User's Manual

EValuateNY: Assessing New York State's Electric Vehicle Market

Version 9.0 | April 2022

NEW YORK STATE | **Charge NY**

Release Notes

Version	Release Date	Author	Organization	Notes
1.0	1/15/2016	Nick Nigro	Atlas Public Policy	Initial release
2.0	11/27/2016	Nick Nigro	Atlas Public Policy	Upgrades to add ZIP Code-level data for demographics, reorganize database to be more manageable, add new source for vehicle registrations, add electricity price data, add solar data, add State-level data where possible.
3.0	4/10/2018	Nick Nigro	Atlas Public Policy	Added employment demographics and National Highway System. Streamlined data integration.
3.1	8/3/2018	Nick Nigro	Atlas Public Policy	Updated vehicle registration data, charging station data, and energy price data through July 2018.
4.0	10/13/2018	Nick Nigro	Atlas Public Policy	Updated vehicle registration data, charging station data, and energy price data through September 2018. Revised template and formatting.
5.0	3/6/2019	Nick Nigro	Atlas Public Policy	Updated vehicle registration data, charging station data, and energy price data through February 2019. Added dashboard for charging use of NYSERDA-funded stations through February 2019.
6.0	11/14/2019	Nick Nigro	Atlas Public Policy	Updated vehicle registration data, charging station data, charging use, and energy price data through October 2019. Added EV market share by ZIP Code beginning in 2016 (when DMV snapshots were archived by Atlas). Updated demographics data through 2017. Refresh user experience. Removed solar data since it is no longer supported.
7.0	5/7/2020	Nick Nigro	Atlas Public Policy	Updated vehicle registration data, charging station data, charging use, and energy price data through April 2020. Updated demographics data through 2018. Removed historical datasets no longer kept up to date. Added REDC regions.

Release Notes continued

8.0	2/3/2021	Josh Rosenberg	Atlas Public Policy	Added Drive Clean Vehicle Rebate Data. Improved accuracy of charging location data, particularly for sites with more than one charging connector. Updated data to include Medium-Duty (Class 3-6) and Heavy-Duty (Class 6-8). Added capability to filter by vehicle class, category, and use case. Updated dashboard to include the share of EV original registrations. Updated vehicle registration and charging location data through February 1, 2021.
9.0	3/23/2022	Josh Rosenberg	Atlas Public Policy	Added disadvantaged community mapping. Refreshed dashboard layout. Updated calculation for vehicle registrations for investor-owned utility territories to be proportional to customers per ZIP Code. Updated calculation for light-duty vehicles to improve accuracy. Updated charging use graphics to show average utilization.

Table of Contents

List of Figures

1 Introduction

This user manual for the EValuateNY tool consists of three sections. The first provides a brief introduction to its purpose and uses. The second defines the types of data included and offers a brief rationale for the information as well as explains the tool's architecture. The third section provides a brief overview of EValuateNY's key features and considerations for when to use the Microsoft Excel and Power BI interfaces. Appendices offer detailed explanations of EValuateNY's relational database and additional details on data sources.

1.1 About EValuateNY

EValuateNY is a software application that was developed to help identify the most effective activities to accelerate plug-in electric vehicle (EV) adoption in New York State. The tool relies on Microsoft Excel and Power BI (a data visualization application that can be downloaded free online) to assemble several categories of information related to the EV market. The EValuateNY tool can be used through Microsoft Excel and Power BI on Windows or Mac computers, as well as tablets and mobile devices. Both the Excel and Power BI interfaces allow users to access information stored in several individual databases through a relational database. A relational database is a method of organizing and comparing data from varying sources to connect each data type (see section 2 for an explanation of relational databases). Visit https://www.nyserda.ny.gov/All-Programs/ChargeNY/Support-Electric/Data-on-Electric-Vehicles-and-Charging-Stations to download EValuateNY as an Excel or Power BI file.

Through a user-friendly interface, EValuateNY allows users to create compelling visualizations and gain insights about the EV market. Users may also access the raw data stored in its relational database to construct new visualizations and gain further insights. The tool was also designed to be easily updated to provide timely information and add new types of data. Power BI is a relatively new program that Microsoft updates monthly. Access to EValuateNY through Power BI requires the Power BI desktop application or a user account, both of which are free.

The tool is intended for policymakers, businesses, academics, and research analysts. Both interfaces allow novice users to interact with the data through existing charts and tables and to derive unique insights into New York State's EV market. Expert analysts may dive deeper to create and customize charts and tables that offer new and innovative insights.

1.2 How EValuateNY Got Started

Currently, EVs are promoted through various public and private initiatives, such as financial incentives for charging equipment and city requirements to provide EV charging access in new garages and parking lots. These initiatives are intended to encourage the growth of the EV market, yet the efficacy of each approach is uncertain. Policymakers and businesses looking to grow the EV market are wondering what's working, and the tool aims to provide data-based answers to that key

question. For example, the tool may offer insights into how New York City's requirement that 20 percent of new eligible parking be ready for charging has affected EV adoption in and around the City.

EValuateNY organizes information about public and private State initiatives to promote EVs with relevant market data in a single database. The information is sorted by location over time to facilitate a user's understanding of the EV market's evolution. Specifically, the tool contains data on EV registrations, charging infrastructure availability and usage, and demographics. Much of the information it contains is publicly available. Information can be accessed directly through the inherent functions of Excel, or by using Power BI for additional data visualization to create sophisticated charts, maps, and tables as shown in Figure 1.

