

TRANSPORTATION DIVISION DATA ANALYTICS TOOLKIT



- **Data Processing**

- Excel
- Python
- R

- **Geospatial Data Processing**

- ArcGIS
- QGIS

- **Interactive Chart**

- Tableau
- Plotly

- **Interactive Map**

- ArcGIS Online

- Carto

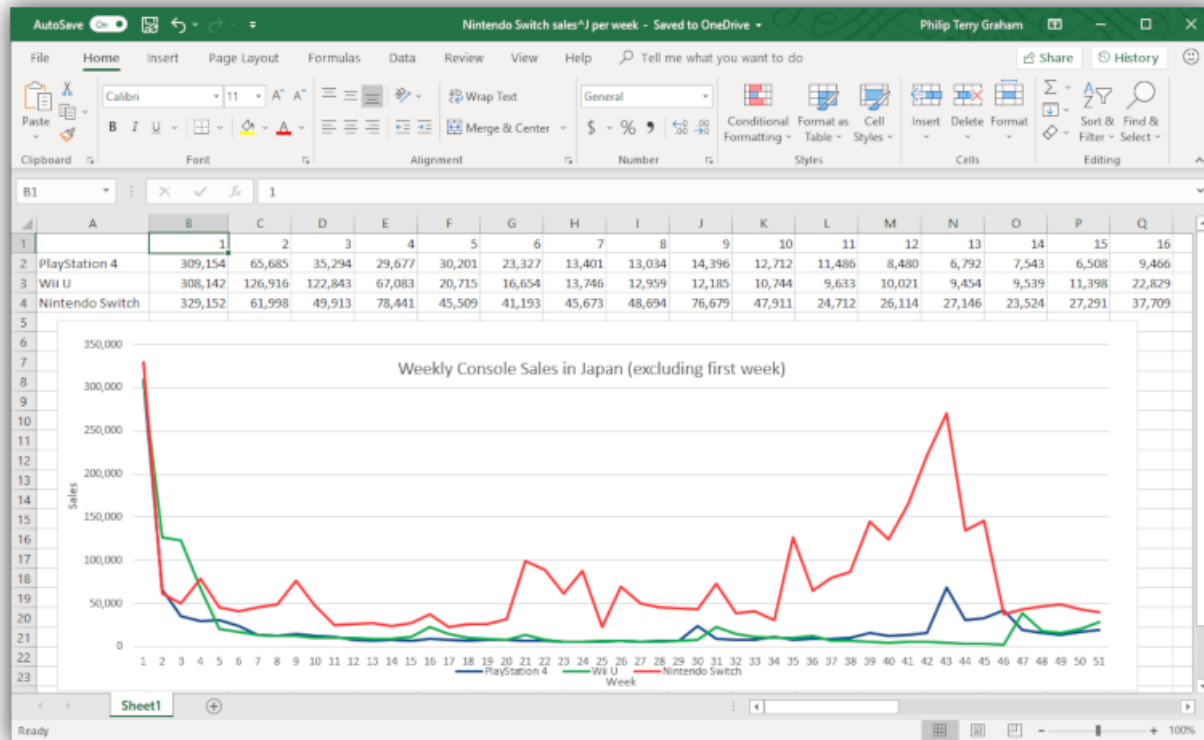
- Mapbox GL JS

- **Database**

- MS SQL Server
- SQLite / SpatiaLite
- PostgreSQL / PostGIS
- DBeaver

- **Other Tools**

- GitHub
- Visual Studio Code
- Geosupport
- DigitalOcean

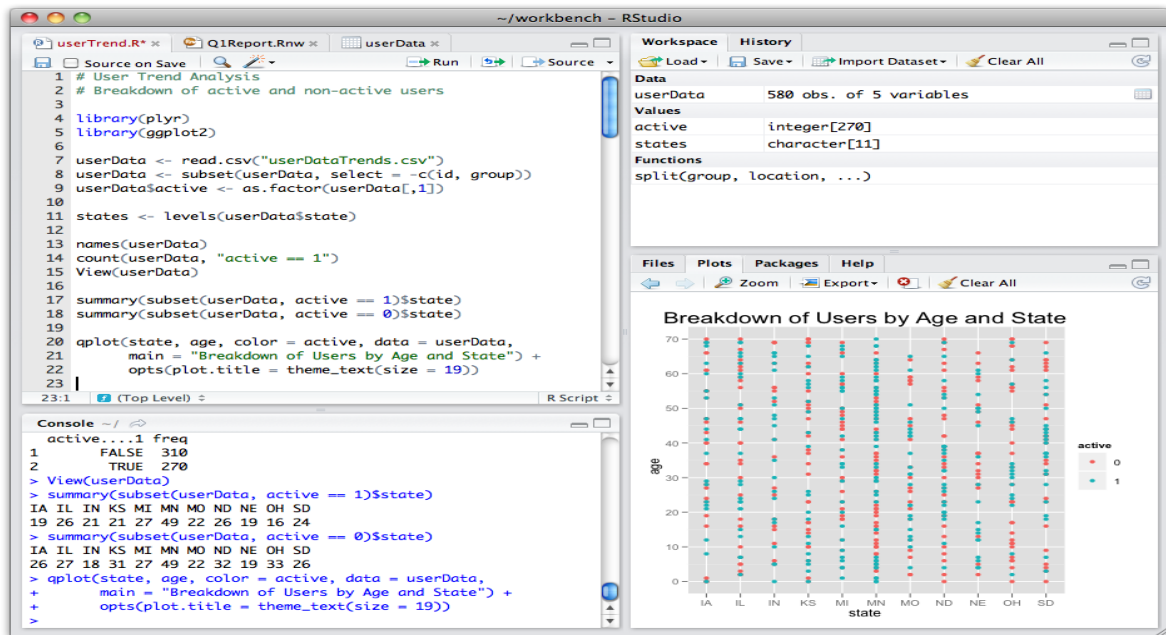


- Data analytics software by Microsoft
- Creating, reading, writing, manipulating, analyzing, and visualizing relatively small datasets
- Formulas (SUM, IF, VLOOKUP, etc.)
- Pivot table
- Static chart / map
- Data Analysis Tools
- Macros with VBA

