Forecasting Bike Accidents in PDX

Author: Ryan Harper

Why?

Portland, OR

Bike Friendly

Biking areas

77 mi Neighborhood Greenways

188 mi Bike lanes

85 mi Bike paths

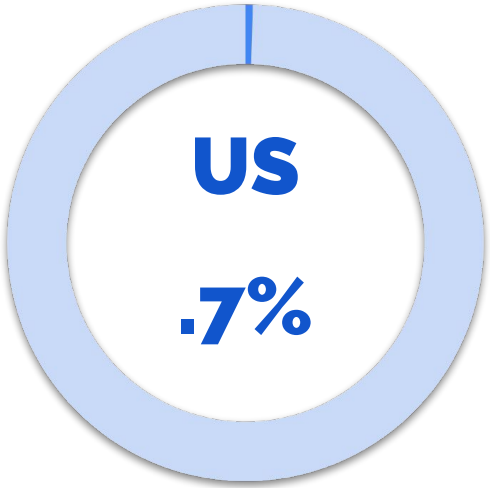
Bike Parking

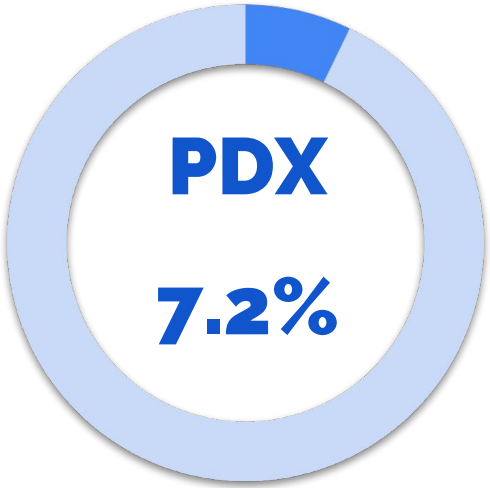
6500 Public Bike Racks

134 Bike Corrals

[PDOT]

commuters who bike



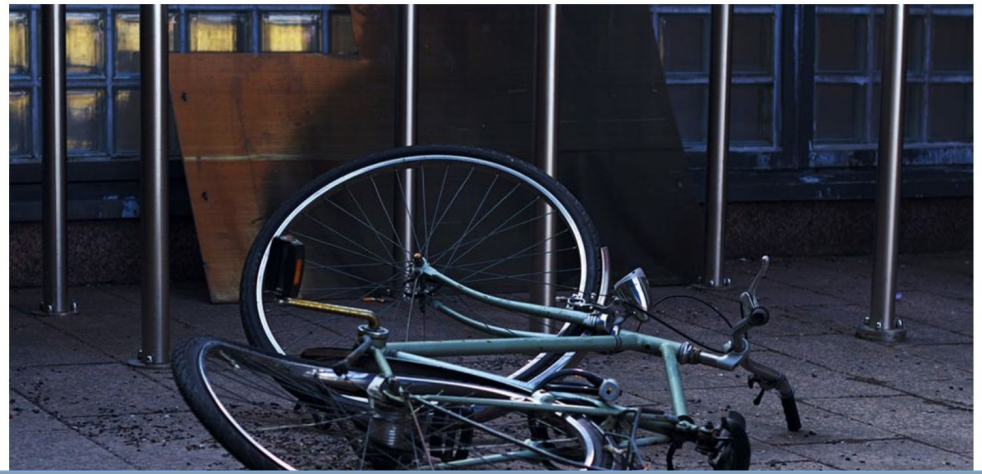


[PDOT]

Portland Cyclist Just Wanted to Ride His Bike Naked. Now He's Suing the City for $566,000.

THE RIDER SUFFERED SERIOUS INJURIES AFTER CRASHING INTO CEMENT BARRIERS IN THE ROAD AS HE RETURNED (FULLY CLOTHED) FROM THE CITY'S WORLD NAKED BIKE RIDE LAST YEAR.

BY MOLLY HURFORD FEB 5, 2018



MORE FROM

News

Methodology

1

Explore

Clean, graph, and visualize the timeline data

2

Trends and Seasonality

Check for directional trends in the data and look for any seasonal patterns

3

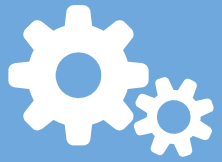
Modeling

Evaluate different regression models via log likelihood, error metrics, and lag plots.

