

Rebalancing citi**bike**

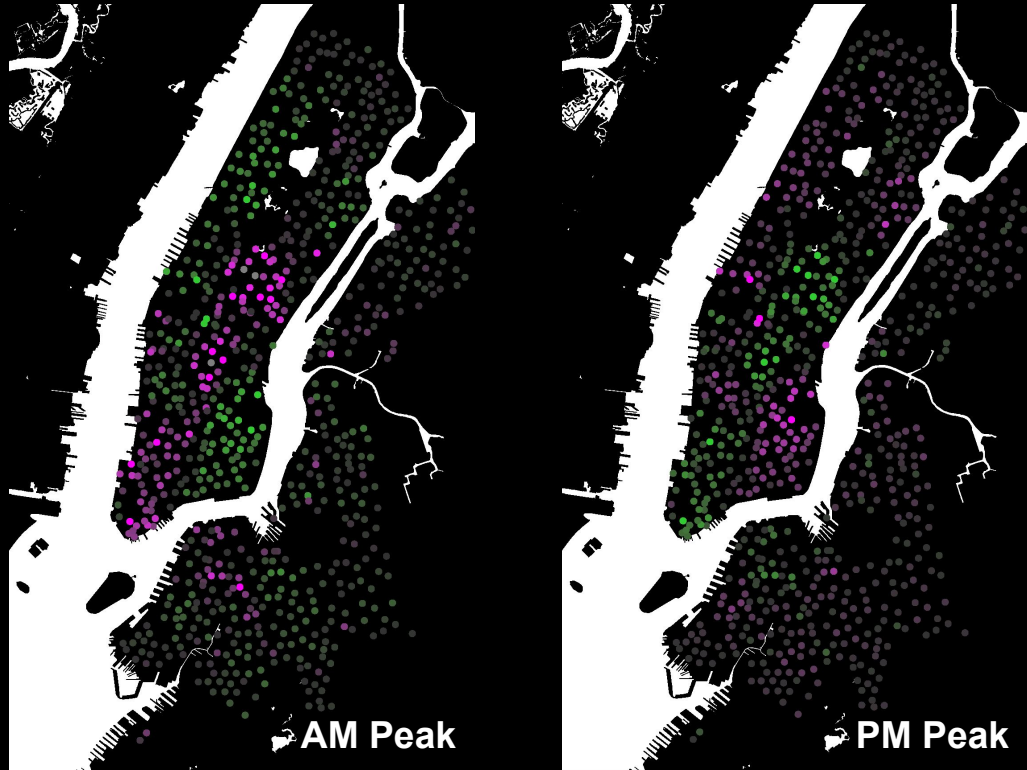
the country's largest
bicycle sharing network.

Nikhil Tangirala
Jonathan Yuan

813 stations (741 studied)
40,000 bikes planned for 2019
1.8+ million trips per month



Asymmetrical demand for bikes.



There is **spatial and temporal asymmetry**, especially across commercial and residential areas. It's at its **worst during rush hour**.