Figure 1. Screenshot of EValuateNY Using the Microsoft Power BI Interface
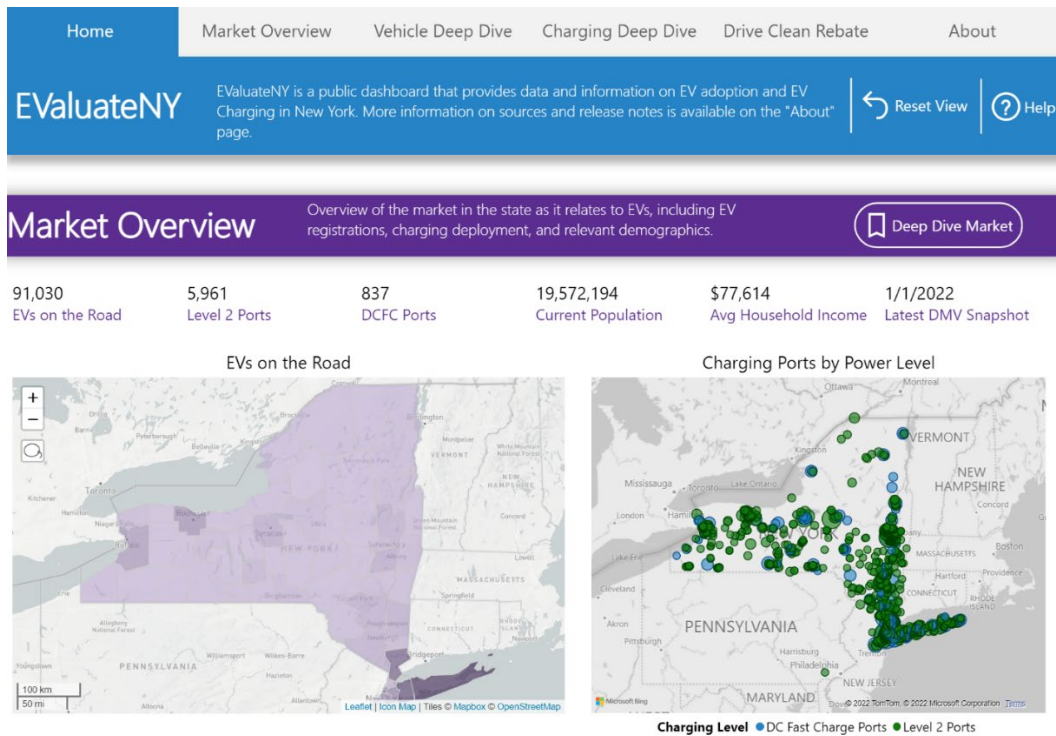


1.3   Contributing Partners

The New York State Energy Research and Development Authority (NYSERDA) provided a grant to develop EValuateNY and helped oversee its development. Atlas Public Policy and the Center for Climate and Energy Solutions (C2ES) created version 1.0 of the tool, and Atlas has maintained the tool since its first release.

Several project partners assisted in identifying and contributing information—some of which is proprietary. Members of the following organizations provided assistance: ChargePoint, General Motors, National Grid, the National Renewable Energy Laboratory (NREL), the New York City Department of Transportation, and the New York Power Authority. The New York State Department of Transportation and the New York State Department of Public Service also provided data for the tool.

Box 1. What's a Relational Database?

A relational database links different data tables by incorporating unique "keys" into the tables. These keys allow users to compare similar information in multiple tables.

For example, a three-column table could describe a state's annual median household income by county. In this example, the columns are year, county, and median household income. There would be a row for each of New York State's counties for every year that data is available. Using a relational data model, this table could be compared to another table that tracked the number of Clean Passes issued by county on a monthly basis. To connect these tables, two additional tables must be created: one table defines all the dates in the median household and EV deployment tables, and the other defines all the counties in the State. See the figure below for an illustration of these connections:
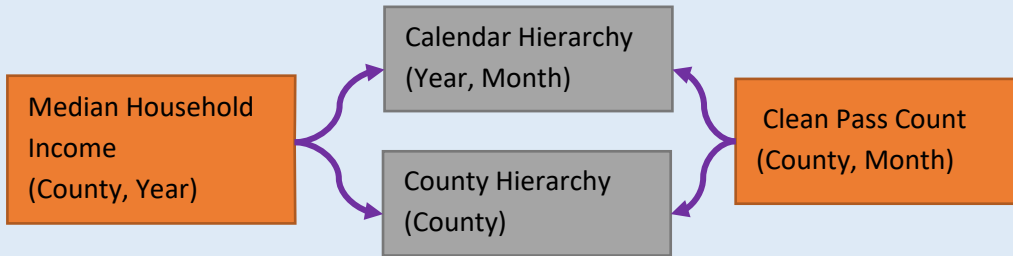
In this example, dates connect the time series data and counties are used to connect the geolocation data. The figure below illustrates the power of this concept using Microsoft Excel. Using relational databases allows different types of data to be compared quickly and easily. And, since each data type is stored in its own Excel table, data maintenance is easier.

Source: https://en.wikipedia.org/wiki/Relational_database
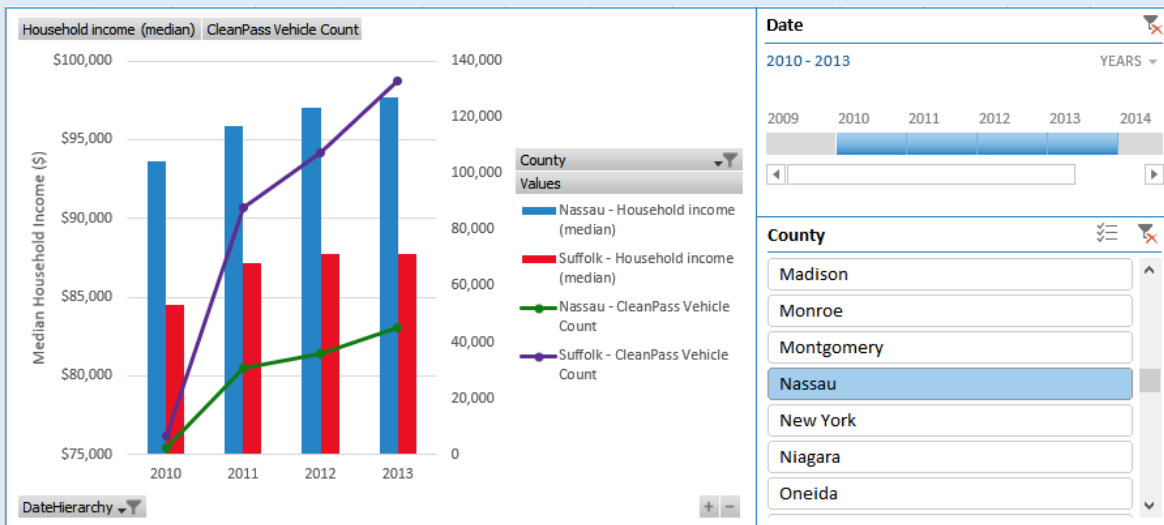
2   Design and Construction

This section describes the architecture of EValuateNY, provides an overview of the data contained in the tool, and examines possible future enhancements to EValuateNY.

2.1   Architecture

EValuateNY provides two ways to interact with the information stored in its comprehensive database as illustrated in Figure 2. The tool can be used through Microsoft Excel and Power BI on Windows or Mac computers, as well as tablets and mobile devices. Both the Excel and Power BI interfaces allow users to access information stored in several individual databases through a relational database. A relational database is a method of organizing and comparing data from varying sources to connect each data type (see Box 1 for an explanation of relational databases). Data in the tool is stored in a single Microsoft Excel file or retrieved directly from web-based application programming interfaces (APIs), simplifying maintenance and use. These data files are filtered through the relational database and made available within the two main interfaces.

