Bikeshare Trip Duration and Nearby Infrastructure in San Francisco

CPLN 505 Spring 2020

Chen (Chelsea) Zhang & Dennis (Jiazheng) Zhu

Agenda

1. Introduction
2. Data Collection
3. Data Analysis
4. Results and Interpretation
5. Conclusion and Limitations

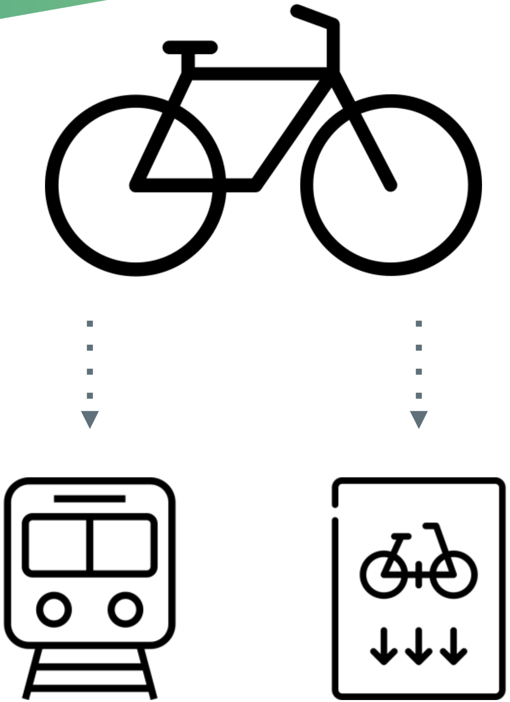
Introduction

Research Question

Does infrastructure around bikeshare stations affect bikeshare trip duration in San Francisco?

Literature Review

"**Public transit** usage is significantly positively associated with bikesharing usage."



"Cyclists most preferred to ride on dedicated **bike infrastructure**, with physically separated lanes and paths."

Data Collection

Joining datasets & Cleaning Data

Datasets

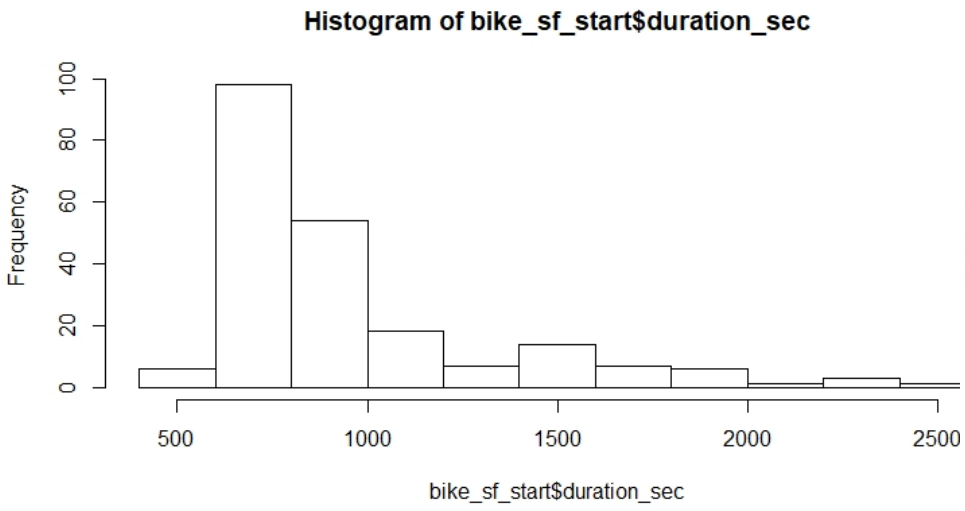
- Data SF (datasf.org)
  - SFMTA Bikeway Network
  - Bicycle Parking
  - Transit stations
- Lyft Bay Wheels Trip Data
  - 10/2019 - 12/2019

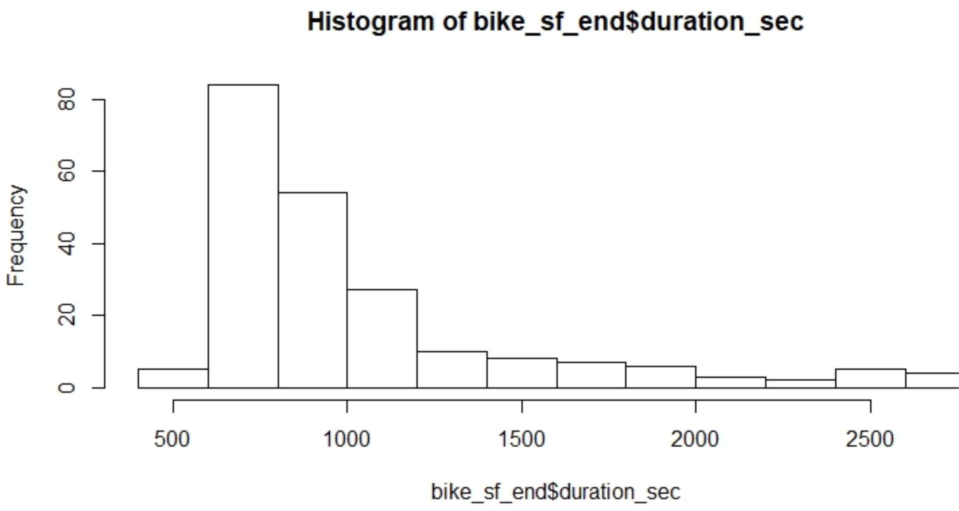
Unit of Analysis

bay wheels

- Each Bay Wheels station in SF
- Average trip duration for all trips that end at a given station
  - Same-day trip > 3 mins

