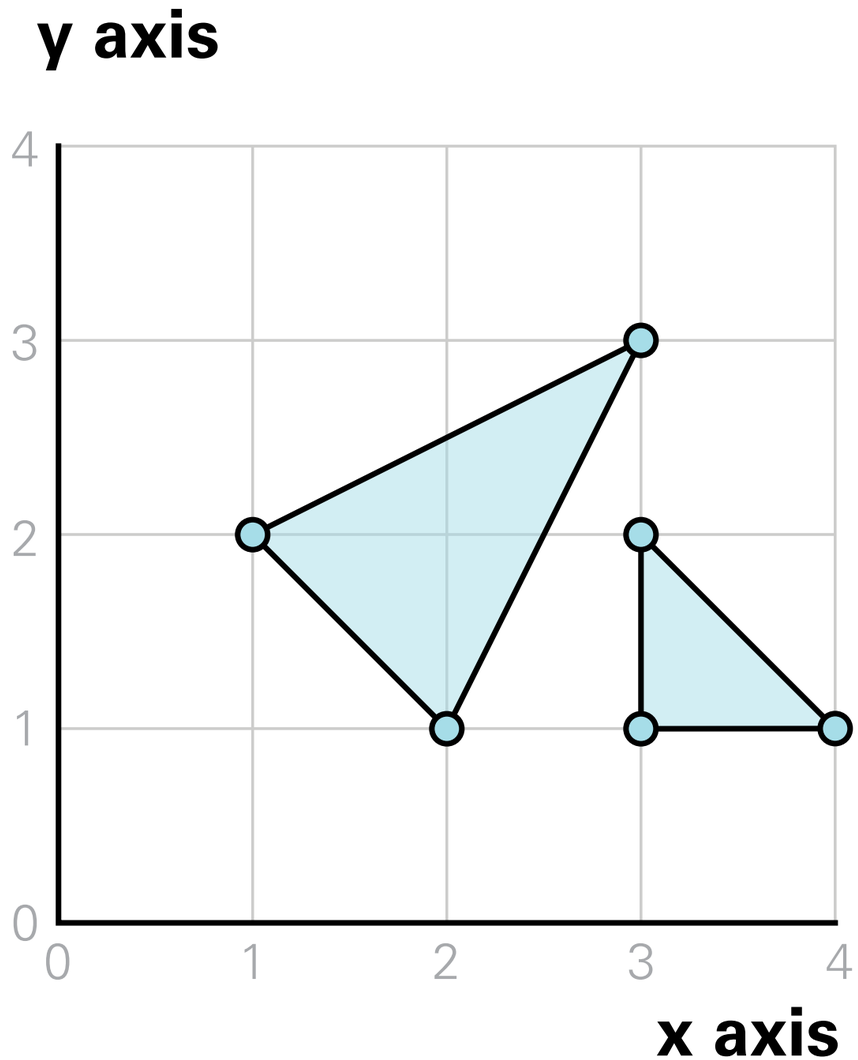
POLYGON
((1 2), (3 3), (2 1),
(1 2))



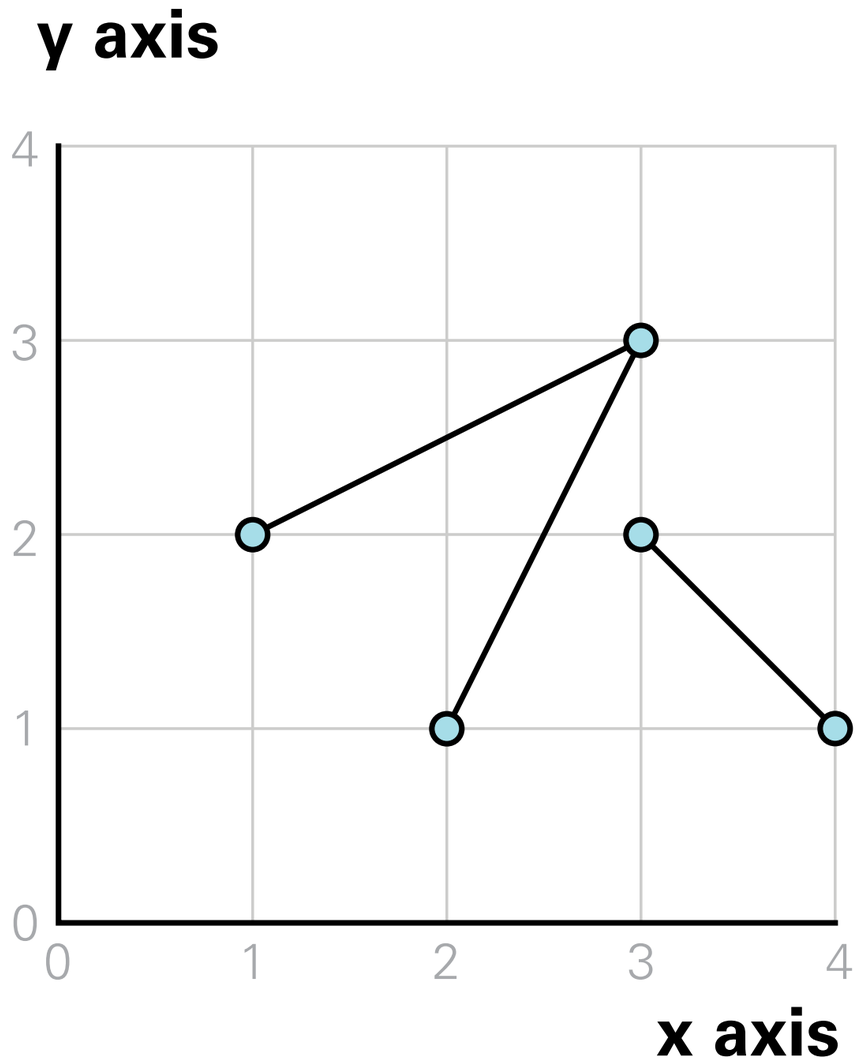
POLYGON
Exterior ring
((1 2), (3 3), (2 1),
(1 2))
Interior ring
((1.5 2), (2 2), (2 1.5),
(1.5 2))



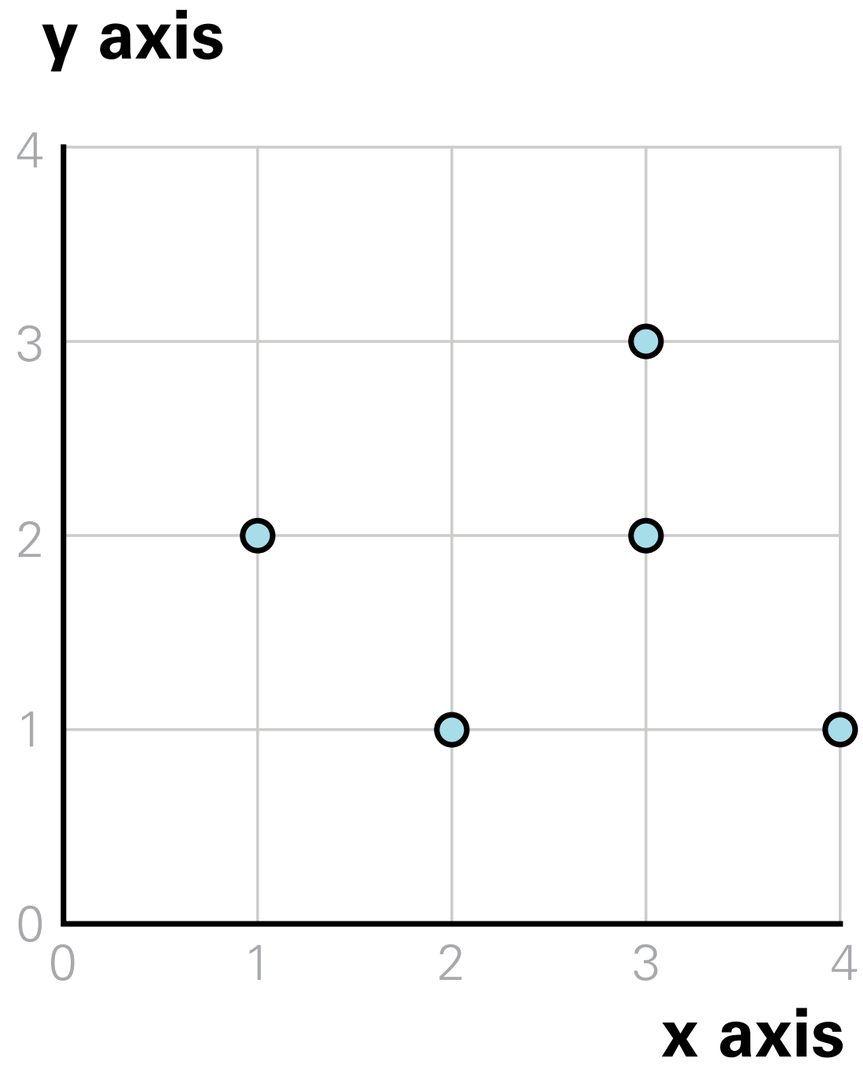
MULTIPOLYGON
((1 2), (3 3), (2 1),
(1 2)),
((3 1), (3 2), (4 1),
(3 1))

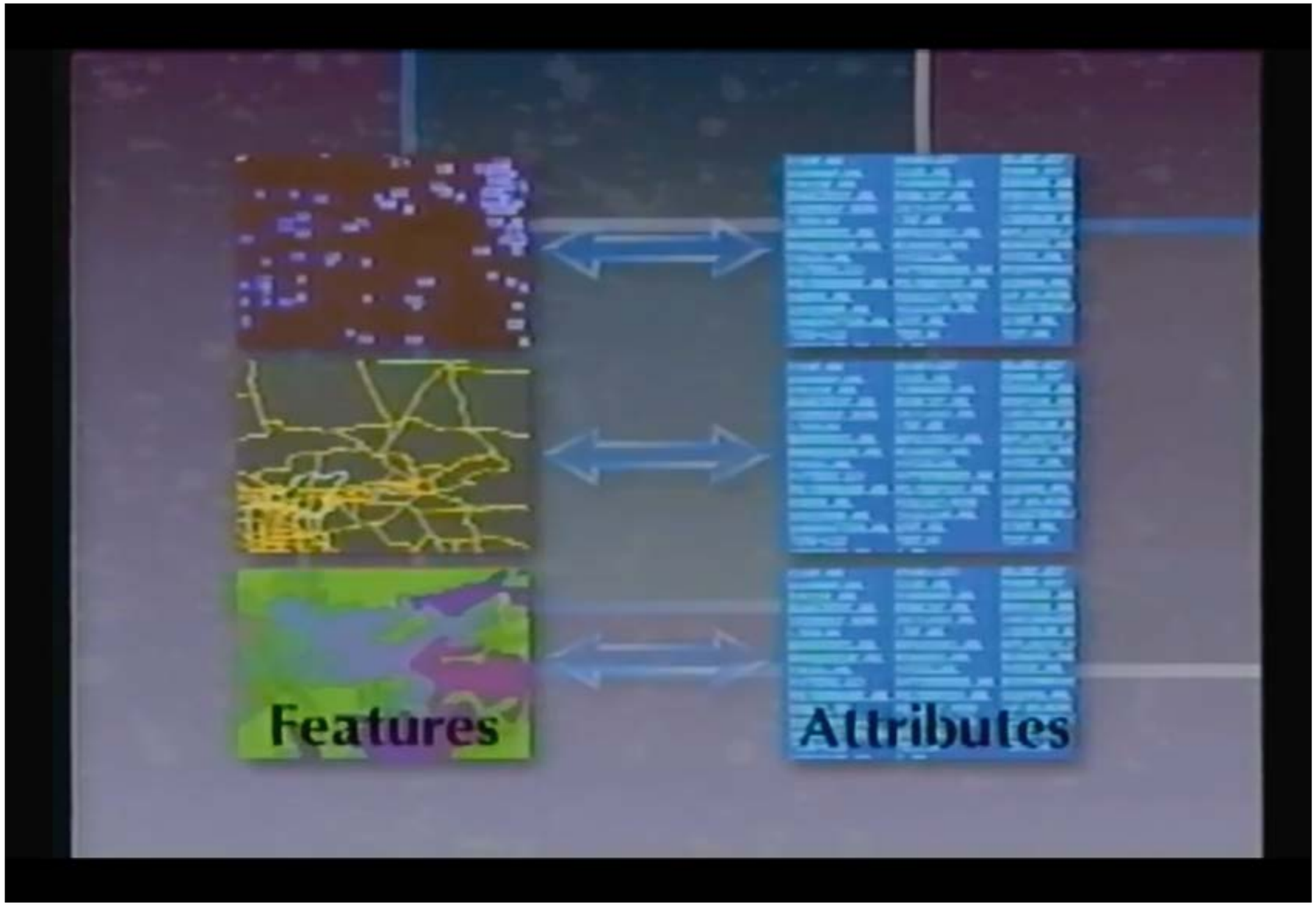


MULTILINESTRING
((1 2), (3 3), (2 1)),
((1.5 2), (2 2))



MULTIPOINT
(1 2), (3 3), (2 1),
(1.5 2), (2 2)



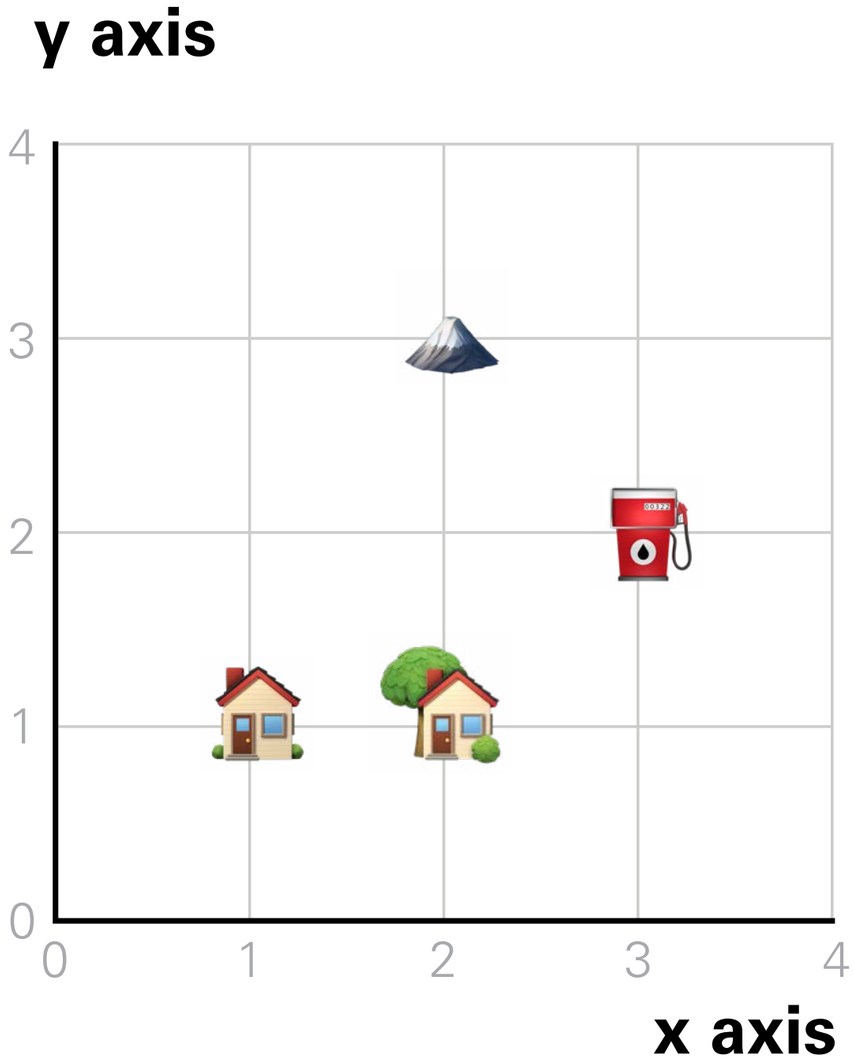


Computer Graphics: GIS Geographic Information System (database, mapping, ARCinfo, ARCview) 1988

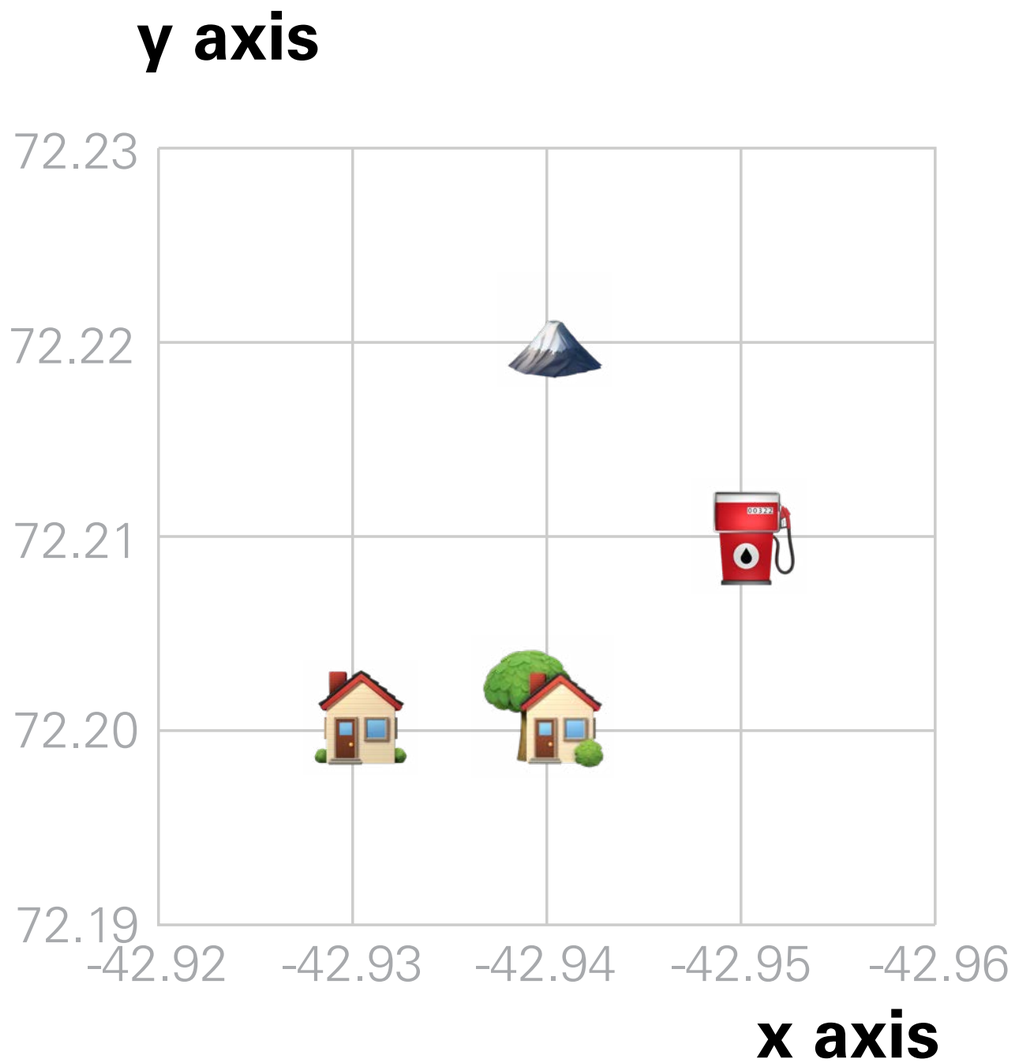
Spatial data

	Description	Geom
🏠	House	Point 1 1
🏡	Garden House	Point 2 1
🗻	Mt. Fuji	Point 2 3
⛽	Fuel pump	Point 3 2

attributes | features



🚨 **Note!**
In practice, geometries are encoded using a **Coordinate Reference System** (CRS).



Coordinate Reference Systems (CRS)
There are two types of CRS:

1. Geographic
2. Projected

Geographic CRSs are used to map data across the **entire planet**.