4

Forecasting

Find the best model for predicting weekly accident count in 2016



1

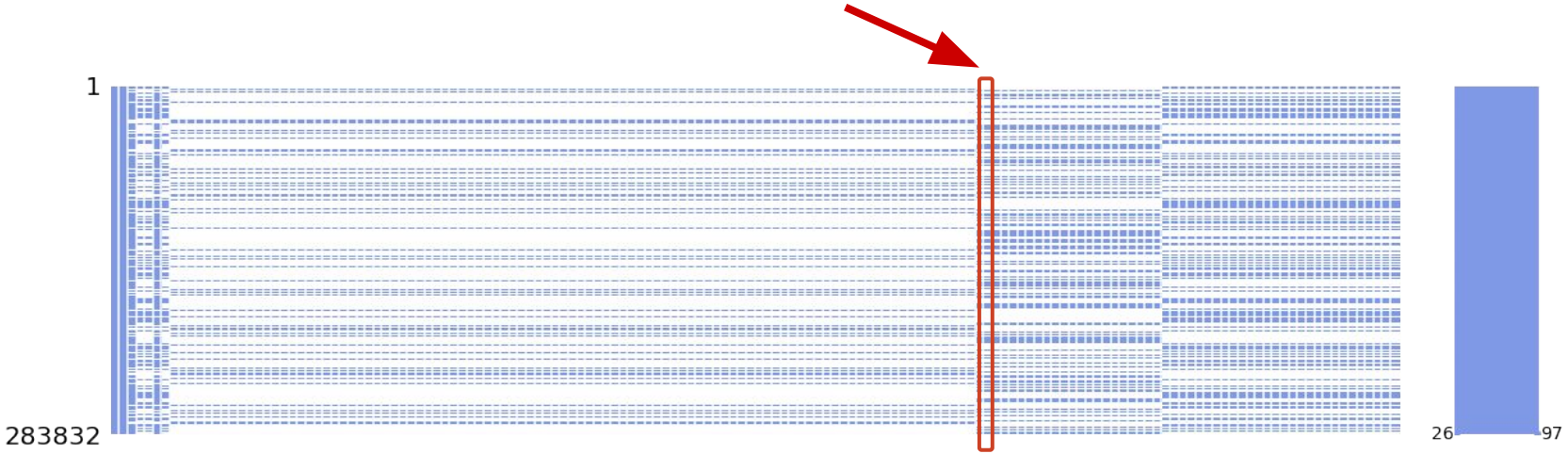
Explore

ODOT Multnomah Dataset

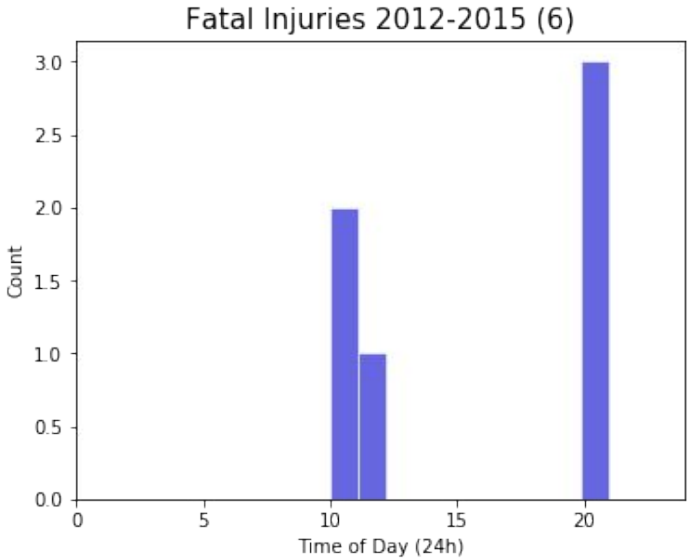
Crash ID | Record Type | Vehicle ID | Participant ID | Participant Display Seq# | Vehicle Coded Seq# | Participant Vehicle Seq# | Serial # | Crash Month | Crash Day | Crash Year | Week Day Code | Crash Hour | County Code | City Section ID | Urban Area Code | Functional Class Code | NHS Flag | Highway Number | Highway Suffix | Roadway Number | Highway Component | Mileage Type | Connection Number | Linear Reference System (LRS) | Latitude Degrees | Latitude Minutes | Latitude Seconds | Longitude Degrees | Longitude Minutes | Longitude Seconds | Special Jurisdiction | Jurisdiction Group | Street Number | Nearest Intersecting Street Number | Intersection Sequence Number | Distance from Intersection | Direction From Intersection | Milepoint | Posted Speed Limit | Road Character | Off Roadway Flag | Intersection Type | Intersection Related Flag | Roundabout Flag | Driveway Related Flag | Number of Lanes | Number of Turning Legs | Median Type | Impact Location | Crash Type | Collision Type | Crash Severity | Weather Condition | Road Surface Condition | Light Condition | Traffic Control Device (TCD) | TCD Functional Flag | Investigating Agency | Crash Level Event 1 Code | Crash Level Event 2 Code | Crash Level Event 3 Code | Crash Level Cause 1 Code | Crash Level Cause 2 Code | Crash Level Cause 3 Code | Work Zone Indicator | Alcohol-Involved Flag | Drugs Involved Flag | Speed Involved Flag | Population Range Code | Road Control | Route Type | Route Number | Region ID | District ID | Segment Marker ID | Segment Point LRS Measure | Unlocatable Flag | Total Vehicle Count | Total Fatality Count | Total Serious Injury (Inj-A) Count | Total Moderate Injury (Inj-B) Count | Total Minor Injury (Inj-C) Count | Total Non-Fatal Injury Count | Total Count of Un-Injured Children Age 00-04 | Total Count of Un-Injured Persons | Total Pedestrian Count | Total Pedestrian Fatality Count | Total Pedestrian Non-Fatal Injury Count | Total Pedalcyclist Count | Total Pedalcyclist Fatality Count | Total Pedalcyclist Non-Fatal Injury Count | Total Unknown Non-Motorist Count | Total Unknown Non-Motorist Fatality Count | Total Unknown Non-Motorist Injury Count | Total Vehicle Occupant Count | Total Count of Persons Involved | Total Quantity of Persons Using Safety Equipment ...

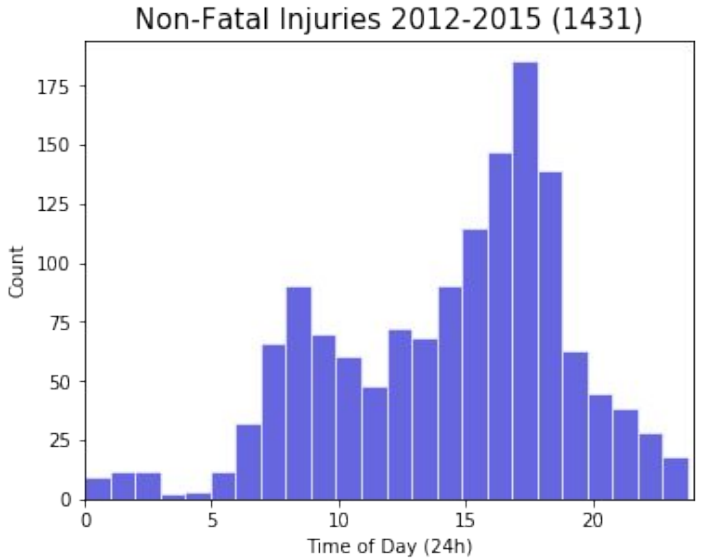
SO MUCH DATA!

