

NY Taxi Review



Ben Kuehner, Tim Papich, Steve Putt

The Dataset

NYC Taxi and Limousine Commission:

<https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>

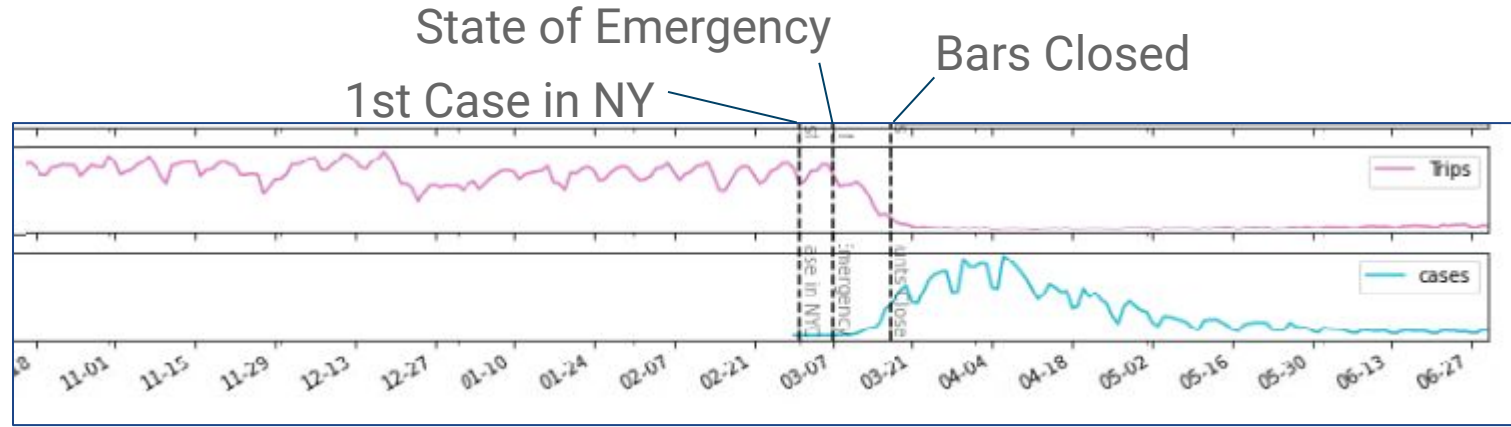
	VendorID	tpep_pickup_datetime	tpep_dropoff_datetime	passenger_count	trip_distance	RatecodeID	store_and_fwd_flag	PULocationID	DOLocationID	payment_type
0	1.0	2020-01-01 00:28:15	2020-01-01 00:33:03	1.0	1.20	1.0	N	238	239	
1	1.0	2020-01-01 00:35:39	2020-01-01 00:43:04	1.0	1.20	1.0	N	239	238	
2	1.0	2020-01-01 00:47:41	2020-01-01 00:53:52	1.0	0.60	1.0	N	238	238	
3	1.0	2020-01-01 00:55:23	2020-01-01 01:00:14	1.0	0.80	1.0	N	238	151	
4	2.0	2020-01-01 00:01:58	2020-01-01 00:04:16	1.0	0.00	1.0	N	193	193	
...
6405003	NaN	2020-01-31 22:51:00	2020-01-31 23:22:00	NaN	3.24	NaN	NaN	237	234	
6405004	NaN	2020-01-31 22:10:00	2020-01-31 23:26:00	NaN	22.13	NaN	NaN	259	45	
6405005	NaN	2020-01-31 22:50:07	2020-01-31 23:17:57	NaN	10.51	NaN	NaN	137	169	
6405006	NaN	2020-01-31 22:25:53	2020-01-31 22:48:32	NaN	5.49	NaN	NaN	50	42	
6405007	NaN	2020-01-31 22:44:00	2020-01-31 23:06:00	NaN	11.60	NaN	NaN	179	205	

6405008 rows × 18 columns

Taxi Data Stats: Jan 2019 - June 2020

Total Rides:	101,246,797	
Sample Size:	1,011,711	Random Sampling
After Cleaning:	997,966	Remove Outliers and Errors
Daily Averages:	547	Aggregation

The Timeline of Events



January
21, 2020

First confirmed COVID-19 case in the U.S.