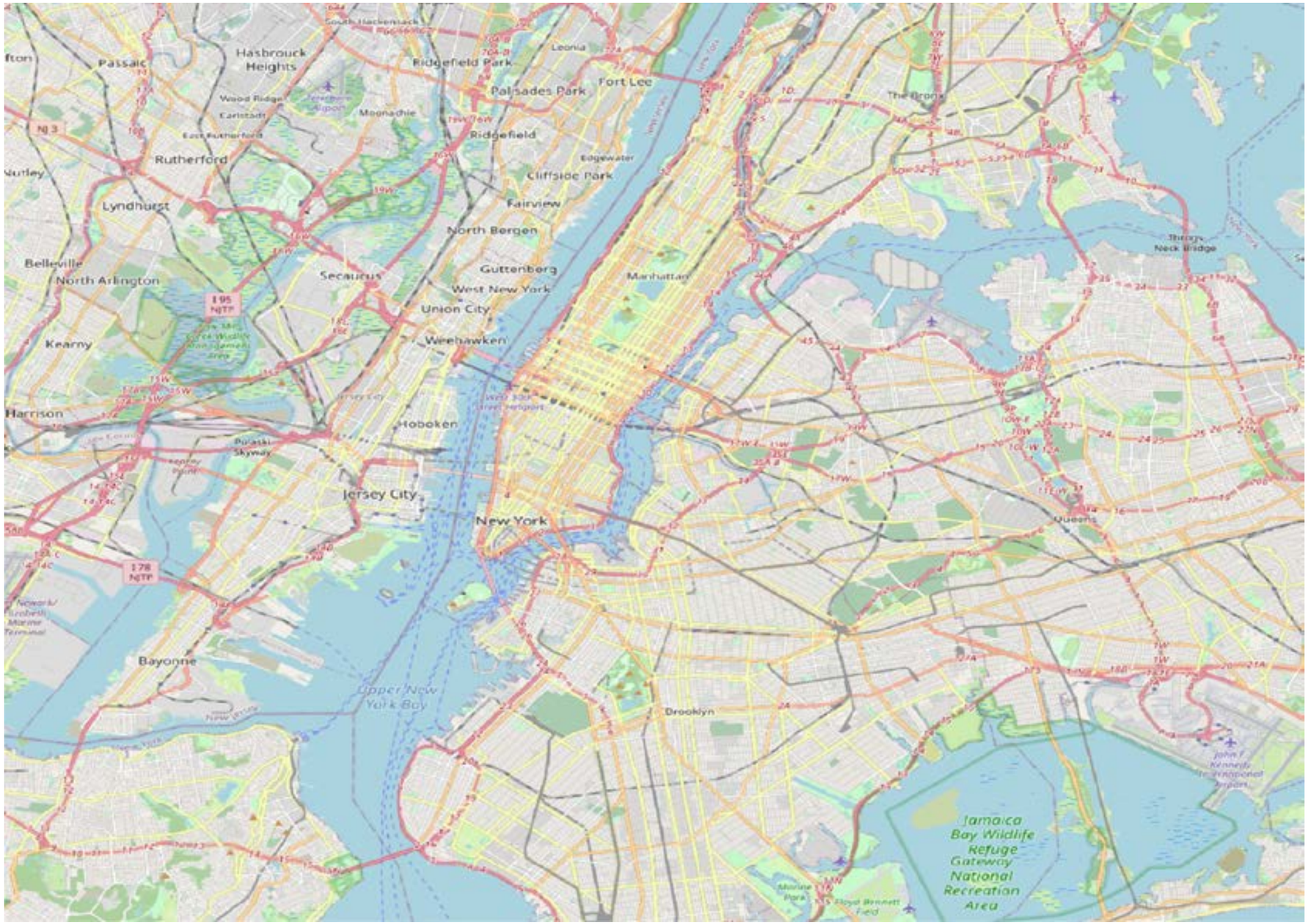
Projected CRSs are used for **specific regions**, to minimize local visual distortion.

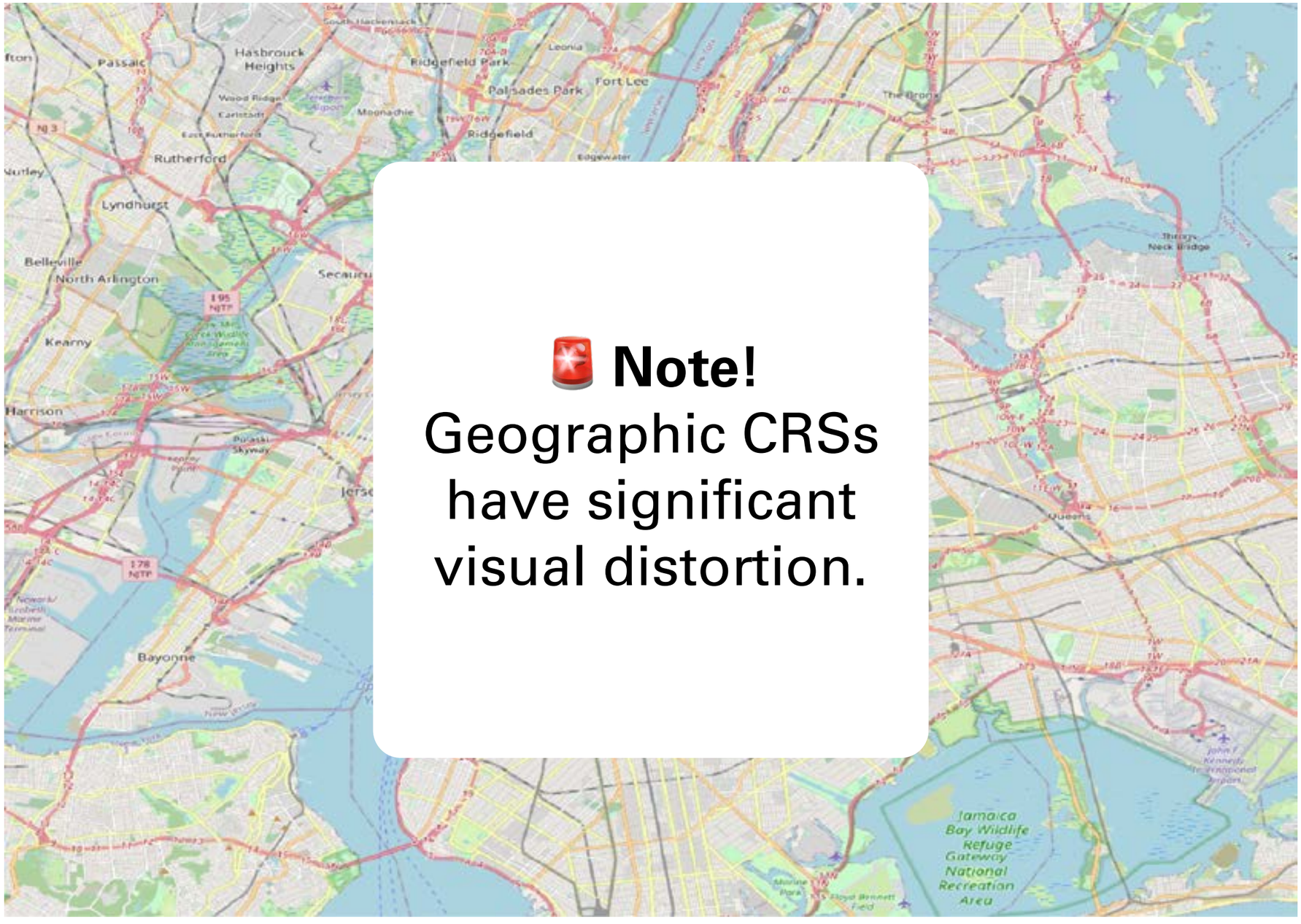
CRSs

Latitude-longitude coordinates used by GPS are part of the World Geodetic System (WGS84) CRS

CRSs

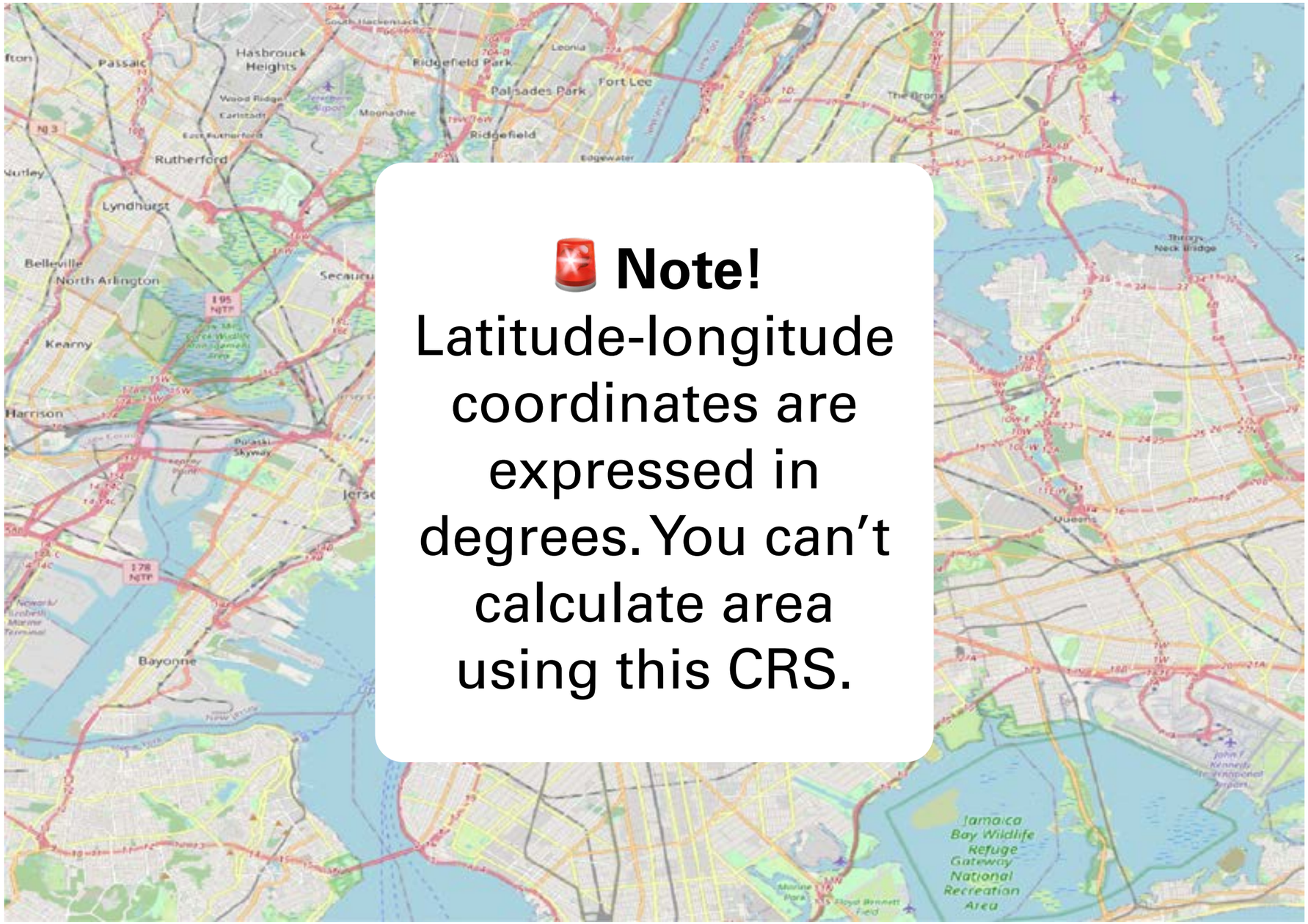
Latitude-longitude coordinates used by GPS are part of the World Geodetic System (WGS84) CRS


Passaic
Hasbrouck Heights
Ridgefield Park
Palisades Park
Fort Lee
Rutherford
Lyndhurst
Ridgefield
Cliffside Park
Fairview
North Bergen
Belleville
North Arlington
Secaucus
Guttenberg
West New York
Union City
Weehawken
Kearny
Harrison
Hoboken
Jersey City
New York
Manhattan
Bayonne
Upper New York Bay
Brooklyn
Queens
Jamaica Bay Wildlife Refuge
Gateway National Recreation Area



🚨 **Note!**

Geographic CRSs have significant visual distortion.

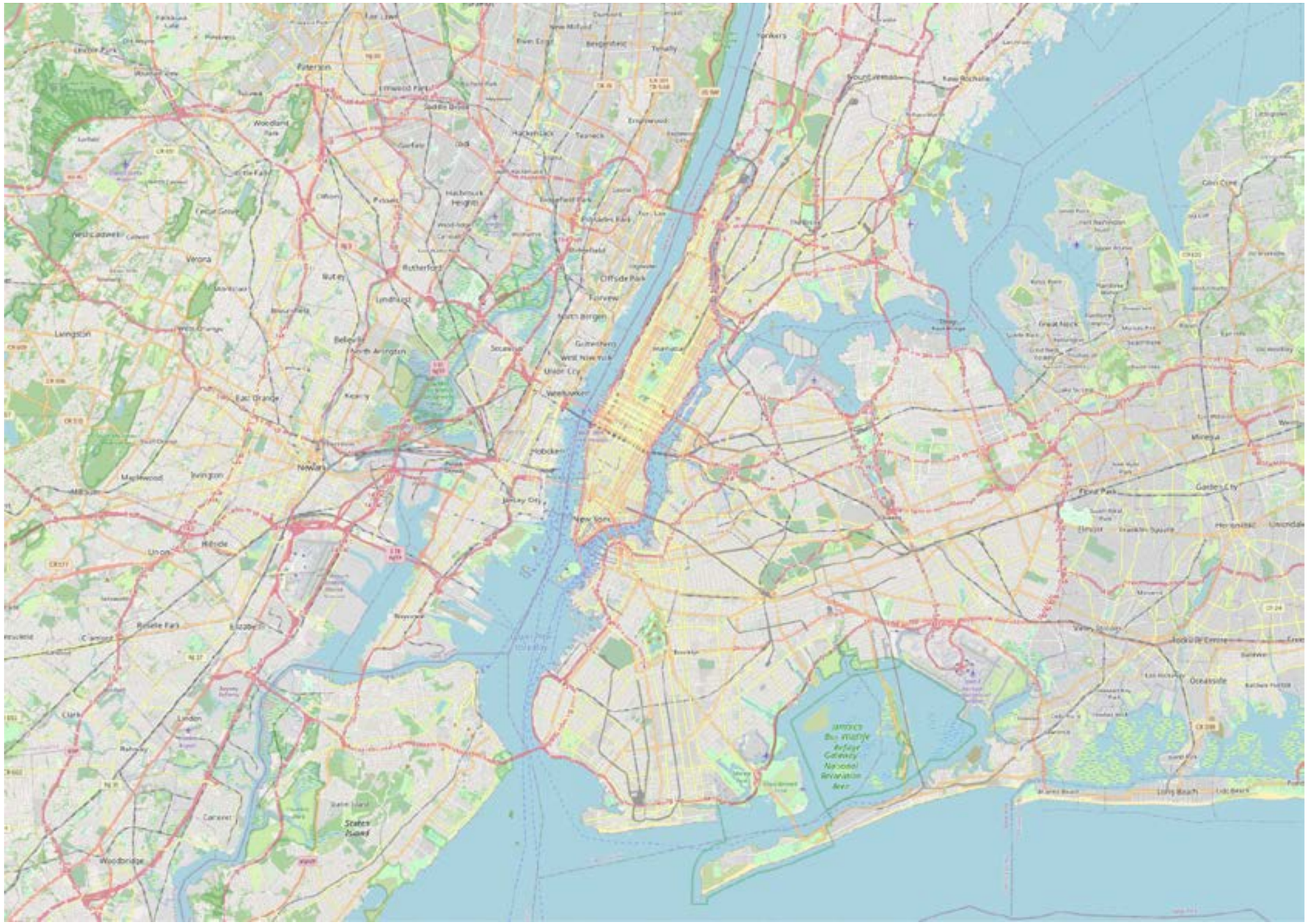


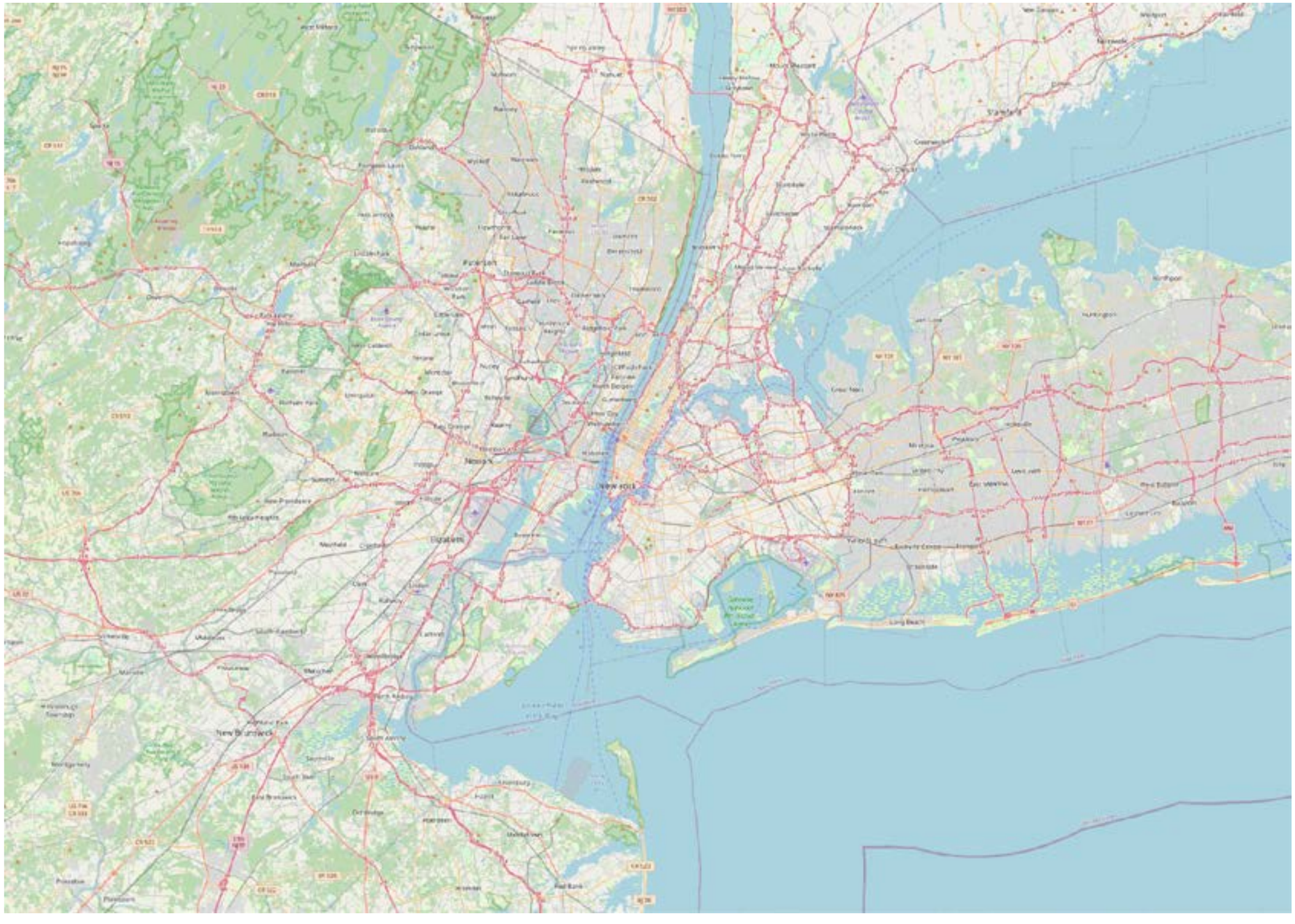
🚨 **Note!**
Latitude-longitude coordinates are expressed in degrees. You can’t calculate area using this CRS.

