

# Modeling Urban Capacity with Public Data: Helping Realize Universal Prekindergarten in New York City

Richard Dunks • Ravi Shroff • Jeongki Lim • Miguel Castro • Haozhe Wang

New York University

Center for Urban Science and Progress

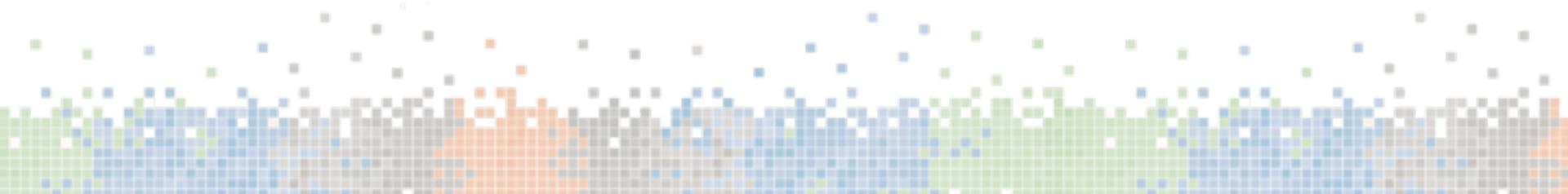
Workshop on Urban Big Data and Informatics

August 11, 2014

# Research Question

How can the location of new Universal Prekindergarten sites be optimized to ensure every 4 year old in New York City has the opportunity to attend Pre-K within a reasonable travel distance?

# Current Situation



# Current Situation

- **73,250** Four year olds in NYC in need of access to full-day pre-K (Source NYC Dept of Education)

# Current Situation

- **73,250** Four year olds in NYC in need of access to full-day pre-K (Source NYC Dept of Education)
- **41,000** Children without full day prekindergarten

# Current Situation

- **73,250** Four year olds in NYC in need of access to full-day pre-K (Source NYC Dept of Education)
- **41,000** Children without full day prekindergarten
- Prekindergarten programs run in DOE schools and DOE-contracted community based organizations (CBOs)

# **Current Situation**

Two-Year Plan:

# Current Situation

## Two-Year Plan:

- Convert **27,000** part-time prekindergarten seats

# Current Situation

## Two-Year Plan:

- Convert **27,000** part-time prekindergarten seats
- Create **14,000** new prekindergarten seats

# Data Sources



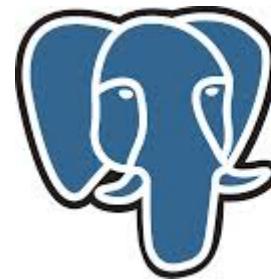
# Data Sources

- Census 2012 American Community Survey 5-year aggregated estimates

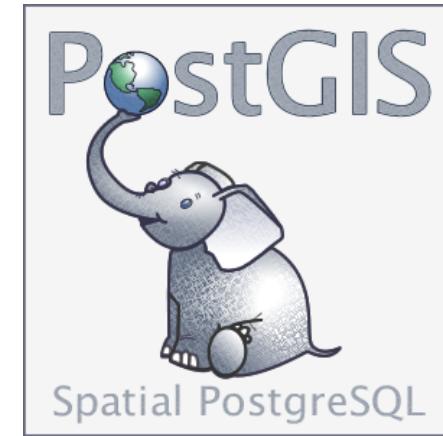
# Data Sources

- Census 2012 American Community Survey 5-year aggregated estimates
- NYC Department of Education and PediaCities prekindergarten site data