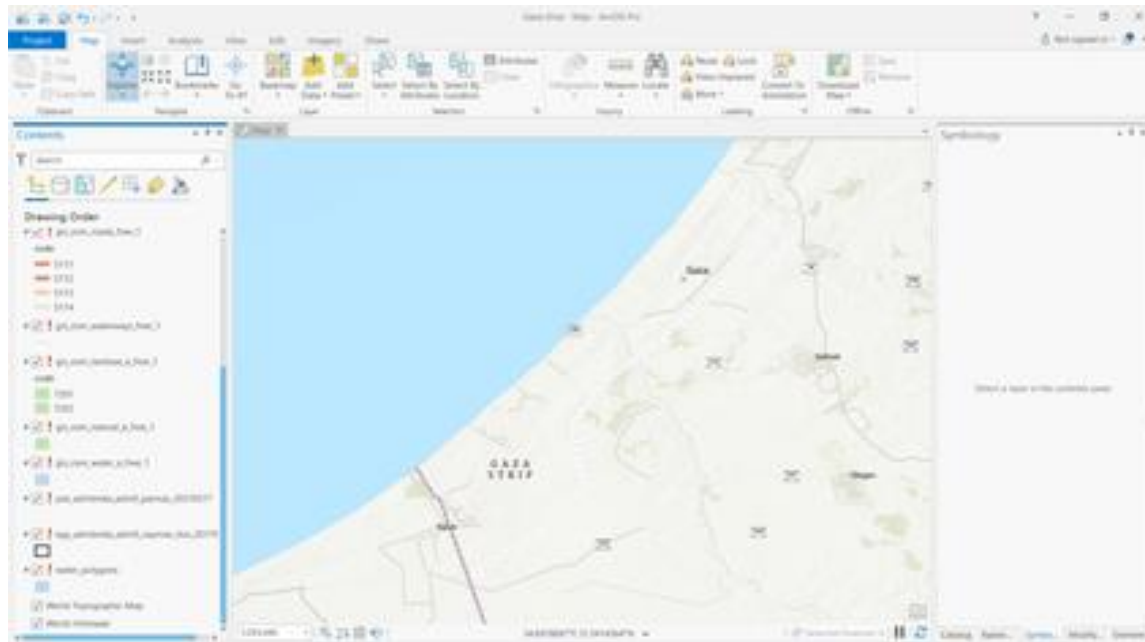


```
1 #%% Settings
2 import requests
3 import pandas as pd
4 import numpy as np
5 import geopandas as gpd
6
7
8 pd.set_option('display_max_columns', None)
9 path='C:/Users/mayij/Desktop/'
10 path='C:/Users/Y_Ma2/Desktop/'
11
12 apikey=pd.read_csv(path+'GITHUB/td-acsapi/secrets.csv', dtype=str).loc[0, 'value']
13 usernm=pd.read_csv(path+'GITHUB/td-acsapi/secrets.csv', dtype=str).loc[1, 'value']
14 passwd=pd.read_csv(path+'GITHUB/td-acsapi/secrets.csv', dtype=str).loc[2, 'value']
15
16 p={'http': 'http://'+str(usernm)+''+str(passwd)+'@dcproxy1.dcp.nycnet:8080',
17   'https': 'http://'+str(usernm)+''+str(passwd)+'@dcproxy1.dcp.nycnet:8080'}
18
19
20
21
22 #%% County List
23 tp=[]
24 for i in ['times square', 'dbk', 'broadway junction', 'bhub', 'fordham', 'morris park', 'e14st', 'e23st', 'e125st', 'w125s']
25     k=pd.read_csv(path+'Blueprint/travelshed/'+str(i)+'/'+str(i)+'.csv')
26     k['county']=[str(x)[0:5] for x in k['tractid']]
27     kek['county'].unique()
28     tp+=list(k)
29 tpset(tp)
30
31
32 #%% All Locations
33 df=pd.DataFrame()
34 quadstatectclipped=gpd.read_file(path+'Blueprint/travelshed/quadstatectclipped.geojson')
35 for i in ['dbk', 'broadway junction', 'bhub', 'fordham', 'morris park', 'e14st', 'e23st', 'e125st', 'w125s', 'lic', 'jamaic']
36     tpsd=pd.read_csv(path+'Blueprint/travelshed/'+str(i)+'/'+str(i)+'.csv', dtype=str)
37     tps=[t['tractid']]
38     dfpd.concat([df, tps], axis=0, ignore_index=True)
39 df=df.drop_duplicates(keep='first').reset_index(drop=True)
40 df.to_csv(path+'Blueprint/travelshed/all/all.csv', index=False)
41 dfpd.merge(quadstatectclipped, df, how='inner', on='tractid')
42 df.to_file(path+'Blueprint/travelshed/all/all.geojson', driver='GeoJSON')
```