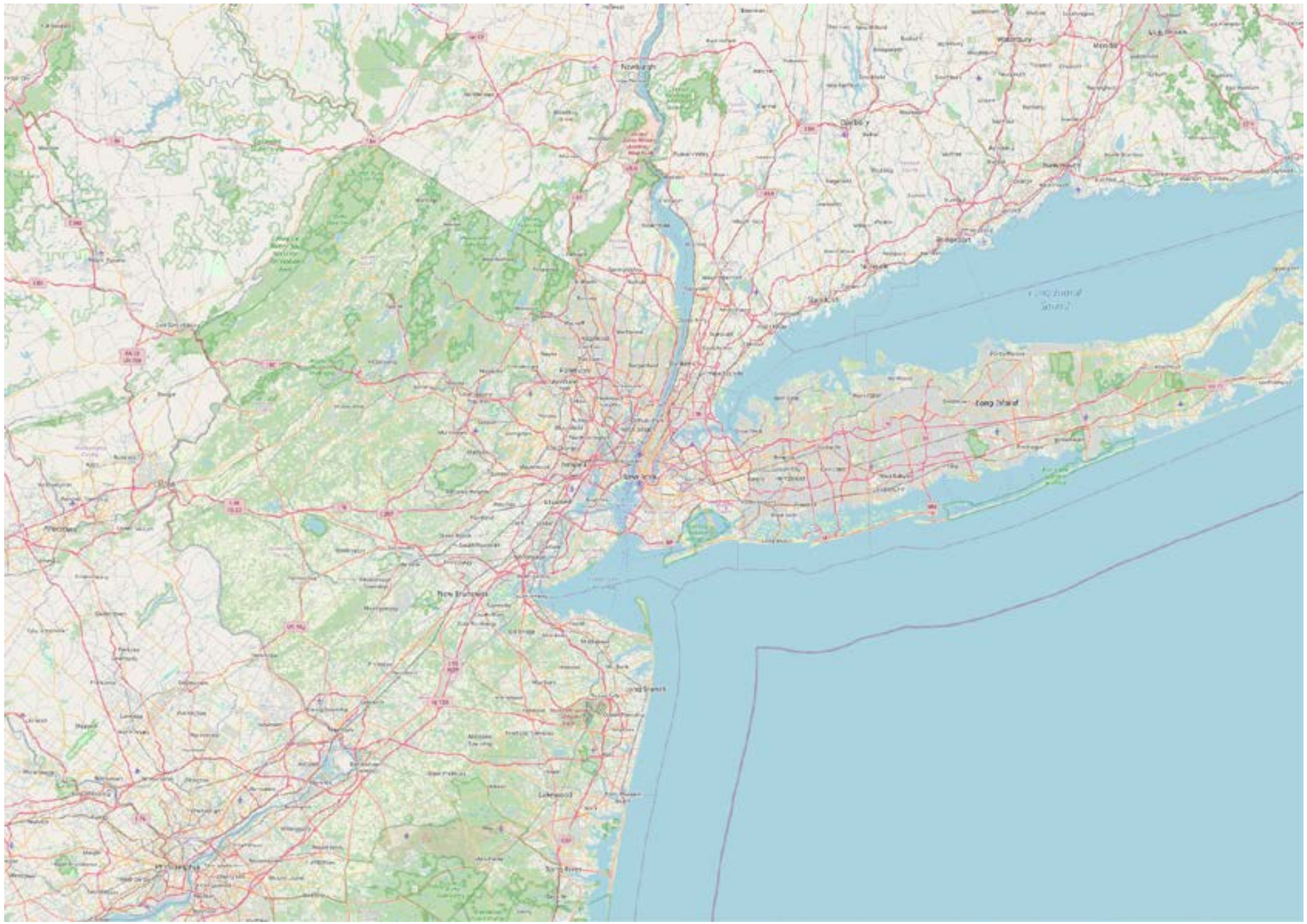
👍 Projected CRSs

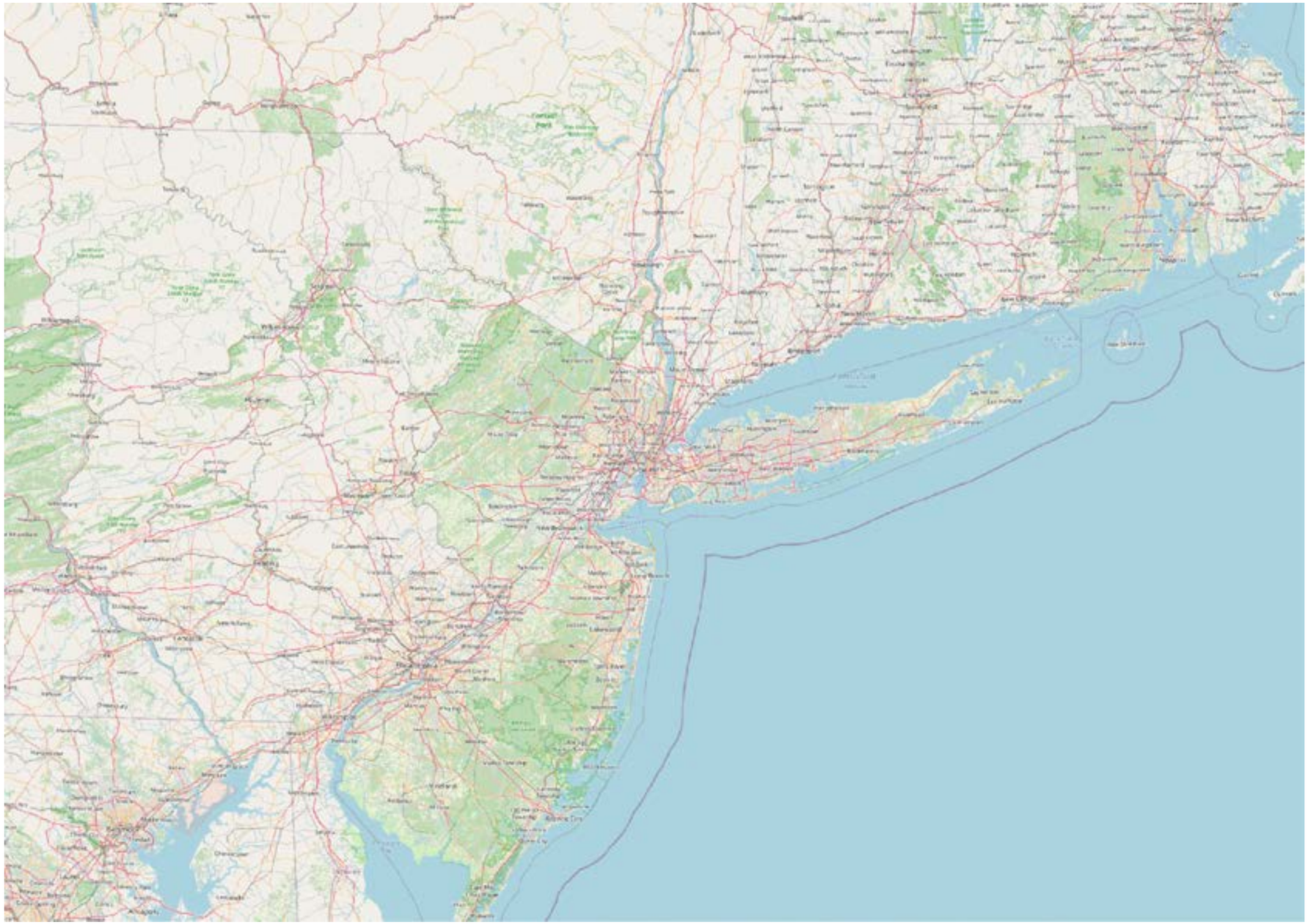
Let’s try mapping NYC with a projected CRS, to minimize visual distortion.

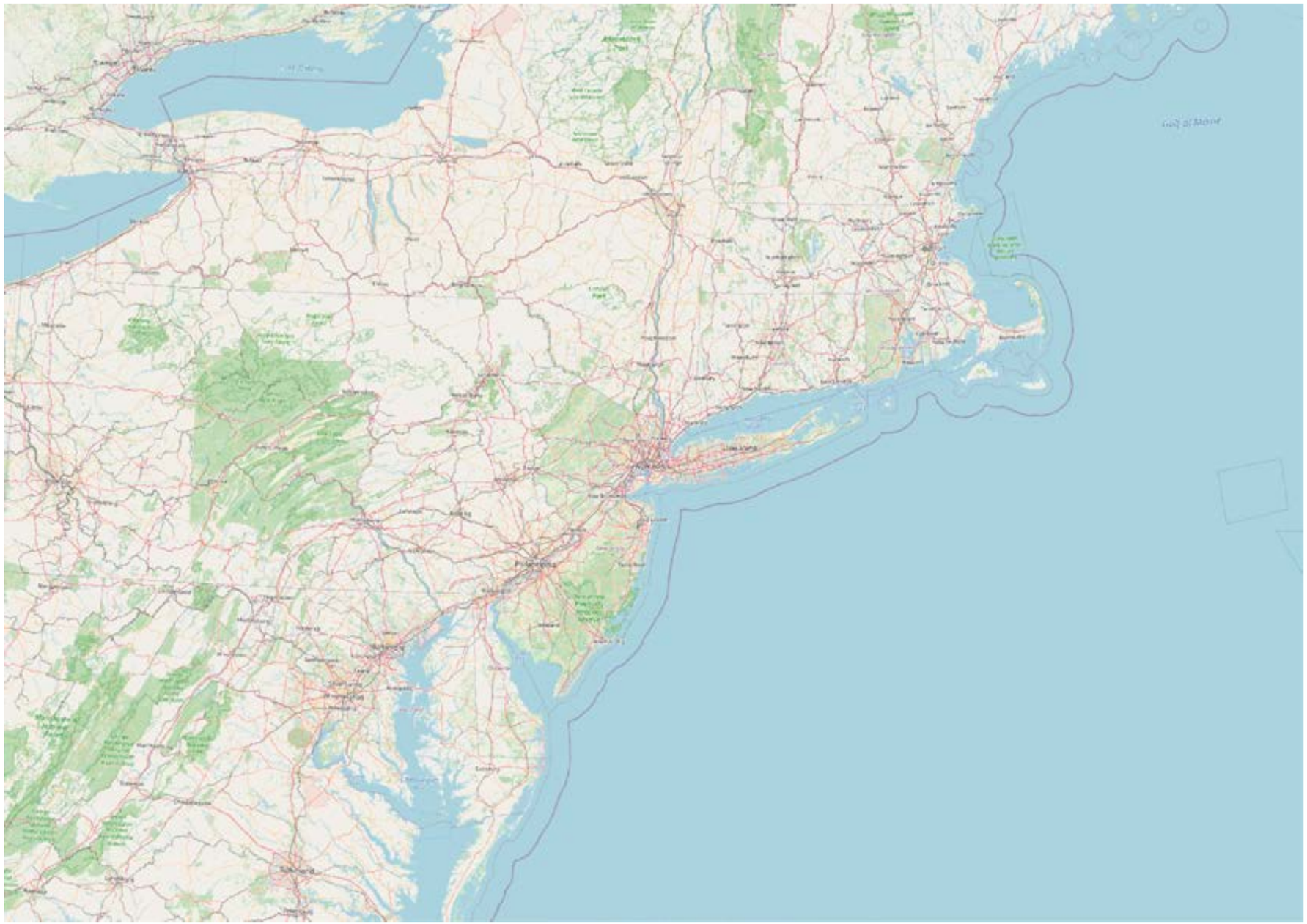
In NYC, use **NAD83 / New York Long Island (ftUS)**. This projection is part of the State Plane Coordinate System (SPCS), a set of projected CRSs covering the U.S.

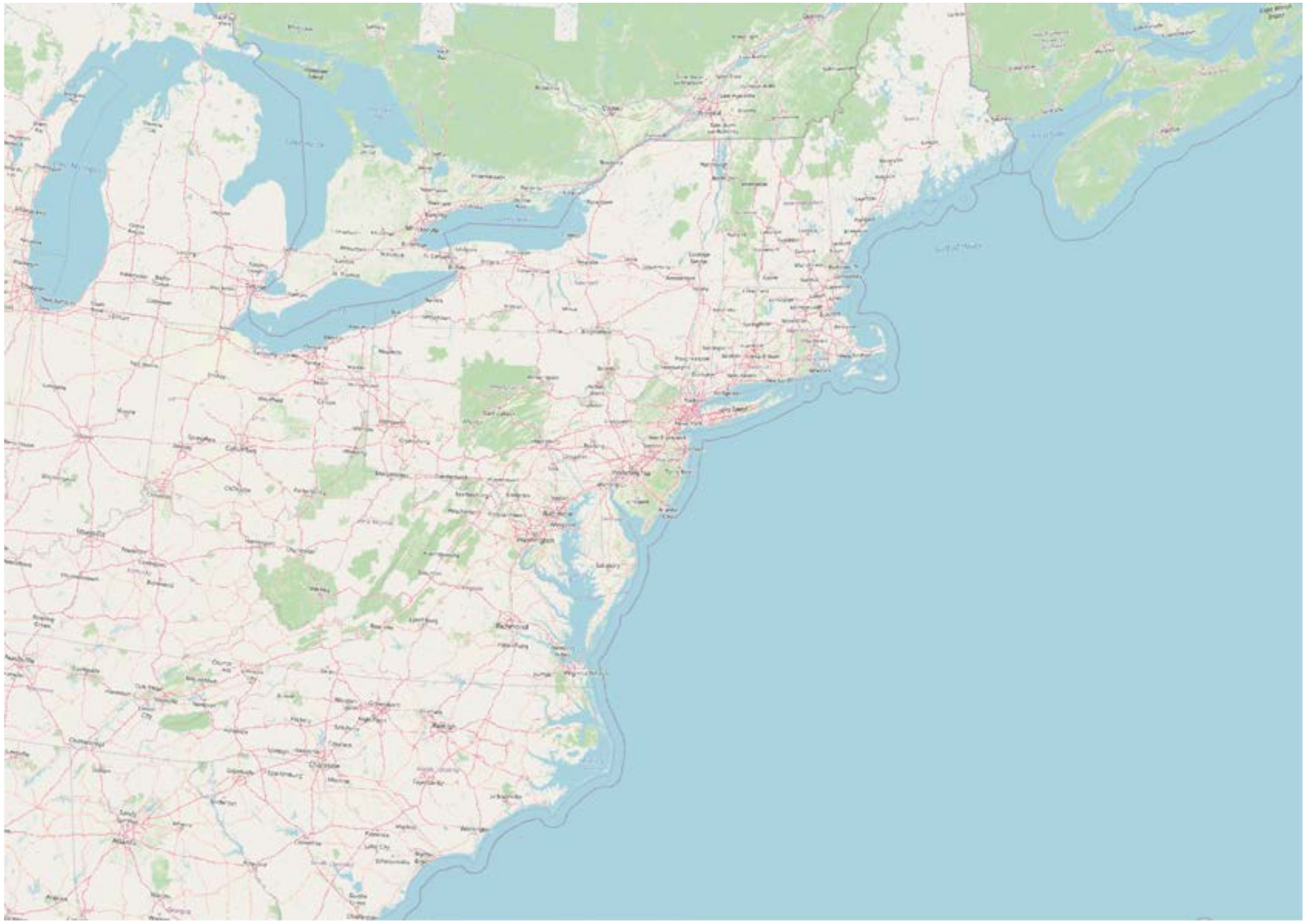

Paterson
Hackensack
Rutherford
Secaucus
Union City
Weehawken
Hoboken
Jersey City
New York
Newark
Bayonne
Elizabeth
Linden
Carteret
Woodbridge
Staten Island
Jamaica Bay Wildlife Refuge Gateway National Recreation Area
Oceanside
Long Beach

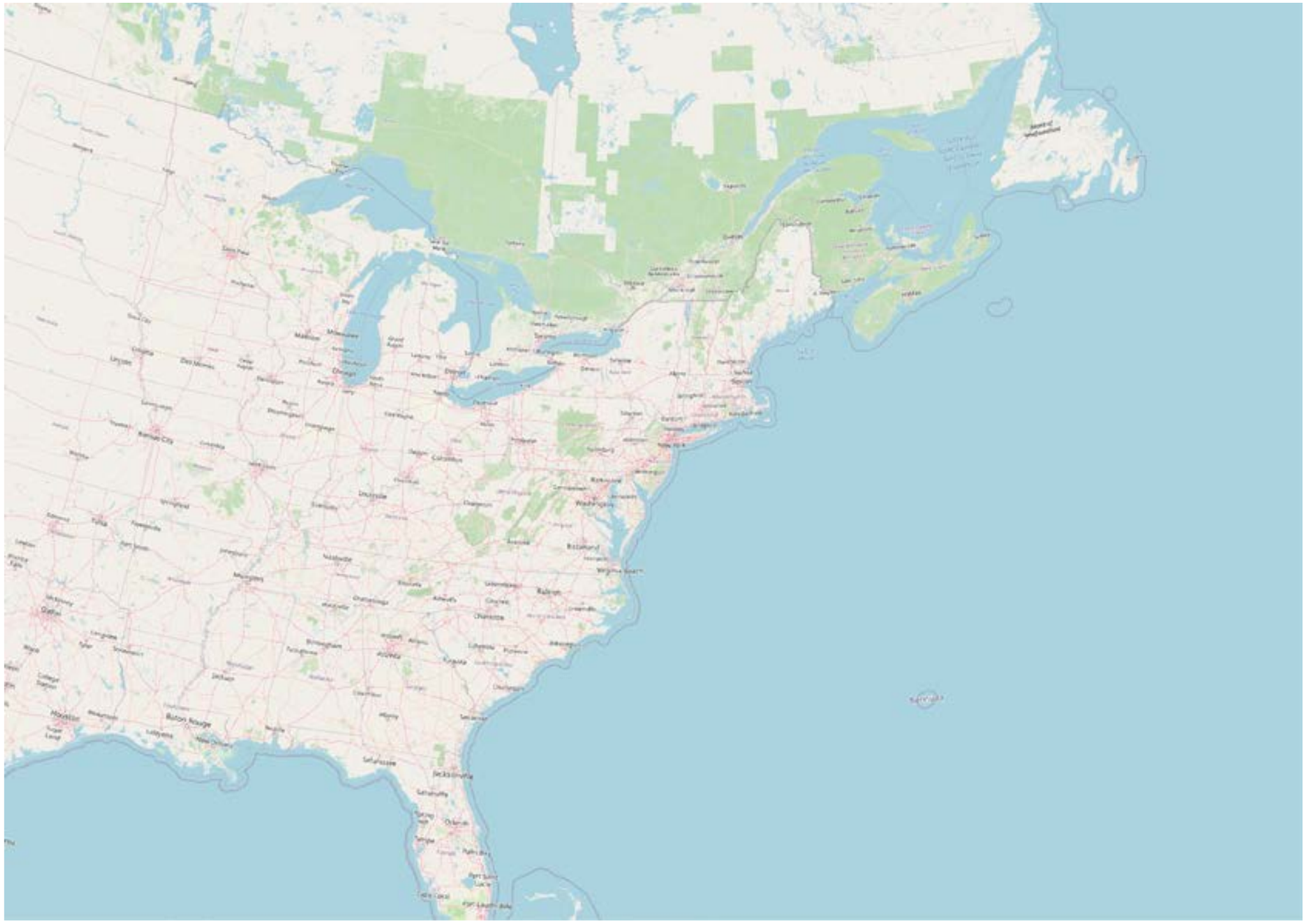

Newark
New York
Elizabeth
New Brunswick
Paterson
Yonkers
Stamford
Hempstead
Long Beach






Gulf of Maine






Acapulco
Puebla
Guatemala
Honduras
Nicaragua
Managua
Jamaica
Cuba

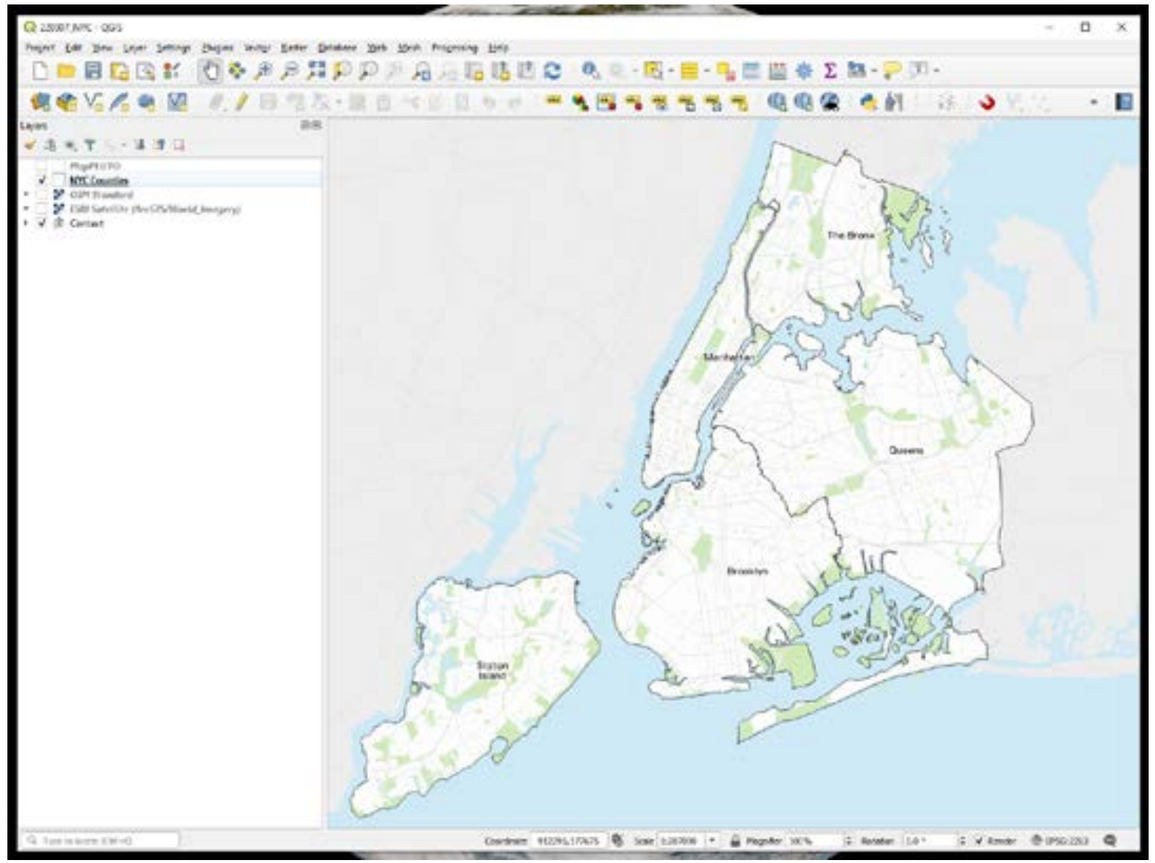

🚨 Note!
Geographic CRSs are locally useful, but significantly distorted at large scales.

