

Forecasting Street and Sidewalk Cleaning Services in San Francisco

Jared Schober and Peter Amerkhanian

San Francisco 311

“Street and sidewalk cleaning requests are generated internally and **through calls received by the City's 311 customer service center.** [...] Public Works' Radio Room **triages the request to the appropriate crew.**” ([source](#))

We examine call counts **at the hour level** between 1/1/2009 and 12/31/2022:

```
----- Descriptive Statistics -----
count      122712.000000
mean         17.626622
std          22.591759
min           0.000000
25%           2.000000
50%           8.000000
75%          25.000000
max          553.000000
Name: calls, dtype: float64
Sum: 2162998
```



Goals and Applications

Our goal: Use machine learning to predict the number of requests that the city will receive in any given hour.

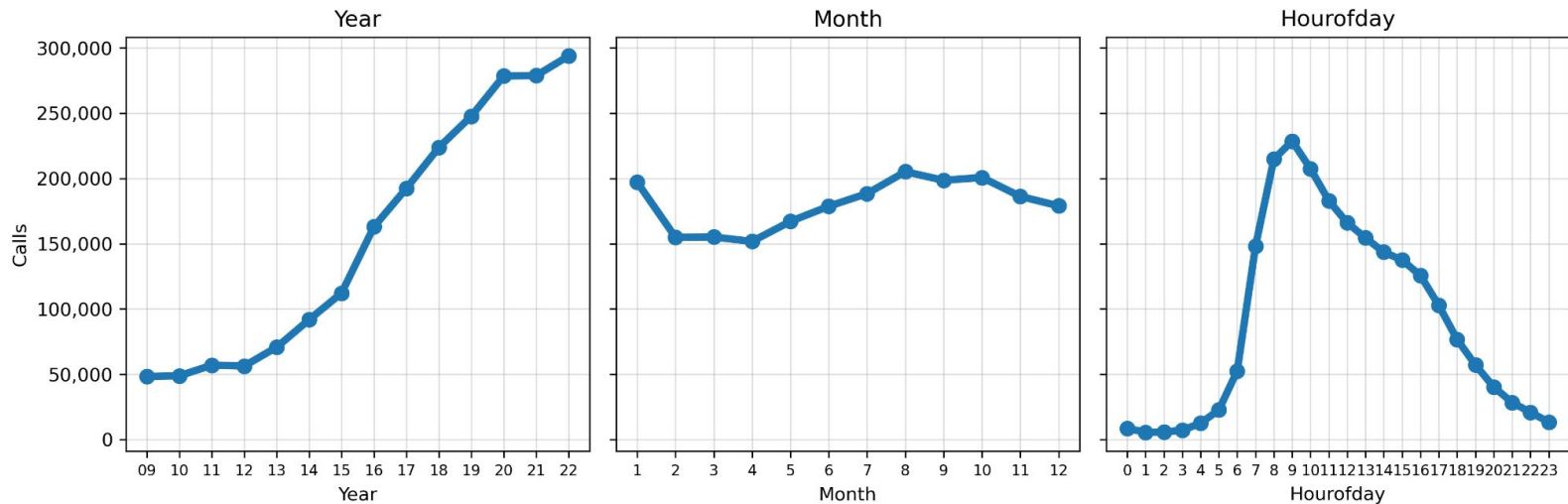
Policy applications: Understand how much staffing will be needed for the 311 call center and public works street cleaning teams at a given time, anticipate sidewalk and street cleaning costs.



[SF Public Works Clean Corridors](#)

