ESRI International User Conference 2014, San Diego, CA
“Travel Patterns of NYC’s Yellow taxis: Routing, Activity and Results”

Commissioner Polly Trottenberg, New York City Department of Transportation (NYCDOT)
Prepared by: Stanislav Parfenov, Andrew Weeks and Zamir Alam
Presented by: Stanislav Parfenov, GISP on July 16, 2014

NEW YORK CITY DOT

Outline

1. Introduction
2. Study Area Map
3. NYC Medallion Taxi GPS and Taxi GPS Breadcrumb
4. SSI Calendar
5. Data Analysis
6. Data Visualization
7. Findings
8. Future Steps
9. Conclusions
10. Gratitude

Photo credit: Rodney Stiles, NYCTLC

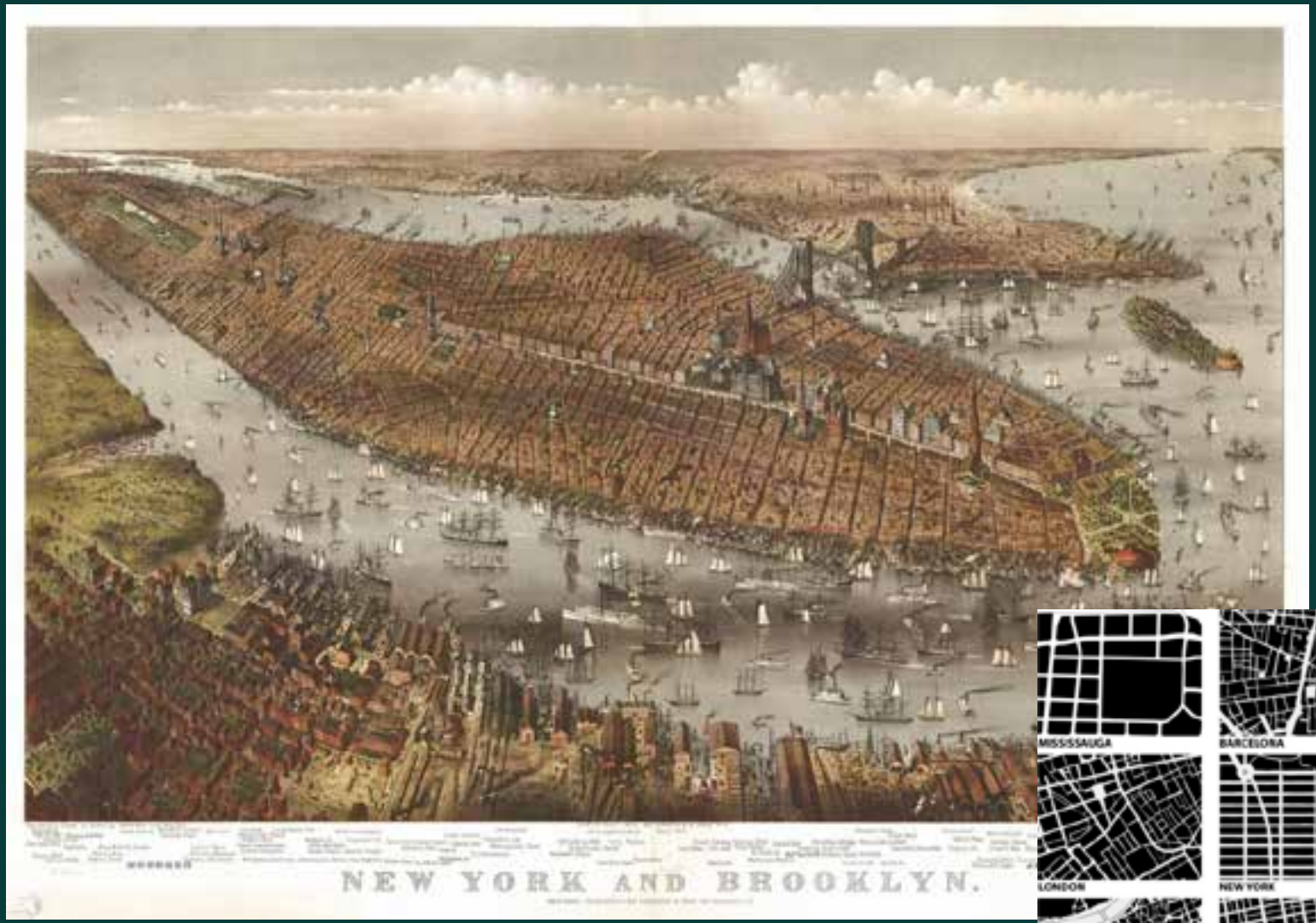
Introduction

- Project’s Main Points:

  - The project focuses on use of taxi GPS technology to analyze travel patterns and travel behavior.
  - Firstly, the research identifies the seasonality in the activity of taxi users in Manhattan’s Central Business District (CBD).
  - Secondly, Manhattan’s grid network system is used to quantify the differences in seasonality and travel behavior of the taxi drivers, along with ESRI’s Network Analyst tool.
  - Finally, taxi GPS data visualizations are prepared to support the research’s findings.