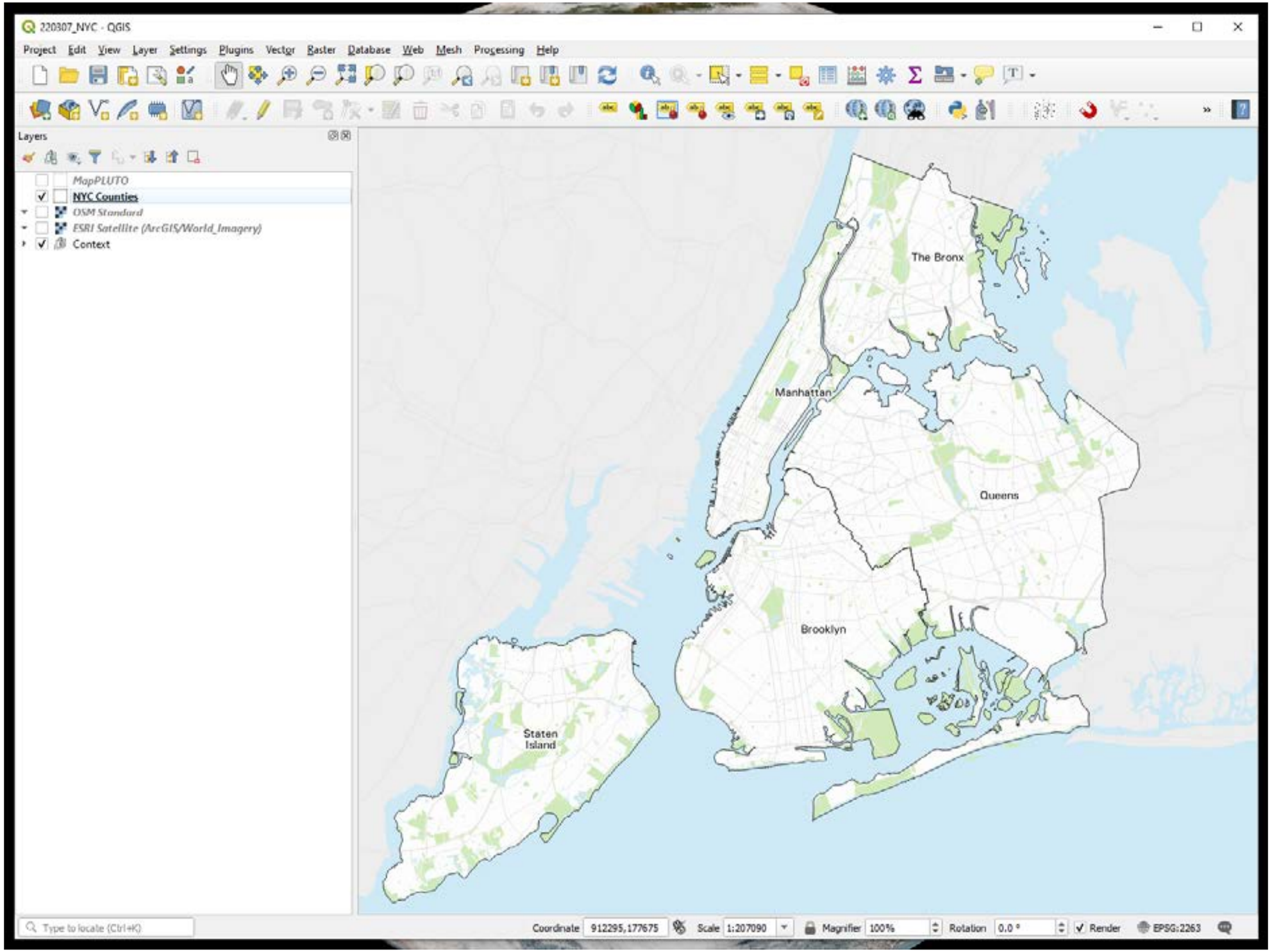
Open **QGIS** and create an **empty project**.

QGIS

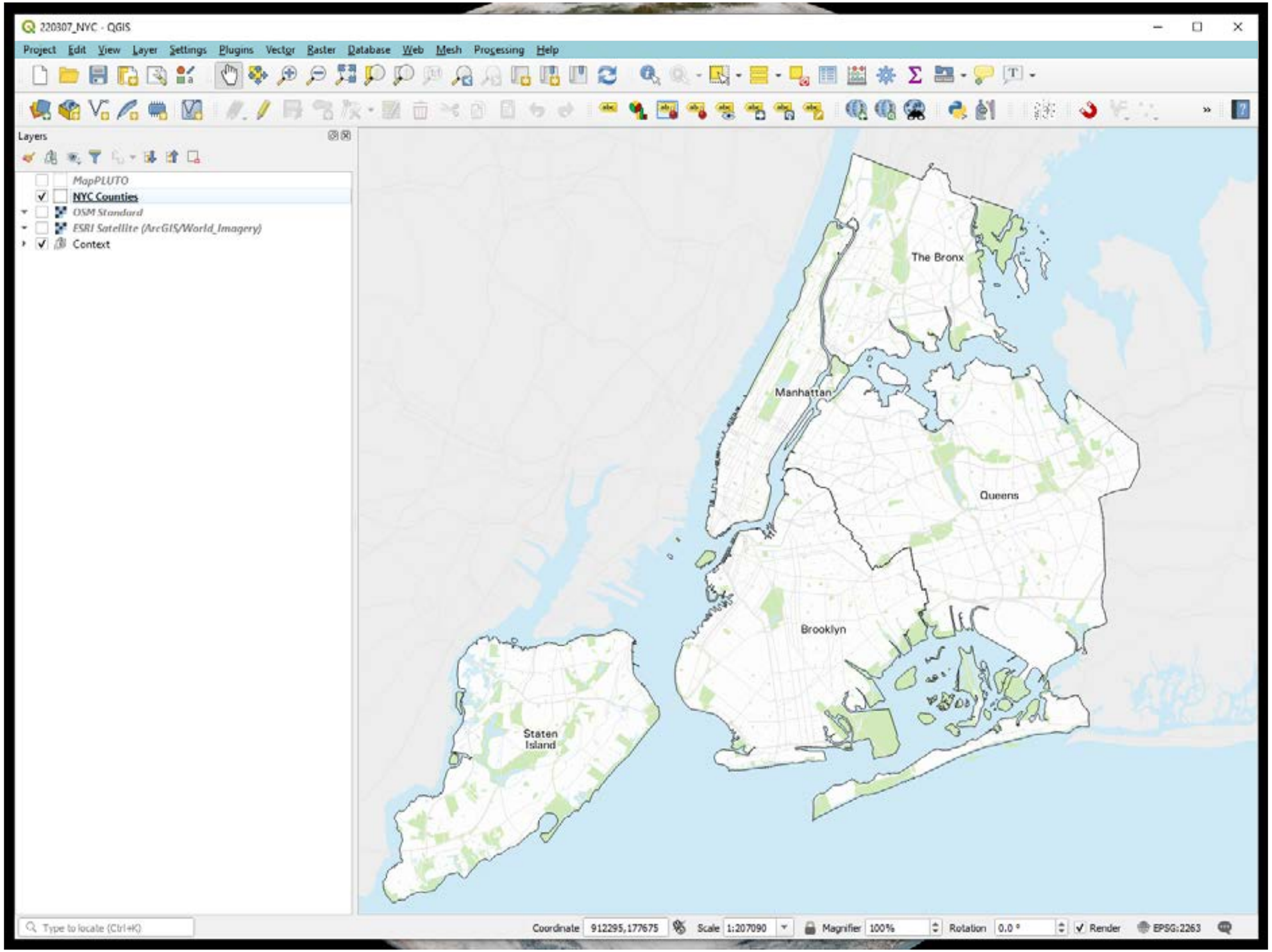
Free & Open Source GIS

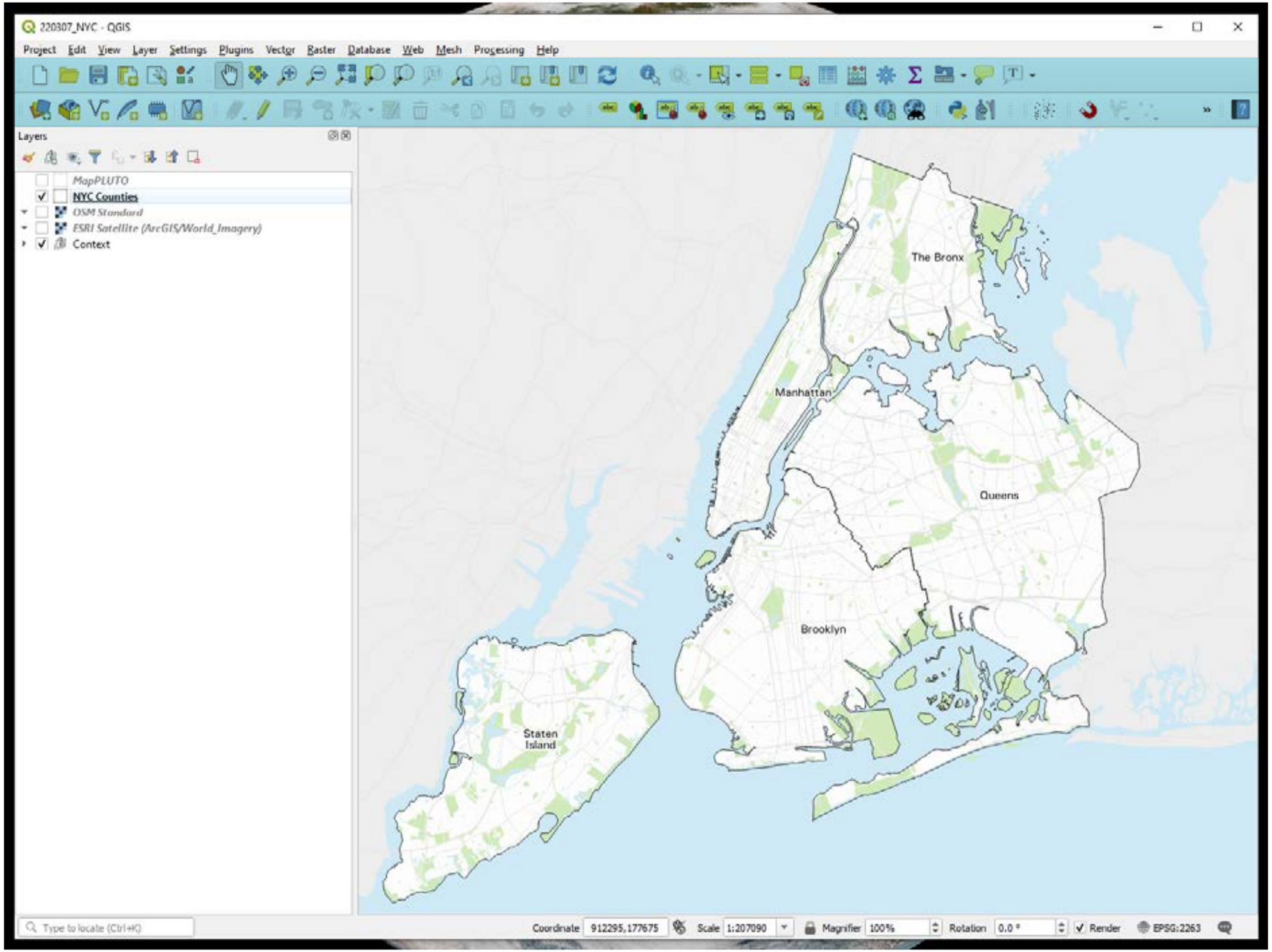
Download https://www.qgis.org/en/site/



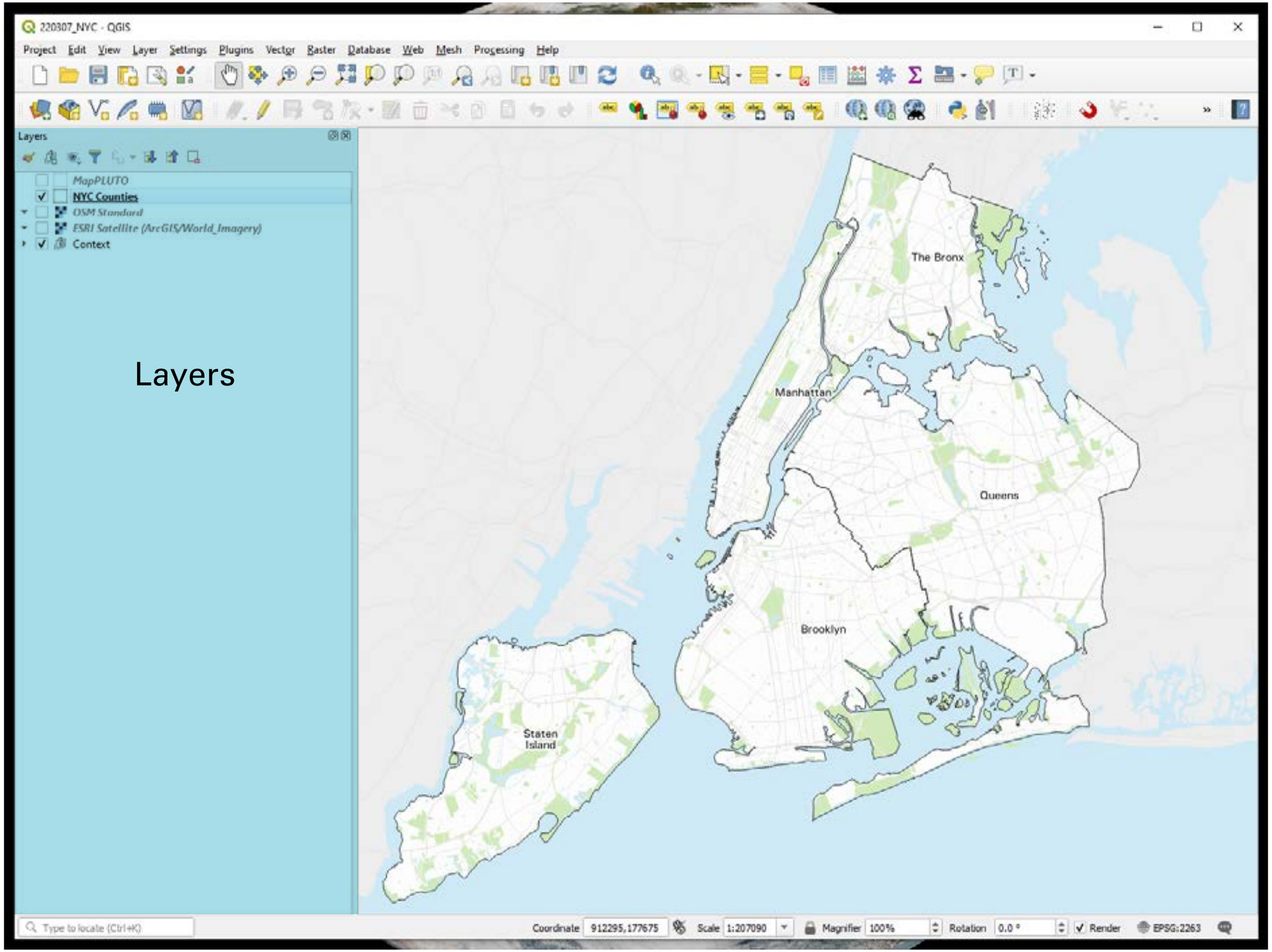

220307_NYC - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
MapPLUTO
NYC Counties
OSM Standard
ESRI Satellite (ArcGIS/World_Imagery)
Context
The Bronx
Manhattan
Queens
Brooklyn
Staten Island
Type to locate (Ctrl+K)
Coordinate 912295,177675
Scale 1:207090
Magnifier 100%
Rotation 0.0 °
Render
EPSG:2263

