



# NYC Taxi Data Exercise

## Modeling Airport Pickups and Drop-offs

June 1 2018

Dan Finkel

[Dan.Finkel@gmail.com](mailto:Dan.Finkel@gmail.com)

617.640.9968

<https://www.github.com/danfinkel/>

# Executive Summary

## Data Exploration

- Geospatial taxi behaviors well captured in public NYC-provided datasets
- Datasets highlight similarities and differences in green and yellow taxi behavior

## Airport Taxi Queue Analysis

- Airport taxi arrivals and departures fit traditional Poisson-based queuing models
- Modeling can be used to characterize idle times, manage parking space and improve customer wait times

## Next Steps

- Additional information about taxi trip (e.g., car ID, coordinate info) can improve model assumptions
- Additional model complexity can be added to address more challenging questions and airport routing schemes

## Problem

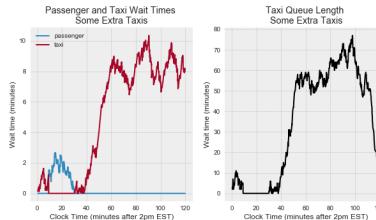


## Dataset

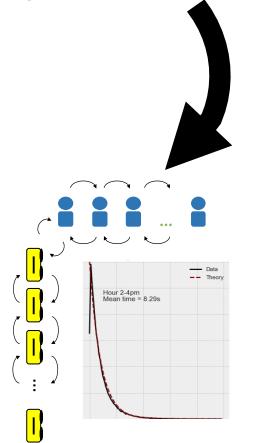
**NYC** Taxi & Limousine Commission



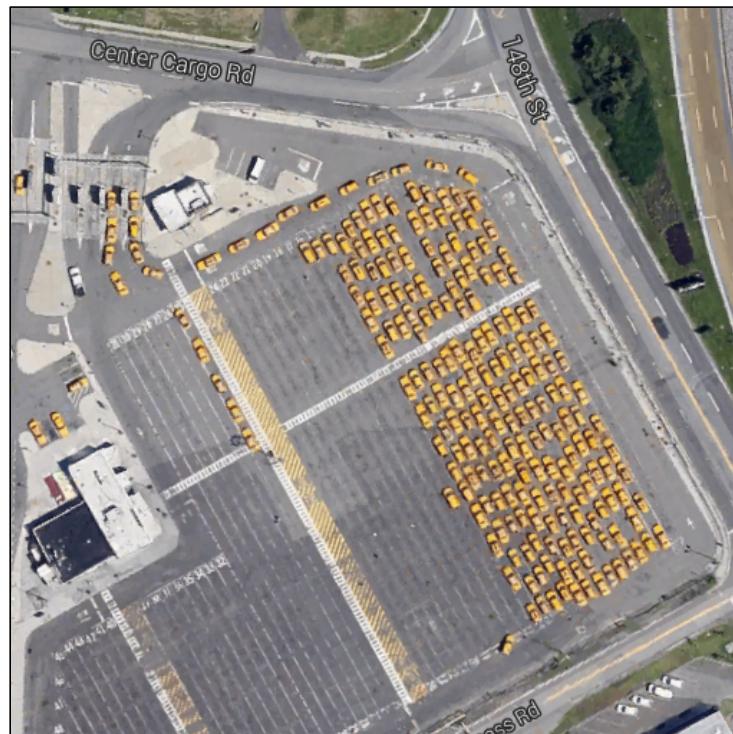
## Product



## Model



# Problem Description



Taxi Parking Lot at LaGuardia Airport NYC

## Questions (Policy-Motivated)

- How is taxi queuing management aligned with customer demand?
- Are there opportunities to improve efficiency?