Origin vs. Destination

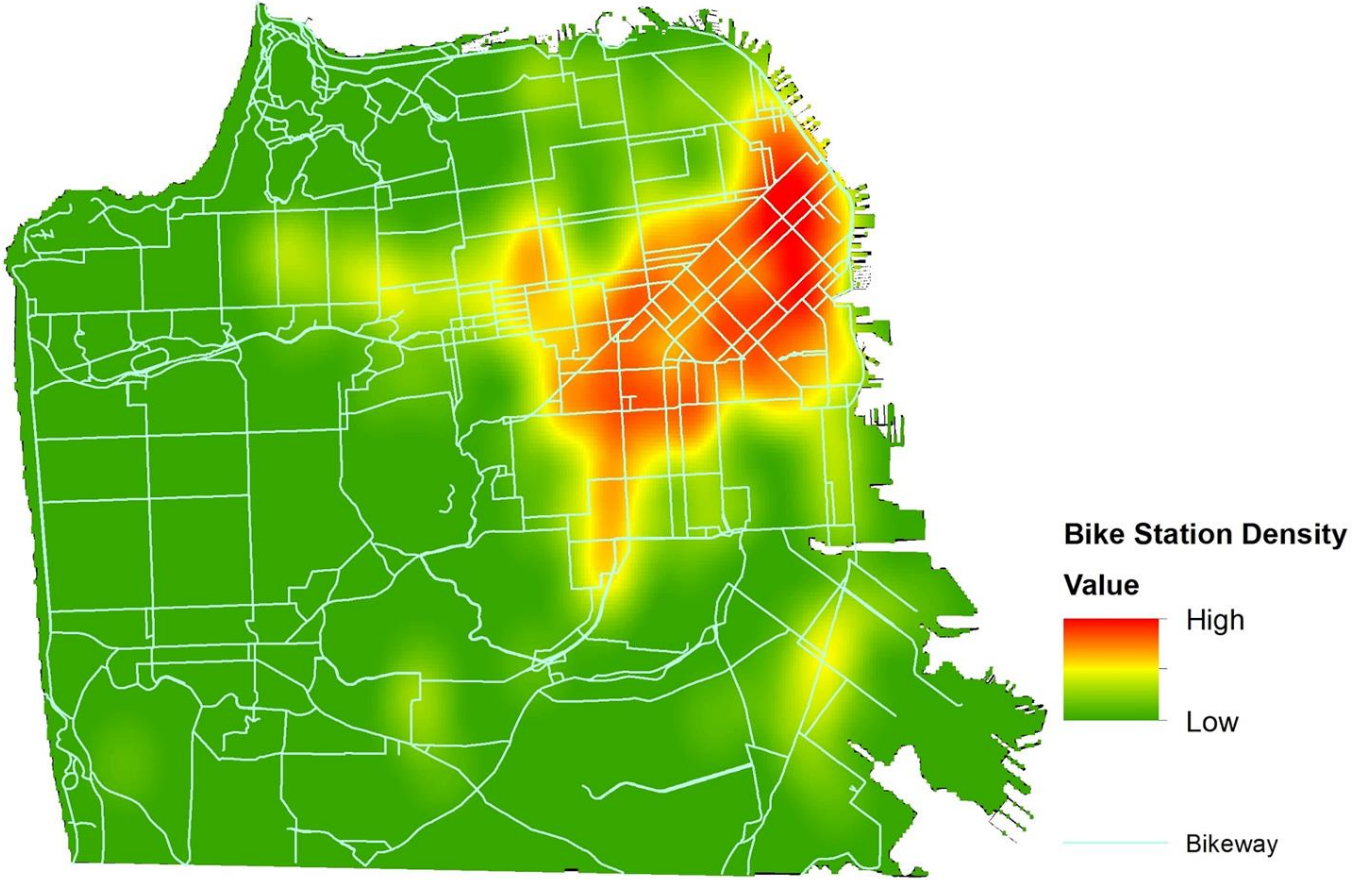




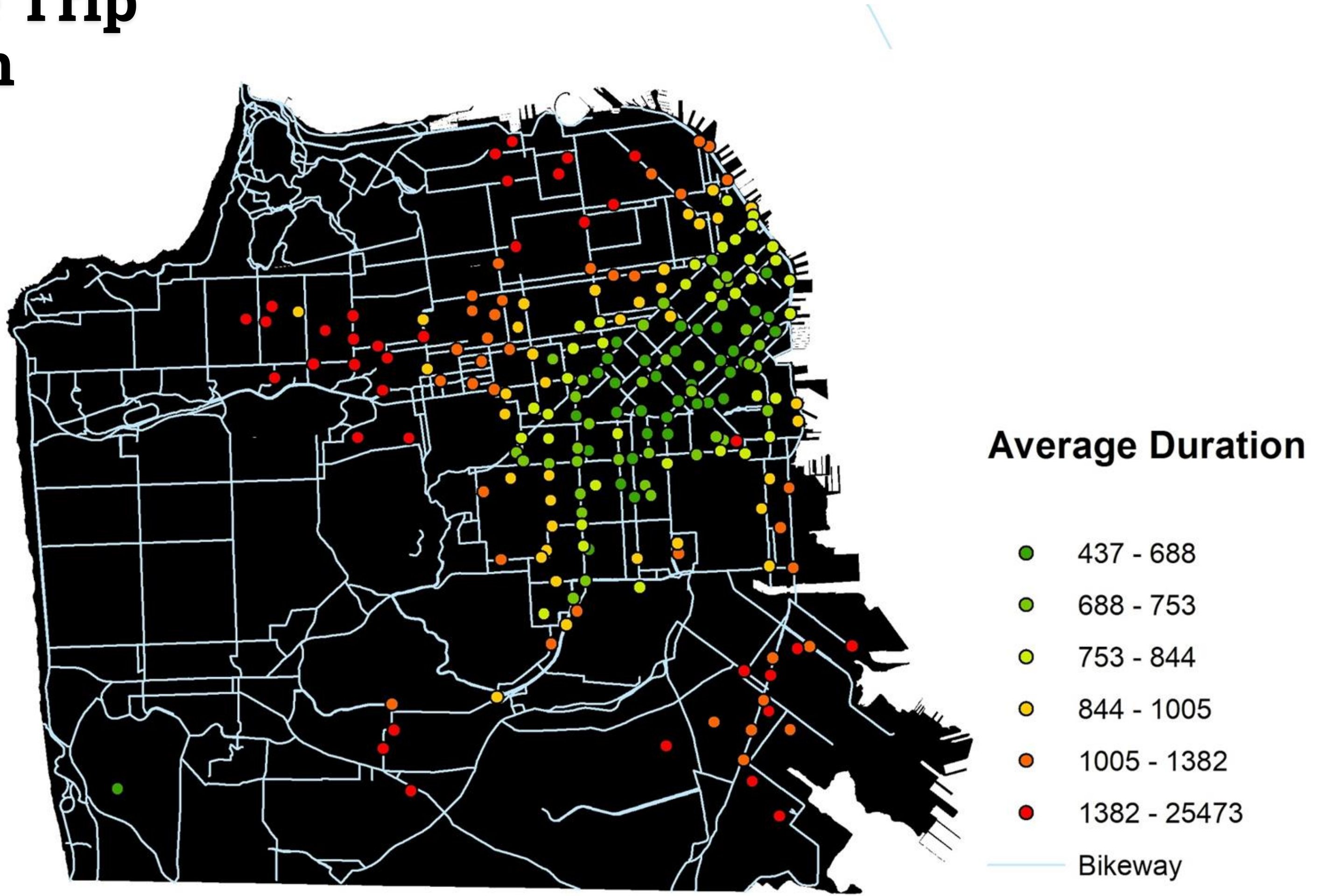
Data Analysis

Variable Selection & Analysis Methods

Bikeshare Station Density