March 1,
2020

First COVID-19 case in New York State

March 7,
2020

NY Governor Andrew Cuomo declares a state of emergency

March
14, 2020

First two COVID-19 deaths in NYS

March
16, 2020

NYC public schools close

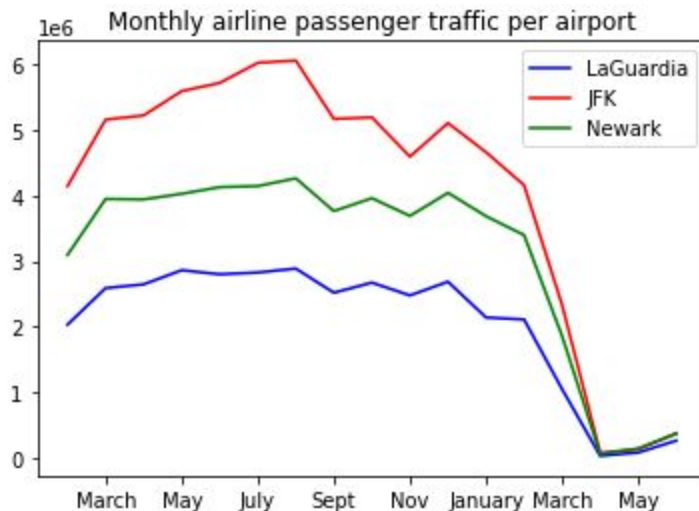
March
17, 2020

NYC bars and restaurants close, except for delivery

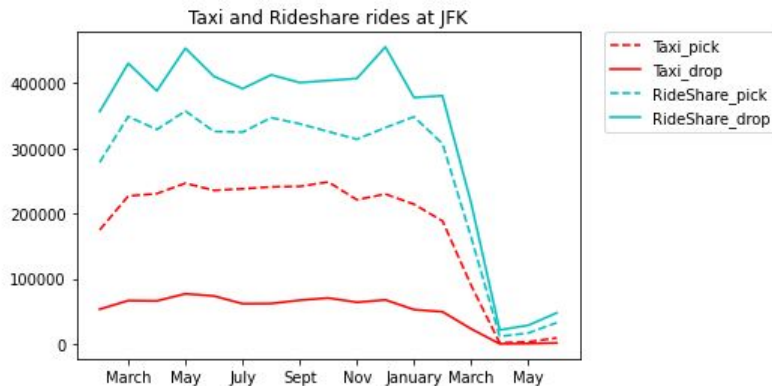
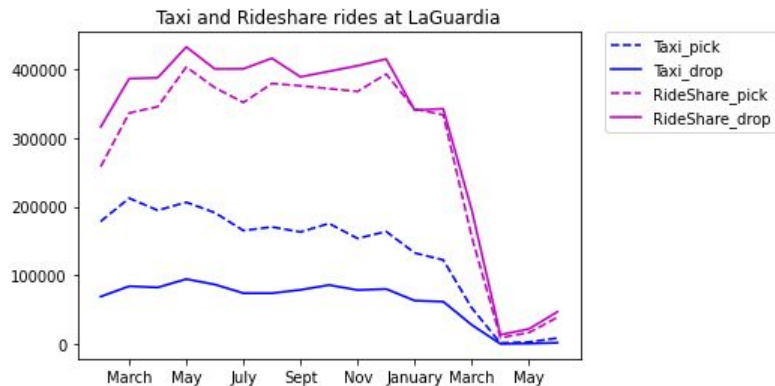
Questions to Answer

- Taxis vs. Ride-sharing Services
- Public Behavior in Response to Legislation
- Measuring Tipping and Charitableness
- Geographical Comparisons
- Effects on Airline Industry?
- Economic Impact