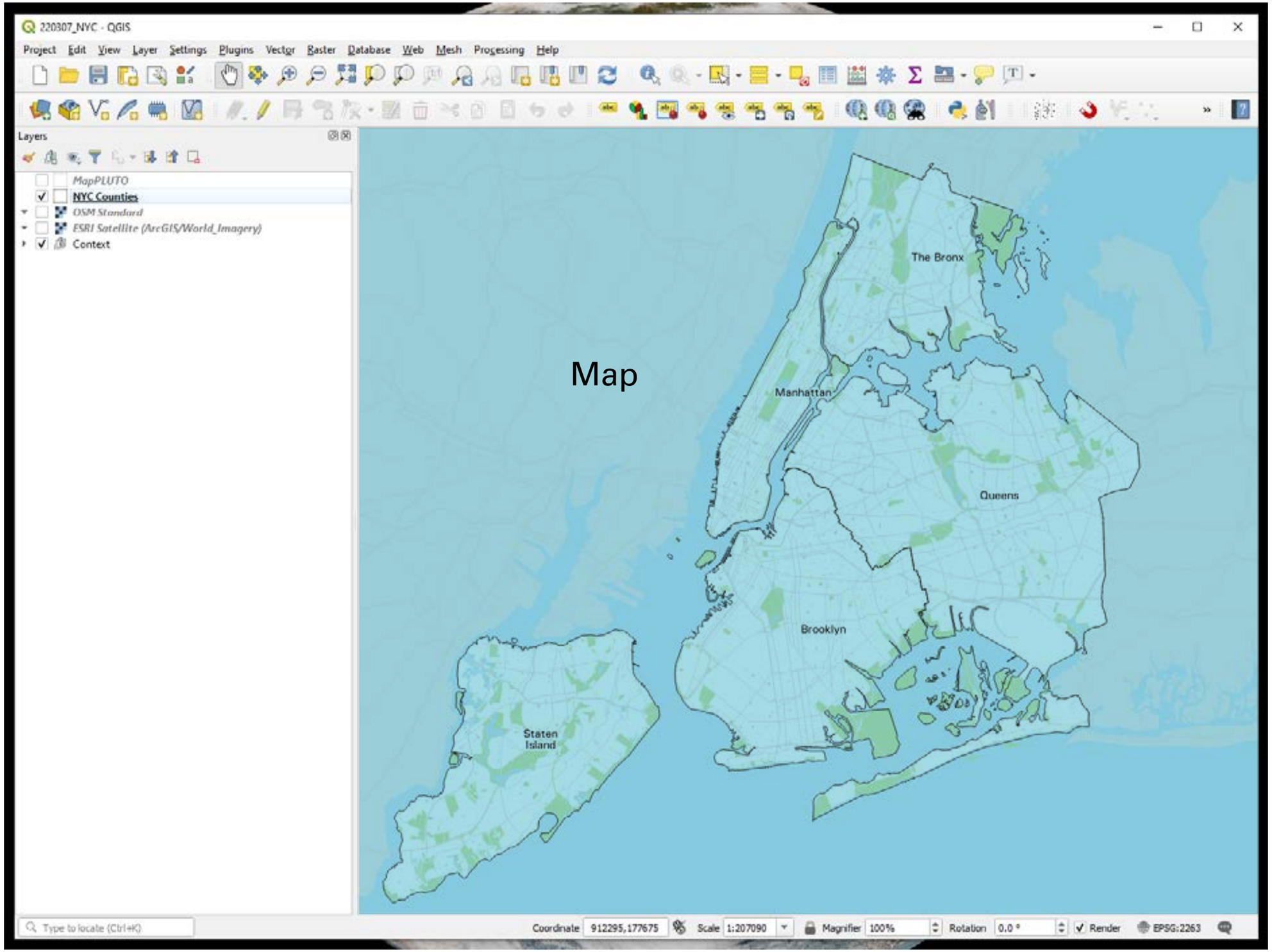
Menu

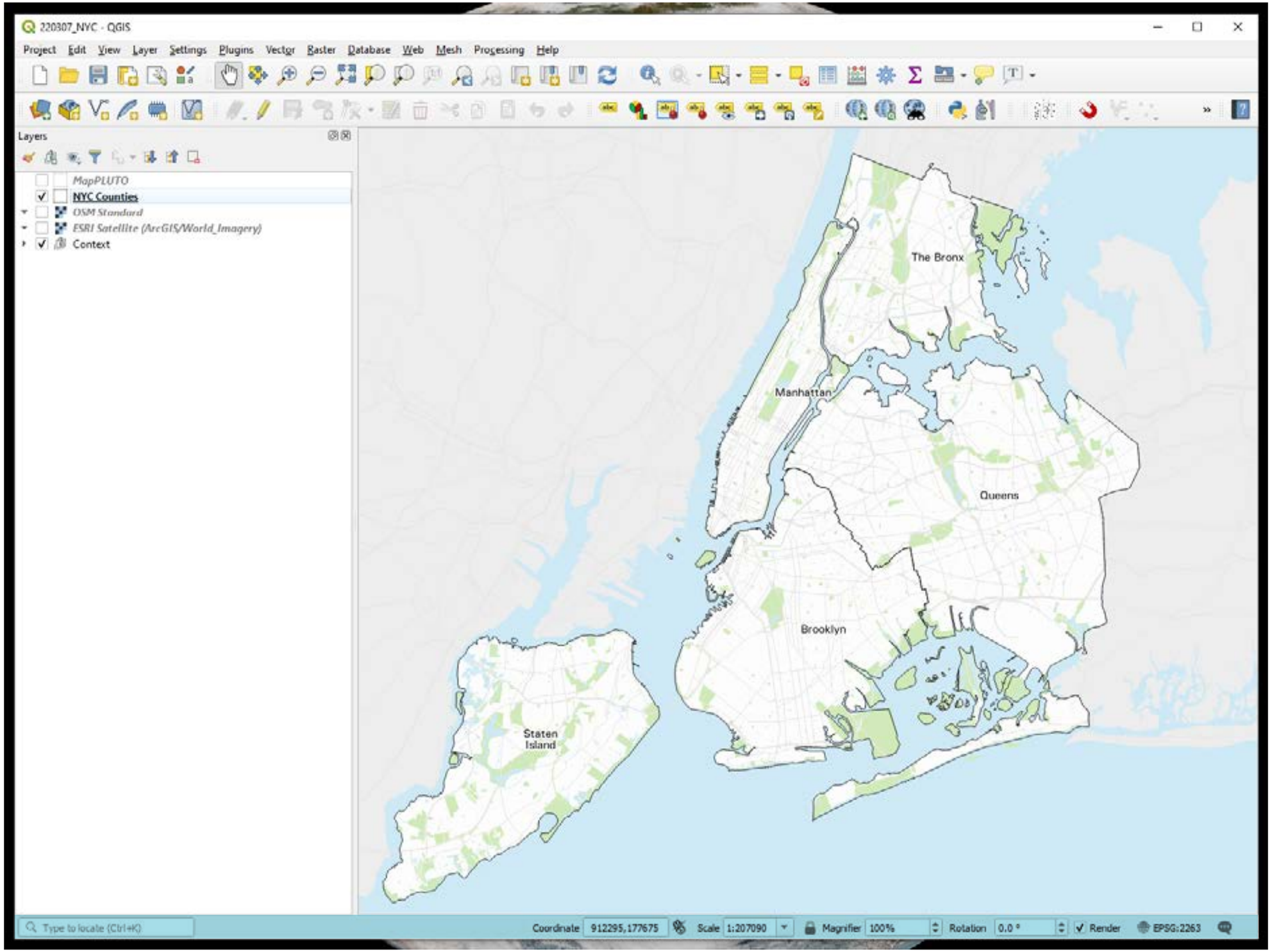

220307_NYC - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
MapPLUTO
NYC Counties
OSM Standard
ESRI Satellite (ArcGIS/World_Imagery)
Context
The Bronx
Manhattan
Queens
Brooklyn
Staten Island
Type to locate (Ctrl+K)
Coordinate 912295,177675
Scale 1:207090
Magnifier 100%
Rotation 0.0 °
Render
EPSG:2263


220307_NYC - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
MapPLUTO
NYC Counties
OSM Standard
ESRI Satellite (ArcGIS/World_Imagery)
Context
The Bronx
Manhattan
Queens
Brooklyn
Staten Island
Type to locate (Ctrl+K)
Coordinate 912295,177675
Scale 1:207090
Magnifier 100%
Rotation 0.0 °
Render
EPSG:2263

Toolbar

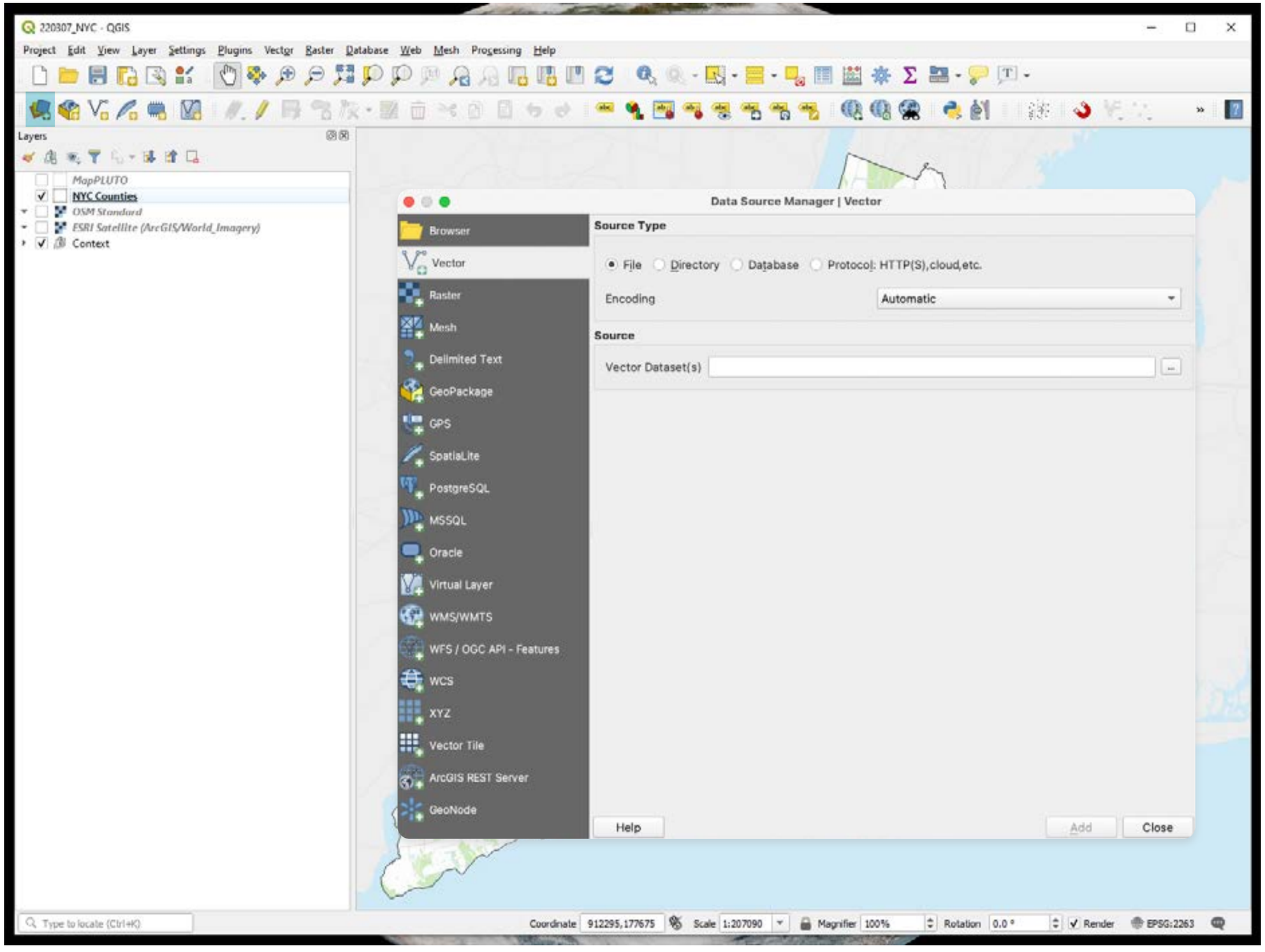

220307_NYC - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
MapPLUTO
NYC Counties
OSM Standard
ESRI Satellite (ArcGIS/World_Imagery)
Context
Layers
The Bronx
Manhattan
Queens
Brooklyn
Staten Island
Type to locate (Ctrl+K)
Coordinate 912295,177675
Scale 1:207090
Magnifier 100%
Rotation 0.0 °
Render
EPSG:2263


220307_NYC - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
MapPLUTO
NYC Counties
OSM Standard
ESRI Satellite (ArcGIS/World_Imagery)
Context
Map
The Bronx
Manhattan
Queens
Brooklyn
Staten Island
Type to locate (Ctrl+K)
Coordinate 912295,177675
Scale 1:207090
Magnifier 100%
Rotation 0.0 °
Render
EPSG:2263


220307_NYC - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
MapPLUTO
NYC Counties
OSM Standard
ESRI Satellite (ArcGIS/World_Imagery)
Context
The Bronx
Manhattan
Queens
Brooklyn
Staten Island
Type to locate (Ctrl+K)
Coordinate 912295,177675
Scale 1:207090
Magnifier 100%
Rotation 0.0 °
Render
EPSG:2263

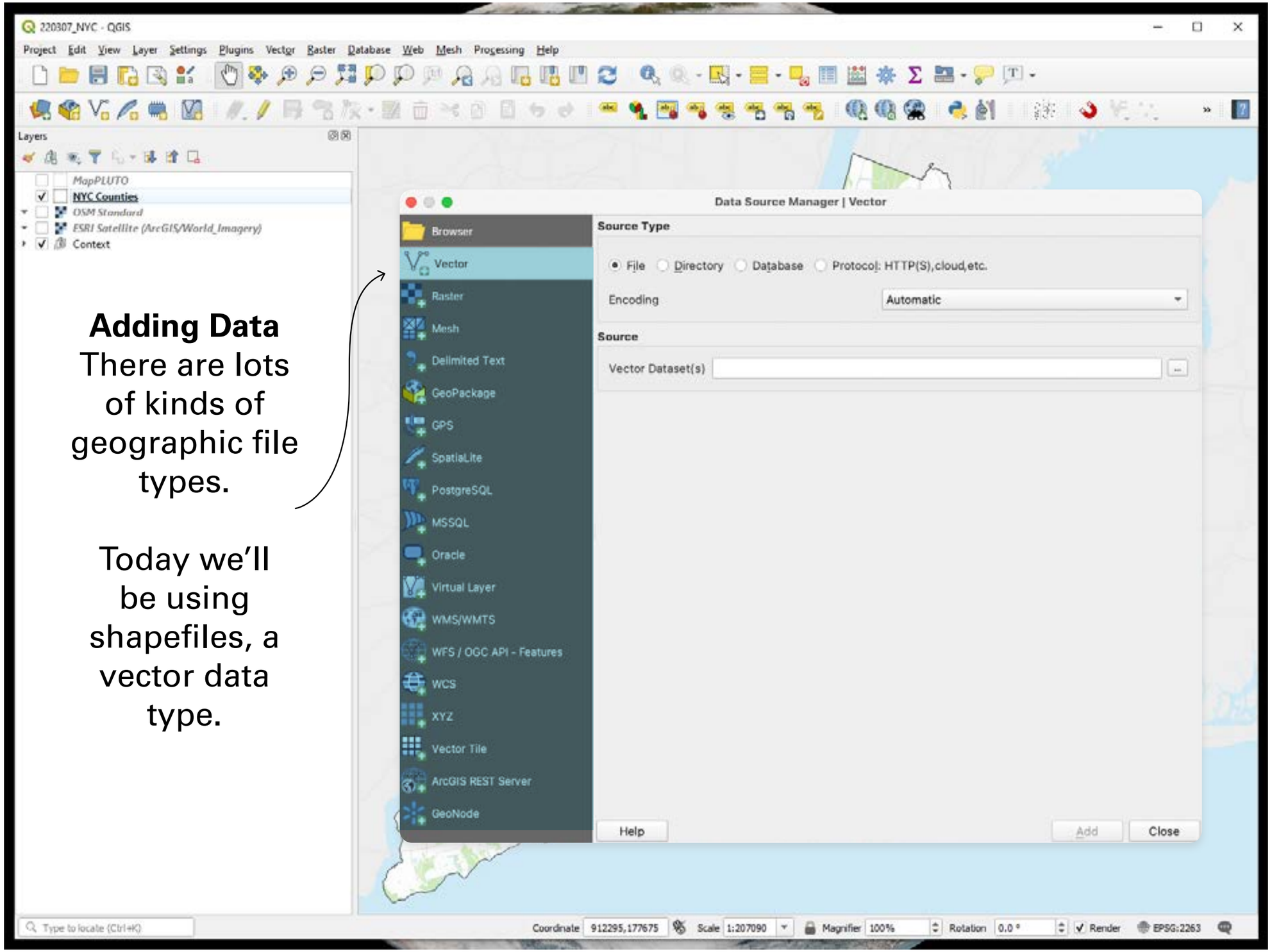
Utilities

Add data

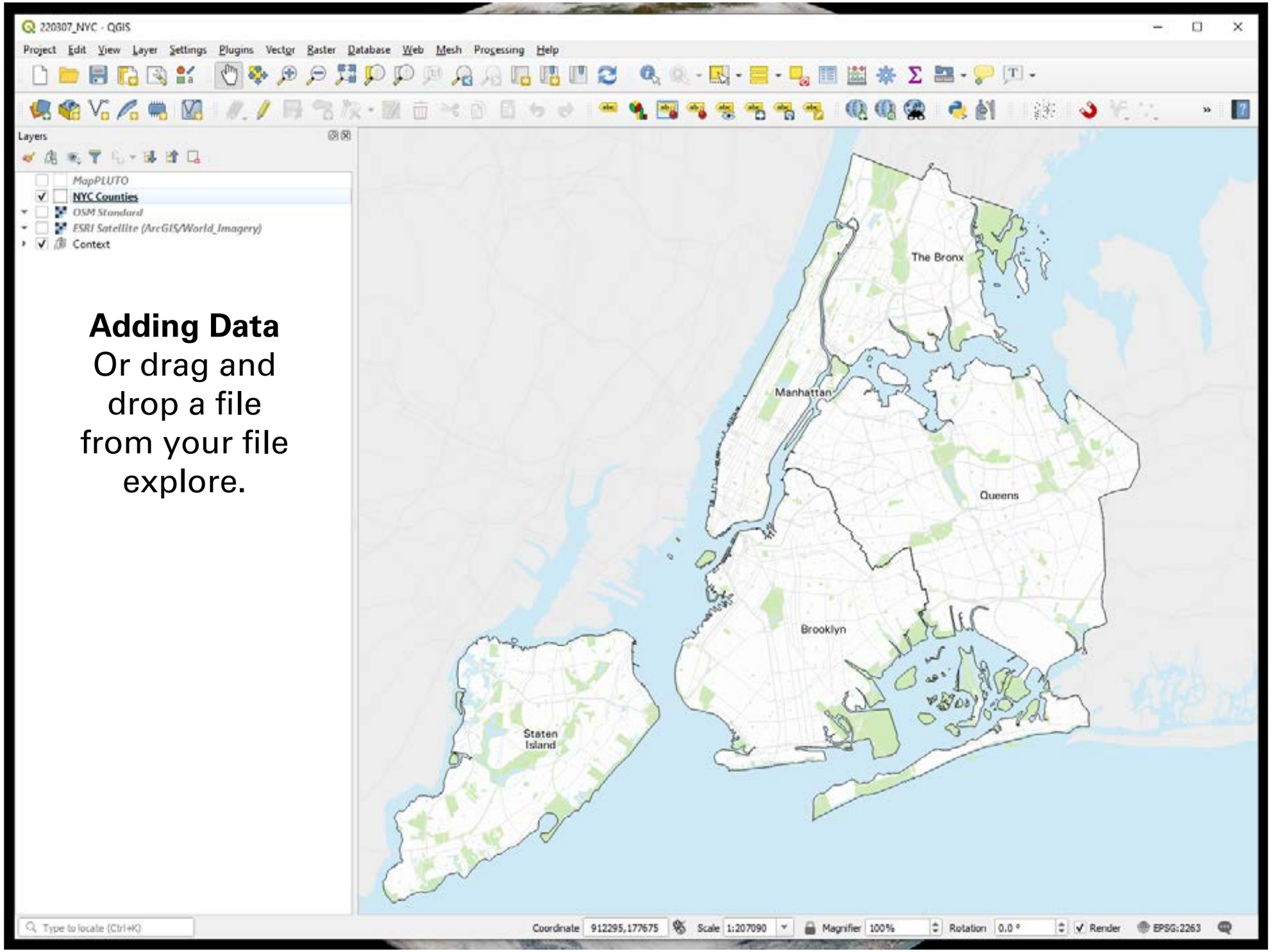

220307_NYC - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
MapPLUTO
NYC Counties
OSM Standard
ESRI Satellite (ArcGIS/World_Imagery)
Context
Data Source Manager | Vector
Browser
Vector
Raster
Mesh
Delimited Text
GeoPackage
GPS
SpatiaLite
PostgreSQL
MSSQL
Oracle
Virtual Layer
WMS/WMTS
WFS / OGC API - Features
WCS
XYZ
Vector Tile
ArcGIS REST Server
GeoNode
Source Type
File Directory Database Protocol: HTTP(S),cloud,etc.
Encoding
Automatic
Source
Vector Dataset(s)
Help
Add
Close
Type to locate (Ctrl+K)
Coordinate 912295,177675
Scale 1:207090
Magnifier 100%
Rotation 0.0 °
Render
EPSG:2263