Figure 2. Architecture for EValuateNY
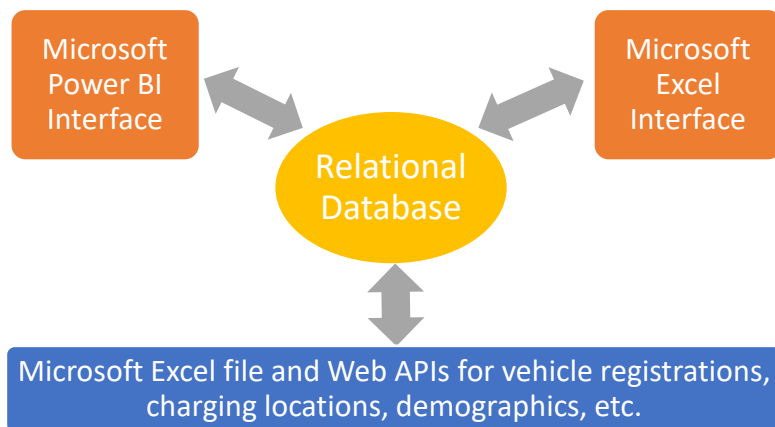
The relational database in EValuateNY is a comprehensive collection of policies, programs, demographics, and market data related to EV sales in New York State. Because the data are organized by location and time, users can explore factors specific to a city or county over time. For example, users can explore the growth of recent EV registrations near a charging station immediately after it was installed. The relationships that connect the data tables enable these kinds of analyses.

In addition, geolocation and time series data is hierarchical, meaning data can be automatically aggregated to gain broader perspectives and insights. For example, monthly data can be aggregated to quarters or calendar years to give users quick access to larger market trends. Similarly, data at the

ZIP Code level can be aggregated to the city, county, utility, or state level. Of course, the opposite is not true. For example, if the most basic data available is stored at the county level, it cannot be explored at the city or ZIP Code level.

Three kinds of relationships exist in EValuateNY:

- **Geolocation relationships** map to either ZIP Codes, cities, counties, or utilities, depending on the level of detail available. For example, vehicle registration data is available at the ZIP Code level.

- **Date relationships** are stored in a monthly or yearly format. For example, data is available to compare the annual growth of charging locations and vehicle registrations for any county.

- **Lookup tables** allow users to access information and improve the user experience by eliminating duplicated data. For example, a lookup table for vehicle make and model provides users with detailed descriptions of each plug-in electric vehicle available.

See appendix A for a diagram of the relationships and data tables contained in EValuateNY.

2.2 Data Contained in EValuateNY

This section summarizes the categories of data contained in the relational database, sources for these data, rationale for including them, and example insights that could be gleaned from exploring the data. For each data category, the source is either a publicly available Web API or a single Microsoft Excel source file as noted.

Electric Vehicle Registrations. A mix of monthly and annual vehicle registration data are available for battery electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs) by vehicle identification number (VIN). Including vehicle registration data enables tool users to follow EV ownership trends by vehicle make and model. Establishing sales trends by date and location allows users to compare how interventions may have affected EV sales at a particular date and location. For example, a lack of public charging access is often cited as a critical barrier to EV deployment. The deployment of new public charging stations may coincide with new EV sales in a particular ZIP Code. Annual data is available from 2010 to 2015 and periodic snapshots of vehicle registrations are available thereafter.[1]

Charging Station Locations. The tool includes the characteristics of nearly all charging locations, as provided by the U.S. Department of Energy Alternative Fuel Data Center. As mentioned previously, access to charging is a commonly cited barrier to EV ownership or use. Nearby charging stations may increase consumer confidence in the EV market, resulting in increased vehicle sales. Also, with this information, the ratio of charging stations to EVs in various locations could be explored to identify opportunities for additional infrastructure deployment.[2]

Charging Station Use. The tool includes the charging use data for the more than 200 stations funded by NYSERDA going back to the beginning of 2013. Monthly charging use is aggregated to the ZIP Code level and categorized by the charging network.[3]

Demographics. ZIP Code-level demographic data from the U.S. Census Bureau is included in the tool.[4] Descriptive statistics for New York State's population may provide valuable context for EV sales. Household income or commute time may influence EV sales just as significantly as dedicated EV promotional activities. The tool's demographic data offer insights into many different aspects of New York State's communities, such as population, the number and percentage of households that own their houses (and may therefore have an easier path to installing a home charging station), the number of vehicles per household, and more.[5]

DC Fast Charging and Level 2 Charging Station Usage. Monthly usage data is available from all public charging stations on the ChargePoint network through the second quarter of 2015 and many stations funded by NYSERDA. Public charging station usage may vary for a number of reasons, such as proximity to registered EVs or siting by regular travel routes. These data will show how drivers use the publicly available charging infrastructure that has already been installed. EV drivers' charging patterns may differ significantly between locations and by charging-type level (i.e., Level 2 or DC fast charging). The usage patterns may provide an understanding of the most effective methods of providing public EV charging in New York State, which would help charging service providers avoid stranding assets or making otherwise ineffective charging investments.[6]

Drive Clean Rebate. The Drive Clean Rebate is a vehicle incentive for passenger electric vehicles. The program has been in place since 2017. The rebate varies among vehicle technologies. The data used in EValuateNY includes the date, location, vehicle make and model, estimates on petroleum use reduction, and estimates on greenhouse gas savings.[7]

Geolocation and Time Resources. Atlas Public Policy created several tables for geolocation and time data for use by EValuateNY. These data include tables to match ZIP Codes to cities, counties, and utilities. Daily and monthly calendars for time series analysis were created automatically using facilities in Microsoft Excel and Power BI.[8]

2.3 Possible Future Enhancements for the Tool

EValuateNY demonstrates the power of relational databases to assess EV market development. Microsoft Excel and Power BI contain several features that could be incorporated into future versions of the tool to empower users to dive deeper into the data and gain greater insights. Possible future enhancements may include:

- Measures. Excel and Power BI contain a feature called "measures," which are formulas that combine and interpret the data to enable comparisons of data using simple or complex mathematical formulas. These measures can be plotted by county, city, or ZIP Code to gain insights into areas where publicly available charging infrastructure may be inadequate. Simple measures such as the ratio of EVs to charging ports or locations is available. More complex measures could be created to gain additional insights.

- Additional Data Sources. EValuateNY was designed to be expanded with data from additional sources that may enhance a future version of the application. For example, auto dealer promotions (e.g., vehicle discounts or ride-and-drives) could be cataloged by date and location and added in a similar way to the public policy data. Another potential data source is the location of key stakeholder groups, including EV advocacy organizations in New York State and a compilation of these groups' activities.

3   How to Use EValuateNY

EValuateNY is a flexible application that can offer insights into the many factors that affect New York State's EV market. This section provides a high-level overview of the Excel and Power BI interfaces for EValuateNY.