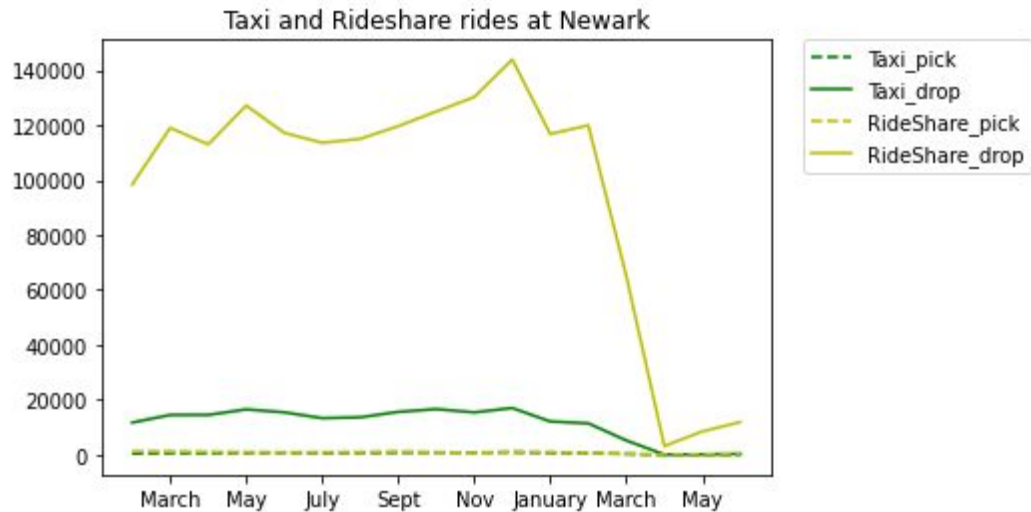
Monthly Airport Passenger Volume



Taxi and Rideshare Volume



Taxi and Rideshare Volume



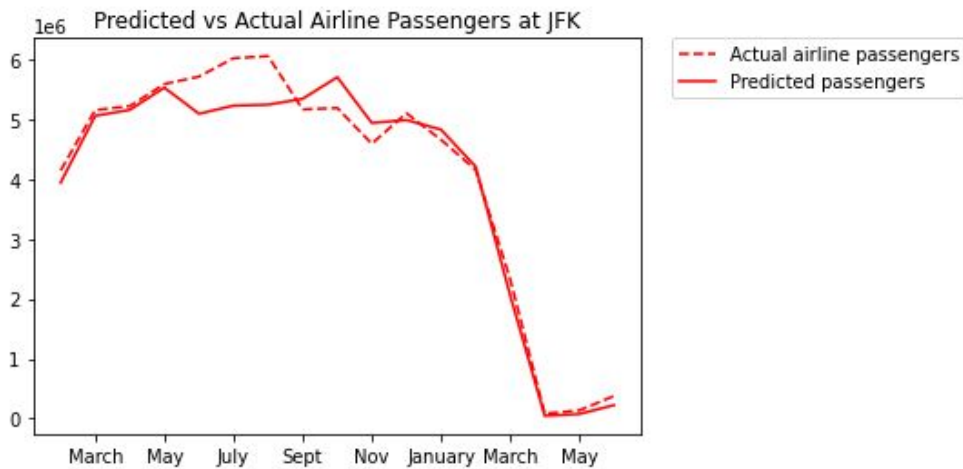
Linear Regression Models

- LaGuardia: Air = $250.12 \times \text{Taxi_Pick} + 479.49 \times \text{Ride_Pick} + 5085.62$
- $R^2 = 0.983$, P-values: Taxi_Pick = 0.003, Ride_Pick = 0.000



