

capstone_prj_scrub_part3-px

July 18, 2021

1 Data Science Project

- Name: Author Name
- Email:

1.1 TABLE OF CONTENTS

- Section 2
- Section 3
- Section 4
- Section 5
- Section 6
- Section 7
- Section ?? _____

2 INTRODUCTION

Explain the point of your project and what question you are trying to answer with your modeling.

3 OBTAIN

If you are running this notebook without restarting the kernel replace ‘%load_ext autoreload’ in imports with ‘%reload_ext autoreload’

3.1 Imports and Functions

```
[1]: # Importing packages
import pandas as pd
import numpy as np
import matplotlib
import matplotlib.pyplot as plt
import seaborn as sns
import itertools
import statsmodels
import statsmodels.tsa.api as tsa
import plotly.express as px
import plotly.io as pio
```

```

import math
from math import sqrt
import holidays
import pmdarima as pm

from statsmodels.tsa.stattools import adfuller, acf, pacf
from statsmodels.tsa.seasonal import seasonal_decompose
from statsmodels.tsa.arima.model import ARIMA
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from statsmodels.tsa.statespace.sarimax import SARIMAX
from sklearn.metrics import mean_squared_error

from pmdarima.arima.stationarity import ADFTest
from pmdarima.arima.utils import ndiffs
from pmdarima.arima.utils import nsdiffs

import pickle
import os
import json

from pathlib import Path
import subprocess
import io
import warnings
warnings.filterwarnings(action='ignore', category=FutureWarning)

from functions_all import *

%load_ext autoreload
%autoreload 2
%matplotlib inline

```

3.2 Data

3.2.1 Data source and data description

Data is from FBI Crime Data Explorer NIBRS data for Colorado from 2009-2019

The [data dictionary](#) is and a [record description](#) are available.

The description of the main and reference tables is in data/README.md file. The agency implemented some changes to the files structure in 2016 and removed the sqlite create and load scripts from the zip directories. Another fact worth mentioning is that files ‘nibrs_property_desc.csv’ from 2014 and 2015 have duplicated nibrs_property_desc_ids (unique identifier in the nibrs_property_desc table) which complicated the loading of the data.

The rest of the original data description is in description is in the [notebook](#) with the first part of data pre-processing.

3.2.2 Using an already created sqlite database

The notebook with database creation is [here](#). The referenced database is in *data/sqlite/db/production1 db*. It takes 2.5 minutes to run the database creation script.

4 SCRUB

4.1 Part I, pre-processing the data in SQL database

The first part of the scrubbing process (working with sqlite3 database, production1) is in [this notebook](#). It takes about 12 minutes to run the code in part1 notebook. The following code is using dataframes created in part I.

In part I the following dataframes have been created and saved in the pickle files:

1. df_incident: data/pickled_dataframes/incident.pickle; main incident DF with date/time of an incident
2. df_offense: data/pickled_dataframes/offense.pickle; main offense DF with offense names and offense codes
3. df_offender: data/pickled_dataframes/offender.pickle; main offender DF with demographic info
4. df_victim: data/pickled_dataframes/victim.pickle; main victim DF with demographic info
5. df_weapon: data/pickled_dataframes/weapon.pickle; main weapon DF with a weapon category used
6. df_bias: data/pickled_dataframes/bias.pickle; main bias DF with offense bias motivation
7. df_rel: data/pickled_dataframes/relationship.pickle; main victim-offender relationship DF with relationship type

4.2 Part II, scrubbing the data in DataFrames

The next step is uploading, displaying info and scrubbing the dataframes in [part 2 notebook](#)

1. Offense, incident, bias and weapon DataFrames are combined into one for the Times-series analysis
2. Offender, victim and relationship DataFrames are set aside for the future dashboard.
The cleaned-up dataframes can be found here:

1. df_incident: data/pickled_dataframes/incident_clean.pickle; main cleaned-up incident DF with incident details
2. df_offense: data/pickled_dataframes/offense_clean.pickle; main cleaned-up offense DF with offense details
3. df_offender: data/pickled_dataframes/offender_clean.pickle; main cleaned-up offender DF with offender details
4. df_victim: data/pickled_dataframes/victim_clean.pickle; main cleaned-up victim DF with victim details
5. df_weapon: data/pickled_dataframes/weapon_clean.pickle; main cleaned-up weapon DF with a weapon category
6. df_bias: data/pickled_dataframes/bias_clean.pickle; main cleaned-up bias DF with offense bias motivation
7. df_rel: data/pickled_dataframes/rel_clean.pickle; main cleaned-up victim-offender relationship DF with relationship type

4.2.1 Combining Incident, Offense, Bias and Weapon DataFrames

The cleaned-up FULL dataframes can be found here:

```
df_full: 'data/pickled_dataframes/df_full_clean.pickle'; main cleaned-up combined DF for time-series analysis
```

```
[2]: with open('data/pickled_dataframes/df_full_clean.pickle', 'rb') as f:  
    df_full=pickle.load(f)
```

5 EXPLORE

5.1 EDA

All EDA part is in the notebook with Part II [part 2 notebook](#)

The dictionaries with timeseries of various crime categories and crime locations can be found here:

1. TS_crime_category: data/pickled_ts_dict/TS_crime_category.pickle; a dictionary with crime categories
2. TS_crime_against: data/pickled_ts_dict/TS_crime_against.pickle; a dictionary with crime against
3. TS_crime_location: data/pickled_ts_dict/TS_crime_location.pickle; a dictionary with crime locations

6 MODEL

6.1 Splitting into a training and a test sets

I am cutting off a ~10% tail of my data to create a test set.

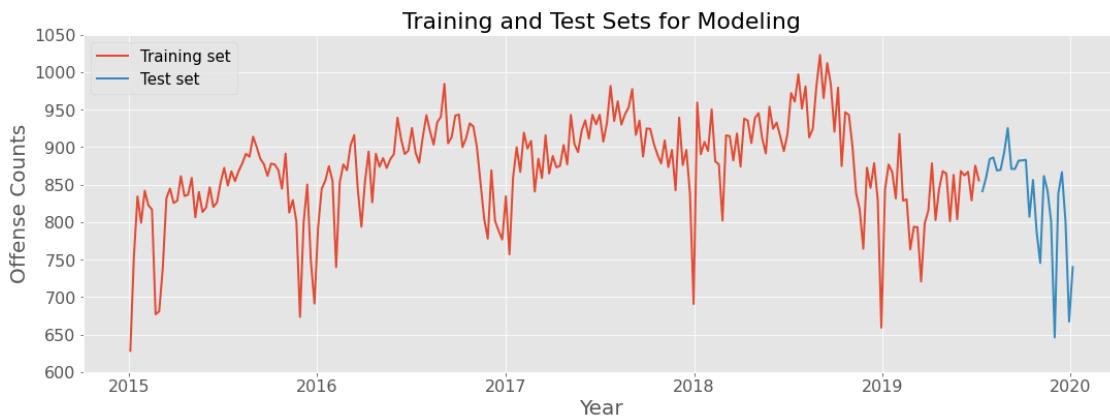
```
[3]: with open('data/pickled_ts/ts_weekly.pickle', 'rb') as f:  
    ts_weekly=pickle.load(f)
```

```
[4]: train_size = round(len(ts_weekly) * 0.90)  
ts_train, ts_test = ts_weekly[:train_size], ts_weekly[train_size:]  
print('Observations: %d weeks' % (len(ts_weekly)))  
print('Training Observations: %d weeks' % (len(ts_train)))  
print('Testing Observations: %d weeks' % (len(ts_test)))  
  
fig=display_figure_w_TSs(ts_train, ts_test, 'Training set', 'Test set',  
    ↴'Training and Test Sets for Modeling')
```

Observations: 262 weeks

Training Observations: 236 weeks

Testing Observations: 26 weeks



6.2 General Crime Rate Modeling

All modeling of General Crime rate is in the notebook with Part II [part 2 notebook](#)

6.3 Crime Rate per Offense Category Modeling

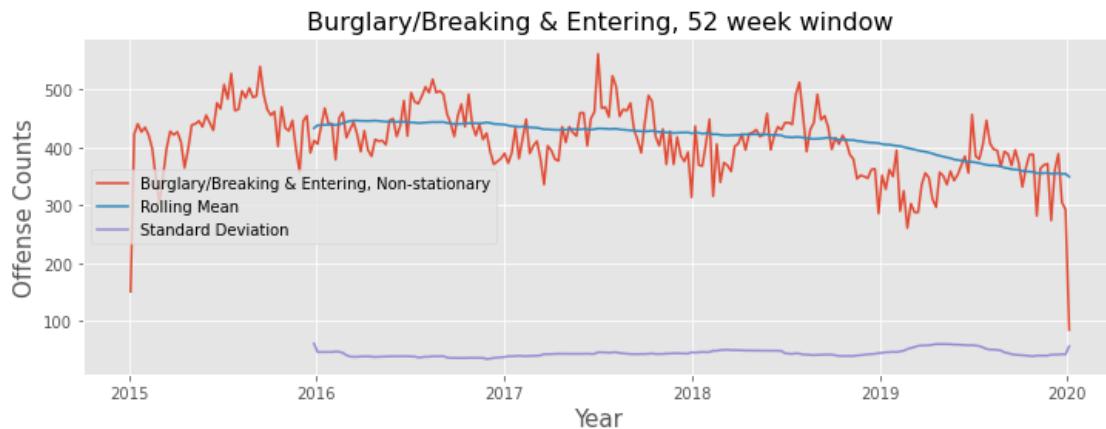
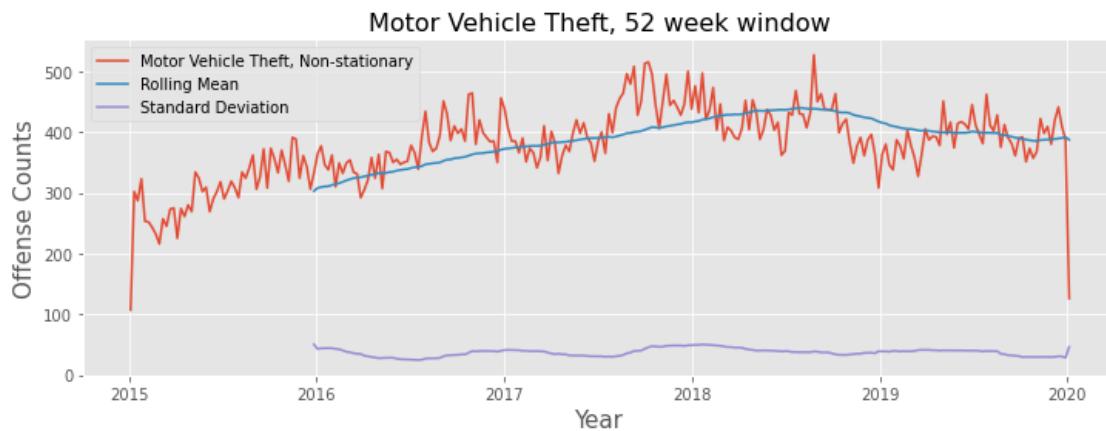
6.3.1 Loading the dictionaries with time-series

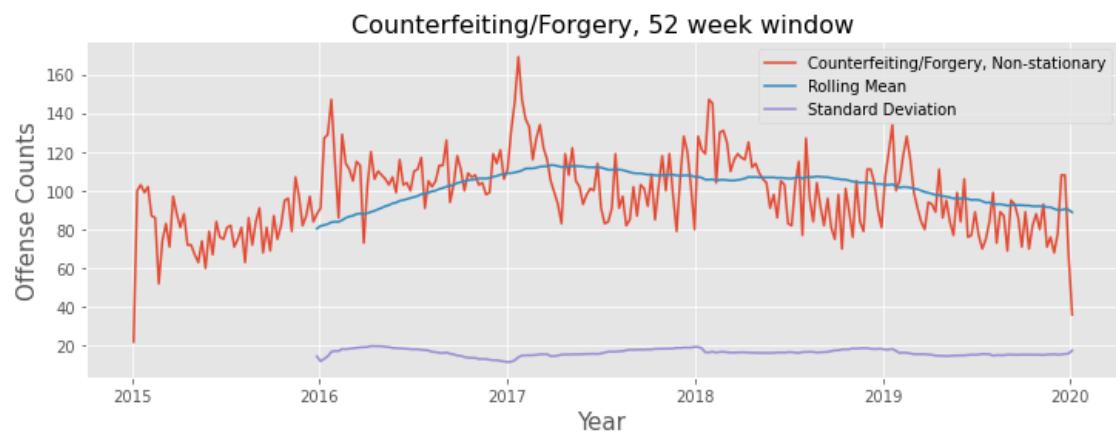
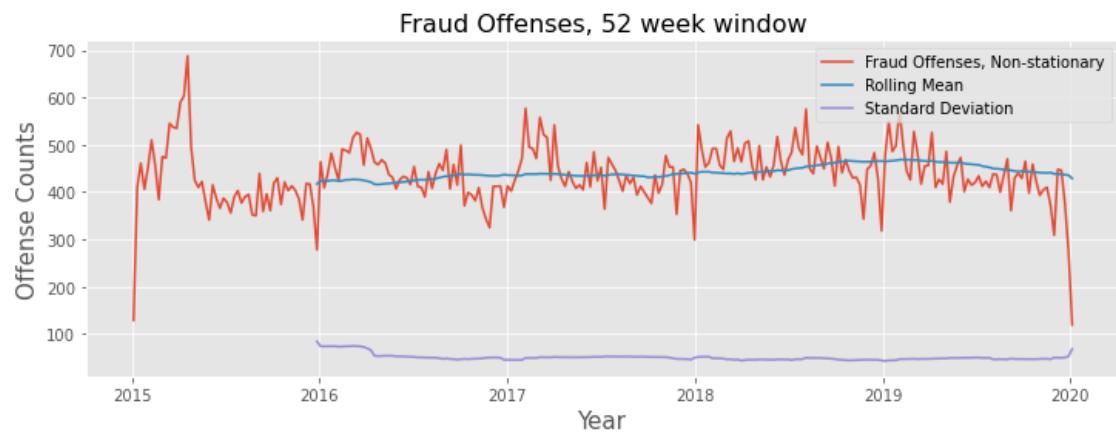
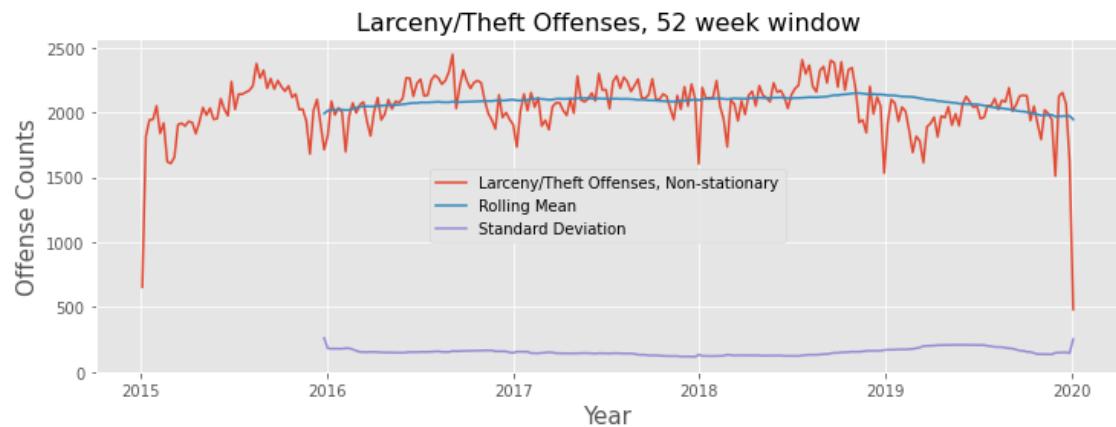
```
[5]: with open('data/pickled_ts/TS_crime_category.pickle', 'rb') as f:  
    TS_crime_category=pickle.load(f)  
  
with open('data/pickled_ts/TS_crime_against.pickle', 'rb') as f:  
    TS_crime_against=pickle.load(f)  
  
with open('data/pickled_ts/TS_crime_location.pickle', 'rb') as f:  
    TS_crime_location=pickle.load(f)
```

6.3.2 Checking for stationarity

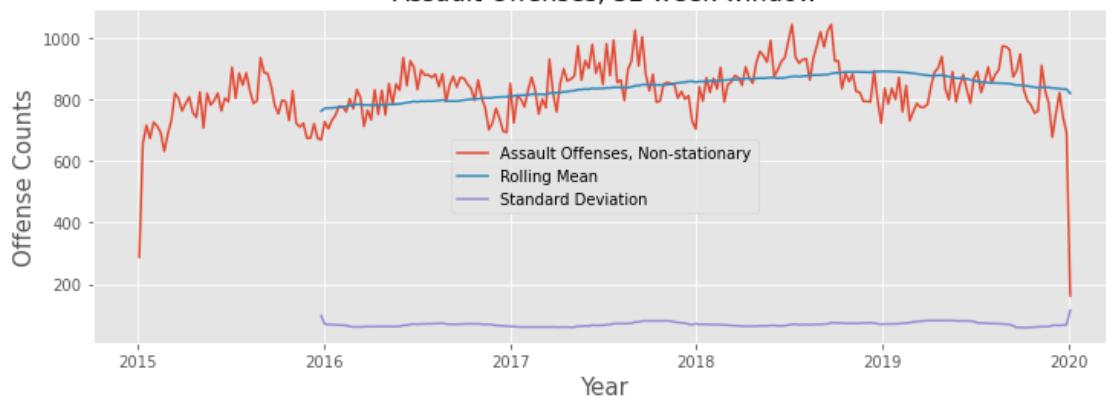
Checking the stationarity of the time-series in Offense category dictionary

```
[6]: df_results1, ts_stationary1, ▾  
    ↪ts_non_stationary_diff1=check_stationarity_multiple(TS_crime_category, ▾  
    ↪window=52, plot=True)
```

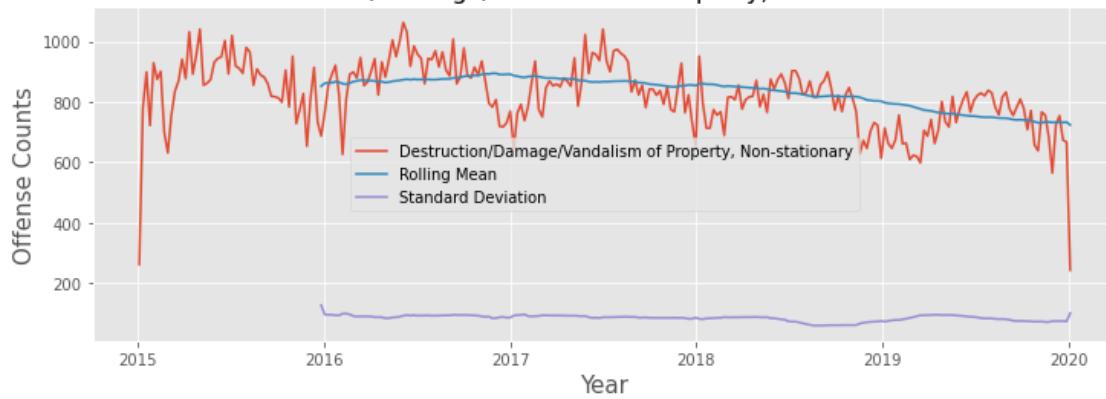




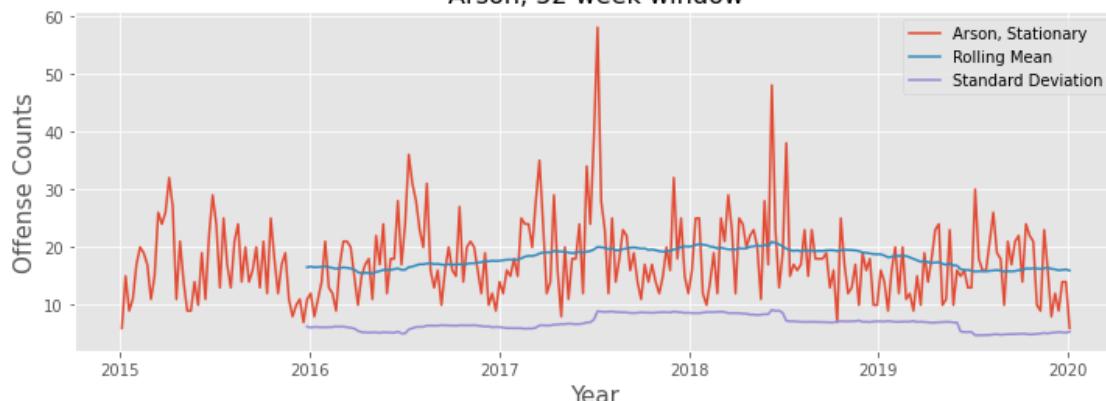
Assault Offenses, 52 week window



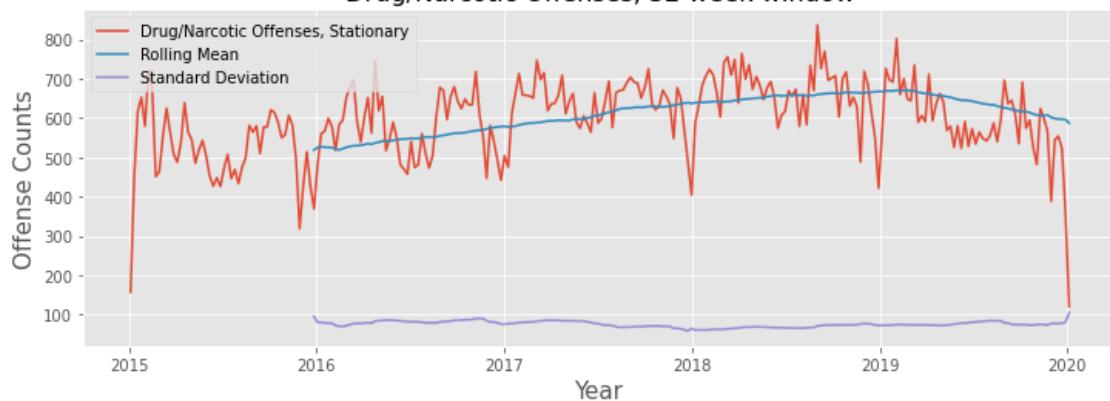
Destruction/Damage/Vandalism of Property, 52 week window