ODOT Multnomah Dataset

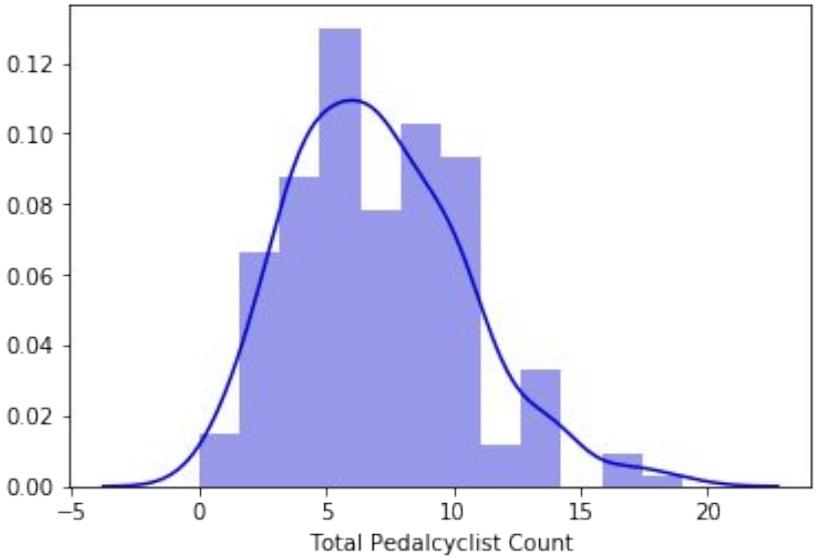


Distribution:
Hourly Accidents

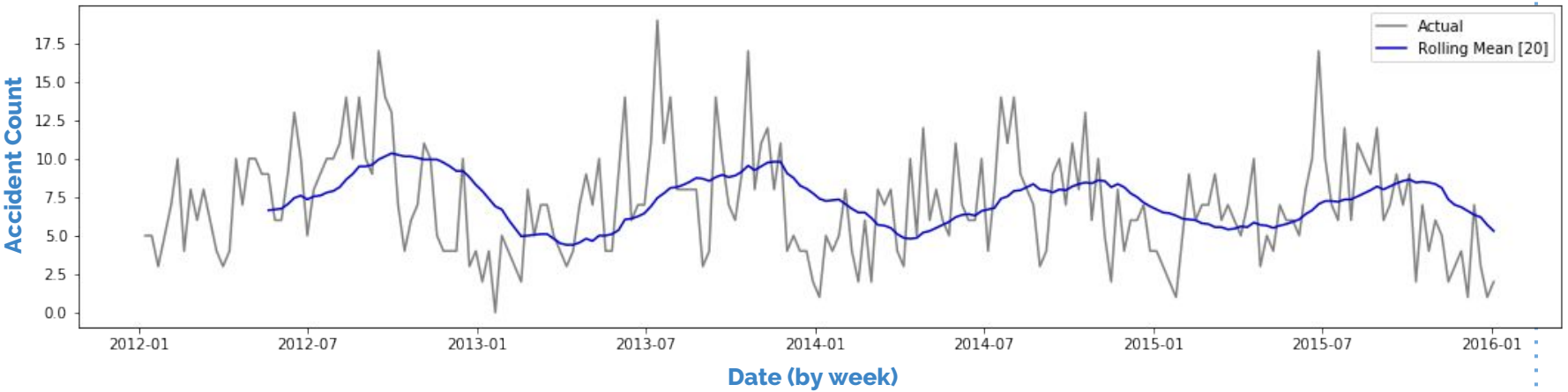




Distribution:
Total Accident Count for 1 Week



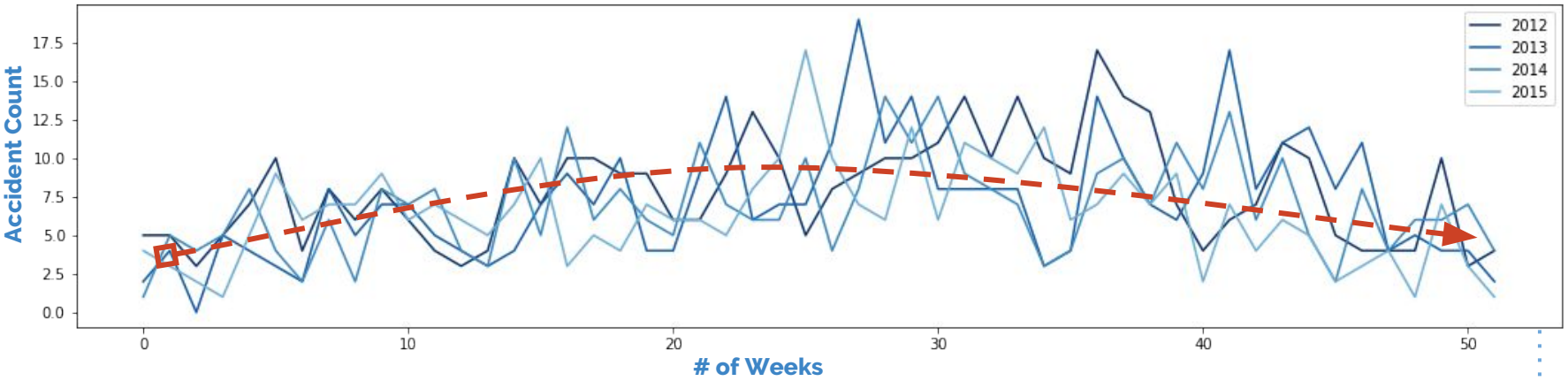
Bike Accidents 2012-2015



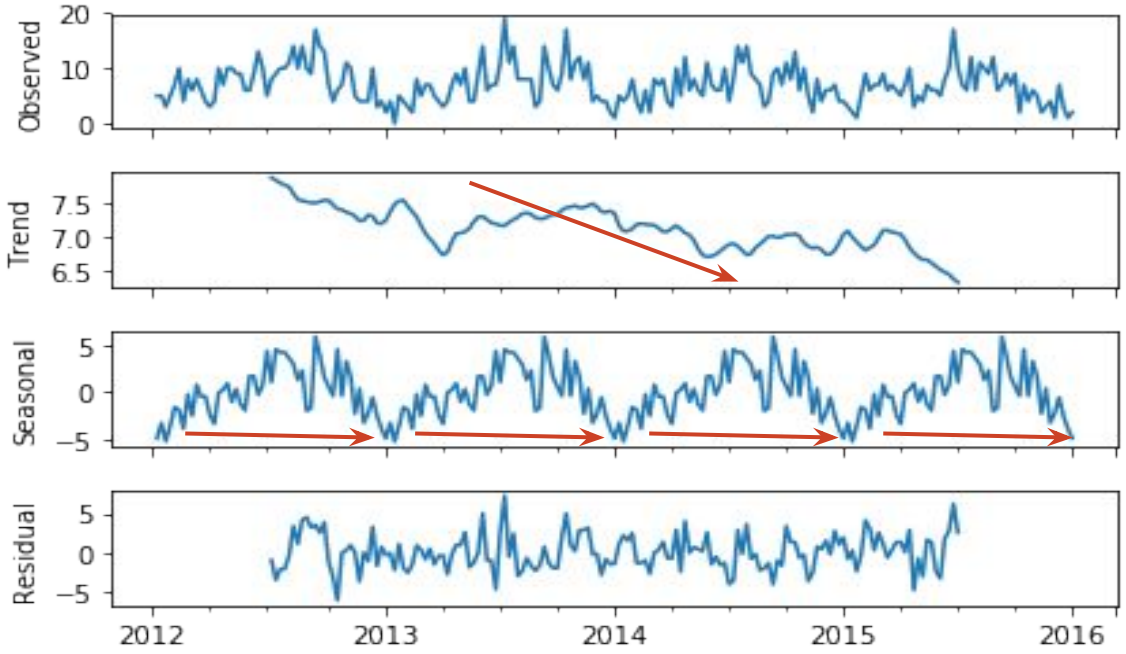