Study Area Map

NYC in 1875



http://sunnycv.com/steve/maps/list-birdseye.html

http://helmofthepublicrealm.com

Study Area Map

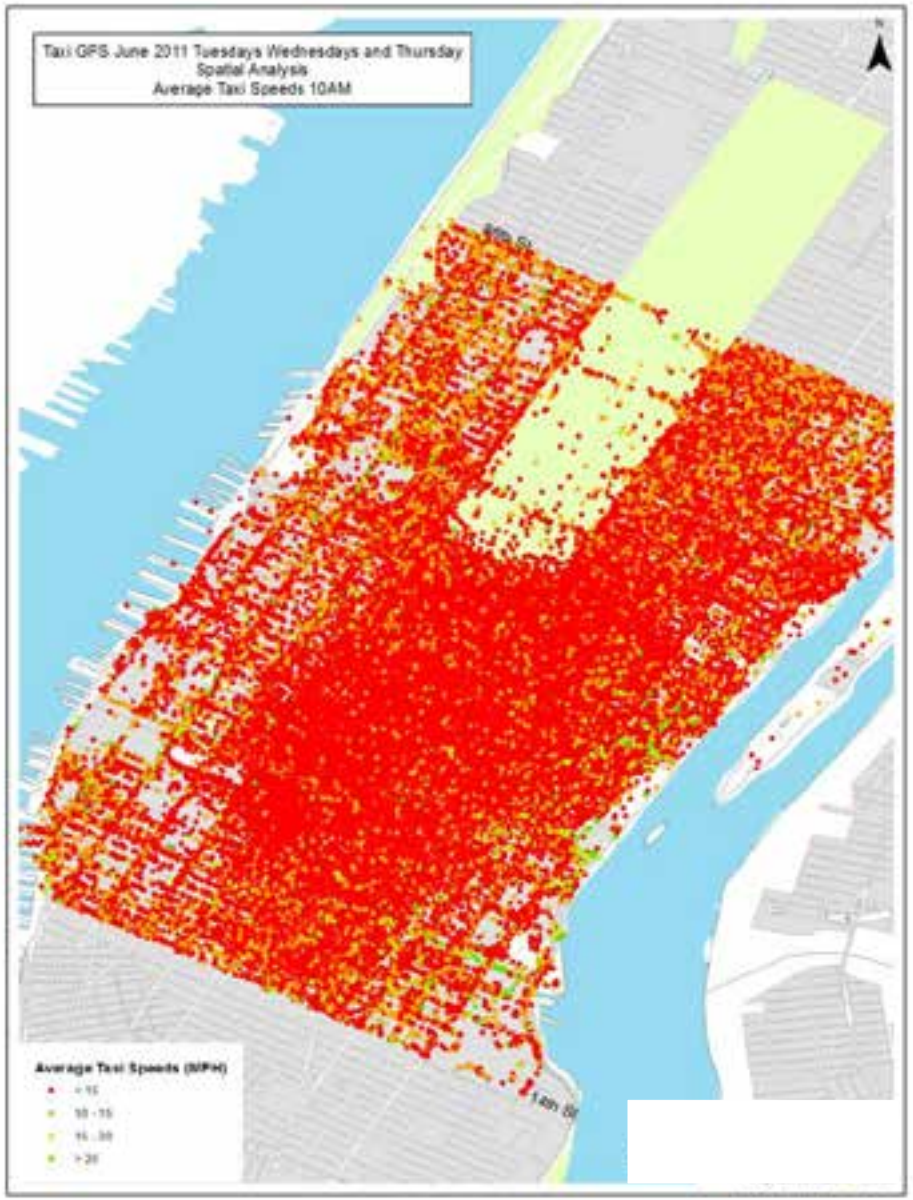
- CBD Major Destinations:
  - Financial District
  - New York University
  - World Trade Center
  - Chelsea Piers
  - Times Square
  - Broadway theaters
  - Little Italy and Chinatown
  - Statue of Liberty and Battery Park

Taxi GPS Info

- Currently, NYC Taxi and Limousine Commission (TLC) collects travel data from all medallion taxi vehicles (yellow taxis)
- Each month has about 13 million trips
- Data is spatially georeferenced for each trip
- Travel data that is collected for each trip includes, but is not limited to: travel time and distance, speed, fare and tolls paid, and other variables
- Data is processed for each month of the year and scrubbed for data errors and outliers

- For this project, taxi GPS data from 2011 through 2014 were used.

Data Visualization



Color is showing:
Red - speed below 10 MPH
Green – speed above 25 MPH

AM Peak Period
6 – 10 AM

Taxi GPS Info

FareAmt	TipAmt	Surcharge	TotalAmt	TripTime	TripDistance	StartLon	StartLat	EndLon	EndLat	Speed	VendorName	TripNumber	PickU
6.90	.00	.00	7.40	10.00	1.18	-73.980107	40.746412	-73.984002	40.754977	7.08	VTS	9830	5A
4.90	60	.00	6.00	6.00	.89	-73.984600	40.780058	-73.987670	40.770133	8.90	VTS	31201	6A
6.90	2.00	.00	9.40	11.00	.88	-73.985150	40.756980	-73.975238	40.750480	4.80	VTS	26559	4C
5.70	.00	.00	6.20	8.00	.84	-73.982442	40.767328	-73.969525	40.760705	6.30	VTS	26768	4C
11.30	.00	.00	11.80	15.00	3.09	-73.982540	40.740647	-73.960962	40.777542	12.36	VTS	19813	5A
7.70	4.00	.00	12.20	12.00	1.28	-73.982422	40.768145	-73.974557	40.757210	6.40	VTS	21295	4C
6.50	3.00	.00	10.00	9.00	1.02	-73.990005	40.752143	-73.985945	40.758812	6.80	VTS	7431	4B
7.70	.00	.00	8.20	12.00	1.68	-73.974607	40.783158	-73.981818	40.766228	8.40	VTS	25993	6A
6.90	00	.00	7.40	10.00	1.04	-73.991047	40.750822	-73.985628	40.740010	6.24	VTS	20090	4A
6.10	00	.00	6.60	9.00	.58	-73.982290	40.768172	-73.973342	40.763710	3.87	VTS	26008	4C
7.30	.00	.00	7.80	10.00	1.54	-73.976700	40.788392	-73.988935	40.771930	9.24	VTS	7047	6A
6.10	.00	.00	6.60	8.00	1.26	-73.974993	40.760872	-73.974245	40.750442	9.45	VTS	14552	4C
10.10	.00	.00	10.60	19.00	1.27	-73.962283	40.779012	-73.968763	40.764188	4.01	VTS	13506	6A
11.30	1.00	.00	12.80	19.00	2.28	-73.979493	40.784570	-73.960507	40.762547	7.20	VTS	2572	6A
4.50	2.00	.00	7.00	5.00	.78	-73.984807	40.741782	-73.984102	40.735522	9.36	VTS	18073	5A
6.90	2.00	.00	9.40	10.00	1.27	-73.978567	40.762497	-73.992973	40.757815	7.62	VTS	24874	4C
5.30	1.00	.00	6.80	6.00	.55	-73.982068	40.770407	-73.978467	40.774490	5.50	VTS	7602	6A
8.10	.00	.00	8.60	12.00	1.57	-73.972312	40.765318	-73.973048	40.780147	7.85	VTS	2041	6A
7.70	00	.00	8.20	11.00	1.74	-73.977447	40.753127	-73.982525	40.771555	9.49	VTS	26896	5C
12.50	00	.00	13.00	25.00	82	-73.963745	40.777147	-73.971753	40.766525	1.97	VTS	24845	6A
10.90	00	.00	11.40	18.00	2.27	-73.998287	40.744540	-73.973320	40.756943	7.57	VTS	22364	2B

Sample Monthly Medallion Taxi GPS Data Table

Taxi GPS Zones

Taxi GPS Zones Manhattan

Taxi GPS Zones NYC

Taxi GPS Analysis

- Do taxis behave differently depending on time of year in Manhattan?
- Are some parts of the city more impacted during summer-like and non-summer like conditions?

- Taxi GPS data from representative summer and non-summer months were used for the analysis (based on 3 years of data or about 470 million records) to compare those conditions.
- NYC Taxi Zones were used as analysis geometries.