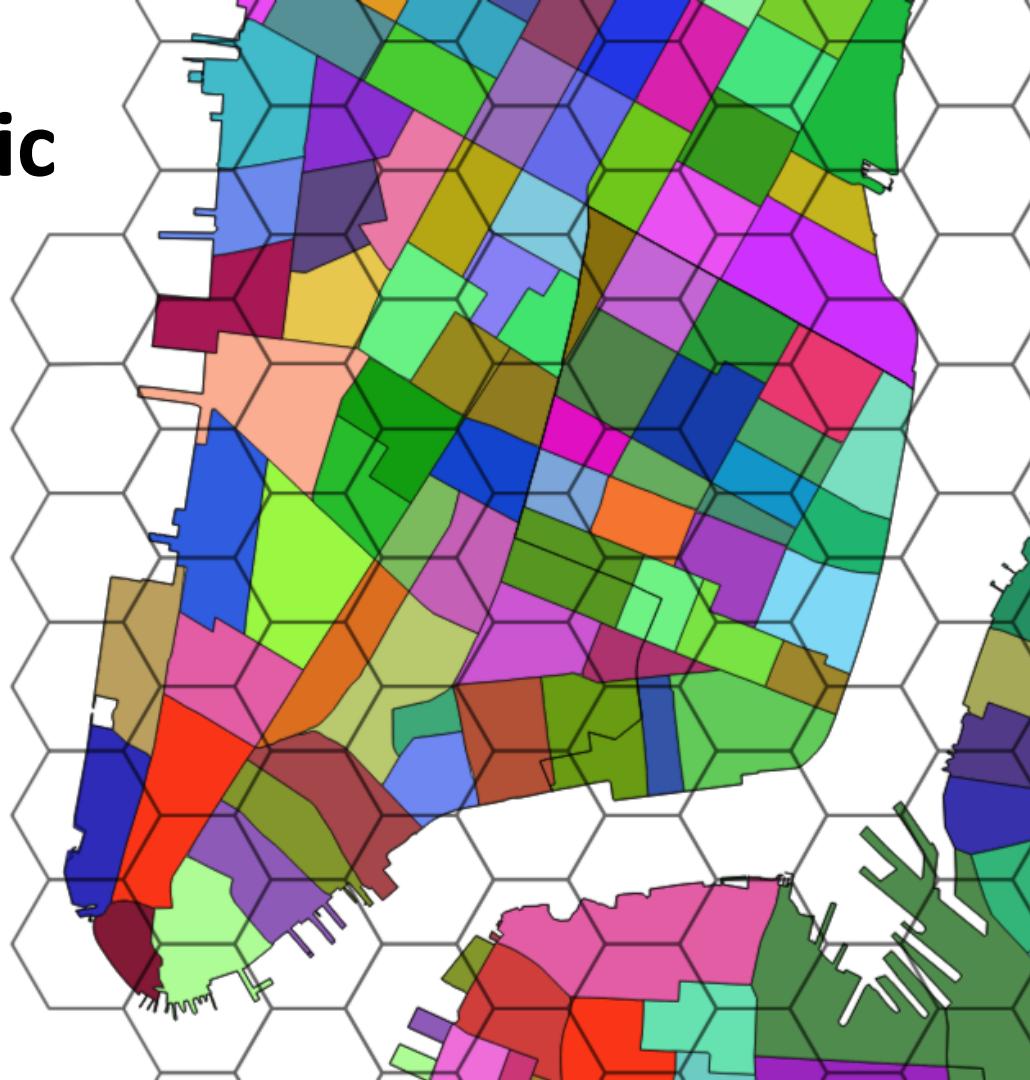
# Tools



**PostgreSQL**  
the world's most advanced open source database

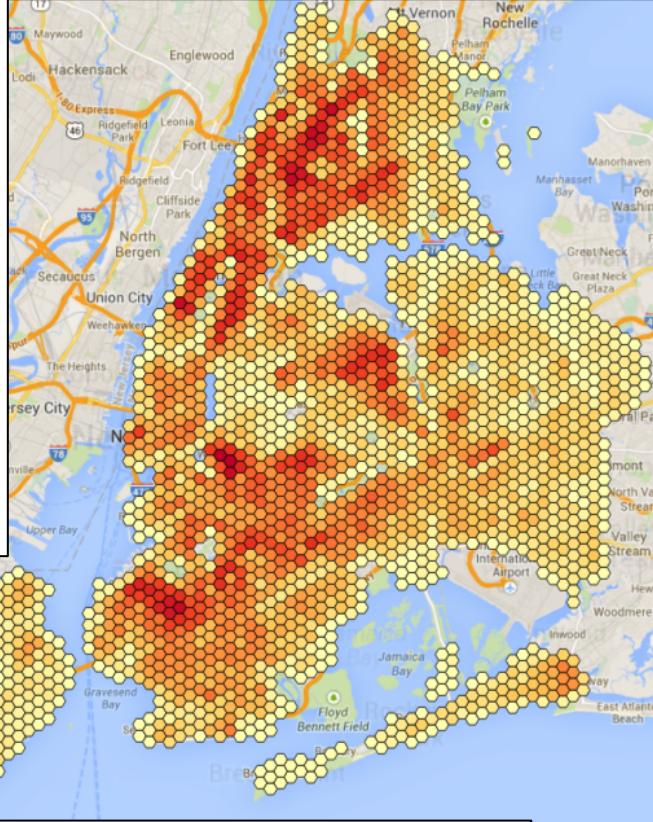
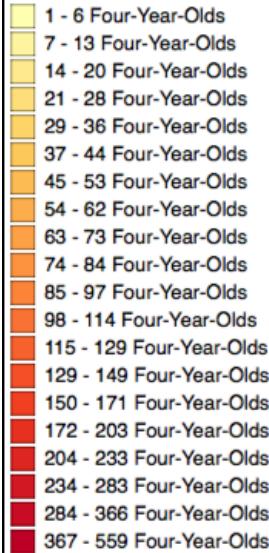


# Dasymetric Mapping

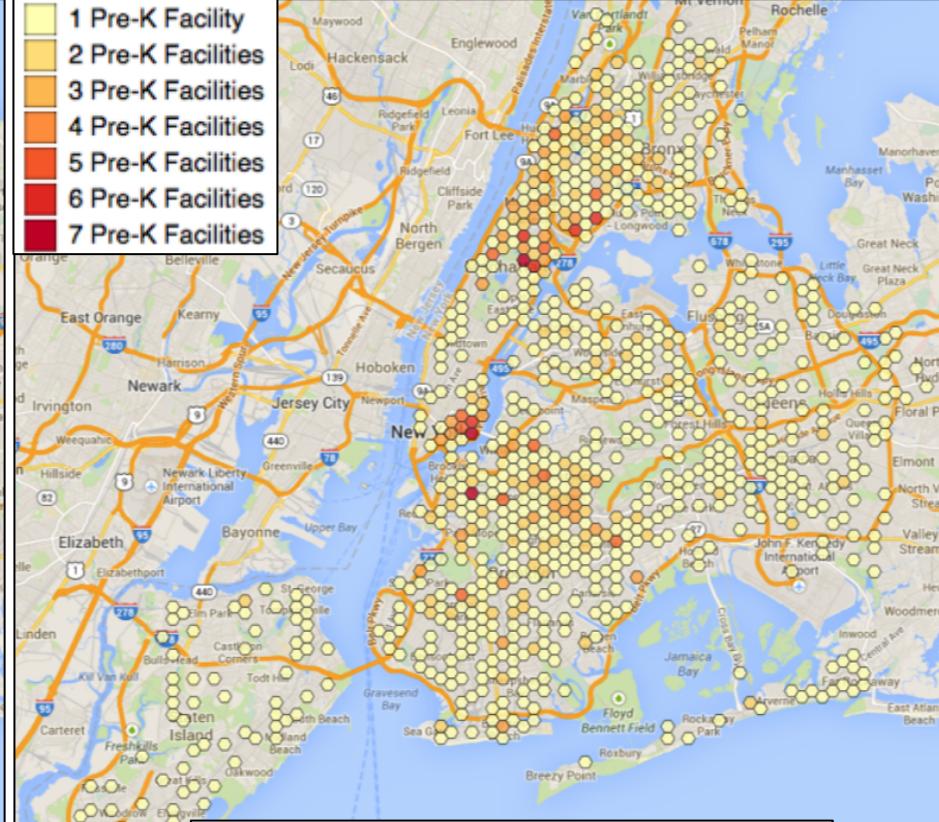


2010 Census Tracts  
with 650 meter  
hexagons overlaid

# Current Situation

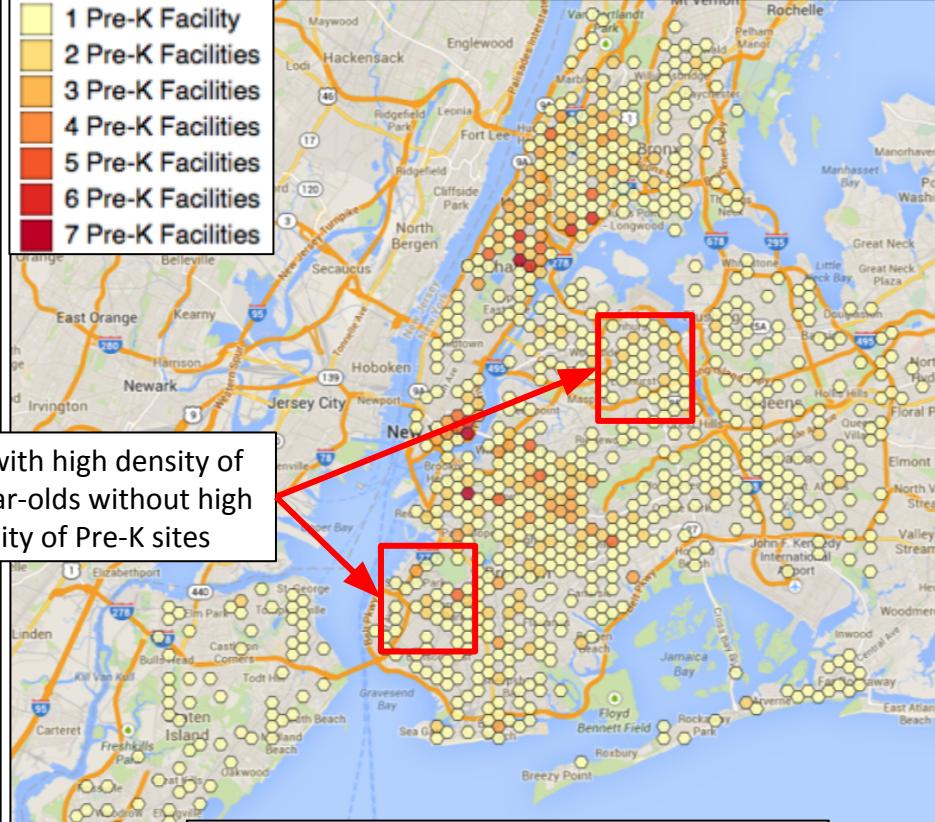
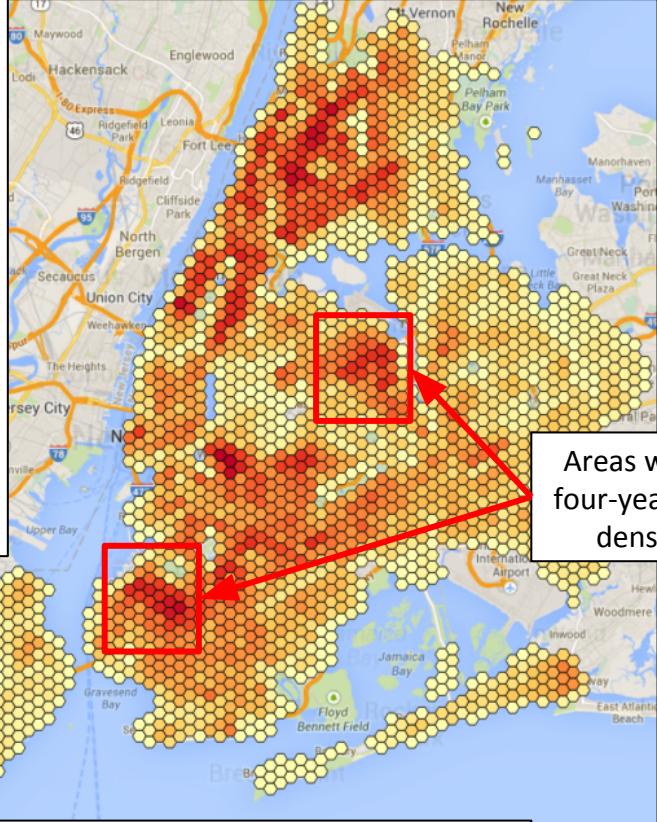
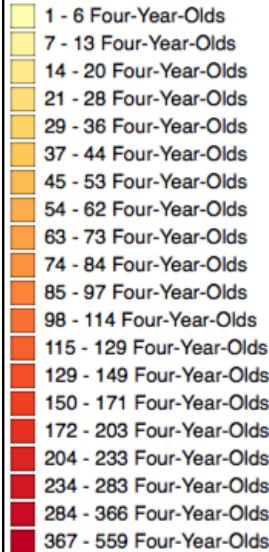