## Analytic Approach

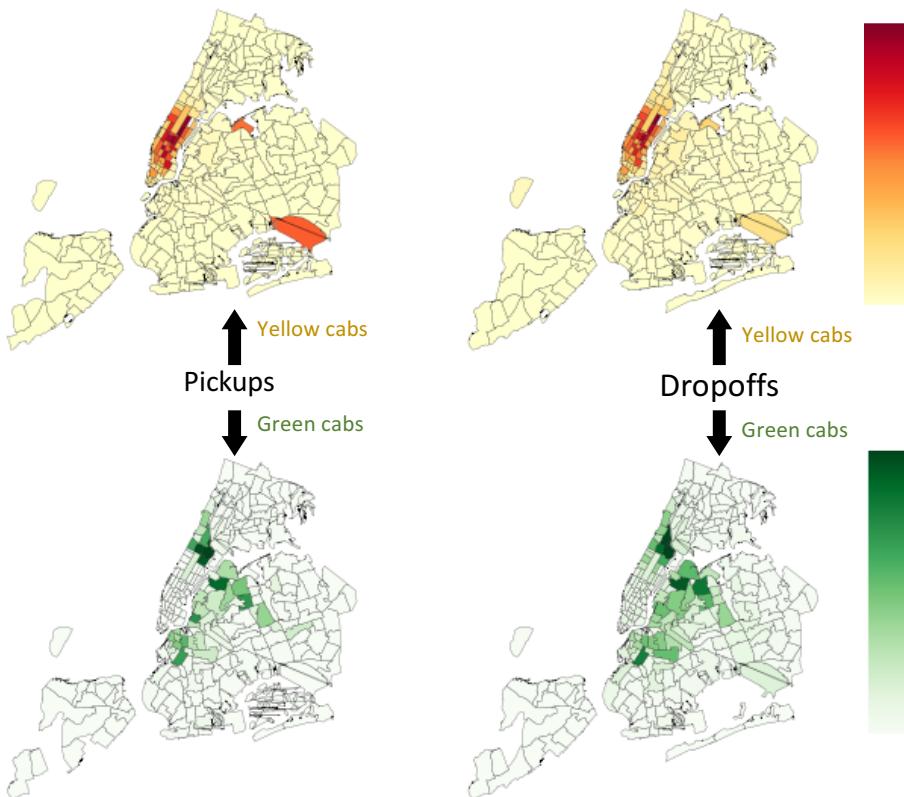
- Utilize NYC taxi data (Jan – Mar 2017)
- Develop simple queuing model and validate with data
- Utilize Monte Carlo simulations to characterize queue behaviors for different parameter settings

## The Washington Post

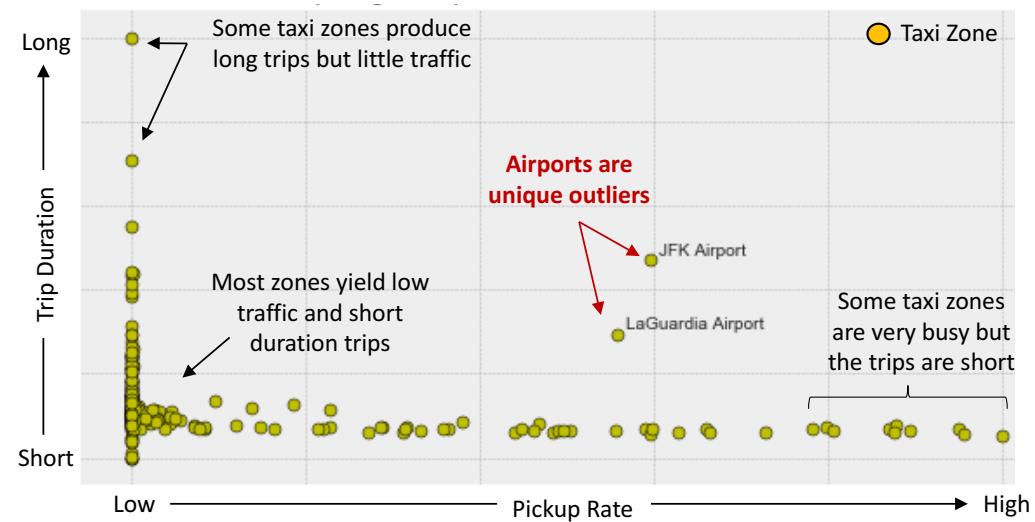
"The impact of ride-booking services on the taxi industry has been well-documented, but it is the demand for parking that may be of most concern for airport officials since it is one of the airports' biggest moneymakers." – Sept 25 2017 (url in notes)