2

Trends and Seasonality

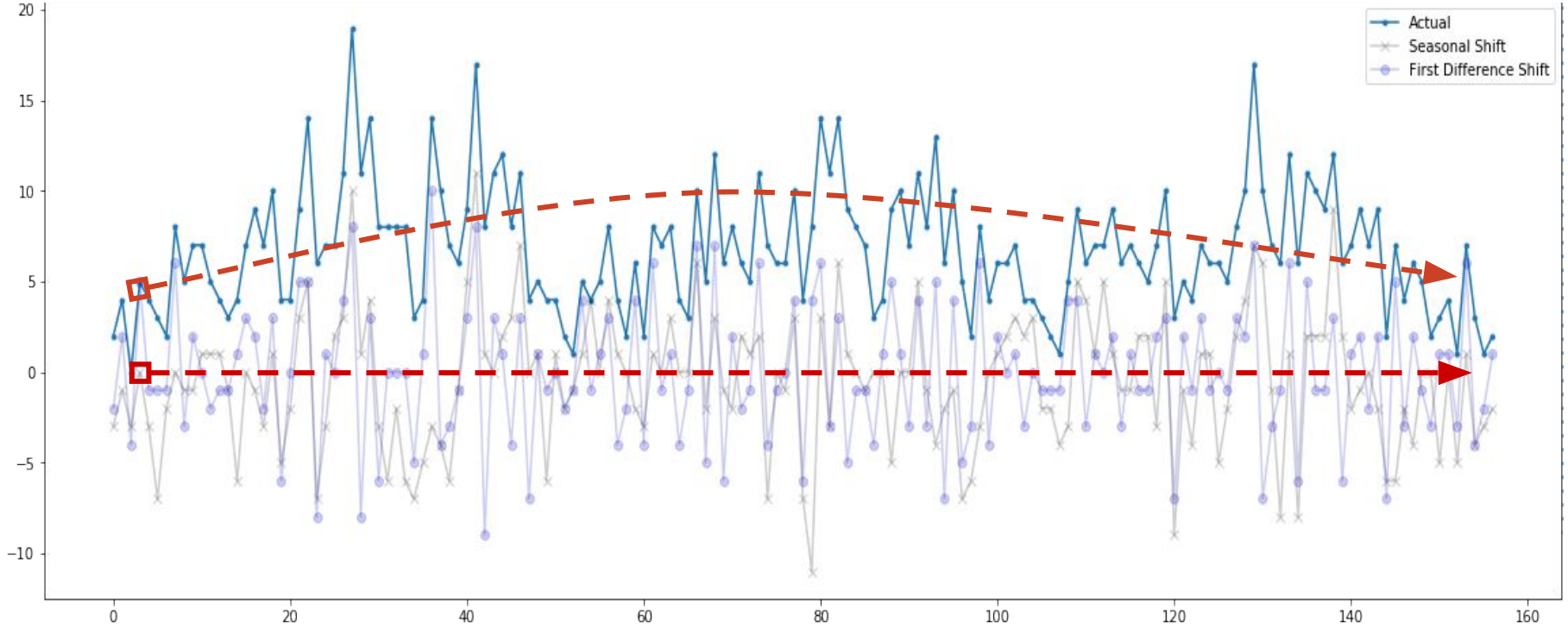
Accidents by year

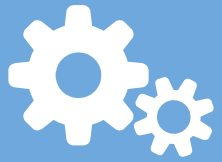


Decomposition



Stationarity



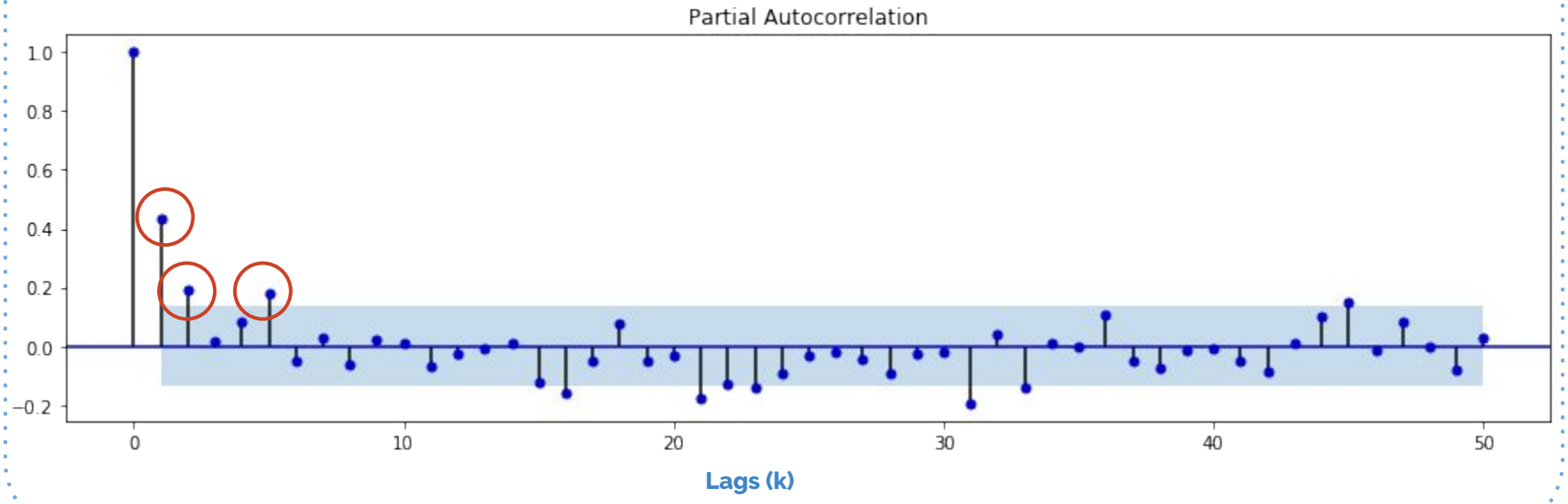


3

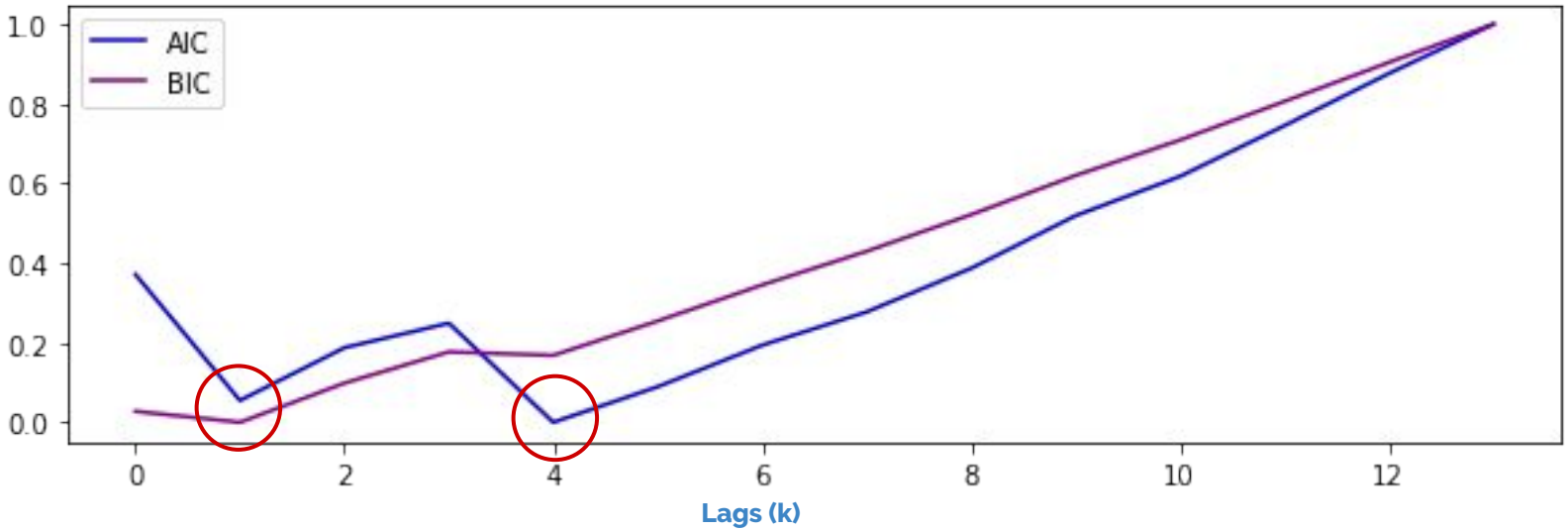
Modeling

Partial Autocorrelation:

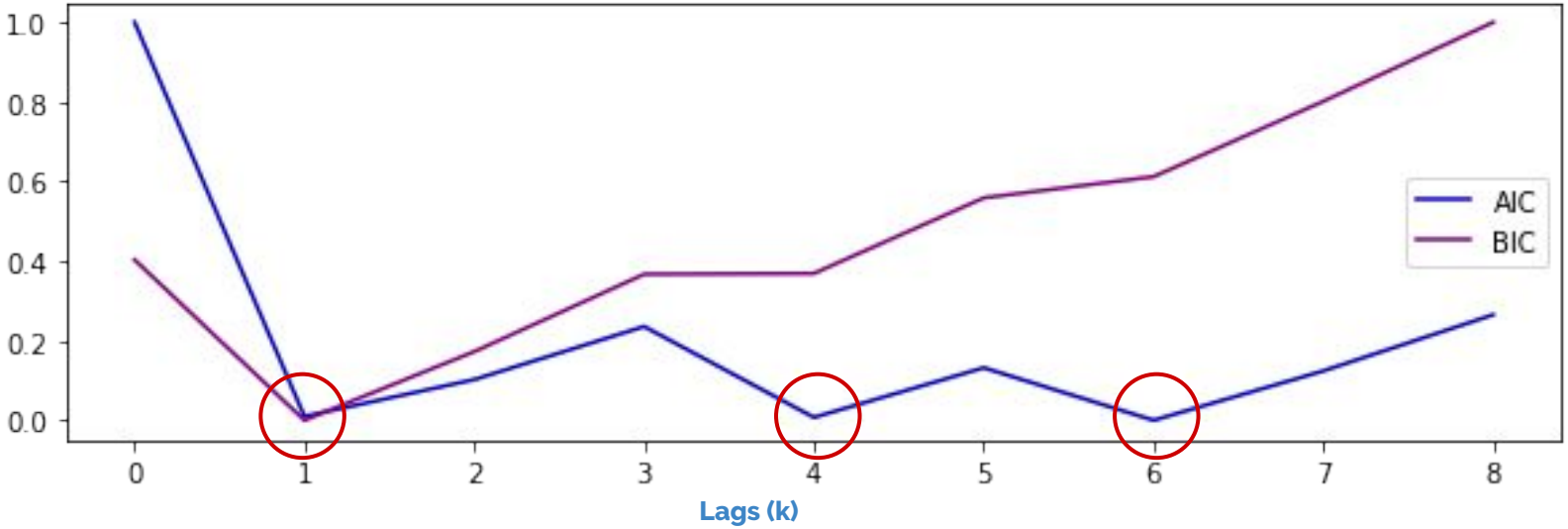
For choosing AR



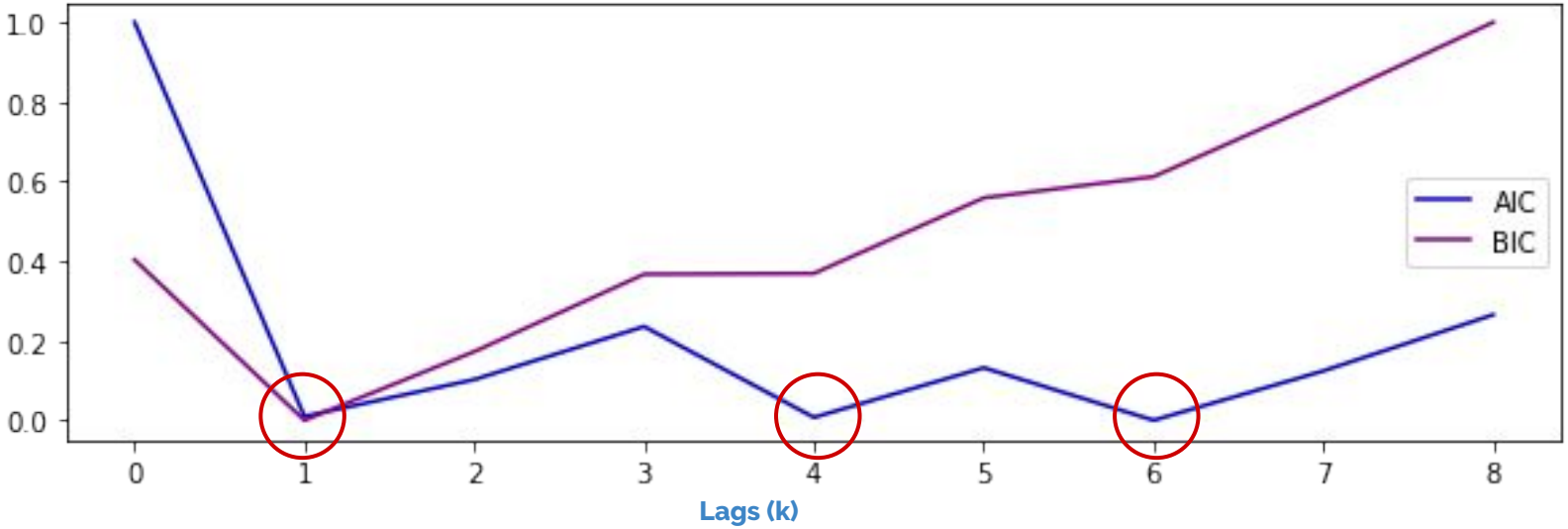
Likelihood:
AR Lags



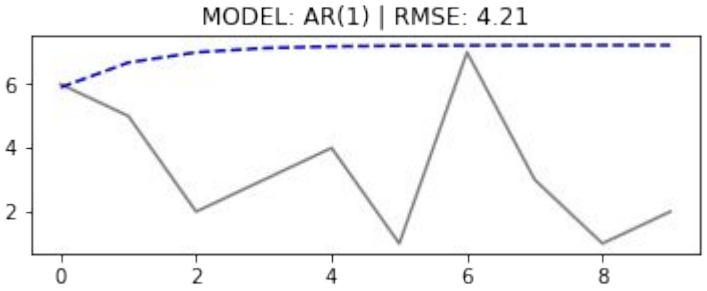
Likelihood:
MA Lags

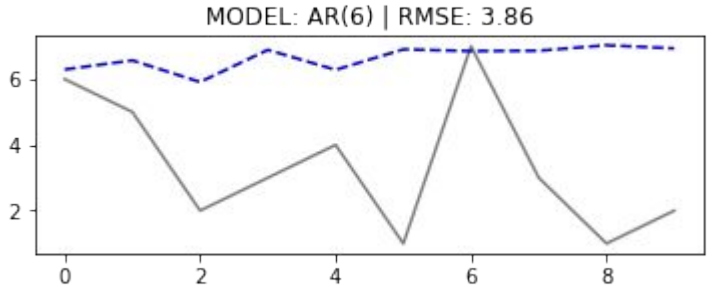


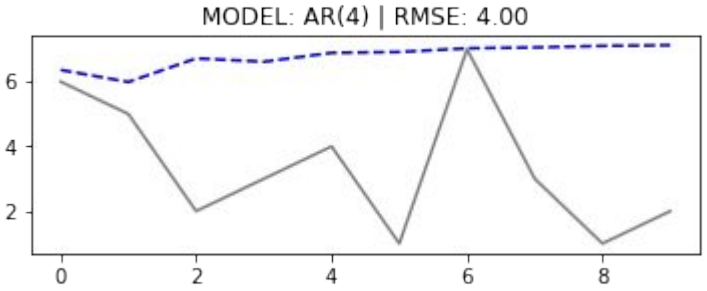
Likelihood:
MA Lags

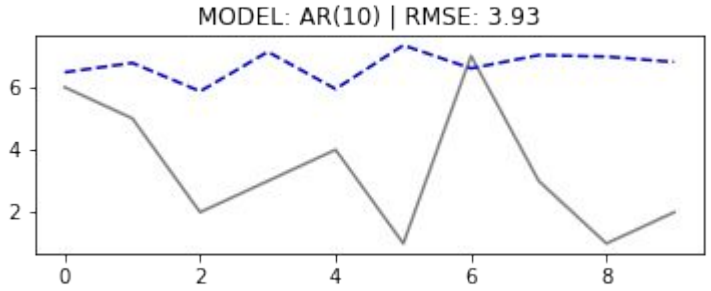


Models: AR

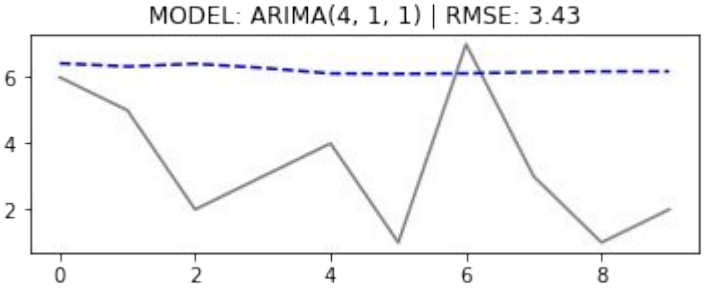