Average Trip Duration



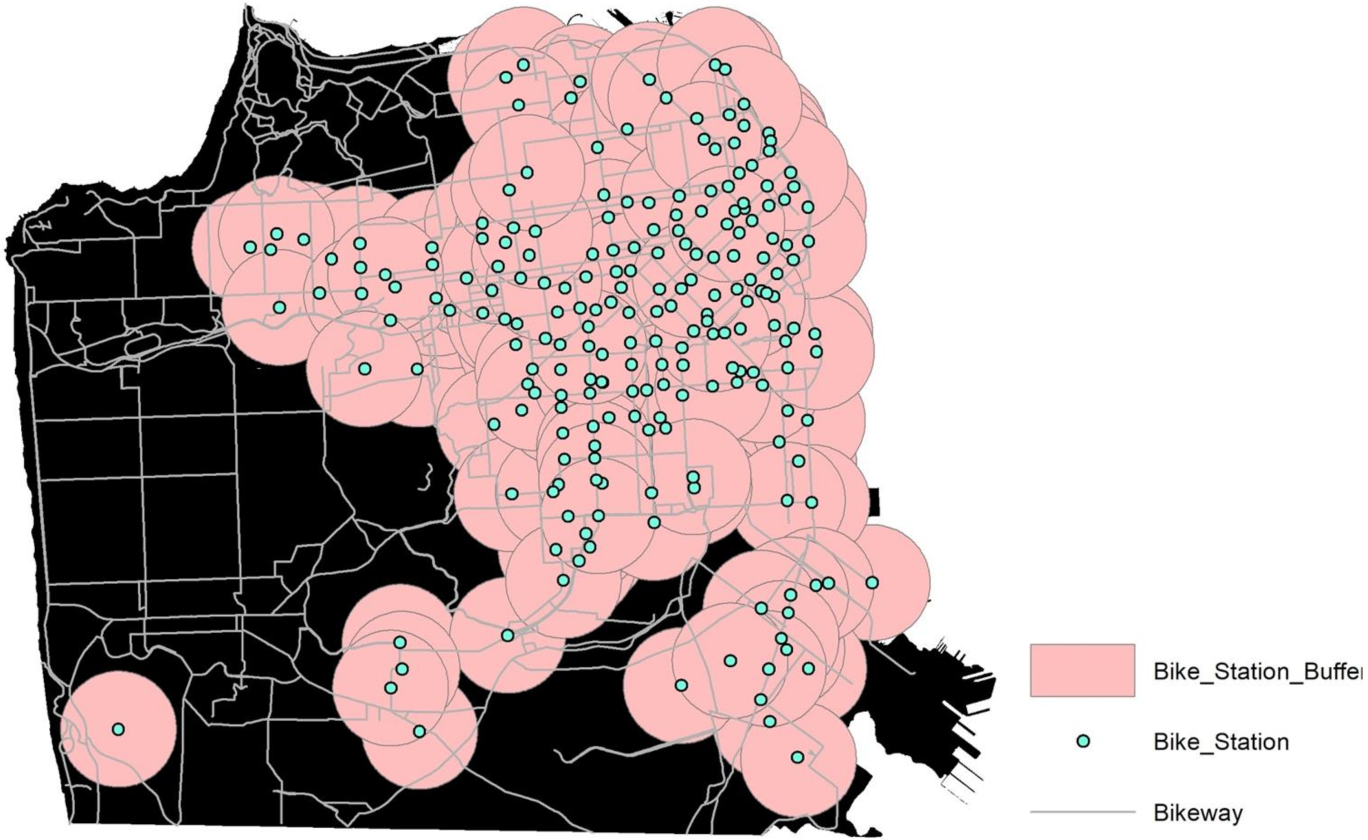
Transit Stations

Commuter Rail(Caltrain)

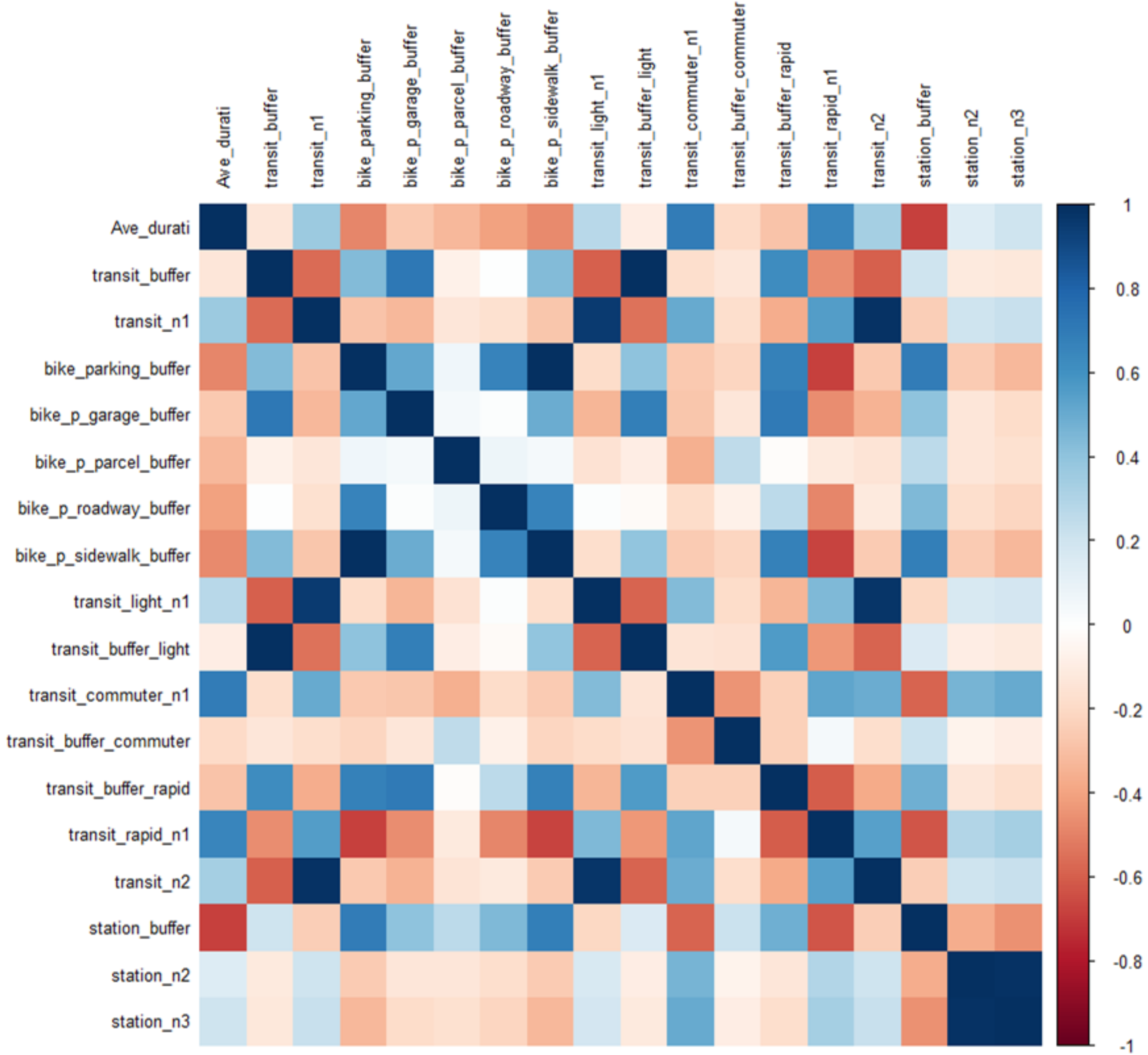
Rapid Rail (BART)