Distribution of four-year-olds in New York City  
(Source: US Census Bureau)



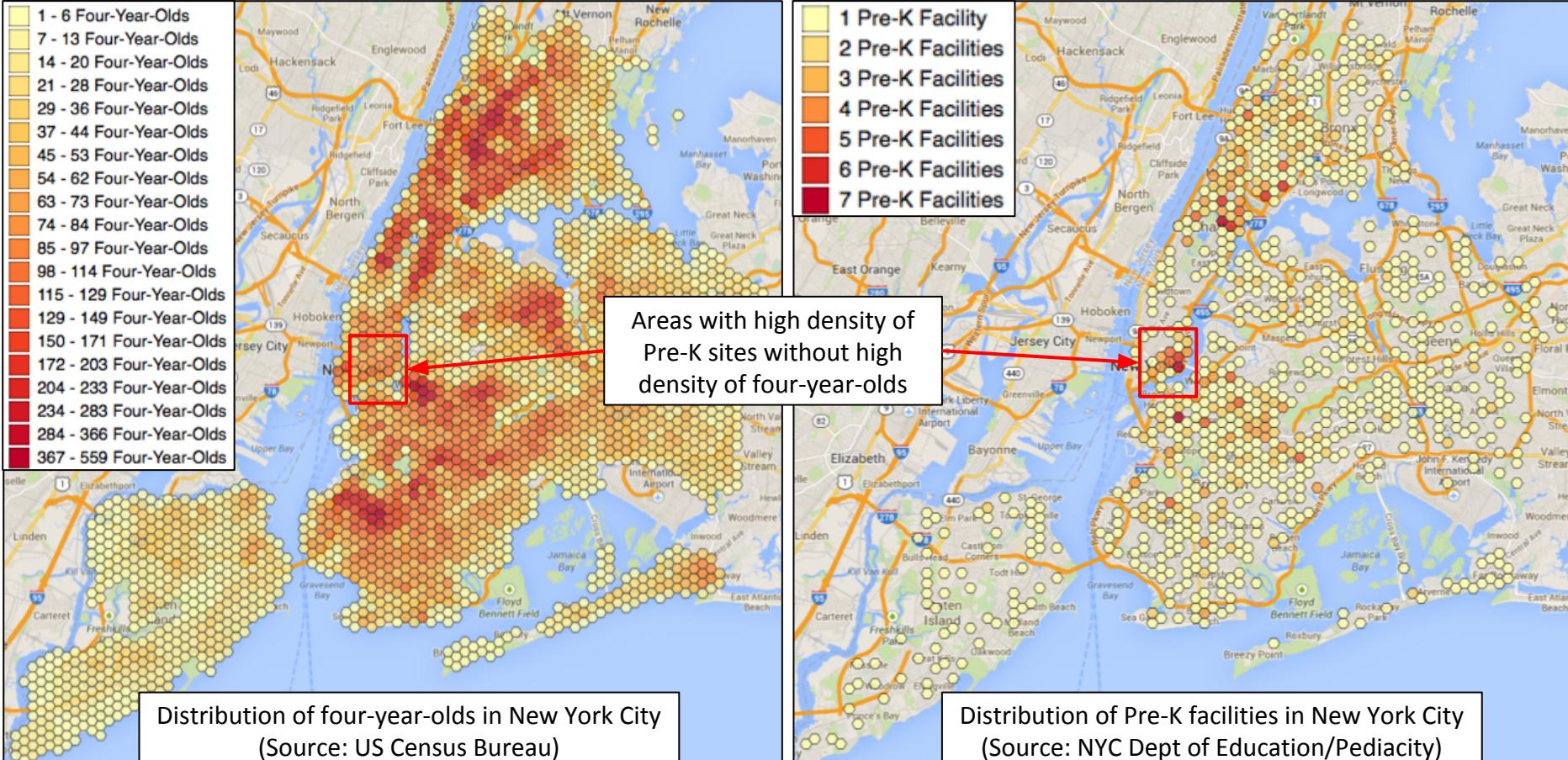
Distribution of Pre-K facilities in New York City  
(Source: NYC Dept of Education/Pediacity)

# Current Situation



Areas with high density of four-year-olds without high density of Pre-K sites

# Current Situation



# Our Approach

Use US Bureau of the Census population data and NYC Department of Education Pre-K site data, with estimates of travel time, to generate an index of underserved areas in New York City



# Allocation Algorithm

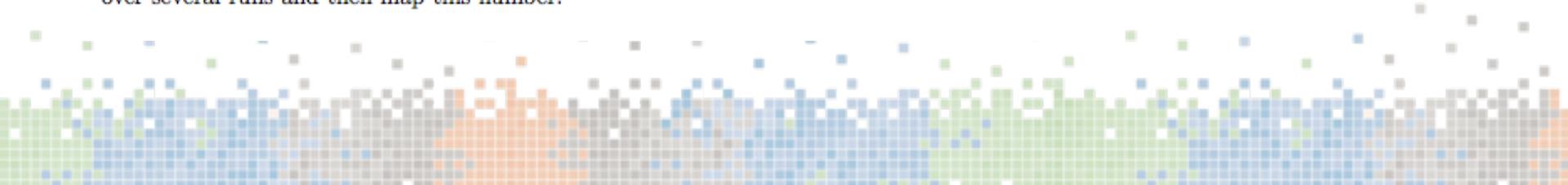
$$H_i = (P_i, S_i, C_i, N_i)$$

- $P_i$  = Population of children in  $H_i$
- $S_i$  = Current number of children assigned to  $H_i$ , initialized to 0.
- $C_i$  = Total number of slots (capacity) in  $H_i$ .
- $N_i$  = Accessible hexagons from  $H_i$ ,  $\{H_{i_1}, \dots, H_{i_n}\}$ .