MODEL: AR(1) | RMSE: 4.21
6
4
2
0
2
4
6
8

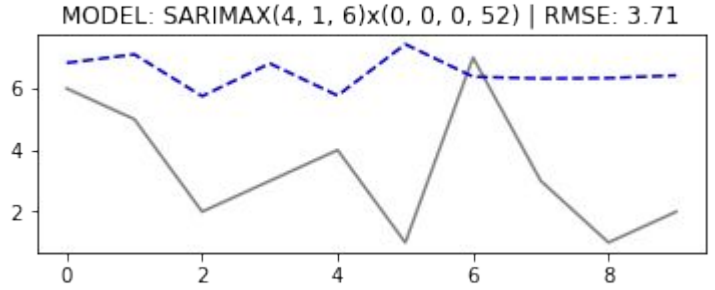

MODEL: AR(6) | RMSE: 3.86
6
4
2
0
2
4
6
8

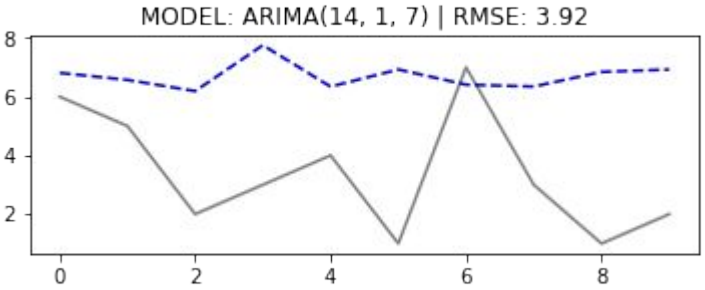

MODEL: AR(4) | RMSE: 4.00
6
4
2
0
2
4
6
8

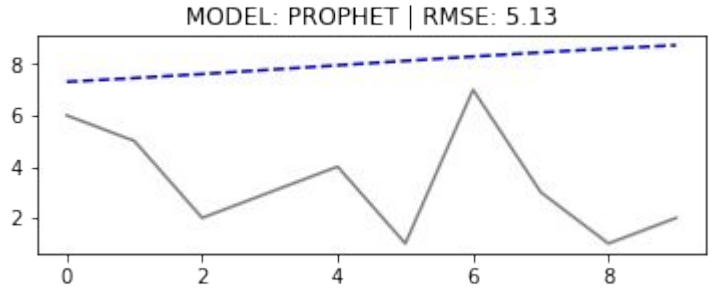

MODEL: AR(10) | RMSE: 3.93
6
4
2
0
2
4
6
8

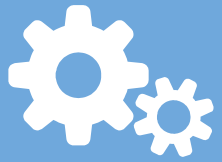
Models: ARIMA, SARIMAX, Prophet