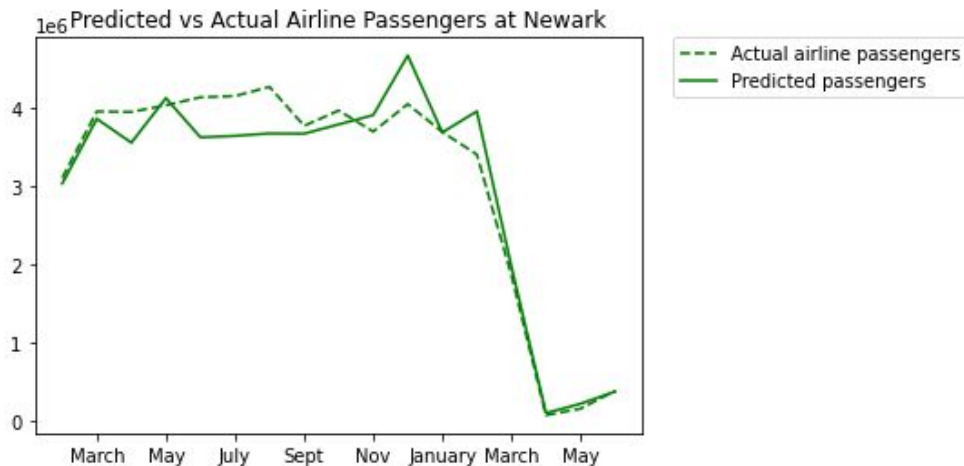
Linear Regression Models

- JFK: Air = $1108.79 \times \text{Taxi_Pick} + 118,700$
- $R^2 = 0.97$, P-value: Taxi_Pick = 0.000



Linear Regression Models

- Newark: Air = $3111.38 \times \text{Ride_Drop} + 46,600$
- $R^2 = 0.949$, P-value: Ride_Drop = 0.000