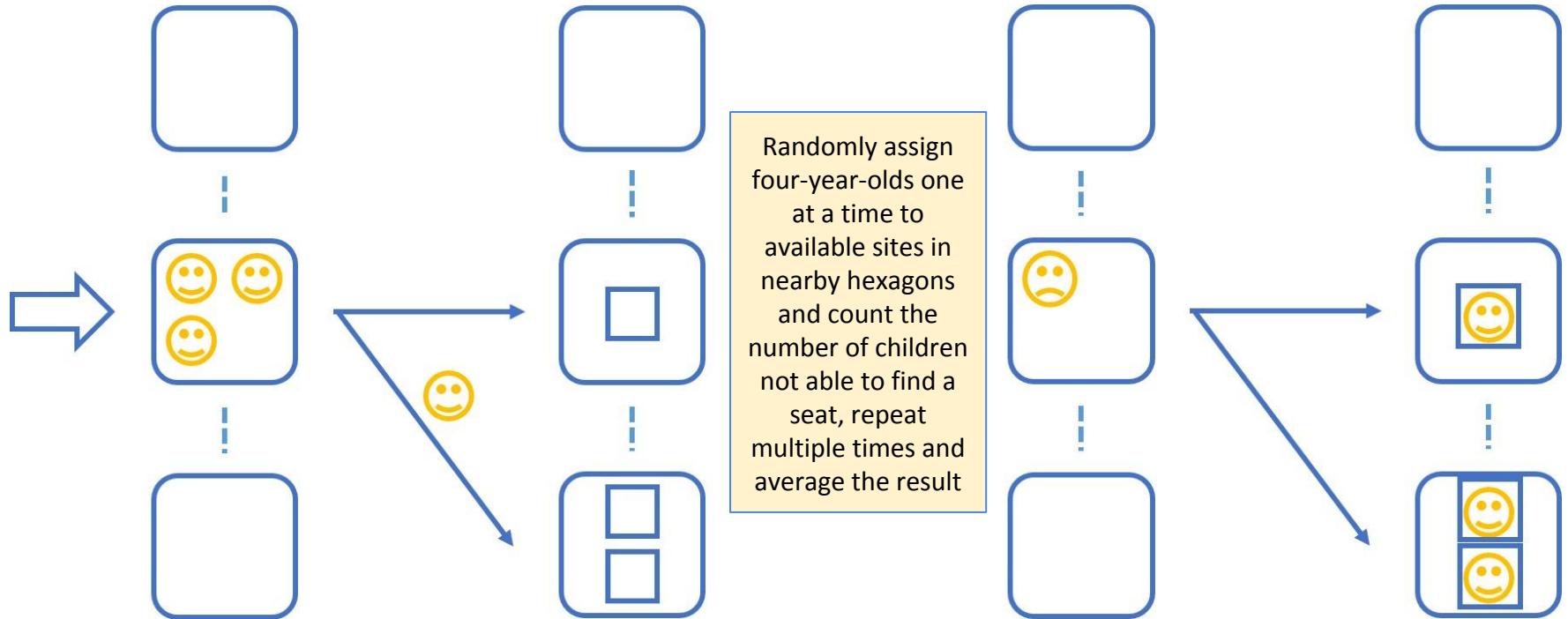
$H_i$  is usable if  $P_i > 0$  and there is some  $H_{i_j}$  which has available capacity (i.e.  $S_{i_j} < C_{i_j}$ ).

- While there exists at least one usable hexagon:
- Choose usable  $H_i$  randomly, and (non-full)  $H_{i_j}$  randomly from  $N_i$ .
- $P_i = P_i - 1$  and  $S_{i_j} = S_{i_j} + 1$ .

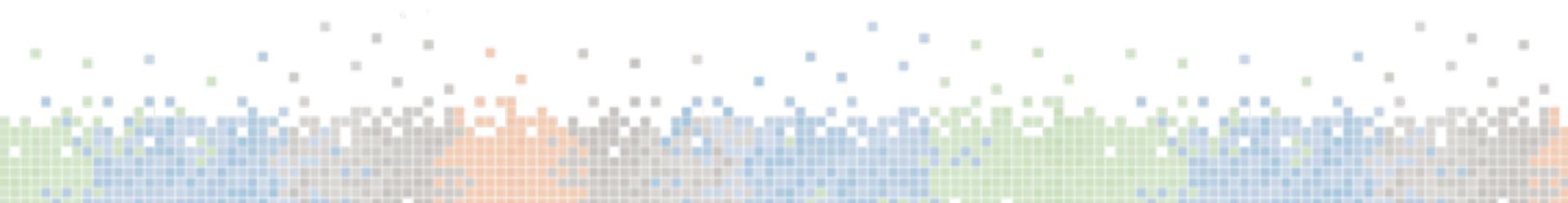
Finally, divide leftover children by initial population to give a number for each hex. Average over several runs and then map this number.