Light Rail (Muni Mero Light Rail)

Half-Mile Buffer



Correlation Matrix



Correlation Matrix

Distance to nearest transit station

Bike parking within each buffer

Bike parcel parking within each buffer

Bike roadway parking within each buffer

Bike sidewalk parking within each buffer

Distance to nearest muni station

Distance to nearest caltrain station

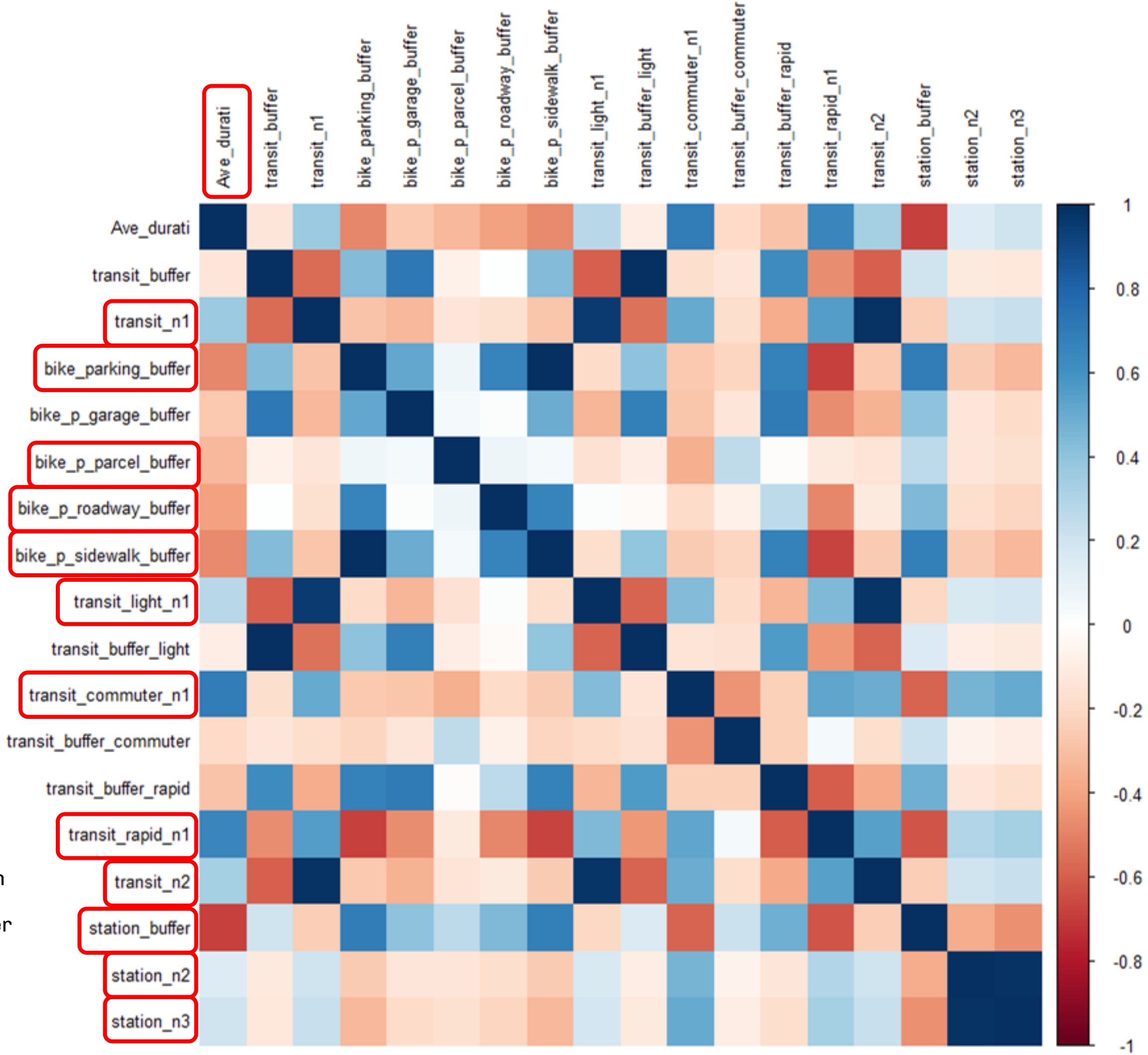
Distance to nearest BART station

Distance to second nearest transit station

Number of bike stations within each buffer

Distance to nearest bike station

Distance to second nearest bike station



Infrastructure Variables

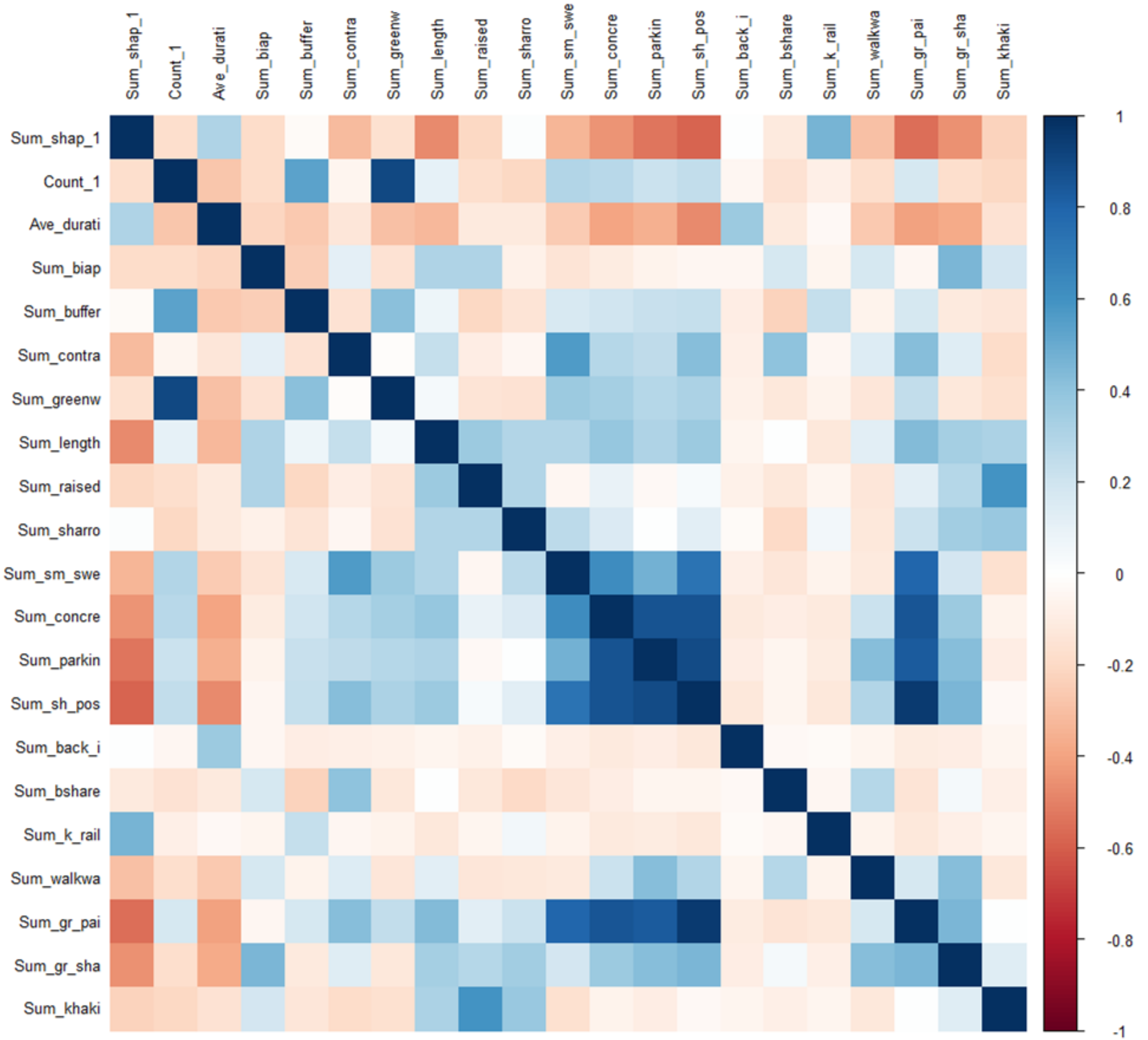
*within 0.5 mile of a bikeshare station

Safe-Hit Posts

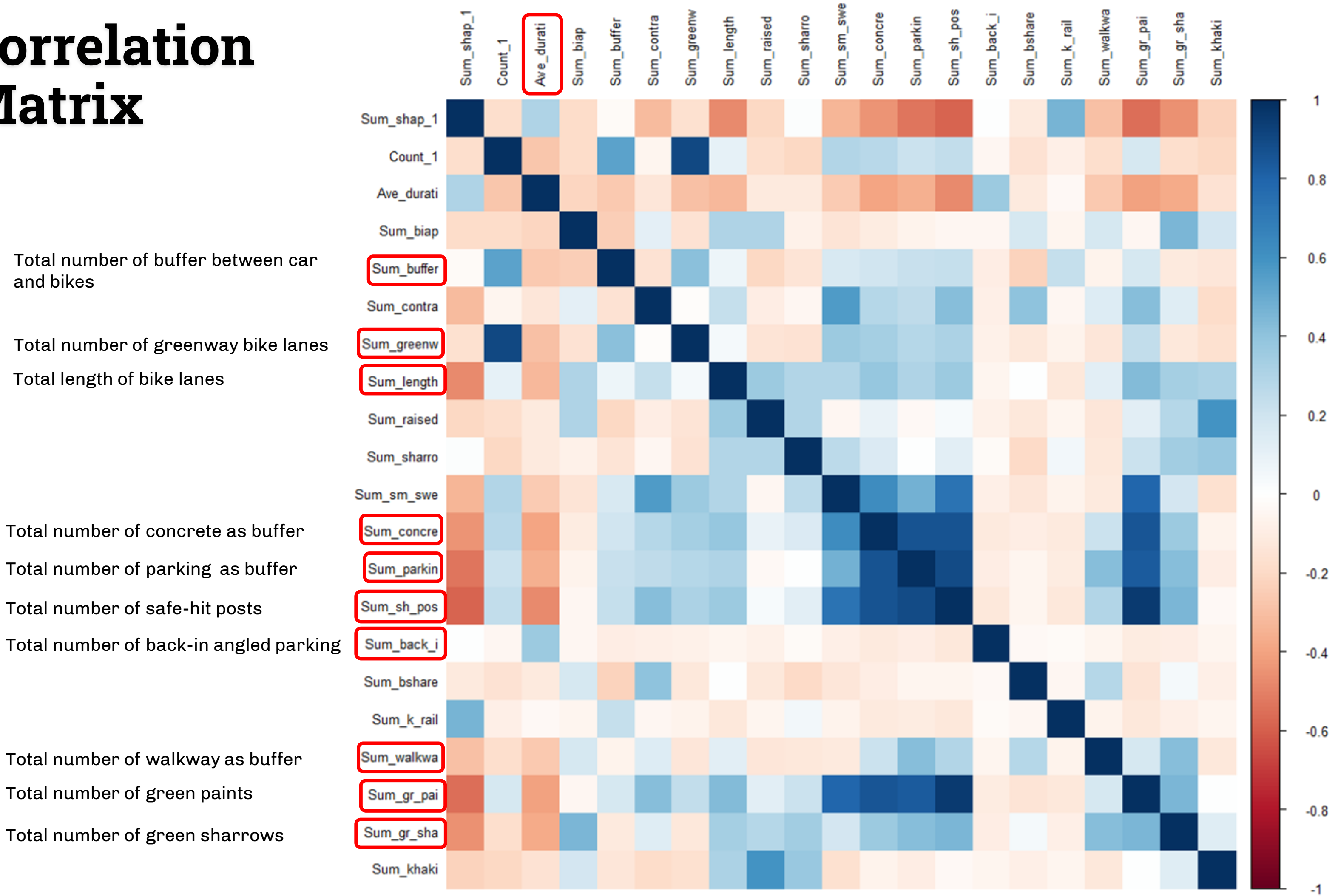
Green Sharrows

Back-In Angled Parking

Correlation Matrix



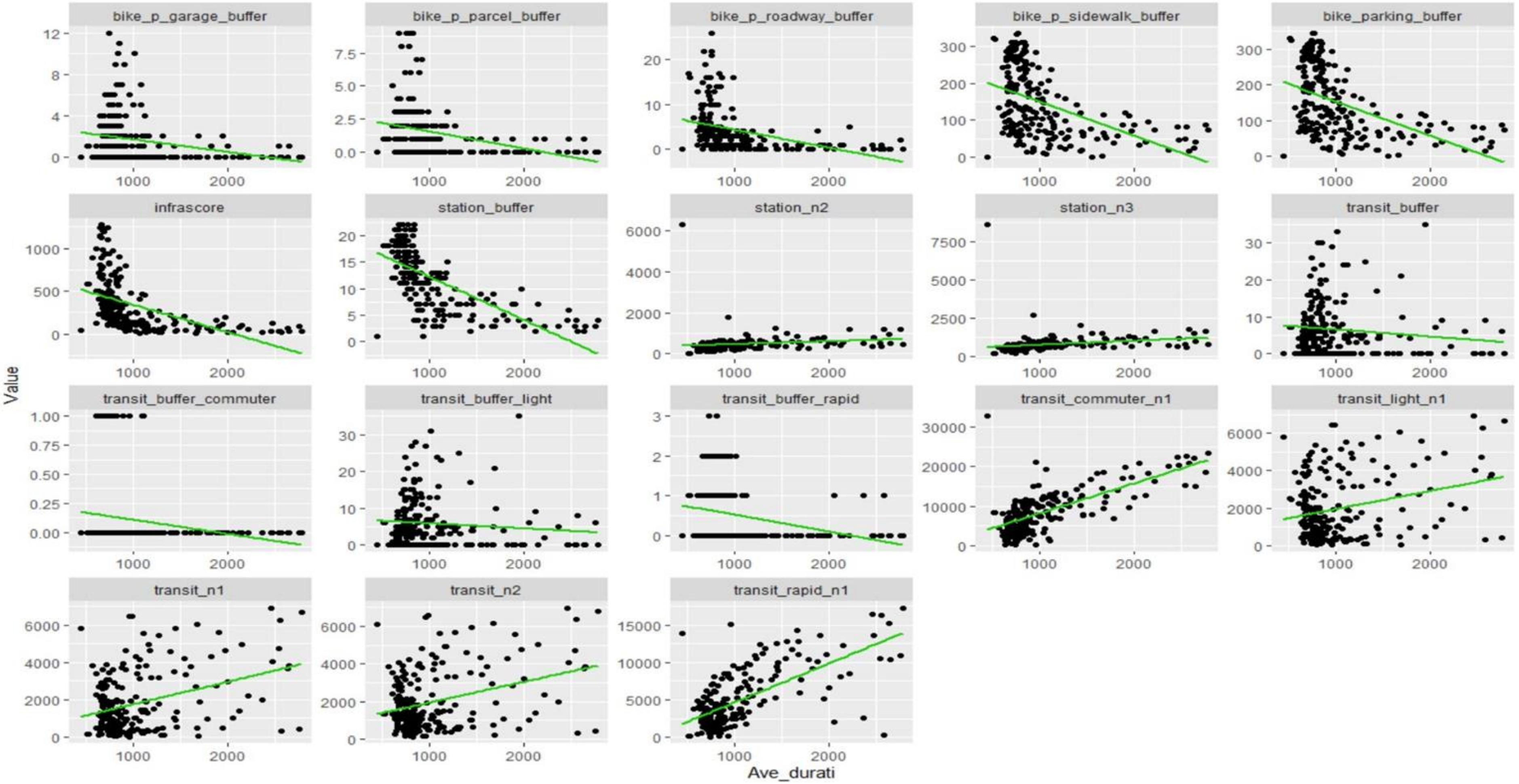
Correlation Matrix



Infrastructure Index

```
dat$infrascore <-
    5*dat$Sum_sh_pos + 4*dat$Sum_concre +
    4*dat$Sum_parkin + 4*dat$Sum_gr_pai +
    4*dat$Sum_gr_sha + 3*dat$Sum_length +
    3*dat$Sum_buffer + 3*dat$Sum_greenw +
    3*dat$Sum_walkwa - 4*dat$Sum_back_i
```

Correlation Analysis



Results & Interpretation

Regression

```
model <- lm(Ave_durati ~ transit_buffer +