Taxi GPS Calendar

Through NYCDOT's Sustainable Street Index program, a Taxi GPS is developed for each day, week, month, season, and year:

2012 Taxi GPS Calendar

Seasonal Findings

- Summer vs Non-Summer Conditions:

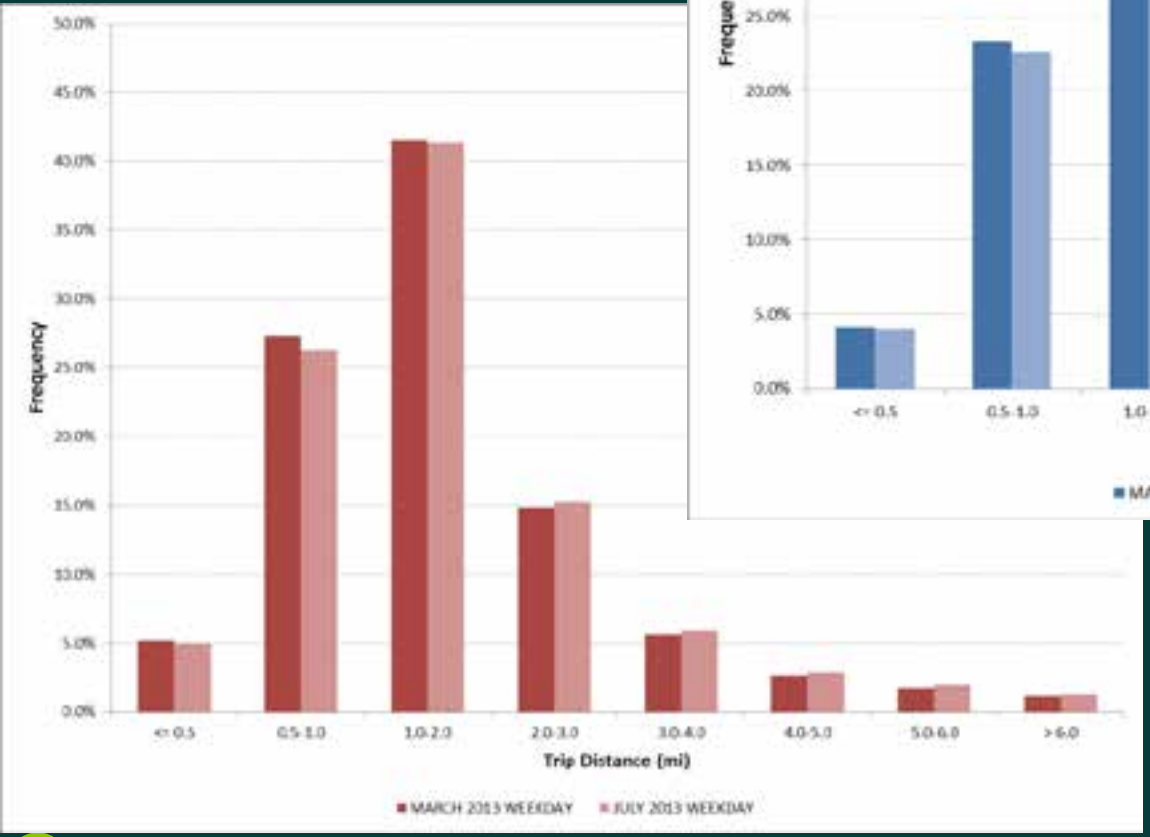
  - According to tourist bureau statistics [1], the peak of tourist activity is in the summer at about 5 million people, while non-summer months are at a little over 3 million
  - Summer trips are about 5% longer in time and only 2% longer in distance, than non-summer trips
  - However, the amount of activity heavily differs – there are 12% more non-summer trips than summer trips or 1.5 Million trips
  - This translates into almost $16 Million less in fare amount and about $1.7 Million less in tips for taxi drivers during summer months

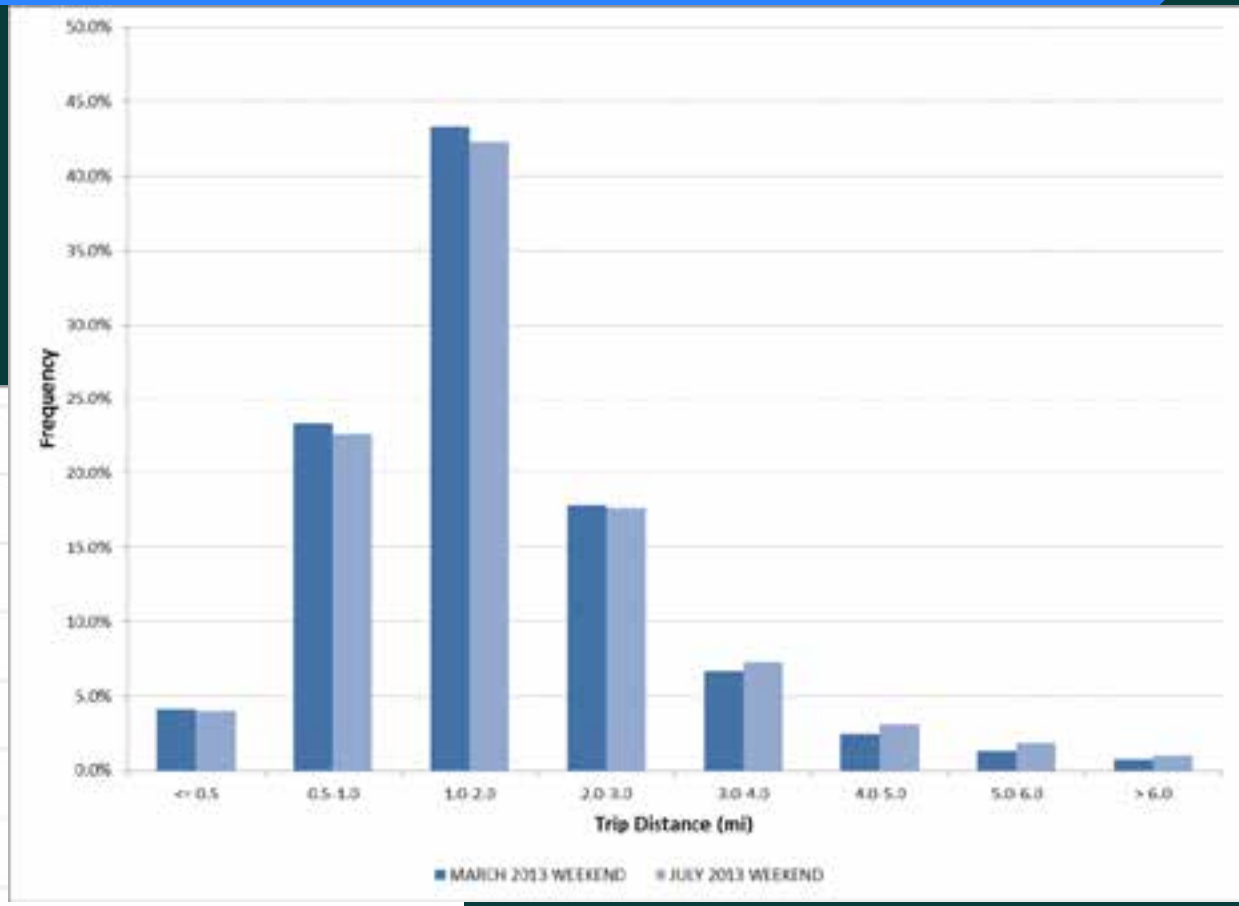
1. http://www.nytix.com/NewYorkCity/articles/temperature.html