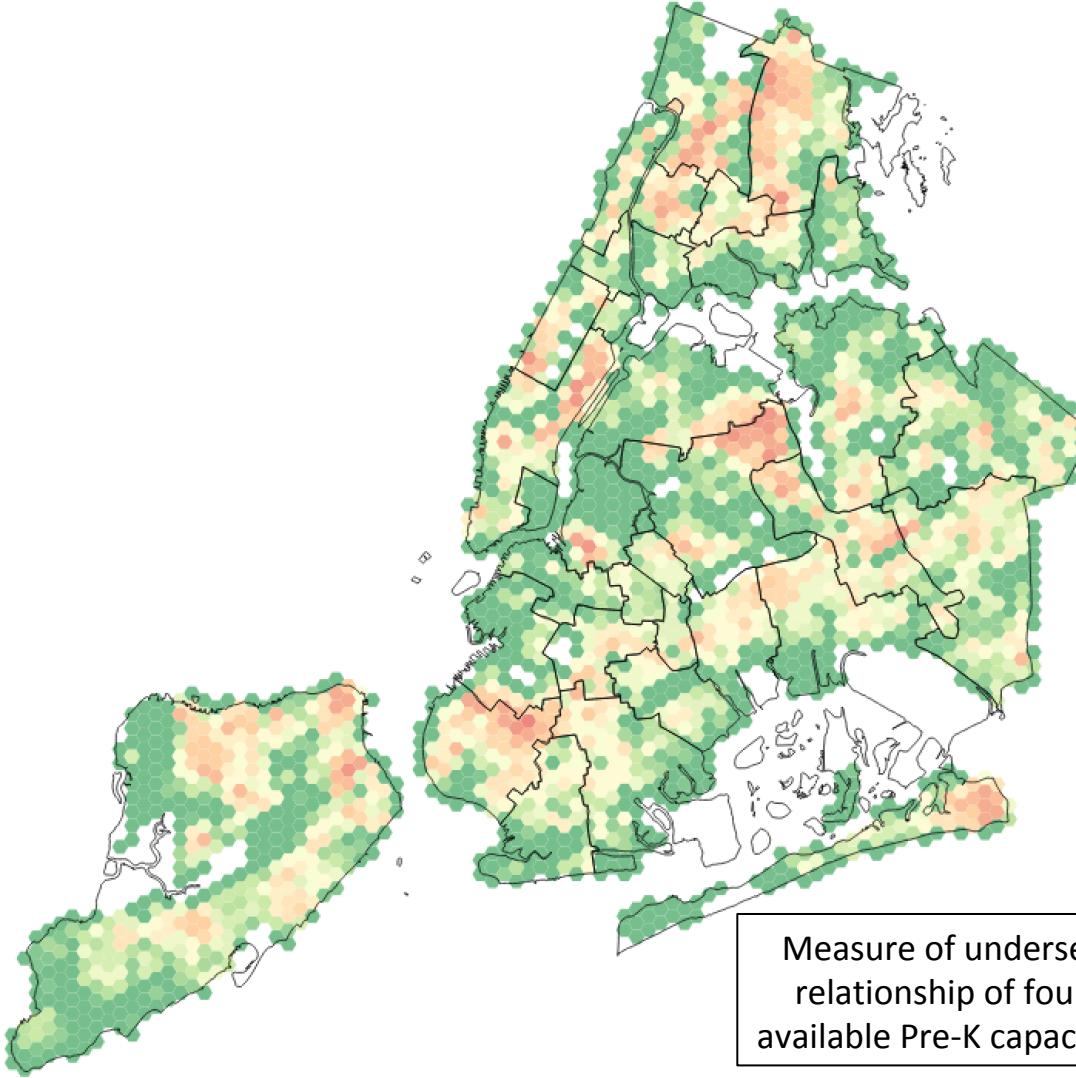
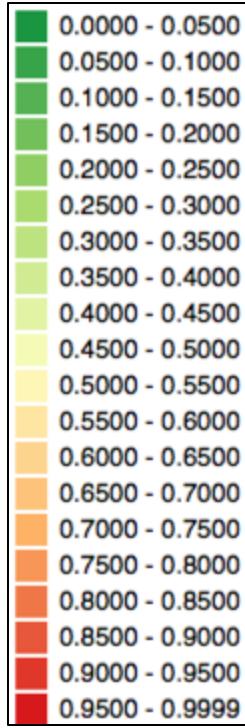


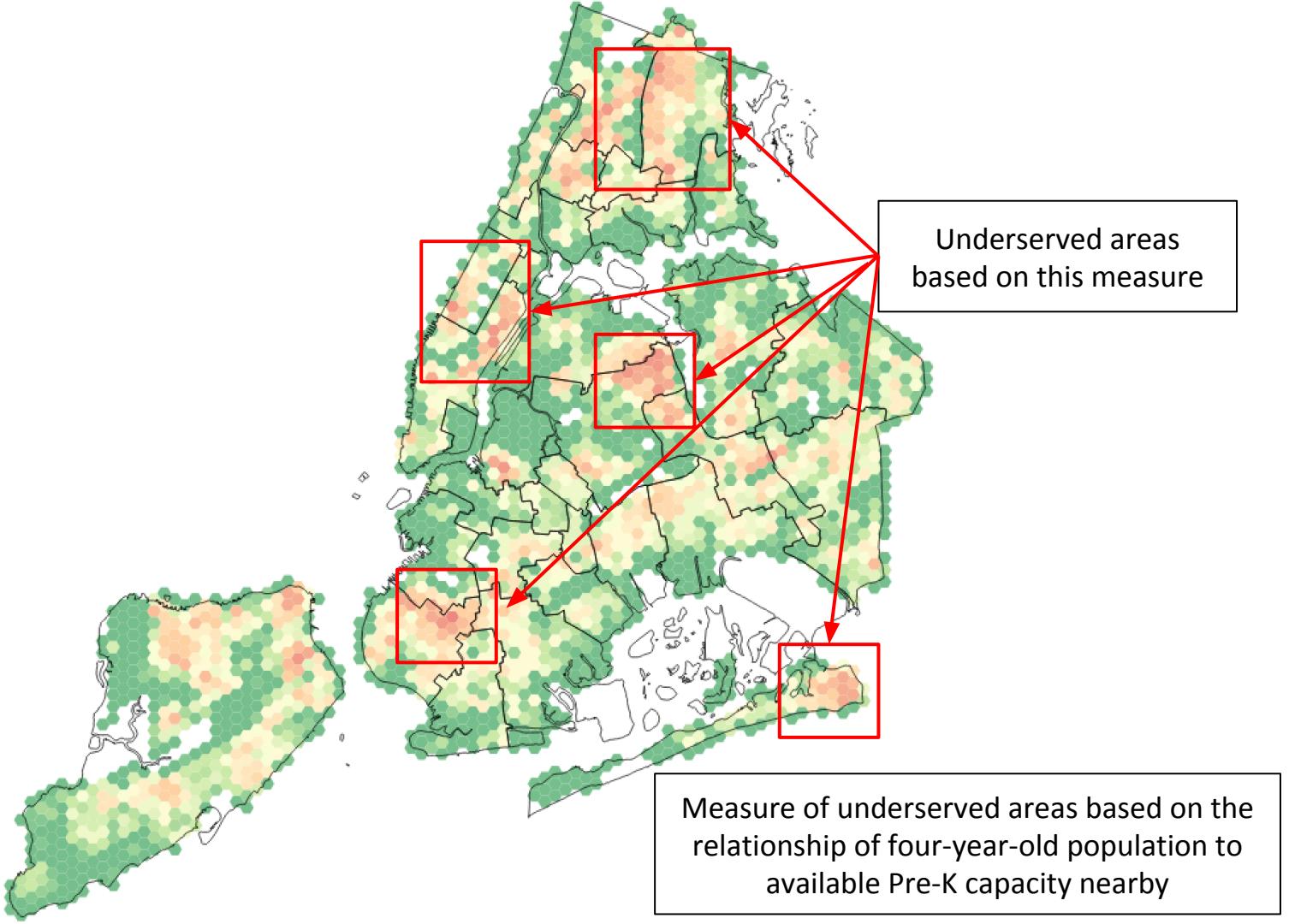
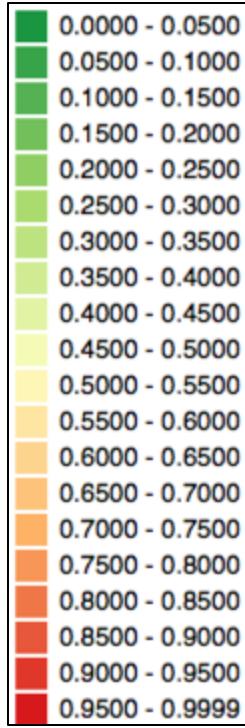
# Allocation Algorithm