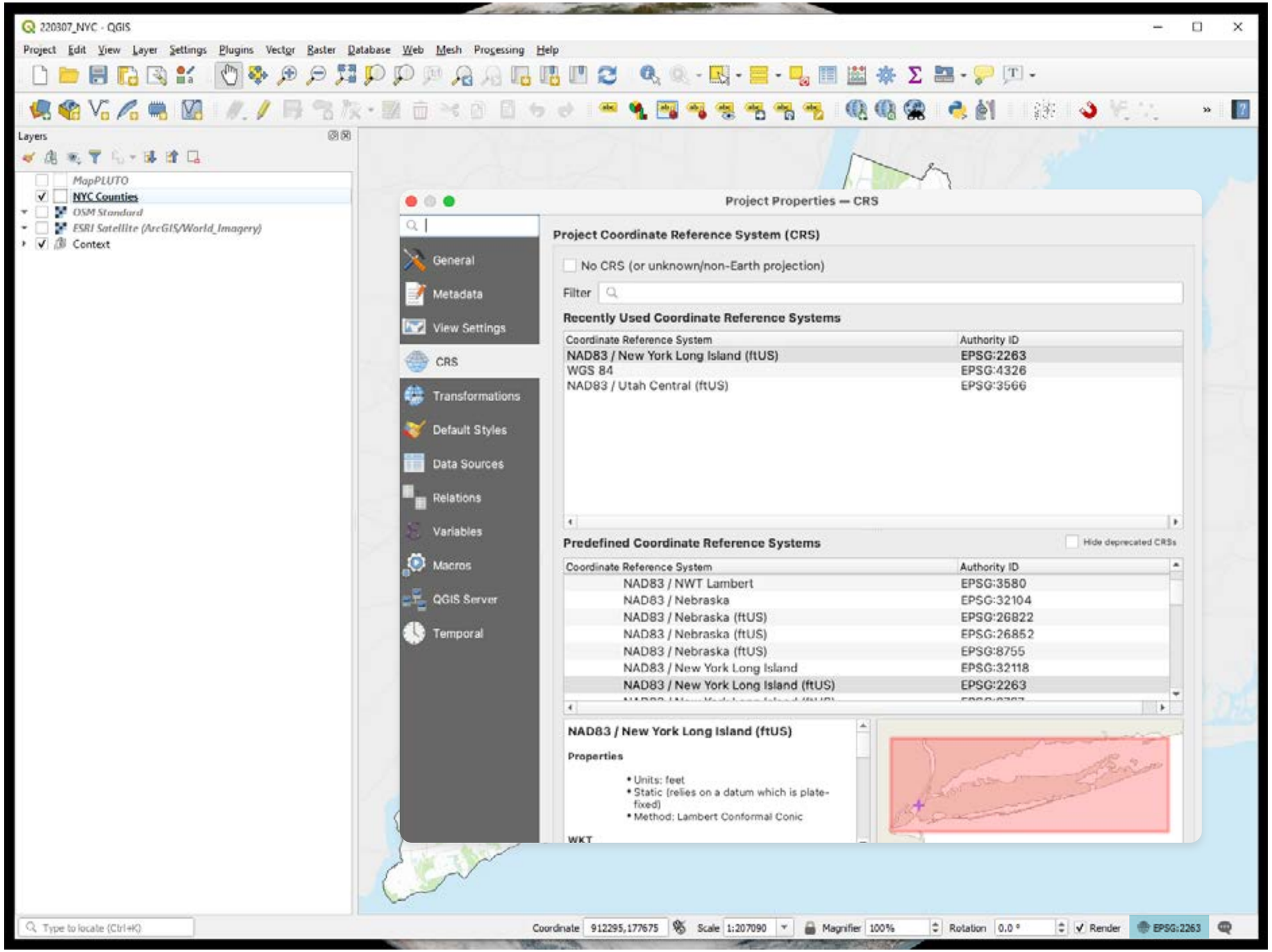
Adding Data
There are lots of kinds of geographic file types.

Today we'll be using shapefiles, a vector data type.

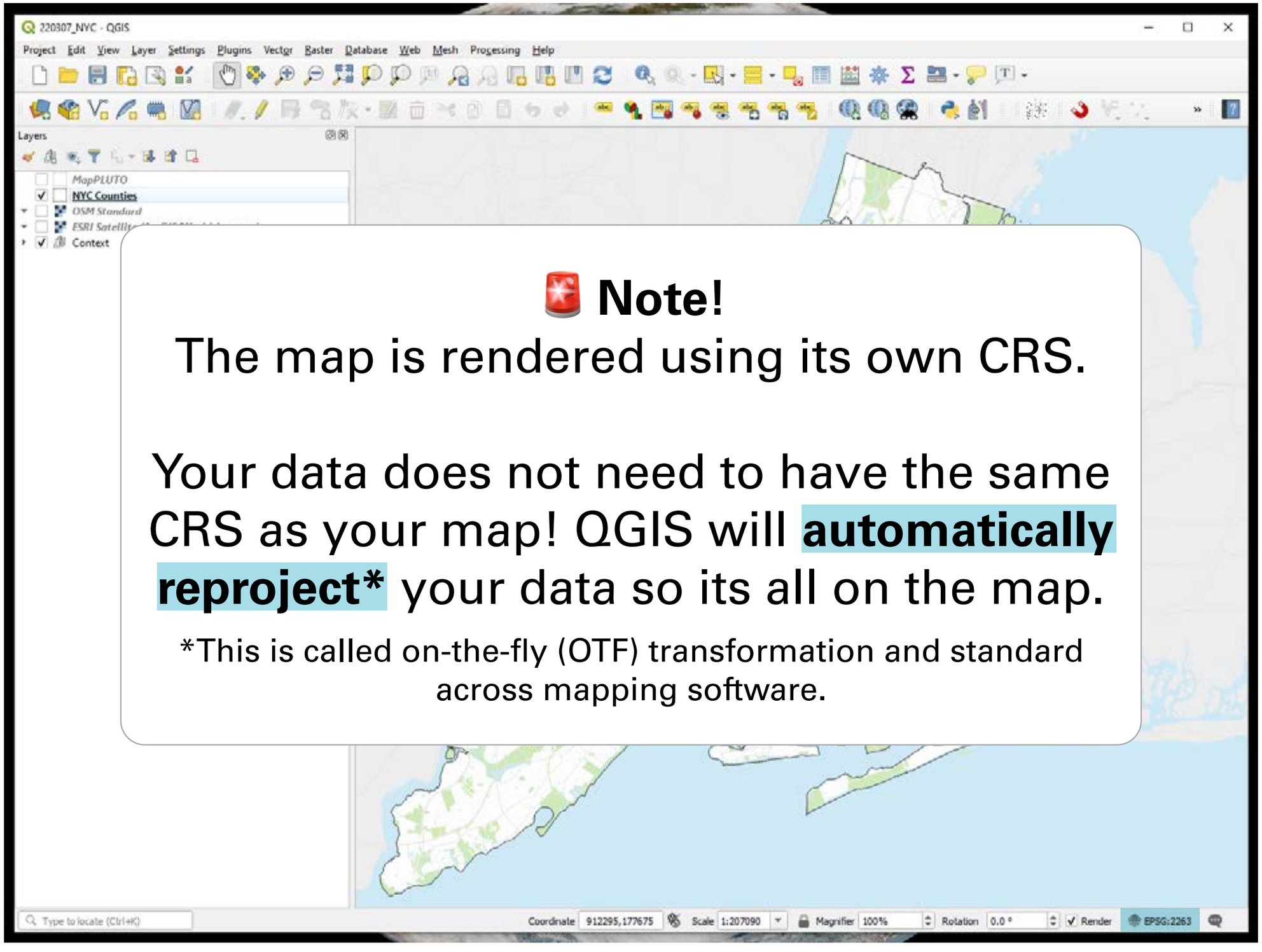


Adding Data
Or drag and drop a file from your file explore.





Change the map CRS



🚨 **Note!**

The map is rendered using its own CRS.

Your data does not need to have the same CRS as your map! QGIS will **automatically reproject*** your data so its all on the map.

*This is called on-the-fly (OTF) transformation and standard across mapping software.

Change the map CRS

Today

- 🗺️ Introduce **GIS** and mapping software
- 🗄️ Highlight **data sources** in New York City
- 🧰 Discuss **common GIS tasks**
- 🏗️ Introduce **site analysis**

🏗️🏢 Land Use, Development, and NYC Data

- NYC Open Data: map layers
- NYC DCP parcel data (PLUTO) + documentation
- NYC DCP zoning data (ZoLa interactive tool)
- Building footprints
- DOB Permit Issuance* (active major construction dashboard)
- Automated City Register Information System (ACRIS)

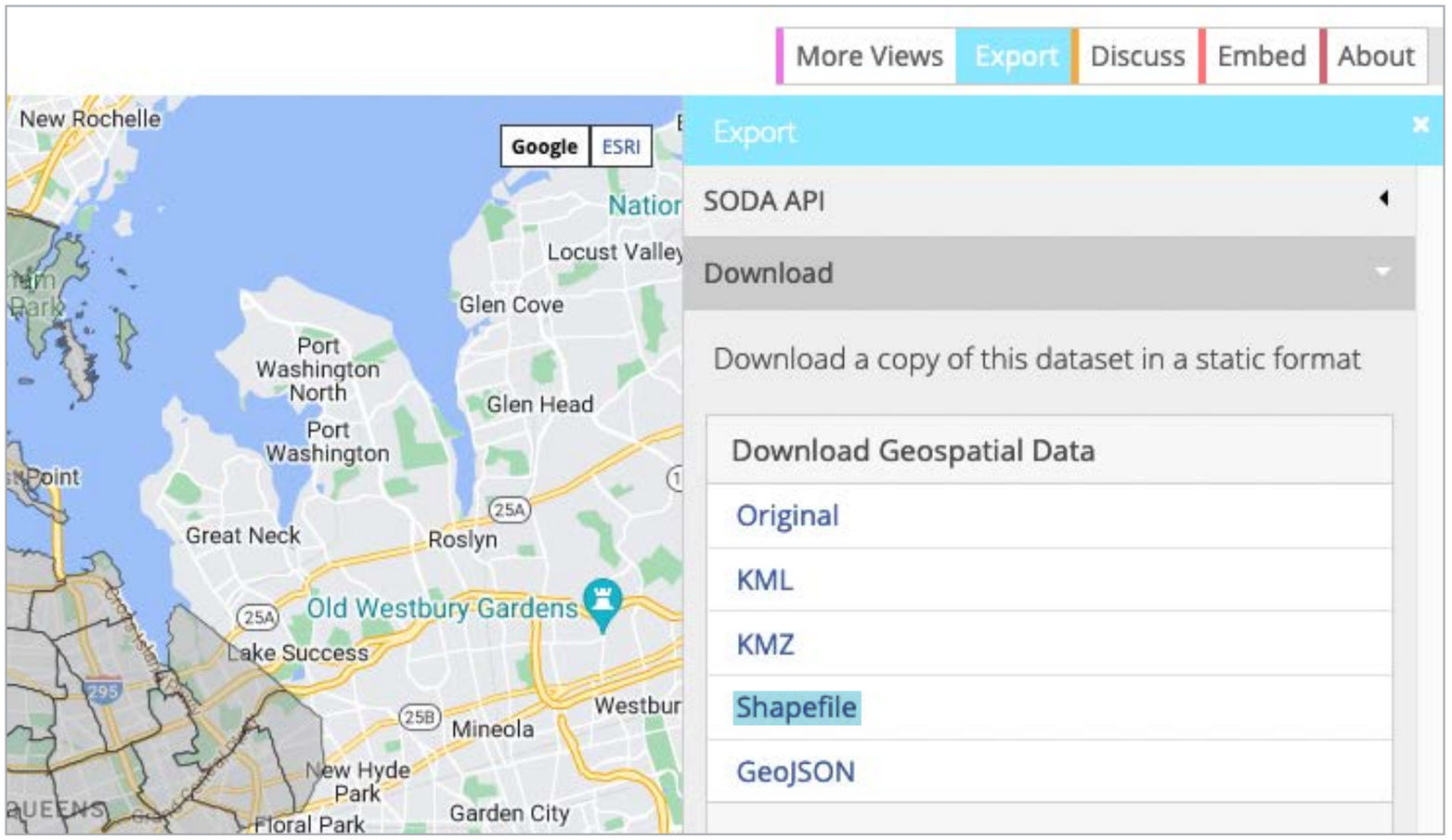
↪️ Exporting data

There are two ways to export data from NYC open data and other Socrata websites

1. Export as a **shapefile** or **file geodatabase**
2. Application Program Interface (**API**)

Usually you will want to export your data as a shapefile. However, when you are dealing with very large data, it can be helpful to use an API if one is available. We'll use the first method today but know that the second is available!

↪️ Exporting data as a shapefile or file geodatabase



Click me!

Go to NYC Open data and download the **Housing Database by NTA** as a shapefile.

Add the layer to your map by dragging the .shp file in or through the data source manager.

⚠

Note: there are two datasets by the same name on NYC Open Data. Use the one that is not labeled “(Map)”

Check the layer CRS of the Housing Database by NTA (Map) data by double-clicking the layer.

Try changing the map CRS (not the layer CRS!) to WGS84 (EPSG 4326).

Now set the map CRS to **NAD83 New York Long Island (ftUS)**(EPSG=2263).

Re-check the layer CRS of the Housing Database by NTA (Map) data by double-clicking the layer.

Confirm that the data source's CRS is still WGS84 (EPSG 4326).

Today

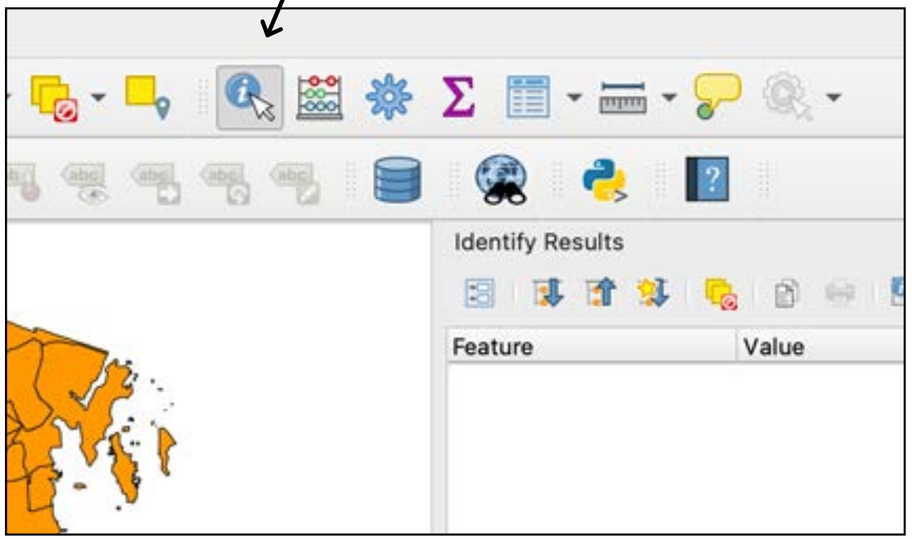
- 🗺️ Introduce **GIS** and mapping software
- 🗄️ Highlight **data sources** in New York City
- 🧰 Discuss **common GIS tasks**
- 🏗️ Introduce **soft site analysis**

🧰 Common GIS tasks 15m

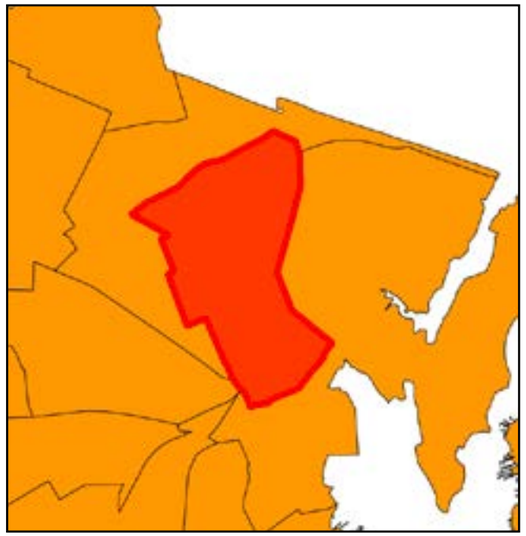
- Selecting and inspecting spatial data
- Filtering data
- Area normalization
- Adding a basemap to your project