CBD Seasonal Findings – Weekdays and Weekend

- Summer compared to Non-Summer Weekdays and Weekends:
  - Weekday trips in the summer have 5.5% slower speeds, with longer trip times and distances by about 8% and 2.5% respectively
  - Results in 106,000 fewer trips
  - Weekends in the summer have 4% faster speeds, but there is no difference in travel time, while distance is 4% longer
  - Results in about 316,000 fewer trips
  - These differences highlight that:
    - Trips are longer as a result of the people's activities and their interaction with land use, and heavy effects of tourism
    - Even though there are fewer business trips on the weekend (faster speeds), demand for taxis remains high

CBD Seasonal Findings – Weekdays and Weekend



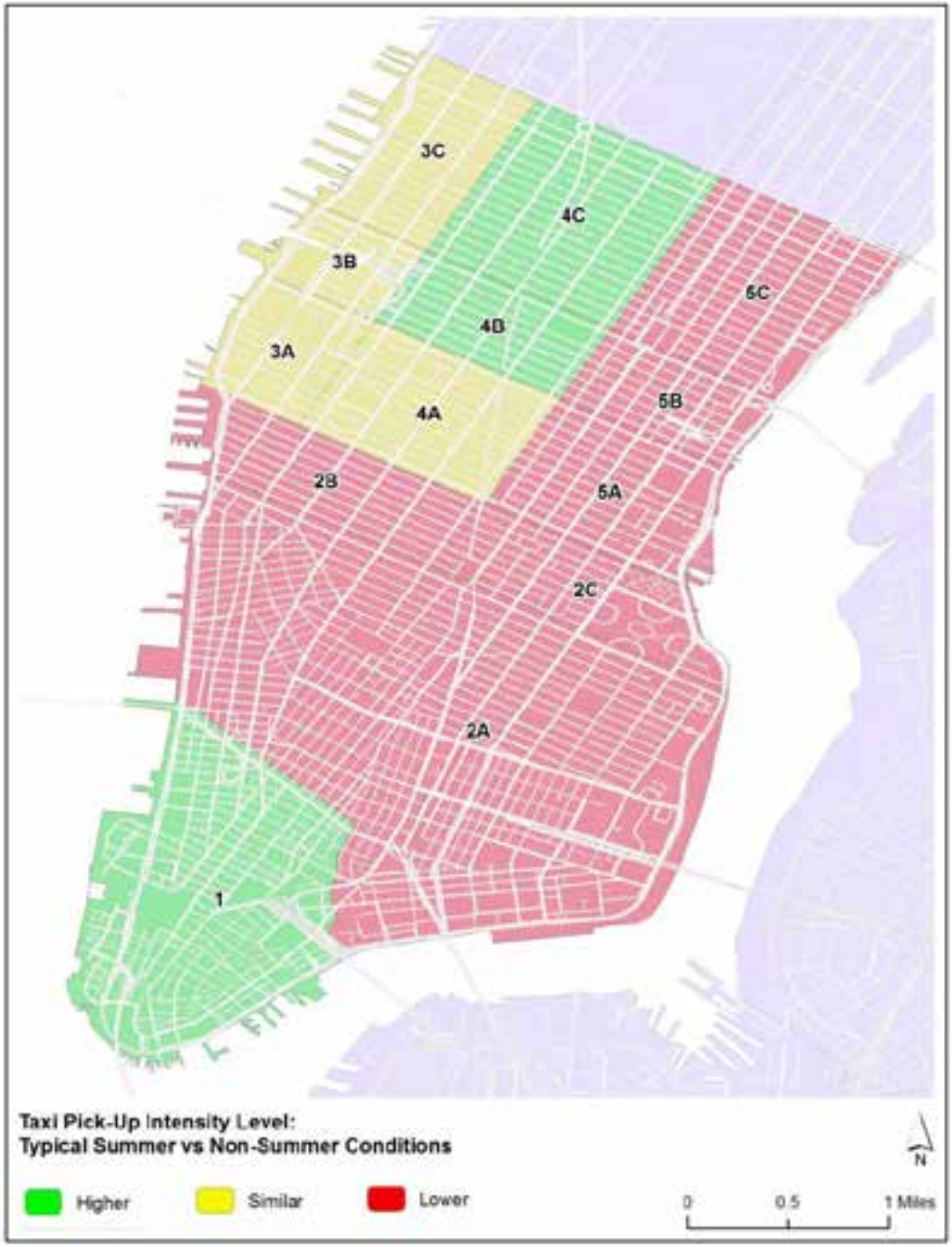


Shift towards longer trips in the summer for both weekends and weekdays.

- Possibly due to tourist activity, with specific POI's in CBD.