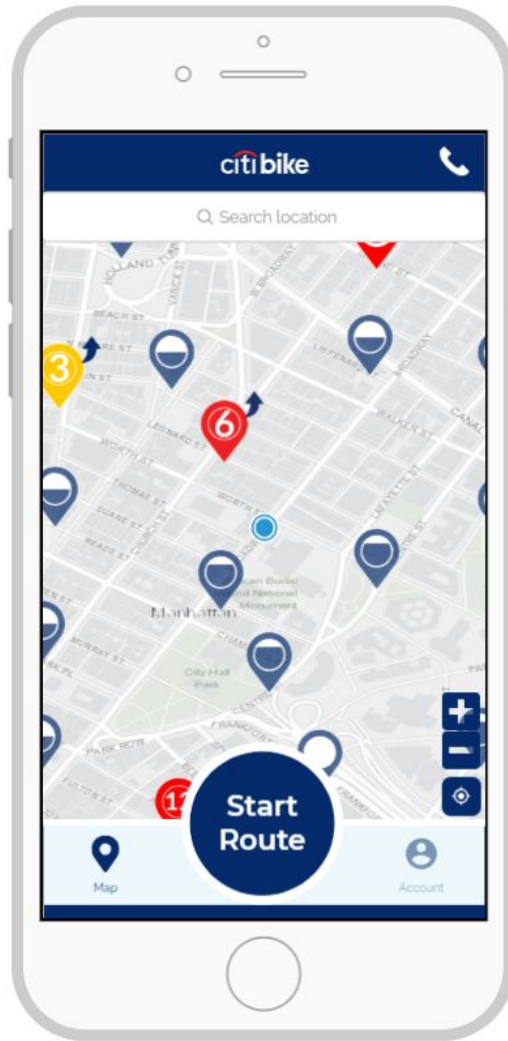


Excess Bikes

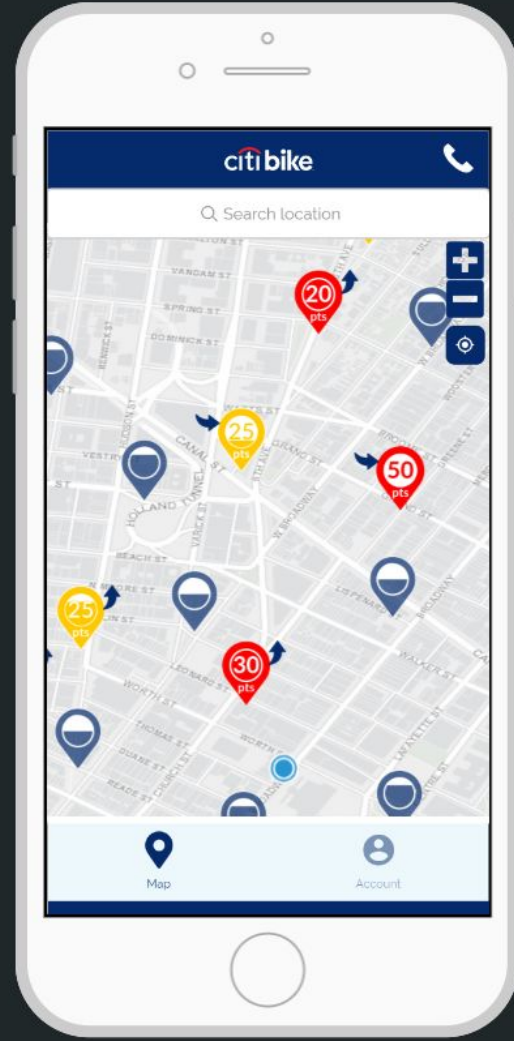
Excess Docks

**One Citi Bike
app to rule
them all—**
Functionality
for operators
and members.

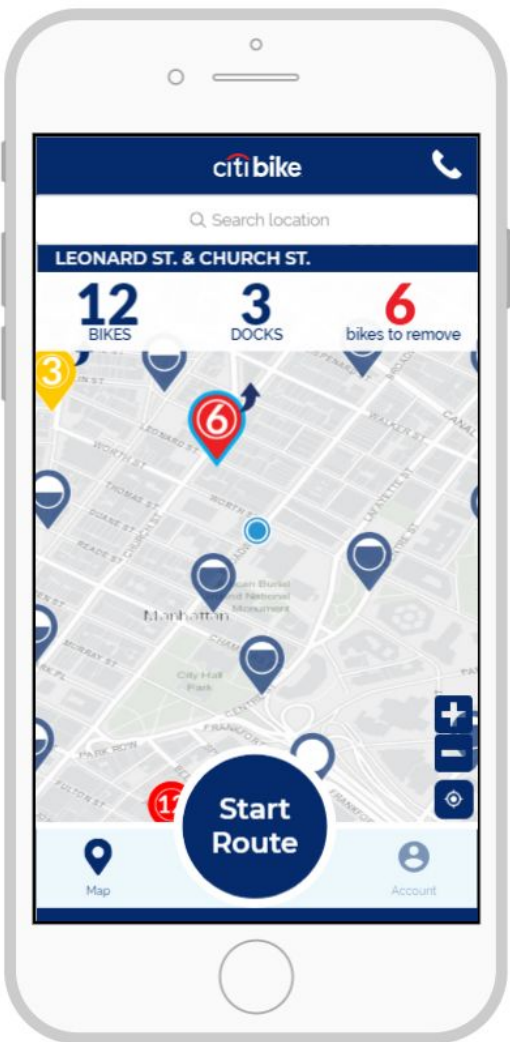