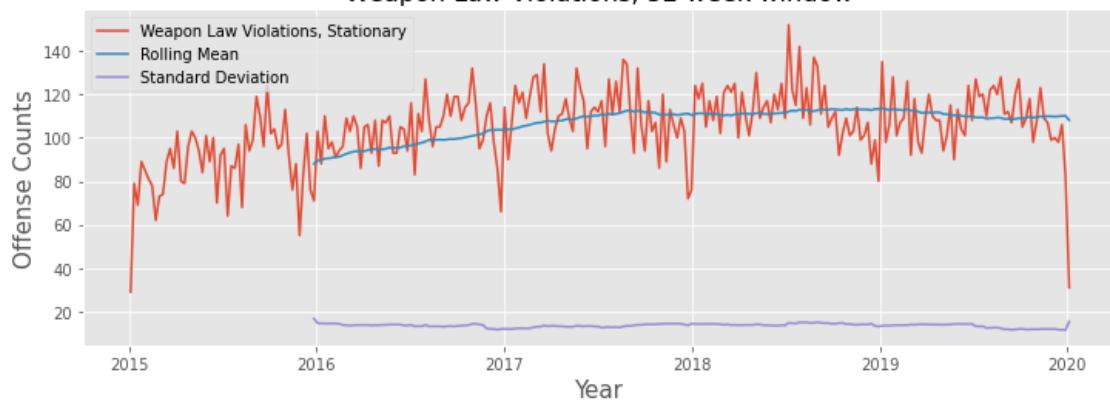
Arson, 52 week window



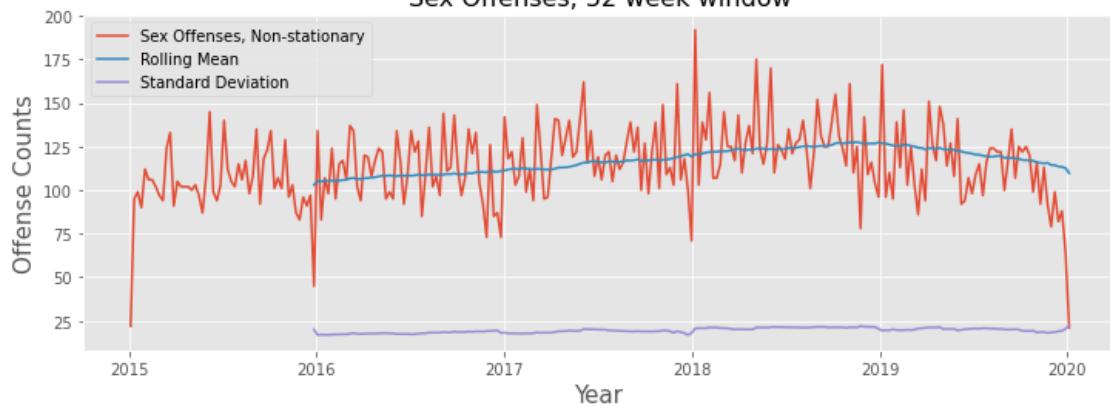
Drug/Narcotic Offenses, 52 week window

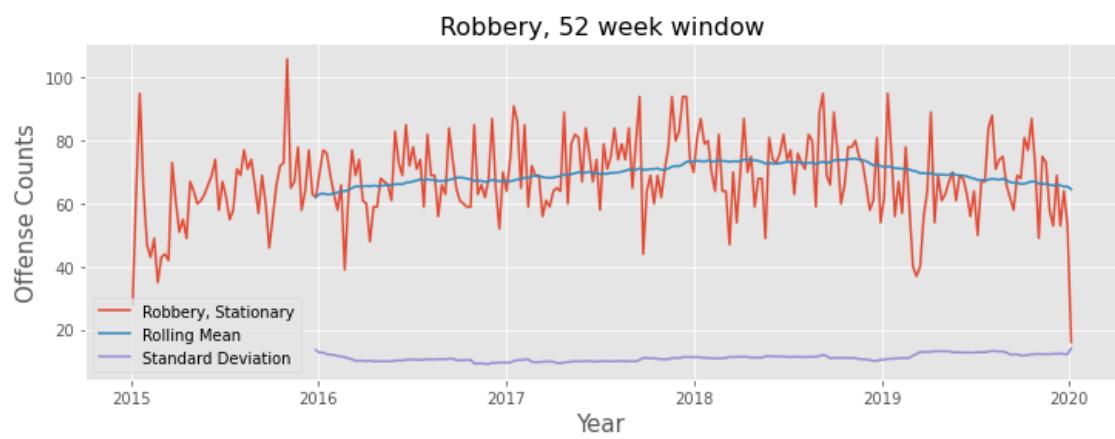
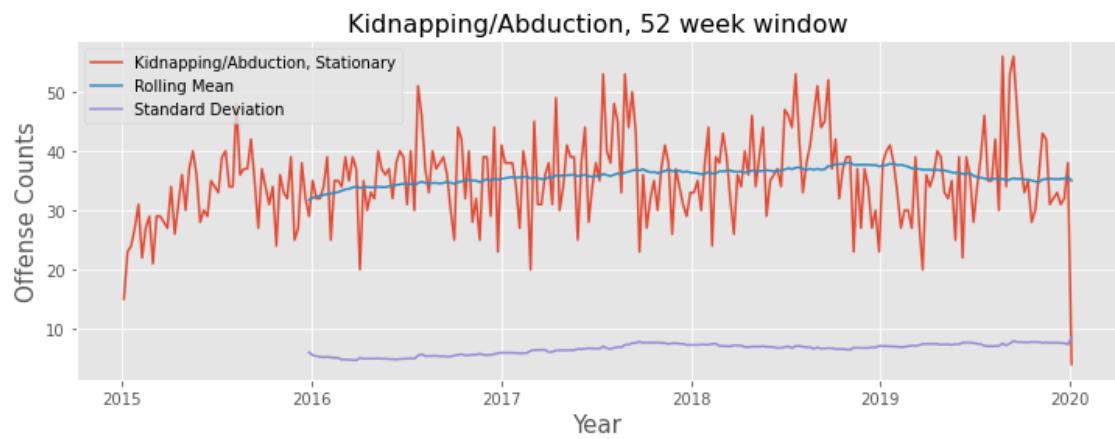


Weapon Law Violations, 52 week window

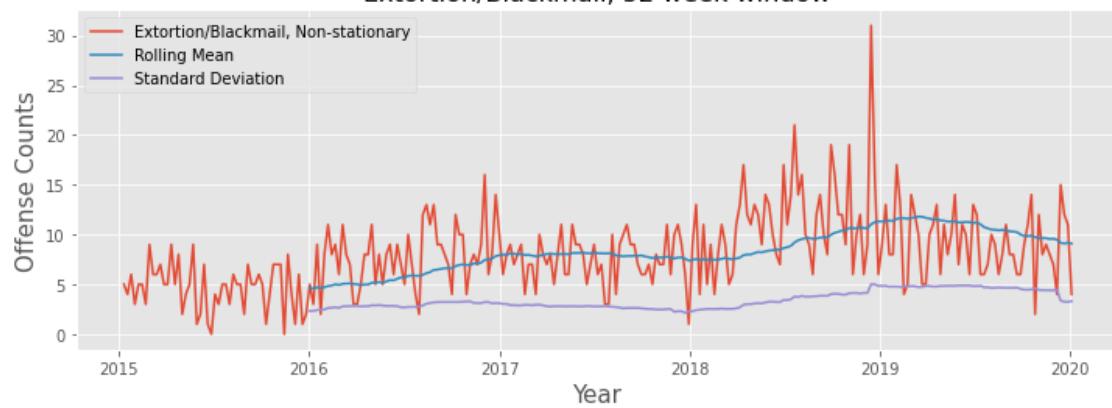


Sex Offenses, 52 week window

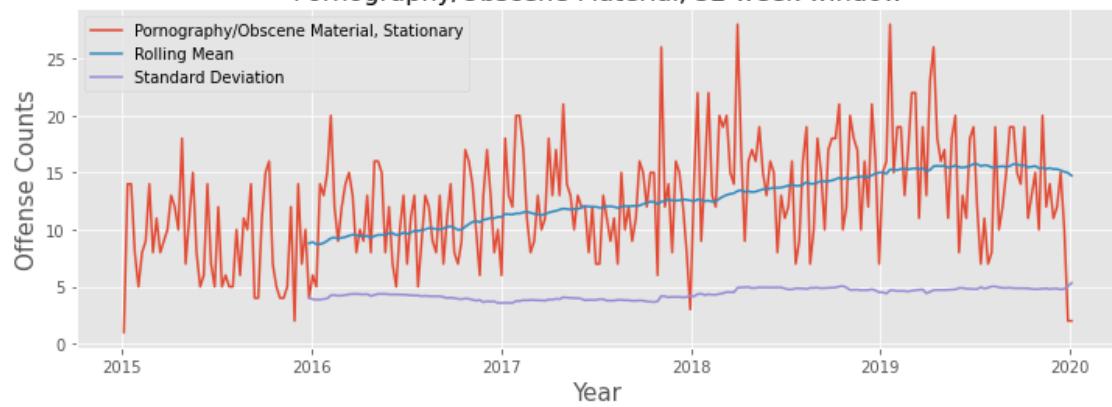




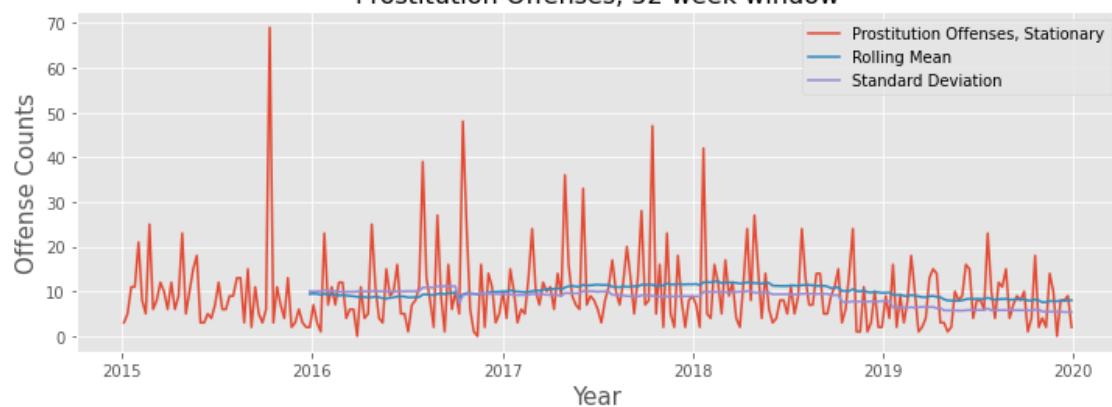
Extortion/Blackmail, 52 week window



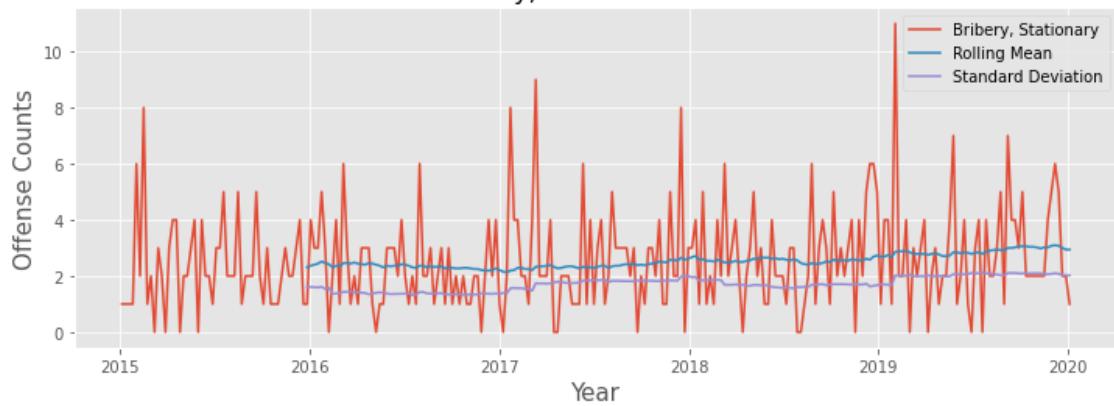
Pornography/Obscene Material, 52 week window



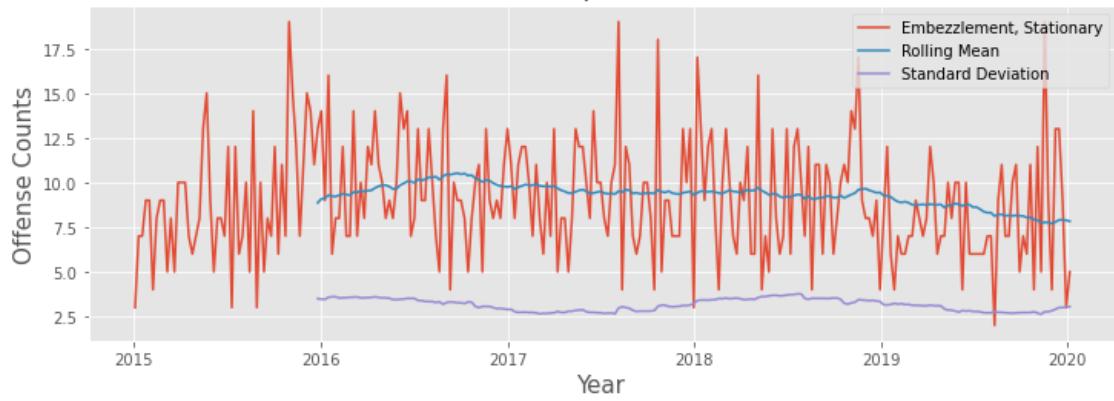
Prostitution Offenses, 52 week window



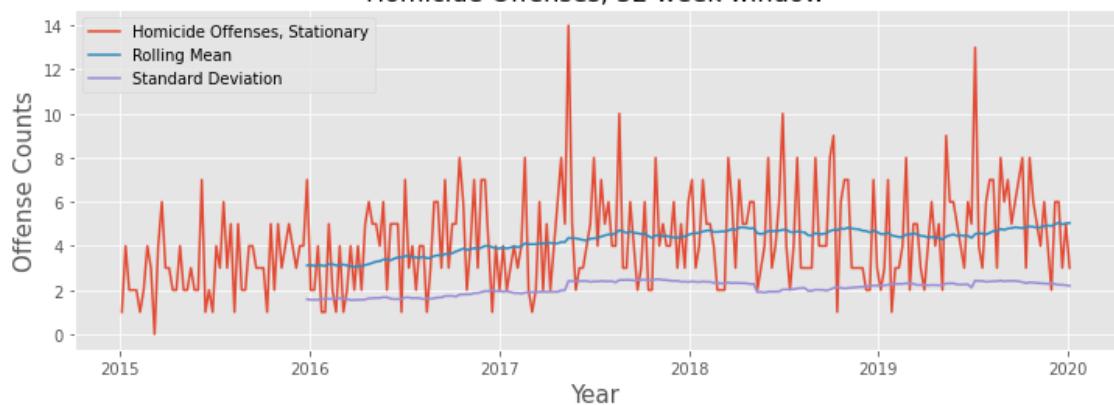
Bribery, 52 week window



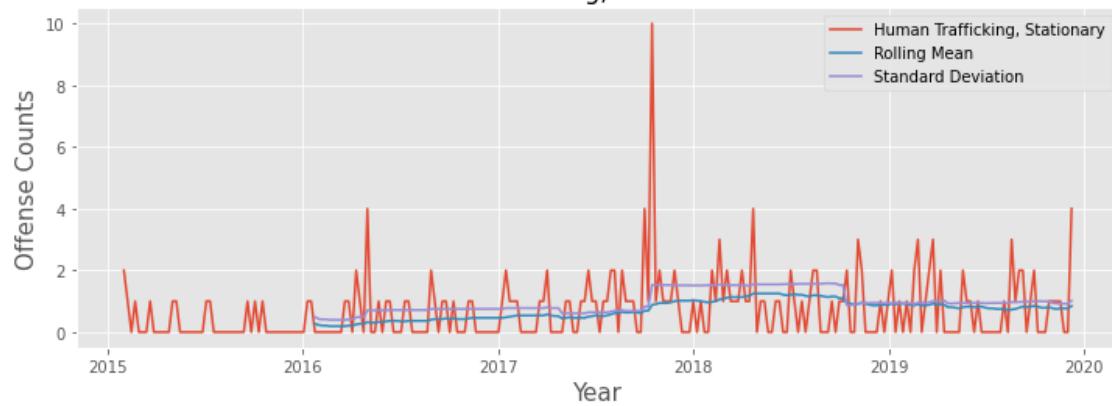
Embezzlement, 52 week window



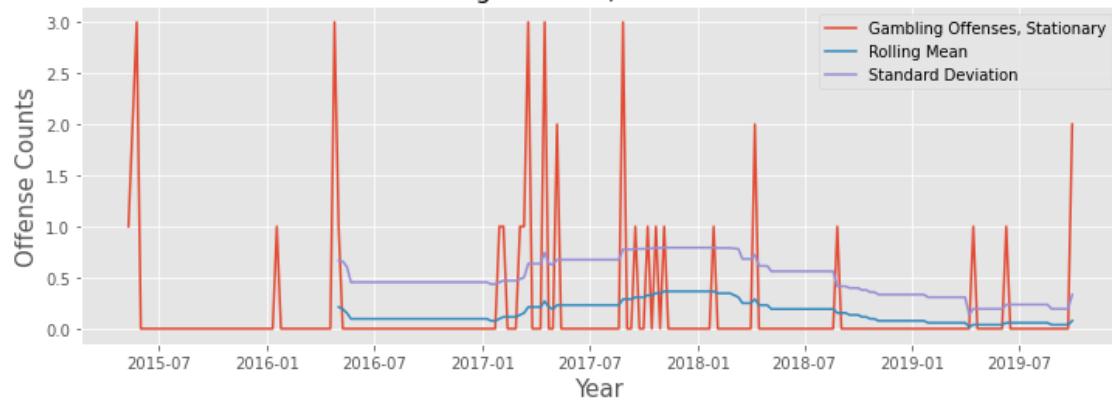
Homicide Offenses, 52 week window



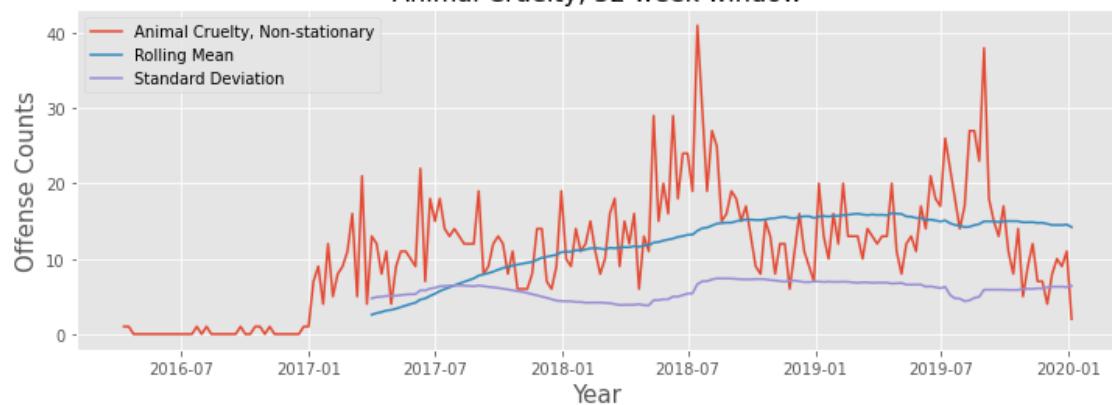
Human Trafficking, 52 week window



Gambling Offenses, 52 week window



Animal Cruelty, 52 week window



[7] : df_results1

[7] :

| Number | Crime Category | Critical Value | P-value | \ |
|--------|--|----------------|-------------|---|
| 1 | Motor Vehicle Theft | -2.47862 | 0.120763 | |
| 2 | Burglary/Breaking & Entering | -1.57887 | 0.494197 | |
| 3 | Larceny/Theft Offenses | -2.07108 | 0.256303 | |
| 4 | Fraud Offenses | -2.34796 | 0.156966 | |
| 5 | Counterfeiting/Forgery | -2.55755 | 0.102093 | |
| 6 | Assault Offenses | -2.36468 | 0.151953 | |
| 7 | Destruction/Damage/Vandalism of Property | -1.47935 | 0.543701 | |
| 8 | Arson | -5.43906 | 2.80274e-06 | |
| 9 | Drug/Narcotic Offenses | -3.0009 | 0.0348101 | |
| 10 | Weapon Law Violations | -3.79849 | 0.00292282 | |
| 11 | Sex Offenses | -1.69734 | 0.432483 | |
| 12 | Stolen Property Offenses | -2.60219 | 0.0925456 | |
| 13 | Kidnapping/Abduction | -5.83445 | 3.90605e-07 | |
| 14 | Robbery | -5.64789 | 1.00156e-06 | |
| 15 | Extortion/Blackmail | -1.88379 | 0.339725 | |
| 16 | Pornography/Obscene Material | -3.26358 | 0.0165871 | |
| 17 | Prostitution Offenses | -16.6825 | 1.51221e-29 | |
| 18 | Bribery | -17.0216 | 8.44947e-30 | |
| 19 | Embezzlement | -6.24598 | 4.57491e-08 | |
| 20 | Homicide Offenses | -5.49933 | 2.08847e-06 | |
| 21 | Human Trafficking | -8.46655 | 1.51701e-13 | |
| 22 | Gambling Offenses | -12.6476 | 1.39041e-23 | |
| 23 | Animal Cruelty | -2.23411 | 0.194099 | |

| Number | Lags | Observations | Critical value, 1% | Critical value, 5% | \ |
|--------|------|--------------|--------------------|--------------------|---|
| 1 | 7 | 254 | -3.45636 | -2.87299 | |
| 2 | 3 | 258 | -3.45595 | -2.87281 | |
| 3 | 4 | 257 | -3.45605 | -2.87285 | |
| 4 | 16 | 245 | -3.45733 | -2.87341 | |
| 5 | 4 | 257 | -3.45605 | -2.87285 | |
| 6 | 3 | 258 | -3.45595 | -2.87281 | |
| 7 | 4 | 257 | -3.45605 | -2.87285 | |
| 8 | 3 | 258 | -3.45595 | -2.87281 | |
| 9 | 3 | 258 | -3.45595 | -2.87281 | |
| 10 | 3 | 258 | -3.45595 | -2.87281 | |
| 11 | 12 | 249 | -3.45689 | -2.87322 | |
| 12 | 5 | 256 | -3.45616 | -2.8729 | |
| 13 | 2 | 259 | -3.45585 | -2.87276 | |
| 14 | 2 | 259 | -3.45585 | -2.87276 | |
| 15 | 10 | 250 | -3.45678 | -2.87317 | |
| 16 | 5 | 256 | -3.45616 | -2.8729 | |
| 17 | 0 | 260 | -3.45575 | -2.87272 | |
| 18 | 0 | 261 | -3.45566 | -2.87268 | |
| 19 | 3 | 258 | -3.45595 | -2.87281 | |

| | | | | |
|----|---|-----|----------|----------|
| 20 | 3 | 258 | -3.45595 | -2.87281 |
| 21 | 1 | 252 | -3.45657 | -2.87308 |
| 22 | 0 | 229 | -3.45923 | -2.87425 |
| 23 | 3 | 192 | -3.46488 | -2.87671 |

Critical value, 10% Stationary?

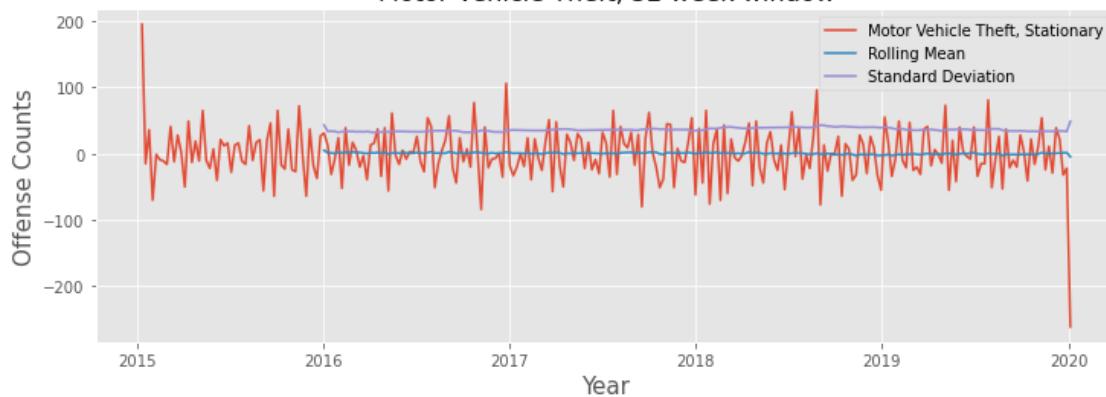
| Number | | | |
|--------|--|----------|-------|
| 1 | | -2.57287 | False |
| 2 | | -2.57277 | False |
| 3 | | -2.5728 | False |
| 4 | | -2.5731 | False |
| 5 | | -2.5728 | False |
| 6 | | -2.57277 | False |
| 7 | | -2.5728 | False |
| 8 | | -2.57277 | True |
| 9 | | -2.57277 | True |
| 10 | | -2.57277 | True |
| 11 | | -2.57299 | False |
| 12 | | -2.57282 | False |
| 13 | | -2.57275 | True |
| 14 | | -2.57275 | True |
| 15 | | -2.57297 | False |
| 16 | | -2.57282 | True |
| 17 | | -2.57273 | True |
| 18 | | -2.57271 | True |
| 19 | | -2.57277 | True |
| 20 | | -2.57277 | True |
| 21 | | -2.57292 | True |
| 22 | | -2.57354 | True |
| 23 | | -2.57486 | False |

There are **12** time-series that are already stationary and **11** that are not and which require additional processing (differencing).