3.1   System Requirements

The Excel interface requires Microsoft Excel 2013 or later, available for purchase on PC, Mac, tablet, or mobile device.

The Power BI interface of EValuateNY can be accessed through a web browser on any device or downloaded as a desktop application (PC only) for free. Users can customize the EValuateNY reports from a web browser or desktop application. The desktop application is required to customize the relational database and use the full feature set of Power BI.

The source data for EValuateNY is accessed directly through web APIs and stored in Microsoft Excel files, also requiring Excel version 2013 or later. Because the EValuateNY uses standard versions of Excel and Power BI, users can leverage the available literature and communities for both applications to operate, maintain, and expand the tool (see Box 2).[9]

Box 2: Microsoft Excel and Power BI Community Resources

Microsoft Excel and the company's new data visualization application, Power BI, are accessible and popular programs for gaining insights from various sources of information. The following online resources are available to learn more about how EValuateNY uses the features of Excel and Power BI:

- PowerBI.com website: Microsoft's official site dedicated to supporting the Power BI application (www.powerbi.com)

- Power Query Help: Power Query is a computer language used by Microsoft Excel and Power BI to get and transform data from the Web APIs, external files, etc.

3.2   Choosing an Interface

The user interface for EValuateNY relies on features native to Microsoft Excel and Power BI. With Excel, users can engage with familiar pivot tables and charts to create custom visualizations (in addition to the tables and charts that come with the tool. The Excel interface provides users with a familiar way to interact and export the data contained in the tool. Through Power BI, users can create compelling, interactive visualizations. Users looking to engage with the data through a Web browser should select the Power BI interface.

The existing interfaces demonstrate the EValuateNY's utility and familiarize users with its functions. The Excel interface consists of worksheets containing tables and charts that are all responsive to user input. Users familiar with pivot tables and charts can quickly customize these worksheets to suit their needs. Though Power BI is a new application that may not be familiar to most users, the interface is intuitive. Many users may also choose to expand the existing functionality by making new connections through the relational database.

3.3   Excel Interface

EValuateNY's Excel interface consists of four worksheets designed to demonstrate its capabilities. The worksheet contents are defined in Table 1. Users may add new pivot tables or charts in new worksheets or by altering existing worksheets.

Table 1. Worksheets in the Excel Version of EValuateNY

Worksheet	Description
About	Introductory text about EValuateNY, including diagrams of the tool's architecture and relational database.
Pivot Table	Demonstration of pivot tables using measures for vehicle registration data.
Release Notes	EValuateNY version history.

3.3.1   Key Features

EValuateNY's Excel interface contains noteworthy features designed to enable the easy filtering and comparison of data. The screenshot in Figure 4 highlights key data analysis features in Excel: pivot tables and filters.

Figure 3. Demonstration of a Pivot Table with Filters

Feature 1. "Pivot tables" aggregate and compare data from multiple sources. The information in pivot tables can also be filtered through slicers and timelines. The pivot table highlighted by the dotted box labeled (1) is a summary of EV registrations by location.

Feature 2. "Slicers" allow the user to filter information on the page by individual variables, such as date or location. The dotted box labeled (2) highlights a slicer to filter data in the pivot table by county.

Feature 3. "Timelines" are a type of slicer that exclusively filter information by date. The dotted box labeled (3) highlights timeline filter data in the pivot table by date.

3.3.2 Exporting Work from Excel

Users can export pivot tables through the typical means familiar to Excel users, including the "clipboard" or taking a screenshot. Exported objects can be imported into existing Microsoft Office documents or saved separately as an image file.

3.4 Power BI Interface

EValuateNY's Power BI interface provides a fully immersive look at the data, including interactive visualizations, tables, and maps. Because Power BI is a newer program that typical users will be less familiar with than Microsoft Excel, users are encouraged to explore the wealth of online tutorials and other help from Power BI's website at www.powerbi.com.

Similar to the Excel interface, the Power BI interface consists of a number of pages that present different visualizations focused on vehicles, charging, demographics, and public policies. Because Power BI contains more types of visualizations than Excel, such as tree maps and interactive state maps, users may be able to derive more varied insights using Power BI.

1.1.1　Key Features

Users can click on any information in a Power BI report and instantly filter by that selection. For example, a user can drill down from information about all-electric utility service territories to explore data specific to one utility's territory as Figure 6 demonstrates. The left side of Figure 6 is a screenshot of the vehicles report page and the right side is the same page after the user has filtered information specific to Suffolk County. This feature allows users to quickly explore data specific to a location, date, or another category.

Figure 4. Interface for Vehicles Report Page Before and After a Filter is Applied
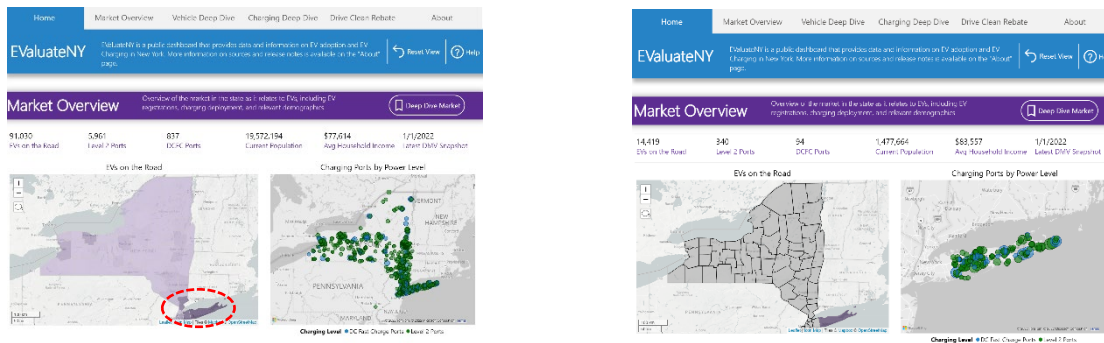
Power BI also includes filters at the visualization, page, and report level similar to the slicers and timelines found in the Excel interface. These filters, located on the far right of a report page, allow users to more precisely filter selections, such as focusing on individual counties or selecting information within a particular year. Users can filter visualizations in three ways: (1) by clicking on individual visualizations and selecting the visualization-level filters, (2) filtering at the page level, or (3) filtering at the report level (see Figure 7).

Figure 5. Filters for a Single Visual on the Landing Page
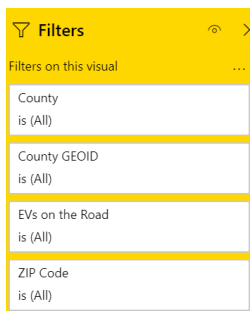
Exporting Work from Power BI

EValuateNY can be embedded on public websites allowing users to create and share interactive dashboards using the underlying data. The reports can also be shared with other users within an organization that is hosting the tool on www.powerbi.com. Currently, Power BI does not offer a

method to save visualizations as image files. Users interested in sharing visualizations will need to take screenshots. Users, however, can export raw data from visualizations. Power BI's feature set expands monthly, so a more convenient method to share visualizations could be implemented in the future.