Operator View



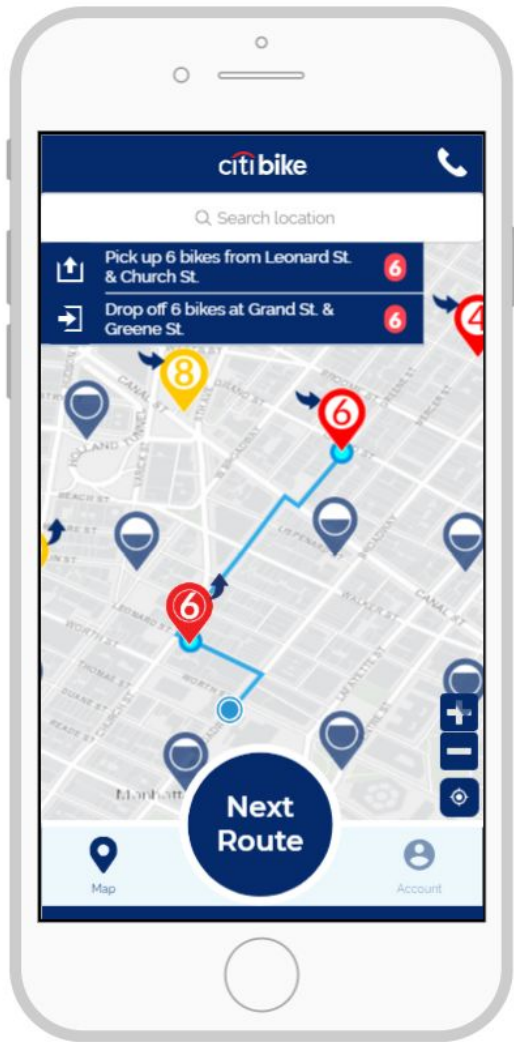
User View



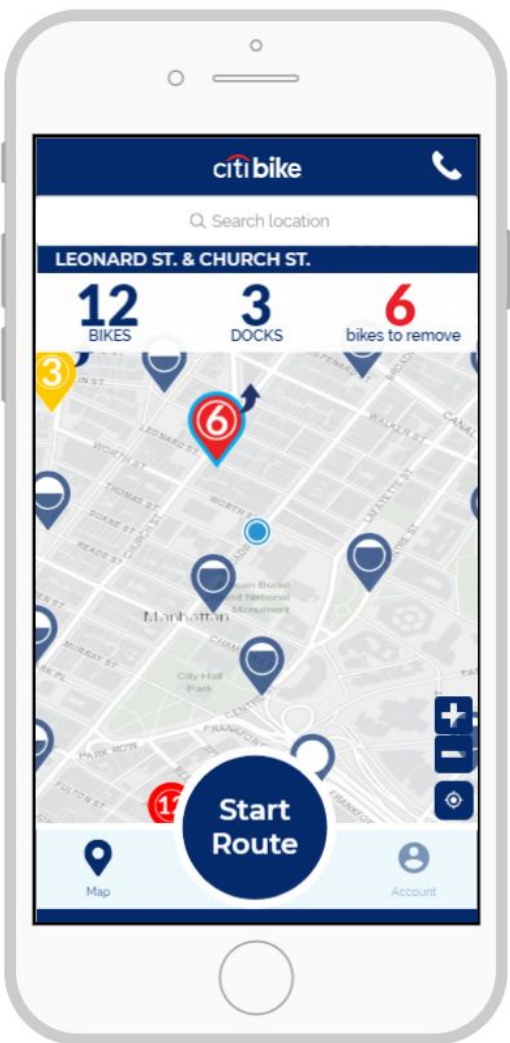
Operator View



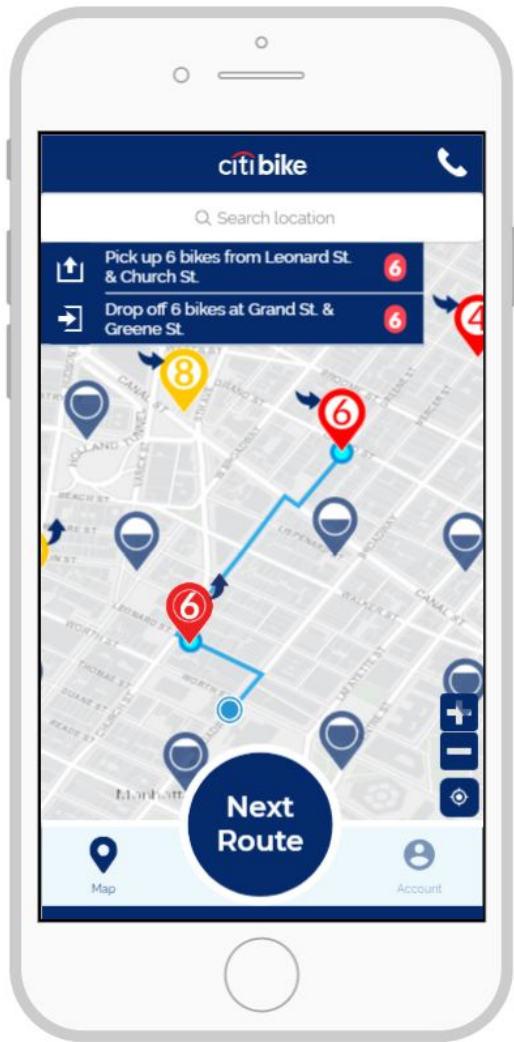
Operator View