🔍 Selecting and inspecting spatial data

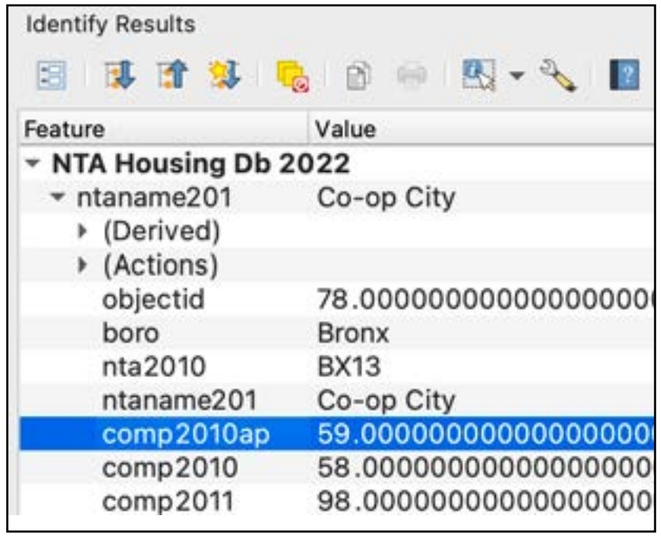
Select the identify tool



Click on a feature

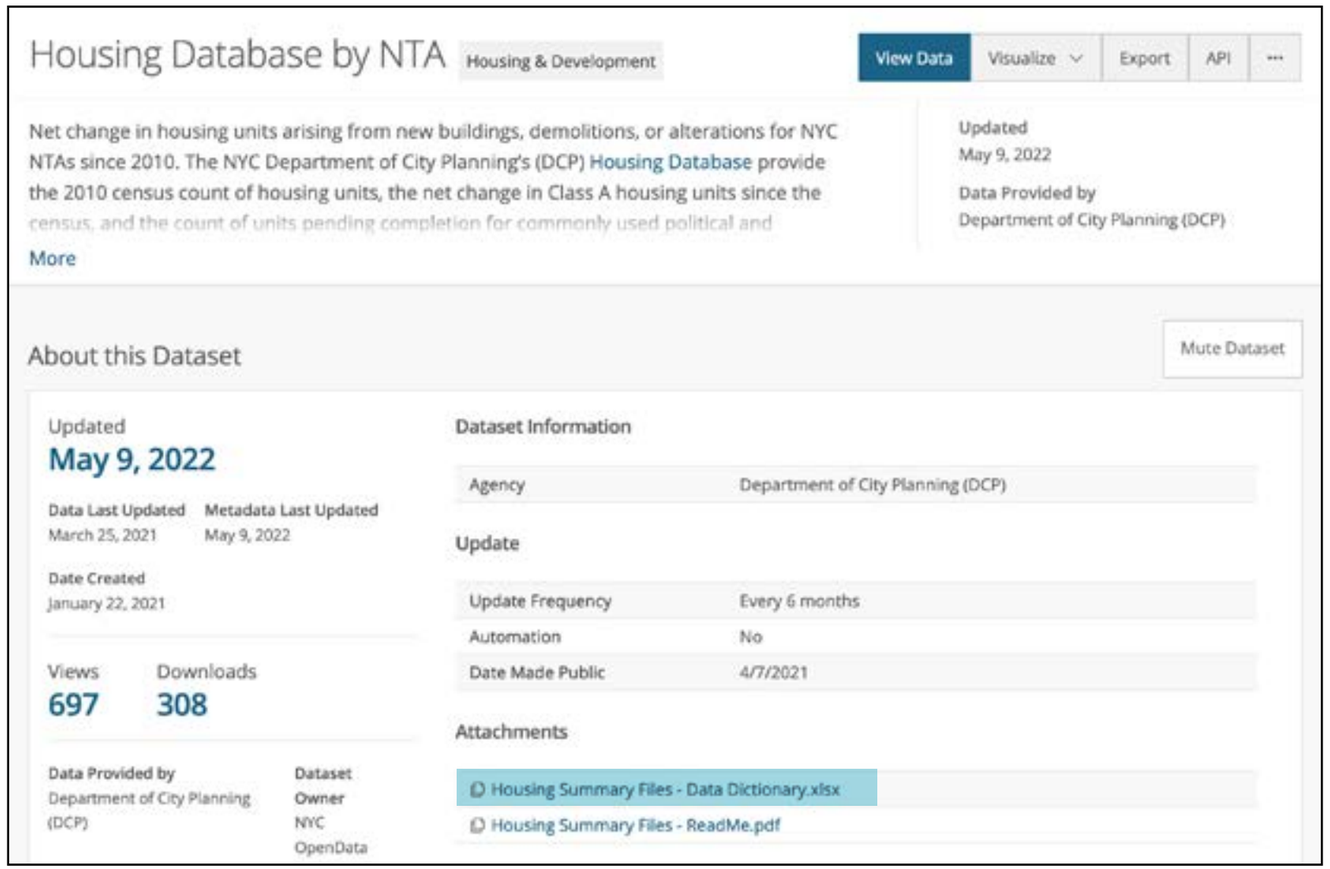


Inspect feature attributes



🔍 Selecting and inspecting spatial data

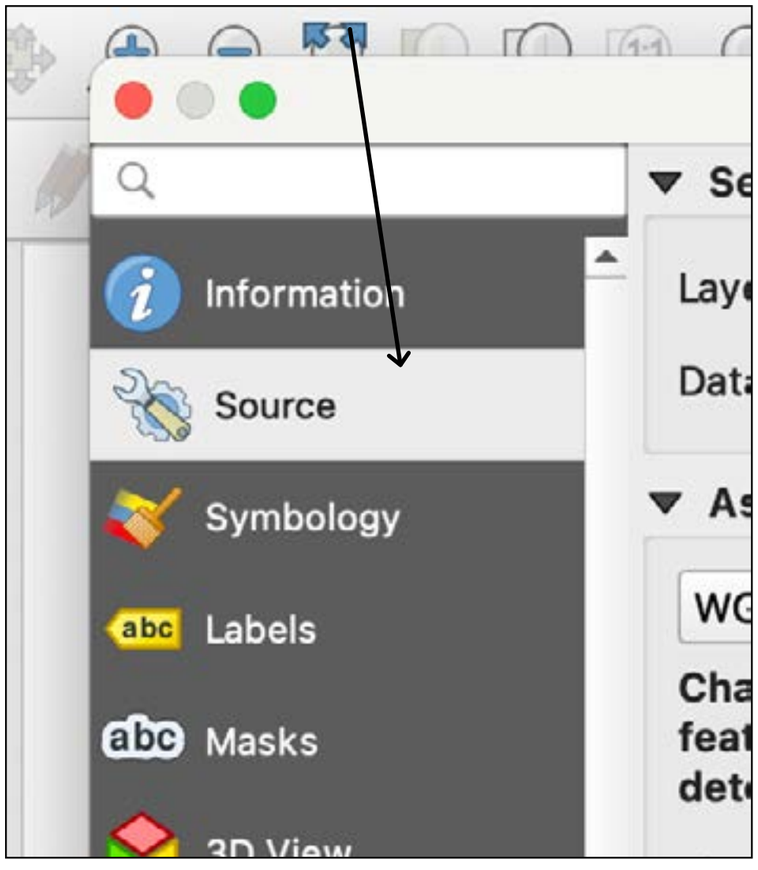
Finally, reference the documentation on NYC Open Data to understand your data!



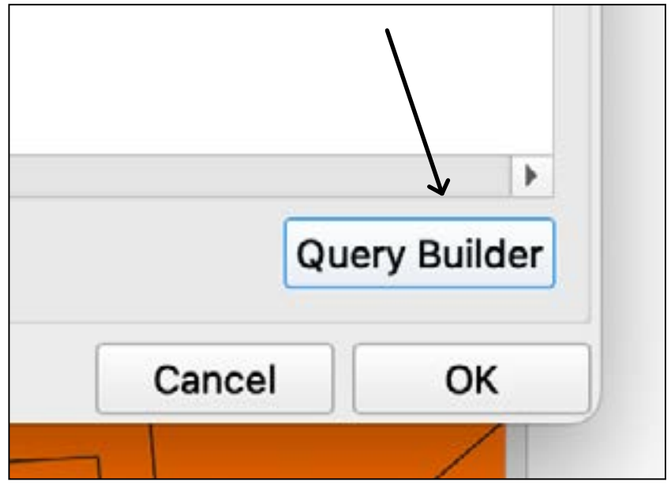
🔍 Filtering data

1. Double-click the layer name

2. Click 'Source' in the sidebar

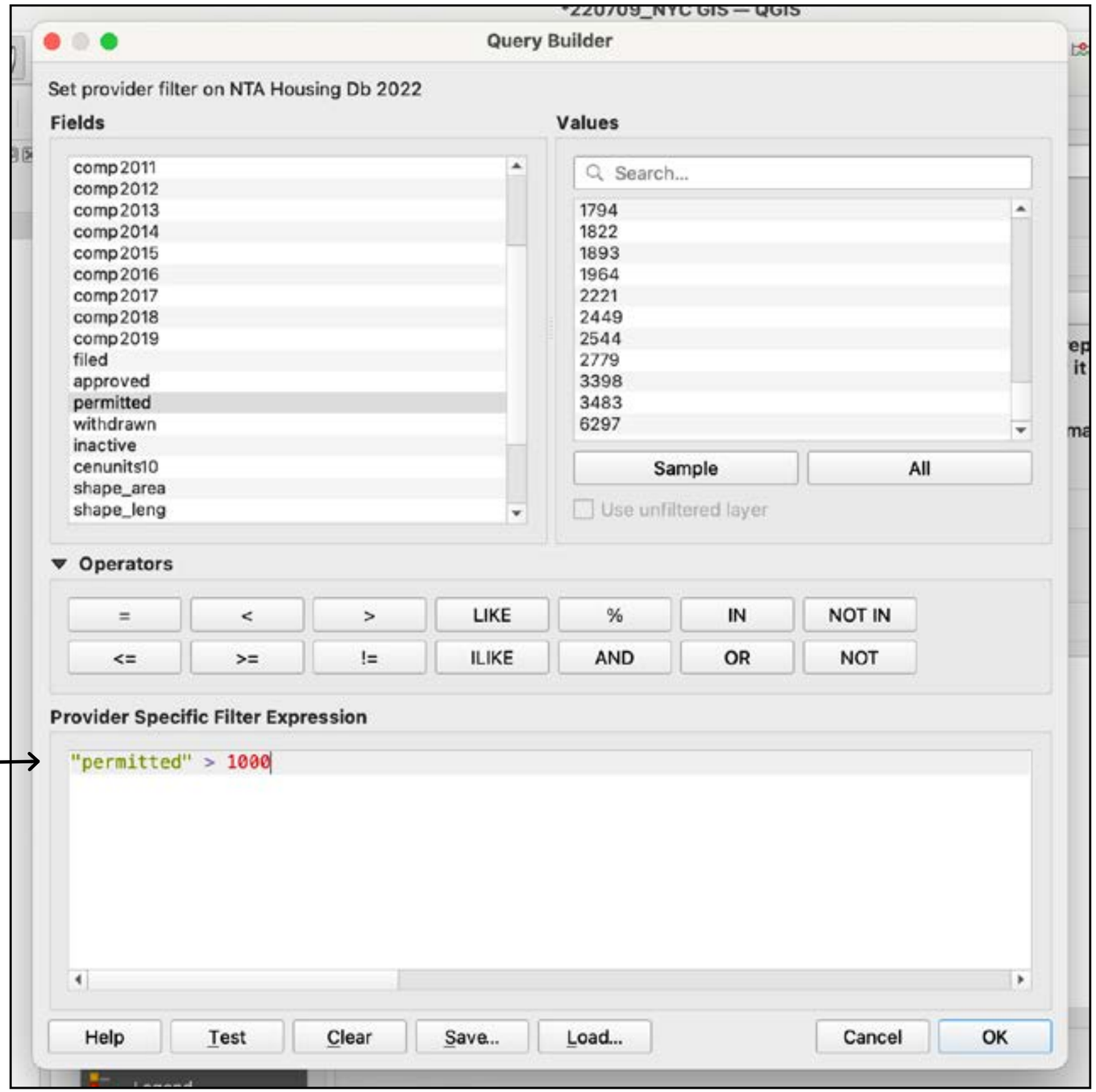


3. Then 'Query Builder'



🔍 Filtering data

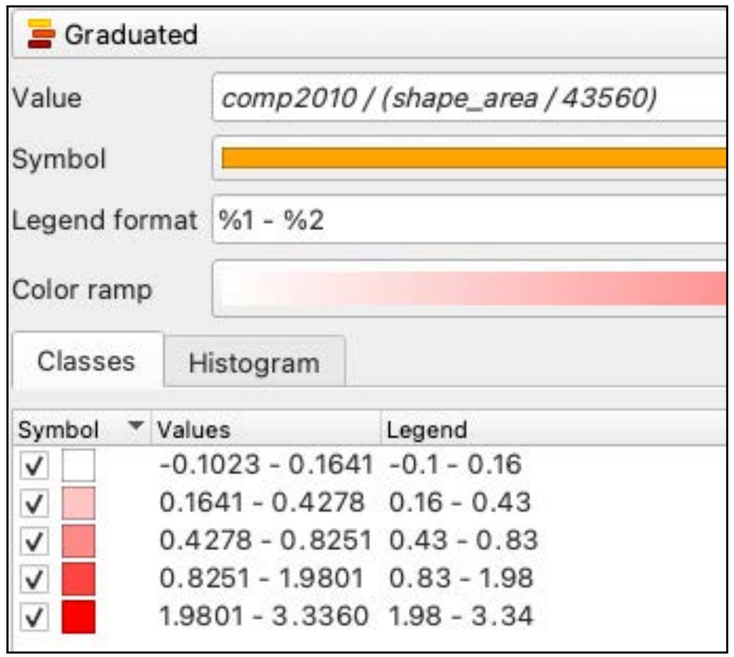
4. Write in the following formula →



Apply a filter of your choice to the NTA data.

Area Normalization

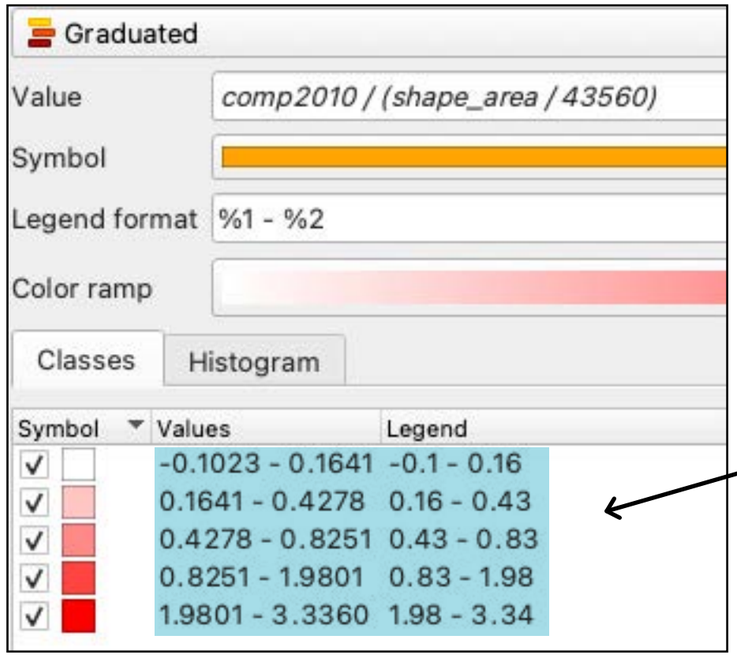
Option 1. In the layer panel's symbology pane, enter:



The “shape_area” column included in the dataset is in square feet. To convert these to acres, divide by 43560.

Area Normalization

Option 1. In the layer panel's symbology pane, enter:

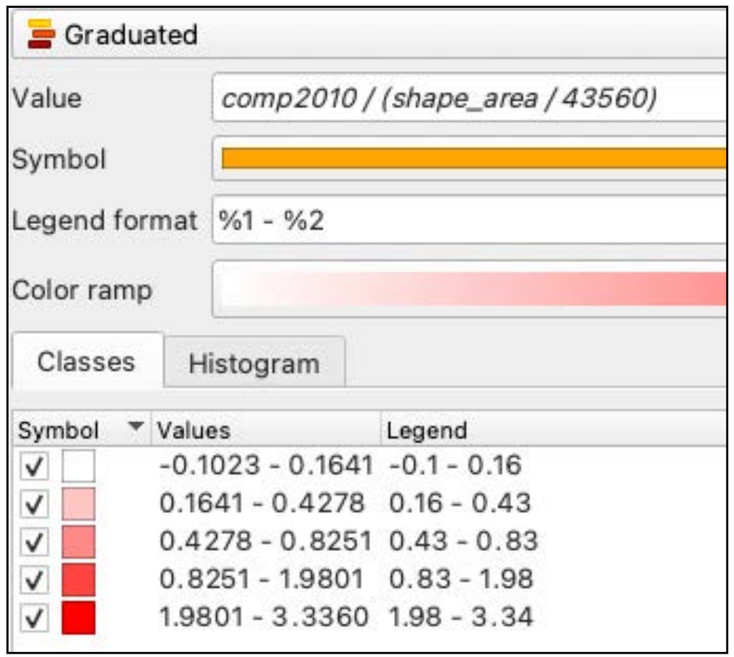


The “shape_area” column included in the dataset is in square feet. To convert these to acres, divide by 43560.

Net number of class A housing units **per acre** completed in 2010, accounting for new buildings, alterations, and demolitions

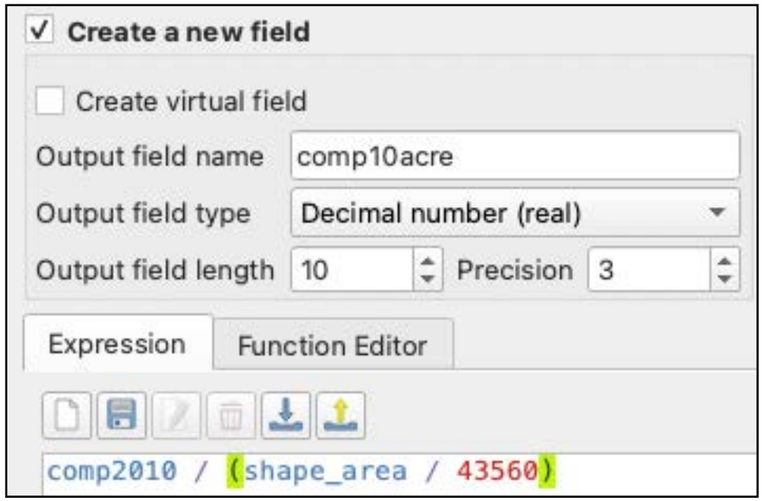
Area Normalization

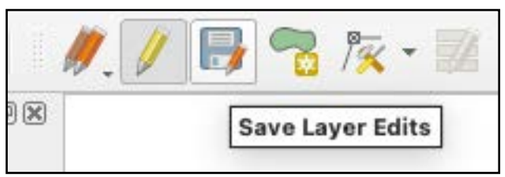
Option 1. Back in symbology, enter the following formula*



*The “shape_area” column included in the dataset is in square feet. To convert these to acres, divide by 43560.

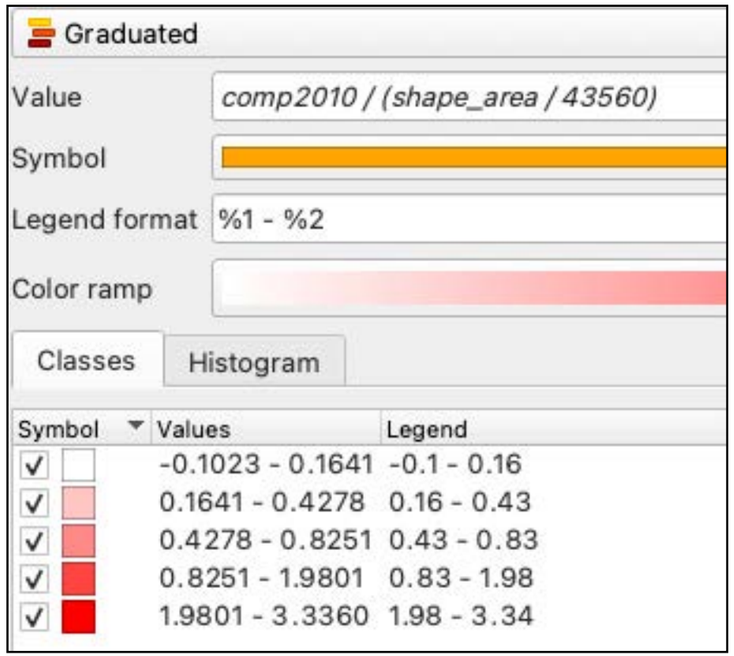
Option 2. Calculate a new column using the same formula





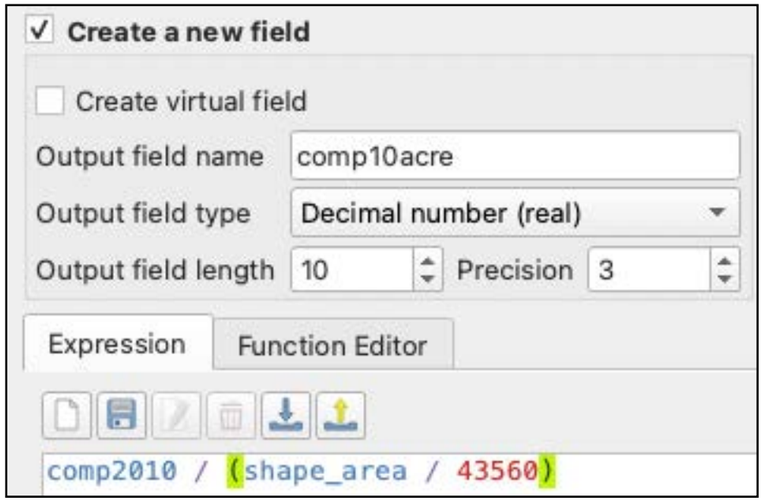
Area Normalization

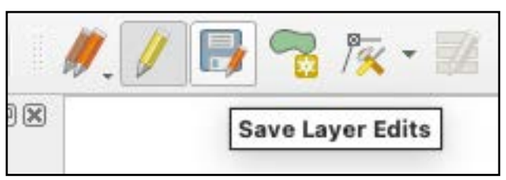
Option 1. Back in symbology, enter the following formula*



*The “shape_area” column included in the dataset is in square feet. To convert these to acres, divide by 43560.

Option 2. Calculate a new column using the same formula



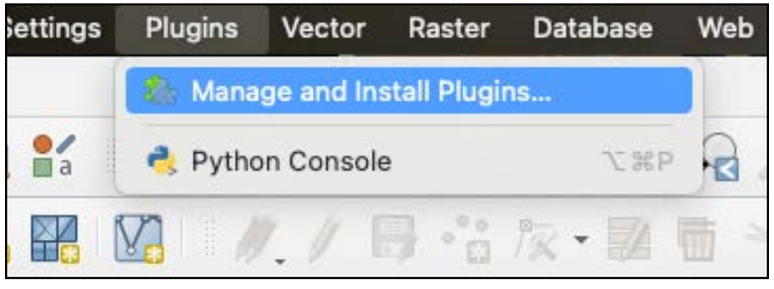


Don't forget to save your edits after creating the column, then turn off editing.

Normalize the "comp2010" column by area.

Using Basemaps

First, we need to install the QuickMapServices Plugin. Open the plugin installer.

