Visualizing and Predicting Ridership

It can be tough to find a Bikeshare bike when you really need one.

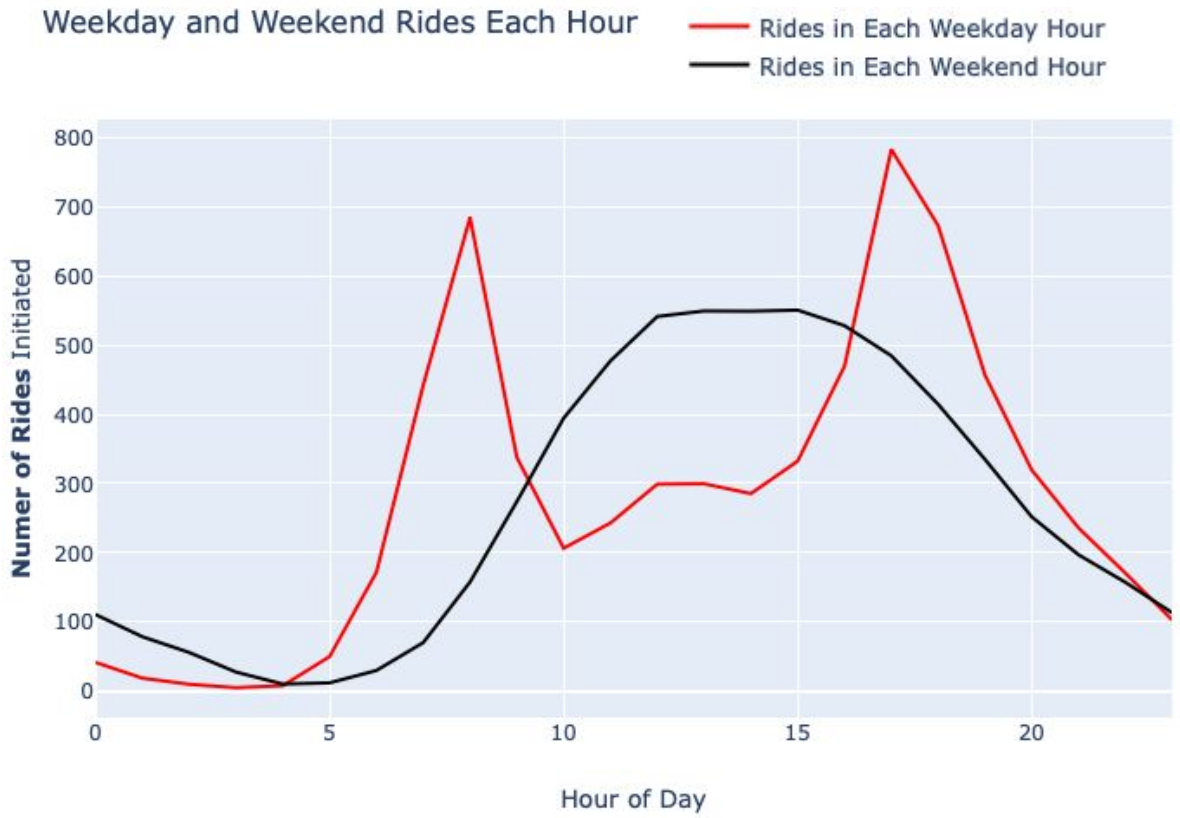
Fortunately, there is plenty of data out there to help us understand the issue.

- Trip data
  - Ten years of rides
- Dock data
  - Three weeks of dock statuses

- Weather data
  - Ten years of temperatures and rain
- Neighborhood data
  - Current neighborhood clusters

I set out to determine what drives ridership.
Here we see strong seasonality.

Here we see that riders behave differently on weekdays and weekends.



I built a model that predicts CB's hourly ridership.