The following are helpful links on how to export data and share Power BI visualizations:

- Publish Power BI reports to the web: https://powerbi.microsoft.com/en-us/publishtoweb

- Sharing a dashboard within an organization:
  https://docs.microsoft.com/en-us/power-bi/service-share-dashboards

- Export raw data from Power BI visualization:
  https://docs.microsoft.com/en-us/power-bi/visuals/power-bi-visualization-export

Appendix A: The Relational Database in EValuateNY

The diagram depicts the layout of the relational database in EValuateNY. Each box is a data table. Data tables can provide geolocation information in a time series format (e.g., monthly or yearly). In addition, to data tables, lookup tables can help avoid data duplication and provide rich information, such as a detailed description of each public policy. Finally, hierarchies exist for geolocation data and time series data allowing users to "drill down" from high-level to granular data.

Charging Locations
(ZIP Code, Open Date)

Charging Use
(ZIP Code, Monthly
Time Series)

Demographics
(ZIP Code, Yearly Time
Series)

ZIP Code Hierarchy
(Utility, County, City,
ZIP Code)

Calendar Hierarchy
(Year, Quarter, Month)

State Hierarchy
(Census Bureau, EIA,
ZEV, State)

National Highway
System
(County, Road)

Vehicles
(N/A)

Automakers
(N/A)

Charging Locations
(Ports)
(N/A)

EV Registrations (ZIP
Code, Monthly Time
Series)

Drive Clean Rebate (ZIP
Code, Monthly Time
Series)

- Market Data
- Actions, Policies, & Market Factors
- Geolocation & Date
- Lookup table

- Map to geolocation
- Map to date
- Map to lookup table

Appendix B: Data Tables and Fields

The following table summarizes each data table and field included in EValuateNY.

Table B-1. Data Tables and Fields

Data Table, Geolocation, Time Series	Data Field	Description
Charging Locations, ZIP Code, Open Date	Access	Public or private access.
	Charging Complex	Whether the charging site has more than one port for any fast charging connector.
	EV Network	Charging network that operates the charging station.
	Latitude	Latitude of charging station location.
	Longitude	Longitude of charging station location.
	Open Date	Date charging state was first opened.
	Station Name	Descriptive name of the charging site.
	Street Address	Street address of charging site.
Charging Locations (Ports) [Lookup Table]	Charging Level	Level 2 or DC fast charging connection at listed ports.
	EV Connector Type	Type of DC fast charging connector (CHAdeMO, SAE Combo, Tesla) or Level 2 (J1772) charging available at the charging location.
	Ports	Number of unique points that connect EVs to charging stations. Stations may have multiple ports. Valid for DC fast charging and Level 2.
Charging Use, ZIP Code, Monthly	Charging Time (hours)	Total charging time in hours.
	Energy (kWh)	Total electricity delivered to the EV in kilowatt-hours.
	Sessions	Total number of charging sessions.
	Statewide Active Network Port Count	Number of unique points that connect EVs to charging stations statewide by charging network.
	Total Duration (Hours)	Total charging session (includes time connected and not charging) length in hours.

Table B-1. continued

Data Table, Geolocation, Time Series	Data Field	Description
Demographics, ZIP Code, Monthly	Educational Attainment (Associate's Degree)	Total number of people with an associate's degree.
	Educational Attainment (Bachelor's degree)	Total number of people with a bachelor's degree.
	Educational Attainment (Doctorate degree)	Total number of people with a doctorate degree.
	Educational Attainment (High School or Less)	Total number of people with high school or less.
	Educational Attainment (Master's degree)	Total number of people with a master's degree.
	Educational Attainment (Professional school degree)	Total number of people with a professional school degree.
	Educational Attainment (Some college)	Total number of people with some college.
	Educational Attainment (Total)	Total number of people.
	Housing Units (1, Detached)	Number of single-family detached housing units.
	Housing Units (10 to 19)	Number of housing units with 10 to 19 households.
	Housing Units (2)	Number of housing units with two households.
	Housing Units (20 to 49)	Number of housing units with 20 to 49 households.
	Housing Units (3 or 4)	Number of housing units with three to four households.

Table B-1. continued

Data Table, Geolocation, Time Series	Data Field	Description
	Housing Units (5 to 9)	Number of housing units with five to nine households.
	Housing Units (50 or More)	Number of housing units with 50 or more households.
	Housing Units (Boat, RV, Van, etc.)	Number of housing units that are boat, RV, van, etc.
	Housing Units (Mobile Home)	Number of housing units that are mobile homes.
Demographics, ZIP Code, Monthly	Median Household Income in Past 12 Months	Average household income.
	Occupied Housing Units	Total occupied housing units.
	Occupied Housing Units (Owner Occupied)	Number of housing units that are owner occupied.
	Occupied Housing Units (Renter Occupied)	Number of housing units that are renter occupied.
	Population	Total population
	Travel Time to Work (10 to 14 minutes)	Number of people whose commute is between 10 and 14 minutes.
	Travel Time to Work (15 to 19 minutes)	Number of people whose commute is between 15 and 19 minutes
	Travel Time to Work (20 to 24 minutes)	Number of people whose commute is between 20 and 24 minutes.
	Travel Time to Work (25 to 29 minutes)	Number of people whose commute is between 25 and 29 minutes.
	Travel Time to Work (30 to 34 minutes)	Number of people whose commute is between 30 and 34 minutes.

Table B-1. continued

Data Table, Geolocation, Time Series	Data Field	Description
	Travel Time to Work (35 to 44 minutes)	Number of people whose commute is between 35 and 44 minutes.
	Travel Time to Work (45 to 59 minutes)	Number of people whose commute is between 45 and 59 minutes.
	Travel Time to Work (60 or more minutes)	Number of people whose commute is 60 or more minutes.
	Travel Time to Work (Less than 10 minutes)	Number of people whose commute is less than 10 minutes.
	Travel Time to Work (Total Workers)	Total workers
Demographics, ZIP Code, Monthly	Workers 16 Years and Over	Total workers 16 years and over.
	Workers 16 Years and Over (Bicycle)	Total workers 16 years and over who commute by bicycle.
	Workers 16 Years and Over (Carpooled)	Total workers 16 years and over who commute by carpool.
	Workers 16 Years and Over (Drove Alone)	Total workers 16 years and over who commute by driving alone.
	Workers 16 Years and Over (Public Transportation (Excluding Taxicab))	Total workers 16 years and over who commute by public transportation.
	Workers 16 Years and Over (Taxicab, Motorcycle, or Other Means)	Total workers 16 years and over who commute by taxicab, motorcycle, or other means.
	Workers 16 Years and Over (Walked)	Total workers 16 years and over who commute by walking.

