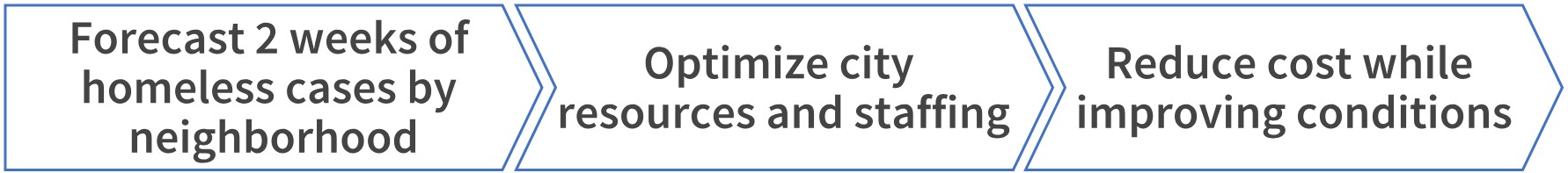
Forecasting San Francisco's Homeless Crisis

Harmeet Hora

Motivation

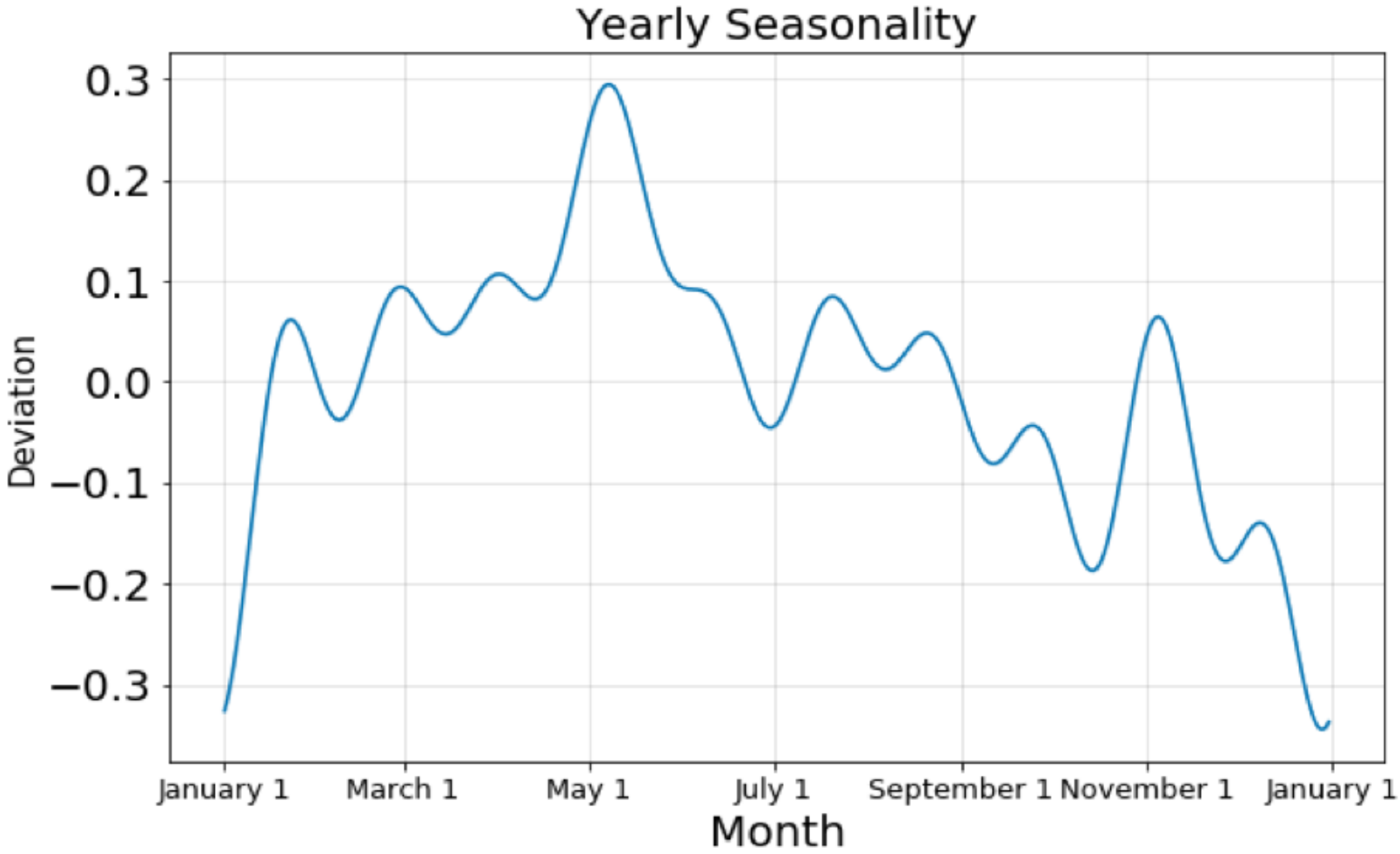
- $250 million spent annually
- Optimize resource allocation
- Predict expected cases by day & neighborhood