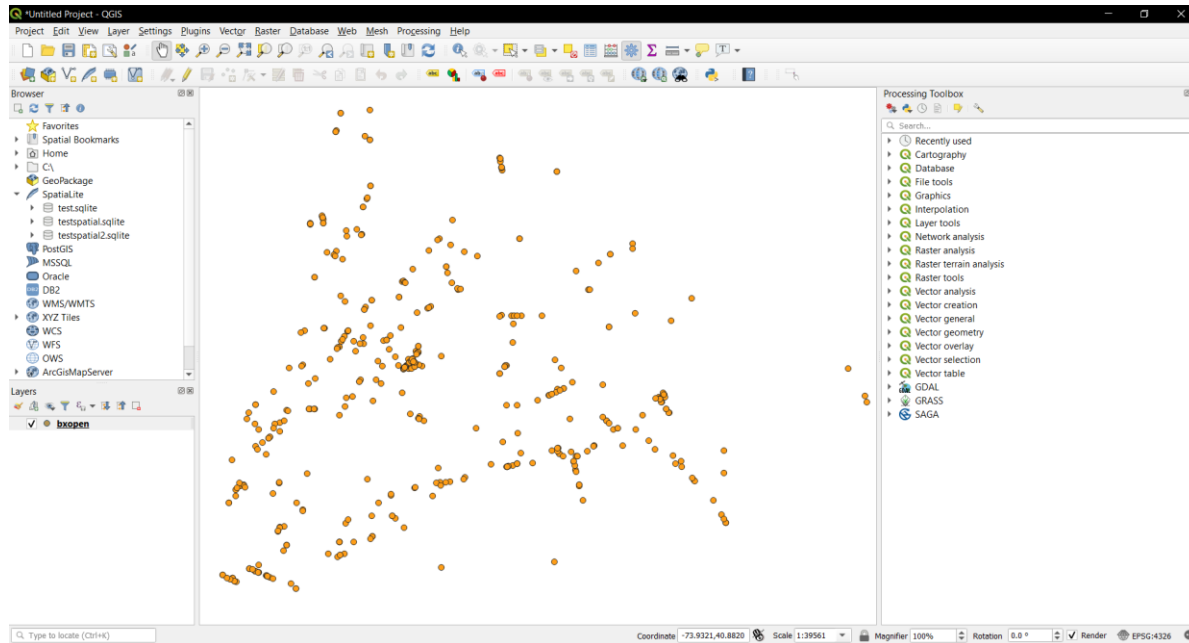
- Open-source general-purpose programming language
- Reading, writing, manipulating, analyzing, and visualizing relatively big data
- Reproducibility and transparency
- IDE: Spyder / Jupyter
- Data wrangling: Pandas
- Spatial data: GeoPandas
- Interactive visualization: Plotly
- Dashboard / App: Streamlit / Dash



- Open-source programming language optimized for statisticians
- Reading, writing, manipulating, analyzing, and visualizing relatively big data
- IDE: RStudio
- Data wrangling: tidyverse (dplyr)
- Spatial data: sf
- Interactive visualization: Plotly
- Dashboard / App: R Markdown / R Shiny



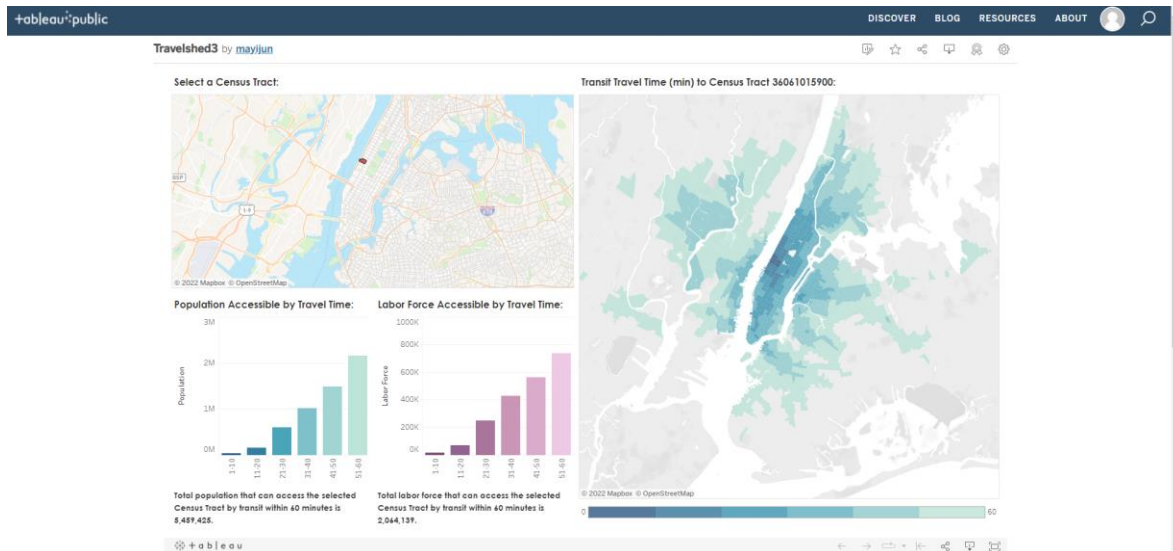
- Proprietary GIS software by ESRI
- Creating, reading, writing, manipulating, analyzing, and visualizing geospatial data
- Data format: Shapefile, Geodatabase, etc.
- ArcGIS Pro replacing ArcMap



- Open-source GIS software
- Creating, reading, writing, manipulating, analyzing, and visualizing geospatial data
- Data format: Shapefile, GeoJSON, etc.
- Toolbox and plugins for geoprocessing