transit_n1 + bike_p_parcel_buffer +
bike_p_sidewalk_buffer + transit_light_n1 +
transit_buffer_light + transit_commuter_n1 +
transit_rapid_n1 + transit_buffer + transit_n1 +
transit_n2 + station_buffer + station_n2 +
station_n3 + infrascore, data = mod)
```

```
Call:
lm(formula = Ave_durati ~ transit_buffer + transit_n1 + bike_p_parcel_buffer +
    bike_p_sidewalk_buffer + transit_light_n1 + transit_buffer_light +
    transit_commuter_n1 + transit_rapid_n1 + transit_buffer +
    transit_n1 + transit_n2 + station_buffer + station_n2 + station_n3 +
    infrascore, data = mod)

Residuals:
     Min       1Q   Median       3Q      Max
-1150.00  -112.14   -10.57    94.08   834.57

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)            9.877e+02  1.111e+02   8.893 3.46e-16 ***
transit_buffer         1.728e+02  3.407e+01   5.071 8.95e-07 ***
transit_n1             1.592e-01  8.128e-02   1.959  0.05147 .
bike_p_parcel_buffer  -2.011e+01  9.360e+00  -2.148  0.03287 *
bike_p_sidewalk_buffer -1.205e+00  3.659e-01  -3.294  0.00117 **
transit_light_n1       1.575e-01  5.973e-02   2.637  0.00901 **
transit_buffer_light  -1.668e+02  3.517e+01  -4.744 3.97e-06 ***
transit_commuter_n1    5.255e-02  5.441e-03   9.658  < 2e-16 ***