# Data Overview

Popular NYC Taxi Drop-off and Pick-up Locations (Jan 2017)



Comparing NYC Taxi Zones by Volume and Mean Trip Duration

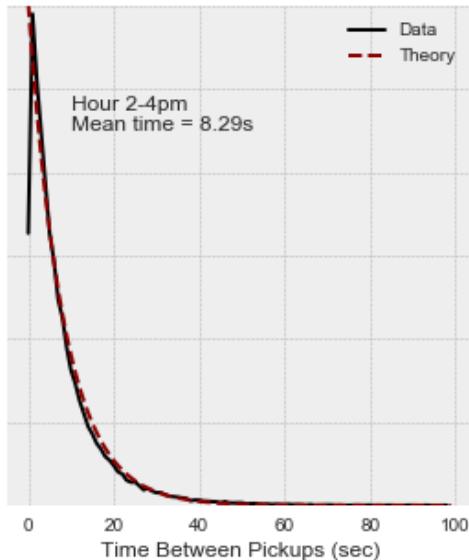


NYC Airports are Unique Locations for Taxi Traffic

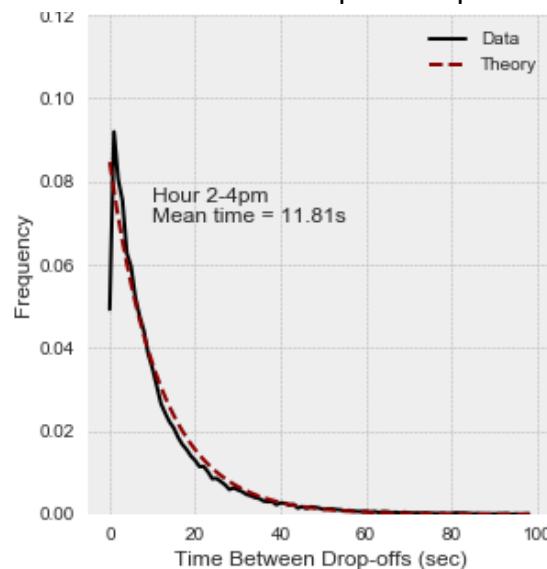
# Model Overview

## Comparing Open Source Data to Theoretical Models

Time Between Airport Pickups



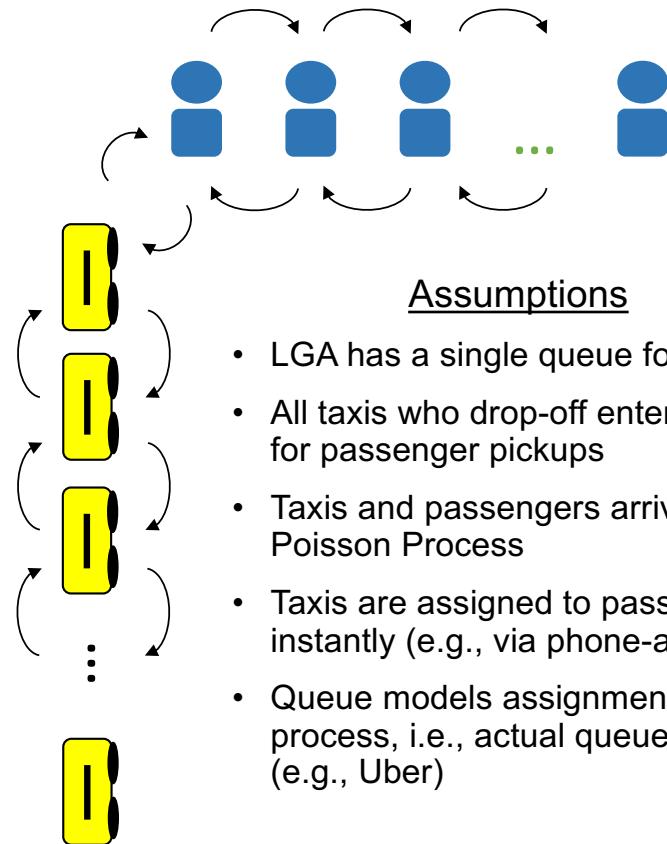
Time Between Airport Drop-offs



Taxi Customer Pickups and Drop-offs Strongly  
Fit\* a Poisson Counting Process Model