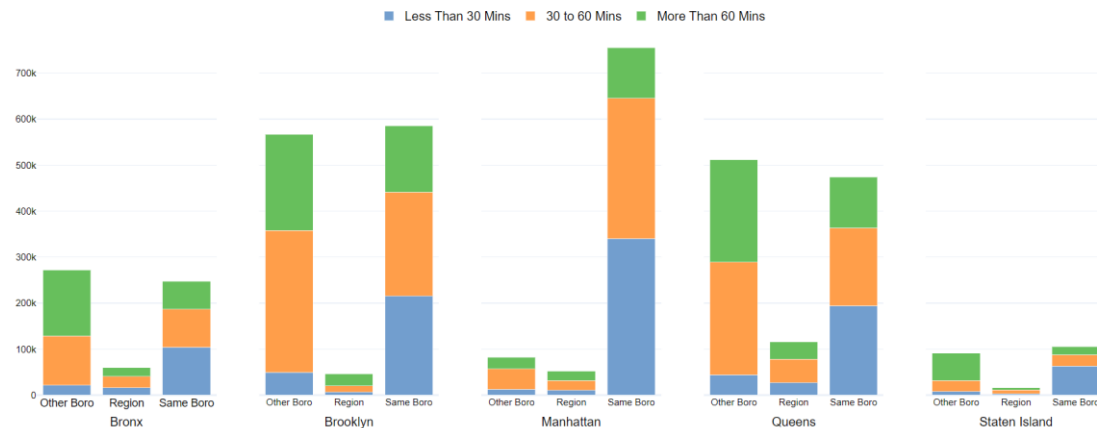


- Data visualization software owned by Salesforce
- Creating dashboards with interactive charts and maps
- Tableau Public (free) for completely public data visualization
- Tableau Desktop / Online (require license) for non-public data





Destination of Work by Borough of Residence for NYC Commuters



- Open-source data visualization tools
- Creating interactive charts and maps
- Multi-platform graphing libraries: Python, R, JavaScript, etc.
- Dash for creating dashboards / apps
- Hosting platform: GitHub Pages



ArcGIS Online

Mapping and analysis: location intelligence for everyone

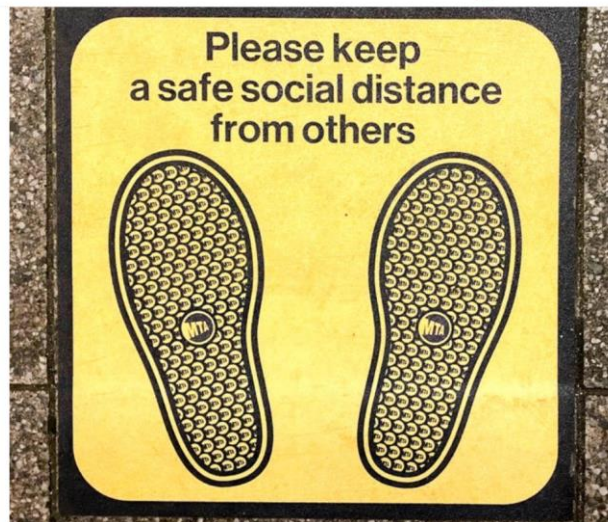
2020 Travel Trends



2020 Travel Trends

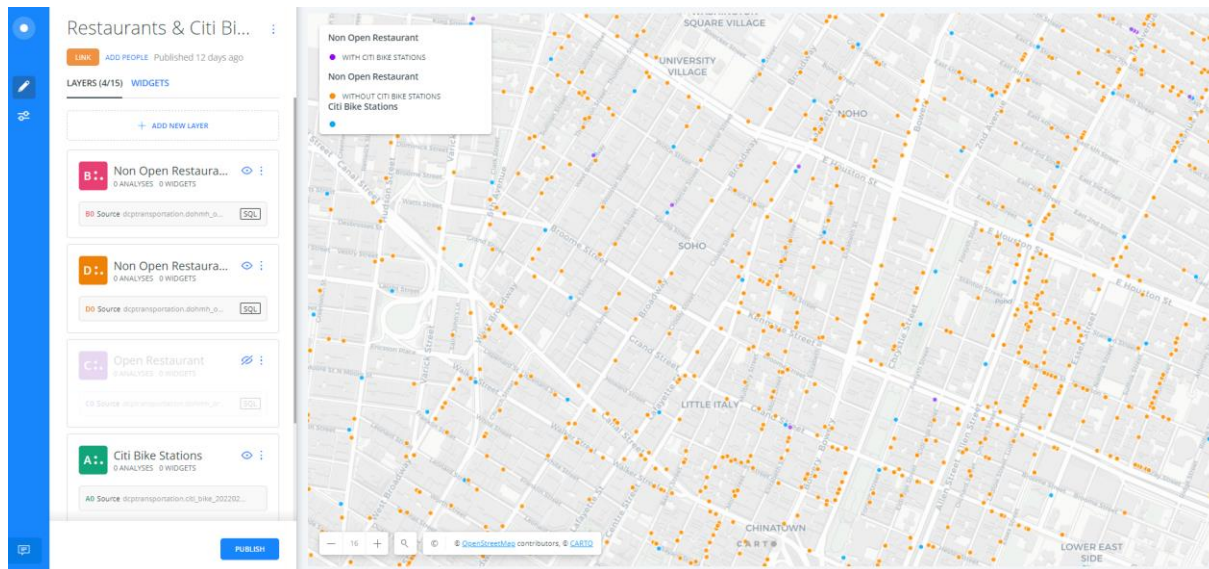
How the COVID-19 pandemic affected mobility in New York City

