Predicting Crimes in Philadelphia

Fiona Chin, Erik Enriquez, Yumeng (Rainy) Niu

Problem Statement

- Predict the number of different crimes in Philadelphia using different indicators
- Our indicators include weather data (i.e temperature, snowfall, precipitation...etc) and city statistical indicators (i.e unemployment rates, cancer mortality, homicide mortality..etc)

Data

- Our Philadelphia crime data was extracted from OpenDailyPhilly.org containing yearly data from 2006 to 2016 and consists of 14 features and 2,154,123 examples of different crime incidents
- The yearly statistic indicators were extracted from the Community Health Assessment containing yearly data from 2000 to 2012 (15 examples, 8 features)
- Our weather dataset comes from weather.gov and contains data from 2006 to 2015

Cleaning and Reshaping

- NA Values - 0.08% had missing values and no significant outliers
- Reshaping - Aggregating by months and type of crime
- Joining Data Sets - Weather Statistics and City Indicators
- Selecting 2006 - 2012 Data
  - 2006 - 2011 Data is used for Train set
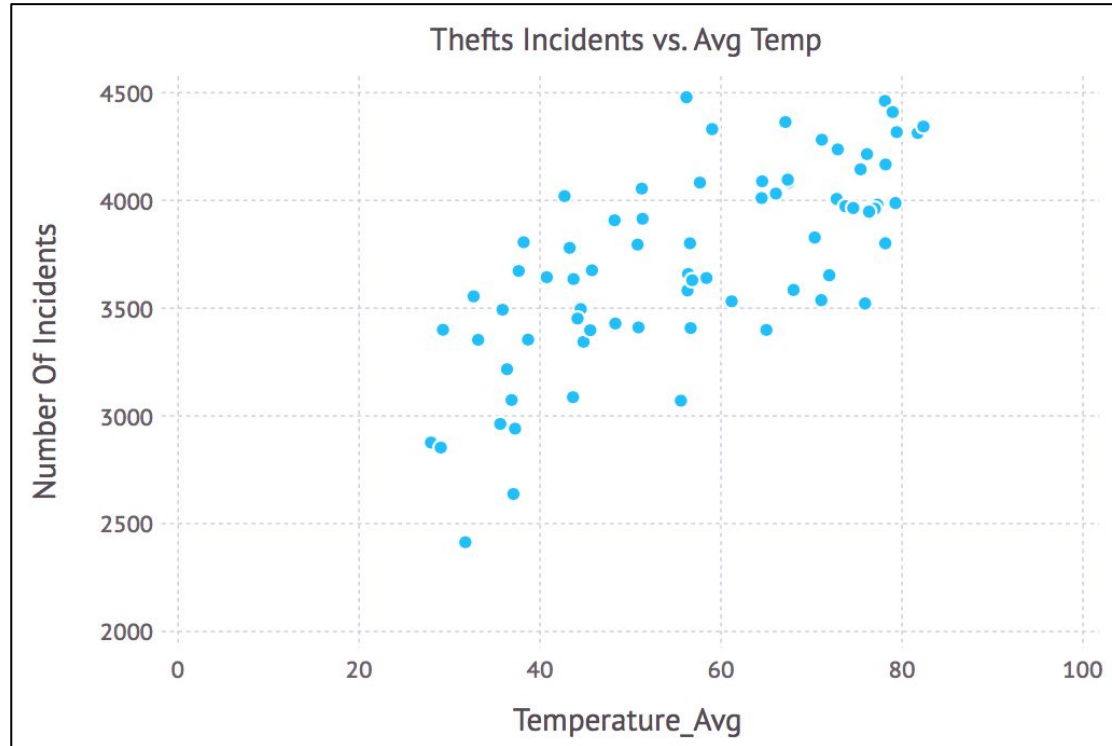 - 2012 is used for Test set (Auto-Regression)
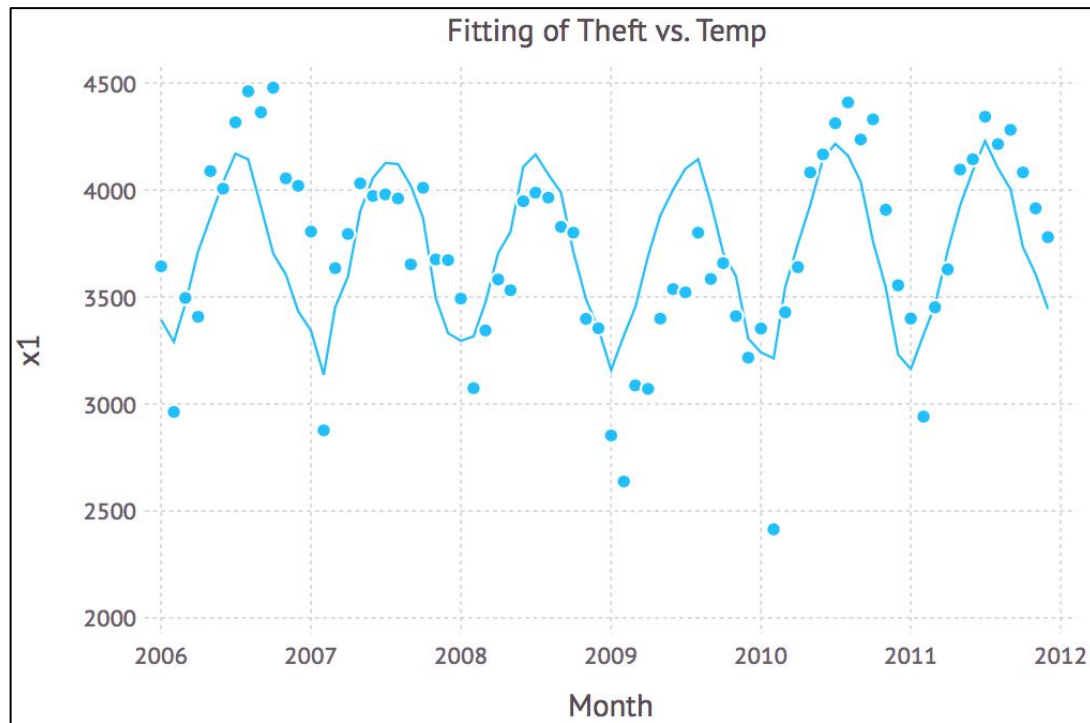
Modeling and Results - Data Exploration