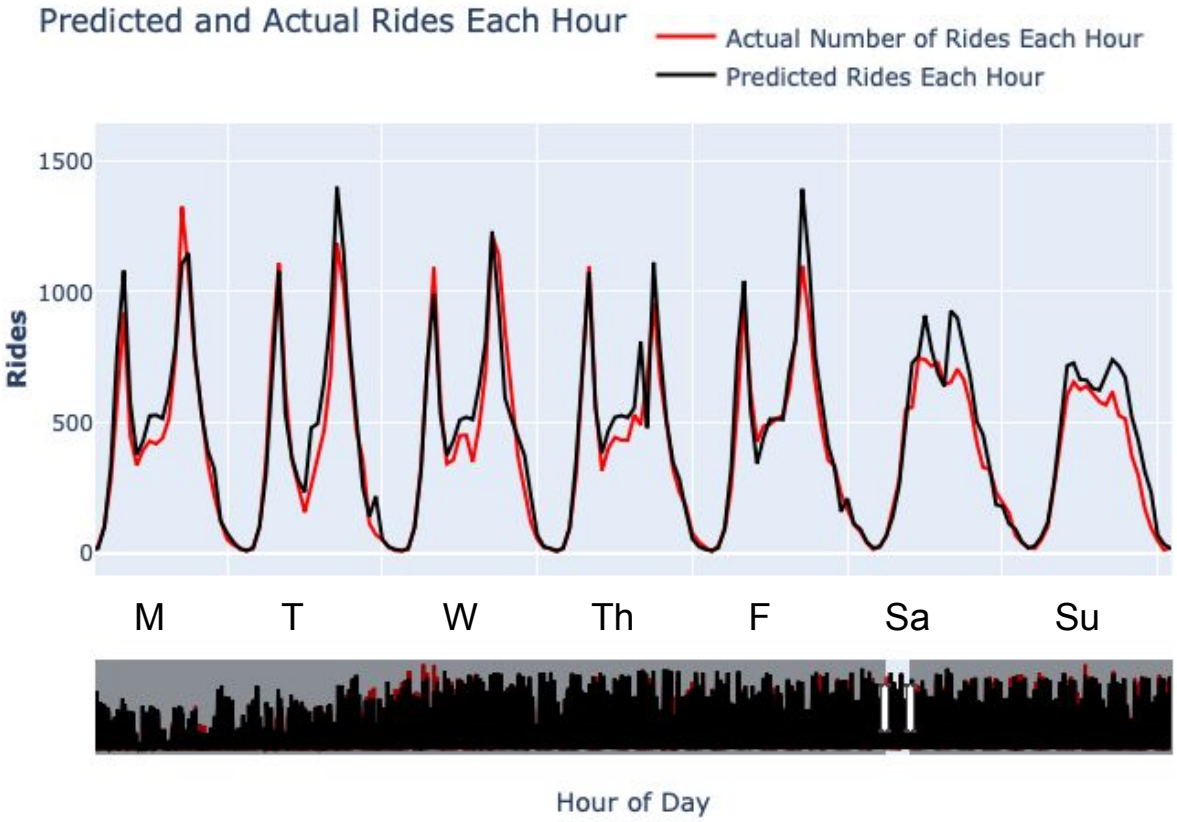
In

1. Temperature
2. Hour
3. Day of Week
4. Rain
5. Month

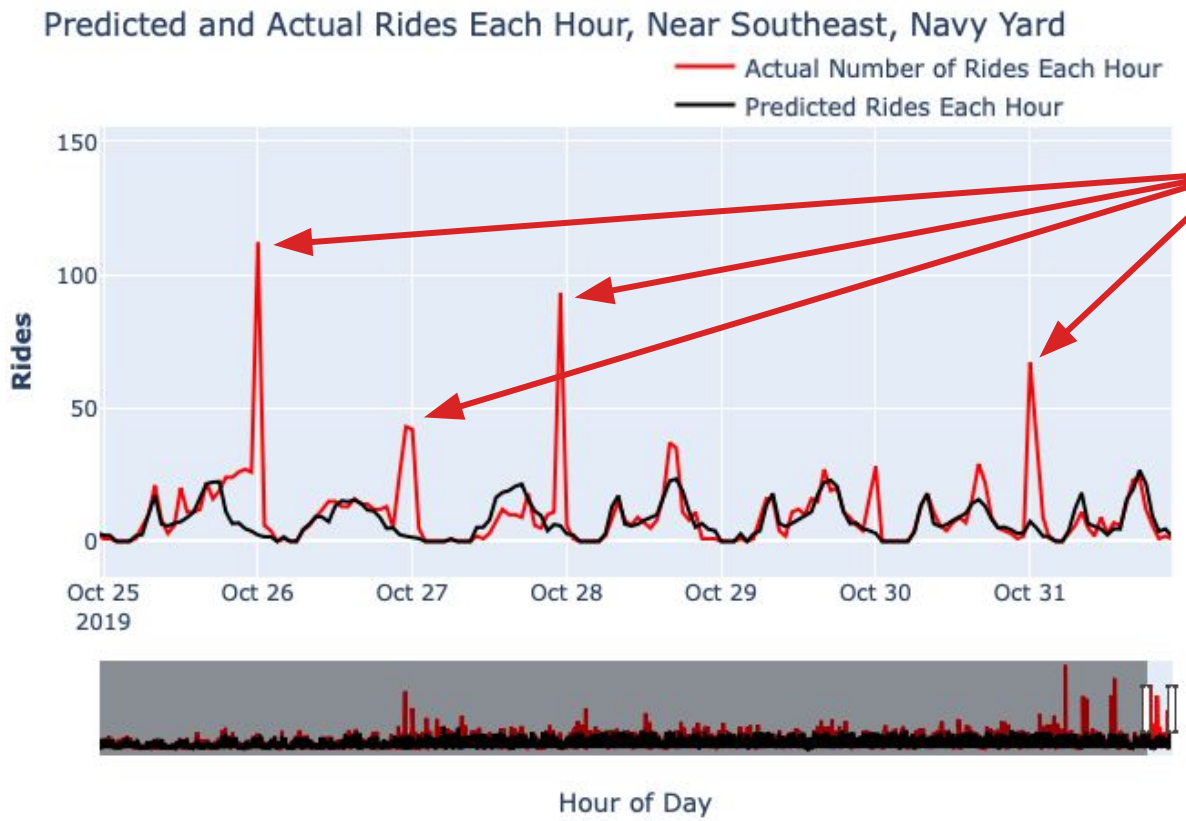
Out

You should expect _____ CB riders during an hour of those conditions.