NYC Department of City Planning, March 2021

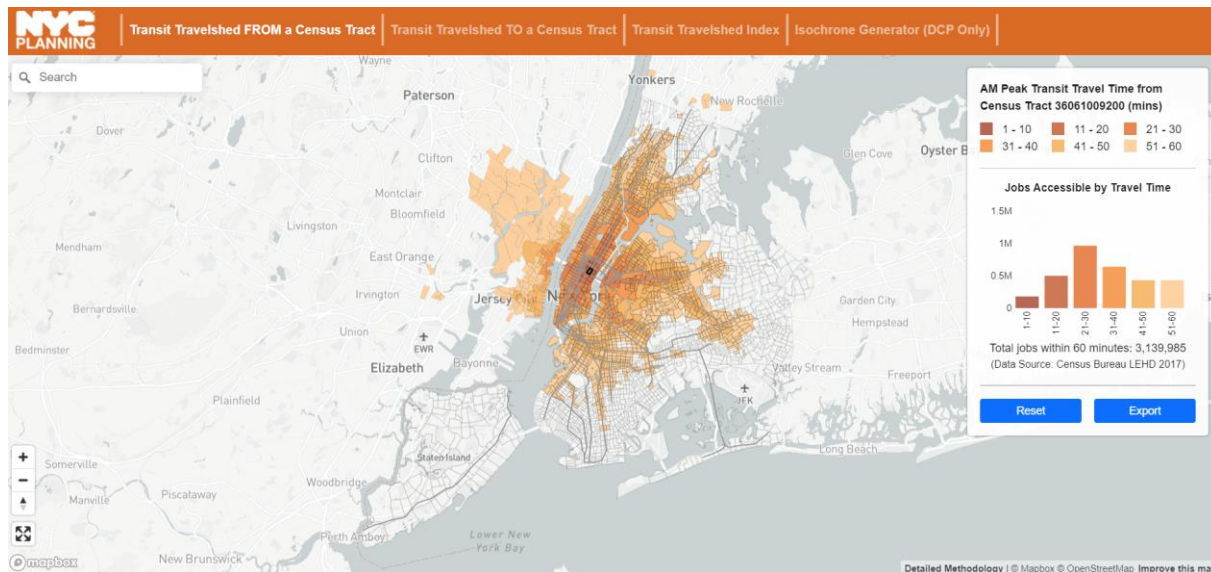


- **Web-based mapping software by ESRI**
- **Creating interactive charts, maps, dashboards, story maps, etc.**
- **Using StoryMaps to organize final report with interactive charts / maps from other platforms embedded**

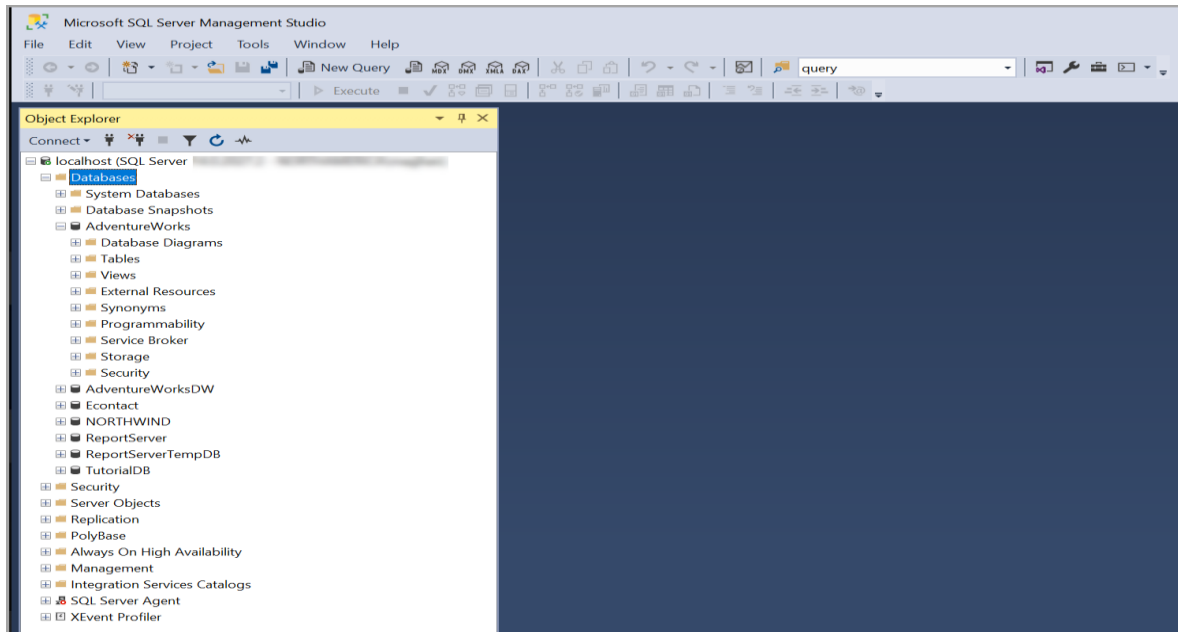
CARTO



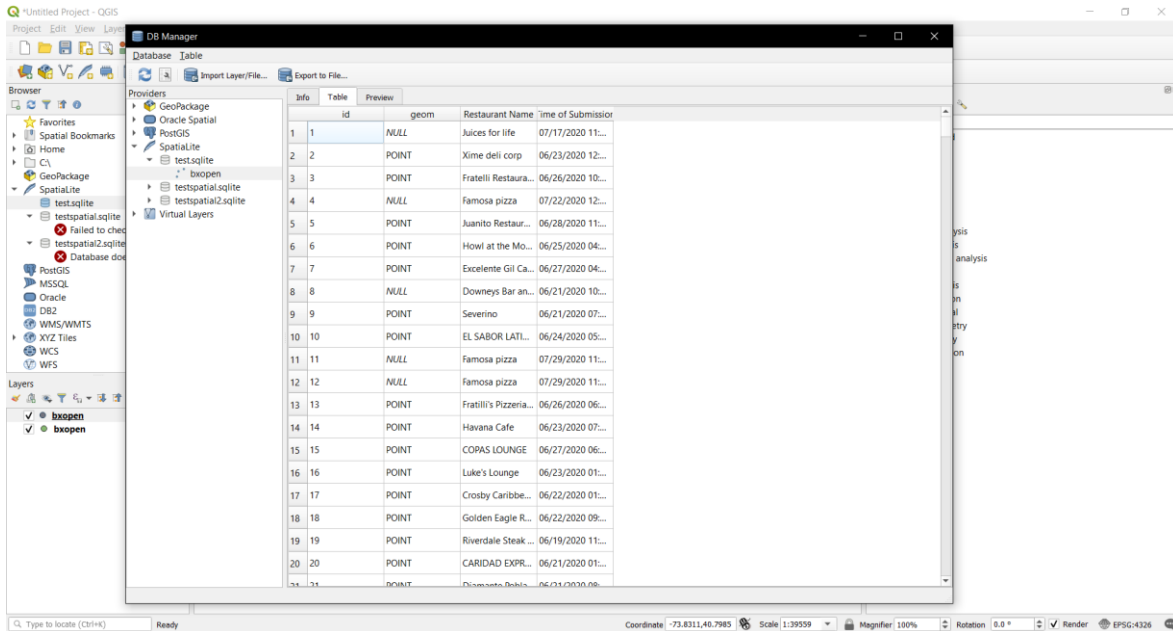
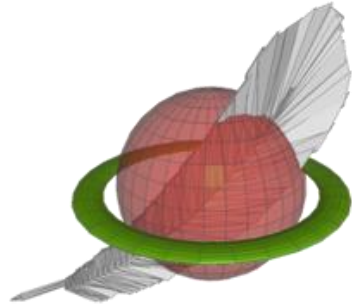
- Web-based mapping platform
- Storing geospatial data and creating interactive maps
- Customization: SQL for data processing, CSS for styling, and HTML for legend editing