Number of Incident By Month and Type
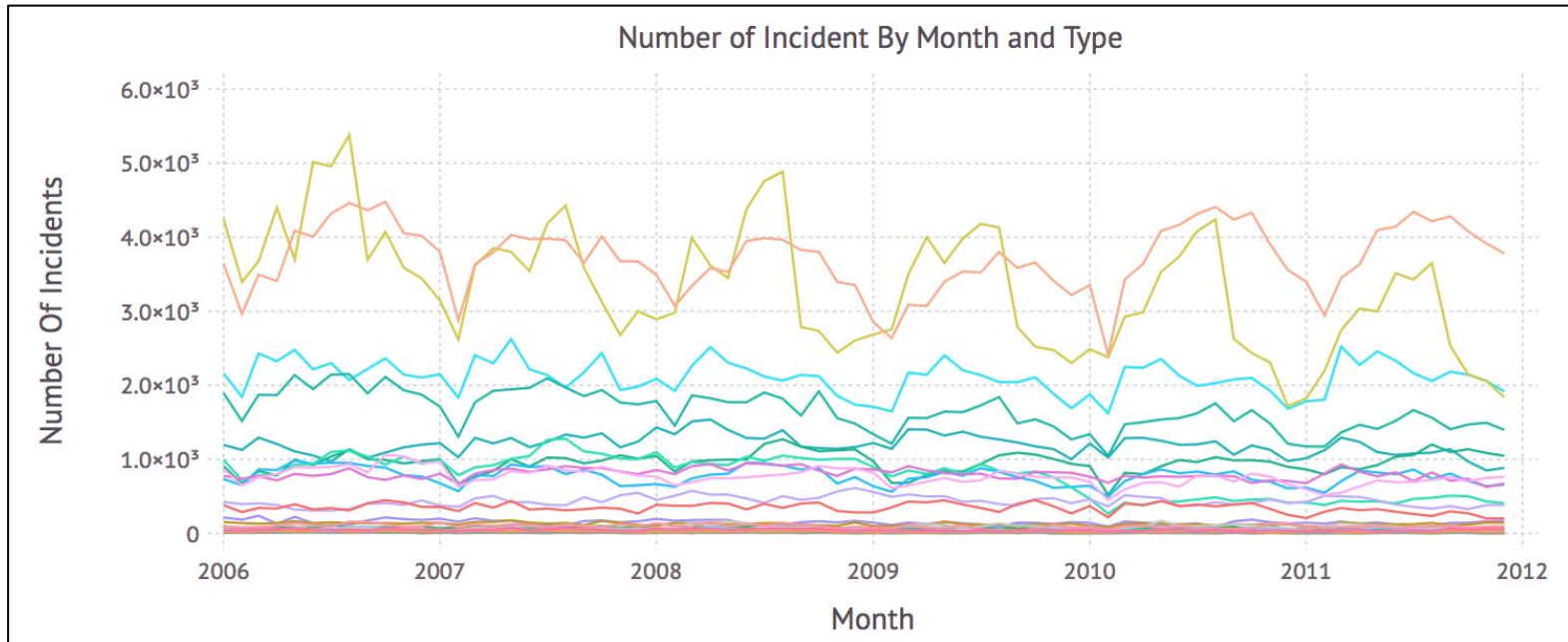
Modeling and Results - Seasonal Trend