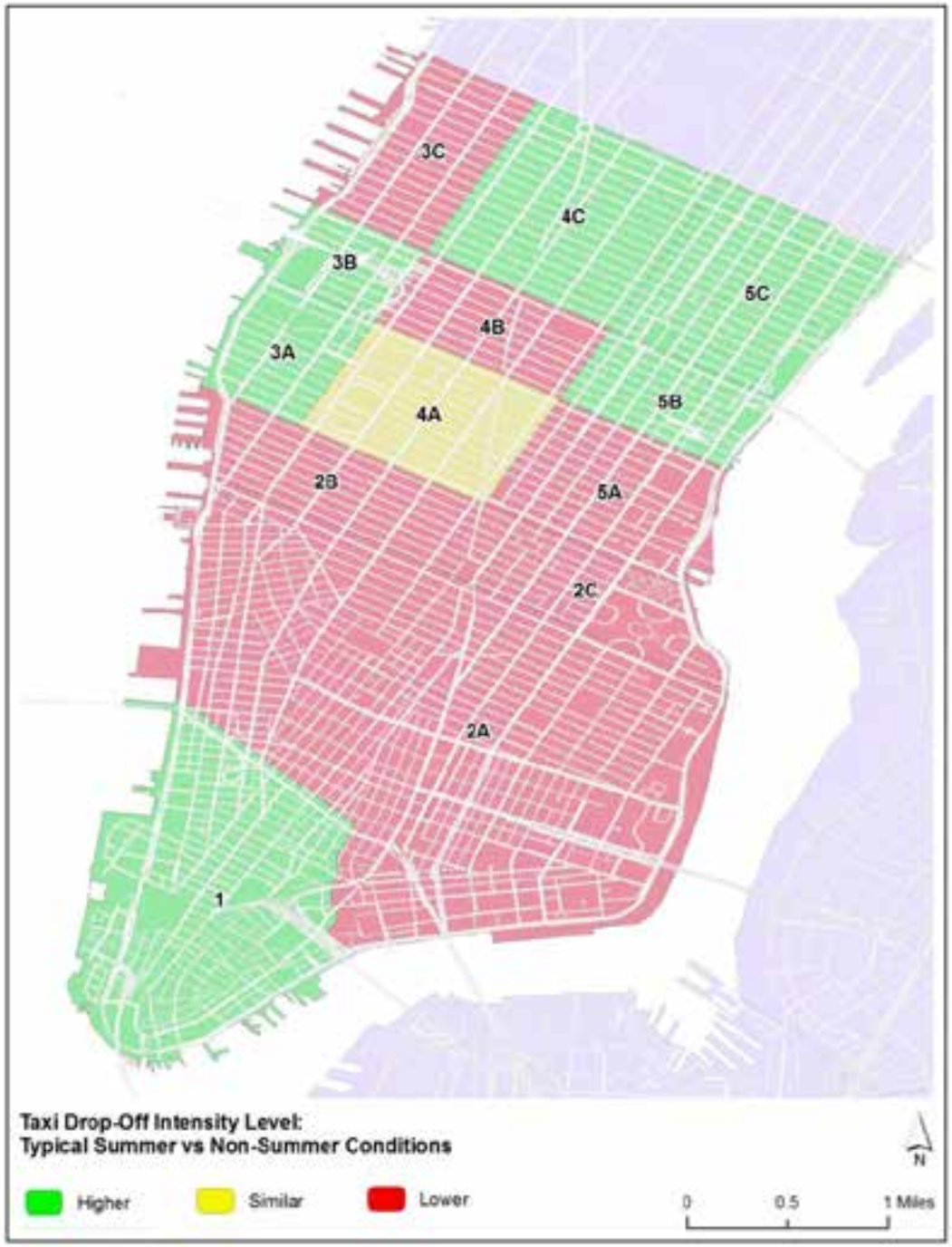
CBD Seasonal Findings – Weekdays and Weekend

- These differences highlight that:
  - People's behavior is driven by seasonality, with warmer summer weather allowing people to rely on other modes of transport rather than costly taxi rides (e.g. walking, bicycling, etc.,).
  - So what about spatial distribution?
  - Do certain parts of the city have different activity levels when we take seasonality into account?



Destinations such as:

- Lower Manhattan – Financial District, Statue of Liberty and Ground Zero.
- Midtown – Carnegie Hall, Museum of Modern Art, Times Square, Rockefeller Center, Empire State Building, Macy's Store, Broadway Theaters.

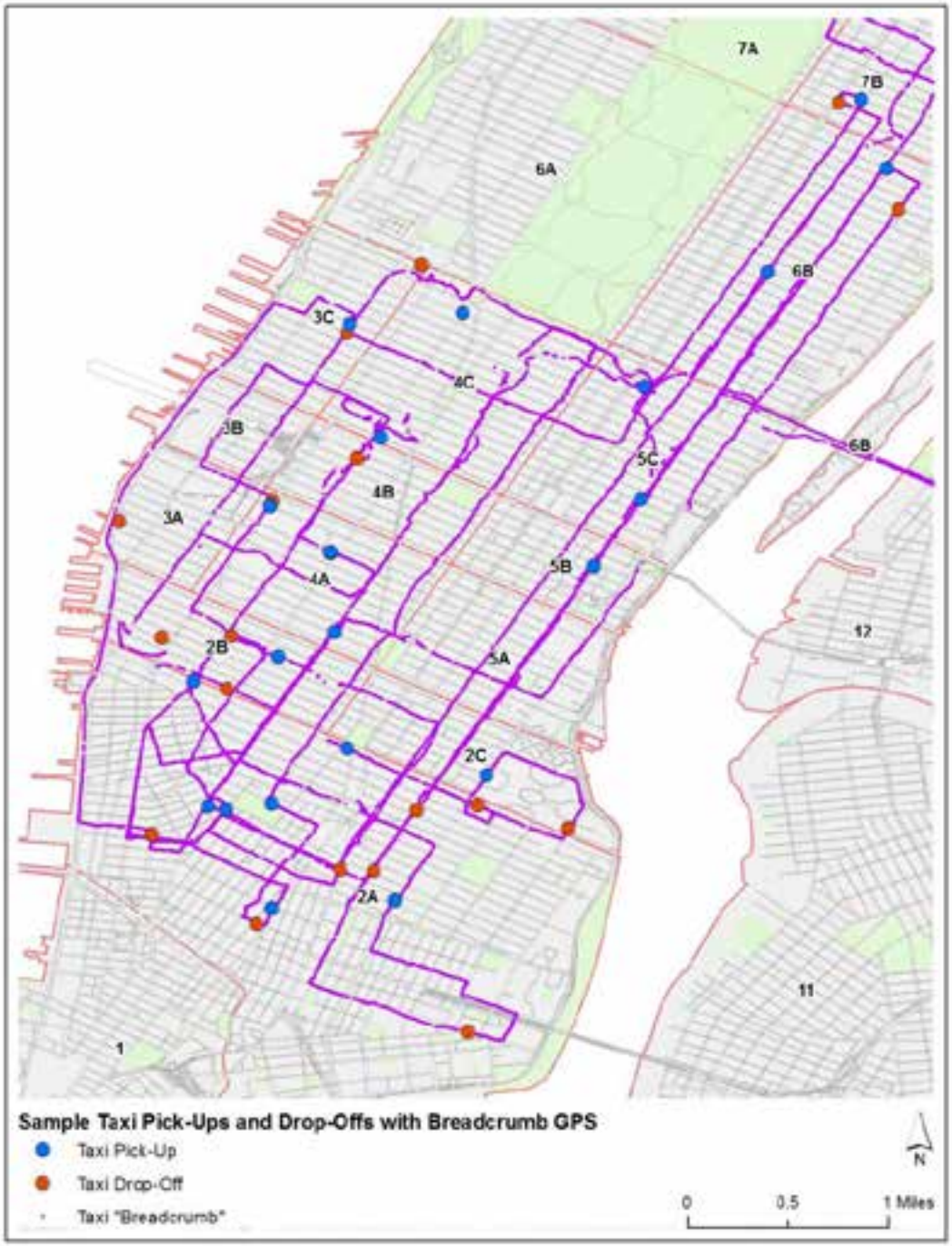


Besides aforementioned Lower Manhattan and Midtown Locations:

- In 3A and 3B:
  - Port Authority Bus Terminal and PATH Train (Connection to NJ).
  - Penn Station
- In 5C and 5B:
  - UN Headquarters
  - Grand Central Station
  - Dense residential housing developments
  - Queens Midtown Tunnel
  - Queensboro Bridge