Table B-1 continued

Data Table, Geolocation, Time Series	Data Field	Description
DMV Snapshots [Lookup Table]	DMV ID	Unique identifier for a snapshot of the vehicle registration database from the New York State DMV.
	DMV Snapshot (Last Registration Date)	Description of DMV snapshot, including the last date of a vehicle registration.
	Link	Source for DMV snapshot.
Drive Clean Rebate, ZIP Code, Monthly	Annual GHG Emissions (MT CO2e)	Estimated annual savings in greenhouse gas emission from EVs.
	Annual Petroleum Reduction (Gallons)	Estimated annual reduction in petroleum use from EVs.
	Rebate ($)	Rebate amount in dollars.
	Transaction Type	Own or Lease
Employment, ZIP Code, Yearly	Employment Size of Establishment	Categorical description on the number of employees.
	Title of NAICS Industry Code	Description of business type.
	Total Establishments	Number of businesses.
EV Registrations, ZIP Code, Monthly	DMV Direct	Flag where 1 means the data is from the DMV directly and 0 means data was provided by an indirect source.
	DMV ID	Identifier from when DMV data from http://data.ny.gov was integrated (1=1/05/2016, 2=3/11/2016, and 3=7/15/2016). Results are then sorted from highest to lowest to enable an accurate calculation of the current EVs on the road.
	First Characters	First eight digits of the VIN.
	VIN	Vehicle identification number.
Location	Area Land Sq Miles	Area of ZIP Code in square miles.
	CBSA Title	Core-based statistical area.
	Census Bureau Division	Division from the U.S. Census Bureau.
	Classification	Urban-rural classification from the U.S. Census Bureau.
	County	County

Table B-1 continued

Data Table, Geolocation, Time Series	Data Field	Description
Location (continued)	County GEOID	Unique identifier for county.
	CSA Title	Combined Statistical Area from the U.S. Census Bureau.
	EIA	U.S. Energy Information Administration region.
	EIA Utility ID	Unique identifier for electric utility from the U.S. EIA.
	EIA Utility Name	Utility name according to the U.S. EIA.
	Latitude	Latitude
	Longitude	Longitude
	Primary City	The primary city for this ZIP Code.
	Primary eGRID Subregion	Primary eGRID region from the U.S. Environmental Protection Agency.
	Secondary eGRID Subregion	Secondary eGRID region from the U.S. Environmental Protection Agency.
	State	U.S. state.
	State Abbreviation	U.S. state abbreviation.
	Tertiary eGRID Subregion	Primary eGRID region from the U.S. Environmental Protection Agency.
	Type	Category for ZIP Code.
	Utility	Familiar electric utility name for investor-owned utilities.
	ZEV State	Zero-emission vehicle state.
	ZIP Code	Zip Code.
Location (County-Level) [Lookup Table]	County	County.
	County GEOID	Unique identifier for county.
	State	U.S. state.
National Highway System [Lookup Table]	Annual Average Daily Traffic	Average daily traffic for road.
	Functional Class	Code for the road classification.
	ID	Unique identifier for road.
	Local Street Name	Local street name.
	Road Sign	Road sign.
	Road Type	Type of road.
	Status	Status of road.

Table B-1 continued

Data Table, Geolocation, Time Series	Data Field	Description
National OEM Sales, National, Monthly	Sales	Monthly light-duty vehicle sales at the national level.
	Vehicle Manufacturer	Automaker
OEMs [Lookup Table]	Familiar	Familiar name automaker.
	Parent Company	Parent company of the automaker.
	Vehicle Manufacturer	Automaker
Technology [Lookup Table]	Definition	Description of vehicle technology.
	Technology	Type of vehicle drivetrain.
Utility [Lookup Table]	EIA Utility ID	Unique identifier from U.S. EIA.
	EIA Utility Name	Name of utility according to the U.S. EIA.
	Familiar	Familiar name for investor-owned utility.
	Intermediate Holding Company	Intermediate holding company for investor-owned utility.
	Operating Company	Name of operating company for investor-owned utility.
	Ownership Type	Type of ownership for utility.
	Parent Company	Parent company for investor-owned utility.
Vehicle Identification Number [Lookup Table]	Make	Vehicle make.
	Model	Vehicle model.
	Vehicle Manufacturer	Vehicle manufacturer.
	Vehicle Name	Vehicle make and model.
	VIN(1-8)	First eight characters of VIN.
Vehicles [Lookup Table]	Dedicated Platform	Whether vehicle is built on a dedicated platform.
	Make	Vehicle make.
	Model	Vehicle model.
	Technology	Vehicle drivetrain technology.
	Vehicle Manufacturer	Vehicle manufacturer.
	Vehicle Name	Vehicle make and model.

Table B-1 continued

Data Table, Geolocation, Time Series	Data Field	Description
	VIN Alternate Name	Alternate name that could be used for VIN matching.
	VIN Vehicle Name	Primary name that is used for VIN matching.
	Use Case	Vehicle use case based on the body class of the vehicle.
	Vehicle Class	The class of the vehicle (1-8) based on the gross vehicle weight rating (GVWR).
	Vehicle Category	The category of the vehicle based on the class (Light Duty, Medium Duty, Heavy Duty).

Appendix C: Methodology for Using DMV Data

One of the most valuable data sets in the transportation energy field is the vehicle registration database, typically maintained by a state department of motor vehicles. In 2016, New York State began publishing the complete vehicle registration database for use by the public on data.ny.gov

The methodology in this section describes how to convert New York State's vehicle registration database into small, usable Microsoft Excel tables containing vehicle registrations for plug-in electric vehicles by ZIP Code and registration date for inclusion in EValuateNY.

Below are terms used in this appendix:

- **VIN**: The Vehicle Identification Number has a unique code containing the vehicle make, model, and other characteristics.

- **Power Query**: A component of Microsoft Excel that allows for the importing and manipulating of large external data sets (i.e., "big data"). Power Query is a computer scripting language used to organize data sets into a format that is usable with other features of Excel, including pivot tables and charts.

- **Data Model**: A component of Microsoft Excel that manages data sets larger than what can typically be managed in Excel worksheets (e.g., tables with more than 1 million rows). The Data Model is a relational database, which links different data sets using unique "keys" in each table. These keys allow users to conduct analyses on multiple data sets as if the data were stored in one large data set.

- **Macro**: A component of Excel that allows for task automation using the Visual Basic computer language.