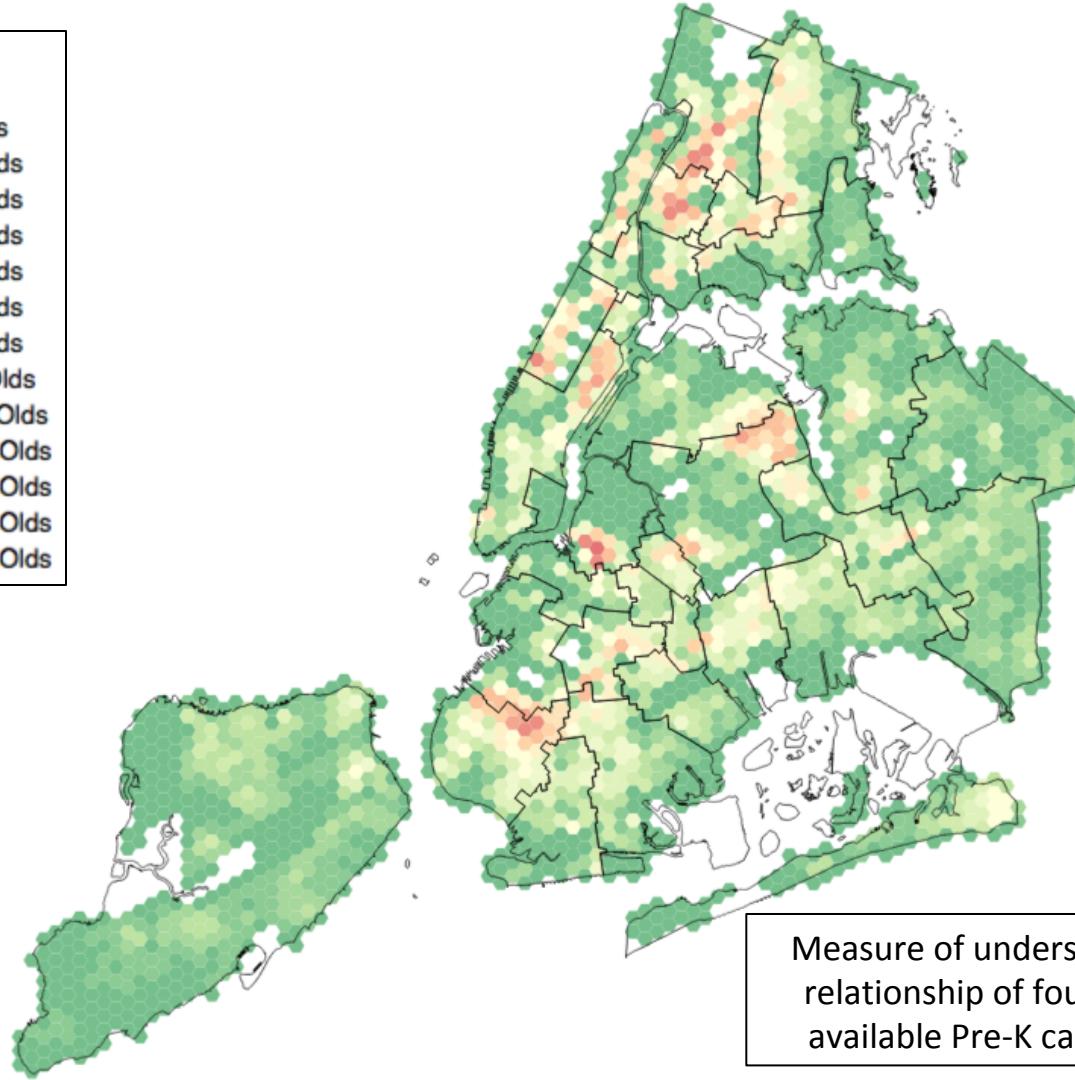
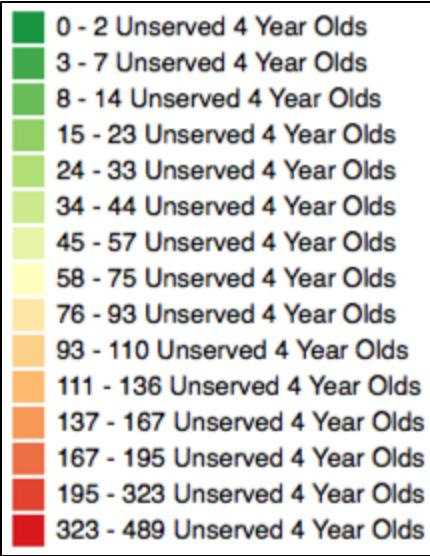