Thefts By Month
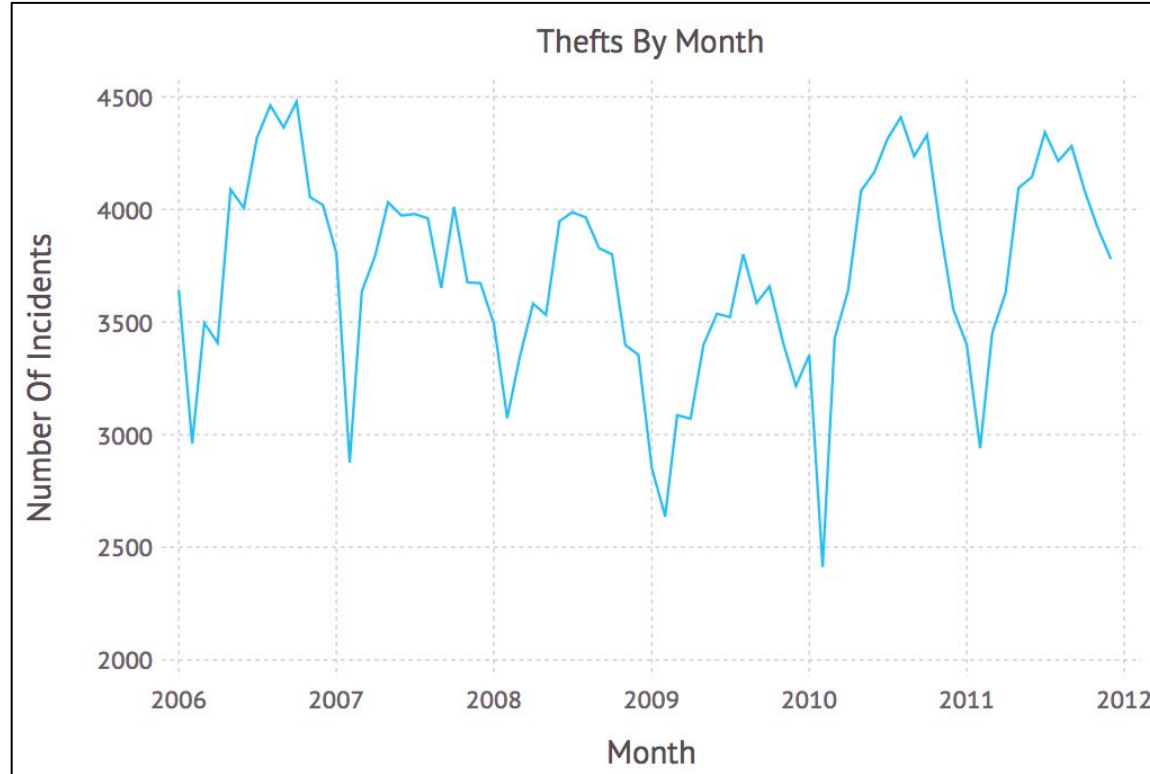
Modeling and Results – vs. Temperature

