Modeling parking demand in San Francisco

Study of how space/time features influence parking behaviors and their implications on planning practice

Bingchu Chen & “Gillian” Xuezhu Zhao
Planning by Numbers | Spring 2021

Use case

Parking in SF

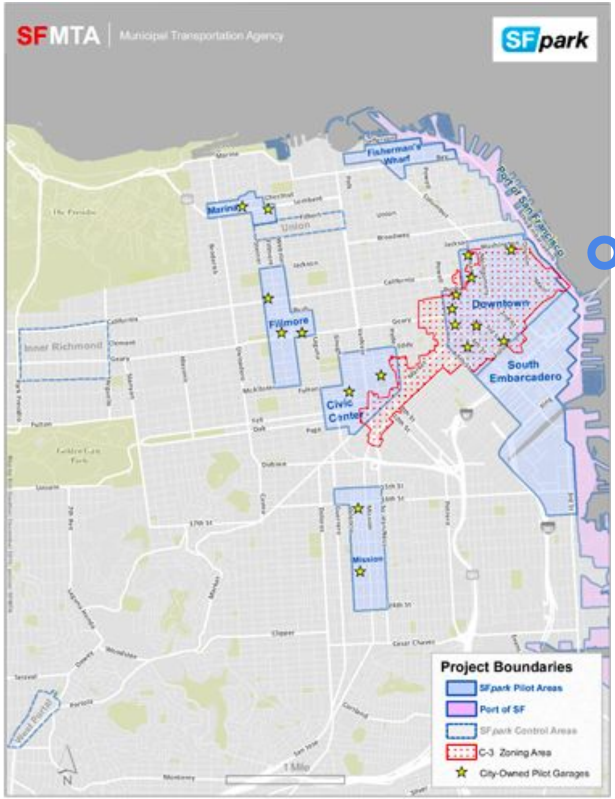
Source: sfmta.com

28,000 meters

Most enforced from 9 a.m. to 6 p.m.
Monday through Saturday
Except for holidays

Use case

Smart parking
Demand-responsive pricing



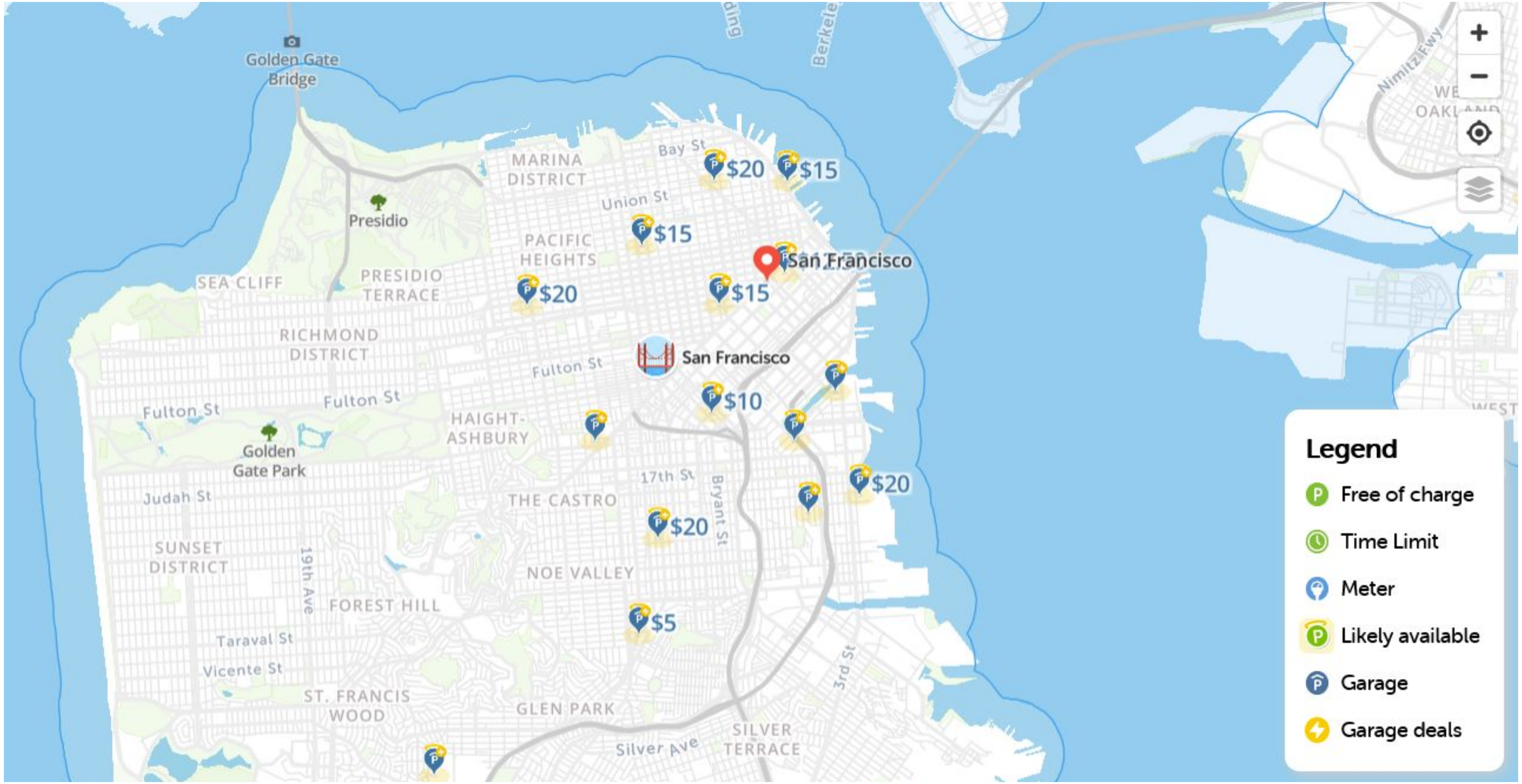
Source: Federal Highway Administration

Aug 2011 SFpark pilot

Dec 2017 City-wide program

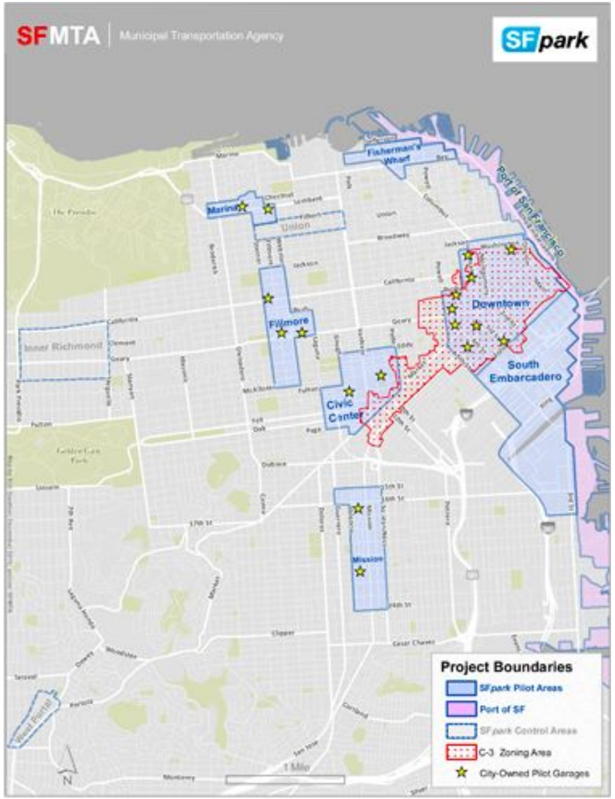
Encourages people to park in underutilized blocks and garages, helping to open up spaces in busy areas and at busy times

Use case



Source: https://www.spotangels.com/san-francisco-parking

Use case



Source: Federal Highway Administration; sfmta.com

SFpark evaluation

Increased business for local businesses

Sales tax revenues rose over 35% in SFpark areas during the compared to less than 20% in the other parts of the city.