Approach and Challenges

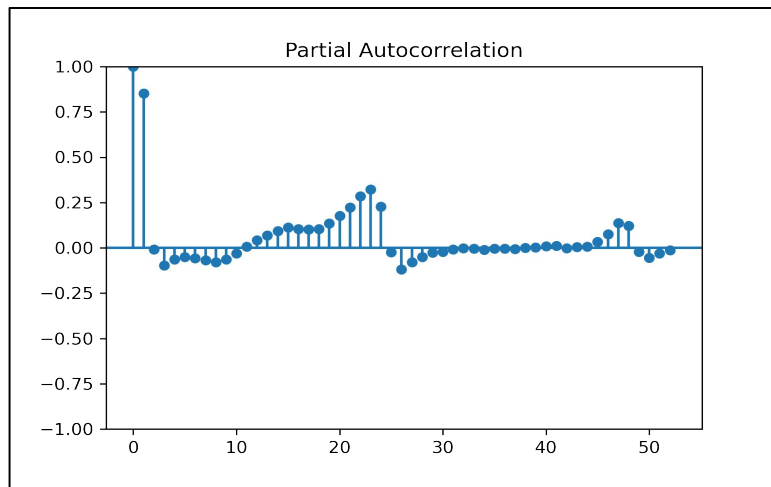
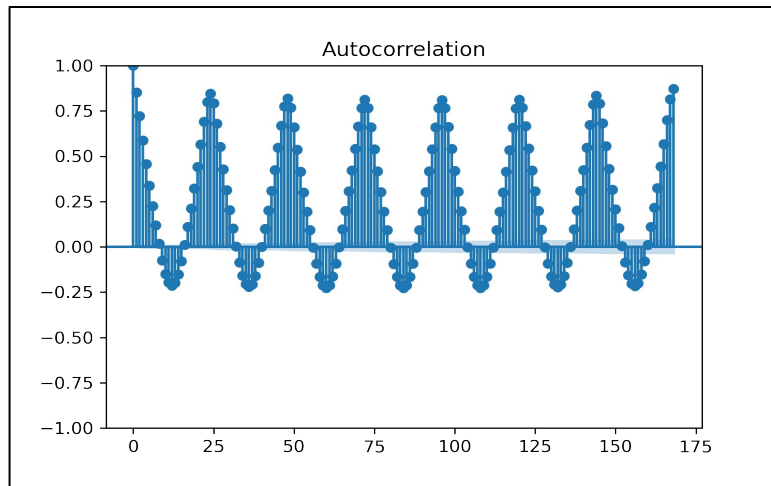
- **311 grew** considerably during this time period.
- We were concerned it would be difficult to forecast within year trends with such a strong secular trend across years.
- We focused on capturing hour-level patterns in our modeling



Feature engineering

- Autocorrelation - significant hourly at the **week level** (and beyond)
- Partial autocorrelation - significant hourly up to about **2 days**

Set up models to use 48 hours of previous call data.



Data and Cross Validation Approach

Lagged Dataset (48 hours)

	calls	lag_1	lag_2	lag_3	lag_4	lag_5	lag_6	lag_7	lag_8	lag_9	...	lag_39	lag_40	lag_41	lag_42	lag_43	lag_44	lag_45	lag_46	lag_47	lag_48
datetime																					
2009-01-03 00:00:00	0	6.0	0.0	1.0	1.0	3.0	5.0	19.0	20.0	19.0	...	6.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2009-01-03 01:00:00	0	0.0	6.0	0.0	1.0	1.0	3.0	5.0	19.0	20.0	...	4.0	6.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2009-01-03 02:00:00	1	0.0	0.0	6.0	0.0	1.0	1.0	3.0	5.0	19.0	...	10.0	4.0	6.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
2009-01-03 03:00:00	3	1.0	0.0	0.0	6.0	0.0	1.0	1.0	3.0	5.0	...	8.0	10.0	4.0	6.0	1.0	0.0	0.0	0.0	0.0	0.0
2009-01-03 04:00:00	0	3.0	1.0	0.0	0.0	6.0	0.0	1.0	1.0	3.0	...	5.0	8.0	10.0	4.0	6.0	1.0	0.0	0.0	0.0	0.0