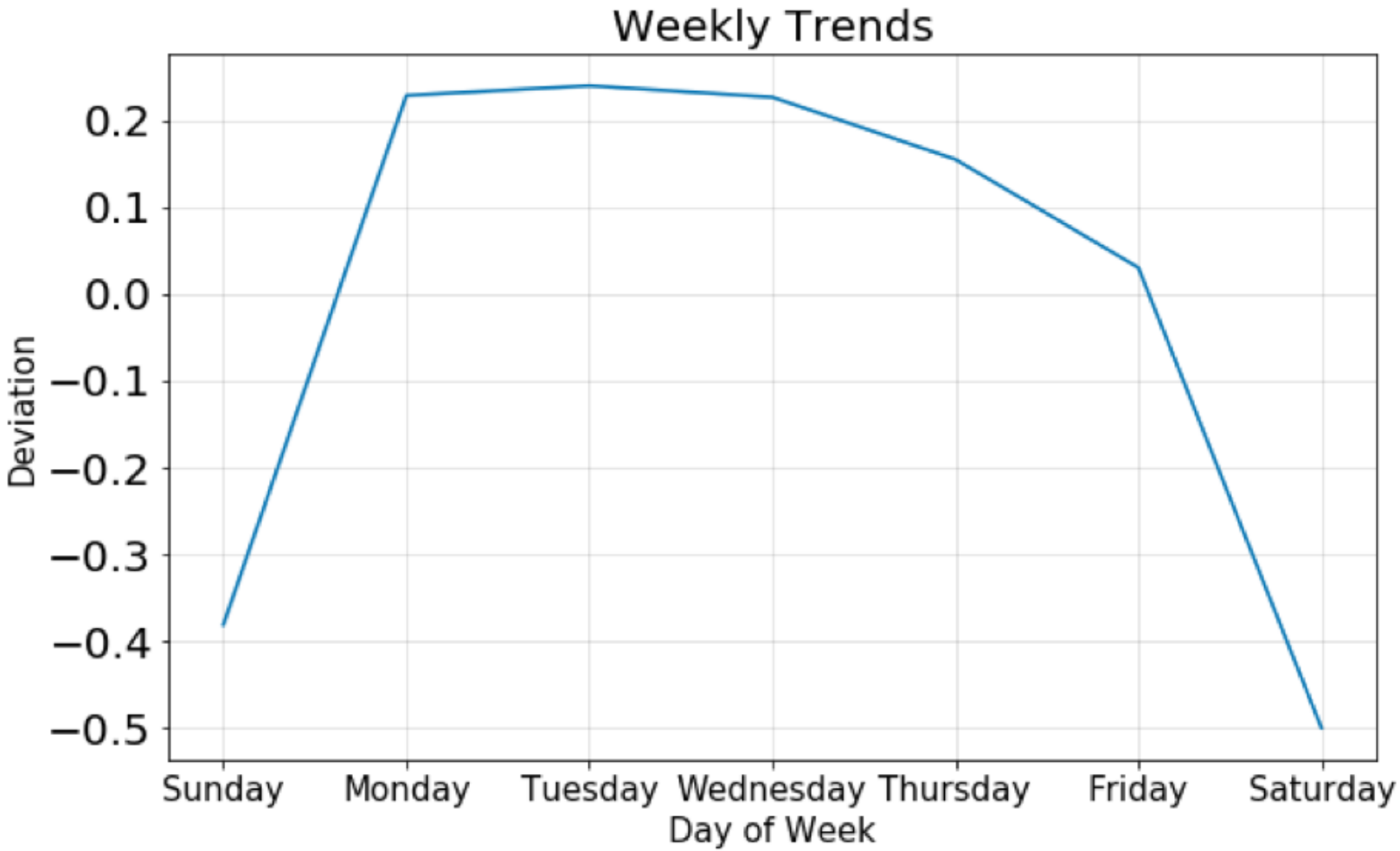
Use Case

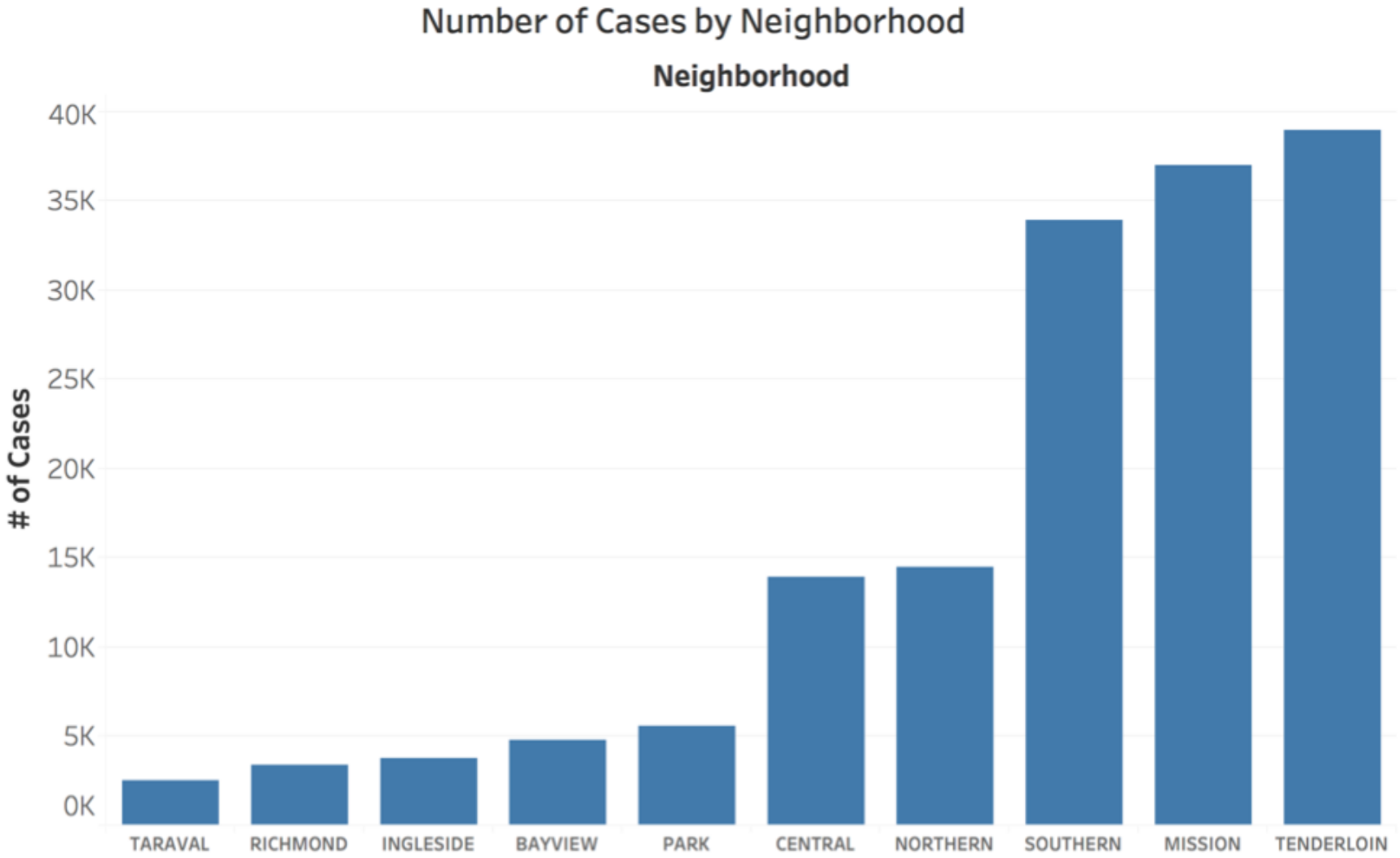

Forecast 2 weeks of homeless cases by neighborhood
Optimize city resources and staffing
Reduce cost while improving conditions