Thefts Incidents vs. Avg Temp

Modeling and Results – Model 1
Theft and Temperature

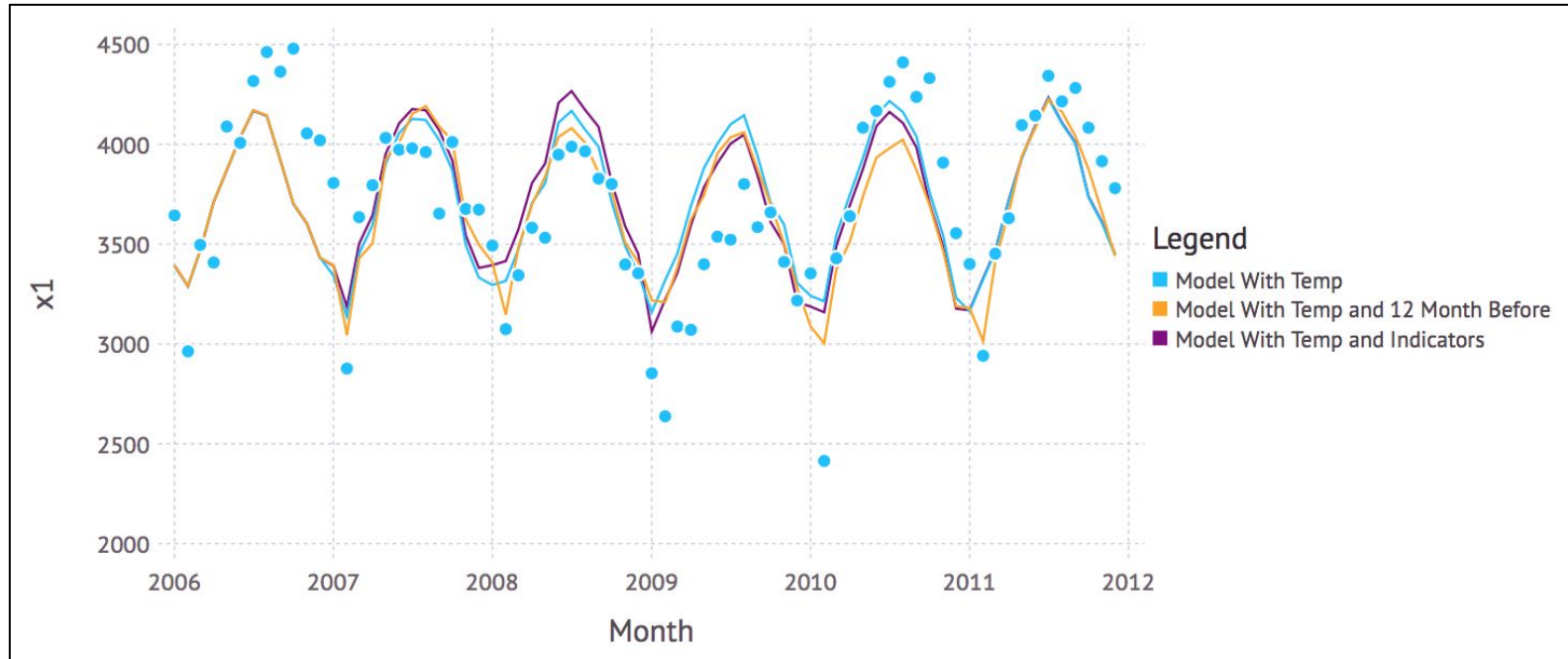Fitting of Theft vs. Temp

Modeling and Results – Model 2