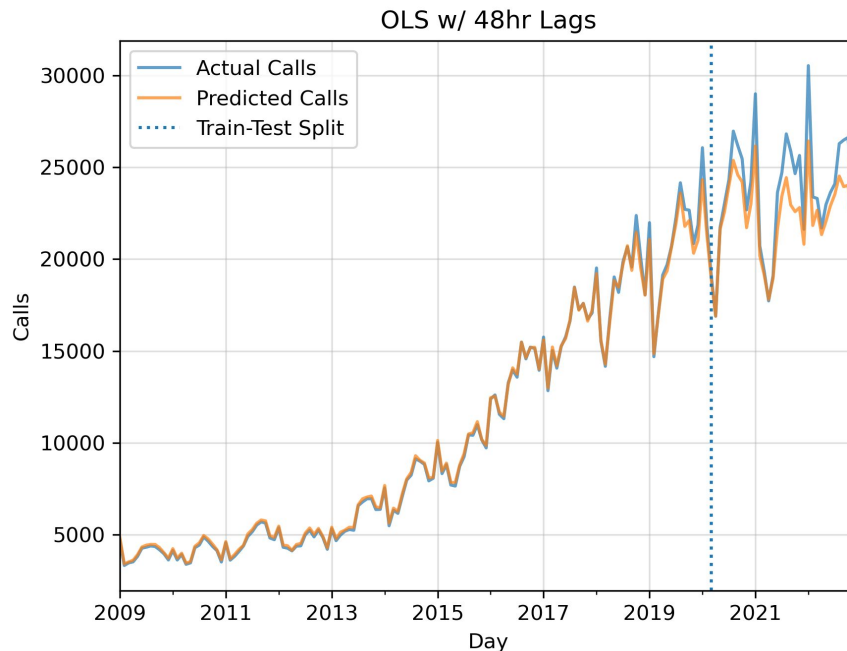
Cross Validation: **Train**, **Test**, standardize, clip at $\pm 3\text{std}$, compute metrics

etc...

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Fold Size	8712	8760	8760	8784	8760	8760	8760	8784	8760	8760	8760	8784	8760	8760

Next-Hour Forecasting

- Given the last 48 hours of data, how many street and sidewalk cleaning calls will come in the next hour?
- This short-term forecasting would be used to prepare dispatchers and cleaning crews



(80-20 train test split pictured for clarity)