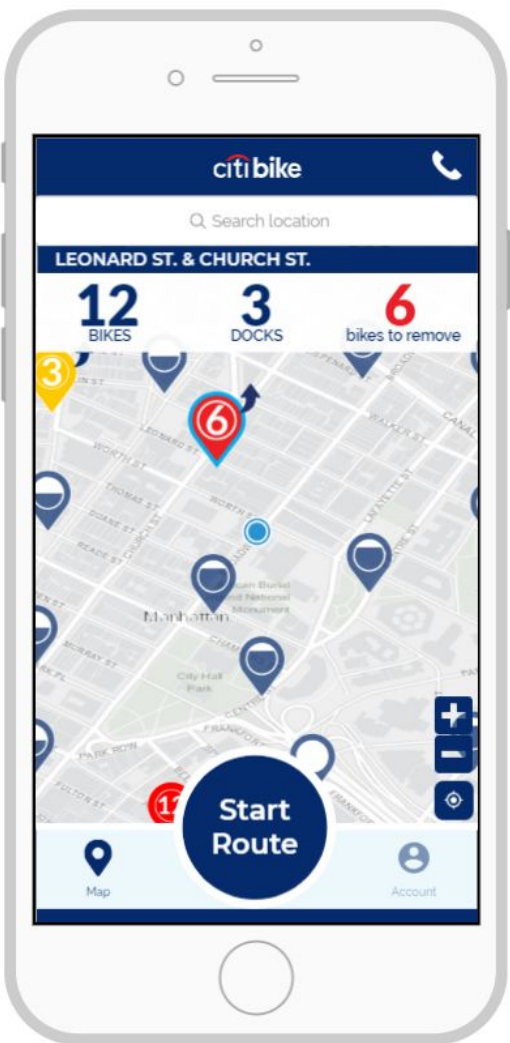
Operator View



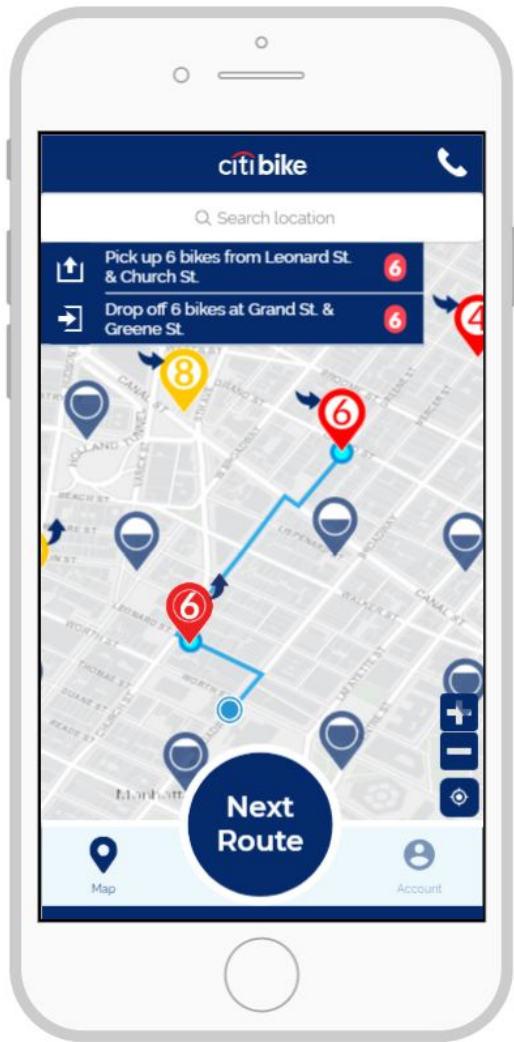
Operator View



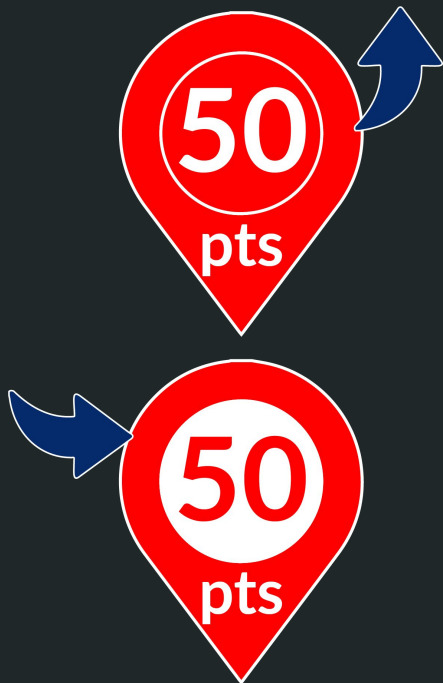
Operator View



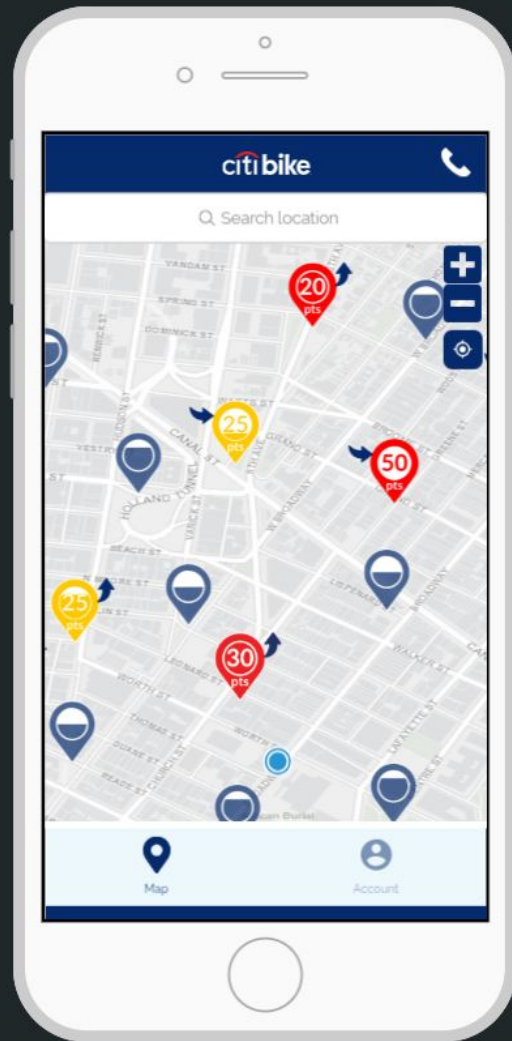
Operator View