transit_rapid_n1       5.276e-02  7.756e-03   6.803 1.15e-10 ***
transit_n2            -3.429e-01  1.180e-01  -2.905  0.00409 **
station_buffer        -3.217e+01  6.266e+00  -5.135 6.65e-07 ***
station_n2             7.595e-02  2.301e-01   0.330  0.74172
station_n3            -3.249e-01  1.755e-01  -1.852  0.06555 .
infrascore             2.620e-01  8.110e-02   3.230  0.00145 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 238.9 on 201 degrees of freedom
Multiple R-squared:  0.7729,	Adjusted R-squared:  0.7582
F-statistic: 52.63 on 13 and 201 DF,  p-value: < 2.2e-16
```

Regression (Stepwise)

```
model_sw <- step(lm(Ave_durati ~ transit_buffer +
transit_n1 + bike_p_parcel_buffer +
bike_p_sidewalk_buffer + transit_light_n1 +
transit_buffer_light + transit_commuter_n1 +
transit_rapid_n1 + transit_buffer + transit_n1 +
transit_n2 + station_buffer + station_n2 +
station_n3 + infrascore, data = mod),
direction= "backward")
```

Call:

```
lm(formula = Ave_durati ~
transit_buffer +
transit_n1 +
bike_p_parcel_buffer +
bike_p_sidewalk_buffer +
transit_light_n1 +
transit_buffer_light +
transit_commuter_n1 +
transit_rapid_n1 +
transit_n2 +
station_buffer +
station_n3 +
infrascore,
data = mod)
```