# Algorithm Result



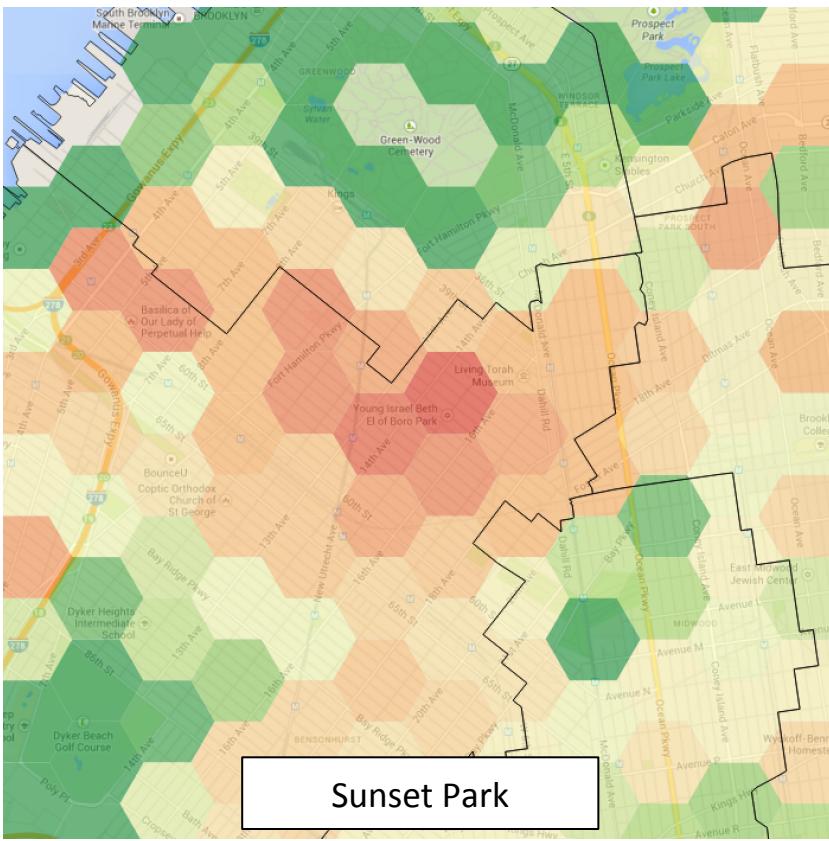




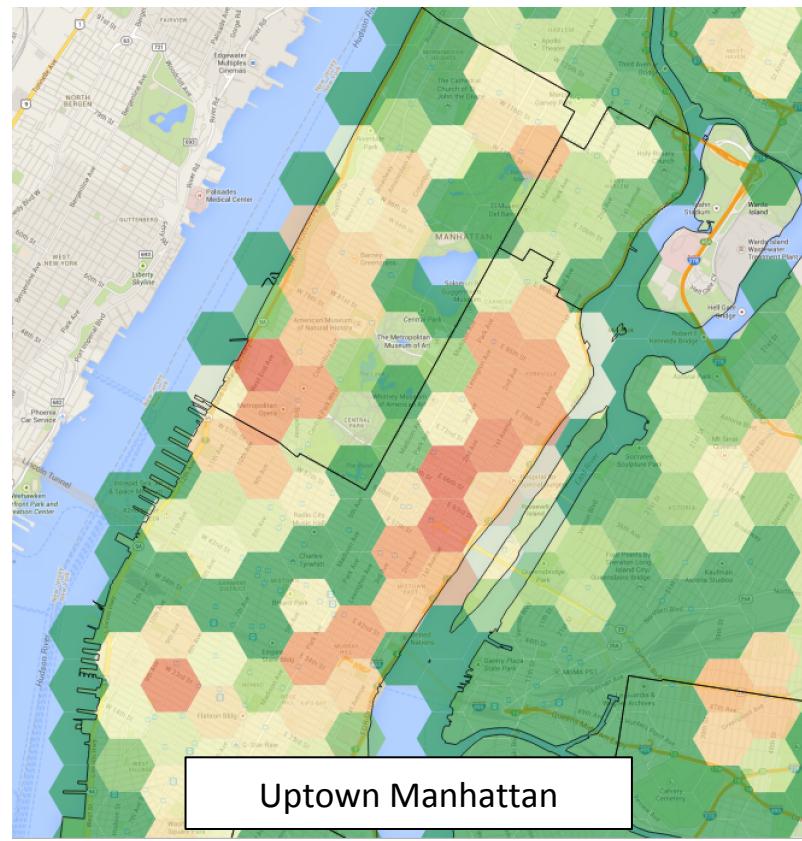


Measure of underserved areas based on the relationship of four-year-old population to available Pre-K capacity nearby (by count)

# Example of Underserved Areas

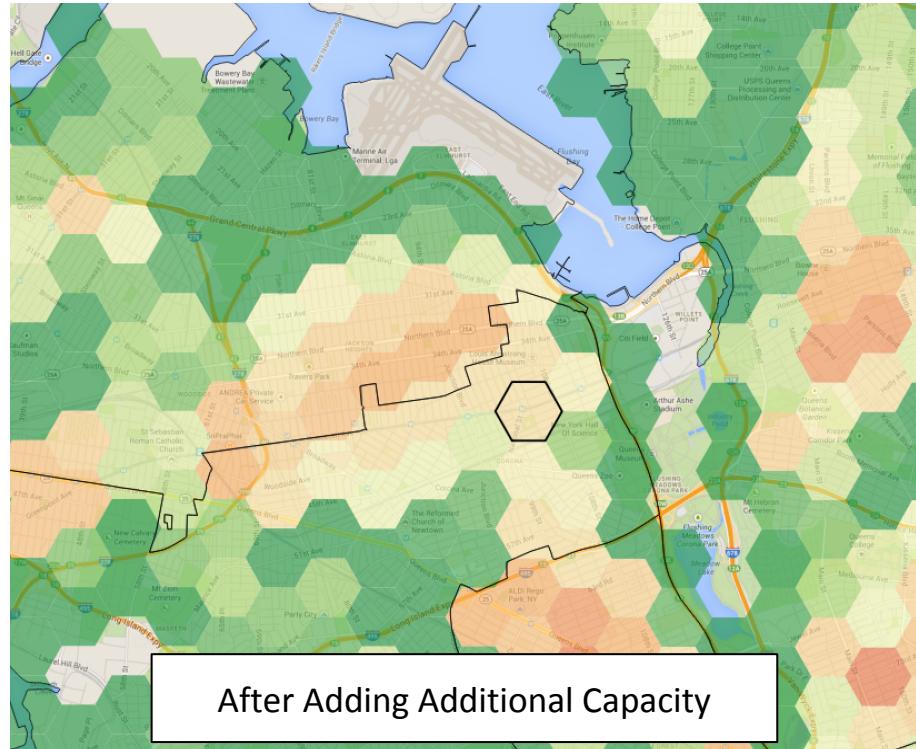
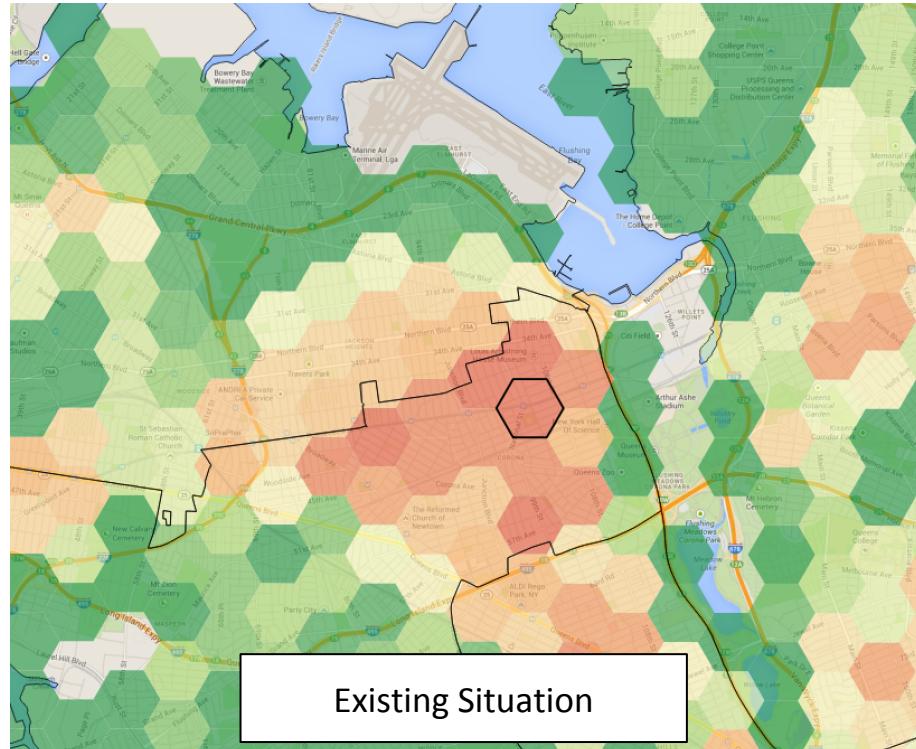