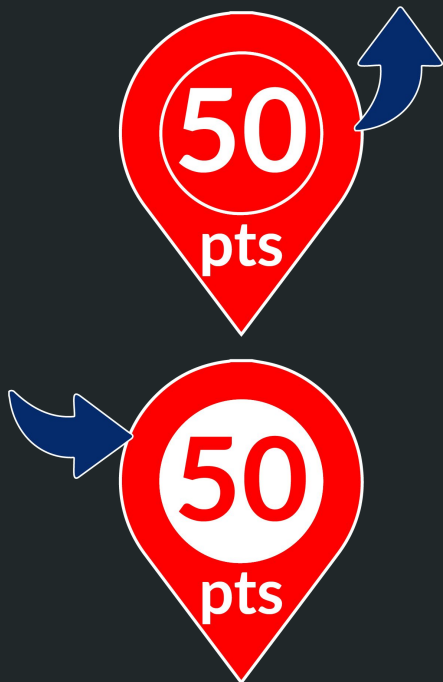
Annual members
collect points for
perks.



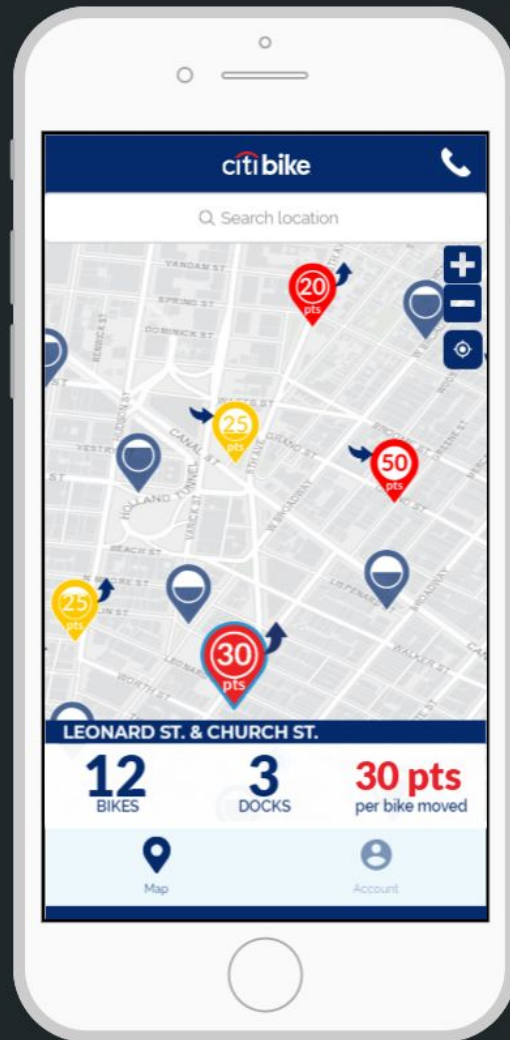
User View



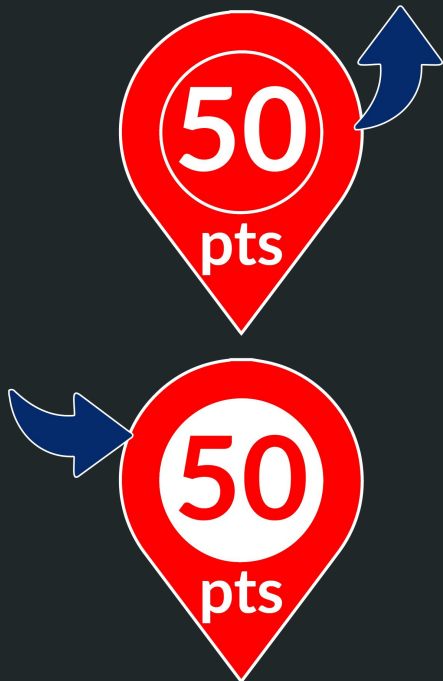
Annual members
collect points for
perks.