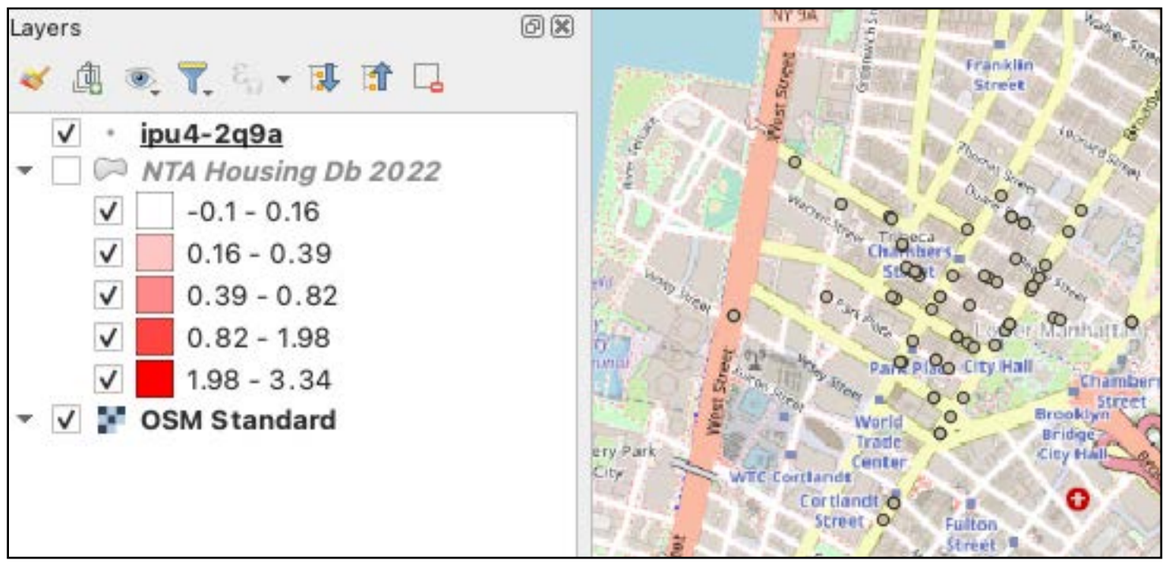


Search for an install the QMS plugin.

🛰️ Using Basemaps

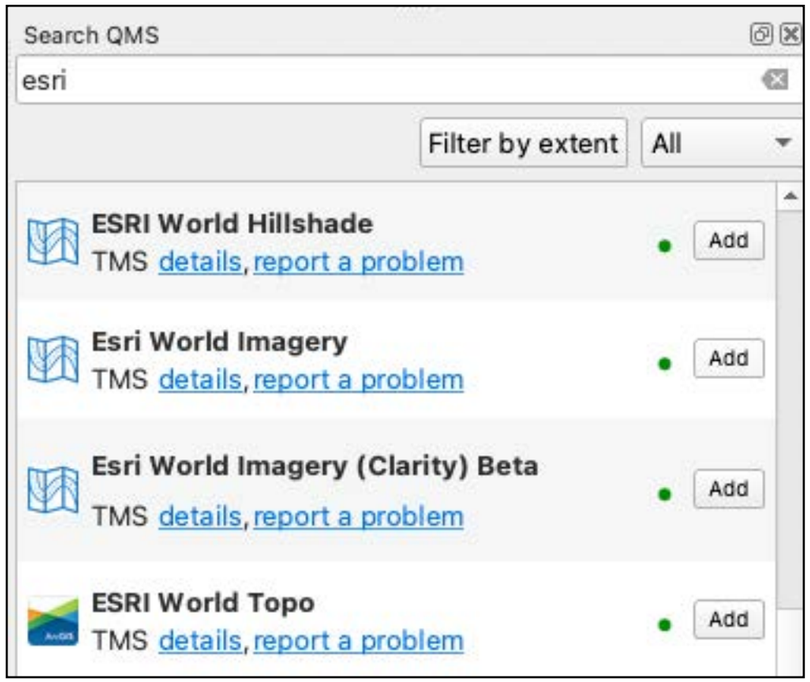
Next, add the OSM standard basemap.

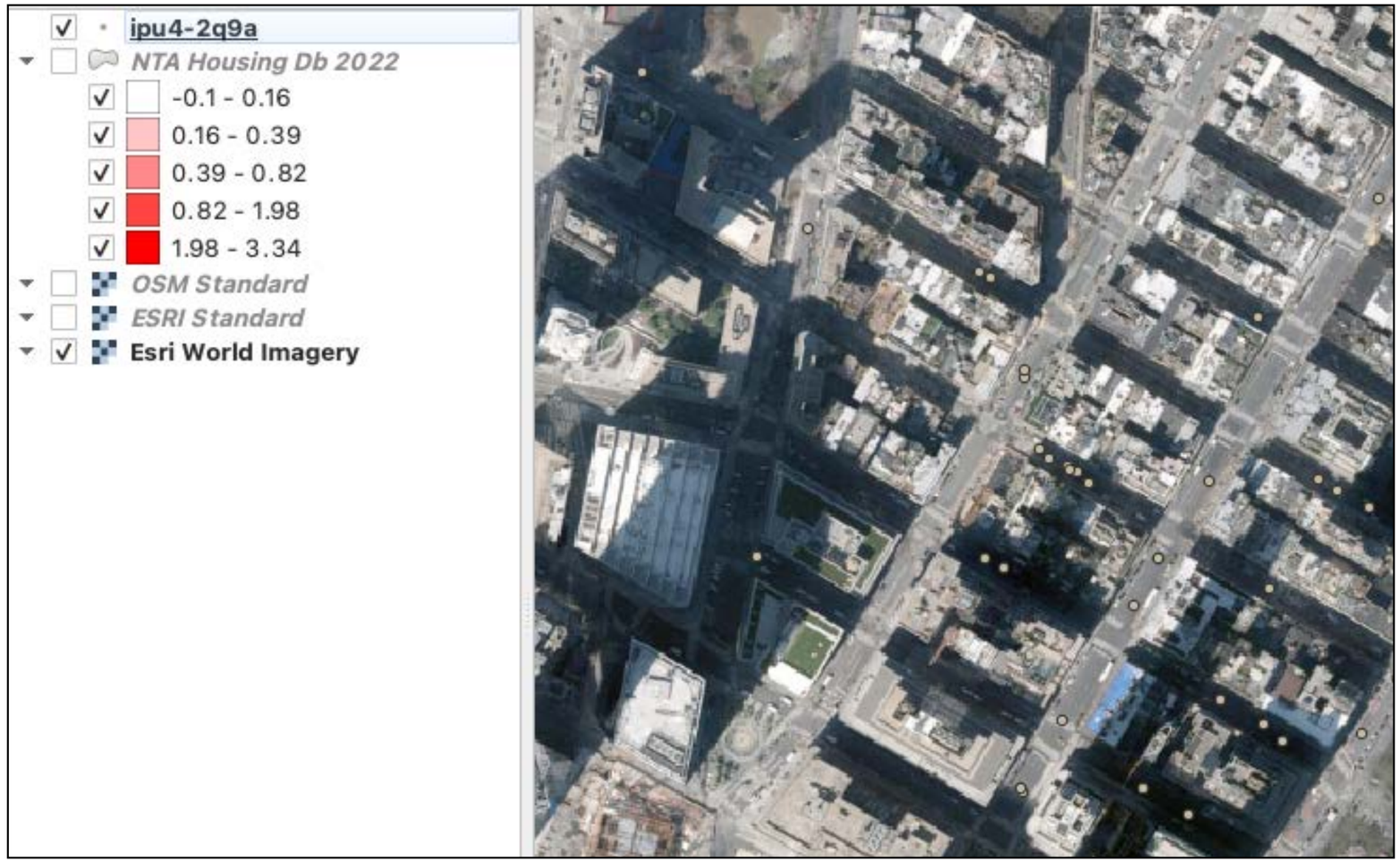
Now we can explore our data with some context.



Using Basemaps

Using the “Search QMS” panel you can add lots of different kids of basemaps including hillshade data, satellite imagery, topography maps, and more.





Add a basemap to your project.

Today

- 🗺️ Introduce **GIS** and mapping software
- 🗄️ Highlight **data sources** in New York City
- 🧰 Discuss **common GIS tasks**
- 🏗️ Introduce **soft site analysis**

859,068

Parcels in NYC (as of June 14 2023)

How can we identify strong development opportunities in this vast city?

How can we identify strong development opportunities in this vast city?

Soft site analysis!

🏗️ What is a soft site?

A **soft site** is a site that represents a strong development opportunity. Whether a site is "soft" will depend on the targeted development product.

Soft site analysis is a body of techniques used to filter out sites unsuitable for development and prioritize others.

🏗️ What is a soft site?

In New York City, a "soft site" also has a specific definition, used in environmental review.

SOFT SITES OR NO-ACTION SITES

Sometimes, projections of development on "soft sites" are appropriate. Soft sites are sites where a specific development is not currently proposed or being planned, but may reasonably be expected to occur by the projected build year. In other words, it may be appropriate to project that development would occur on a site under existing zoning on an "as-of-right" basis in the future No-Action condi-

NYC CEQR Technical Manual, 2021
https://www1.nyc.gov/assets/oec/technical-manual/02_Establishing_the_Analysis_Framework_2021.pdf

🏗️ What is a soft site?

In New York City, a "soft site" also has a specific definition, used in environmental review.

The term "soft site" this term is used by both city government to determine potential development capacity (and their impacts) and by developers looking for opportunity sites.

Finding soft sites is like sifting through recycling

SIMS Recycling
6SQFT

Depending on our criteria, we'll use different tools and identify different types of soft sites

SIMS Recycling
6SQFT

🏗️ Filtering Sites

For example, what filters can we apply to identify sites suitable for medium- to high- density residential development?

🏗️ Filtering Sites

For example, what filters can we apply to identify sites suitable for medium- to high- density residential development?

Strong characteristics / sites to prioritize

- Sites that are zoned as-of-right for medium- to high- density residential (R5 and denser, for example).
- Sites currently vacant or with low existing built density (e.g. surface lots).
- Sites with attractive adjacencies (e.g., retail corridors, transit, schools).
- By specific neighborhood / geography of interest (e.g., based on local market conditions)

Weak characteristics / sites to exclude

- "Overbuilt" sites
- Small sites, which can't accommodate the desired housing product
- Landmarked sites
- Sites that are likely to be rent-stabilized (see NYC CEQR Technical Manual)
- Sites without appropriate frontage or depth
- Newly constructed buildings, particularly with utility uses
- Longstanding institutional uses (e.g., parks, universities, houses of worship)
- Grade challenges (too steed)
- Environmental contamination
- Climate risk / insurance burden (e.g., sites threatened by fire in CA, sea level rise in NYC)
- Irregularly shaped sites

🏗️ What is an "overbuilt" or "underbuilt" site?

Sites that are overbuilt have a built density that exceeds their allowable density.

% Built = Built Density / Allowable Density * 100

🏗️ What is an "overbuilt" or "underbuilt" site?

Sites that are overbuilt have a built density that exceeds their allowable density.

% Built = Built Density / Allowable Density * 100

Sites with that are overbuilt or near their maximum allowable built density are not likely to represent a strong development opportunity.

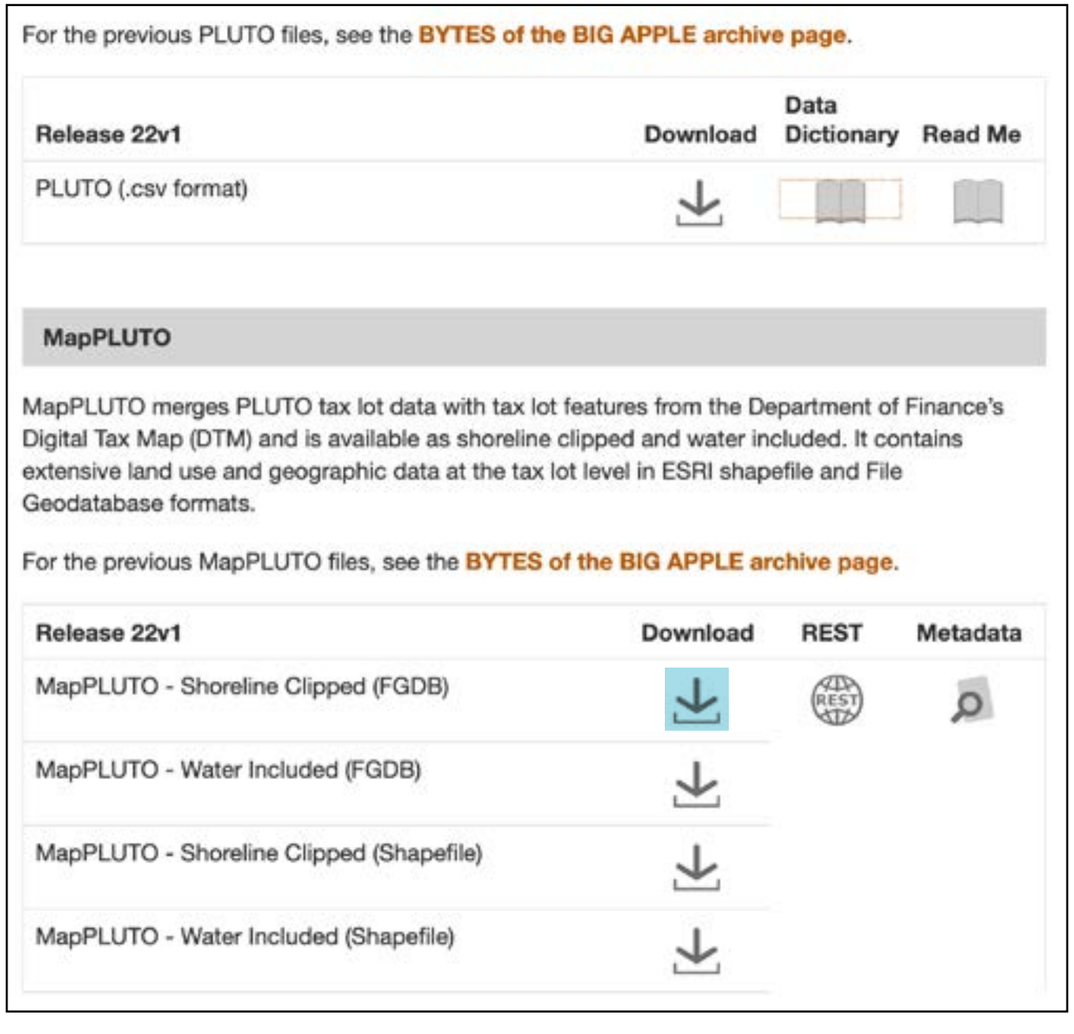
🏗️ What is an "overbuilt" or "underbuilt" site?

Sites with that are underbuilt have an allowable built density that greatly exceeds the existing built density.

% Built = Built Density / Allowable Density * 100

🔢 Download PLUTO data

Let's download every parcel in the City: https://www1.nyc.gov/site/planning/data-maps/open-data/dwn-pluto-mappluto.page



For the previous PLUTO files, see the **BYTES of the BIG APPLE archive page.**

Release 22v1	Download	Data Dictionary	Read Me
PLUTO (.csv format)			

MapPLUTO

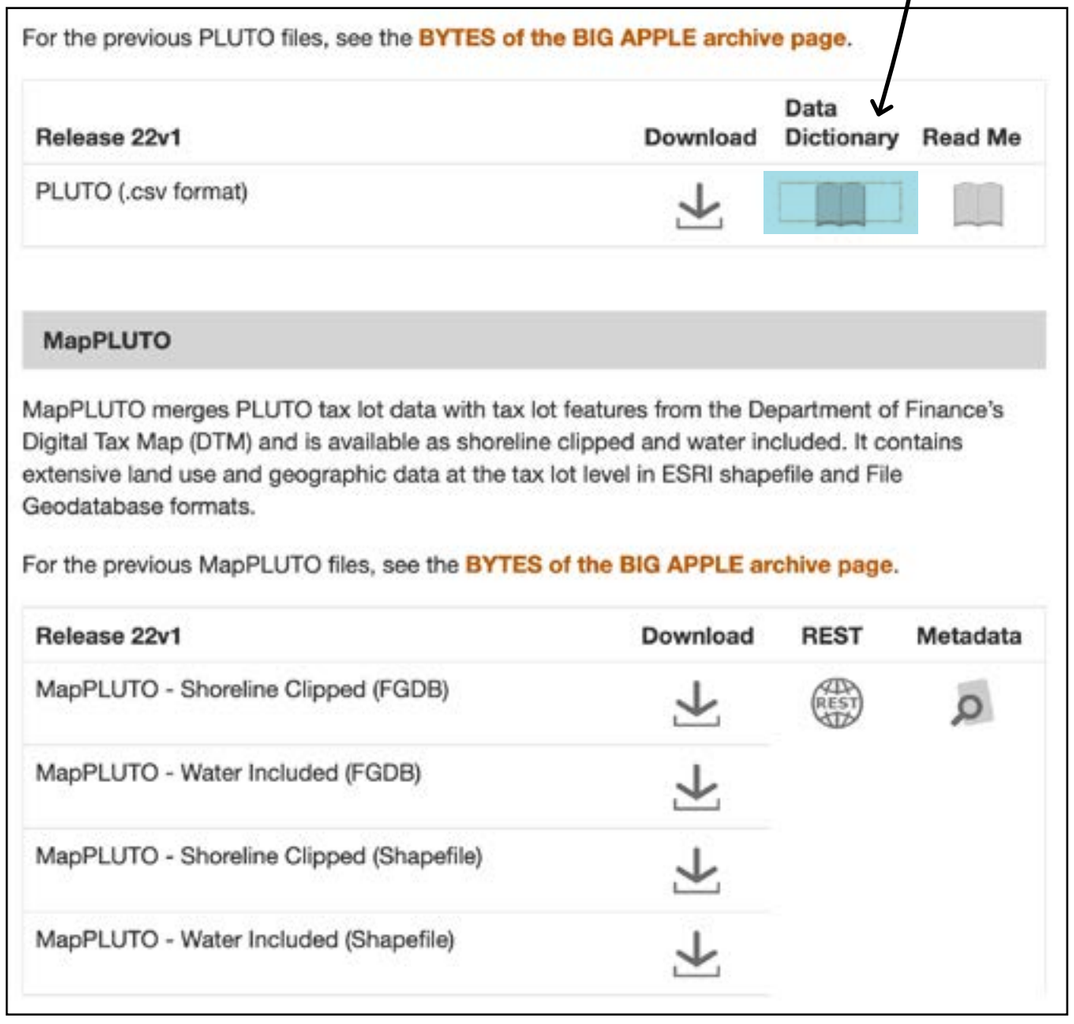
MapPLUTO merges PLUTO tax lot data with tax lot features from the Department of Finance's Digital Tax Map (DTM) and is available as shoreline clipped and water included. It contains extensive land use and geographic data at the tax lot level in ESRI shapefile and File Geodatabase formats.

For the previous MapPLUTO files, see the **BYTES of the BIG APPLE archive page.**

Release 22v1	Download	REST	Metadata
MapPLUTO - Shoreline Clipped (FGDB)			
MapPLUTO - Water Included (FGDB)			
MapPLUTO - Shoreline Clipped (Shapefile)			
MapPLUTO - Water Included (Shapefile)			

🔢 Download PLUTO data

Don’t forget to download PLUTO’s documentation



For the previous PLUTO files, see the **BYTES of the BIG APPLE archive page**.

Release 22v1	Download	Data Dictionary	Read Me
PLUTO (.csv format)			

MapPLUTO

MapPLUTO merges PLUTO tax lot data with tax lot features from the Department of Finance’s Digital Tax Map (DTM) and is available as shoreline clipped and water included. It contains extensive land use and geographic data at the tax lot level in ESRI shapefile and File Geodatabase formats.

For the previous MapPLUTO files, see the **BYTES of the BIG APPLE archive page**.

Release 22v1	Download	REST	Metadata
MapPLUTO - Shoreline Clipped (FGDB)			
MapPLUTO - Water Included (FGDB)			
MapPLUTO - Shoreline Clipped (Shapefile)			
MapPLUTO - Water Included (Shapefile)			

Download the PLUTO dataset here.

🏗️ Overbuilt in NYC

The PLUTO dataset provides a set of columns that make it easy to calculate whether a site is underbuilt or overbuilt. For residential development, the formula is:

% Built = "BldgArea" / ("ResidFAR" * "LotArea") * 100

Built Floor Area (SqFt)

Estimate of allowable Zoning Floor Area (ZFA in SqFt)

🤹Thank you!

Please don’t hesitate to reach out with questions
raphael@wxystudio.com

Questions

Assignment