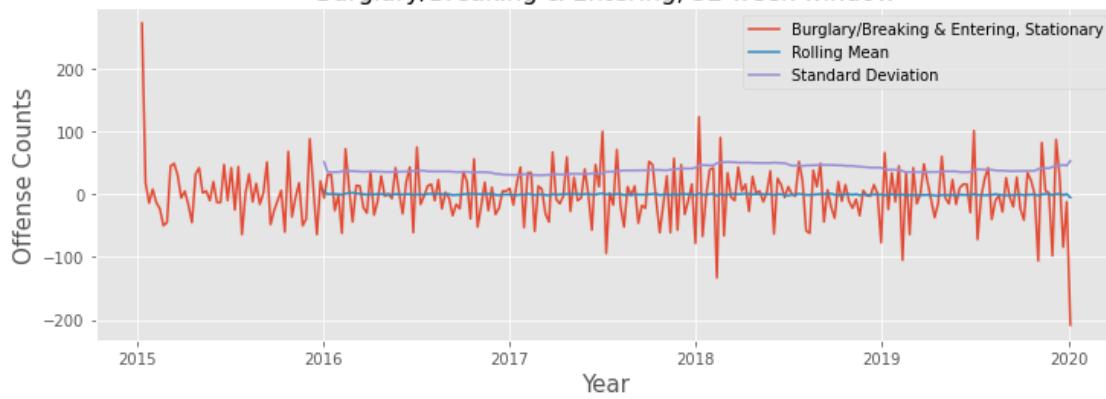
Differencing the time series that are not stationary

```
[8]: df_results2, ts_stationary2, ts_non_stationary_diff2=check_stationarity_multiple(ts_non_stationary_diff1,
    window=52, plot=True)
```

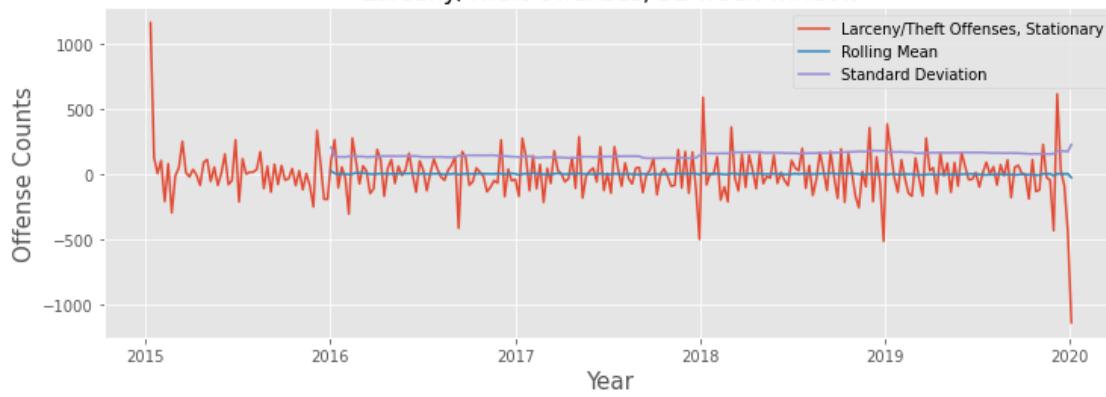
Motor Vehicle Theft, 52 week window

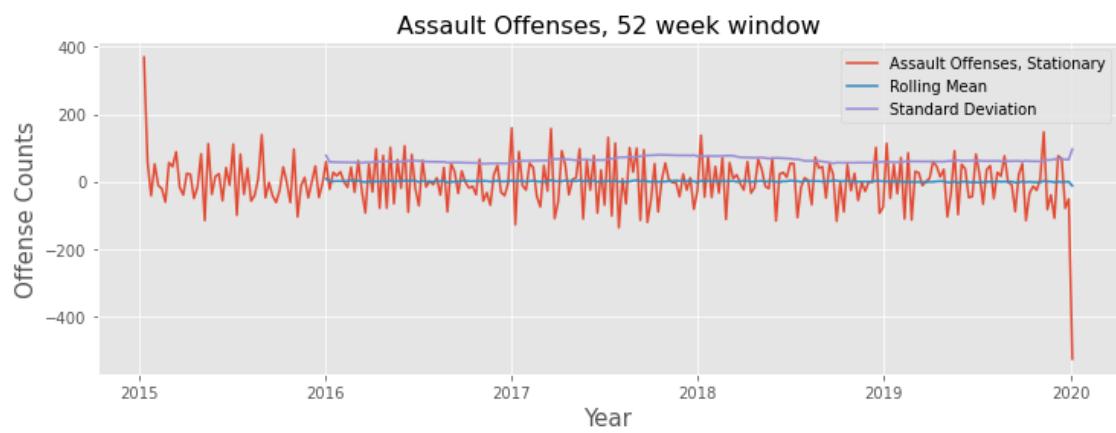
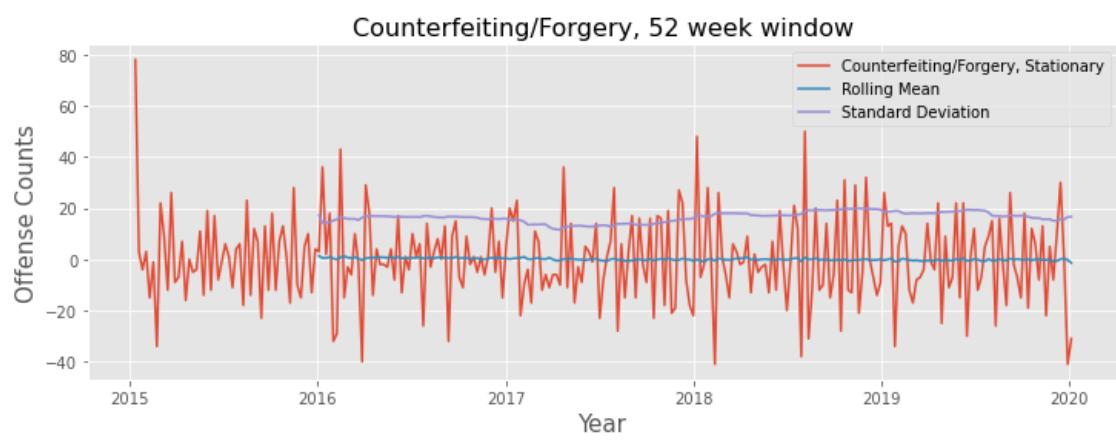
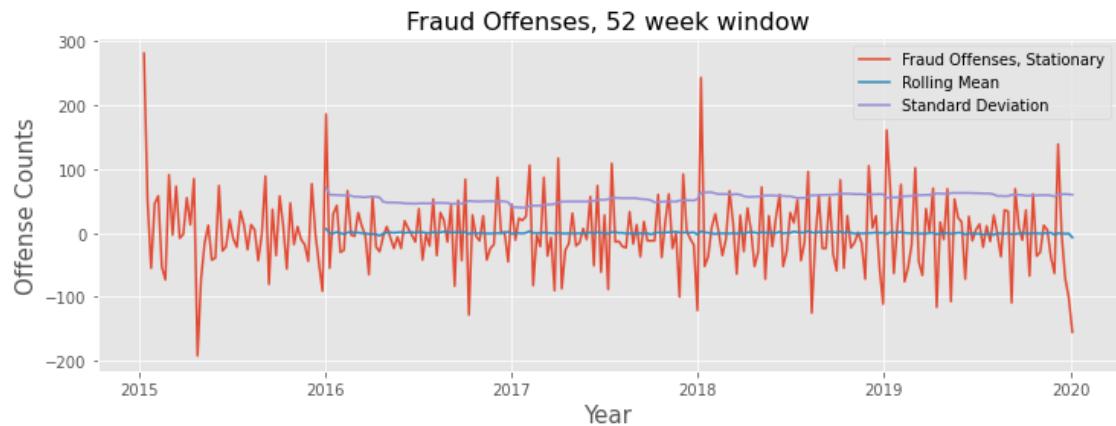


Burglary/Breaking & Entering, 52 week window

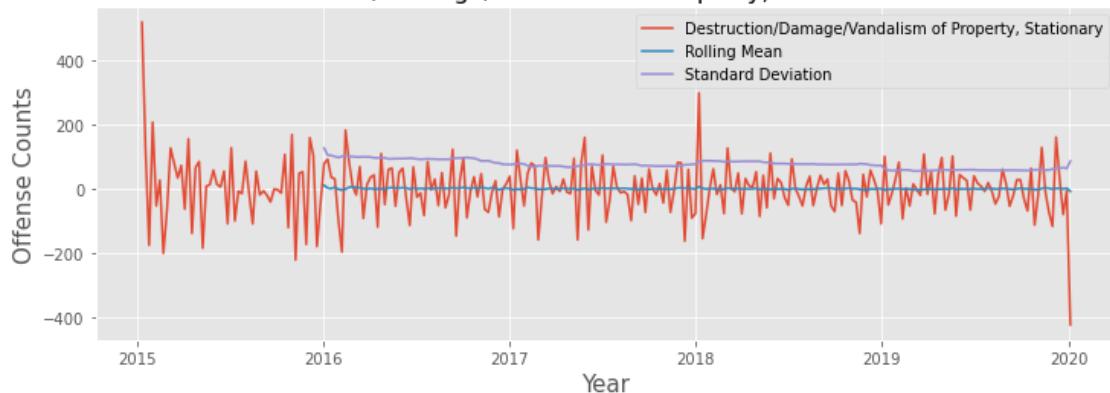


Larceny/Theft Offenses, 52 week window

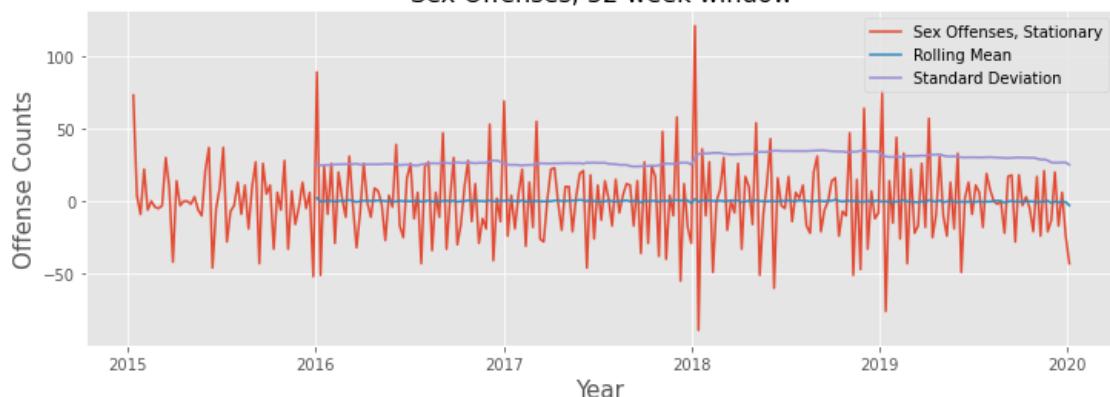




Destruction/Damage/Vandalism of Property, 52 week window

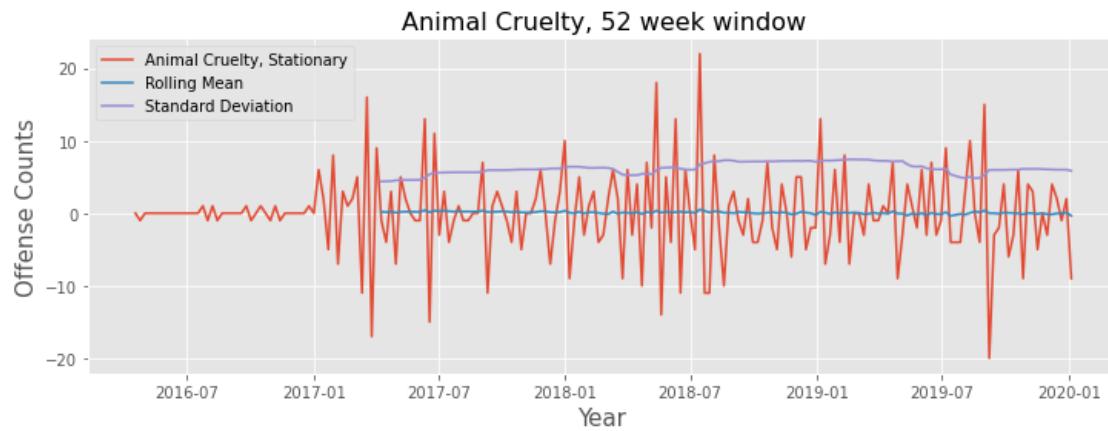
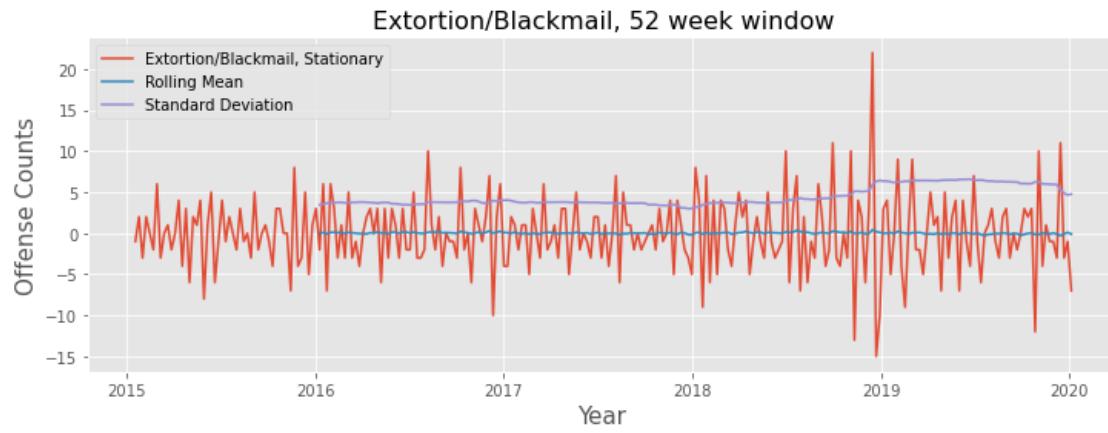


Sex Offenses, 52 week window



Stolen Property Offenses, 52 week window





```
[9]: df_results2
```

| | Crime Category | Critical Value | P-value | \ |
|--------|--|----------------|-------------|---|
| Number | | | | |
| 1 | Motor Vehicle Theft | -6.17315 | 6.72911e-08 | |
| 2 | Burglary/Breaking & Entering | -12.8616 | 5.0966e-24 | |
| 3 | Larceny/Theft Offenses | -11.1342 | 3.22534e-20 | |
| 4 | Fraud Offenses | -6.08654 | 1.06105e-07 | |
| 5 | Counterfeiting/Forgery | -10.6294 | 5.26217e-19 | |
| 6 | Assault Offenses | -10.7223 | 3.12766e-19 | |
| 7 | Destruction/Damage/Vandalism of Property | -12.0371 | 2.78743e-22 | |
| 8 | Sex Offenses | -6.52936 | 9.95943e-09 | |
| 9 | Stolen Property Offenses | -10.3858 | 2.08208e-18 | |
| 10 | Extortion/Blackmail | -9.33958 | 8.87139e-16 | |
| 11 | Animal Cruelty | -12.7603 | 8.16952e-24 | |

| Number | Lags | Observations | Critical value, 1% | Critical value, 5% | \ |
|--------|------|--------------|--------------------|--------------------|---|
| 1 | 7 | 253 | -3.45646 | -2.87303 | |
| 2 | 2 | 258 | -3.45595 | -2.87281 | |
| 3 | 3 | 257 | -3.45605 | -2.87285 | |
| 4 | 15 | 245 | -3.45733 | -2.87341 | |
| 5 | 5 | 255 | -3.45626 | -2.87294 | |
| 6 | 2 | 258 | -3.45595 | -2.87281 | |
| 7 | 3 | 257 | -3.45605 | -2.87285 | |
| 8 | 11 | 249 | -3.45689 | -2.87322 | |
| 9 | 5 | 255 | -3.45626 | -2.87294 | |
| 10 | 9 | 250 | -3.45678 | -2.87317 | |
| 11 | 2 | 192 | -3.46488 | -2.87671 | |

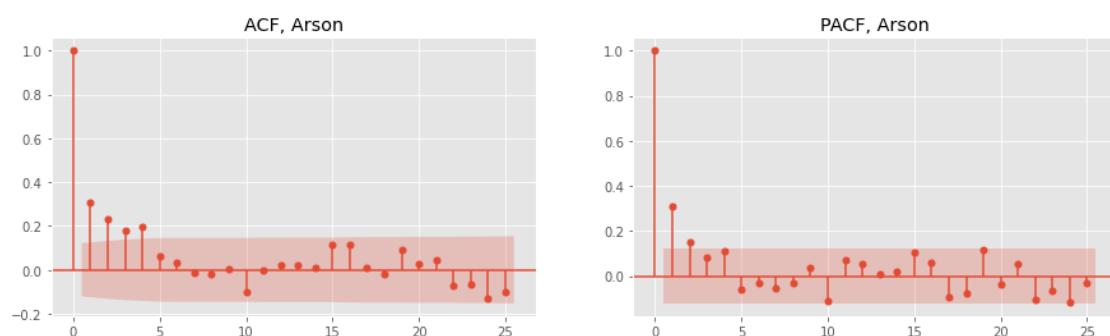
Critical value, 10% Stationary?

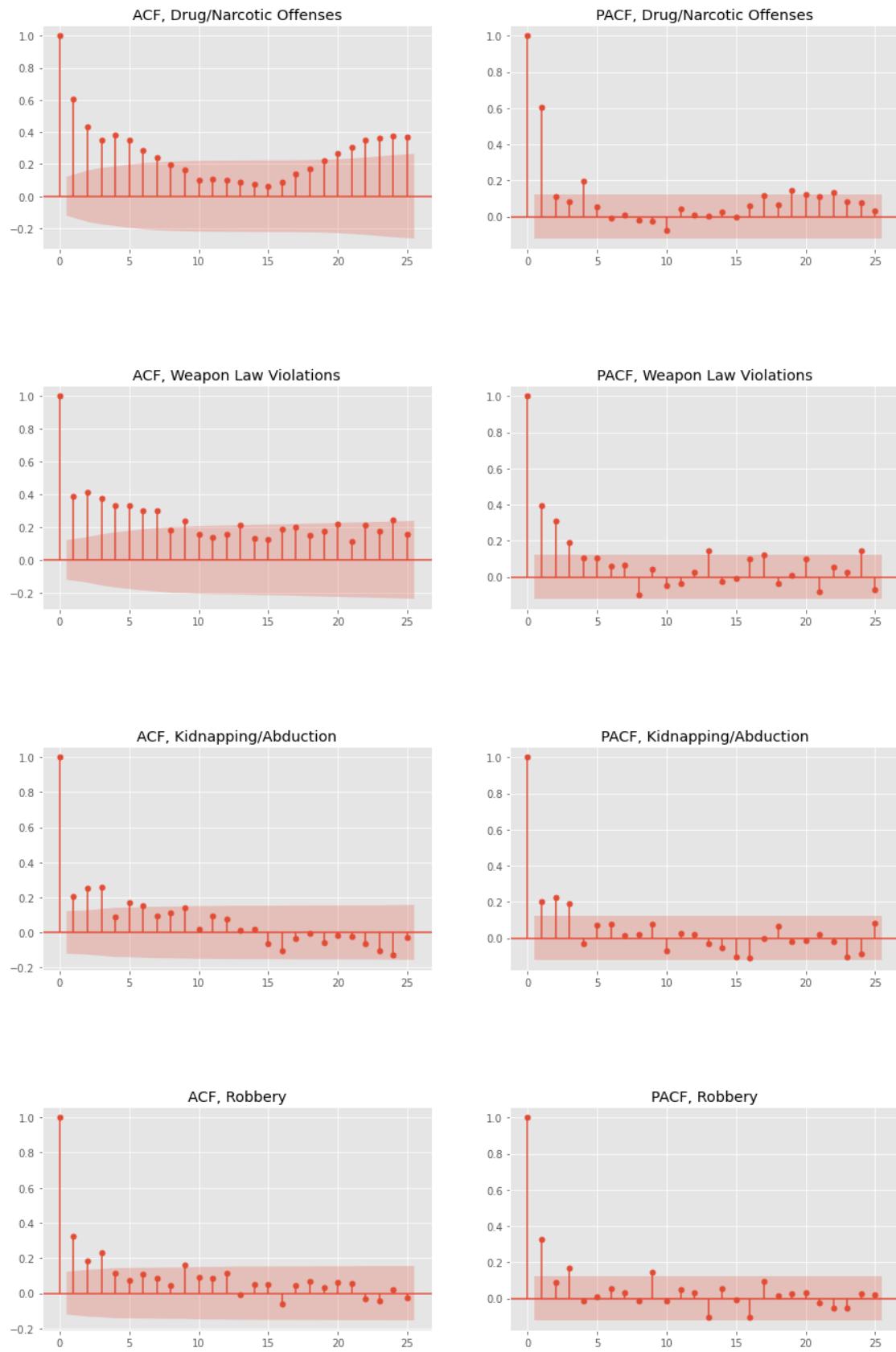
| Number | Critical value, 10% | Stationary? |
|--------|---------------------|-------------|
| 1 | -2.57289 | True |
| 2 | -2.57277 | True |
| 3 | -2.5728 | True |
| 4 | -2.5731 | True |
| 5 | -2.57285 | True |
| 6 | -2.57277 | True |
| 7 | -2.5728 | True |
| 8 | -2.57299 | True |
| 9 | -2.57285 | True |
| 10 | -2.57297 | True |
| 11 | -2.57486 | True |

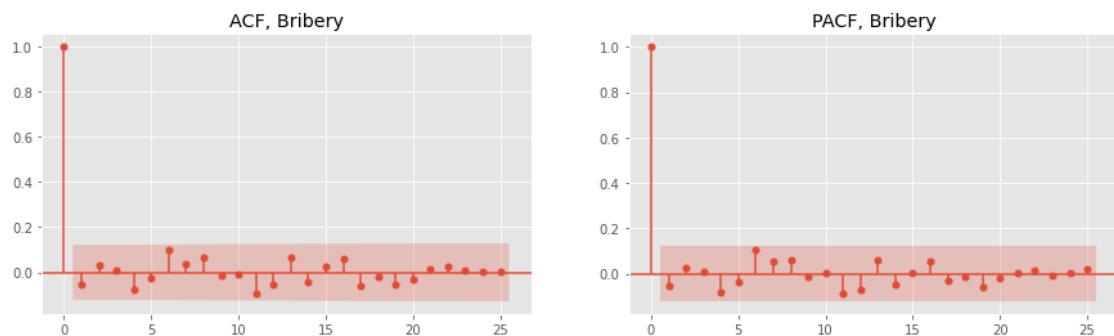
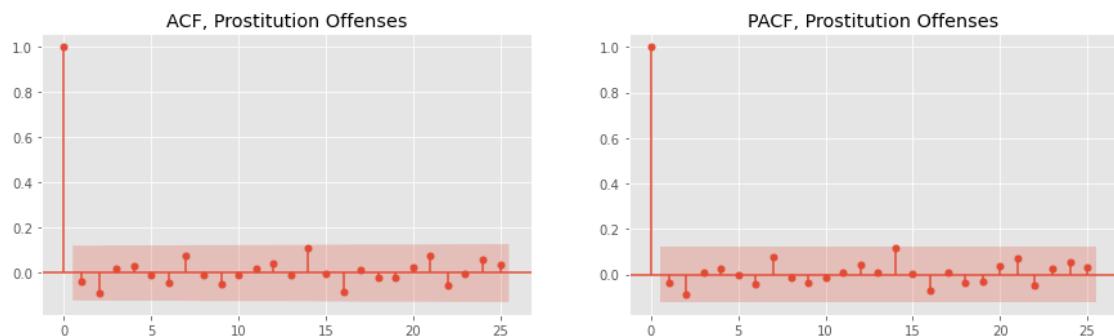
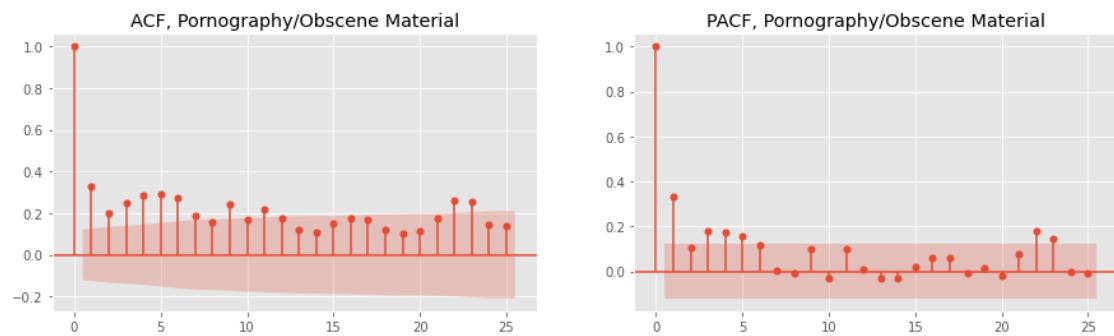
All **11** time-series got stationarized by the first differencing. There are two separate dictionaries for offenses categories: one with 12 original time-series, that were stationary from the get go and another one with 11 time-series that were pre-processed with the first differencing. The next step is to explore ACFs and PACFs of the timeframe and make a decision on the pdq and PDQs orders.