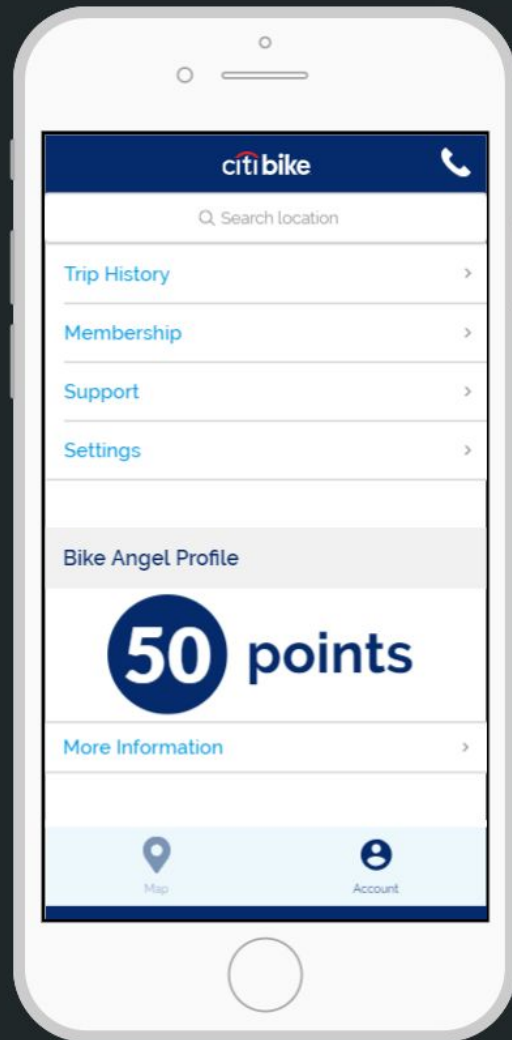
User View



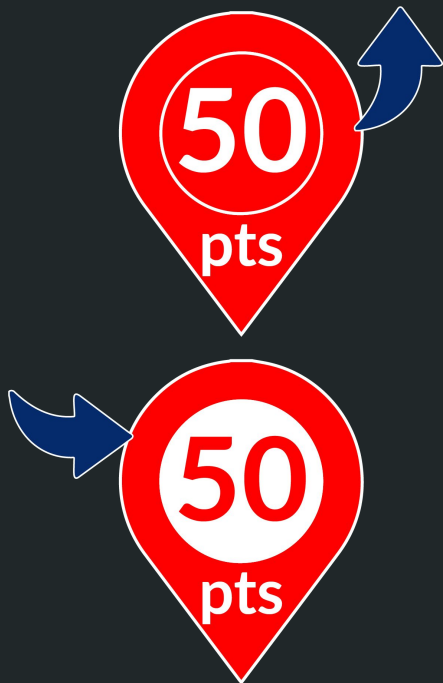
Annual members
collect points for
perks.



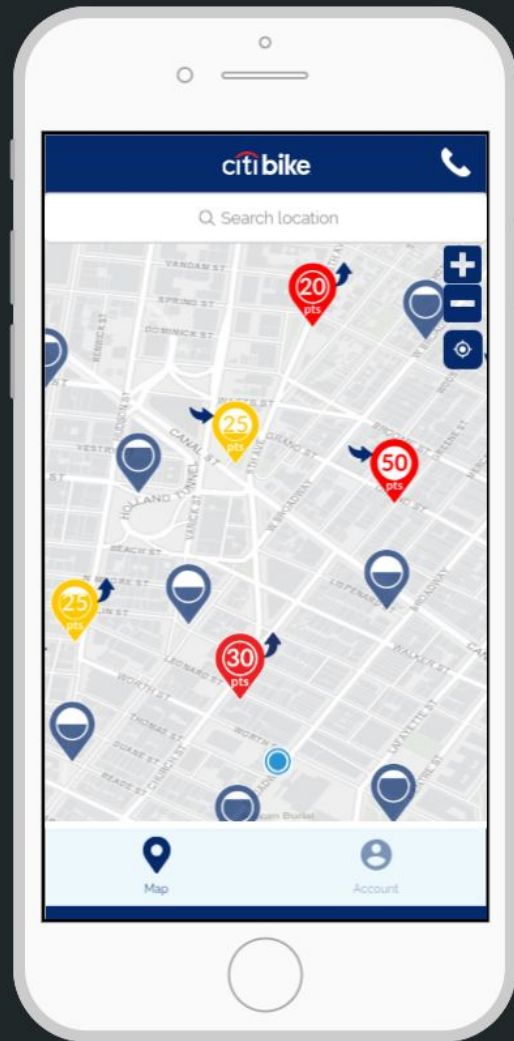
User View



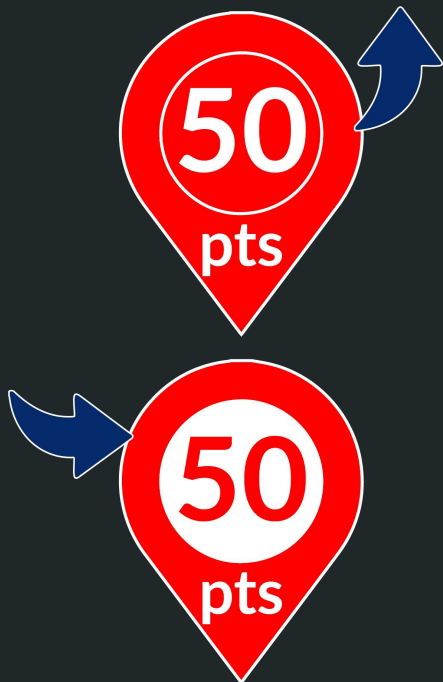
Annual members
collect points for
perks.