Theft, Temperature, and 12 Month Auto-Regression
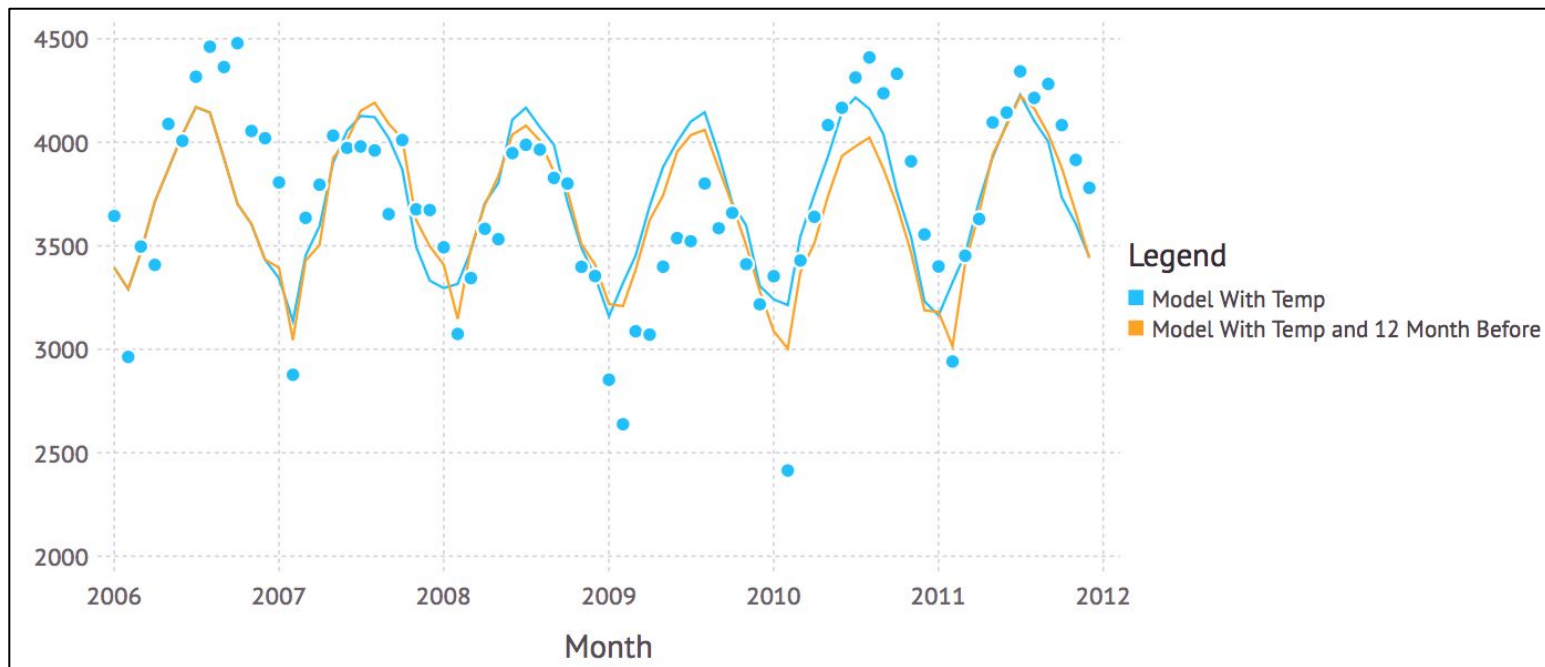
Modeling and Results – Model 3
Theft, Temperature, and City Indicators

Future Study