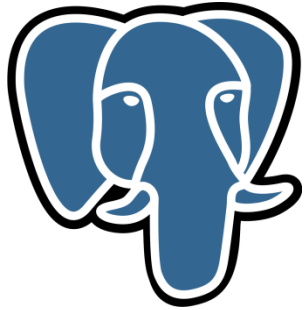
- JavaScript mapping library
- Creating interactive maps and highly customized web mapping apps
- Editing tool: Visual Studio Code
- Languages used: HTML, CSS, and JavaScript
- Hosting platform: GitHub Pages



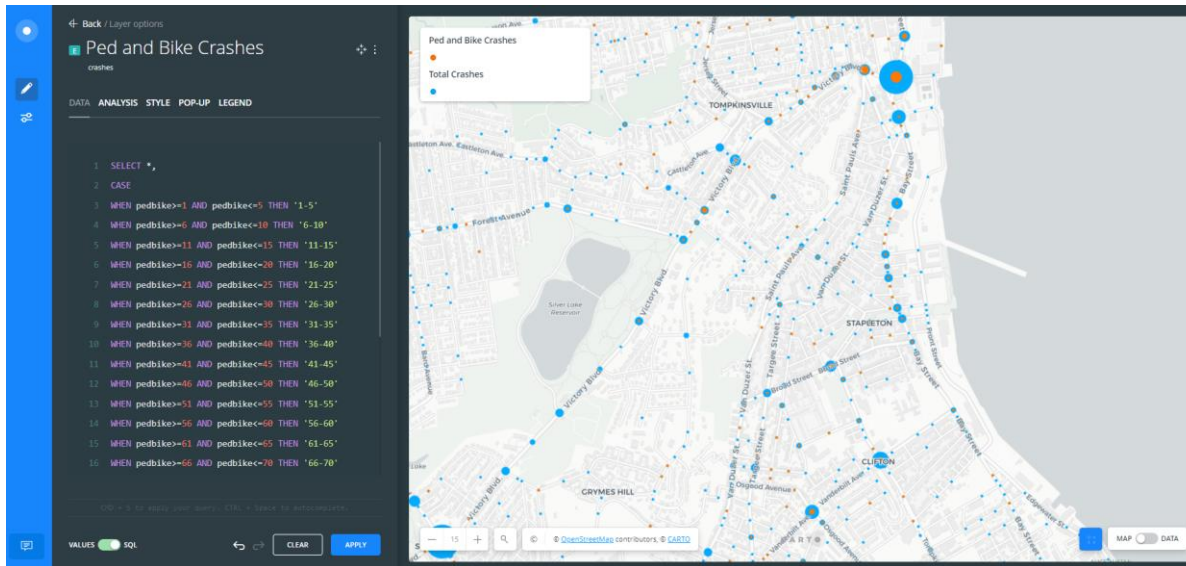
- Commercial client/server relational database management system by Microsoft
- Deployed and managed by ITD
- Storing confidential datasets (TLC data)
- Interface: SQL Server Management Studio / DBeaver