This is how the actual values and the model's predicted values compare in a typical week.



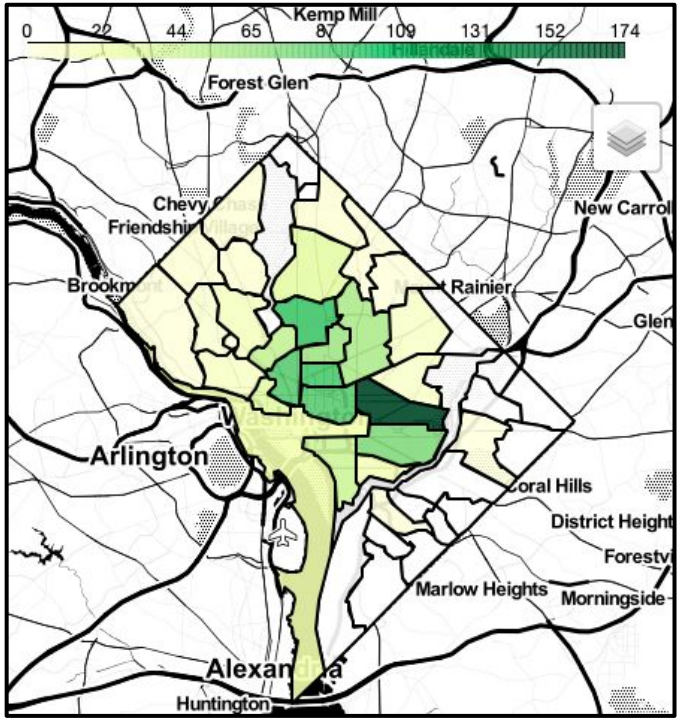
There are some fun anomalies that could be accounted for in a more sophisticated model.