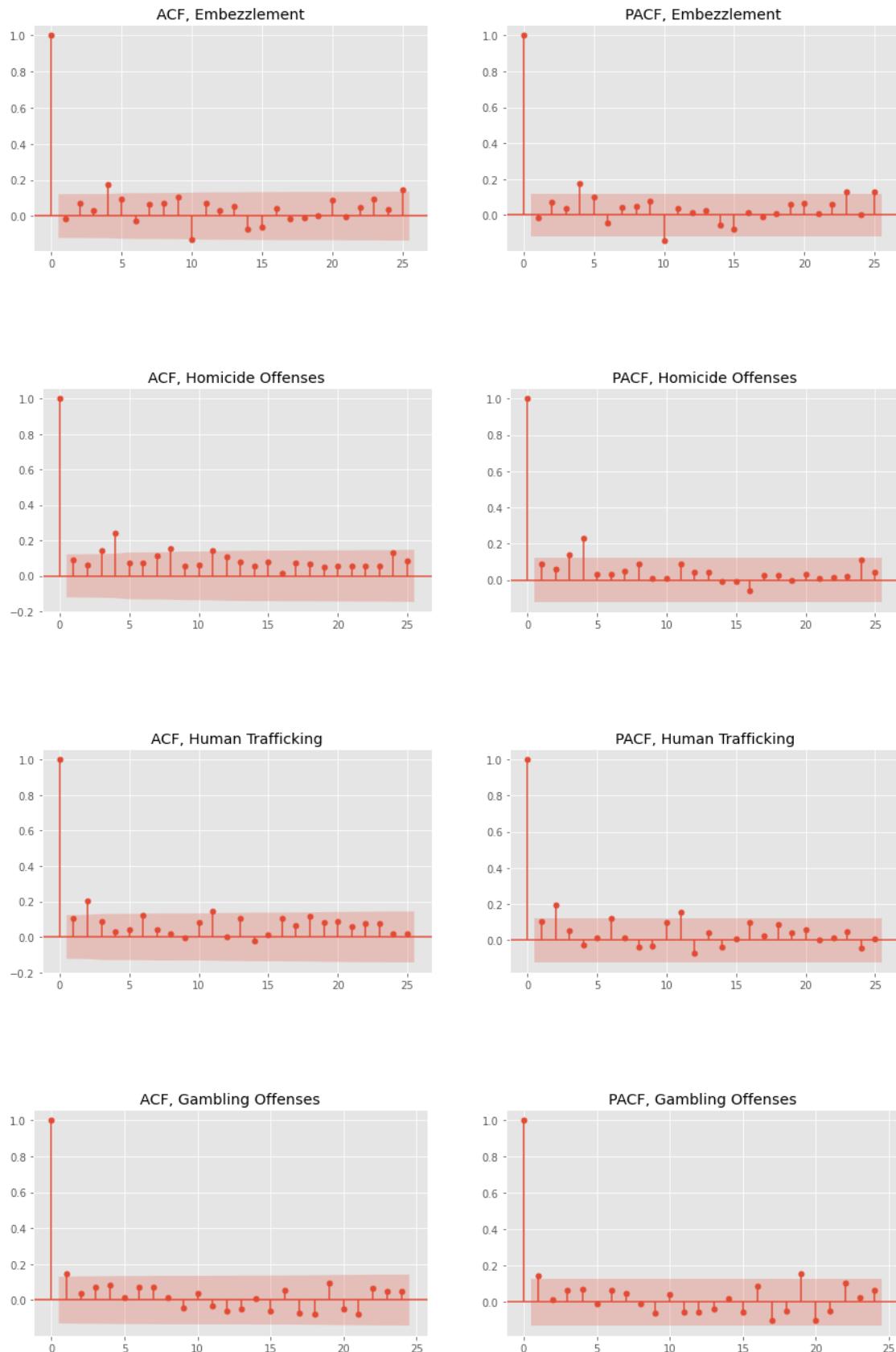
Exploring ACFs and PACFs of the originally stationary time-series

[10]: `ACF_PACF_multiple(ts_stationary1);`

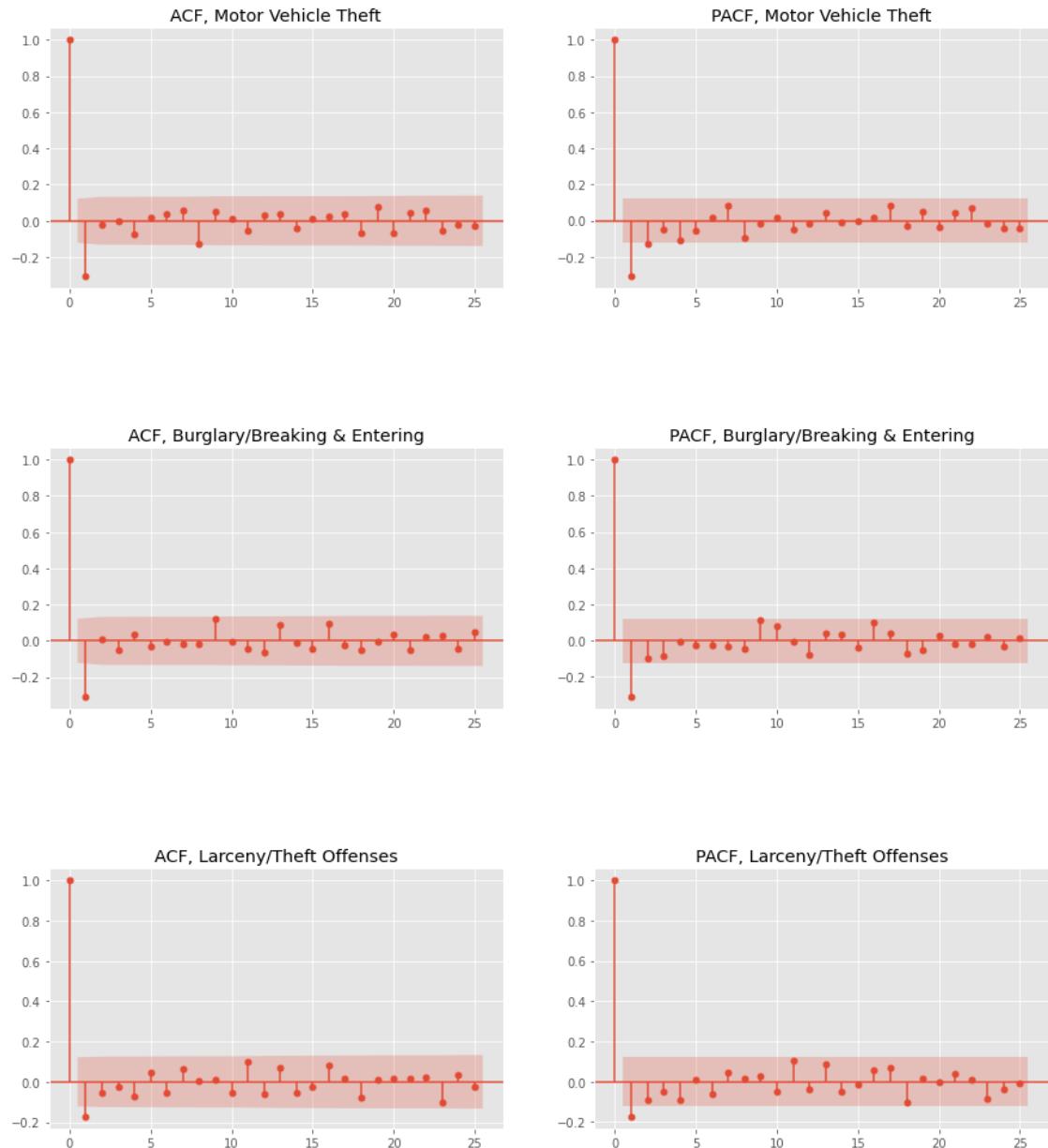


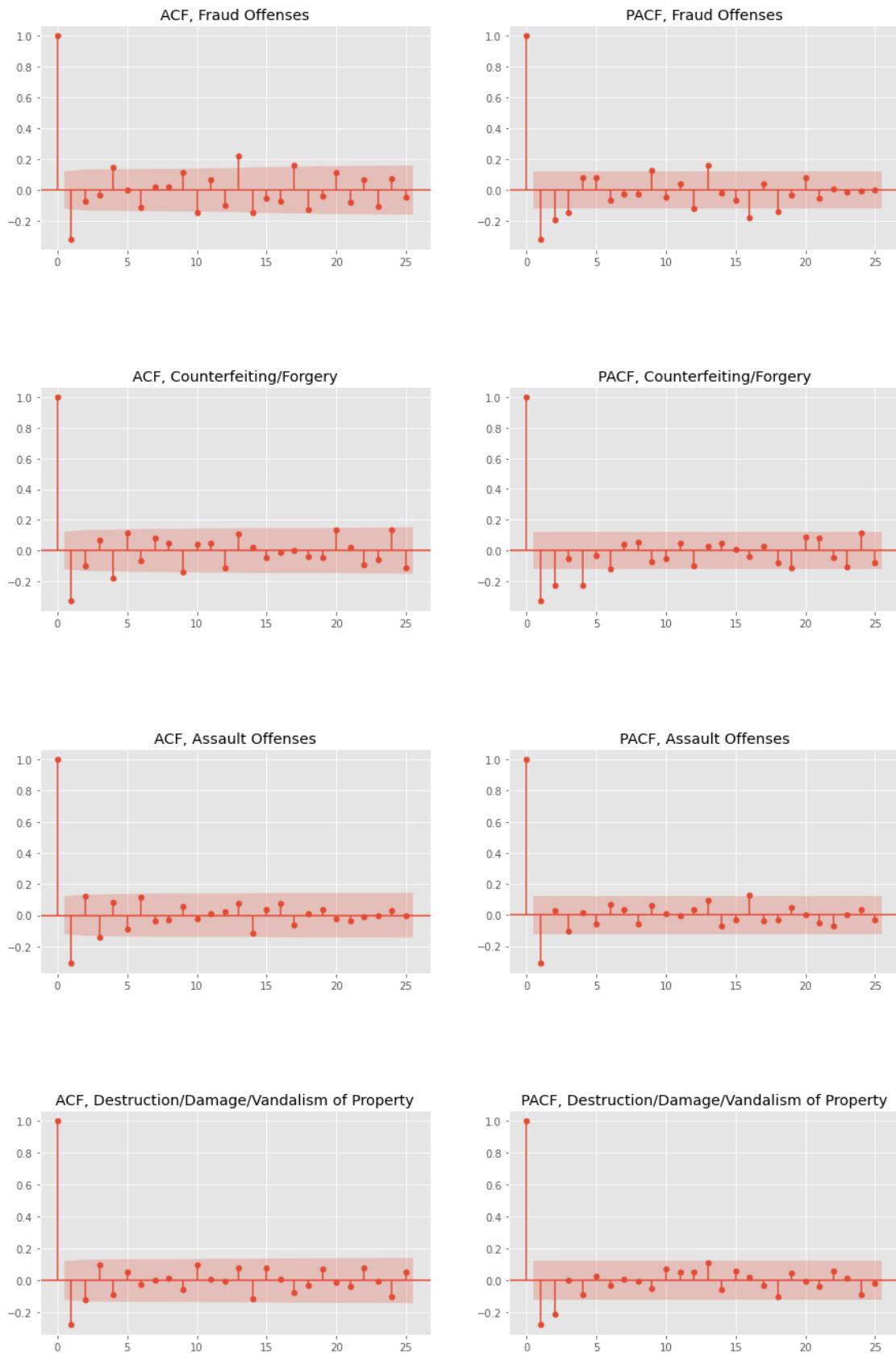


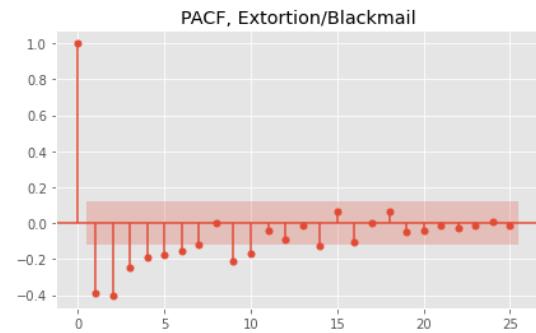
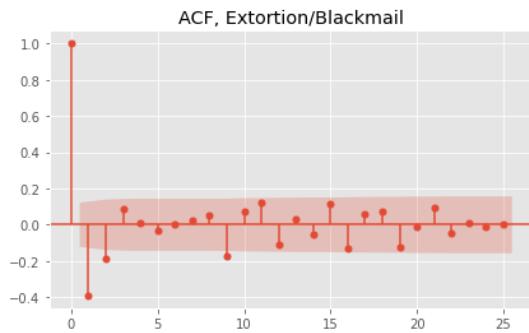
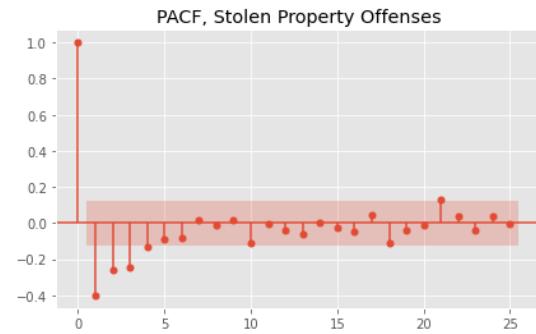
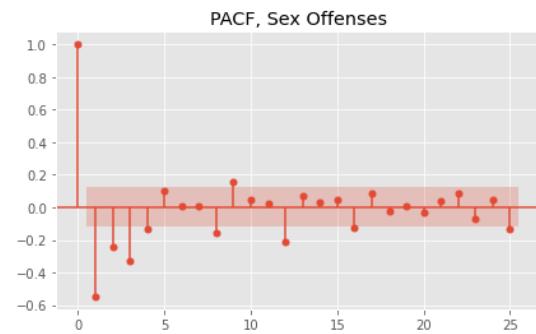
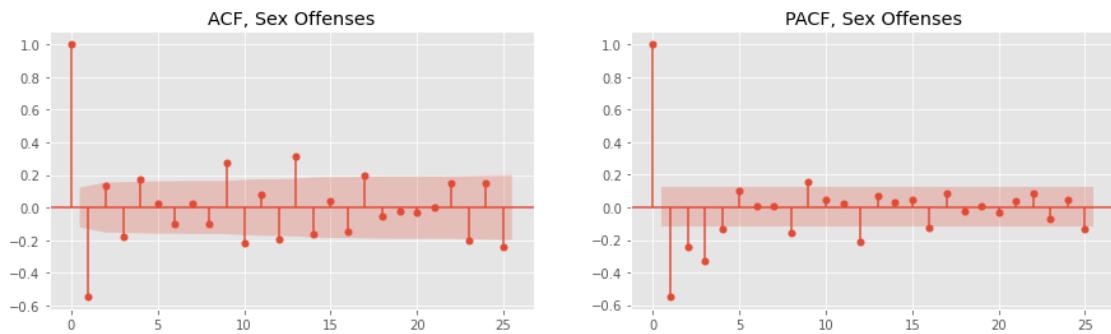


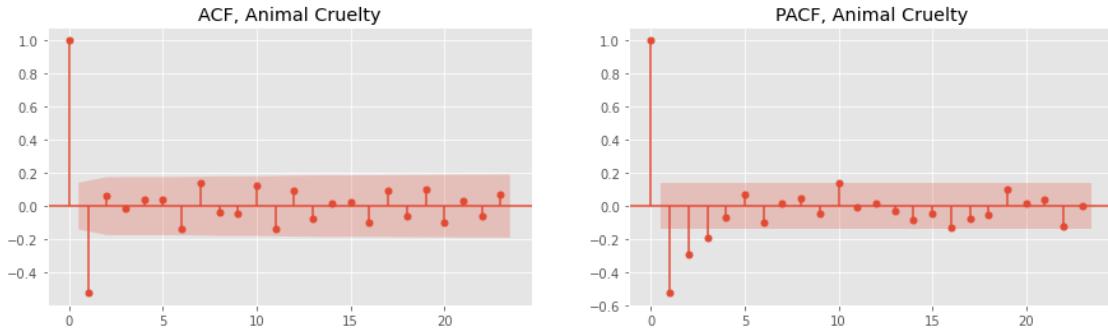


```
[11]: ACF_PACF_multiple(ts_stationary2);
```









6.3.3 Auto ARIMA for multiple categories of offenses

```
[12]: # RESULTS = []
# for crime, ts in TS_crime_category.items():
#     crime_categories_results = []

#     # Splitting it up
#     print('===='*20)
#     print('GRIDSEARCHING FOR {}'.format(crime))
#     train_size = round(len(ts) * 0.90)
#     ts_train, ts_test = ts[:train_size], ts[train_size:]

#     predictions_fig=display_figure_w_TSs(ts_train, ts_test, 'Training set', 'Test set',
#                                         'Training and Test Sets for Modeling {}'.format(crime), limit_=False)

#     #### Gridsearch

#     auto_model_train = pm.auto_arima(ts_train,
#                                     start_p=0,start_q=0, d=1,
#                                     start_P=1, start_Q=1, D=1,
#                                     max_p=2, max_q=2,
#                                     max_P=2, max_Q=2,
#                                     m=52, maxiter=150,
#                                     trace=False,verbose=True)

#     ## Fit SARIMAX with best parmas and compare forecast vs test
#     best_model = tsa.SARIMAX(ts_train,order=auto_model_train.order,
#                             seasonal_order = auto_model_train.
#                             seasonal_order,
#                             enforce_invertibility=False).fit()
```

```

#     ## Use diagnostics
#     diagnostics(best_model)

#     ## Prediction comparison
#     plt.style.use('ggplot')
#     y_hat_train=best_model.predict(typ='levels')
#     y_hat_test=best_model.predict(start=ts_test.index[0], end=ts_test.
#     ↪index[-1], typ='levels')

#     rmse = np.sqrt(mean_squared_error(ts_test, y_hat_test))
#     print('RMSE of the {} model for {}'.format(crime, round(rmse,2)))

#     predictions_fig=display_figure_w_TSs(ts_train, ts_test, 'Train set', ↪
#     'Test set',
#                                         'Training and Test Sets Raw Values and ↪
#     ↪Predictions, {}'.format(crime),
#     #                                         n=4, ts3=y_hat_test,
#     #                                         ts4=y_hat_train, label3='Prediction for Test ↪
#     ↪set', label4='Prediction for Training set',
#     #                                         limit_=False)

#     print('\tFINAL MODEL:')

#     final_model = tsa.SARIMAX(ts, order=auto_model_train.order,
#     seasonal_order = auto_model_train.seasonal_order,
#     enforce_invertibility=False).fit()

#     ## Plot forecast
#     forecast_fig=plot_predictions(ts, final_model, 'Forecast For Two Years ↪
#     ↪Forward, {}'.format(crime),
#     #                                         steps=104, xmin='2015')

#     ## Fill in results and
#     crime_categories_results['final_model'] = final_model
#     crime_categories_results['predict_fig'] = predictions_fig
#     crime_categories_results['forecast_fig'] = forecast_fig

#     ## Saving results to RESULTS dict
#     RESULTS[crime] = crime_categories_results

#     print("\n\n")

```

[13]: # cropped_RESULTS = {key:val for key, val in RESULTS.items() if ((key != 'Sex' ↪
Offenses')&(key != 'Weapon Law Violations'))}

```

[14]: # cropped_RESULTS.keys()

[15]: # with open('data/pickled_models/RESULTS1.pickle', 'wb') as f:
#       pickle.dump(cropped_RESULTS, f)

[16]: # TS_crime_category_to_rerun1={}
# TS_crime_category_to_rerun1['Sex Offenses']=TS_crime_category['Sex Offenses'].copy()
# TS_crime_category_to_rerun1['Weapon Law Violations']=TS_crime_category['Weapon Law Violations'].copy()

[17]: # RESULTS_second_run = {}
# for crime, ts in TS_crime_category_to_rerun1.items():
#     crime_categories_results = {}

#     # Splitting it up
#     print('===='*20)
#     print('GRIDSEARCHING FOR {}'.format(crime))
#     train_size = round(len(ts) * 0.90)
#     ts_train, ts_test = ts[:train_size], ts[train_size:]

#     predictions_fig=display_figure_w_TSs(ts_train, ts_test, 'Training set', 'Test set',
#                                         'Training and Test Sets for Modeling {}'.format(crime), limit_=False)

#     ##### Gridsearch
#     auto_model_train = pm.auto_arima(ts_train,
#                                     start_p=0,start_q=0, d=0,
#                                     start_P=0, start_Q=0, D=0,
#                                     max_p=2, max_q=2, max_d=1,
#                                     max_P=2, max_Q=2, max_D=1,
#                                     m=52, maxiter=300,
#                                     trace=False,verbose=True)

#     ## Fit SARIMAX with best parmas and compare forecast vs test
#     best_model = tsa.SARIMAX(ts_train,order=auto_model_train.order,
#                             seasonal_order = auto_model_train.
#                             seasonal_order,
#                             enforce_invertibility=False).fit()

#     ## Use diagnostics

```

```

#     diagnostics(best_model)

#     ## Prediction comparison
#     plt.style.use('ggplot')
#     y_hat_train=best_model.predict(typ='levels')
#     y_hat_test=best_model.predict(start=ts_test.index[0], end=ts_test.
#     ↪index[-1], typ='levels')

#     rmse = np.sqrt(mean_squared_error(ts_test, y_hat_test))
#     print('RMSE of the {} model for {}'.format(crime, round(rmse,2)))

#     predictions_fig=display_figure_w_TSs(ts_train, ts_test, 'Train set', ↪
#     ↪'Test set',
#                                         'Training and Test Sets Raw Values and ↪
#                                         ↪Predictions, {}'.format(crime),
#                                         n=4, ts3=y_hat_test,
#                                         #                                         ts4=y_hat_train, label3='Prediction for Test ↪
#                                         ↪set',
#                                         #                                         label4='Prediction for Training ↪
#                                         ↪set', limit_=False)

#     print('\tFINAL MODEL:')

#     final_model = tsa.SARIMAX(ts, order=auto_model_train.order,
#                             seasonal_order = auto_model_train.seasonal_order,
#                             enforce_invertibility=False).fit()

#     ## Plot forecast
#     forecast_fig=plot_predictions(ts, final_model, 'Forecast For Two Years ↪
#     ↪Forward, {}'.format(crime),
#                                     steps=104, xmin='2015')

#     ## Fill in results and
#     crime_categories_results['final_model'] = final_model
#     crime_categories_results['predict_fig'] = predictions_fig
#     crime_categories_results['forecast_fig'] = forecast_fig

#     ## Saving results to RESULTS dict
#     RESULTS_second_run[crime] = crime_categories_results

#     print("\n\n")

```

```
[18]: # with open('data/pickled_models/RESULTS2.pickle', 'wb') as f:
#     pickle.dump(RESULTS_second_run, f)
```

```
[19]: # def print_out_models(dictionary):
#     for crime, dict_ in dictionary.items():
#         print('OFFENSE CATEGORY: '+ crime)
#         for key, value in dict_.items():
#             if key=='final_model':
#                 print('\nTHE FINAL MODEL SUMMARY: \n')
#                 display(value.summary());
#                 display(value.plot_diagnostics(figsize=(15,7)));
#             elif key=='predict_fig':
#                 print('\nPREDICTION FOR TRAIN AND TEST sets: \n')
#                 display(value);
#             else:
#                 print('\nFORECAST: \n')
#                 display(value);
#     plt.close()
```

```
[20]: with open('data/pickled_models/RESULTS1.pickle', 'rb') as f:
    results1_back=pickle.load(f)
```

```
[21]: with open('data/pickled_models/RESULTS2.pickle', 'rb') as f:
    results2_back=pickle.load(f)
```

```
[22]: combined_results = {**results1_back, **results2_back}
```

```
[23]: print_out_models(combined_results)
```