Breadcrumb GPS

- For the pilot project, 6 taxis were equipped with GPS loggers, which collected XY location every 1 second for each taxi's work day from Summer 2013 to Spring 2014
- The goal of the project was to gather the taxi route information, since the TLC's Taxi GPS dataset does not include the exact route of each trip and to compare it to the ESRI's Network Analyst's network routing
- OpenStreetMap shapefile and Python scripting were used to code the network to be used with ESRI's Network Analyst
- Travel behavior was analyzed to determine the duration of the average taxi trip, as well as time spent looking for the next passenger
- Seasonality was taken into account and summer vs non-summer conditions were compared

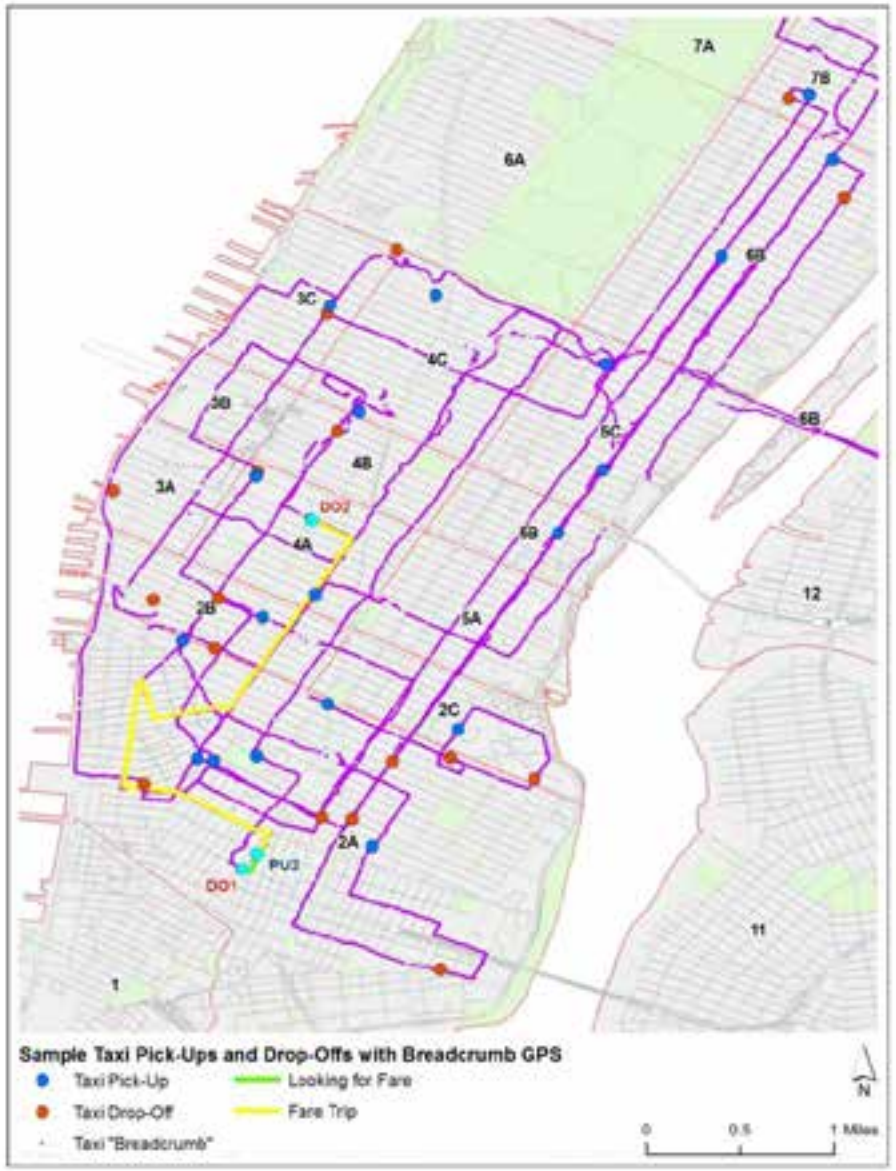

7A
7B
6A
6B
3C
4C
3B
5C
6B
4B
3A
5B
4A
12
2B
5A
2C
2A
11
1
Sample Taxi Pick-Ups and Drop-Offs with Breadcrumb GPS
Taxi Pick-Up
Taxi Drop-Off
Taxi "Breadcrumb"
N
0
0.5
1 Miles