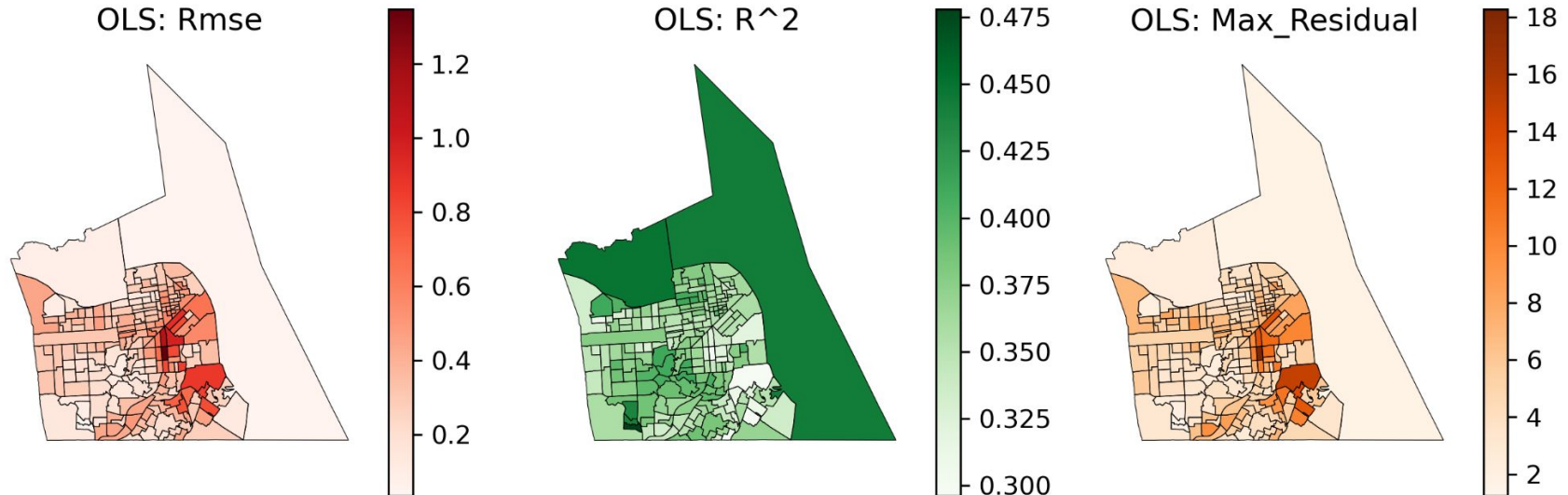
Next-Hour Performance Summary

Model	Avg RMSE	Avg R ²	Avg max residual
Baseline	12.58	0.379	171
Linear regression	8.29	0.595	157.33
Lasso Regression ($\alpha=0.0001$)	8.28	0.594	157.30
Random Forest (<i>max_depth=None, estimators=200, etc.</i>)	7.95	0.606	153.60
RNN	62.13	0.18	398.465

Average metrics are computed across year-folds. All models have optimized hyperparameters for each fold's predictions

Next-Hour Neighborhood-level Forecasting

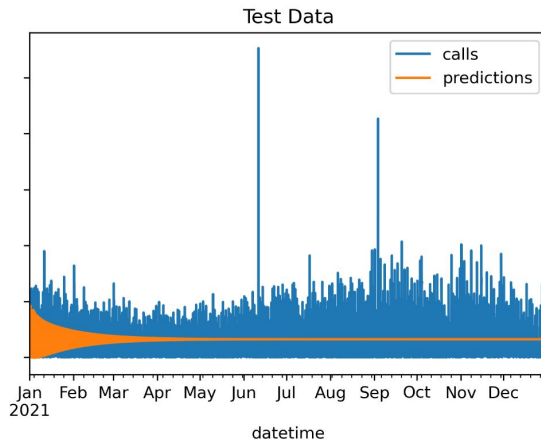
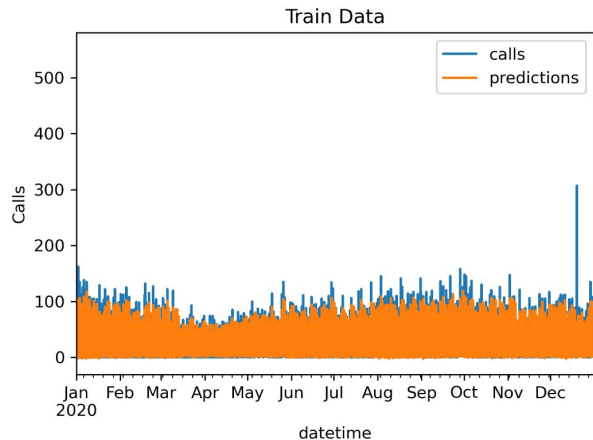
- Given the last 48 hours of data, how many street and sidewalk cleaning calls will come in the next hour, **and where will they be?**
- These forecasts would be used to dispatch cleaning crews preemptively



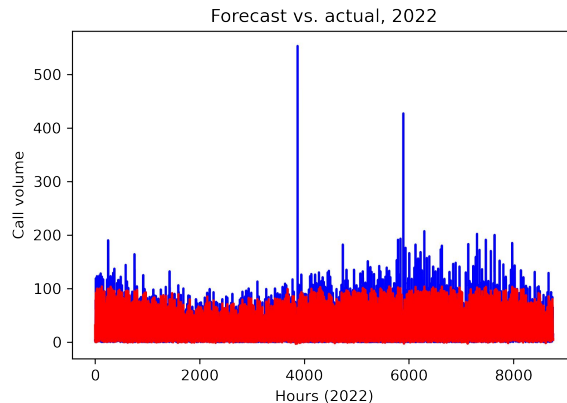
Long-Run Forecasting

- Given the last 48 hours of data, how many street and sidewalk cleaning calls will **come in the next year**?
- This is an exploration of how far we can predict trends before quality degrades significantly

OLS



ARIMA



Long Run Performance Summary

Model	Avg RMSE	Avg R ²	Avg max residual
Baseline	22.97	-0.75	179.33
Linear regression	17.46	-0.041	166.23
Random Forest (<i>max_depth=None,</i> <i>estimators=200, etc.</i>)	20.277	-0.394	170.54
ARIMA	11.37	0.56	161.376

Average metrics are computed across year-folds. All models have optimized hyperparameters for each fold's predictions