OFFENSE CATEGORY: Motor Vehicle Theft

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""

```

SARIMAX Results

| Dep. Variable: | Motor Vehicle Theft | No. Observations: | 262 |
|----------------|---------------------------------|-------------------|-----------|
| Model: | SARIMAX(0, 1, 1)x(2, 1, [], 52) | Log Likelihood | -1087.379 |
| Date: | Fri, 16 Jul 2021 | AIC | 2182.757 |
| Time: | 12:23:58 | BIC | 2196.126 |
| Sample: | 01-04-2015 - 01-05-2020 | HQIC | 2188.162 |

Covariance Type: opg

| | coef | std err | z | P> z | [0.025 | 0.975] |
|----------|---------|---------|---------|-------|--------|--------|
| ma.L1 | -0.7592 | 0.059 | -12.913 | 0.000 | -0.874 | -0.644 |
| ar.S.L52 | -0.5590 | 0.088 | -6.333 | 0.000 | -0.732 | -0.386 |

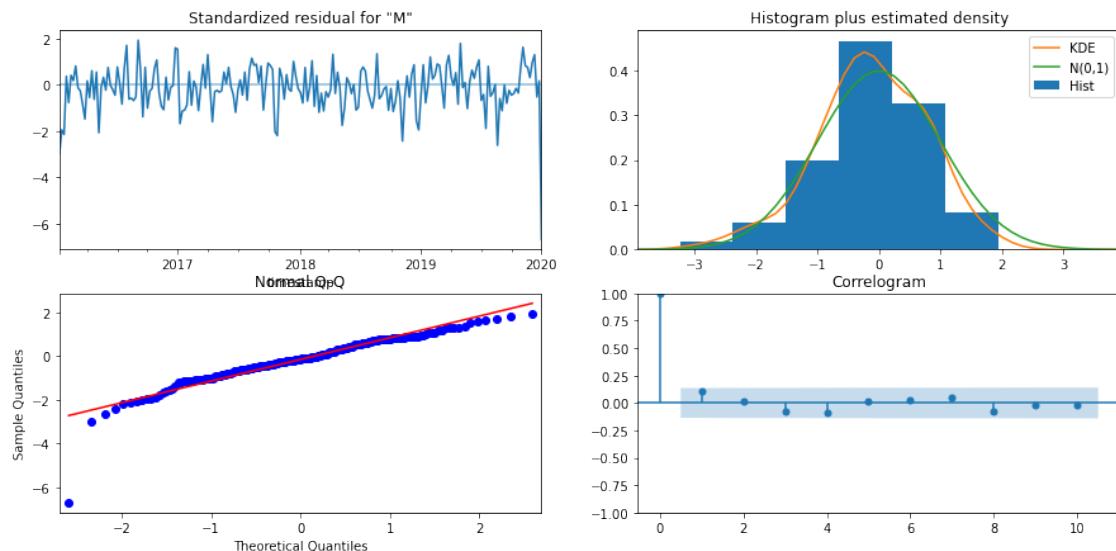
```

ar.S.L104      -0.2983      0.128      -2.327      0.020      -0.550      -0.047
sigma2       1748.7905    94.113     18.582      0.000    1564.332    1933.249
=====
Ljung-Box (L1) (Q):           2.41   Jarque-Bera (JB):        666.52
Prob(Q):                   0.12   Prob(JB):            0.00
Heteroskedasticity (H):      1.53   Skew:                  -1.56
Prob(H) (two-sided):         0.08   Kurtosis:             11.17
=====
```

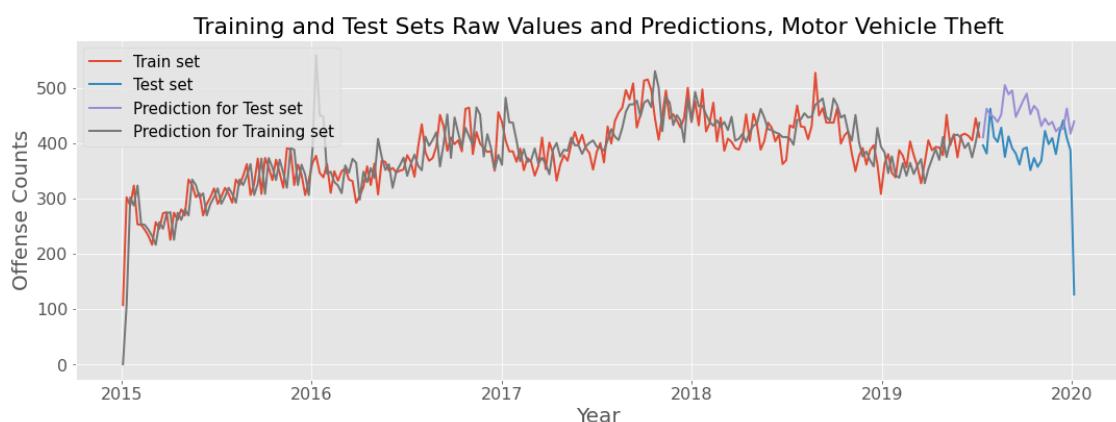
Warnings:

```
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

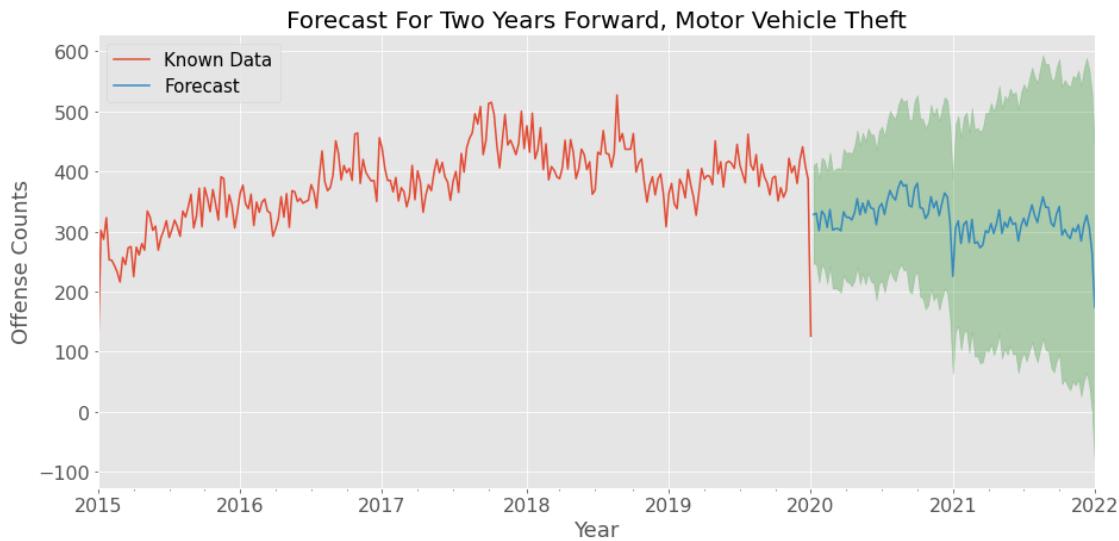
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Burglary/Breaking & Entering

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
                                          SARIMAX Results
=====
Dep. Variable:          Burglary/Breaking & Entering    No. Observations:                  262
Model:                 SARIMAX(0, 1, 2)x(2, 1, [], 52)    Log Likelihood:                -1078.874
Date:                   Fri, 16 Jul 2021      AIC:                            2167.749
Time:                      12:23:59      BIC:                            2184.460
Sample:                 01-04-2015      HQIC:                           2174.505
                           - 01-05-2020
Covariance Type:                    opg
=====
            coef      std err           z      P>|z|      [0.025      0.975]
-----
ma.L1     -0.7629      0.073     -10.485      0.000     -0.906     -0.620
ma.L2     -0.1181      0.074     -1.594      0.111     -0.263      0.027
ar.S.L52   -0.4278      0.080     -5.334      0.000     -0.585     -0.271
ar.S.L104  -0.4115      0.113     -3.644      0.000     -0.633     -0.190
```

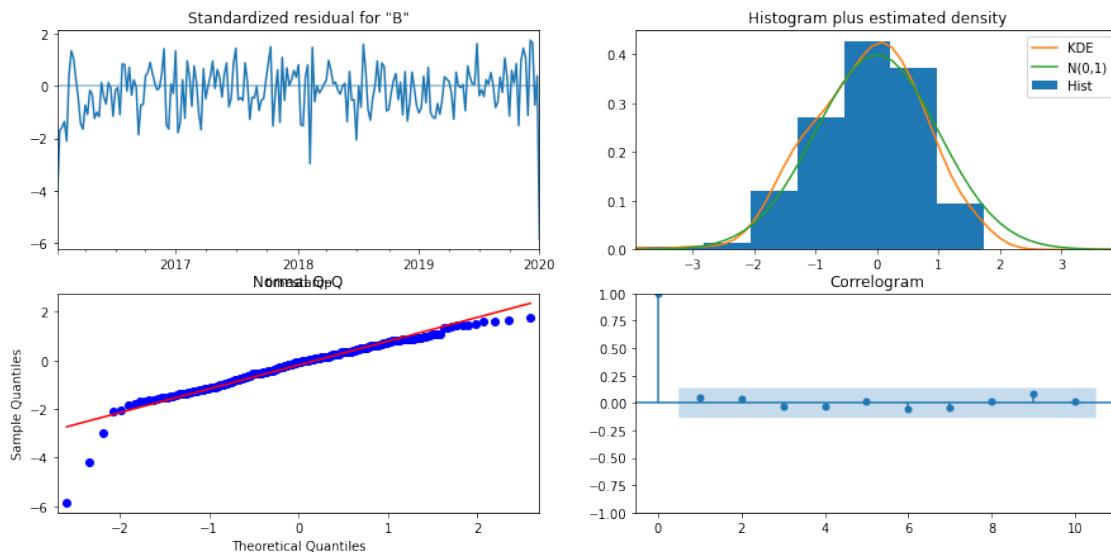
```

sigma2      1577.3598    156.099     10.105      0.000    1271.411    1883.309
=====
Ljung-Box (L1) (Q):           0.63  Jarque-Bera (JB):        291.53
Prob(Q):                   0.43  Prob(JB):            0.00
Heteroskedasticity (H):      0.99  Skew:                -1.27
Prob(H) (two-sided):         0.96  Kurtosis:             8.20
=====
```

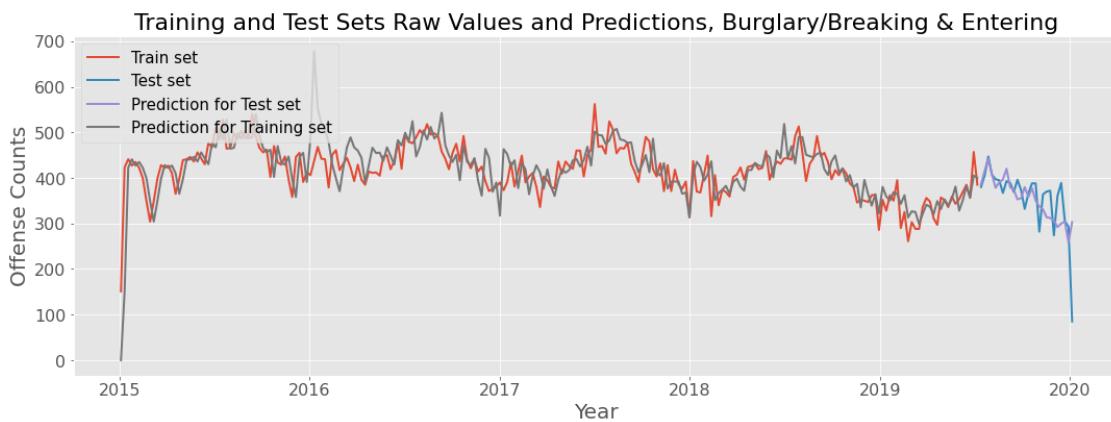
Warnings:

```
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

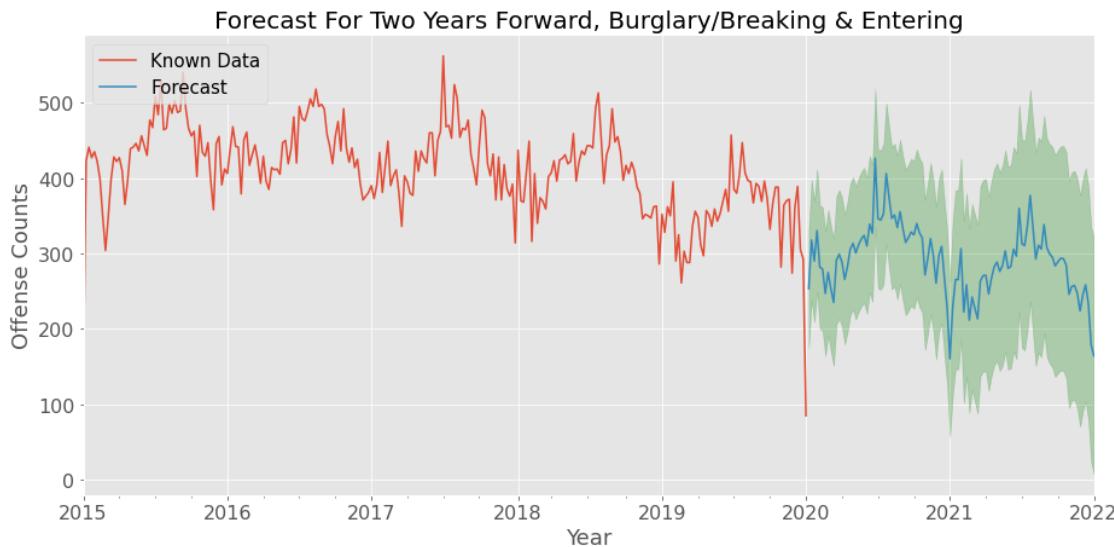
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Larceny/Theft Offenses

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
=====
          SARIMAX Results
=====

Dep. Variable:           Larceny/Theft Offenses    No. Observations:                 262
Model:                  SARIMAX(0, 1, 1)x(2, 1, []), 52    Log Likelihood:                -1381.281
Date:                   Fri, 16 Jul 2021            AIC:                            2770.561
Time:                   12:24:00                    BIC:                            2783.930
Sample:                 01-04-2015 : - 01-05-2020    HQIC:                           2775.966
Covariance Type:             opg
=====
              coef      std err       z   P>|z|      [0.025      0.975]
-----
ma.L1      -0.8804      0.043   -20.302      0.000     -0.965     -0.795
ar.S.L52    -0.5155      0.091    -5.688      0.000     -0.693     -0.338
ar.S.L104   -0.3554      0.147    -2.411      0.016     -0.644     -0.066
```

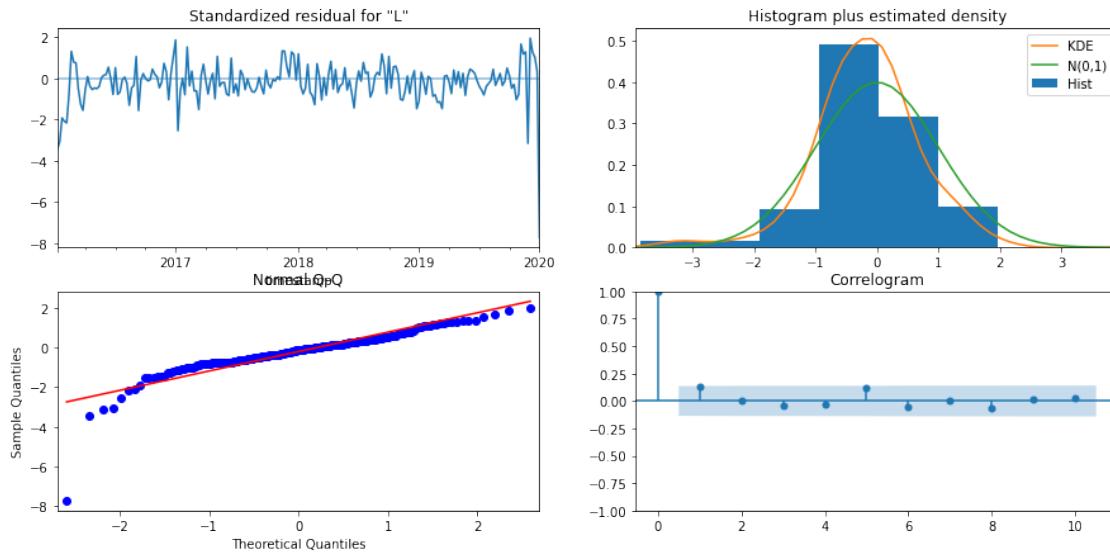
```

sigma2      2.886e+04   2762.625     10.446      0.000    2.34e+04   3.43e+04
=====
Ljung-Box (L1) (Q):           3.34  Jarque-Bera (JB):        2487.91
Prob(Q):                      0.07  Prob(JB):                  0.00
Heteroskedasticity (H):       1.44  Skew:                     -2.43
Prob(H) (two-sided):          0.13  Kurtosis:                 19.19
=====
```

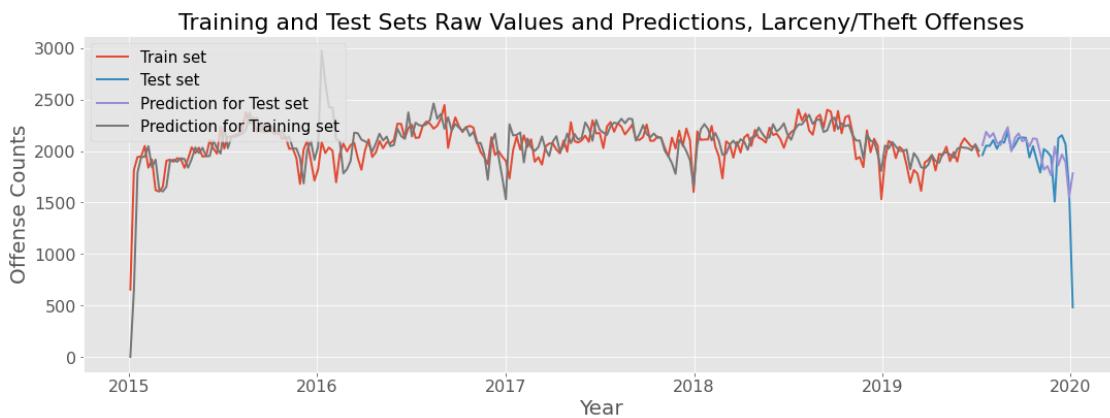
Warnings:

```
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

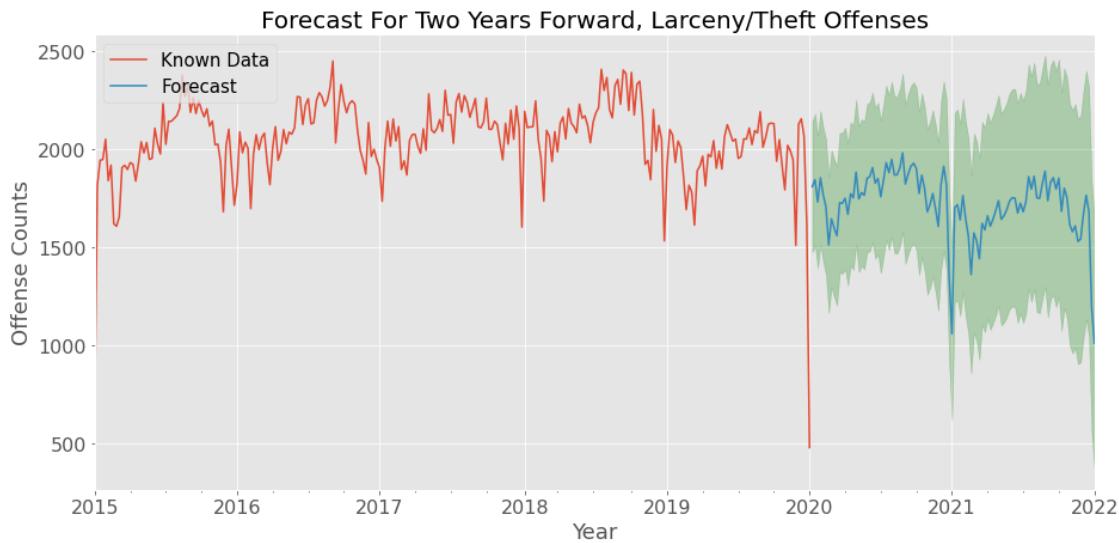
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Fraud Offenses

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
=====
          SARIMAX Results
=====
Dep. Variable:                  Fraud Offenses    No. Observations:                   262
Model: SARIMAX(0, 1, 1)x(2, 1, []) , 52    Log Likelihood:                -1138.743
Date: Fri, 16 Jul 2021            AIC:                         2285.485
Time: 12:24:01                    BIC:                         2298.854
Sample: 01-04-2015 : - 01-05-2020    HQIC:                         2290.890
Covariance Type: opg
=====
              coef      std err           z      P>|z|      [0.025      0.975]
-----
ma.L1     -0.6526      0.050   -12.950      0.000     -0.751     -0.554
ar.S.L52  -0.4171      0.073    -5.698      0.000     -0.561     -0.274
ar.S.L104 -0.4076      0.072    -5.645      0.000     -0.549     -0.266
```

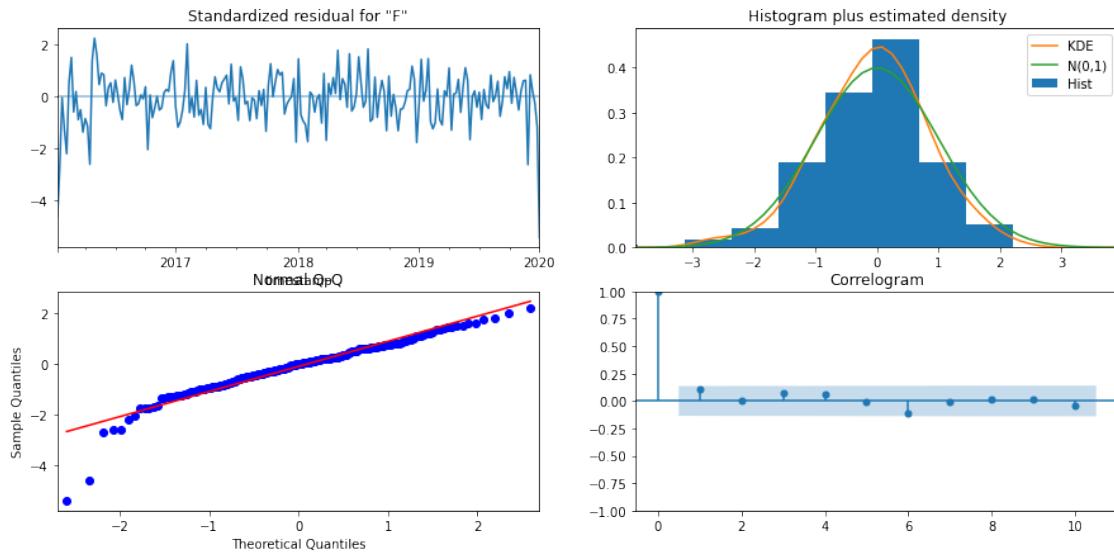
```

sigma2      2819.6525    226.429     12.453      0.000    2375.859    3263.446
=====
Ljung-Box (L1) (Q):           2.22  Jarque-Bera (JB):       258.78
Prob(Q):                   0.14  Prob(JB):            0.00
Heteroskedasticity (H):      0.84  Skew:              -1.23
Prob(H) (two-sided):        0.47  Kurtosis:          7.86
=====
```

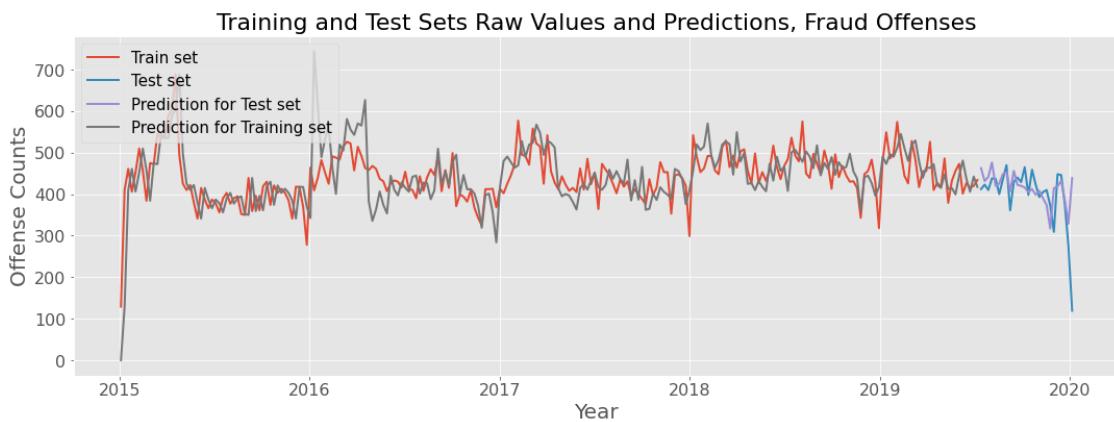
Warnings:

```
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

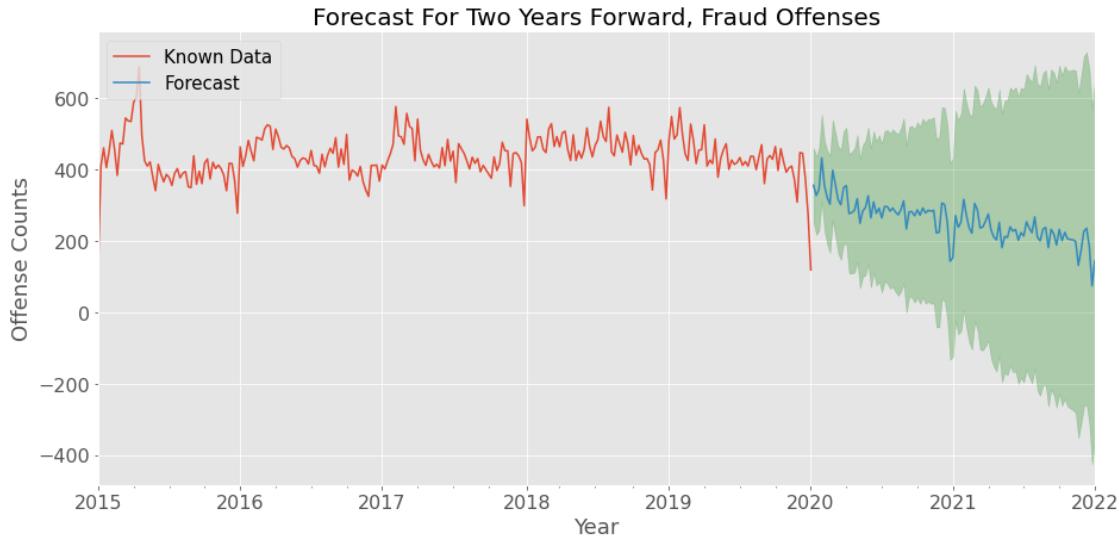
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Counterfeiting/Forgery

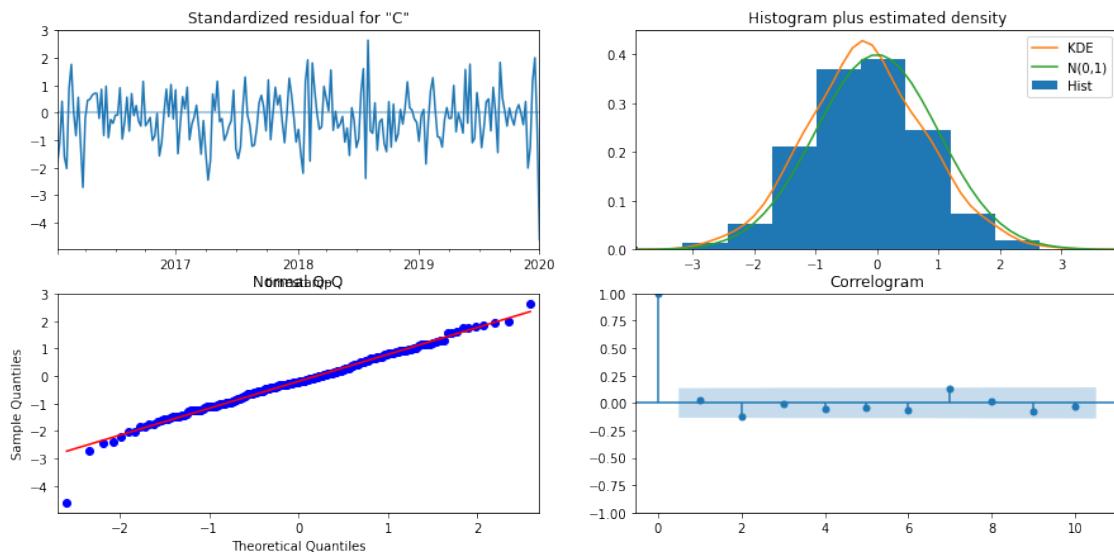
THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
                                          SARIMAX Results
=====
Dep. Variable:          Counterfeiting/Forgery    No. Observations:                 262
Model:                  SARIMAX(0, 1, 1)x(2, 1, [], 52)    Log Likelihood:                -871.182
Date:                   Fri, 16 Jul 2021      AIC:                            1750.364
Time:                     12:24:01        BIC:                            1763.734
Sample:                 01-04-2015      HQIC:                           1755.770
                           - 01-05-2020
Covariance Type:                    opg
=====
              coef    std err         z      P>|z|      [0.025      0.975]
-----
ma.L1       -0.8404     0.041   -20.468      0.000     -0.921     -0.760
ar.S.L52    -0.5933     0.089    -6.697      0.000     -0.767     -0.420
ar.S.L104   -0.3472     0.097    -3.563      0.000     -0.538     -0.156
sigma2      216.0399   16.310    13.246      0.000    184.073    248.007
```

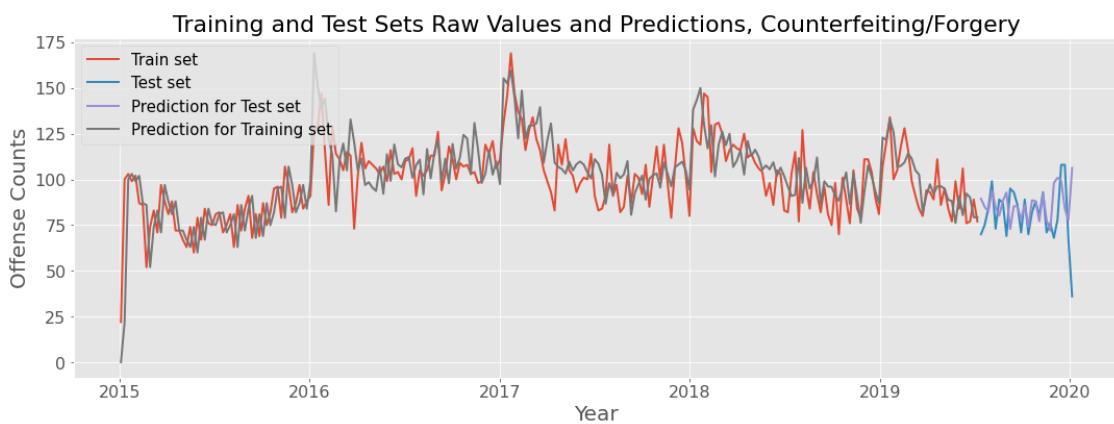
| | | | |
|-------------------------|------|-------------------|-------|
| Ljung-Box (L1) (Q): | 0.11 | Jarque-Bera (JB): | 21.25 |
| Prob(Q): | 0.74 | Prob(JB): | 0.00 |
| Heteroskedasticity (H): | 1.13 | Skew: | -0.31 |
| Prob(H) (two-sided): | 0.61 | Kurtosis: | 4.44 |

Warnings:

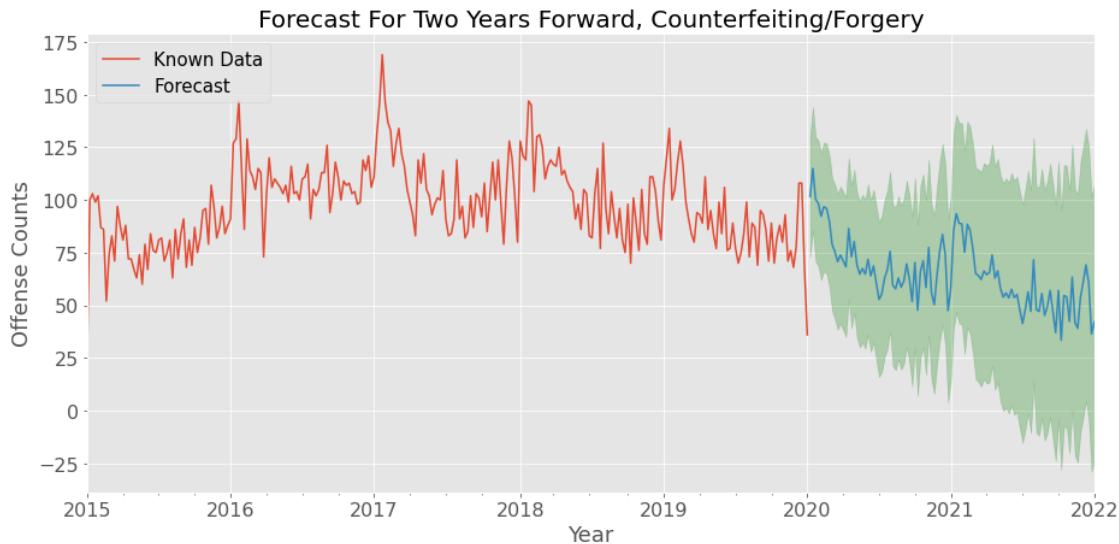
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
 """



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Assault Offenses

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
=====
SARIMAX Results
=====
Dep. Variable:          Assault Offenses    No. Observations:                 262
Model:             SARIMAX(2, 1, 1)x(2, 1, []), 52    Log Likelihood:                -1207.539
Date:                Fri, 16 Jul 2021    AIC:                            2427.078
Time:                      12:24:02    BIC:                            2447.132
Sample:                01-04-2015    HQIC:                           2435.186
                           - 01-05-2020
Covariance Type:                    opg
=====
              coef      std err           z      P>|z|      [0.025      0.975]
-----
ar.L1      0.1210      0.109      1.108      0.268     -0.093      0.335
ar.L2      0.1233      0.111      1.113      0.266     -0.094      0.340
ma.L1     -0.9078      0.063     -14.431      0.000     -1.031     -0.785
ar.S.L52   -0.4029      0.101     -3.985      0.000     -0.601     -0.205
```

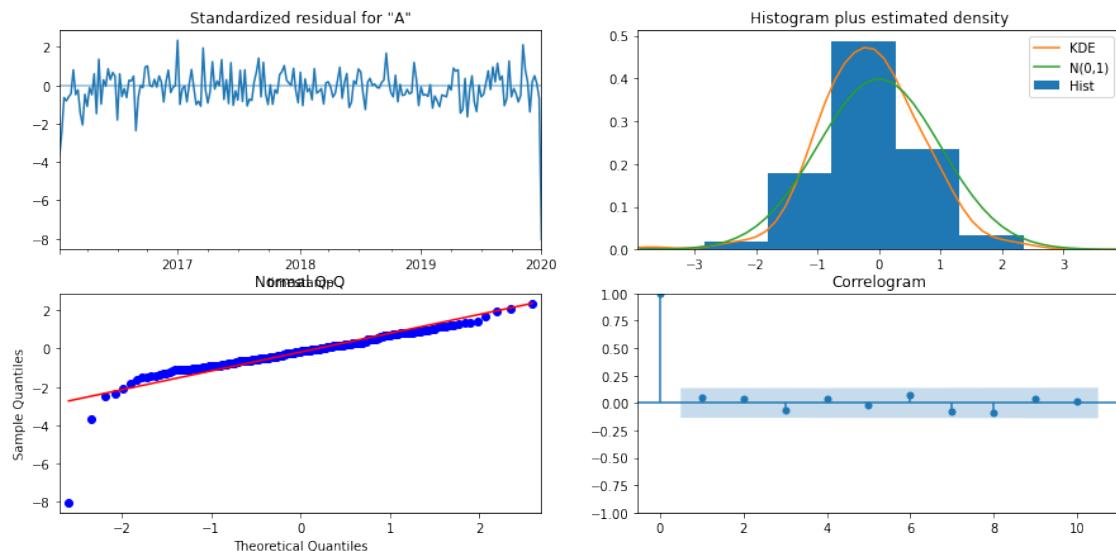
```

ar.S.L104      -0.3213      0.119      -2.700      0.007      -0.555      -0.088
sigma2       5617.1717    295.661     18.999      0.000     5037.687    6196.656
=====
Ljung-Box (L1) (Q):          0.50  Jarque-Bera (JB):        3272.61
Prob(Q):                  0.48  Prob(JB):            0.00
Heteroskedasticity (H):    1.53  Skew:                 -2.49
Prob(H) (two-sided):      0.08  Kurtosis:           21.73
=====
```

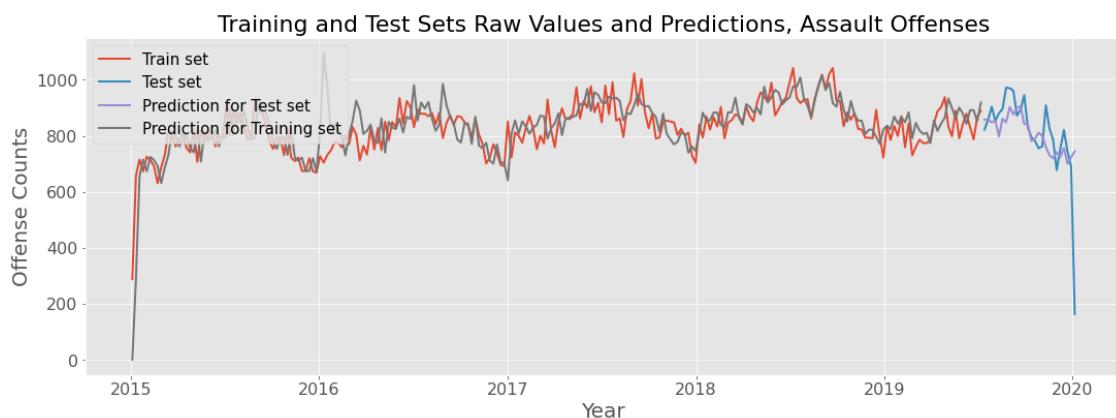
Warnings:

```
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

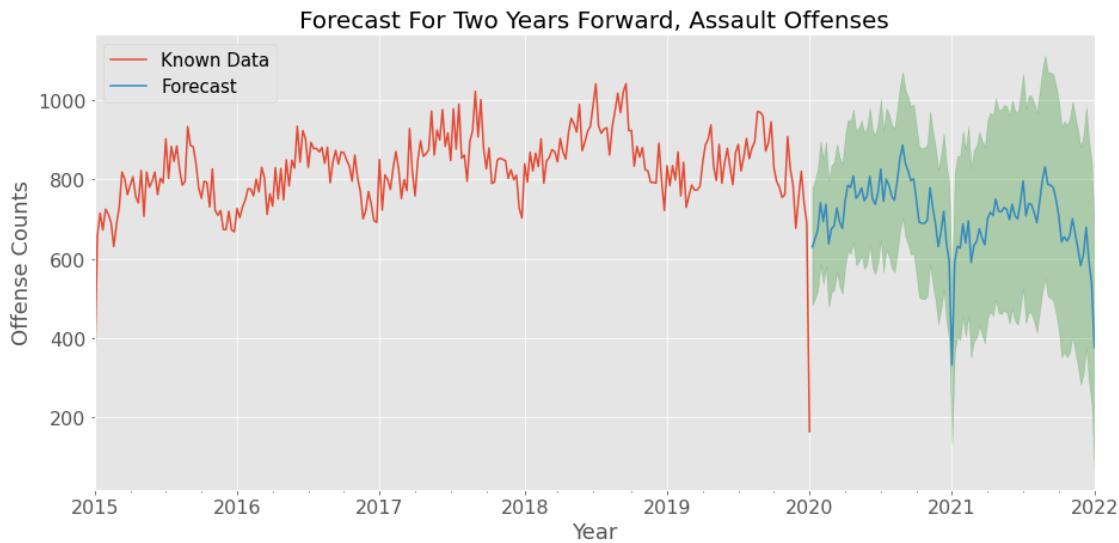
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Destruction/Damage/Vandalism of Property

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
=====
          SARIMAX Results
=====
Dep. Variable:      Destruction/Damage/Vandalism of Property    No. Observations:      -124
Model:                  SARIMAX(1, 1, 2)x(2, 1, [], 52)    Log Likelihood:        248
Date:                      Fri, 16 Jul 2021    AIC:                  256
Time:                          12:24:03    BIC:                  256
Sample:                      01-04-2015    HQIC:                 256
                           - 01-05-2020
Covariance Type:             opg
=====
              coef    std err        z     P>|z|    [0.025    0.975]
-----
ar.L1       -0.4574    0.234   -1.954     0.051    -0.916     0.001
ma.L1       -0.2473    0.229   -1.078     0.281    -0.697     0.202
ma.L2       -0.6048    0.200   -3.030     0.002    -0.996    -0.214
```

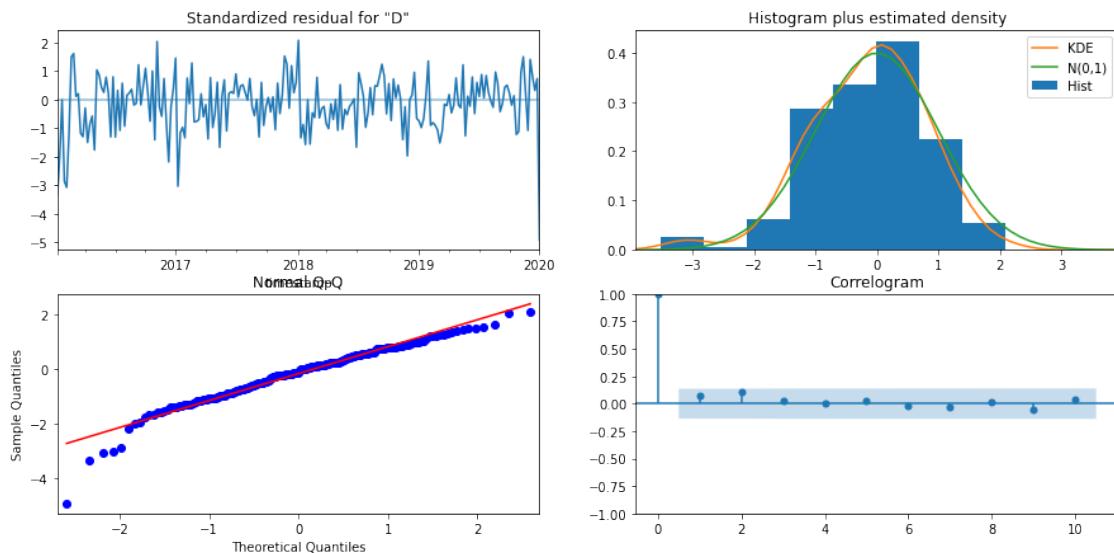
```

ar.S.L52      -0.5324      0.084      -6.323      0.000      -0.697      -0.367
ar.S.L104     -0.3604      0.092      -3.904      0.000      -0.541      -0.179
sigma2       7448.8908    862.754      8.634      0.000    5757.925    9139.857
=====
Ljung-Box (L1) (Q):                  0.98   Jarque-Bera (JB):        78.67
Prob(Q):                           0.32   Prob(JB):            0.00
Heteroskedasticity (H):             0.67   Skew:                 -0.88
Prob(H) (two-sided):                0.10   Kurtosis:           5.44
=====
```

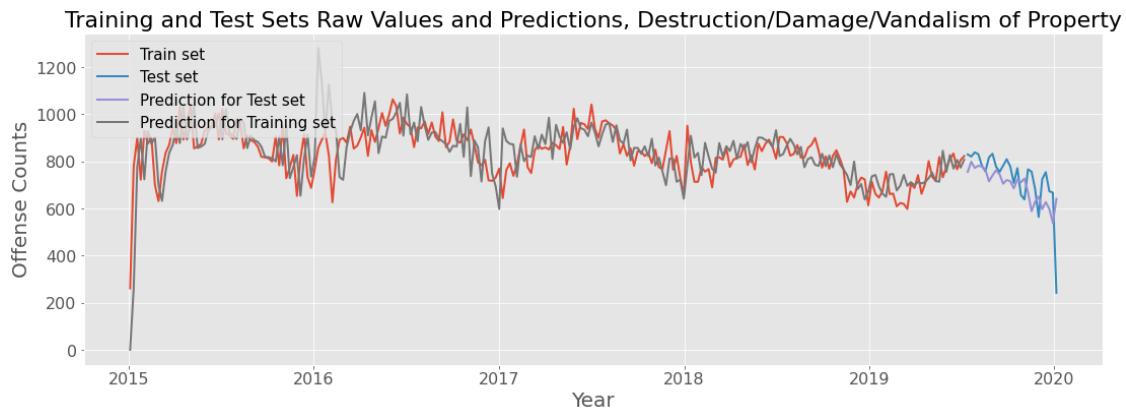
Warnings:

```
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

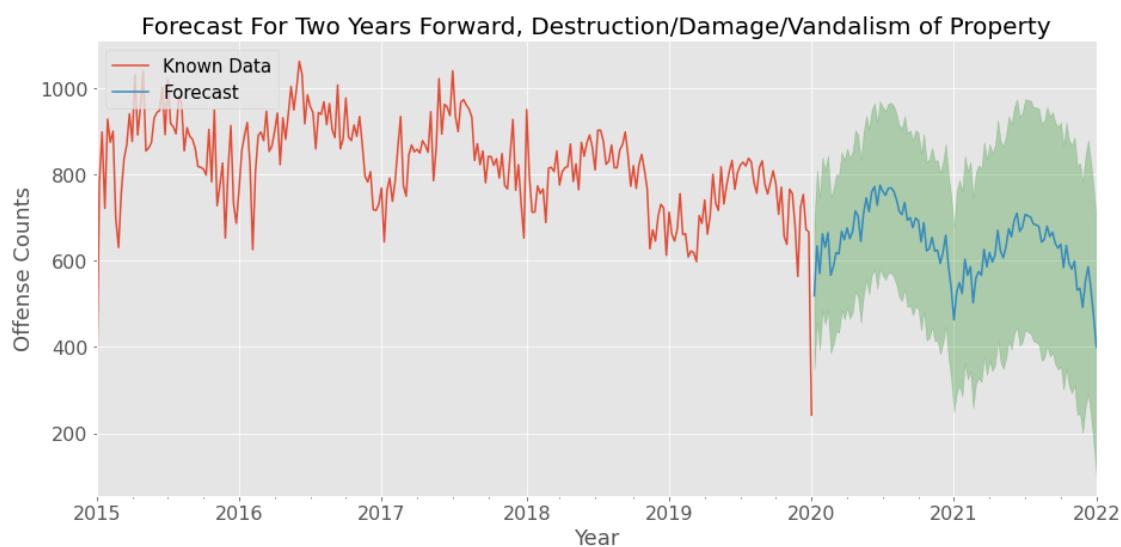
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Arson

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
```

```
"""
```

SARIMAX Results

```
=====
```

Dep. Variable:

Arson No. Observations:

262

```

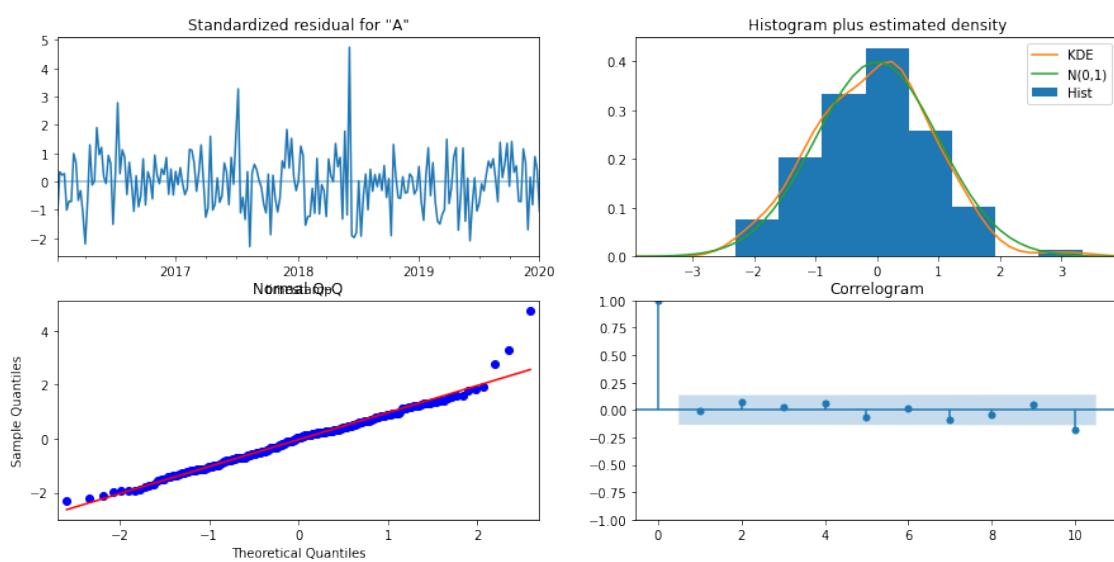
Model: SARIMAX(1, 1, 1)x(2, 1, []) , 52 Log Likelihood           -722.935
Date: Fri, 16 Jul 2021 AIC                                1455.870
Time: 12:24:04 BIC                                1472.582
Sample: 01-04-2015 HQIC                               1462.627
- 01-05-2020
Covariance Type: opg
=====

            coef    std err      z   P>|z|   [0.025   0.975]
-----
ar.L1      0.1322    0.067   1.974    0.048    0.001    0.264
ma.L1     -0.9588    0.028 -33.989    0.000   -1.014   -0.904
ar.S.L52   -0.5102    0.070  -7.311    0.000   -0.647   -0.373
ar.S.L104  -0.3924    0.087  -4.493    0.000   -0.564   -0.221
sigma2     51.9199    3.810  13.626    0.000  44.452  59.388
=====

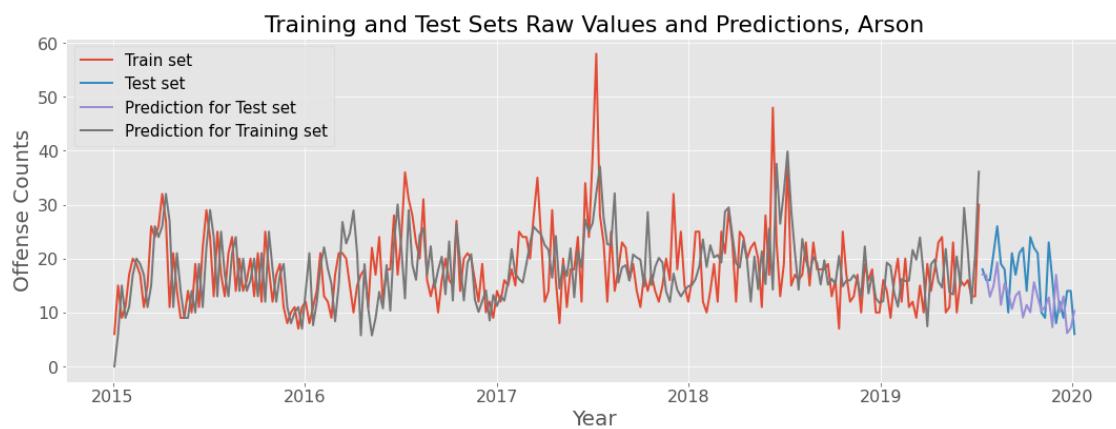
Ljung-Box (L1) (Q):          0.03 Jarque-Bera (JB):        47.51
Prob(Q):                   0.86 Prob(JB):                  0.00
Heteroskedasticity (H):     1.04 Skew:                      0.57
Prob(H) (two-sided):       0.86 Kurtosis:                 5.04
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

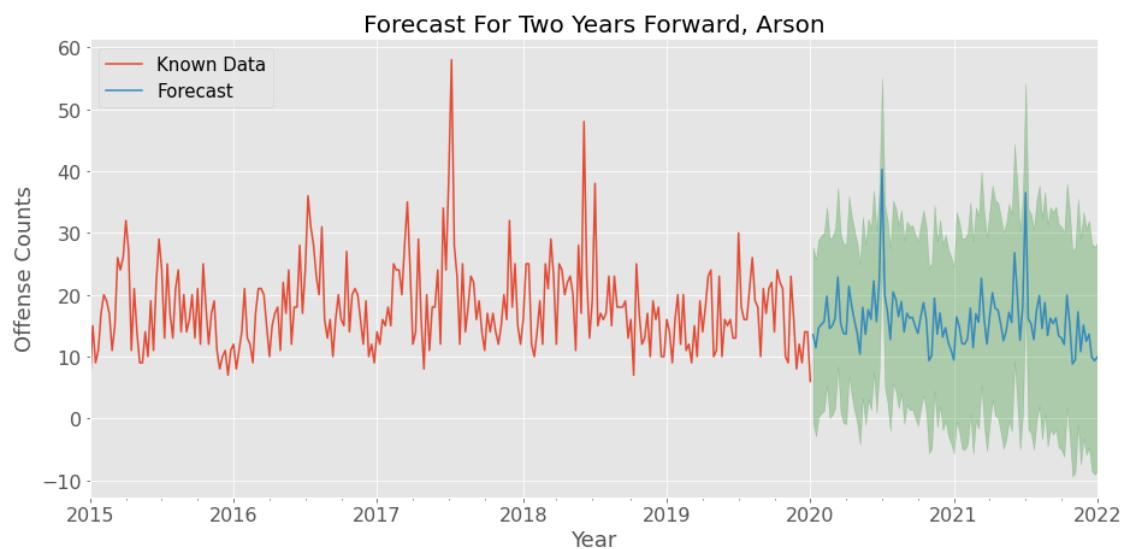
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Drug/Narcotic Offenses

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""

```

SARIMAX Results

Dep. Variable:

Drug/Narcotic Offenses

No. Observations:

262

```

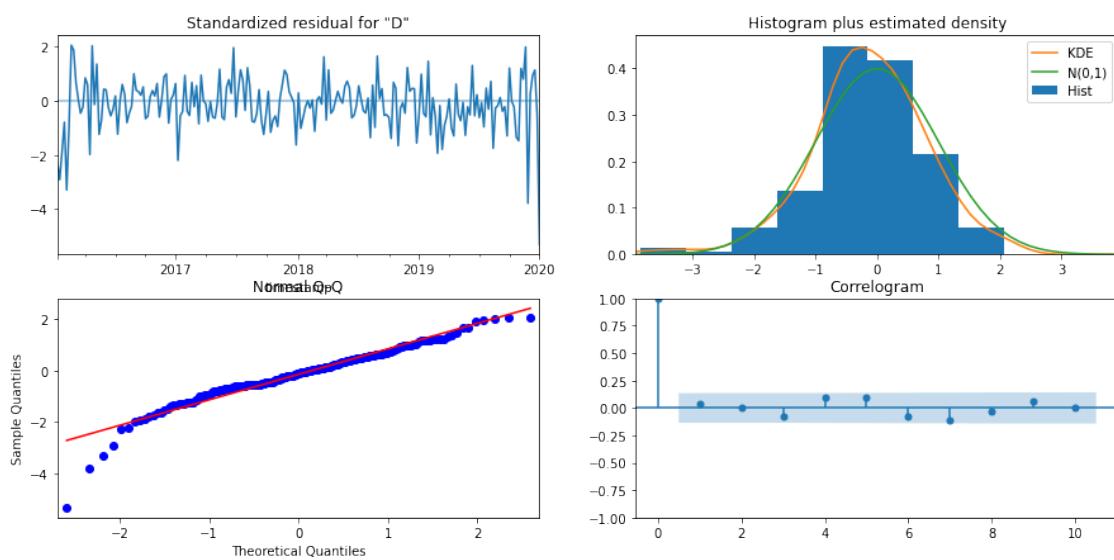
Model: SARIMAX(1, 1, 1)x(2, 1, []) , 52 Log Likelihood           -1184.394
Date: Fri, 16 Jul 2021 AIC                                2378.788
Time: 12:24:04 BIC                                2395.500
Sample: 01-04-2015 HQIC                               2385.545
        - 01-05-2020
Covariance Type: opg
=====

            coef    std err      z   P>|z|    [0.025    0.975]
-----
ar.L1      0.1509    0.076   1.994    0.046    0.003    0.299
ma.L1     -0.8906    0.047 -18.944    0.000   -0.983   -0.798
ar.S.L52   -0.4894    0.099  -4.934    0.000   -0.684   -0.295
ar.S.L104  -0.2756    0.104  -2.642    0.008   -0.480   -0.071
sigma2    4502.3816  326.682 13.782    0.000 3862.097 5142.666
=====

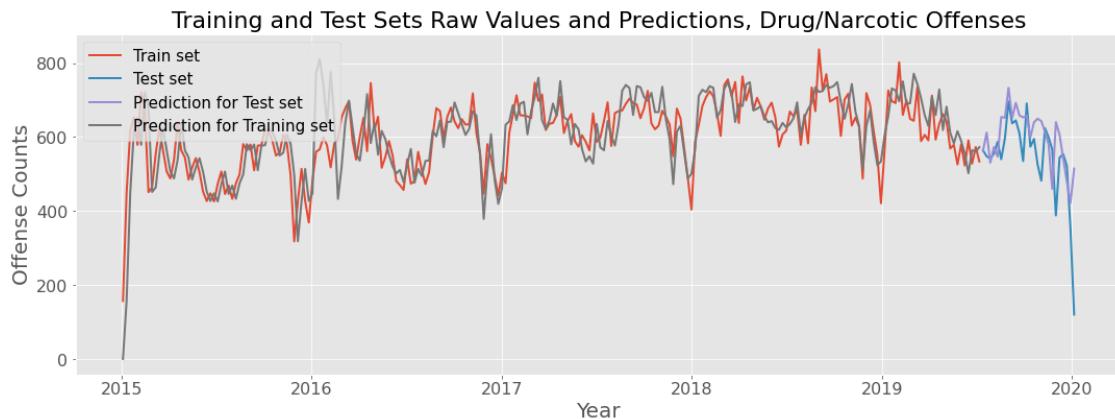
Ljung-Box (L1) (Q):          0.29 Jarque-Bera (JB):       162.75
Prob(Q):                   0.59 Prob(JB):                  0.00
Heteroskedasticity (H):     1.36 Skew:                      -0.99
Prob(H) (two-sided):        0.20 Kurtosis:                 6.85
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

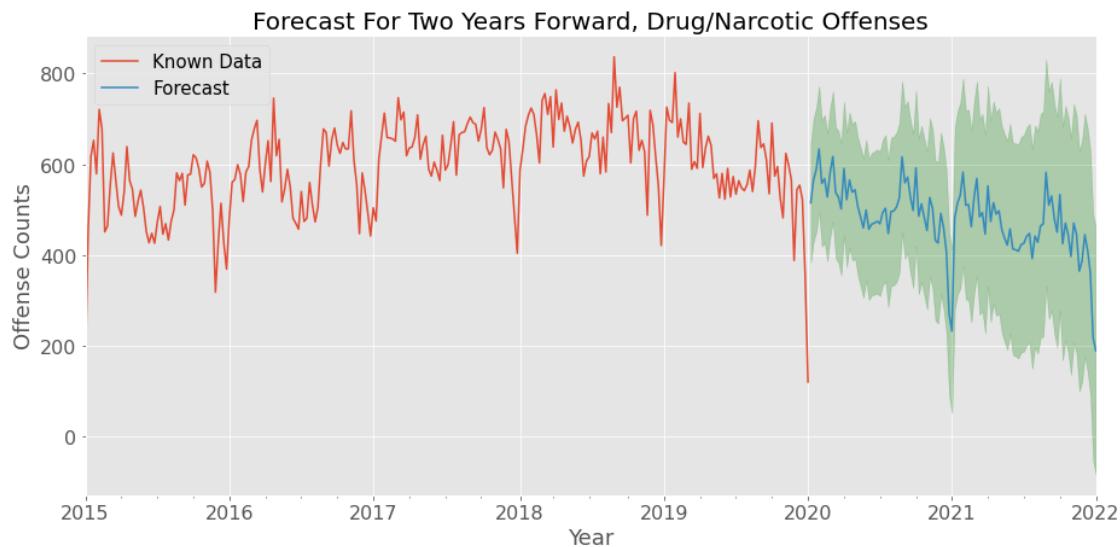
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Stolen Property Offenses

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""

```

SARIMAX Results

```
=====
Dep. Variable:          Stolen Property Offenses    No. Observations:      262
```

262

```

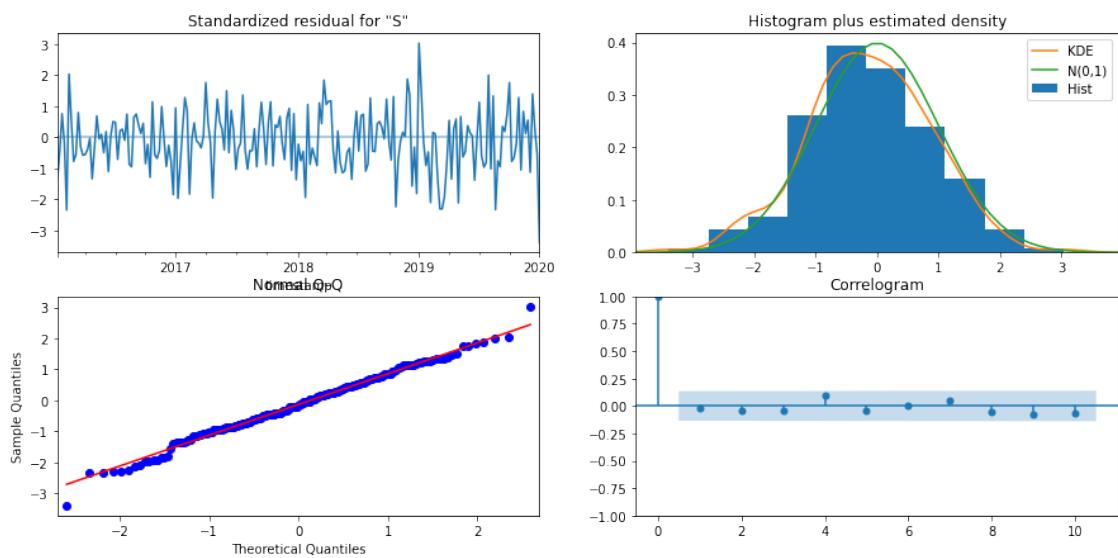
Model: SARIMAX(0, 1, 2)x(2, 1, [], 52) Log Likelihood           -787.419
Date:                               Fri, 16 Jul 2021      AIC                  1584.838
Time:                                12:24:05          BIC                  1601.550
Sample:                             01-04-2015        HQIC                 1591.595
                                      - 01-05-2020
Covariance Type:                    opg
=====

              coef    std err      z   P>|z|    [0.025    0.975]
-----
ma.L1     -0.7840    0.065  -11.977    0.000   -0.912   -0.656
ma.L2     -0.0757    0.068  -1.116    0.264   -0.209    0.057
ar.S.L52   -0.7599    0.069  -11.050    0.000   -0.895   -0.625
ar.S.L104  -0.4050    0.080  -5.093    0.000   -0.561   -0.249
sigma2    91.4605   9.984   9.161    0.000   71.892  111.029
=====

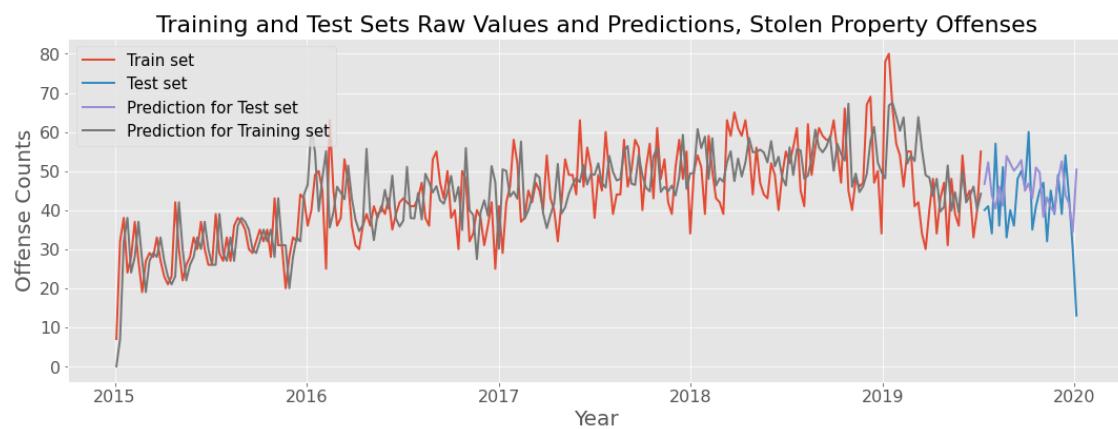
Ljung-Box (L1) (Q):                0.05  Jarque-Bera (JB):            0.86
Prob(Q):                           0.82  Prob(JB):                   0.65
Heteroskedasticity (H):             1.93  Skew:                      -0.10
Prob(H) (two-sided):                0.01  Kurtosis:                  3.25
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

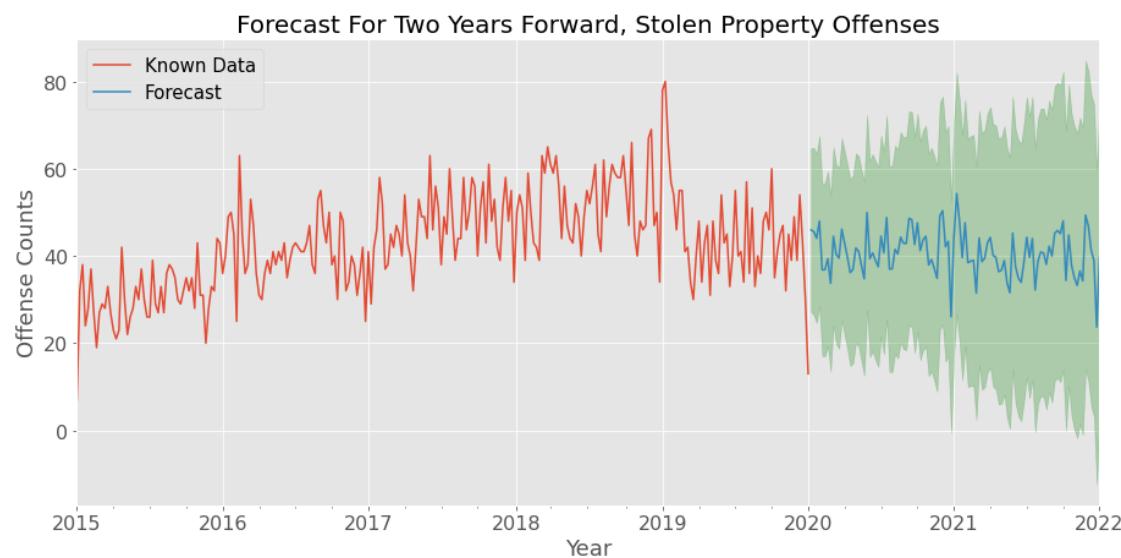
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Kidnapping/Abduction

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""

```

SARIMAX Results

```
=====
Dep. Variable: Kidnapping/Abduction No. Observations: 262
=====
```

```

Model: SARIMAX(1, 1, 1)x(2, 1, []) , 52 Log Likelihood      -726.762
Date: Fri, 16 Jul 2021 AIC                         1463.525
Time: 12:24:06 BIC                         1480.236
Sample: 01-04-2015 HQIC                      1470.281
        - 01-05-2020
Covariance Type: opg
=====

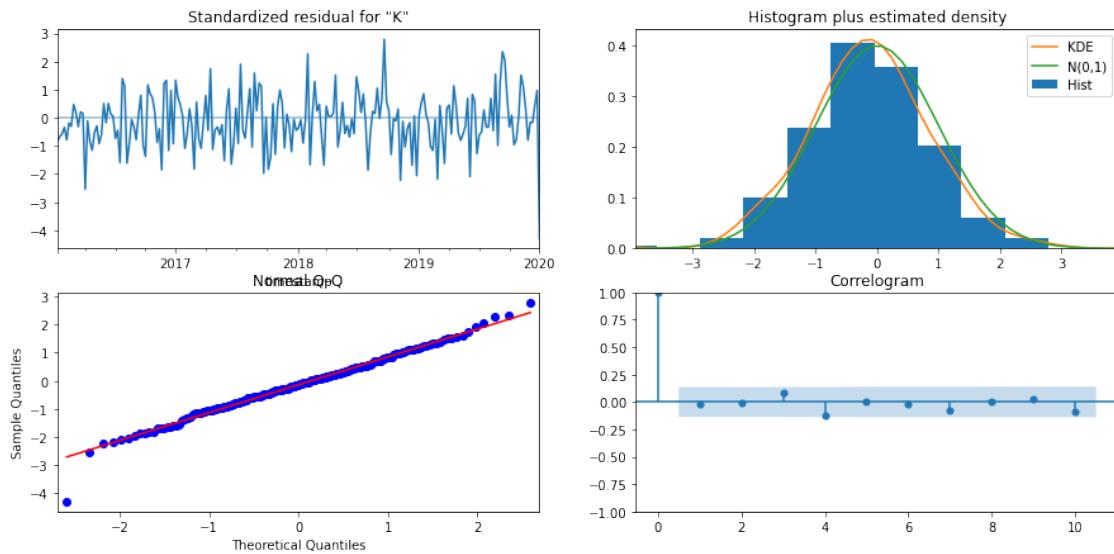
            coef    std err       z   P>|z|    [0.025    0.975]
-----
ar.L1     -0.1379    0.082  -1.687    0.092   -0.298    0.022
ma.L1     -0.9205    0.035 -26.484    0.000   -0.989   -0.852
ar.S.L52   -0.6328    0.076  -8.311    0.000   -0.782   -0.484
ar.S.L104  -0.1894    0.112  -1.688    0.091   -0.409    0.030
sigma2    54.9004    4.521  12.144    0.000   46.040  63.761
=====

Ljung-Box (L1) (Q):          0.08 Jarque-Bera (JB):      10.95
Prob(Q):                   0.78 Prob(JB):           0.00
Heteroskedasticity (H):     1.58 Skew:                 -0.19
Prob(H) (two-sided):        0.06 Kurtosis:             4.05
=====
```

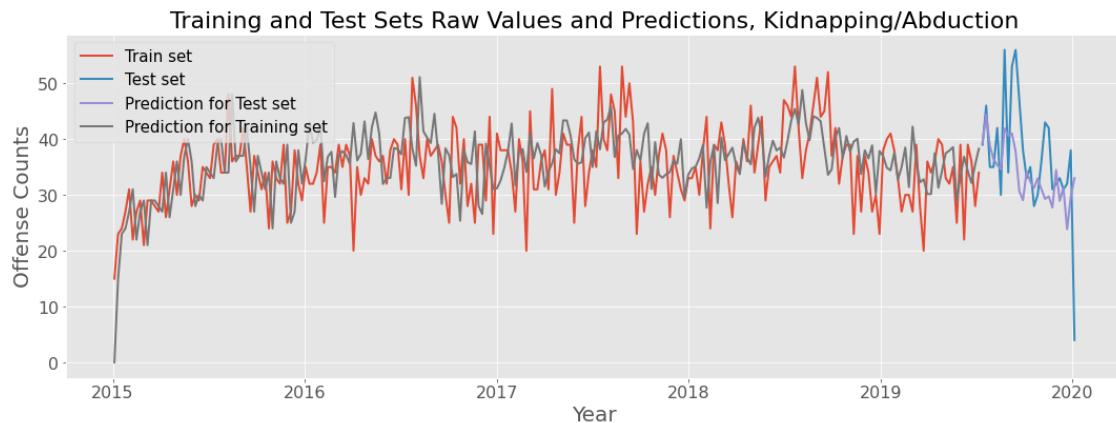
Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

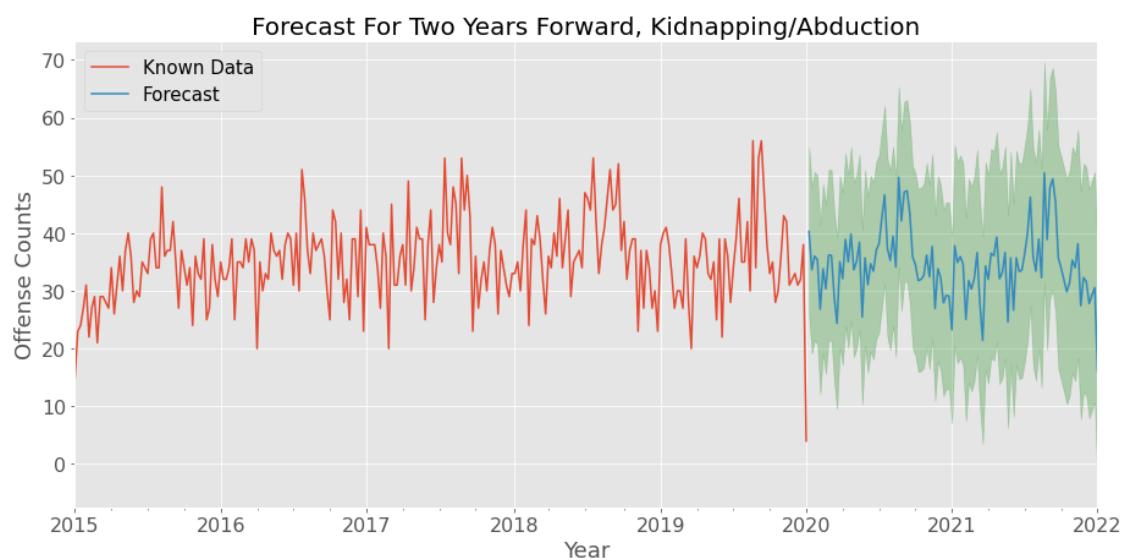
"""