User View



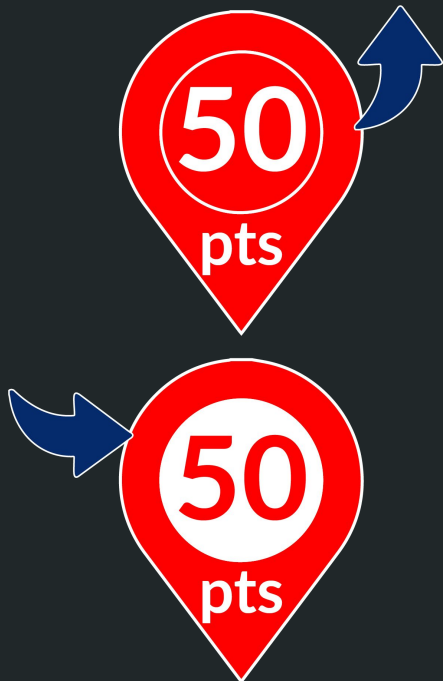
Annual members
collect points for
perks.



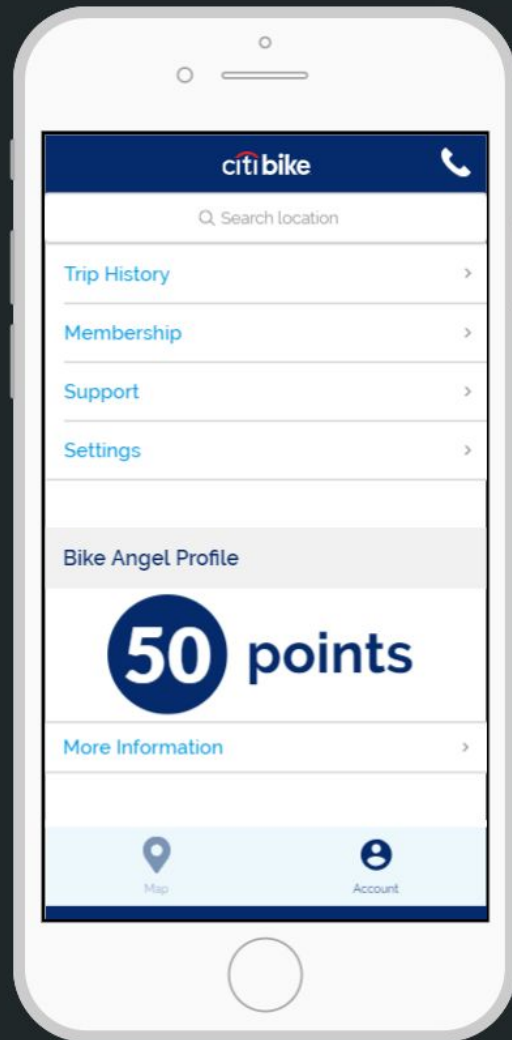
User View



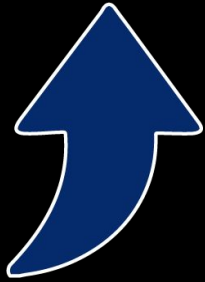
Annual members
collect points for
perks.



User View



Rebalance where supply doesn't meet demand.



Supply vs. Demand

Are there too many bikes or too many docks?



Rebalance Number

How many bikes should the operator move?



Priority Score

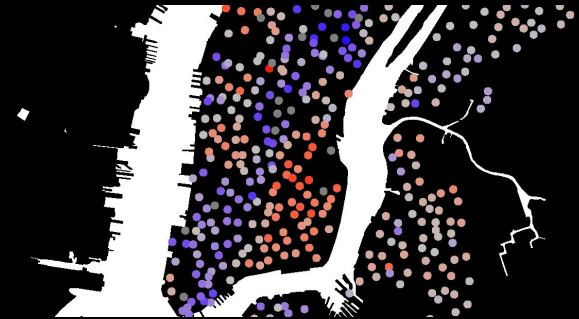
Does this station need bikes immediately?

Predict bike counts using k -fold Poisson regression.