Lower parking rates

Average meter rates were reduced by 4% (down $0.11/hour) in SFpark on-street pilot areas.

Decreased parking search time

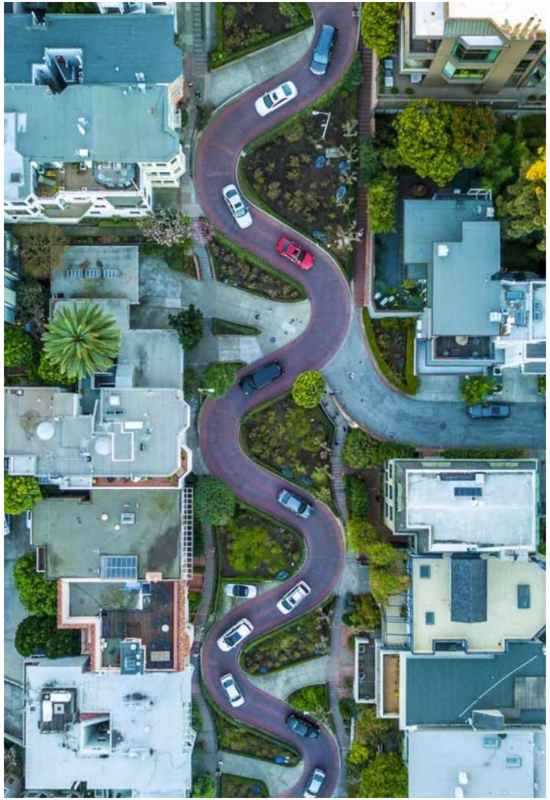
Reported parking search time went down by 43% under the SFpark pilot.

Decreased daily vehicle miles traveled

Reduced circling for parking led to a 30% decrease in miles traveled in SFpark areas.

Use case

Predict **parking demand**



Source: San Francisco Travel Association

Recognize factors associated with parking demand

➔ Help other cities identify high-need areas/time periods to implement parking demand management

➔ Help bottom-up predictions of parking demand and price

Data Processing

SFMTA Parking Meter Detailed Revenue Transactions

DataSF Parking Meters

Sampled 1M transactions between July 4 - August 1, 2017

Joined by POST_ID

Parking Transactions grouped by POST_ID and time

Daily Parking Count per Meter

Environmental Features Data

2013-2017 ACS Data

Feature engineering to create Nearest Neighbors and Binary Variables

Demographic Characteristics

Transportation Behaviors

Spatial Features

Categorical Features

Predictors

Dependent Variable

SFMTA Parking Meter Detailed Revenue Transactions

DataSF Parking Meters

Sampled 1M transactions between July 4 - August 1, 2017

Joined by POST_ID

Parking Transactions grouped by POST_ID and time

Source: https://data.sfgov.org/

Use case

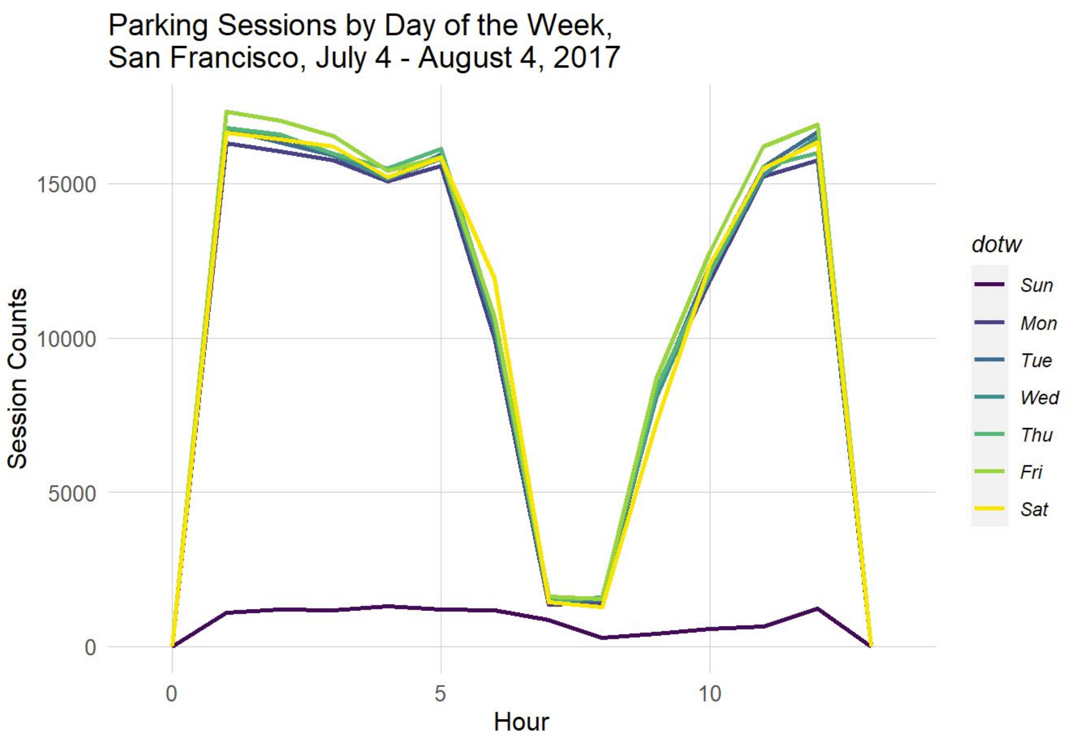
Pro-
cess

Parking Demand

Time:

Day of the week

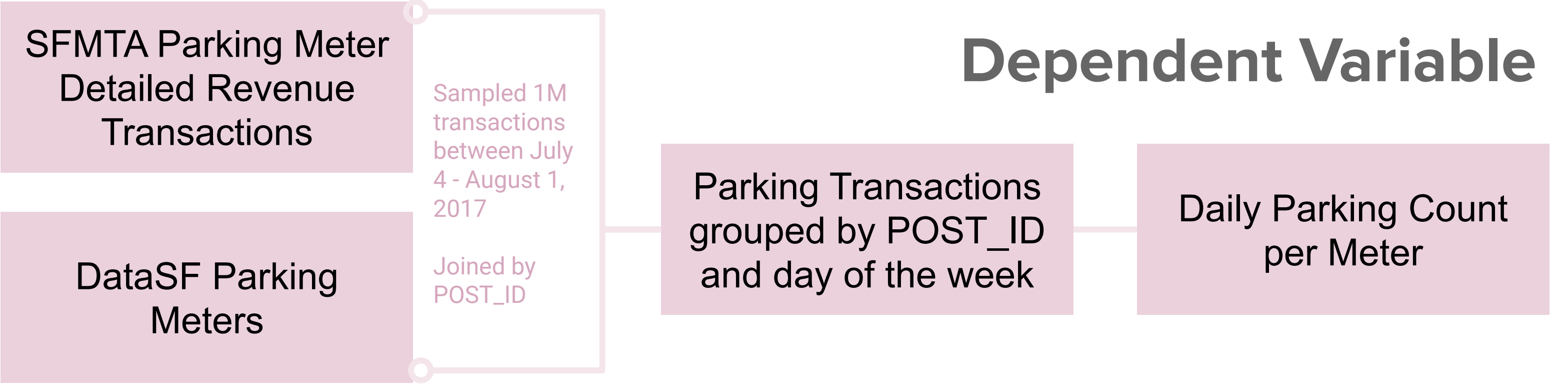
Hour of the day



Use case

Pro-
cess

Dependent Variable



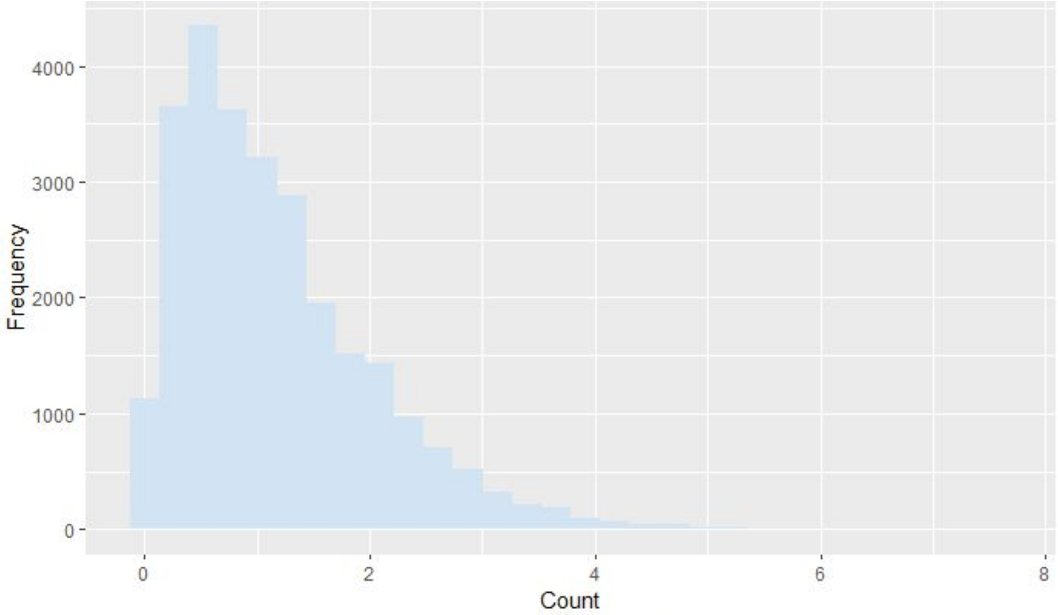
Use case

Pro-cess

Parking Demand

100,000 transactions
26,766 meters

Daily Count of parking transaction per meter



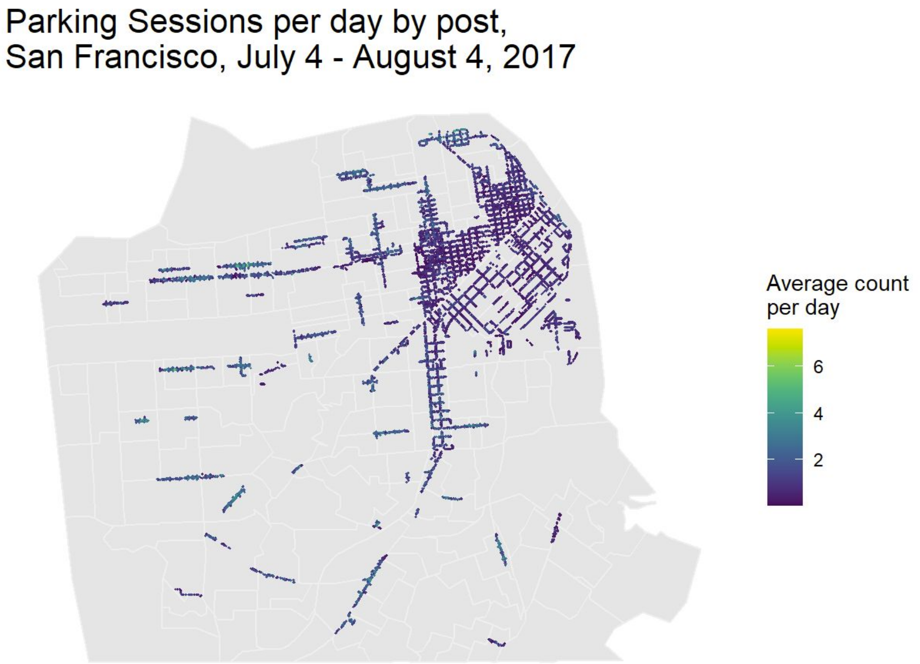
Use case

Pro-cess

Parking Demand

100,000 transactions
26,766 meters

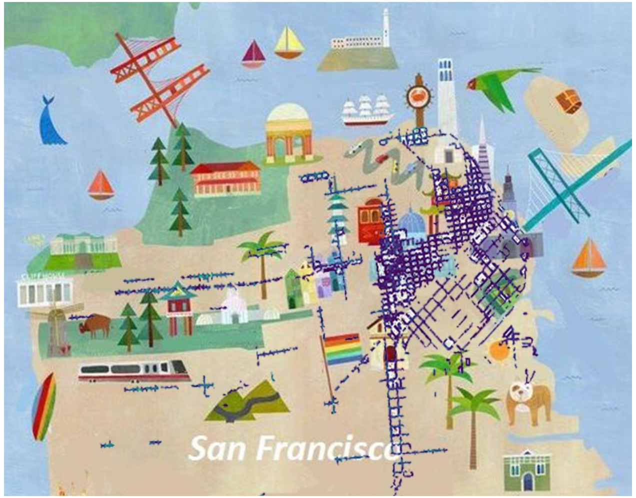
Daily Count of parking transaction per meter