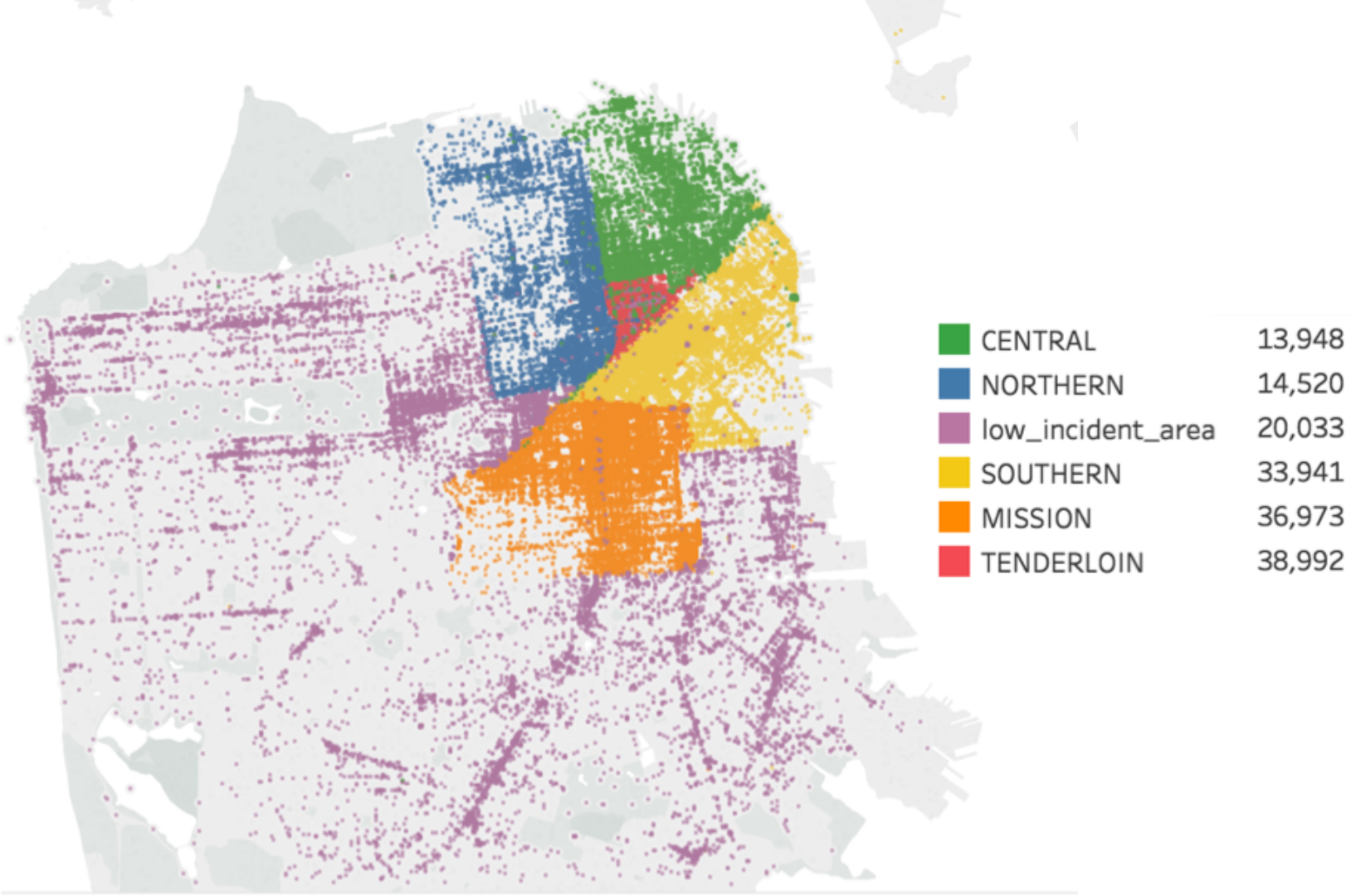
158,409

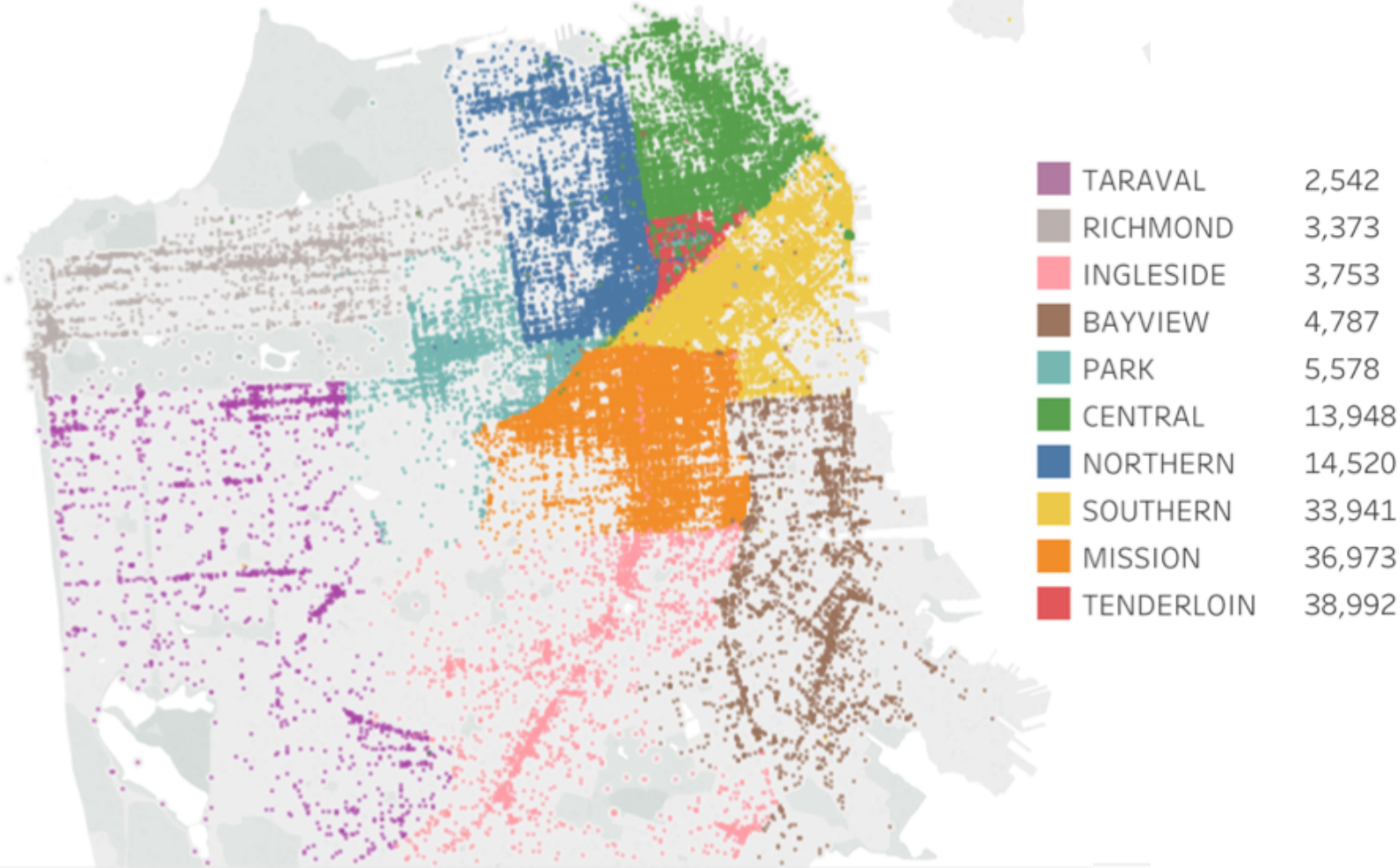
Homeless related 311 cases over 10 years


Yearly Seasonality
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
Deviation
January 1
March 1
May 1
July 1
September 1
November 1
January 1
Month


Weekly Trends
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
Deviation
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of Week


Number of Cases by Neighborhood
Neighborhood
# of Cases
40K
35K
30K
25K
20K
15K
10K
5K
0K
TARAVAL
RICHMOND
INGLESIDE
BAYVIEW
PARK
CENTRAL
NORTHERN
SOUTHERN
MISSION
TENDERLOIN


CENTRAL 13,948
NORTHERN 14,520
low_incident_area 20,033
SOUTHERN 33,941
MISSION 36,973
TENDERLOIN 38,992


TARAVAL 2,542
RICHMOND 3,373
INGLESIDE 3,753
BAYVIEW 4,787
PARK 5,578
CENTRAL 13,948
NORTHERN 14,520
SOUTHERN 33,941
MISSION 36,973
TENDERLOIN 38,992

Method

- ARIMA Forecasting
- Linear Regression
- Facebook Prophet

Model Performance

Neighborhood	Best Model	Mean of test set	MAE	% of Incidents
Tenderloin	ARIMA (6,1,7)	15.4	2.35	24%
Mission	ARIMA(14,1,2)	22.6	2.91	23%
Southern	Random Forest	24	2.75	21%
LIA	Random Forest	17.8	1.93	13%
Central	OLS	14.6	1.46	10%
Northern	Random Forest	14.2	1.57	9%

Next Steps

Incident type

Increase precision of location

LSTM Neural Network

Thank You!
I am Harmeet Hora

You can reach me at:
harmeethora@gmail.com
linkedin.com/in/harmeet-hora/