

NYC Parking Violation Data Set

by Angela Gunn, Louie Viriato, and Chula Watugala

Welcome to the age where governments – federal, state and municipal – are pressured to increase transparency by creating open data portals to the public. Government collected data is often about individuals and even with efforts made to anonymize the data, we show in this paper it is possible to learn more about the individuals captured in the data.

The data we are working with comes from the New York City OpenData. The specific data set we are looking at is the [“Parking Violations Issued – Fiscal Year 2014 \(August 2013 – June 2014\).”](#) This data set shows many fields related to the parking

Sample of Data Fields in NYC Parking Violations Issued	
Registration	Plate Number & State Plate Type (commercial, passenger, etc.) Expiration Date
Vehicle	Body Type Vehicle Make, Year & Color
Violation	Address (of Violation) Date & Time of Violation Issuer Precinct (Issuer and Violation)

violation, from vehicle type to the address where the ticket is cited. The included table on this page shows some of the more interesting pieces of information. The name of the driver or registered owner of the vehicle is not listed.

Many of the license plates are listed for multiple violations. Often the tickets issued to one specific plate are around the same proximity, but others appear to be at very different addresses. One concern we have regarding this is if we can determine who the registered owner of the vehicle is, we will have information on where that person (or at least the person’s

vehicle) has been based on the location and date of the parking tickets. This can lead to speculations about the person's routines or one-off errands.

Driver's Privacy Protection Act

Having the vehicle license plate was first seen as something simple to trace back to the owner. We were thwarted by the Driver's Privacy Protection Act (DPPA). The DPPA prohibits the release or use of personal information about an individual obtained in connection with a motor vehicle record. In other words, unless we have a valid reason for requiring the personal information linked to the license plate, we cannot look up the plate in an easily accessible DMV database.

If you are motivated, there are always ways to find the information you are looking for. There are many websites that will provide this information for a fee. One such website (GovDMVRegistry.com) states they gather their information from signing contracts with 3rd party data providers and thousands of proprietary sources. However, you can also find many warnings to be suspicious of these websites as often the data is outdated. Due to the DPPA, it is not reasonable to believe they are getting their information from any DMV.

We found a website [That's Them](http://That'sThem.com) that returns an abundance of information ranging from name(s), contact information, education and occupation and even financial details. One of the search terms that can be used to search for people on *That's Them* is a VIN. Obtaining the VIN for the license plates we had was a matter of entering them into the website [AutoCheck](http://AutoCheck.com). *AutoCheck* has a limit on the number of free searches you can do on it, but it was enough for us to work with.

Possibilities are endless from there

To prove our method for identifying the individual associated with a parking ticket, we picked a license plate from the data that had a high ticket count (six in four months). We entered the license plate into *AutoCheck*, and then entered the VIN into *That's Them*. The information retrieved from there led to additional searches in other websites to gain a full picture of our individual. We also learned that all of the parking tickets for this individual are in close proximity to the individual's office, so no really interesting information there.

Individual's Information Found		Sources Used:
License Plate GBD2181 NY		
Car	2013 Audi A4 Quattro Premium VIN Style / Body: Sedan 4D Engine: 2.0L I4 DI Country of Assembly: Germany	http://www.autocheck.com/ http://thatsthem.com http://www.peoplefinders.com http://www.whitepages.com https://www.facebook.com
Personal	Name Age / DOB Home Phone Cell Phone Address	
Family	Wife Name (Married Date) Daughter	
Other	Work School (University, High School) Home Town	

Another license plate we searched resulted in an obituary.

Reflection and Summary

Our method in identifying individuals in *the NYC Parking Violations Issued* dataset has many issues. First, the ticket dates range in years from 1970 to 2060. This is either an error in the dataset or an intentional masking of the information. Comments made on the data set suggest NYC OpenData believes it is an issue with the third-party agency's quality control. Second, the registered owner of the vehicle may not be the actual person who parked the vehicle illegally. Therefore, any conclusions we make on the location of the parking tickets

issued may not apply to the individual linked to the vehicle registration. Third, we have reason to believe the information on *That's Them* is not accurate. The individual we explored had information we could use to verify we likely had the correct individual. However when we tried using our own names as the search criteria in *That's Them*, one of us wasn't found, one had many results (though none of them the actual person) and the last person was listed but with inaccurate information such as an old phone number from many years ago.

The fact we could find so much information on one individual shows how much information is out there. Some of the sources we used, such as *Facebook*, can be set to maintain the user's privacy. However other sites such as *That's Them* will only suppress user information if that user completes and submits a detailed form, which includes name, phone number and address as required fields. In other words, unless you know about the website and unless you provide them your personal information, they will continue to display information (accurate or not) to anyone who queries for it.

The *NYC Parking Violations Issued* dataset itself is not providing explicit details about the individuals involved directly or indirectly with the parking violations. Paired with other data sets, a nameless record becomes a vehicle, a name, a person, a family. We know a large amount about our individual, who was selected only because of the volume of parking tickets in a short time period.