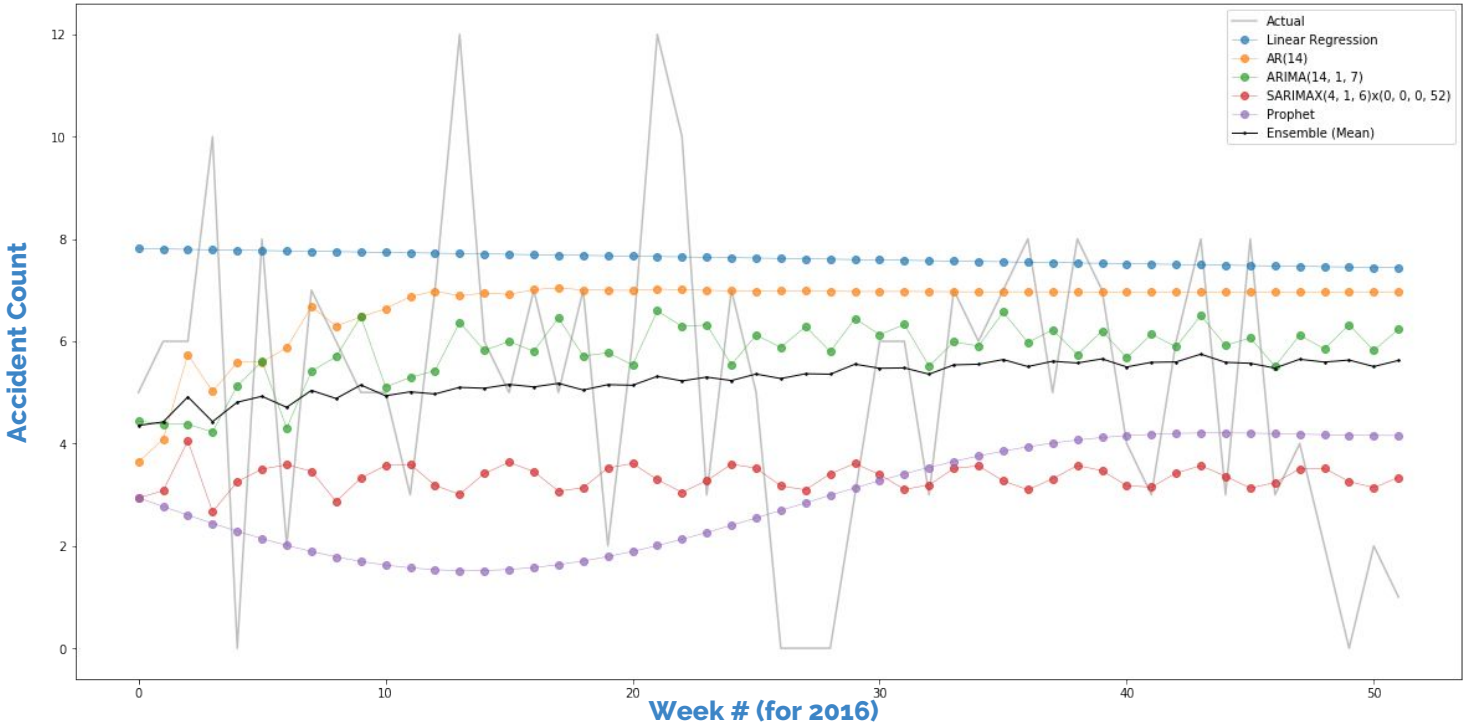


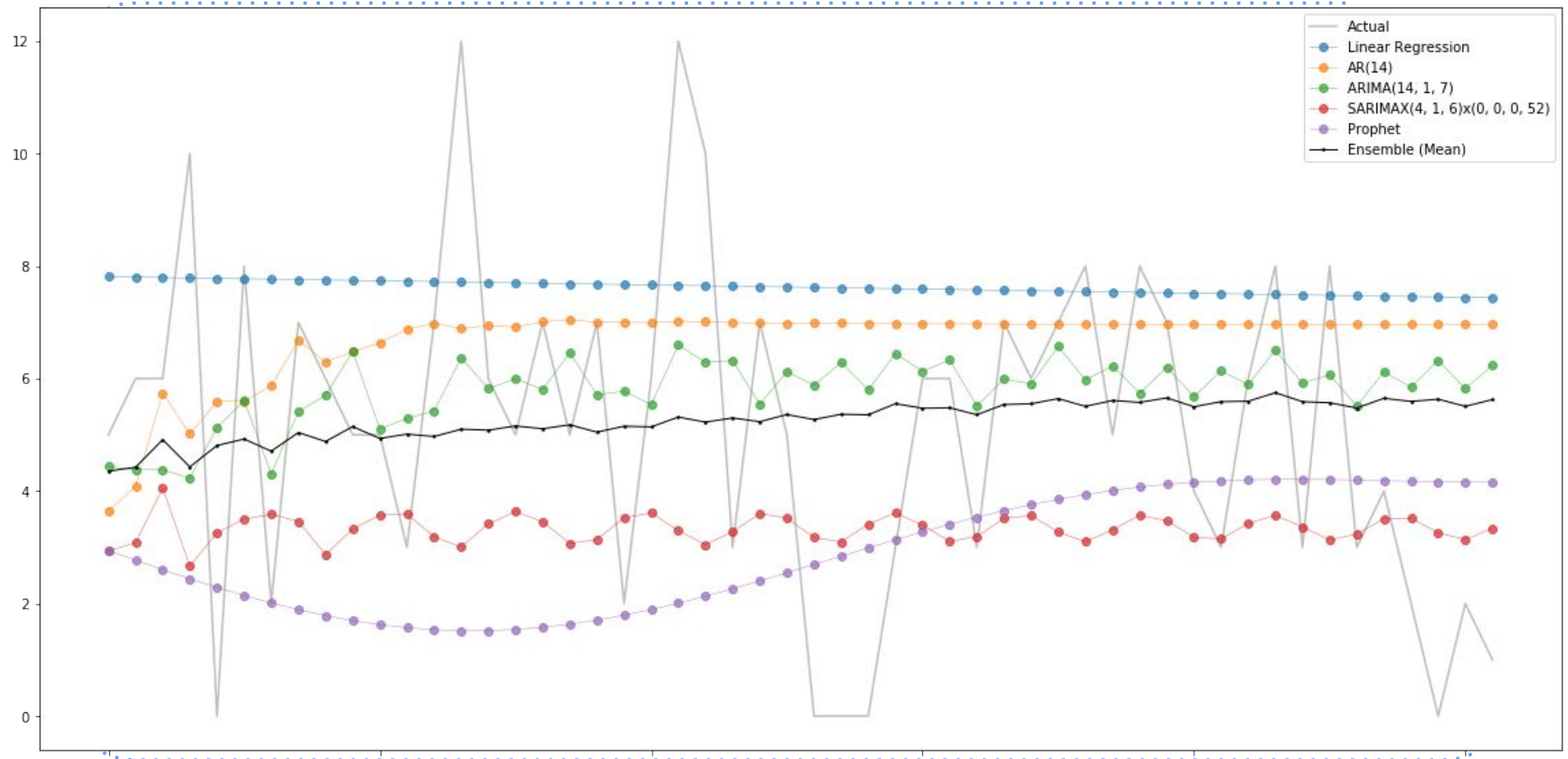


4

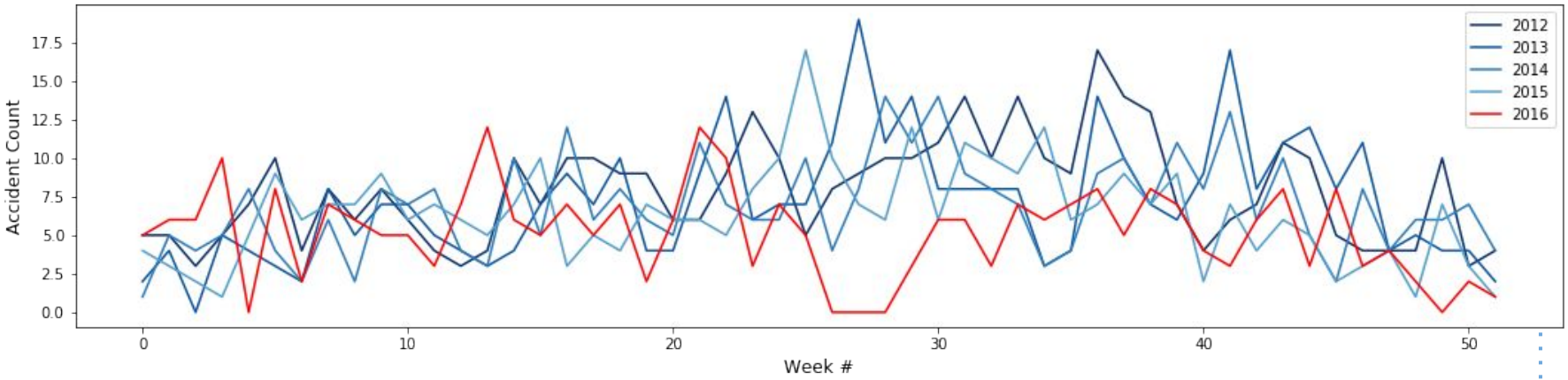
Forecasting

2016 Predictions

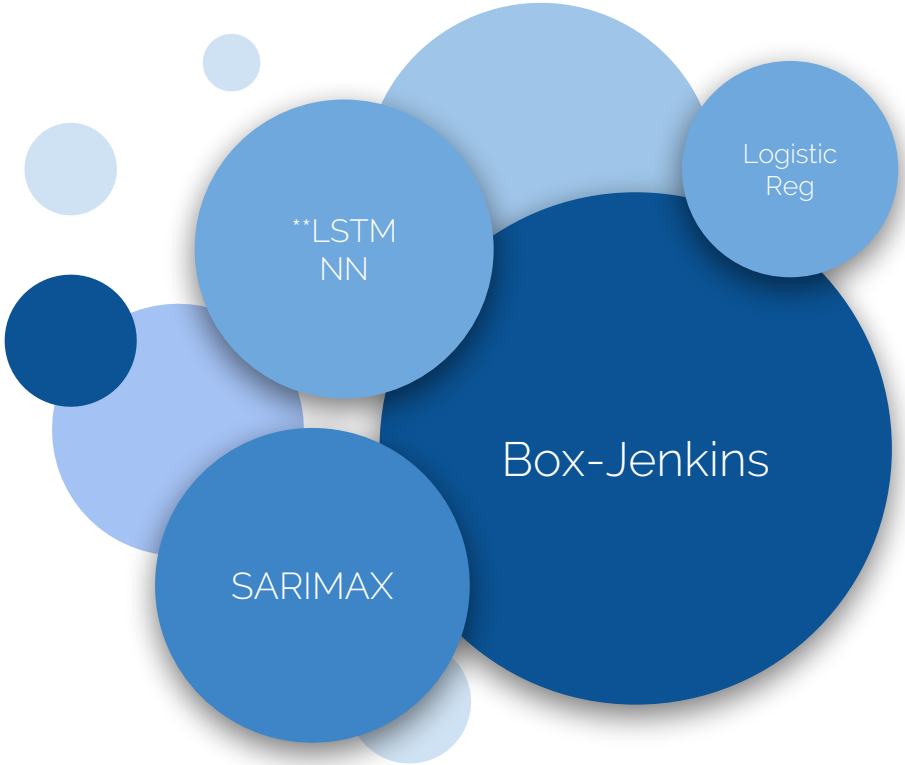


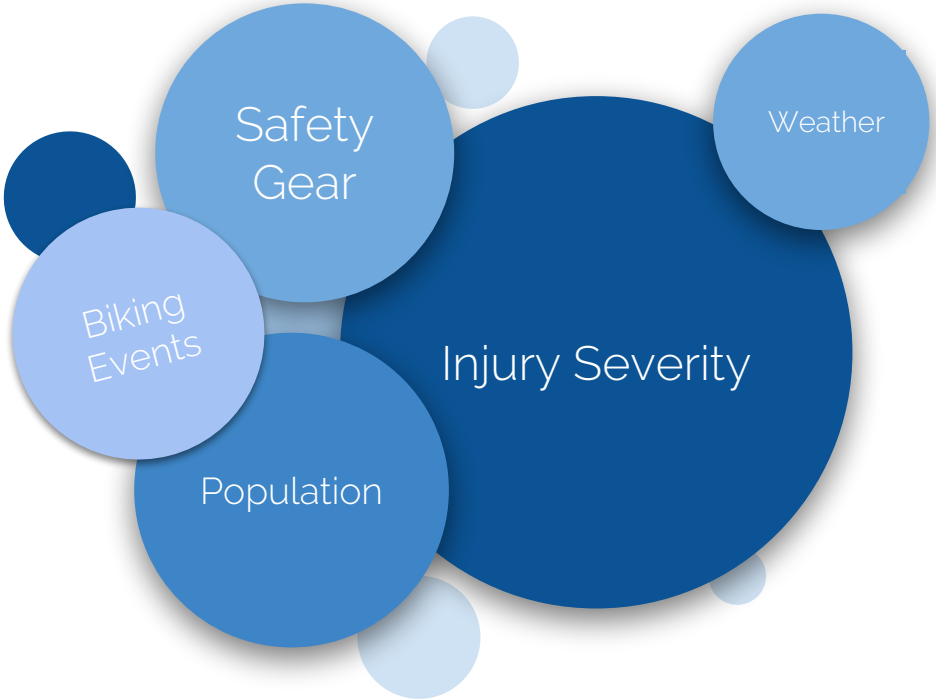

Actual
Linear Regression
AR(14)
ARIMA(14, 1, 7)
SARIMAX(4, 1, 6)x(0, 0, 0, 52)
Prophet
Ensemble (Mean)
12
10
8
6
4
2
0