- **Power BI**: A new, free Microsoft program that allows users to import and manipulate data sets using the Power Query language and Data Model concept. Power BI can make sophisticated visualizations and conduct data analyses of a relational database.

- **SQL Database:** A relational database that allows for structured storage of a large amount of data. A SQL database is used for storing raw registration data received from the state for the entire historical state vehicle population.

The DMV data set is updated frequently, and since January 2016, includes Vehicle Identification Number (VIN), the ZIP Code of the vehicle registration, and registration date. With these three pieces of data, analysts can effectively track the make and model of vehicles by location over time.

Automakers can use one or more VINs for a single make and model, making tracking and decoding VINs into human readable information (i.e., vehicle make and model) challenging. For this effort, NYSERDA provided a table containing the full list of known VINs for plug-in electric vehicles for model years 2010 through 2016. The table included the first 8 digits of the VIN and the corresponding make, model, and vehicle technology (all-electric or plug-in hybrid technology).

NYSERDA provided registration data for plug-in electric vehicles before 2016. New York DMV provided annual snapshots of its vehicle registration database to the New York State Department of Environmental Conservation (NYSDEC) from 2011 through 2015 for program evaluation purposes. The data was provided in a "coded" format and included the VIN number and ZIP Code for each registration; the registration date was not provided. The DEC shared these data sets with NYSERDA in support of this effort; each file was between 600 megabytes and 1.2 gigabytes making it challenging to use the data directly.[10]

Box 3. DMV Registration Decoder Ring

The vehicle registration data provided to the NYSDEC needed to be decoded to identify the VIN and ZIP Code. Below is the format for raw data:

- Model Year = 4 spaces

- VIN = 17 spaces

- Make = 5 spaces

- Fuel = 8 spaces

- Weight = 6 spaces

- County Code = 2 spaces

- ZIP Code = 5 spaces

- License Plate = 8 spaces

- Registration Class = 2 or more spaces (depending on year)

Assembling DMV Historical Data (2011–2015)

Below is the process followed to convert the DMV historical data (2011–2015) into small, manageable Excel files containing plug-in electric vehicle registrations by ZIP Code and year. These files were then merged into a single Excel table for use in EValuateNY. The file created in this procedure is included with EValuateNY and is called *ny_ev_registrations.xslxm*.

Note, this process does not need to be repeated as it was only necessary for including historical data.

1. Create a new Excel file.
2. Copy the VIN decoder table into the Excel file and add it to the file's Data Model and a new Power Query.
3. Create a plug-in electric vehicle registration table for each year through Power Queries called *NY_201x_EVs* and *NY_201x* for each year from 2011 to 2015.
   a. Import yearly registration file into a new Power Query.
   b. Organize the data into a readable table using Power Query commands (e.g., *Table.SplitColumn* and *Table.TransformColumnTypes*) and the decoder ring.

    c.   Add date column (i.e., January 1, 20XX) to identify the approximate date when a vehicle was registered. The resulting table could be used for analyses dependent on the entire vehicle registration database.

    d.   To improve performance and keep the data sets manageable, create a new Power Query and reference the table from step 3c. Merge this table with the VIN decoder table to only retain plug-in electric vehicle registrations.

    e.   Add the table to the Data Model.

    f.   Export each table containing the annual plug-in electric vehicle registrations to a CSV file using a custom Excel Macro.[11]

4. Create a single plug-in electric vehicle registrations table through Power Query called *NY Historical EV Registrations*.

    a.   Merge the annual EV registration files into a single table using a new Power Query that references each CSV file from step 3.

    b.   Add column containing first 8 characters of VIN.

    c.   Add the table to the Excel Data Model of the same file from step 1.

Retrieving DMV Data from data.ny.gov (2016–present)

Below is the procedure to download the DMV database snapshots from data.ny.gov. Steps 1 through 6 of this procedure must be completed each time the New York DMV posts a new database. The procedure described below is included with EValuateNY in a file called *ny_ev_registrations.csv*.

1. Download the Vehicle, Snowmobile, and Boat Registrations database, located at https://data.ny.gov/Transportation/Vehicle-Snowmobile-and-Boat-Registrations/w4pv-hbkt as a CSV file.

2. Load the CSV file into a table in a SQL database that stores all raw snapshots, filtering for only vehicles (VEH).

3. Decode VINs by joining the VIN prefix (first 8 characters) and the model year (10th digit in the VIN) to the same values in the VIN decoder table to extract electric vehicles.

4. Export electric vehicles for all snapshots as a CSV and Excel file with the name ny_ev_registrations.

EValuateNY merges the historical and current EV registrations to create a single database from 2011 to the present. EValuateNY also marks a registration as "Original" on the first appearance in the database and marks a registration as "Renewal" for subsequent registrations.

Appendix D: Methodology to Map ZIP Codes to Electric Utility Territories

ZIP Codes are the base geography for EValuateNY. Electric utility territories in the United States are not neatly divided by ZIP Code and more than one electric utility can operate in the same ZIP Code. EValuateNY uses a three-step process to make a best approximation of the primary electric utility in each ZIP Code.

For investor-owned utilities, each utility shares data with Atlas Public Policy that defines the share of customers for each ZIP Code in which the utility operates. Atlas then uses that data with the methodology below to distribute EVs by ZIP Code:

1. Total customer counts (residential and commercial) by ZIP Code are collected from each investor-owned utility (IOU) in the State.
2. A table is generated that lists for each ZIP Code the IOU and the proportion of customers that utility has in the ZIP Code [Utility Share], where the sum of [Utility Share] for each ZIP Code equals one. This table resides in the *resources.xlsx* file.
3. For ZIP Codes in the State where an IOU does not have any customers, the utility territory is determined using Atlas's previous method described here: https://www.atlasevhub.com/help-center/geography-on-the-ev-hub.
4. For publicly owned, cooperatives, and other utilities that have not provided data on customer counts and share a ZIP Code territory with one of the IOUs, all EVs in that ZIP Code will be *only* assigned to the IOUs based on the steps above. This will result in a decrease of EVs in non-IOU territories.
5. The generated table with the fields [ZIP], [Utility], and [Utility Share] are brought into the EValuateNY model. EVs on the road are then distributed for each ZIP Code to each utility based on the utility share value. When the number of EVs in a ZIP Code for any utility is not a whole number, fractions will be truncated, and one EV will be added to the utility with the highest share of EVs.