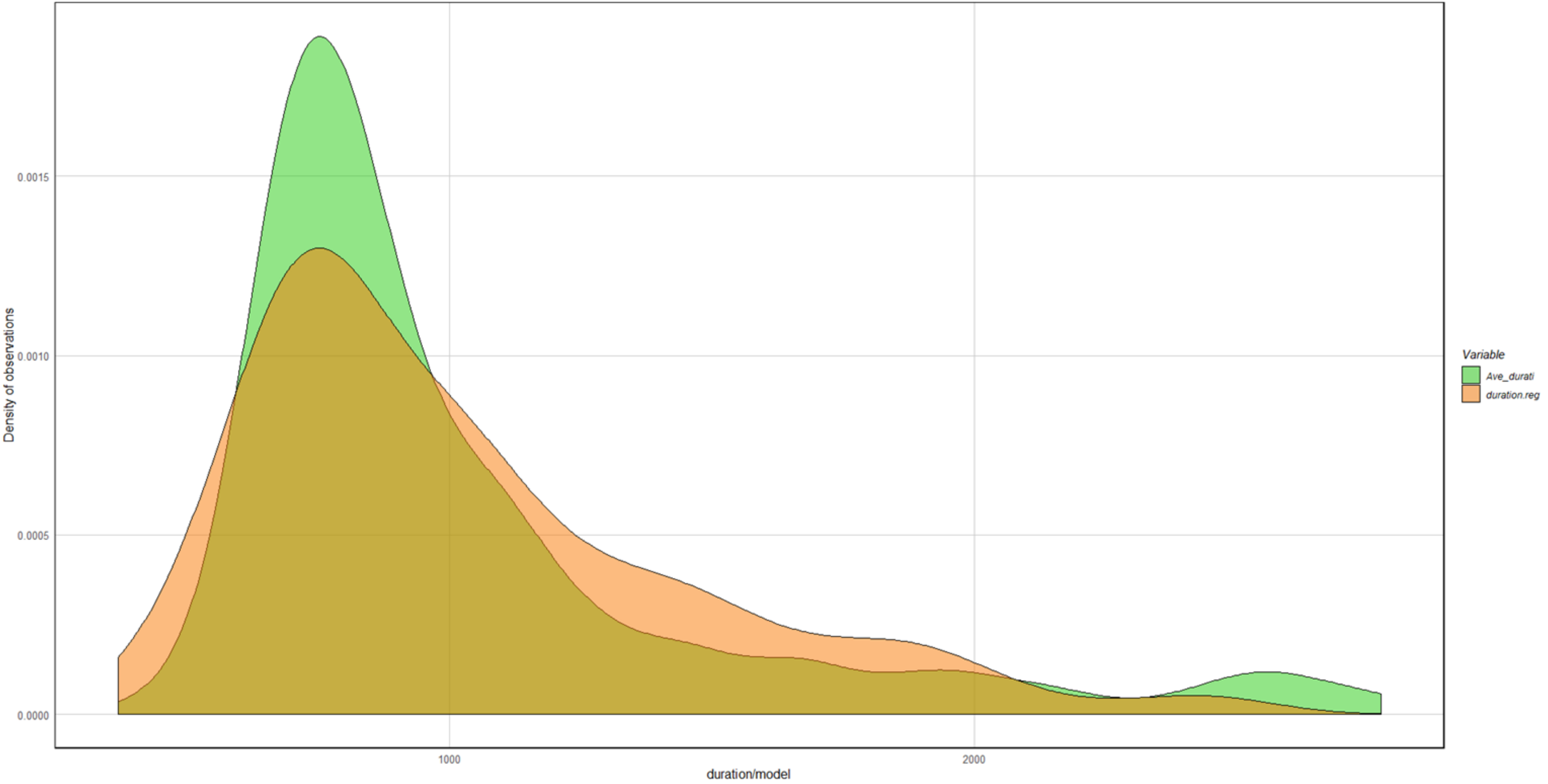
Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	9.716e+02	9.947e+01	9.767	< 2e-16	***
transit_buffer	1.725e+02	3.398e+01	5.076	8.72e-07	***
transit_n1	1.626e-01	8.046e-02	2.021	0.04461	*
bike_p_parcel_buffer	-2.004e+01	9.337e+00	-2.146	0.03304	*
bike_p_sidewalk_buffer	-1.187e+00	3.606e-01	-3.291	0.00118	**
transit_light_n1	1.554e-01	5.924e-02	2.623	0.00939	**
transit_buffer_light	-1.667e+02	3.509e+01	-4.750	3.85e-06	***
transit_commuter_n1	5.252e-02	5.428e-03	9.676	< 2e-16	***
transit_rapid_n1	5.301e-02	7.703e-03	6.882	7.28e-11	***
transit_n2	-3.443e-01	1.177e-01	-2.926	0.00383	**
station_buffer	-3.150e+01	5.908e+00	-5.331	2.60e-07	***
station_n3	-2.680e-01	3.254e-02	-8.235	2.21e-14	***
infrascore	2.595e-01	8.058e-02	3.220	0.00149	**

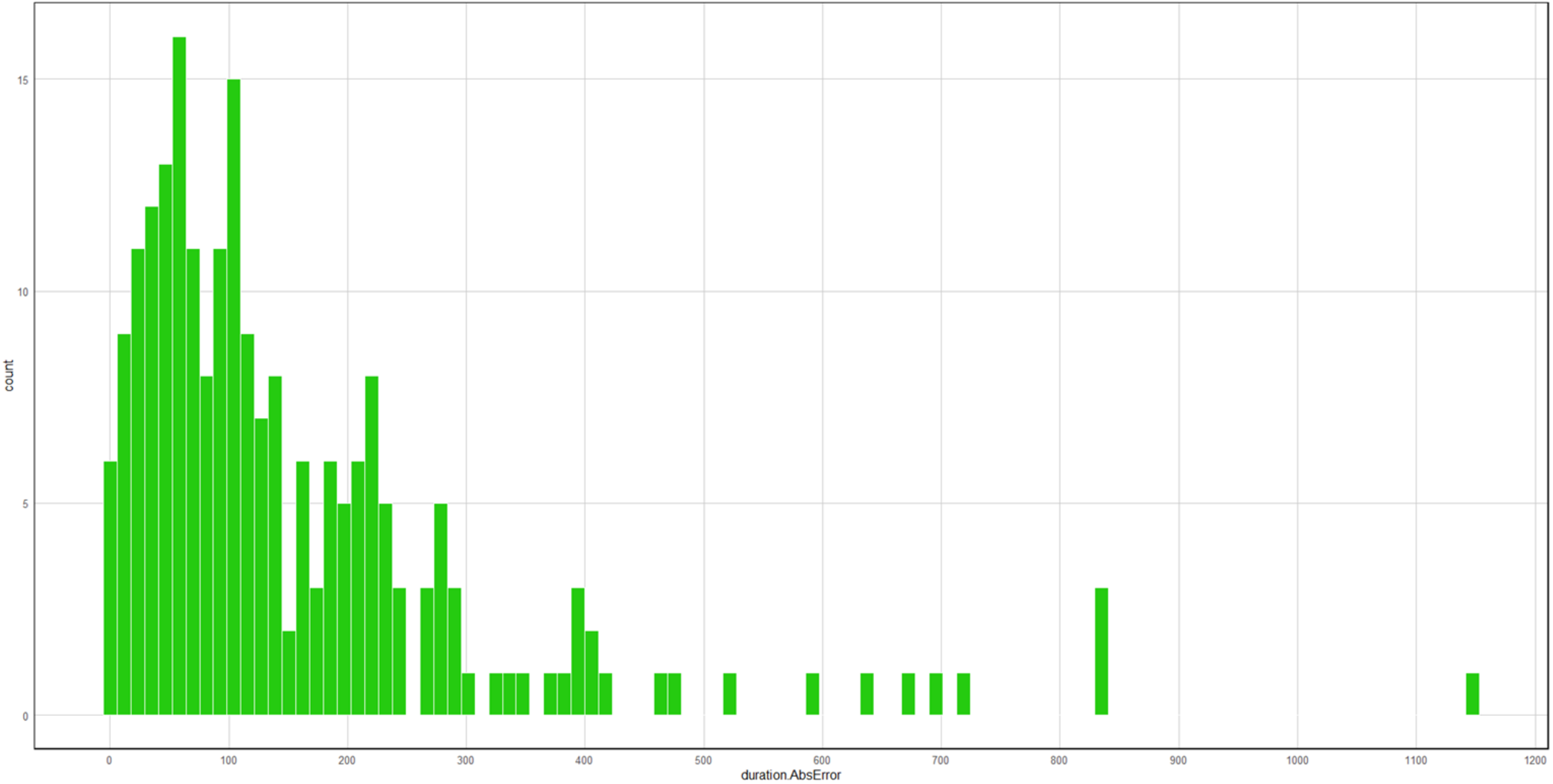
```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 238.3 on 202 degrees of freedom
Multiple R-squared:  0.7728,	Adjusted R-squared:  0.7593
F-statistic: 57.26 on 12 and 202 DF,  p-value: < 2.2e-16
```

