**Project Title**

Analysis of parking violations in New York City

**Background**

New York City issued parking tickets information online for fiscal year of 2017 providing information of parking violations in the city, i.e., 43 variables including summon number, vehicle type, color, registration state, street name, violation time and data contains 10.8 million lines of records. This dataset is valuable for further analysis for the most frequent violation state vehicles, most violation areas and streets so that related parties may take appropriate measures to reduce violation in the future.

**Questions to ask about the parking tickets dataset:**

1. For all tickets, what are the states with most violations where registration plates issued? Which state seems odd compared to other states?

2. which area/streets where most violations happened?

3. Inferential statistics analysis: the relationship between the number of violations and the distance from New York City to state where vehicle was registered.

**Modeling:** linear regression

**Programing**: Python

**Dataset**: <https://data.cityofnewyork.us/City-Government/Parking-Violations-Issued-Fiscal-Year-2017/2bnn-yakx>

**Expected Business value**

We discovered New Jersey is the second most violation state, which suggest in order for reduce of burden of NYPD, for most of highway toll roads that run into New York City, it is wise to impose higher fee.

For commuters to New York city, the streets in Manhattan borough might be cautious to be parked or might be avoided if possible since there are most parking tickets issued on these streets.