0
10
20
30
40
50

Accidents by year + 2016



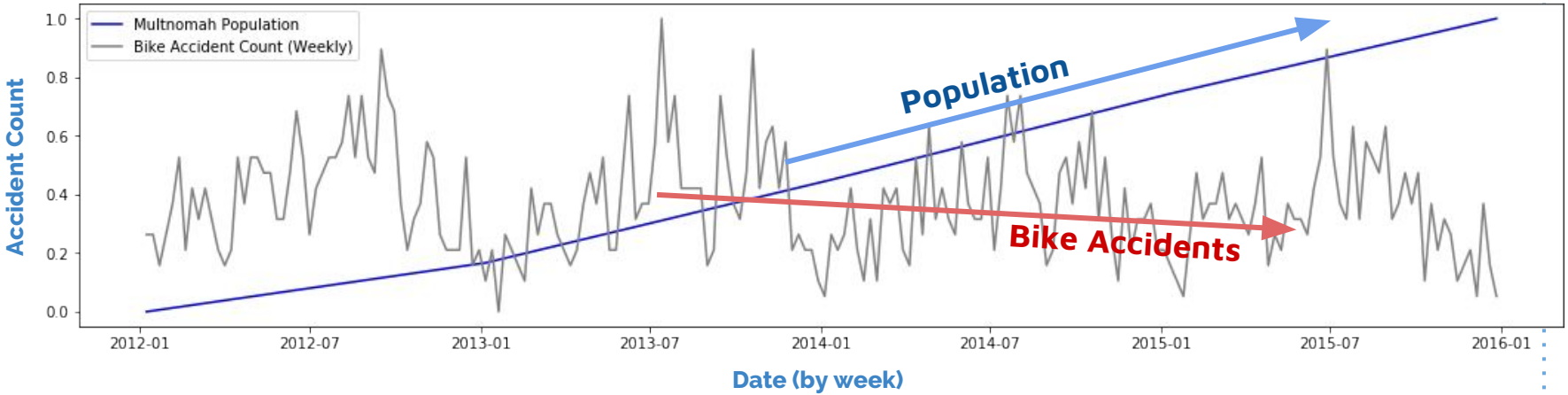
Other models to explore

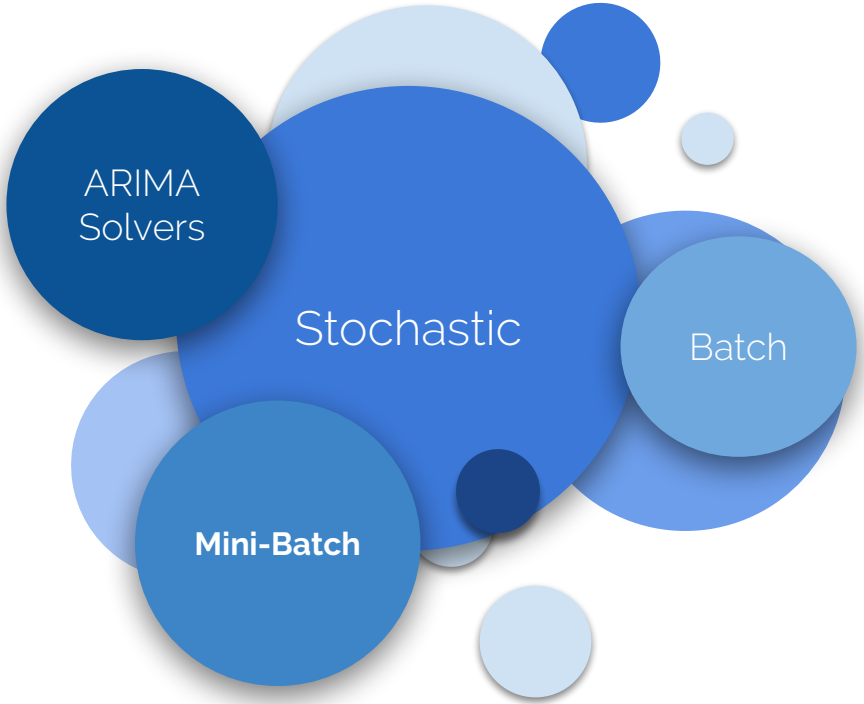




variables to account for

Bike Accidents + Population 2012-2015





weight optimization and hyper parameters

Any Questions?

Special Thanks!!!

Kris Pickrell (Thinkful mentor)

Google Slides Template:

by SlidesCarnival