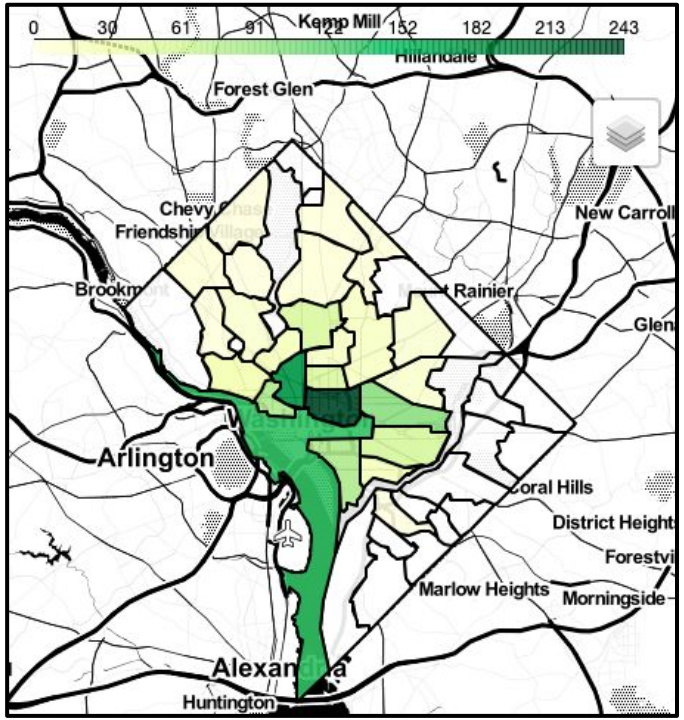


We can map where and when people start and finish rides.