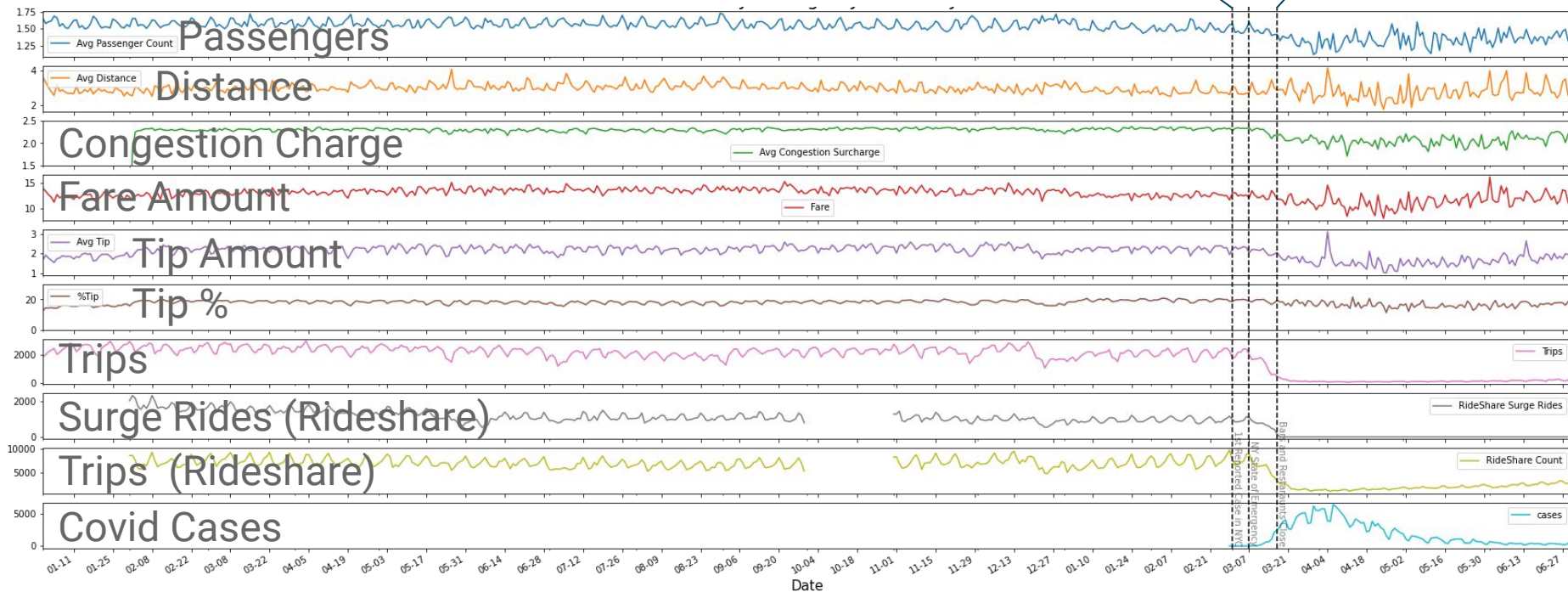


NY Taxi Patterns Over Time

Daily Averages: Jan '19 through Jun '20

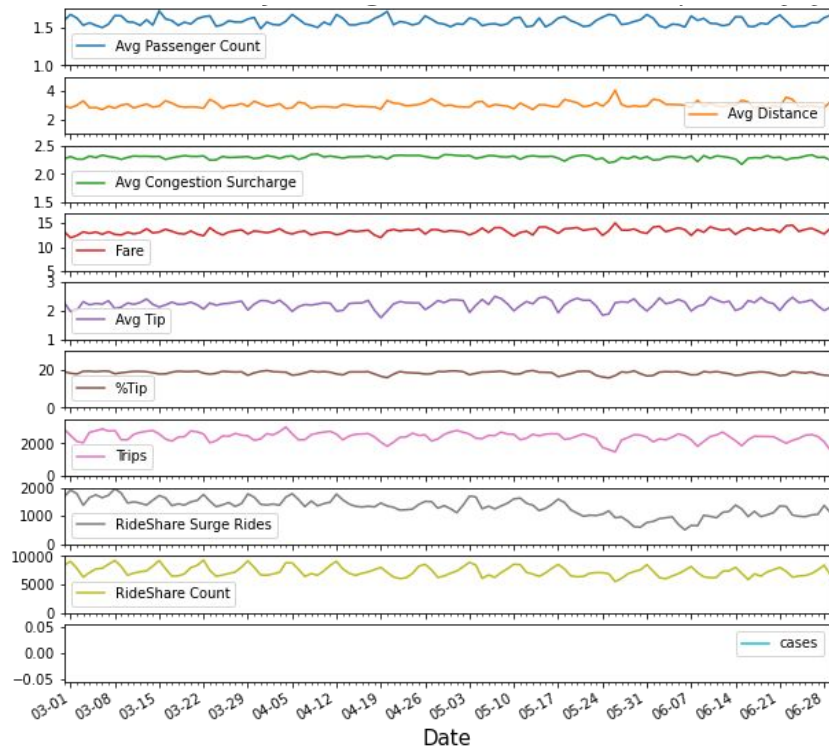
1st Case in NY