For publicly owned, cooperatives, and other utilities that have not provided data on customer counts and do not share a ZIP Code territory with one of the IOUs, we use data from the U.S. Energy Information Administration (EIA-860 and EIA-861), which provides a listing of all electric utilities and counties that they operate in. We also use the utility-ZIP mapping from OpenEI. For each ZIP Code, we first choose the utility defined on OpenEI. If it's not there, we fall back on the utility-county lookup data from the U.S. EIA. If more than one utility operates in a ZIP Code from OpenEI or a county from the EIA surveys, then we'll choose the largest utility as measured in megawatt hours served, according to EIA data. The data is updated annually and the mapping of utility to ZIP Code can be reassigned based on the data from OpenEI and EIA.

Appendix E: Methodology for Tracking EVs and Charging Stations in Disadvantaged Communities

EValuateNY includes a section on disadvantaged communities (DACs) that highlight EV and charging deployment in areas the State has identified as a DAC. The criteria for identifying an area as a disadvantaged community are defined by the State, and users can learn more at: https://www.nyserda.ny.gov/ny/disadvantaged-communities

DACs are identified at the level of Census Block Group, and each metric tracked by DAC has a unique methodology.

Population

Population counts for each DAC are determined based on the Census Block Group, and EValuateNY uses the American Communities Survey (ACS) to provide an estimate of population in a given DAC.

Charging Locations

The counts of charging locations and charging ports are determined by geocoding each charging station based on the latitude and longitude of the station in order to determine the Census Block Group in which it resides. The counts are then mapped back to the DAC Census Block Groups to determine the count of stations or ports in a given DAC or in DACs statewide.

Electric Vehicles on the Road

Registration data in the State is provided at the ZIP Code level, so the count of EVs on the road that fall in a DAC must be estimated. To accomplish this, Atlas uses the methodology defined in appendix F to estimate EVs that are registered with a DAC.

Appendix F: ZIP Code to Census Block Group: Methodology and Applications

Overview

Atlas has created a methodology for distributing Census Block Group populations to ZIP Code populations using information from the U.S. Census Bureau and ProximityOne (http://proximityone.com). This mapping can be used in applications where spatial analysis requires Census Block level information. For EValuateNY, this methodology was used to identify disadvantaged community population in a ZIP Code.

Methodology

1. Atlas downloaded population data by Census Block and ZIP Code tabulation area from the U.S. Census Bureau.
2. Atlas used a publicly available resource from ProximityOne (http://proximityone.com/bg-zip.htm) that distributes the population of each Census Block Group to the ZIP Code in which the population resides. For most block groups, this will be a 1:1 map as the Census Block Group is fully contained within the ZIP code. For Census Block Groups where the boundary is shared with multiple ZIP Codes, the population is split between the ZIP Codes based on a methodology developed by ProximityOne, where population is determined at the street-address level and then mapped to both Census Block Group and ZIP Code.
3. Atlas then uses the ZIP Code-Block Group (ZIP-BG) population to determine the share of people or nominal population that falls within specific Census Block Groups for each ZIP Code.

Example Application: Disadvantaged Communities in New York State

1. To identify the proportion of each ZIP Code that is considered a disadvantaged community (DAC) in New York State, each Census Block Group is flagged as either a DAC or a non-DAC based on the Interim Disadvantaged Communities (DAC) data set from the New York State Energy Research and Development Authority (NYSERDA) (https://data.ny.gov/Energy-Environment/Interim-Disadvantaged-Communities-DAC-2020/t6wd-tdrv).
2. For each ZIP Code, Atlas applies a function to determine the share of the population in the ZIP Code that lie in a DAC:
   a. For each ZIP Code, a list of Census Block Groups and the population of the Census Block Group that fall in the ZIP Code [ZIP Block Population] is generated.

b.   The population is summed for Census Block groups flagged as DAC [DAC Population], and non-DAC Census Block groups [Non-DAC Population].

$$DAC\ Population = \sum_{ZIP} ZIP\ Block\ Population_{Census\ Block\ Group}\ \{DAC = True\}$$

$$Non\ DAC\ Population = \sum_{ZIP} ZIP\ Block\ Population_{Census\ Block\ Group}\ \{DAC = False\}$$

c.   The share of the ZIP Code which is in a DAC is calculated as the DAC Population divided by the total Population in the ZIP Code.

$$DAC\ (\%)_{ZIP} = \frac{DAC\ Population}{DAC\ Population + Non\ DAC\ Population}$$

3.   The DAC (%) field can then be used to determine metrics only available at the ZIP Code level, such as EVs on the Road.

Appendix G: Files Included in EValuateNY

Filename	Description
census_bureau.xlsx	File containing demographic information by ZIP Code from the U.S. Census Bureaus' American Community Survey.
Decoded_Prefixes.csv	File containing decoded information for unique VIN prefixes in the registration data. Vehicles are decoded using data from NHTSA.
EValuateNY.pbix	Microsoft Power BI version of the complete tool.
EValuateNY.xlsx	Microsoft Excel version of the complete tool.
ny_ev_registrations.csv	File used to retrieve historical and current EV registration data.
NY Original Registrations.csv	File containing all historical original registrations. This is a table of unique VINs where one row corresponds to a single VIN at its earlier registration date.
resources.xslx	Data tables used in EValuateNY that are not retrieved from Web APIs except EV registrations, including, but not limited to, charging use and VIN decoder.
Vehicle_Share.csv	File containing the share of historic light-duty (GVWR < 8500) vehicles by snapshot and ZIP Code. This is used to calculate the market share of EVs for all light-duty vehicles.

Endnotes

[1] For vehicle registrations, a special methodology was necessary in order to blend annual data before 2015 that was not available publicly. See appendix A.

[2] Data source: Web API from U.S. Department of Energy available at http://www.afdc.energy.gov/locator/stations

[3] Data source: ChargePoint, GE, and EV Connect

[4] The U.S. Census Bureau uses ZIP Code approximations called ZIP Code tabulation areas. More information on the methodology and any limitations of this approach is available at http://www.census.gov/geo/reference/zctas.html

[5] Data source: Web API from the U.S. Census Bureau available at http://www.census.gov/data/developers/data-sets/acs-5year.html

[6] Data source: Excel file using data from ChargePoint and NYSERDA

[7] Data source: https://data.ny.gov/Energy-Environment/NYSERDA-Electric-Vehicle-Drive-Clean-Rebate-Data-B/thd2-fu8y

[8] Data source: Excel file assembled by Atlas Public Policy analysis of data from http://www.eia.gov/electricity/data/eia861 and http://en.openei.org/datasets/dataset/u-s-electric-utility-companies-and-rates-look-up-by-zipcode-2014

[9] Visit https://www.nyserda.ny.gov/All-Programs/ChargeNY/Support-Electric/Data-on-Electric-Vehicles-and-Charging-Stations to download the Excel and Power BI files.

[10] These large files can be shared upon request.

[11] Exporting CSV files for each year was necessary for performance reasons. When trying to merge the annual files, Excel would try to import the data from each original external file, resulting in considerable delay.