Parking Demand

100,000 transactions
26,766 meters

Most meters located near downtown and attractions



Source: https://www.sporcle.com/games/mucciniale/frisco-tourist-map

Independent Variable

SFMTA Parking Meter Detailed Revenue Transactions

DataSF Parking Meters

Sampled 1M transactions between July 4 - August 1, 2017

Joined by POST_ID

Parking Transactions grouped by POST_ID and day of the week

Daily Parking Count per Meter

Environmental Features Data

Source: https://data.sfgov.org/

2013-2017 ACS Data

Source: https://data.census.gov/

Created Nearest Neighbors and Binary Variables

Demographic Characteristics

Transportation Behaviors

Spatial Features

Categorical Features

Use case

Pro-
cess

Key Predictor

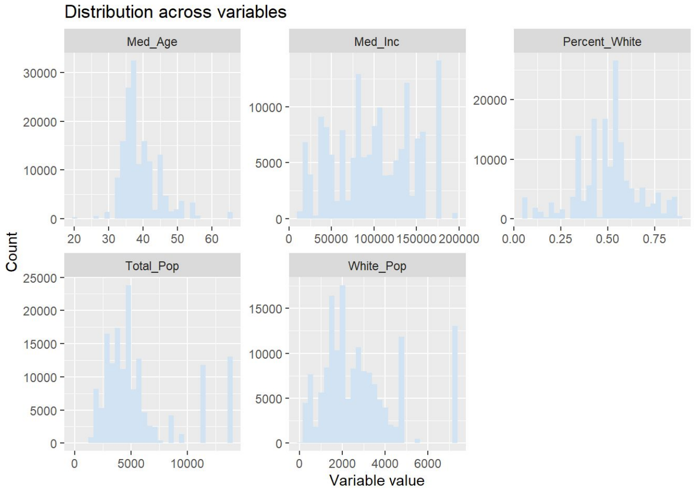
Distance to
Special Event Locations

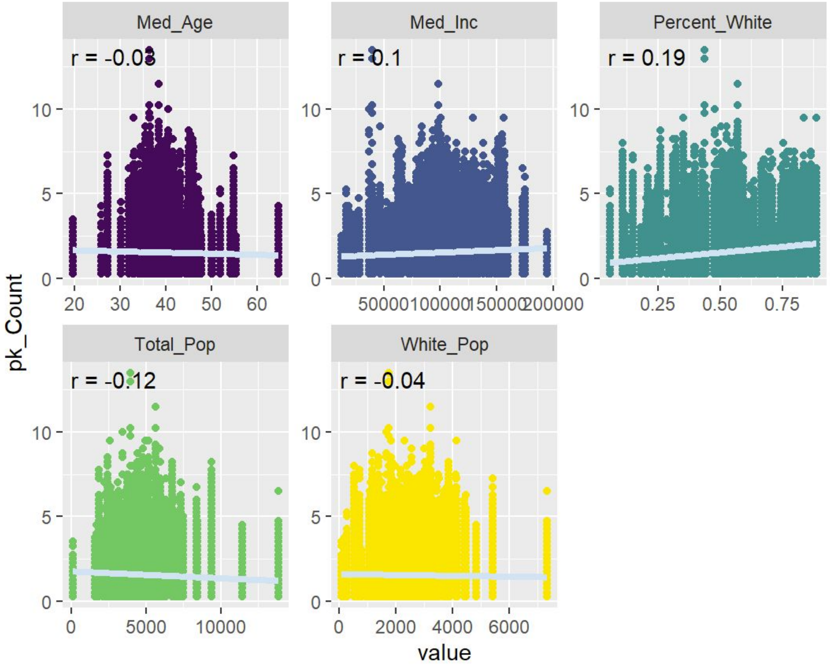
Use case

Pro-
cess

Exploratory Analysis

Demographic variables



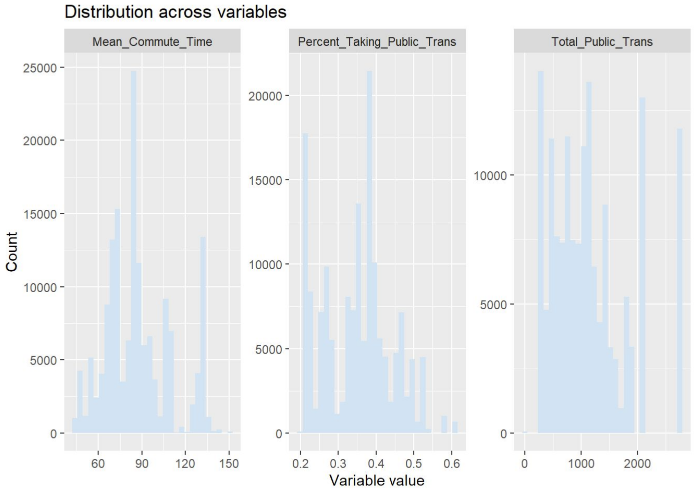