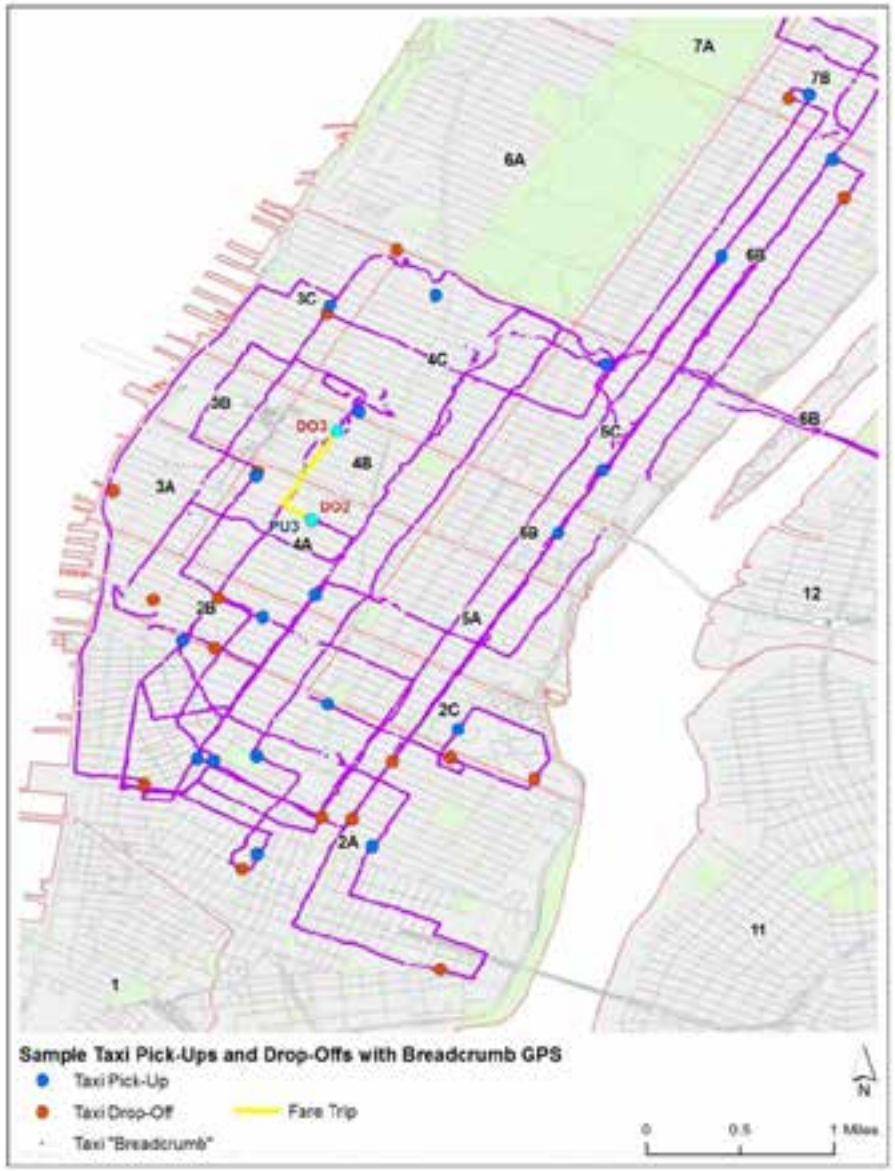
Breadcrumb GPS

- Findings:
  - It took less time to find a passenger in the summer than non-summer months
  - It took longer to find a passenger on the weekday than weekend
  - For both summer and non-summer trips, there is at least 37% probability that if a new passenger was picked up close to the drop off location, that trip will also have a successive quick pick-up

Summer		Non-Summer		
Far	Close	Far	Close	
<20%	>65%	<20%	53%	**Weekend**
21%	37%	9%	41%	**Weekday**

Breadcrumb GPS





Breadcrumb GPS

Breadcrumb GPS

- ESRI Network Analyst Findings:
  - The routes for mostly Avenue trips were easier to predict, overall with 50% accuracy for both summer and non summer, because more trips begin and end on Avenues
  - Mostly Crosstown trips could be predicted with about 40% accuracy for both seasons
  - Combined Avenue-Crosstown routes could be predicted with 13% accuracy in the Summer conditions and 58% for Non-Summer Conditions

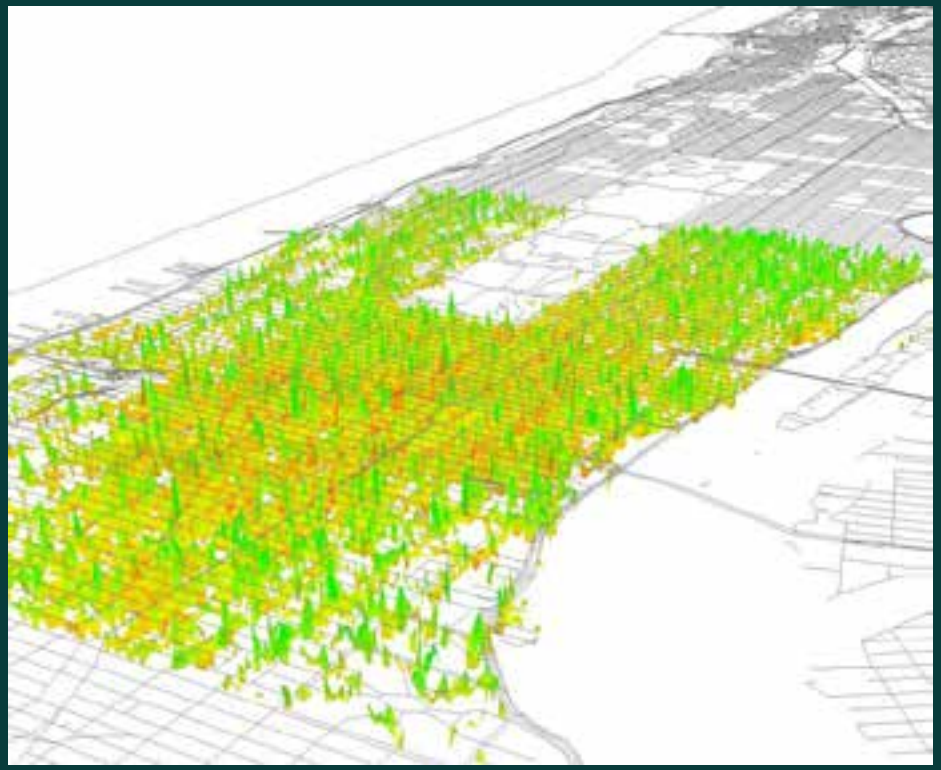
Network Analyst comparison compared to actual taxi trips		Avenue	Crosstown	Avenue-Crosstown
Summer	**Weekday**	50%	38%	13%
	Weekend	50%	50%	N/A
Non-Summer	**Weekday**	67%	25%	N/A
	Weekend	45%	50%	17%
Summer	All Days	50%	44%	13%
Non-Summer	All Days	56%	38%	17%

Findings: Effects of Grid Network

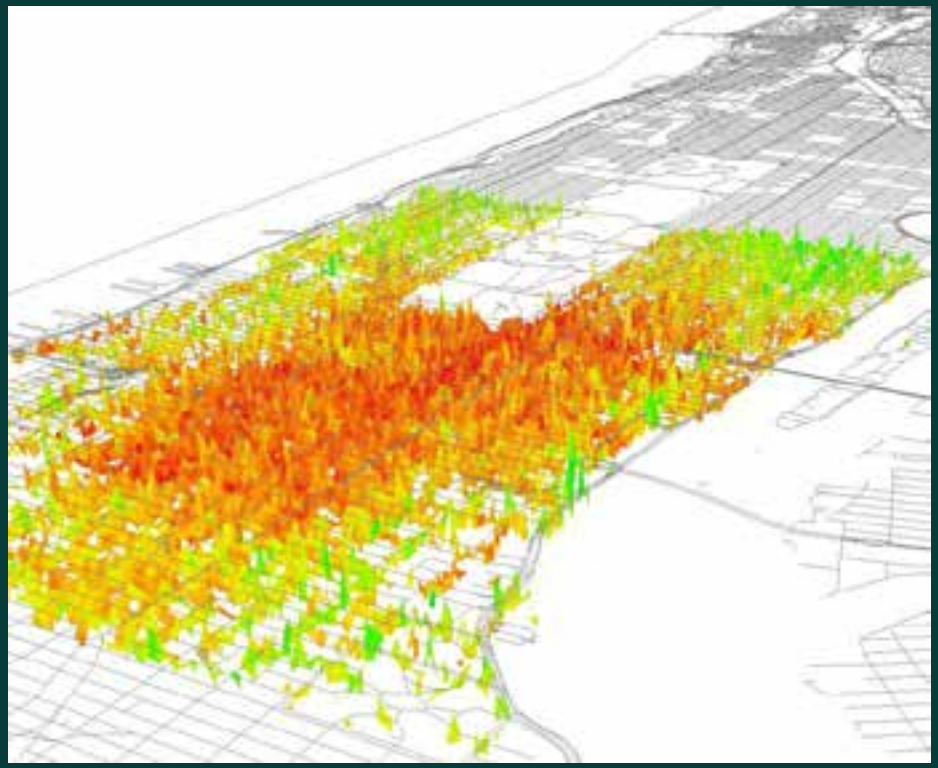
- Grid network provides simple planning and routing capabilities, with option of one-way streets and avenue designs
- However, resulting congestion due to special events might result in additional roadway congestion and driver confusion
- Prohibited right or left turns and U-turns may reduce drivers' options and thus lead to lengthened trips
- It is relatively easy to predict avenue trips, but hard to predict offset crosstown trips, or a combination of avenue-crosstown trips
- ESRI Network Analyst provided shorter routing than the taxi drivers took, which was more pronounced in Non-Summer Weekends
- Special traffic conditions (events, double parked vehicles, road work, etc.,) on a given street may lead to forced route changing