Public Transport Discouraged

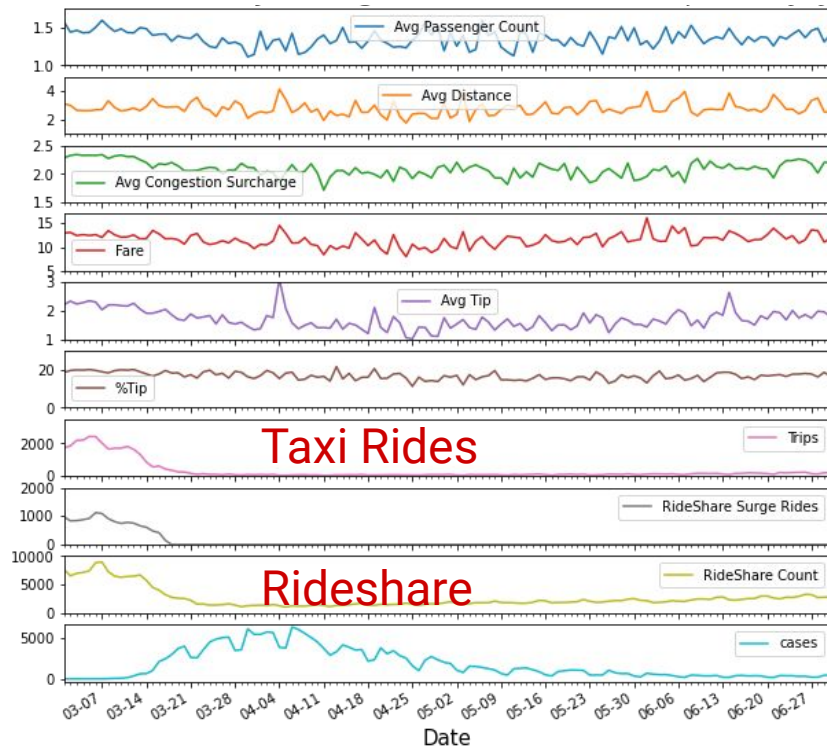


NY Taxi Daily Patterns, Pre and Post Covid

Mar-Jun 2019



Mar-Jun 2020



Short Trips = Higher Tips (%)

Influence (Gain)