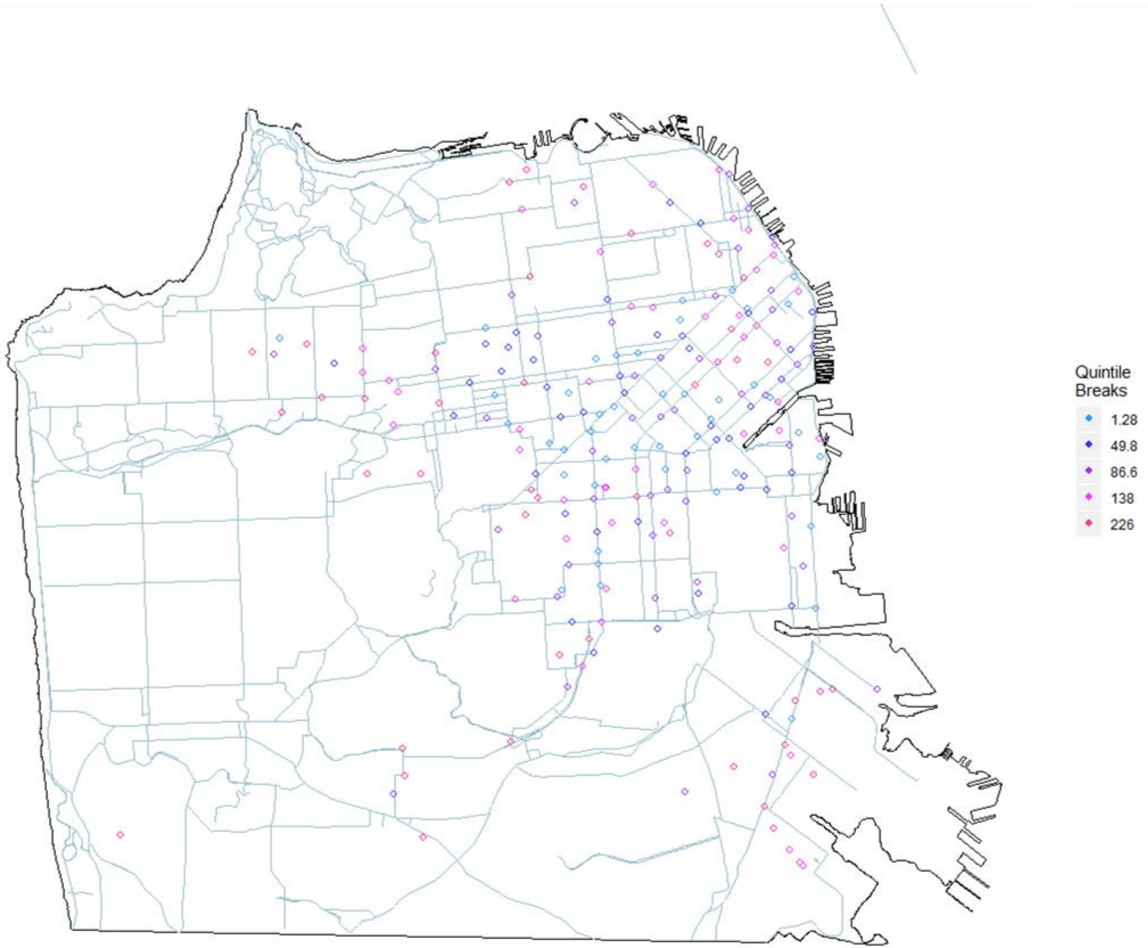
Distribution of Duration



Distribution of Model Error



Duration Absolute Error



Conclusion & Limitation

Conclusion

Shorter trips when there are

- more subway stations
- more bikeshare stations
- more sidewalk/parcel parking spots for bikes

Longer trips when there are

- more Caltrain/BART station
- more bicycle-specific infrastructure

...at the destination bikeshare station.

Limitation

- No route data available
- Infrastructure index could be improved

References

- Open data SF
- Lyft Bay Wheels System Data
- Jaffe, Eric, et al. “The Methodology of Bike-Share Station Placement in New York City.” CityLab, 5 Oct. 2011, www.citylab.com/transportation/2011/10/how-new-york-city-will-choose-its-bike-share-stations/248/.
- Zhang, Yuanyuan, and Yuming Zhang. “Associations between Public Transit Usage and Bikesharing Behaviors in The United States.” Sustainability, vol. 10, no. 6, Apr. 2018, p. 1868., doi:10.3390/su10061868.
- Buehler, Ralph, and Jennifer Dill. “Bikeway Networks: A Review of Effects on Cycling.” Transport Reviews 36, no. 1 (January 2, 2016): 9–27.
- Park, Yujin, and Gulsah Akar. “Why Do Bicyclists Take Detours? A Multilevel Regression Model Using Smartphone GPS Data.” Journal of Transport Geography 74 (2019): 191–200.
- Wergin, Jon, and Ralph Buehler. “Where Do Bikeshare Bikes Actually Go?: Analysis of Capital Bikeshare Trips with GPS Data.” Transportation Research Record: Journal of the Transportation Research Board 2662, no. 1 (2017): 12–21.

Thank you!

Chelsea Zhang:
czhang03@upenn.edu
Dennis Zhu:
djzzhu@upenn.edu