Where Rides Are Starting

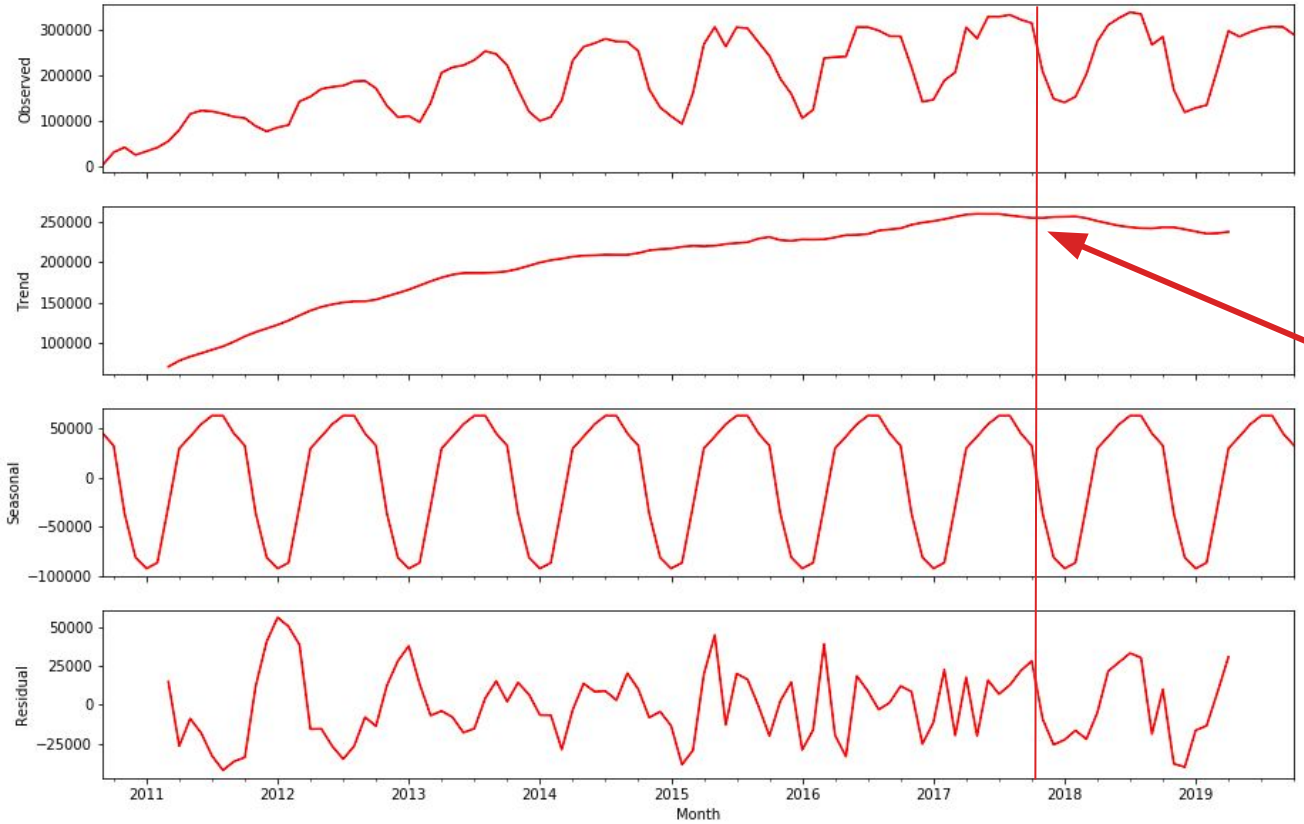
8am Hour





5pm Hour

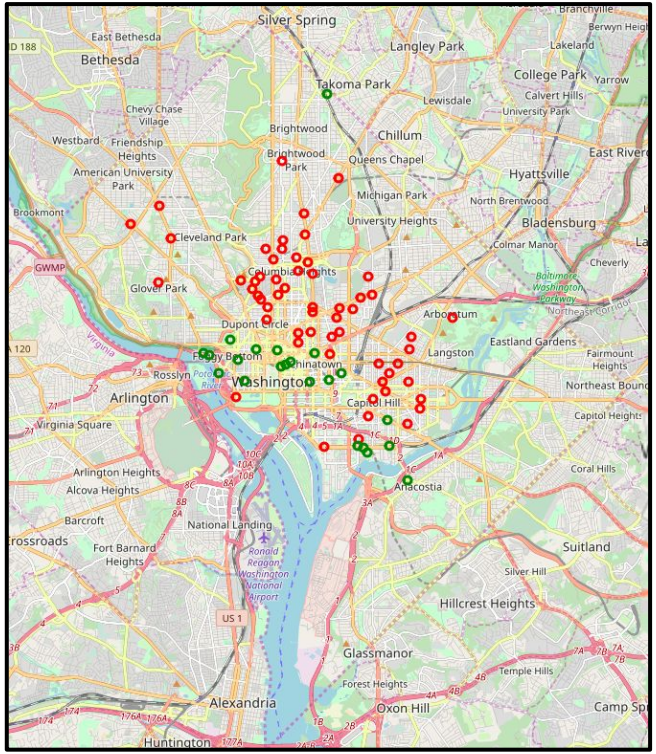
There are also questions of long-term trends.
New travel options impact demand for CB bikes.



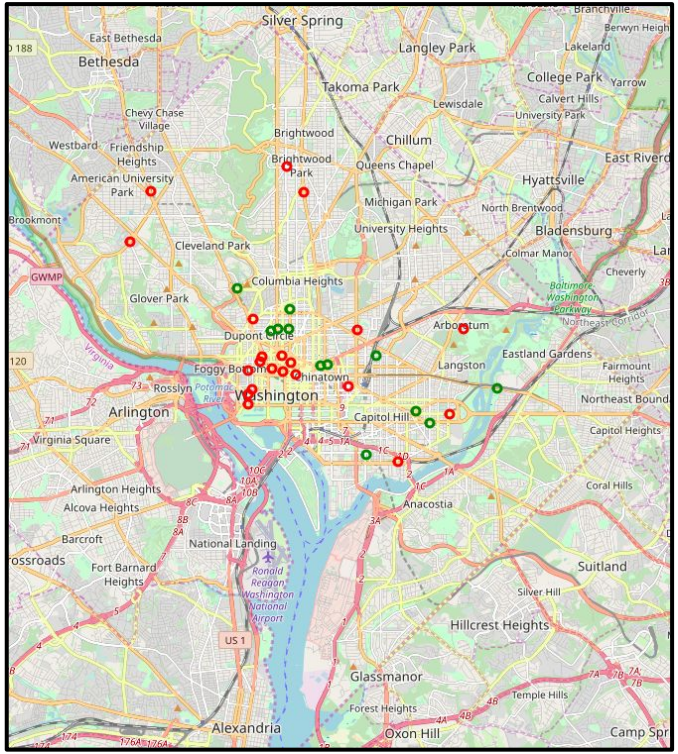
September 2017 - DC launches pilot program for dockless transport

Capital Bikeshare wants to minimize empty stations and full stations.