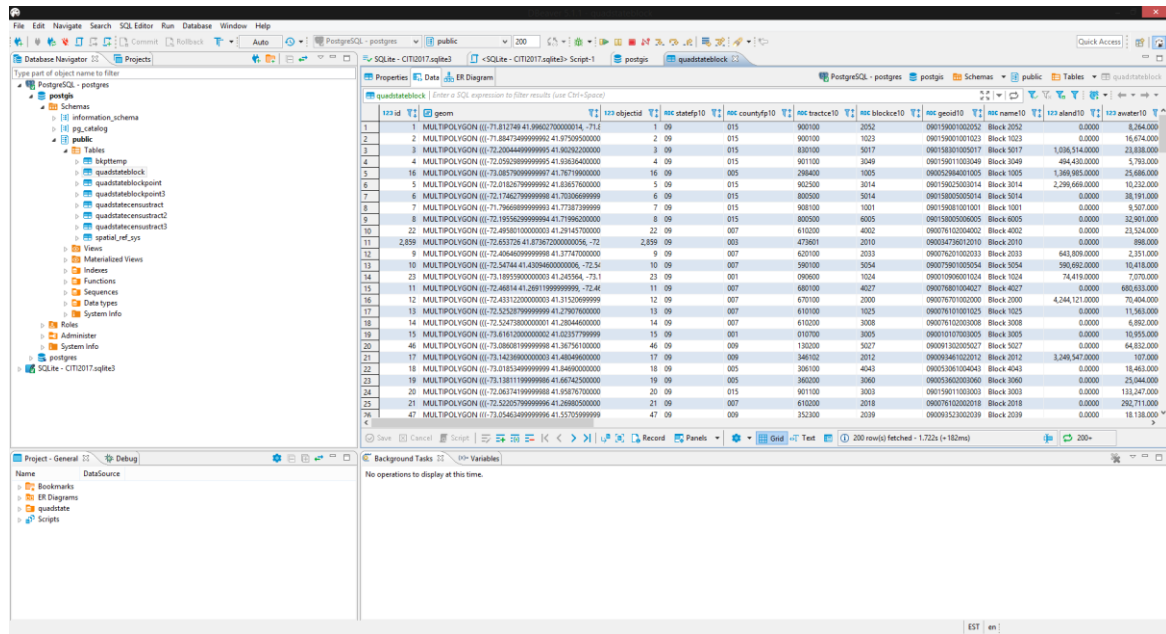


- SQLite: open-source file-based relational database management system
- SpatiaLite: spatial extension to SQLite
- Storing large public datasets into single database
- Interface: DBeaver / QGIS



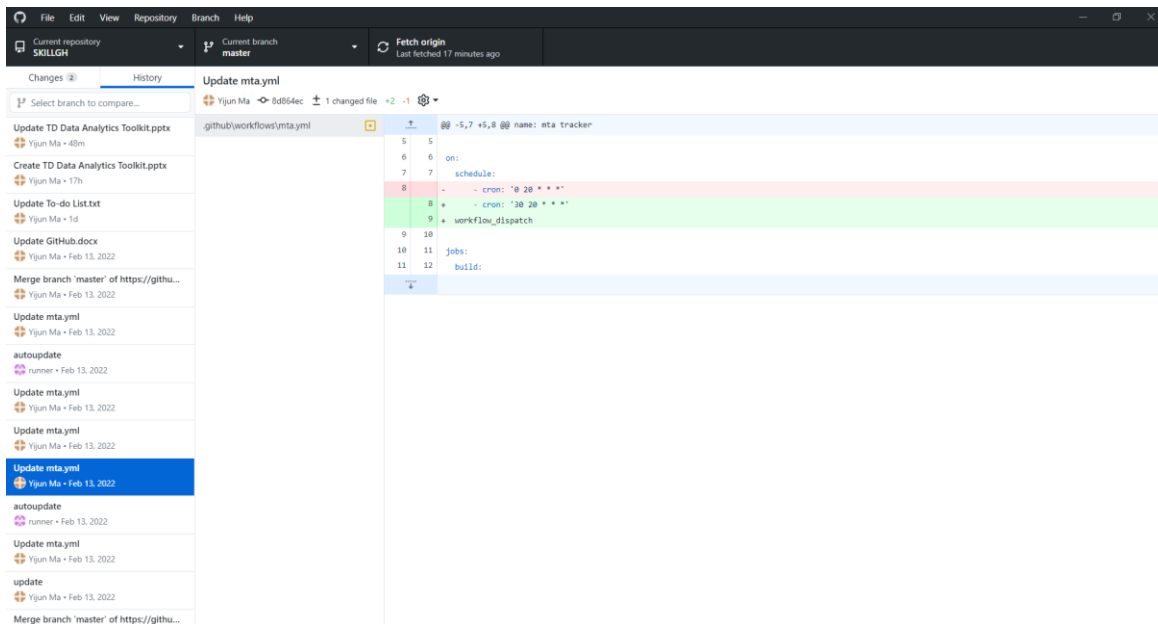
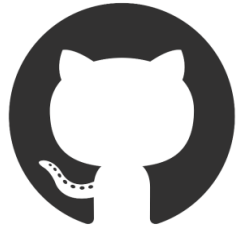
- PostgreSQL: open-source client/server relational database management system
- PostGIS: spatial extension to PostgreSQL
- Storing spatial datasets
- Interface: Carto



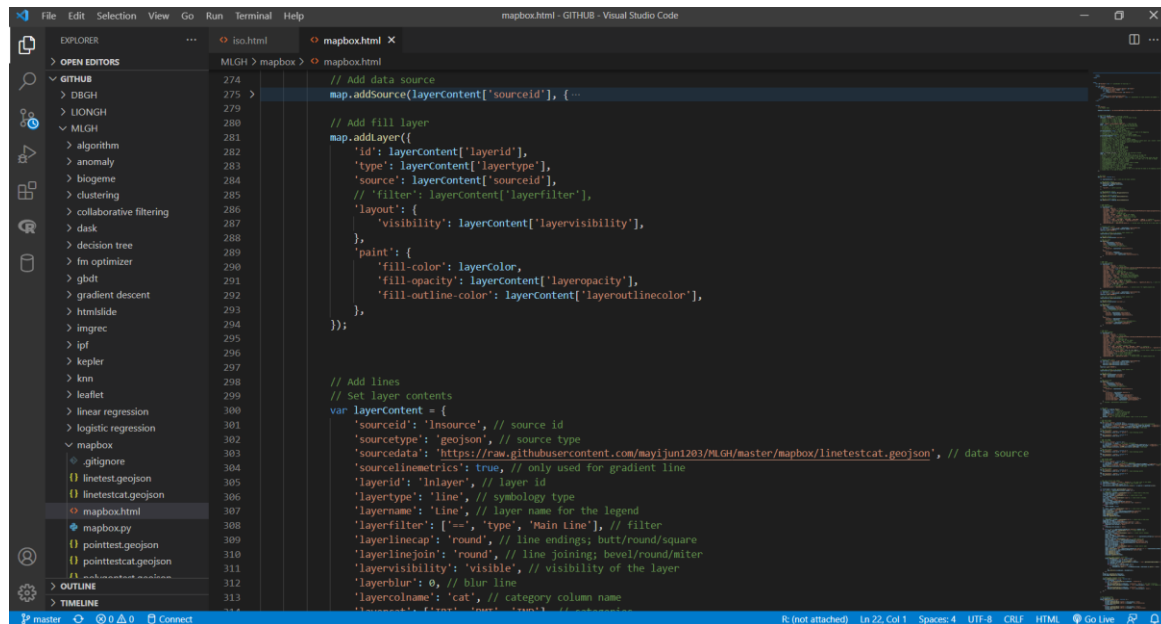
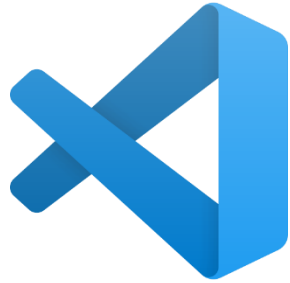


