

# DATA 698 Project v1

## Predicting Citi Bike Availability in NYC

Philip Tanofsky

2022-11-19

# Problem and Objective (1)

Introduce yourselves and describe your problem. Explain your objectives, challenges of your work, proposed methodologies, and the assumptions you made while conducting modeling and/or analysis. Provide an overview of your approach and/or conceptual model (please do not present your code directly). Describe the results you obtain and summarize the current achievements and possibility of future works.

# Challenges and Assumptions (2)

- Too much data
  - Over 3.5 million trips in month of Sept. 2022
- Rebalancing identification
- User friendly approach
  - Inputs to output

## Data Collection (3)

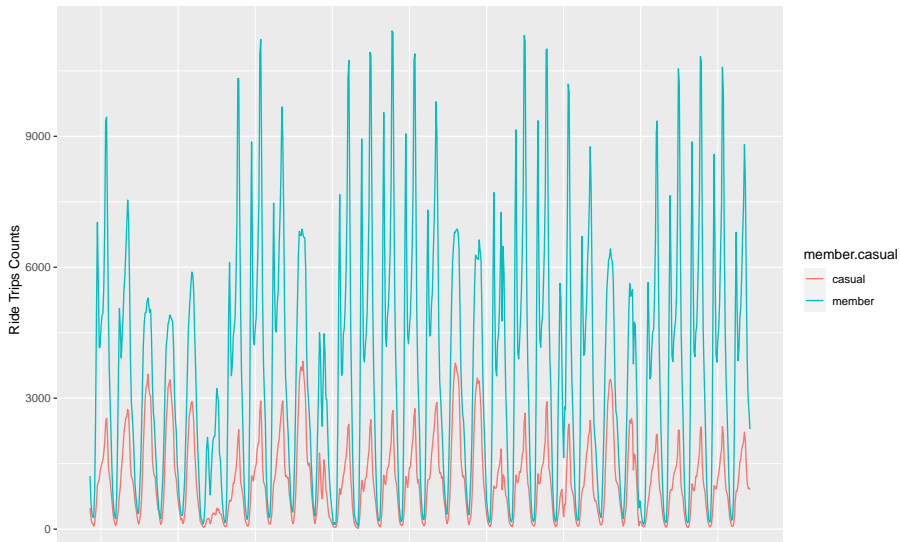
- Creating timetable of surplus/shortage

# Data Analysis (4)

# EDA 1 (5)

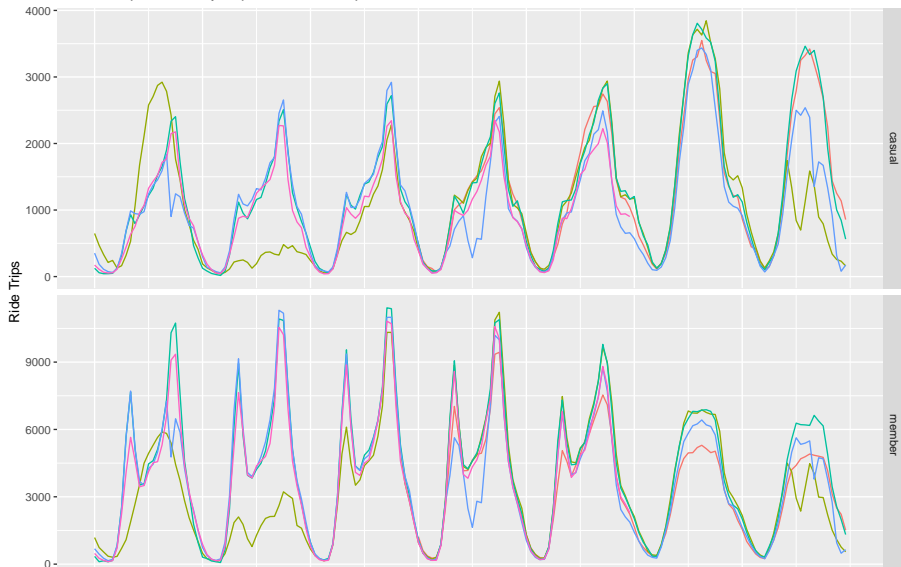
Ride Trips by Hour – Sept. 2022

Citi Bike NYC

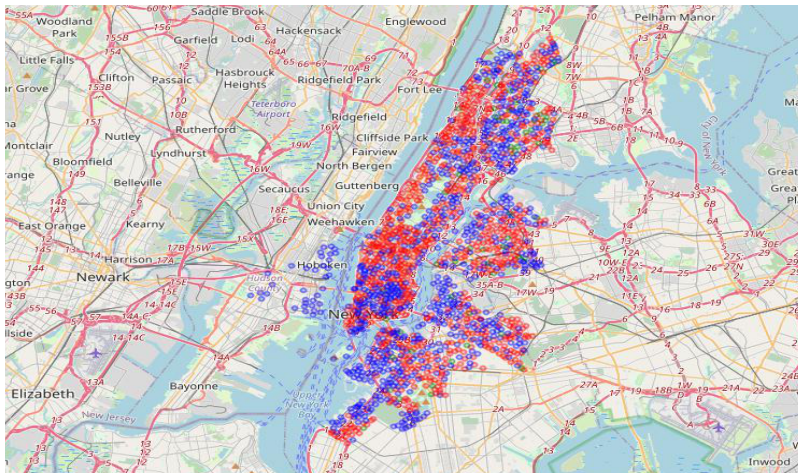


# EDA 2 (6)

Seasonal plot: Weekly Trip Counts for Sept. 2022



## Overall Monthly Surplus by Docking Station





## EDA 4 (8)

## focus on Brooklyn Areas (9)

# Proposed Methodologies (10)

- Inputs
  - Latitude and longitude
  - Day of the Week
  - Time of Day
- Citi Bike offers live map of availability
- Lyft provides real-time availability
- Time series model
- Poisson distribution and Negative Binomial given the over-dispersion

# Overview of Approach (11)

- Data from API call every 15 minutes for two weeks
  - Citi bike availability at each station
  - Two weeks is small interval to predict
    - Valid limitation of model

# Conceptual Model: Step 1 (12)

Clustering

# Conceptual Model: Step 2 (13)

Modeling

# Model Results 1 (14)

- Certain input to model based on 3-hour intervals
- Results table

## Model Results 2 (15)

- Certain input to model based on 15-minute intervals
- Results table



# Model Results Visual (16)

# Prediction Function 17

- Clustering
- Model prediction

# Current Achievements 18

# Future Works 19

- Weather . . . actually, can I predict weather? would that really work?
- Subway stations: Citi Bike offers valet
- Model of all NYC
- Real-time clustering would be better

# Slide 20

# Slide 21