Predictive Factors

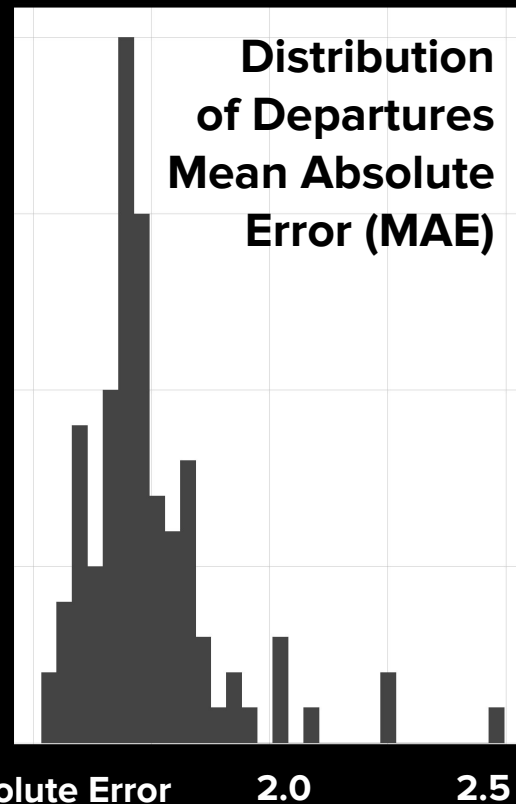
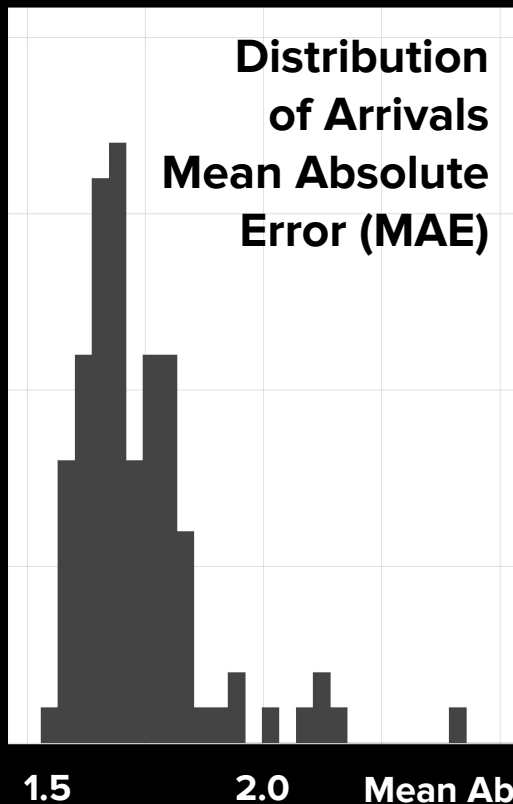
- space and time lag
- jobs and residences
- weather
- physical features



Are the models generalizable over 100 folds?

100-fold
Poisson
regression

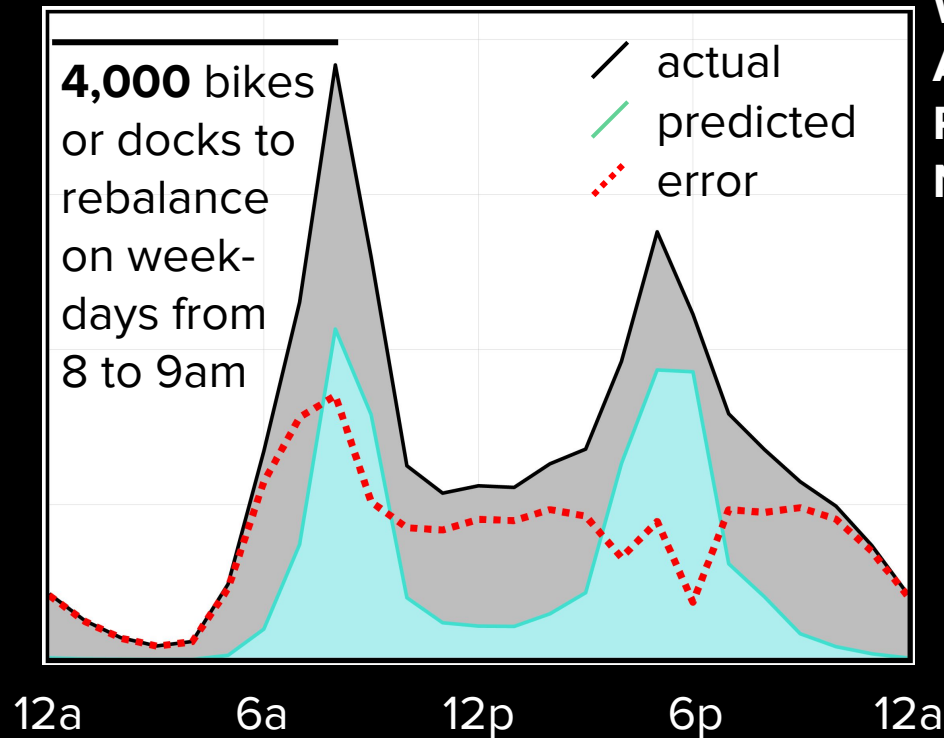
R-squared: 0.53
MAE: 1.74
RMSE: 4.62



100-fold
Poisson
regression

R-squared: 0.52
MAE: 1.74
RMSE: 4.62

But, inconsistencies across space and time indicate an incomplete model.



**Weekday
Aggregate
Rebalance
Numbers**

Higher RMSE
(root mean
square error)
at high-volume
stations near
midtown
Manhattan.



Next steps to improve the model.

- Use **additional predictive factors**
- Build an alternative model **directly predicting excess** between demand and supply of bikes
- Take into account **current available bikes or docks**