Sunset Park

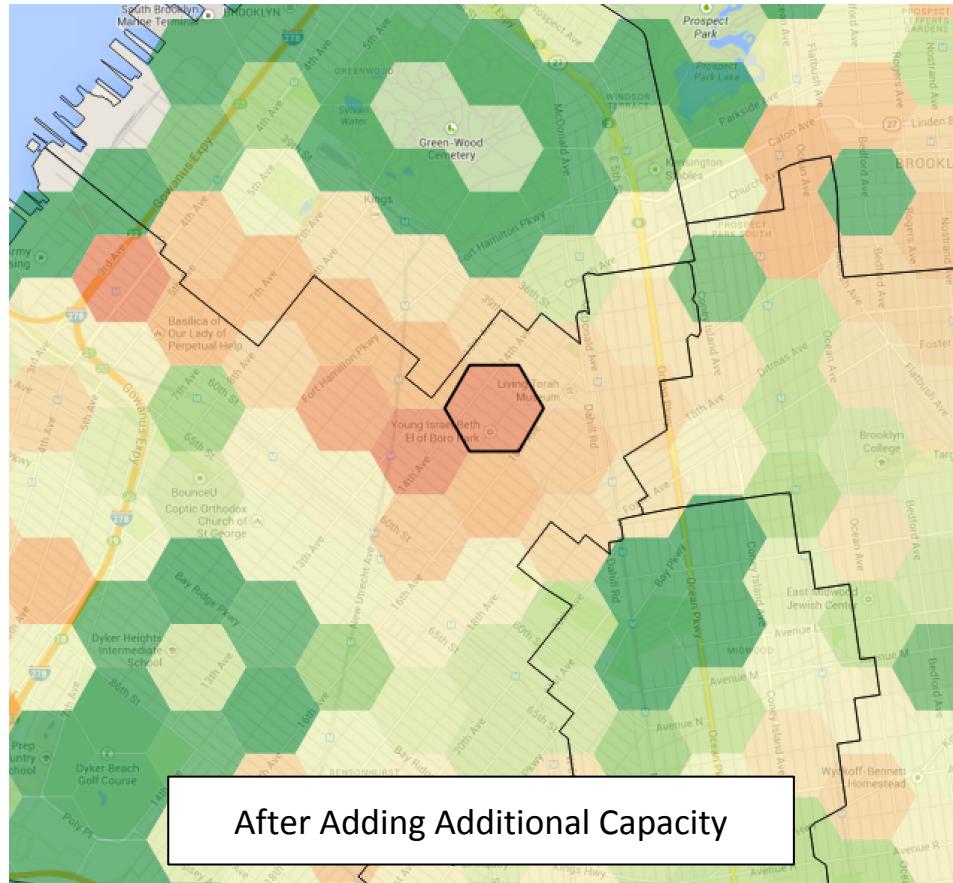
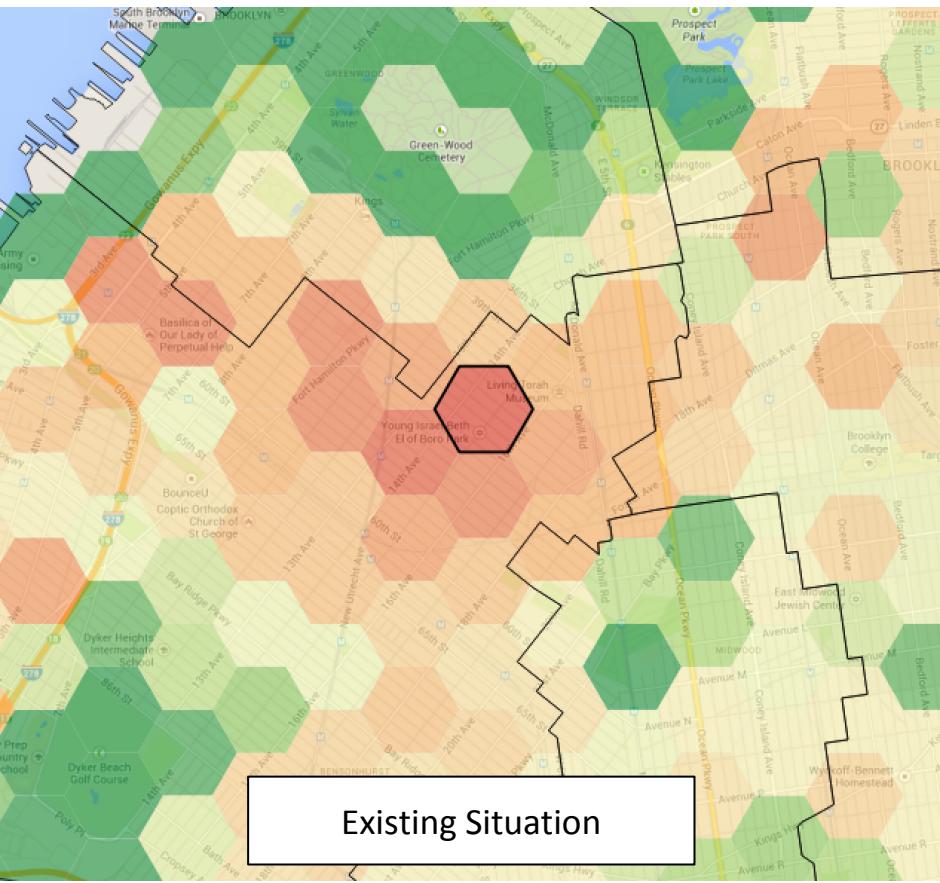


## Uptown Manhattan

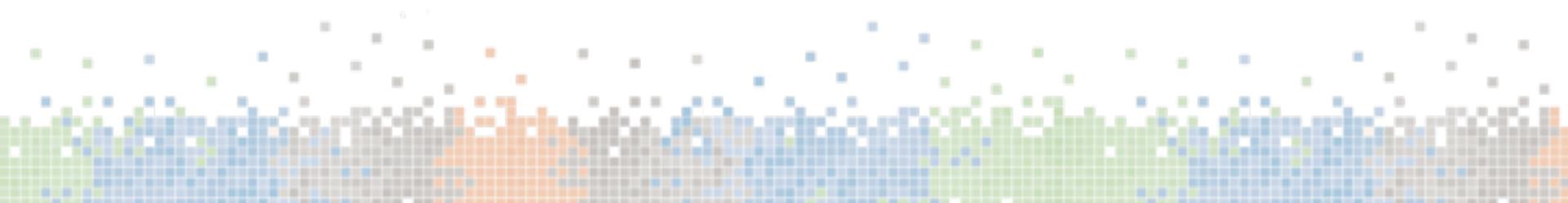
# 1000 Seats in Jackson Heights



# 1000 Seats in Sunset Park



# Challenges



# Challenges

- Missing CBO capacity data

# Challenges

- Missing CBO capacity data
- Computational complexity of allocation algorithm

# Improvements



# Improvements

- Smaller hexagons

# Improvements

- Smaller hexagons
- Better data on CBO capacity

# Improvements

- Smaller hexagons
- Better data on CBO capacity
- Use administrative records data for more precise modeling of population density

# Improvements

- Smaller hexagons
- Better data on CBO capacity
- Use administrative records data for more precise modeling of population density
- Better estimate of travel time using public transportation

# Improvements

- Smaller hexagons
- Better data on CBO capacity
- Use administrative records data for more precise modeling of population density
- Better estimate of travel time using public transportation
- Better estimate travel direction using Census Journey to Work data

# Conclusion

Adding Pre-K capacity will have a positive effect no matter how it is distributed, but optimizing the allocation of capacity to target areas with high demand (lots of four-year-olds) but low supply (number of Pre-K seats) will ensure the goals of Universal Pre-K are met by targeting capacity in areas that need it most

## Acknowledgements:

Dr. Sharad Goel, Microsoft  
Research NYC

Dr. Manuela Veloso, Carnegie  
Mellon University

Dr. Theodoros Damoulas, NYU  
CUSP

Dr. Constantine Kontokosta,  
NYU CUSP

Download the paper:

[http://bit.ly/bduic\\_nyc\\_upk](http://bit.ly/bduic_nyc_upk)

Corresponding authors:

Richard Dunks ([richard.  
dunks@nyu.edu](mailto:richard.dunks@nyu.edu))

Ravi Shroff ([ravi.shroff@nyu.  
edu](mailto:ravi.shroff@nyu.edu))