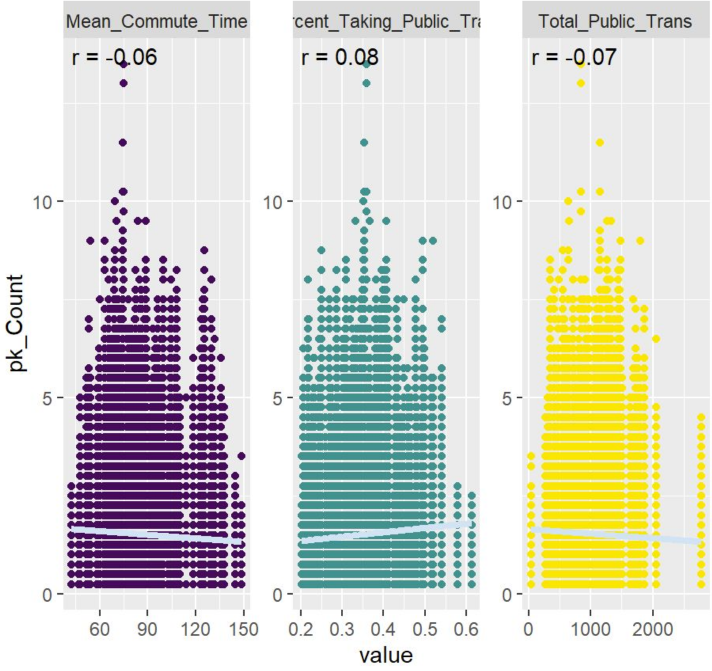
Use case

Pro-cess

Exploratory Analysis

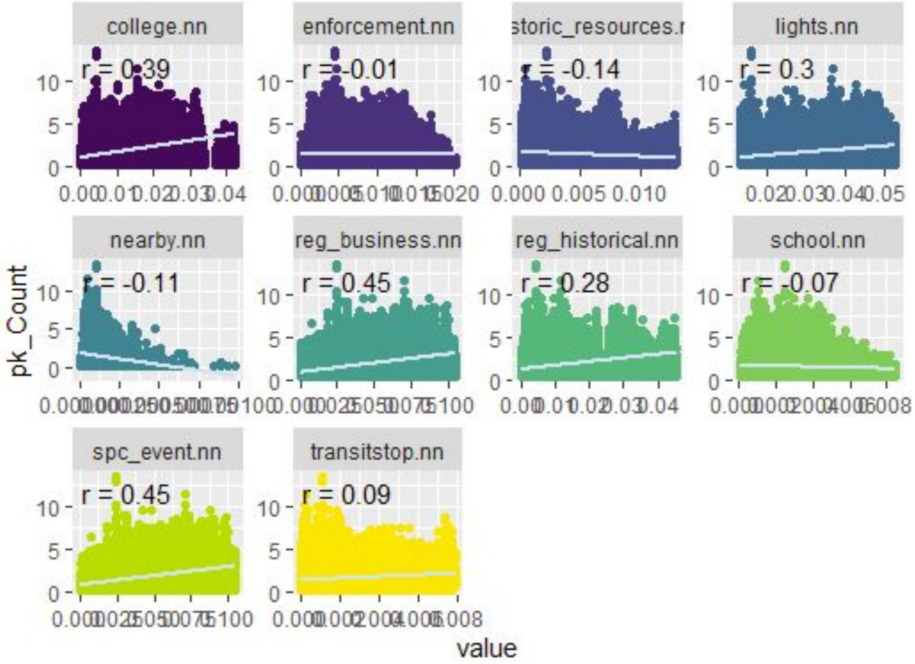
Transportation variables





Exploratory Analysis

Distance to nearest neighbors variables

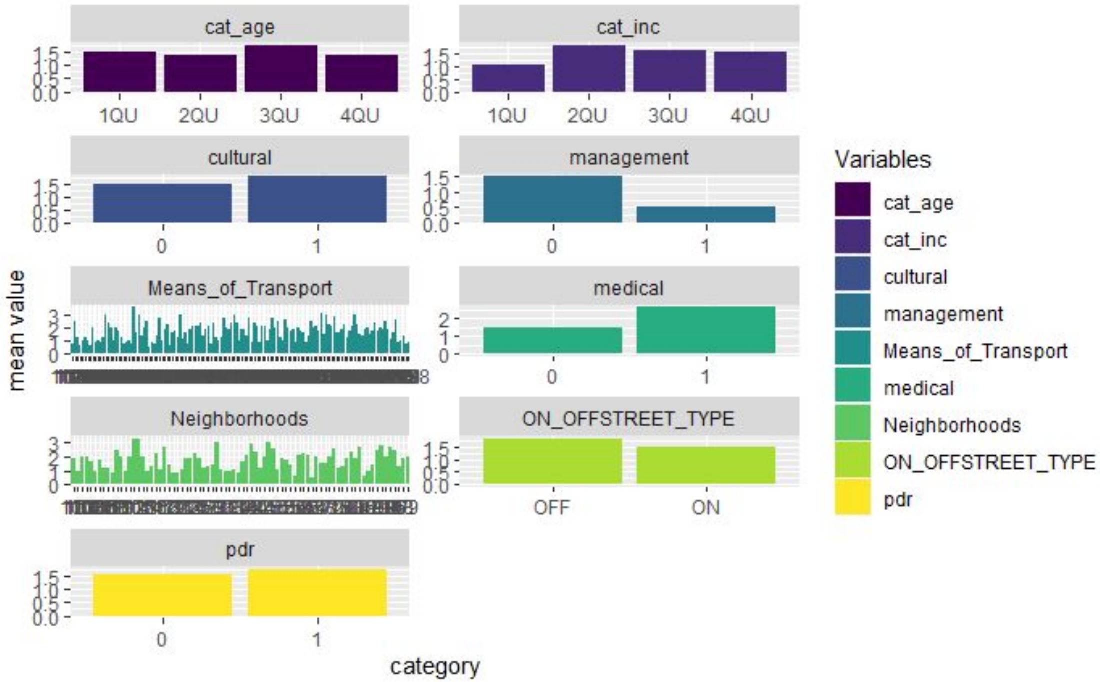


Use case

Pro-cess

Exploratory Analysis

Categorical features

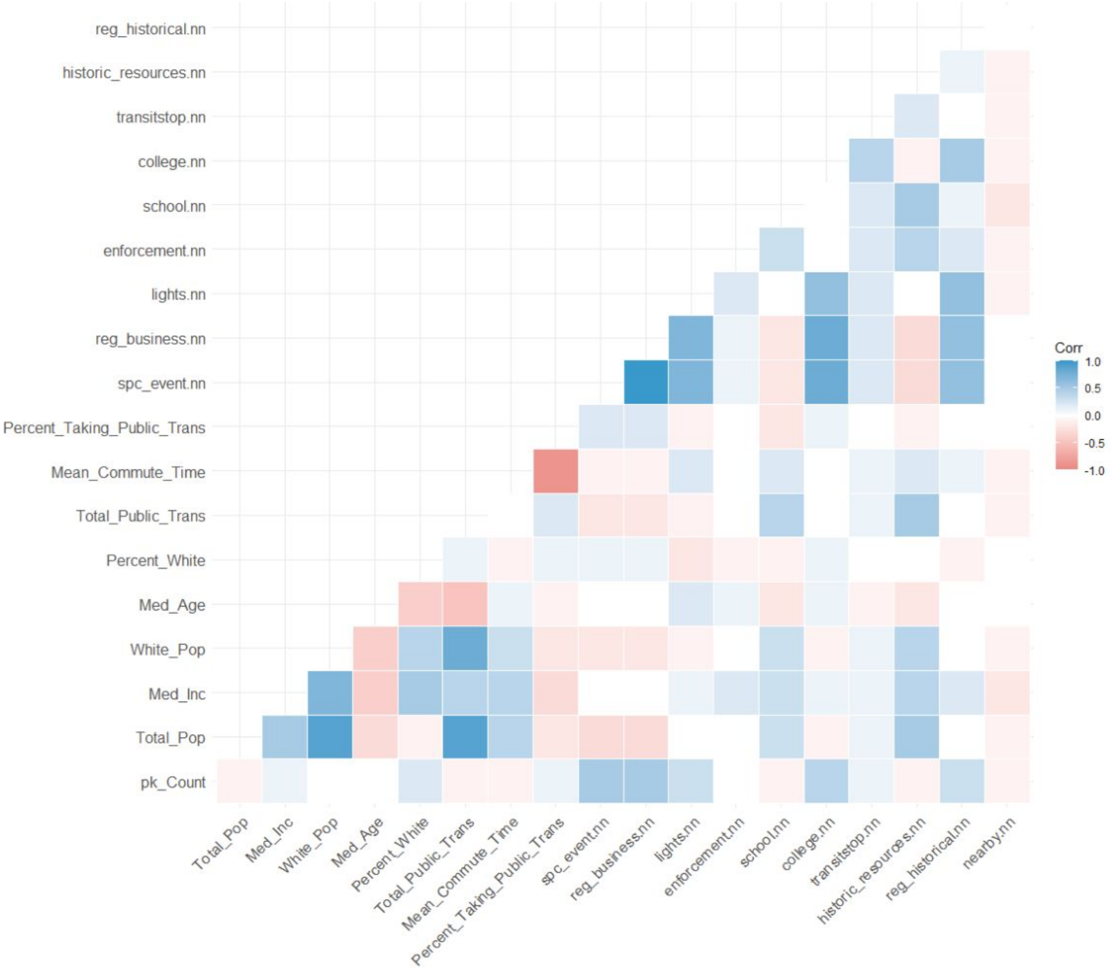


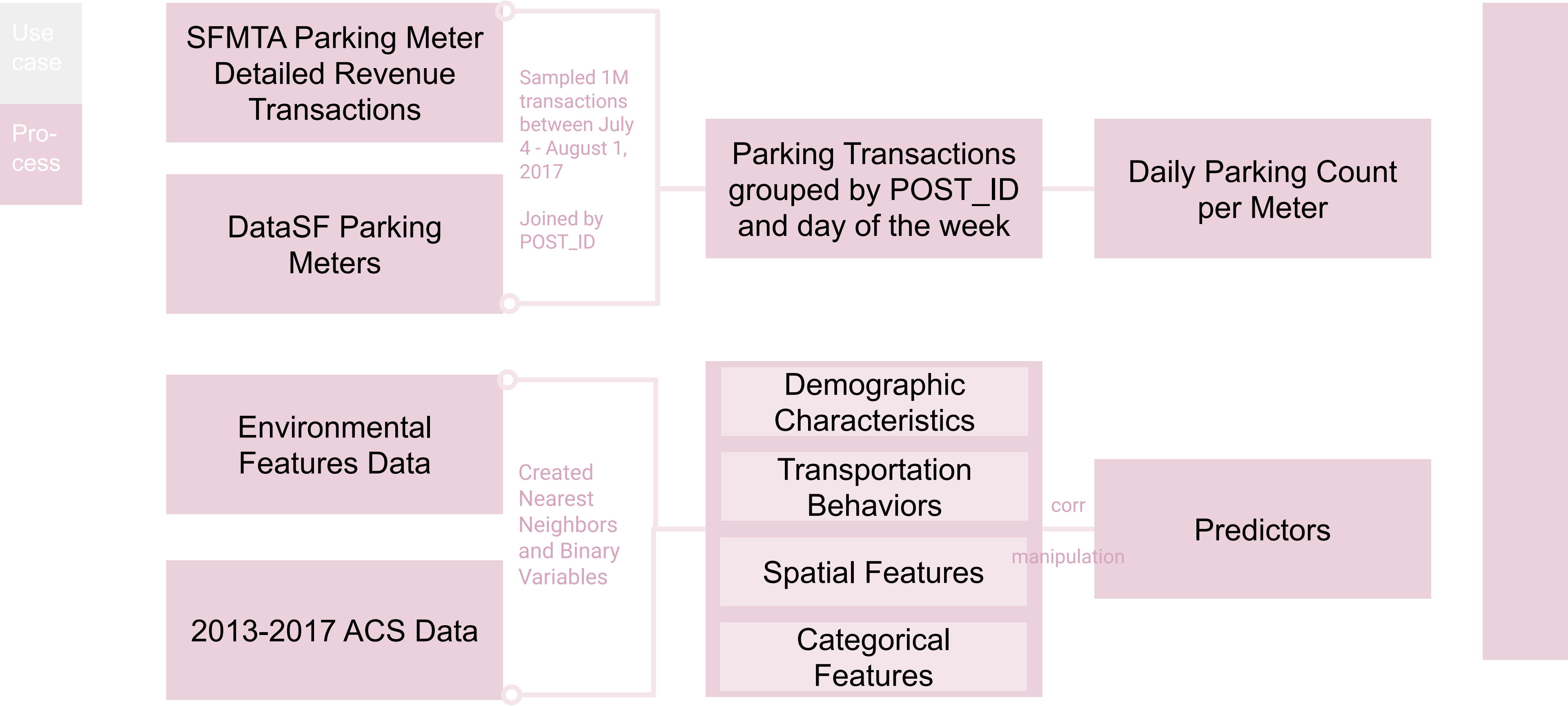
Use case

Pro-cess

Correlation

Between all continuous variables





Use case
Pro-cess
SFMTA Parking Meter Detailed Revenue Transactions
DataSF Parking Meters
Sampled 1M transactions between July 4 - August 1, 2017
Joined by POST_ID
Parking Transactions grouped by POST_ID and day of the week
Daily Parking Count per Meter
Environmental Features Data
2013-2017 ACS Data
Created Nearest Neighbors and Binary Variables
Demographic Characteristics
Transportation Behaviors
Spatial Features
Categorical Features
corr
manipulation
Predictors