- Open-source database administration tool

- Supports CSV, MS SQL Server, SQLite, PostgreSQL, MySQL, Oracle, IBM DB2, MS Access, etc.



- Version control and source code management platform
- Storing scripts, hosting web apps, automating updates, etc.
- GitHub Desktop: simple tool to pull and push scripts to the cloud
- GitHub Pages: place to host interactive charts / maps and web apps for public sharing
- GitHub Actions: tool to automate the workflows and update the data regularly

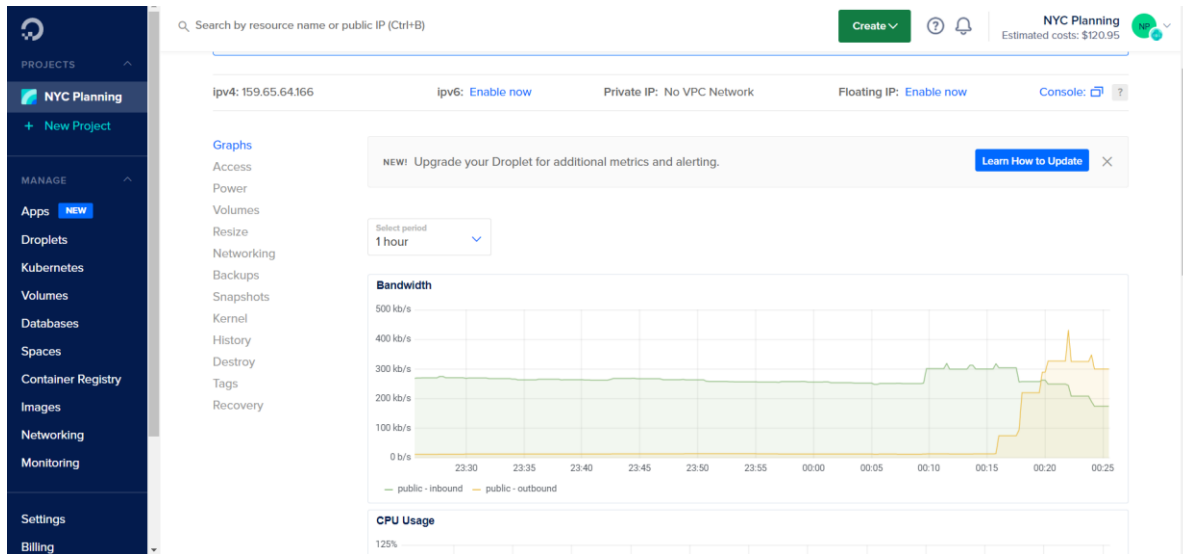


- Source code editor developed by Microsoft
- Editing scripts for creating web apps and testing the apps locally
- Supported languages: HTML, CSS, JavaScript, SQL, Python, R, etc.



- Geocoding package for NYC addresses
- GOAT: web app for geocoding individual address
- GBAT: batch geocoding multiple addresses
- Geoclient (DOITT) / Geoservice (DCP): geocoding API
- Python-Geosupport: python package to geocode large datasets (especially confidential data)

The screenshot shows the GOAT (Geographic Online Address Translator) web application. The header includes the GOAT logo and the text 'Geographic Online Address Translator' on the left, and the 'NYC PLANNING GEOSUPPORT' logo on the right. The main content area is titled 'Display Street and Property Information by Address'. It features a form with three input fields: 'Select a Borough' (a dropdown menu set to 'Manhattan'), 'Address Number' (a text box with '120'), and 'Street or Place Name' (a text box with 'broadway'). There are 'Submit' and 'Create Link' buttons to the right of the 'Street or Place Name' field. Below the form is a 'Unit' input field and a 'Function Options' button. At the bottom of the form, there are several tabs: 'Centerline Level Information', 'City Service Information', 'Political Information', 'Property Level Information', 'Address Range', 'Map', and 'Send Feedback'. The 'Centerline Level Information' tab is selected, showing a result: 'Centerline Level Information for 120 BROADWAY in MANHATTAN'. Below this, there is an 'Orientation' section with the text 'Address is on the right when facing from PINE STREET to THAMES STREET'. On the left side of the interface, there is a vertical navigation menu with options like 'WELCOME', 'ADDRESS FUNCTIONS', 'FUNCTION 1A', 'FUNCTION 1B', 'FUNCTION 1E', 'FUNCTION AP', 'INTERSECTION FUNCTION 2', 'STREET SEGMENT FUNCTION 3', 'STREET STRETCH FUNCTION 3S', and 'BLOCK LOT'.



- Cloud computing service provider
- Streaming data, processing massive datasets, hosting web service, etc.
- Connection tool: PuTTY / SSH
- Data transfer tool: WinSCP / Cyberduck