Data Visualization

Color is showing:
Red - speed below 10 MPH
Green – speed above 25 MPH



June 16, 2011
Typical Thursday



June 23, 2011
Special Event Thursday

Data Visualization

Canyon Effect

Conclusions

- Using a combination of GPS location technologies to supplement each other to predict the travel patterns in the city grid, shows to be an effective approach
- ESRI's Network Analyst proved to be a robust tool for the analysis of the Manhattan's CBD urban grid, however, more detailed street level editing is needed for better route prediction
- There is a strong spatial and seasonal demand for taxis in the city, that is most likely driven by the influx of tourists
- Summer months tend to be more impacted even with fewer taxi trips, but longer trips and roadway activities increase demand for taxis (it takes less time to find a taxi in the summer months than in non-summer months)
- Taxis proved to be an excellent source of probe data, considering there are over 13,000 taxis (vehicles in motion)

Future Steps Analysis

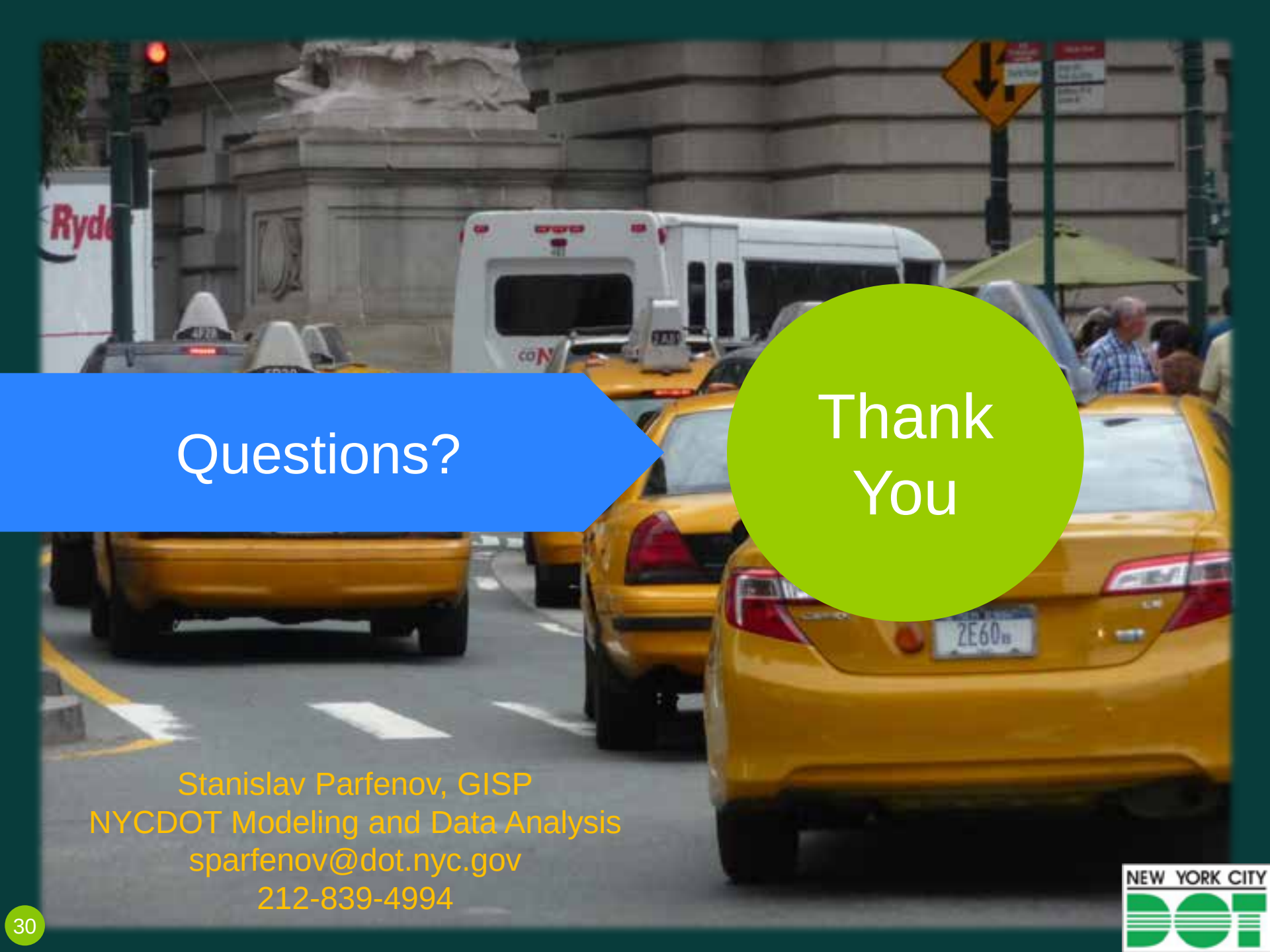
Work will continue, but will not be limited to the following activities:

- Identify road segments that experience congestion more than 2/3 of the time, and build them into the network, which would allow for better route choice prediction
- Identify road segments that experience planned events and road closures for better prediction of route choice
- Integrate lower frequency GPS readings into the workflow – NYCDOT tests show that 10 to 30 second “pings” will be an adequate frequency to more accurately show the trip travel route
- Create more robust algorithm to predict taxi driver behavior, which would allow a more accurate taxi activity, based both on season and time of day
- Work with NYCTLC to secure more frequent breadcrumbs and real-time data

Gratitude

I would like to sincerely express my gratitude to a number of people who I had a pleasure to collaborate and get assistance from to accomplish this project:

- Assistant Commissioner, Mike Marsico, P.E. (NYCDOT)
- Rodney Stiles (NYCTLC) and others at TLC for their continuing support with Taxi GPS Program
- My colleagues Andrew Weeks and Zamir Alam (NYCDOT)
- Other colleagues at NYCDOT
- Alexander Parfenov, Ph.D.



Stanislav Parfenov, GISP
NYCDOT Modeling and Data Analysis
sparfenov@dot.nyc.gov
212-839-4994