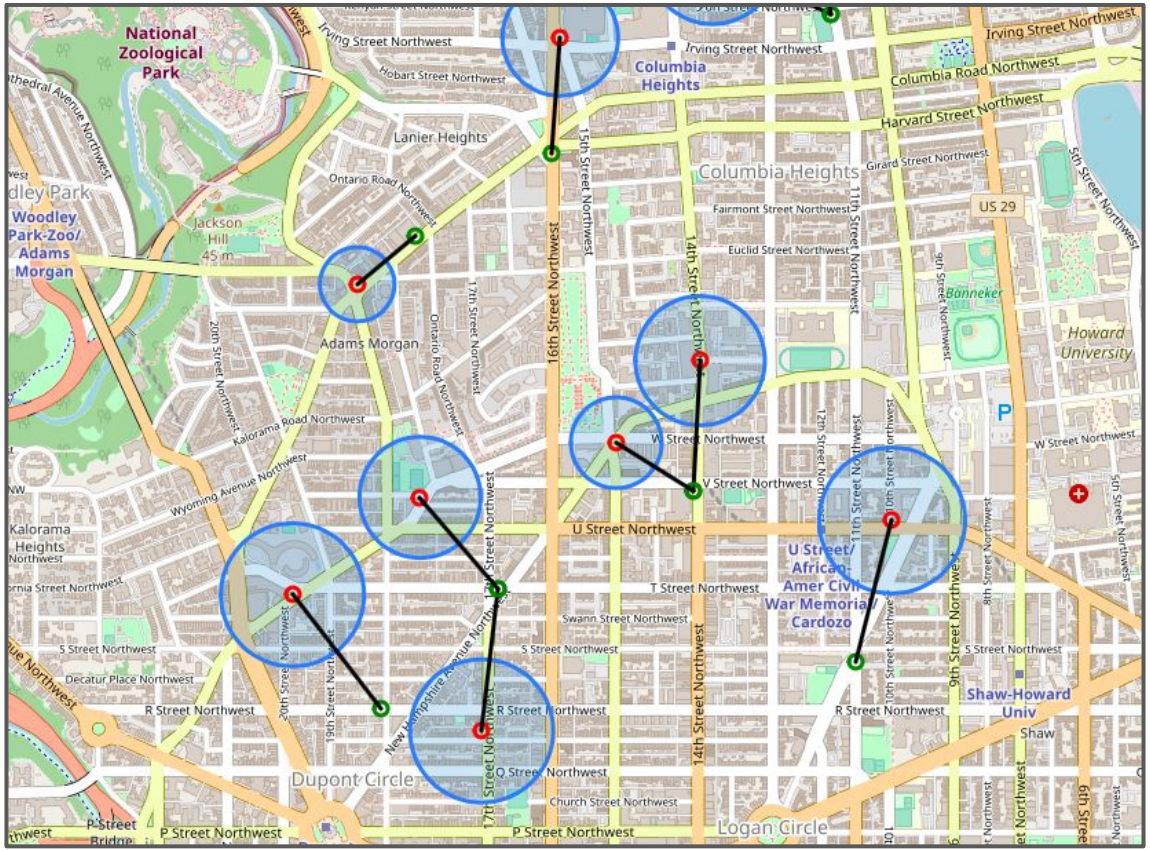
9:30am



7pm



One possibility: incentivize riders to pick up and dock at underutilized stations



Take a look at my web applications for further exploration!

https://obscure-garden-58632.herokuapp.com

Thanks!