Final variables

NUMERIC
After removing correlated variables
---->

BINARY (LAND USE)
PDR (Industrial (Production, Distribution, Repair))

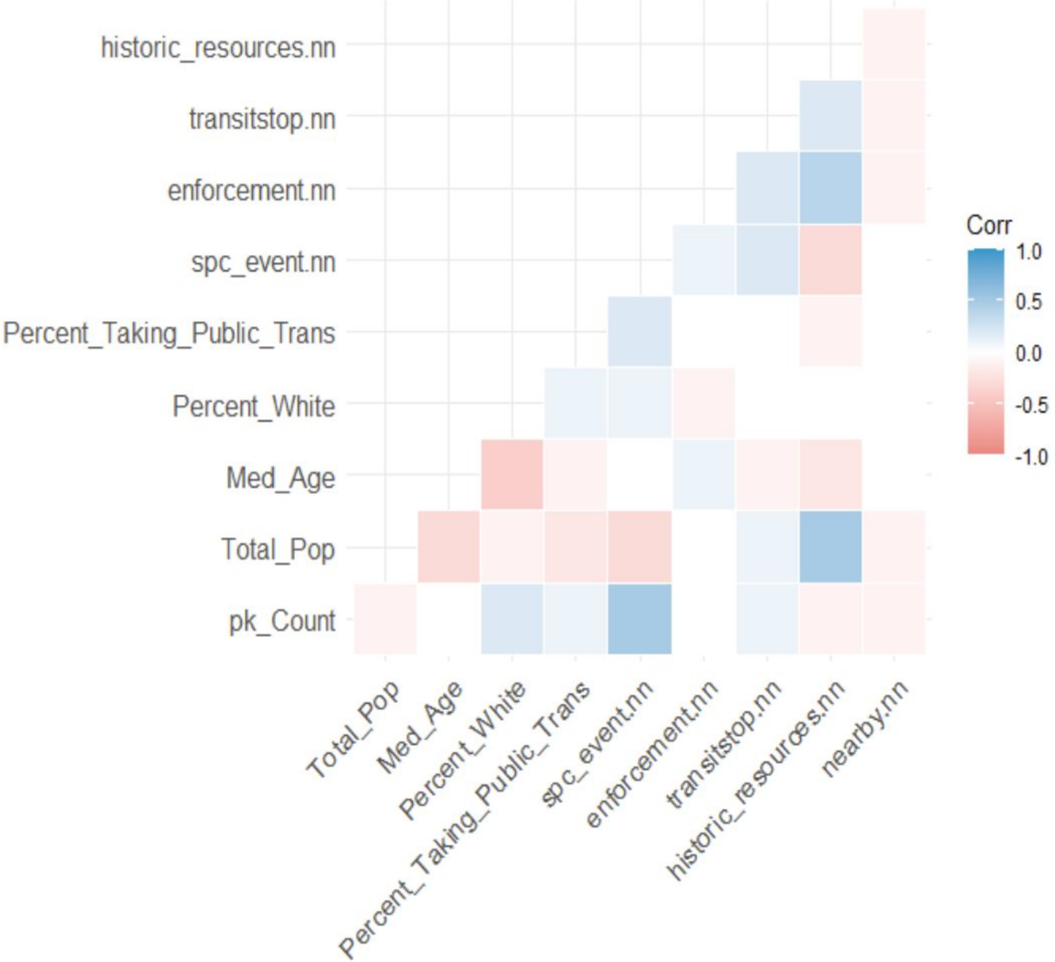
Cultural (Cultural, Institutional, Educational)

Medical

Management (Cultural, Institutional, Educational)

CATEGORICAL
Categorized median household income

Categorized median age



Use case

Pro-cess

OLS

```
pk_Count ~

+	Total_Pop
+	Percent_White
+	Percent_Taking_Public_Trans
+	spc_event.nn
+	enforcement.nn
+	transitstop.nn
+	historic_resources.nn
+	nearby.nn
+	ON_OFFSTREET_TYPE
+	cat_inc
+	cat_age
+	pdr
+	cultural
+	medical
+	management
+	dotw
```

Model building

OLS Regression model #1

Predicting daily parking transaction count per each meter

Use case

Pro-cess

OLS

Model building

OLS Regression model #2

Predicting daily parking transaction count per each meter

pk_Count ~

+ Total_Pop
+ Percent_White
+ Percent_Taking_Public_Trans
+ spc_event.nn
+ enforcement.nn
+ transitstop.nn
+ historic_resources.nn
+ nearby.nn
+ ON_OFFSTREET_TYPE
+ cat_inc
+ cat_age
+ ~~pdr~~
+ cultural
+ medical
+ ~~management~~
+ dotw

Model comparison

multivariate models

	Dependent variable: pk_Count (1)	(2)
Total_Pop	-0.00001***	-0.00001***
	(0.00000)	(0.00000)
Percent_White	0.351***	0.351***
	(0.024)	(0.024)
Percent_Taking_Public_Trans	0.081**	0.081**
	(0.039)	(0.039)
spc_event.nn	19.351***	19.351***
	(0.132)	(0.132)
enforcement.nn	-32.619***	-32.630***
	(1.130)	(1.130)
transitstop.nn	7.121***	7.138***
	(2.617)	(2.617)
historic_resources.nn	-15.543***	-15.538***
	(1.326)	(1.326)
nearby.nn	-2,678.771***	-2,679.260***
	(63.624)	(63.622)
ON_OFFSTREET_TYPEON	-0.298***	-0.298***
	(0.014)	(0.014)
pdr	0.023	
	(0.049)	
cat_inc2QU	0.151***	0.151***
	(0.010)	(0.010)
cat_inc3QU	0.282***	0.282***
	(0.010)	(0.010)
cat_inc4QU	0.352***	0.352***
	(0.013)	(0.013)
cat_age2QU	-0.125***	-0.125***
	(0.008)	(0.008)
cat_age3QU	0.027***	0.027***
	(0.009)	(0.009)
cat_age4QU	-0.067***	-0.067***
	(0.010)	(0.010)
cultural	0.453***	0.453***
	(0.054)	(0.054)
medical	0.568***	0.567***
	(0.208)	(0.208)
management	-0.458	
	(0.454)	
dotwMon	-0.069***	-0.069***
	(0.009)	(0.009)
dotwSat	0.018**	0.018**
	(0.009)	(0.009)
dotwSun	-0.547***	-0.547***
	(0.020)	(0.020)
dotwThu	-0.040***	-0.040***
	(0.009)	(0.009)
dotwTue	-0.056***	-0.056***
	(0.009)	(0.009)
dotwWed	-0.049***	-0.049***
	(0.009)	(0.009)
Constant	1.424***	1.425***
	(0.030)	(0.030)
Observations	156,523	156,523
R2	0.248	0.248
Adjusted R2	0.248	0.248
Residual Std. Error	1.015 (df = 156497)	1.015 (df = 156499)
F Statistic	2,069.943*** (df = 25; 156497)	2,249.896*** (df = 23; 156499)

Note: *p<0.1; **p<0.05; ***p<0.01

Coefficients:

	Estimate	Pr(>|t|)
(Intercept)	1.425e+00	< 2e-16 ***
Total_Pop	-1.155e-05	< 2e-16 ***
Percent_White	3.514e-01	< 2e-16 ***
Percent_Taking_ Public_Trans	8.074e-02	0.03940 *
spc_event.nn	1.935e+01	< 2e-16 ***
enforcement.nn	-3.263e+01	< 2e-16 ***
transitstop.nn	7.138e+00	0.00637 **
historic_resources.nn	-1.554e+01	< 2e-16 ***
nearby.nn	-2.679e+03	< 2e-16 ***
ON_OFFSTREET_TYPE	-2.981e-01	< 2e-16 ***
cat_inc2QU	1.505e-01	< 2e-16 ***
cat_inc3QU	2.815e-01	< 2e-16 ***
cat_inc4QU	3.523e-01	< 2e-16 ***
cat_age2QU	-1.253e-01	< 2e-16 ***
cat_age3QU	2.680e-02	0.00222 **
cat_age4QU	-6.707e-02	7.58e-12 ***
cultural	4.530e-01	< 2e-16 ***
medical	5.673e-01	0.00630 **
dotwMon	-6.881e-02	1.54e-14 ***
dotwSat	1.793e-02	0.04814 *
dotwSun	-5.469e-01	< 2e-16 ***
dotwThu	-3.982e-02	8.31e-06 ***
dotwTue	-5.618e-02	3.02e-10 ***
dotwWed	-4.905e-02	3.93e-08 ***

Interpretation

Continuous variables

↑ 1% white population	↑0.35 count
↑ 1m distance to nearest special event	↑19.35 count
↑ 1m distance to nearest transit stops	↑7.14 count
↑ 1m distance to nearest historic resources	↓15.54 count
↑ 1 m distance to enforcement	↓32.63 count
↑ 1 m distance to nearest meter	↓2679 count

Categorical variables

Higher the income, higher the parking demand
Areas with a median age around 35 have the highest parking demand
Medical land use has 0.56 higher parking demand compared with non-medical land use, controlling all other variables.

Interpretation

Total Population by Tract

Model Result

OLS Regression model

Predicting daily parking transaction count per each meter

```
pk_Count ~

+	Total_Pop
+	Percent_White
+	Percent_Taking_Public_Trans
+	spc_event.nn
+	enforcement.nn
+	transitstop.nn
+	historic_resources.nn
+	nearby.nn
+	ON_OFFSTREET_TYPE
+	cat_inc
+	cat_age
+	pdr
+	cultural
+	medical
+	management
+	dotw
```

Model building

Binomial Regression model #1

Predicting high-demand meters

Model building

```
pk_Count ~

+       Total_Pop
+       Percent_White
+       Percent_Taking_Public_Trans
+       spc_event.nn
+       enforcement.nn
+       transitstop.nn
+       historic_resources.nn
+       nearby.nn
+       ON_OFFSTREET_TYPE
+       cat_inc
+       cat_age
+       pdr
+       cultural
+       medical
+       management
+       dotw
```

Binomial Regression model #2

Predicting high-demand meters

Model comparison

Use case
Pro-cess
OLS
Bino mial

binomial models

	Dependent variable: as.factor(highDemand) (1)	(2)
Total_Pop	-0.00004***	-0.00004***
	(0.00000)	(0.00000)
Percent_White	1.095***	1.094***
	(0.055)	(0.055)
Percent_Taking_Public_Trans	0.045	0.045
	(0.089)	(0.089)
spc_event.nn	30.981***	30.983***
	(0.293)	(0.293)
enforcement.nn	-55.699***	-55.670***
	(2.493)	(2.493)
transitstop.nn	-20.724***	-20.710***
	(5.694)	(5.694)
historic_resources.nn	-42.308***	-42.250***
	(3.034)	(3.034)
nearby.nn	-4,947.835***	-4,948.238***
	(164.276)	(164.270)
ON_OFFSTREET_TYPEON	-0.200***	-0.198***
	(0.034)	(0.034)
pdr	-0.079	
	(0.116)	
cat_inc2QU	0.385***	0.385***
	(0.021)	(0.021)
cat_inc3QU	0.524***	0.523***
	(0.023)	(0.023)
cat_inc4QU	0.588***	0.588***
	(0.029)	(0.029)
cat_age2QU	-0.257***	-0.256***
	(0.018)	(0.018)
cat_age3QU	0.021	0.021
	(0.019)	(0.019)
cat_age4QU	-0.199***	-0.199***
	(0.022)	(0.022)
cultural	1.389***	1.390***
	(0.123)	(0.123)
medical	11.287	11.288
	(40.059)	(40.059)
management	-9.951	
	(88.038)	
dotwMon	-0.158***	-0.158***
	(0.020)	(0.020)
dotwSat	0.091***	0.091***
	(0.020)	(0.020)
dotwSun	-0.991***	-0.991***
	(0.054)	(0.054)
dotwThu	-0.052***	-0.052***
	(0.020)	(0.020)
dotwTue	-0.088***	-0.088***
	(0.020)	(0.020)
dotwWed	-0.091***	-0.091***
	(0.020)	(0.020)
Constant	-1.046***	-1.048***
	(0.069)	(0.069)
Observations	156,523	156,523
Log Likelihood	-86,602.930	-86,604.070
Akaike Inf. Crit.	173,257.900	173,256.100

Note: *p<0.1; **p<0.05; ***p<0.01

Use case

Pro-cess

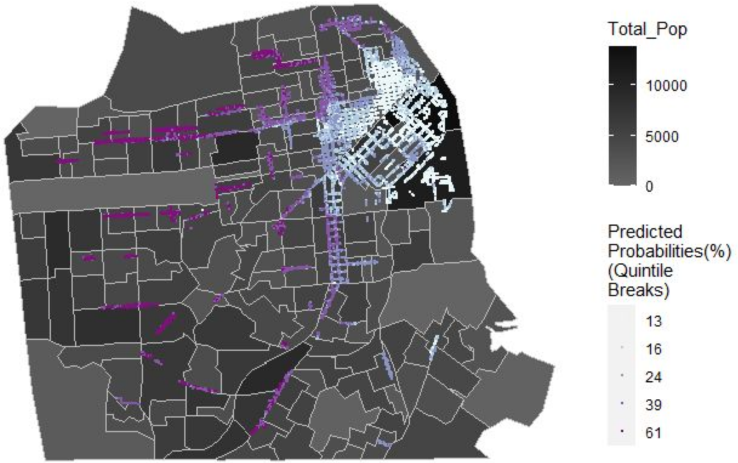
OLS

Bino mial

Model Result

Binomial Regression model

Predicting high-demand meters



Use case

Pro-cess

OLS

Bino mial

Evaluation

Confusion Matrix and Statistics

```
          Reference
Prediction     0     1
         0 67728 14761
         1 31786 42248

         Accuracy : 0.7026
           95% CI : (0.7003, 0.7049)
 No Information Rate : 0.6358
 P-Value [Acc > NIR] : < 2.2e-16

            Kappa : 0.3964

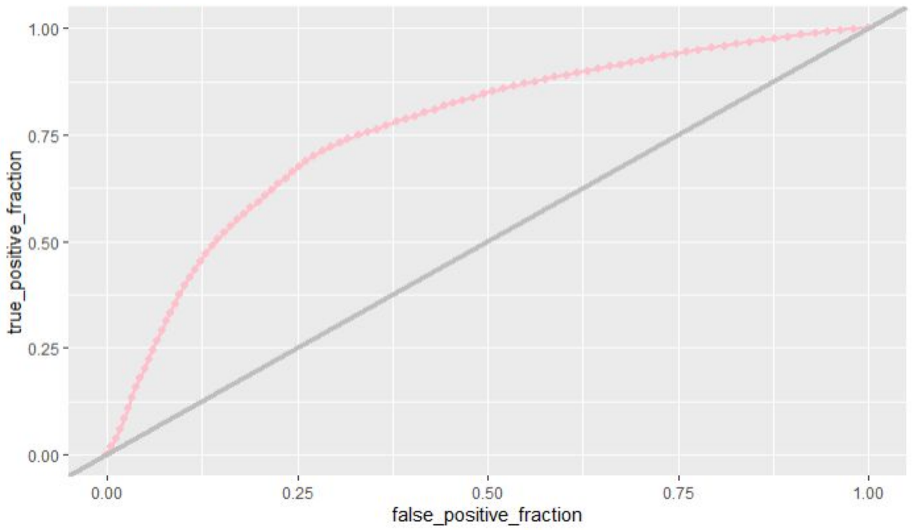
 Mcnemar's Test P-Value : < 2.2e-16

       Sensitivity : 0.7411
       Specificity : 0.6806
    Pos Pred Value : 0.5707
    Neg Pred Value : 0.8211
        Prevalence : 0.3642
    Detection Rate : 0.2699
 Detection Prevalence : 0.4730
 Balanced Accuracy : 0.7108

 'Positive' Class : 1

Area under the curve: 0.763
```