\* Model parameters vary over the course of the day and model assumptions degrade during early am hours

## Simple Queuing Model



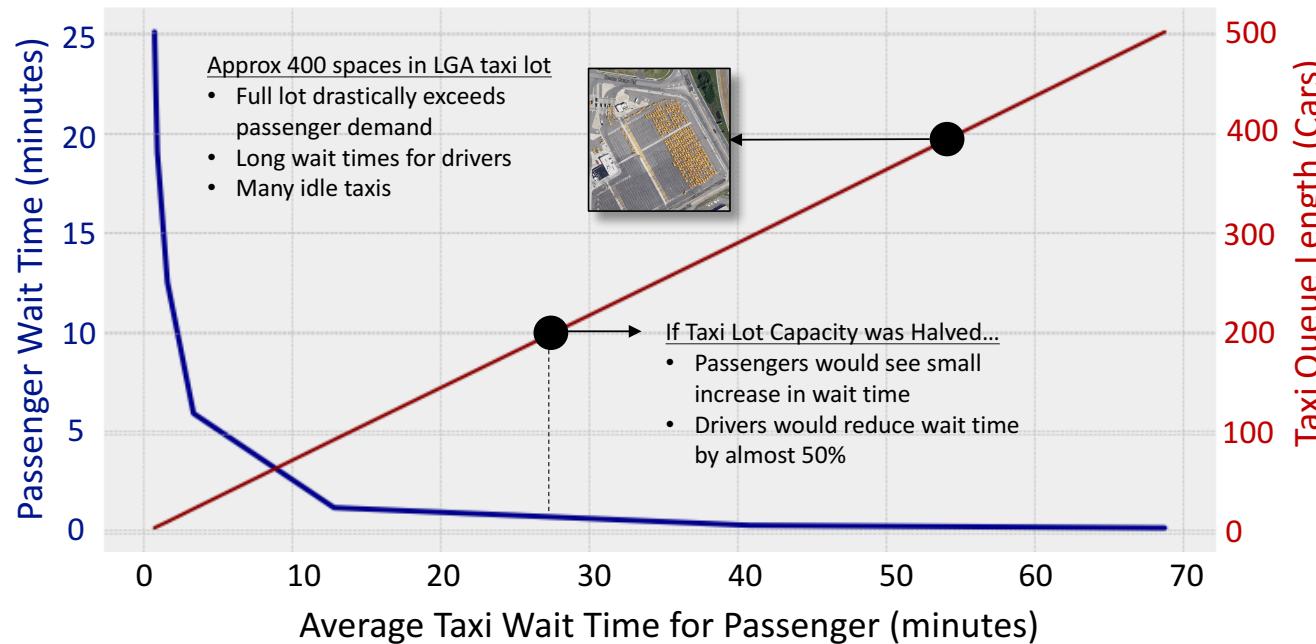
## Assumptions

- LGA has a single queue for taxis
- All taxis who drop-off enter queue for passenger pickups
- Taxis and passengers arrive via Poisson Process
- Taxis are assigned to passengers instantly (e.g., via phone-app)
- Queue models assignment process, i.e., actual queue is virtual (e.g., Uber)

# Results and Next Steps

## PRELIMINARY

Simulation Results – How Much Queue Space is Needed for Taxis?



### Key Takeaways

- Current taxi lot size is too big and may lead to large taxi queues
- Lot could be shrunk to reduce idle taxis and reallocate space
- Results are preliminary and additional fidelity is needed to gain confidence in results

### Next Steps

- Increase complexity of model to gain confidence in results
- Define app-based taxi-passenger assignment
- Add key information to data requests
  - Break airport into terminals
  - Label each taxi ride with car and/or driver ID