Location: 0.23

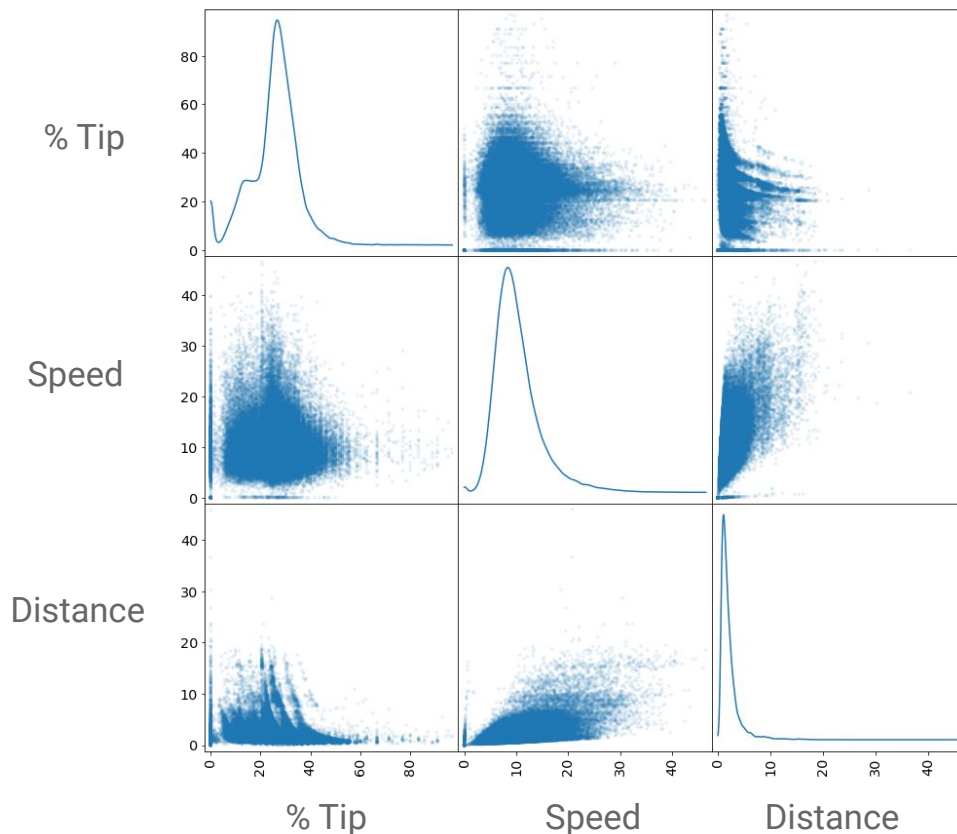
Distance: 0.27

Passengers: 0.20

Avg Speed: 0.22

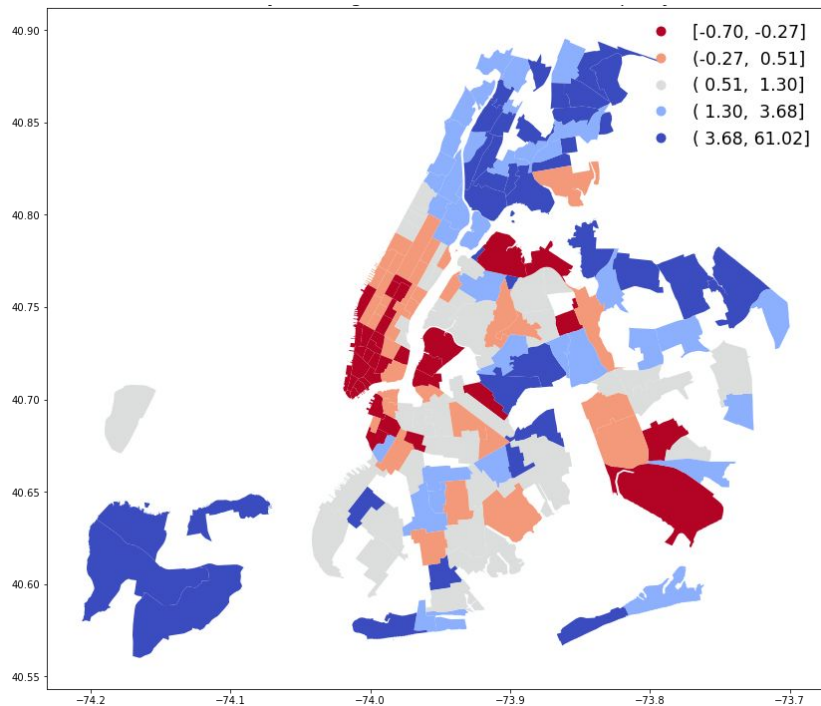
Classification by:

- Tip Size Above or Below 20%
- Distance and Speed Groups
- Passenger Count

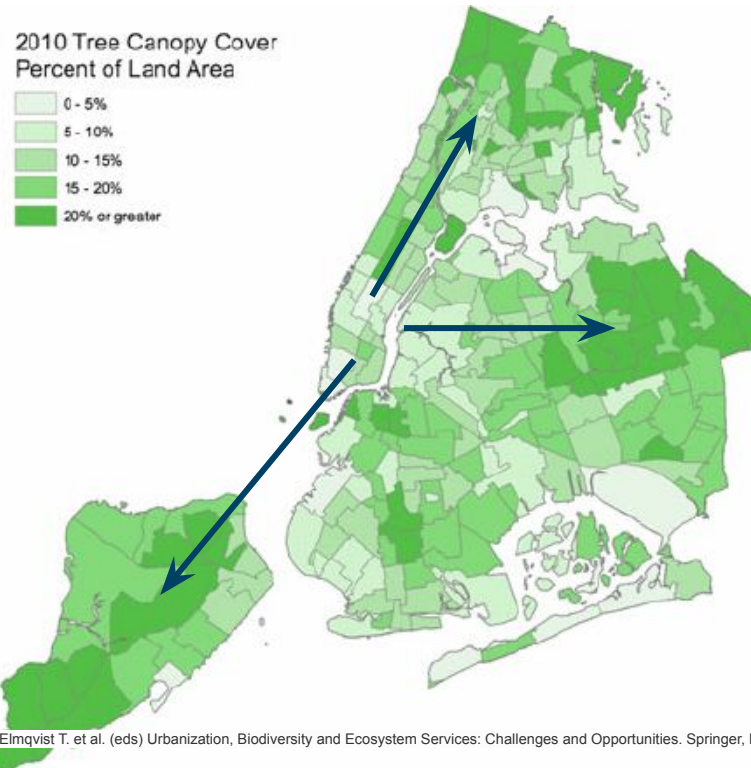


Increased Traffic to Open Space

% Change in Pickups by borough



Tree Canopy Cover*



Economic Impact of Less Taxis

Monthly Tax Revenue

Pre Covid:	\$19,358,000
------------	--------------

Post Covid:	\$446,000
-------------	-----------

Estimated Impact Since March 2020

Tax Revenue	\$ -151,300,000
-------------	-----------------

Tips	\$ -128,610,000
------	-----------------