ROC Curve



Interpretation

```
```{r}
exp(coef(binomialModel1))
```

          (Intercept)   Total_Pop   Percent_White Percent_Taking_Public_Trans   spc_event.nn
         3.506330e-01   9.999626e-01   2.986430e+00   1.045560e+00   2.855971e+13
       enforcement.nn   transitstop.nn historic_resources.nn   nearby.nn   ON_OFFSTREET_TYPEON
         6.653378e-25   1.013361e-09   4.477626e-19   0.000000e+00   8.201234e-01
          cat_inc2QU   cat_inc3QU   cat_inc4QU   cat_age2QU   cat_age3QU
         1.469125e+00   1.687805e+00   1.799545e+00   7.739358e-01   1.020961e+00
          cat_age4QU   cultural   medical   dotwMon   dotwSat
         8.198126e-01   4.013164e+00   7.985880e+04   8.537268e-01   1.094900e+00
             dotwSun   dotwThu   dotwTue   dotwWed
         3.711858e-01   9.489124e-01   9.153746e-01   9.133831e-01
```

<u>Continuous variables</u>

↑ 1% total population	↑89.9% odds of high demand
↑ 1% white population	↑198% odds of high demand
↑ 1% population taking public transit	↑4.56% odds of high demand
↑ 1m distance to nearest special event	↑ 2.855971e+15% odds of high demand

<u>Categorical variables</u>

Higher the income, higher the odds of high demand
Areas with a median age around 35 , the highest parking demand
Cultural land use has 300% odds of being high demand compared with non-cultural land use, controlling all other variables.

Use case
Pro-cess
OLS
Binomial
Findings

Findings

	A: y=OLS	B: y=binomial
(Intercept)	1.42 ***	-1.05 ***
	(0.03)	(0.07)
Total_Pop	-0.00 ***	-0.00 ***
	(0.00)	(0.00)
Percent_White	0.35 ***	1.09 ***
	(0.02)	(0.05)
Percent_Taking_Public_Trans	0.08 *	0.04
	(0.04)	(0.09)
spc_event.nn	19.35 ***	30.98 ***
	(0.13)	(0.29)
enforcement.nn	-32.63 ***	-55.67 ***
	(1.13)	(2.49)
transitstop.nn	7.14 **	-20.71 ***
	(2.62)	(5.69)
historic_resources.nn	-15.54 ***	-42.25 ***
	(1.33)	(3.03)
nearby.nn	-2679.26 ***	-4948.24 ***
	(63.62)	(164.27)
ON_OFFSTREET_TYPEON	-0.30 ***	-0.20 ***
	(0.01)	(0.03)
cat_inc2QU	0.15 ***	0.38 ***
	(0.01)	(0.02)
cat_inc3QU	0.28 ***	0.52 ***
	(0.01)	(0.02)
cat_inc4QU	0.35 ***	0.59 ***
	(0.01)	(0.03)
cat_age2QU	-0.13 ***	-0.26 ***
	(0.01)	(0.02)
cat_age3QU	0.03 **	0.02
	(0.01)	(0.02)
cat_age4QU	-0.07 ***	-0.20 ***
	(0.01)	(0.02)
cultural	0.45 ***	1.39 ***
	(0.05)	(0.12)
medical	0.57 **	11.29
	(0.21)	(40.06)
dotwMon	-0.07 ***	-0.16 ***
	(0.01)	(0.02)
dotwSat	0.02 *	0.09 ***
	(0.01)	(0.02)
dotwSun	-0.55 ***	-0.99 ***
	(0.02)	(0.05)
dotwThu	-0.04 ***	-0.05 **
	(0.01)	(0.02)
dotwTue	-0.06 ***	-0.09 ***
	(0.01)	(0.02)
dotwWed	-0.05 ***	-0.09 ***
	(0.01)	(0.02)
N	156523	156523
R2	0.25	
adj.R2	0.25	
F_test	2249.90	

Model Comparison Table

Findings

The closer to tourist attractions the lower the demand. This might suggest that the Smart Meters method is working in curbing the demand as the prices there are usually higher.

Demand for parking on Sunday is significantly lower.

The ratio of white population in the area is highly associated with its odd of being a high parking demand area.

Higher the median household income, higher the odds of high parking demand in the area.

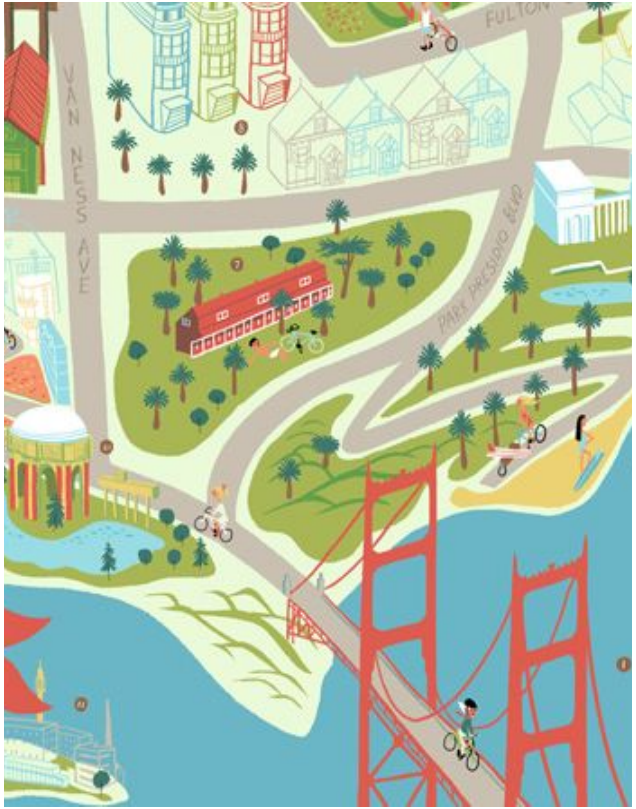
Lands that are identified as Cultural (Cultural, Institutional, Educational) and Medical show a distinct higher association with parking demand compared with other land uses.

Source: https://www.myparkingsign.com/blog/no-free-sunday-parking-san-francisco/

Discussion

- ➔ Low goodness of fit suggest room for influence from other features
  - such as time of the day and whether the meters are demand-responsive smart meters.

- ➔ OLS and binomial
  - IV's association with parking count does not necessarily align with their association with the odds of high parking demand.
  - For example, the distance to the nearest historical resources and to the parking enforcement instances are negatively associated with the parking counts but positively associated with the odds of high parking demand.
  - This may be caused by the limited parking space in these areas, which could indicate the demand of adding new meters.



Next steps

- Shorter time intervals
- Compare between years
- Map out the spatial and temporal mean absolute error (MAE) to identify areas with the highest error and find new variables to narrow the error

Modeling parking demand in San Francisco

Thank you for watching!

Please visit our github repo to see the R code
➔ https://github.com/gxzhao1/parkingSF

Bingchu Chen & “Gillian” Xuezhu Zhao
Planning by Numbers | Spring 2021