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Robbery

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""

```

SARIMAX Results

Dep. Variable:

Robbery No. Observations:

262

```

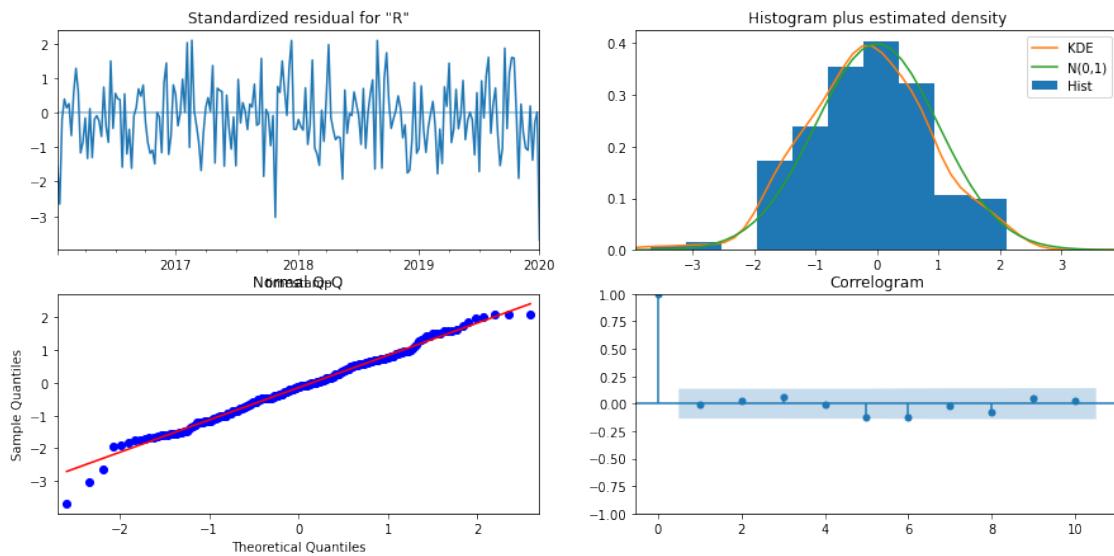
Model: SARIMAX(1, 1, 1)x(1, 1, 0, 52) Log Likelihood      -835.654
Date:             Fri, 16 Jul 2021    AIC                  1679.309
Time:                      12:24:07    BIC                  1692.678
Sample:           01-04-2015    HQIC                 1684.714
                   - 01-05-2020
Covariance Type: opg
=====

            coef   std err      z   P>|z|   [0.025   0.975]
-----
ar.L1       0.1493    0.086   1.738    0.082   -0.019    0.318
ma.L1      -0.9293    0.036 -25.568    0.000   -1.000   -0.858
ar.S.L52    -0.4793    0.065  -7.334    0.000   -0.607   -0.351
sigma2     161.6784   14.925 10.833    0.000  132.425 190.931
=====

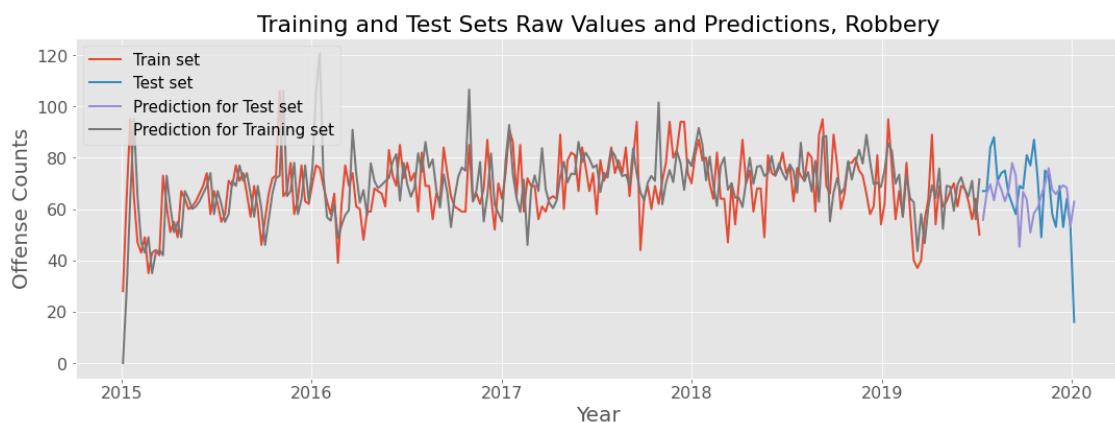
Ljung-Box (L1) (Q):          0.02   Jarque-Bera (JB):        1.50
Prob(Q):                    0.89   Prob(JB):              0.47
Heteroskedasticity (H):     1.23   Skew:                  -0.13
Prob(H) (two-sided):        0.40   Kurtosis:              3.32
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

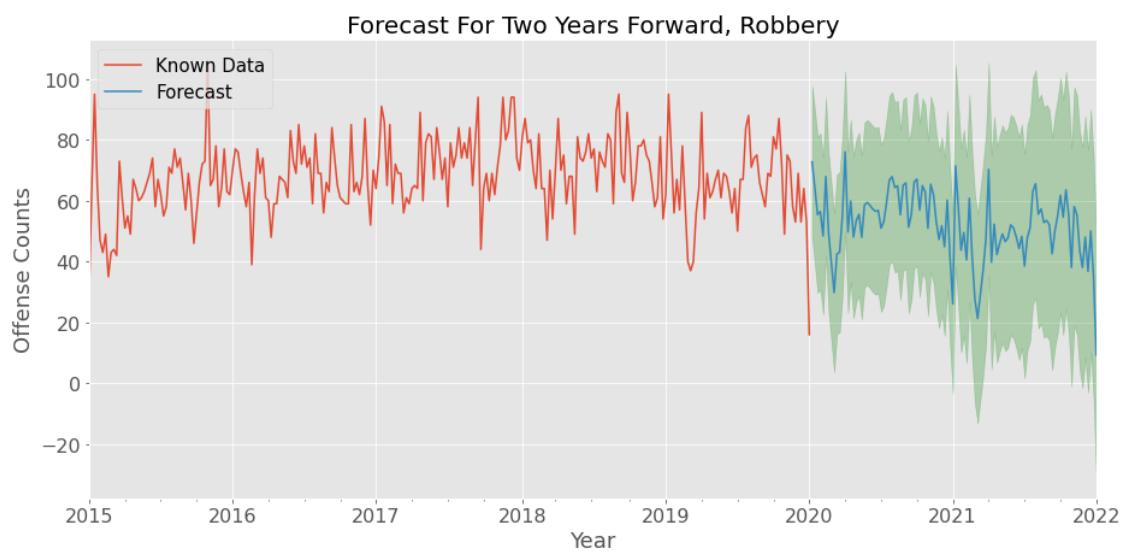
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Extortion/Blackmail

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""

```

SARIMAX Results

Dep. Variable:

Extortion/Blackmail

No. Observations:

261

```

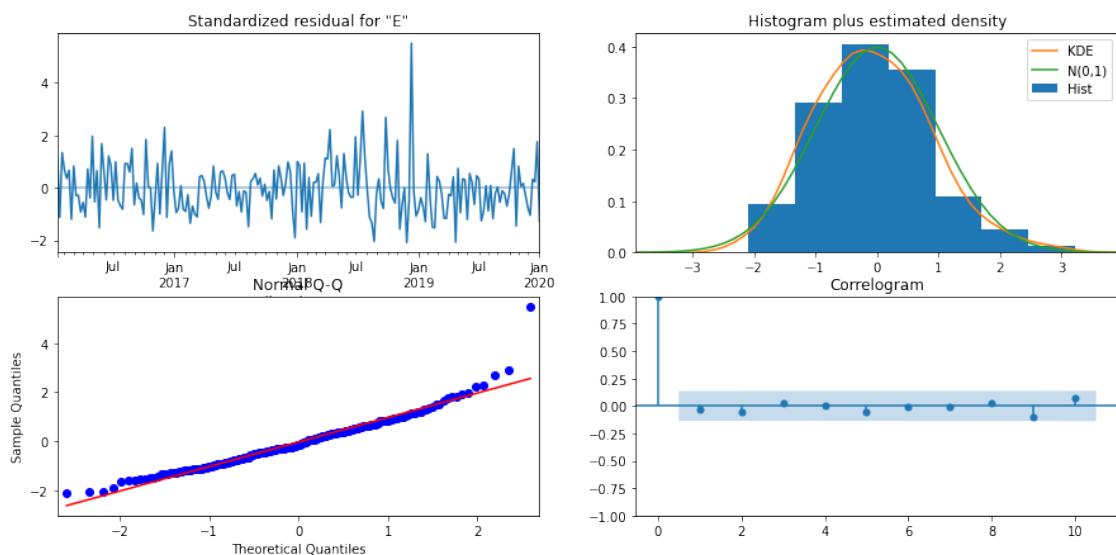
Model: SARIMAX(1, 1, 2)x(2, 1, []) , 52 Log Likelihood      -586.725
Date: Fri, 16 Jul 2021 AIC                         1185.449
Time: 12:24:07 BIC                         1205.474
Sample: 01-11-2015 HQIC                      1193.546
        - 01-05-2020
Covariance Type: opg
=====

            coef    std err       z   P>|z|    [0.025    0.975]
-----
ar.L1     -0.6409    0.214   -2.992    0.003   -1.061   -0.221
ma.L1     -0.1107    0.174   -0.635    0.526   -0.452    0.231
ma.L2     -0.7017    0.156   -4.488    0.000   -1.008   -0.395
ar.S.L52   -0.6926    0.095   -7.254    0.000   -0.880   -0.505
ar.S.L104  -0.3277    0.078   -4.223    0.000   -0.480   -0.176
sigma2    14.2969    0.938   15.246    0.000   12.459   16.135
=====

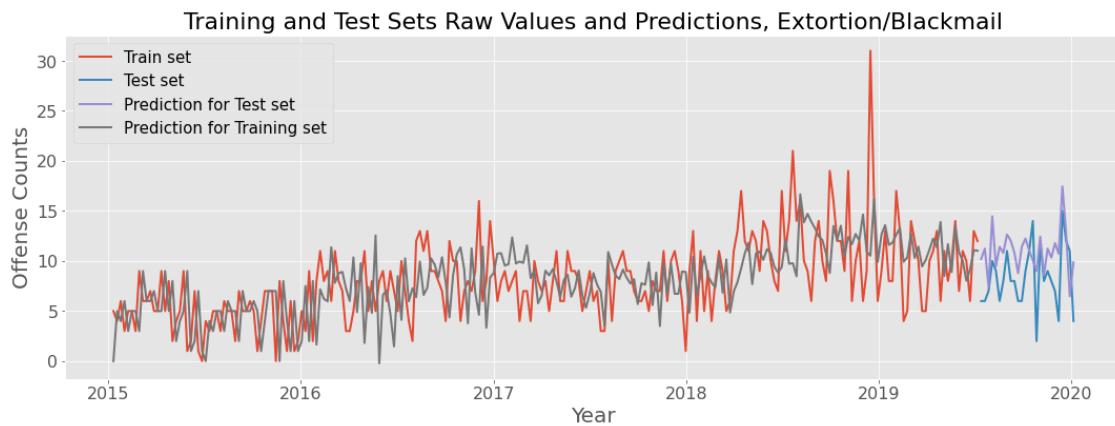
Ljung-Box (L1) (Q):          0.16   Jarque-Bera (JB):      159.88
Prob(Q):                   0.69   Prob(JB):                  0.00
Heteroskedasticity (H):     1.52   Skew:                     1.04
Prob(H) (two-sided):        0.08   Kurtosis:                 6.76
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

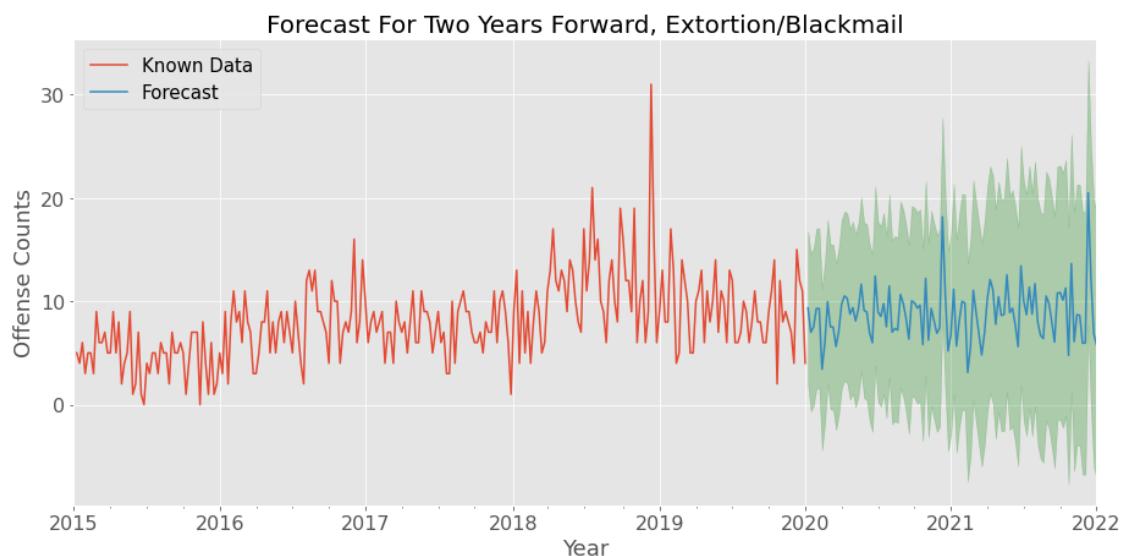
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Pornography/Obscene Material

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
SARIMAX Results
```

```

=====
Dep. Variable: Pornography/Obscene Material No. Observations: 262
Model: SARIMAX(2, 1, 0)x(1, 1, 0, 52) Log Likelihood -670.395
Date: Fri, 16 Jul 2021 AIC 1348.789
Time: 12:24:08 BIC 1362.159
Sample: 01-04-2015 HQIC 1354.195
- 01-05-2020
Covariance Type: opg
=====

```

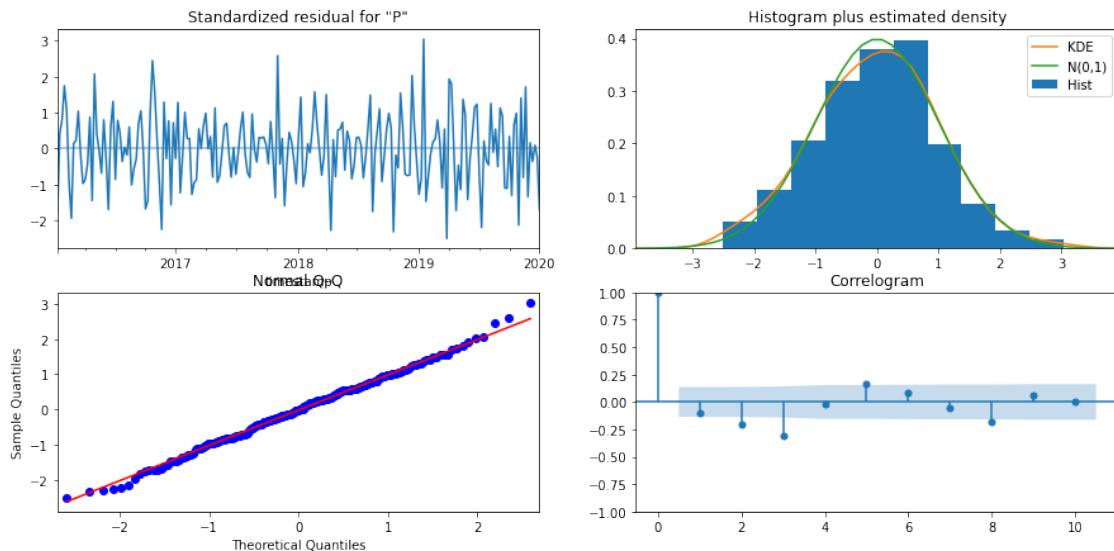
| | coef | std err | z | P> z | [0.025 | 0.975] |
|----------|---------|---------|--------|-------|--------|--------|
| ar.L1 | -0.5977 | 0.065 | -9.236 | 0.000 | -0.724 | -0.471 |
| ar.L2 | -0.3678 | 0.072 | -5.127 | 0.000 | -0.508 | -0.227 |
| ar.S.L52 | -0.4428 | 0.077 | -5.737 | 0.000 | -0.594 | -0.292 |
| sigma2 | 33.8076 | 3.454 | 9.789 | 0.000 | 27.039 | 40.577 |

```

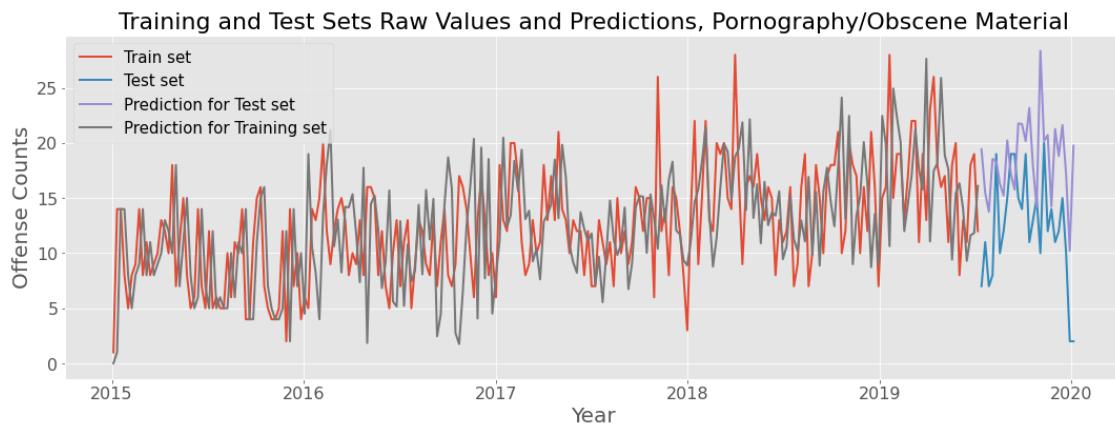
Ljung-Box (L1) (Q): 1.92 Jarque-Bera (JB): 0.01
Prob(Q): 0.17 Prob(JB): 0.99
Heteroskedasticity (H): 1.26 Skew: 0.02
Prob(H) (two-sided): 0.34 Kurtosis: 2.99
=====
```

Warnings:

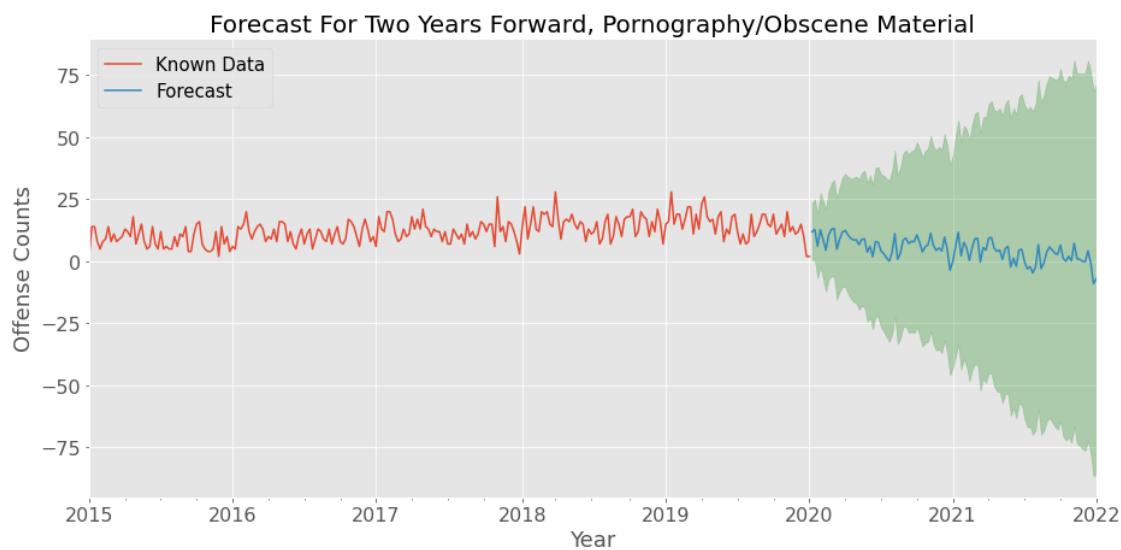
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
 """



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Prostitution Offenses

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
SARIMAX Results
```

```

=====
Dep. Variable: Prostitution Offenses No. Observations: 261
Model: SARIMAX(2, 1, 2)x(2, 1, [], 52) Log Likelihood -768.793
Date: Fri, 16 Jul 2021 AIC 1551.586
Time: 12:24:09 BIC 1574.949
Sample: 01-04-2015 HQIC 1561.033
- 12-29-2019
Covariance Type: opg
=====

```

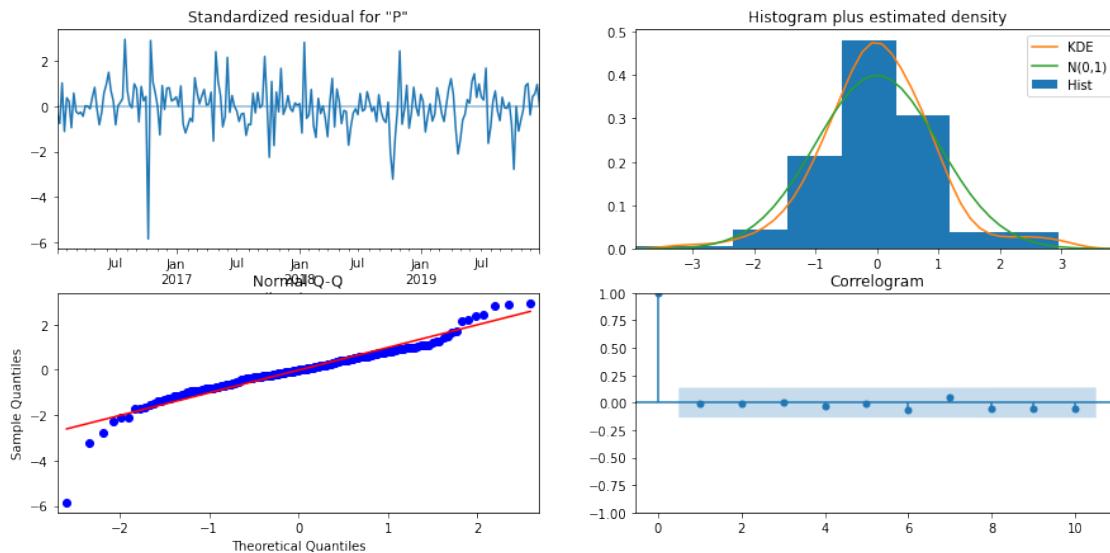
| | coef | std err | z | P> z | [0.025 | 0.975] |
|-----------|---------|---------|---------|-------|--------|--------|
| ar.L1 | 0.0872 | 0.378 | 0.230 | 0.818 | -0.654 | 0.829 |
| ar.L2 | -0.1269 | 0.122 | -1.041 | 0.298 | -0.366 | 0.112 |
| ma.L1 | -1.2842 | 0.375 | -3.426 | 0.001 | -2.019 | -0.550 |
| ma.L2 | 0.3174 | 0.367 | 0.864 | 0.388 | -0.403 | 1.038 |
| ar.S.L52 | -0.5900 | 0.050 | -11.911 | 0.000 | -0.687 | -0.493 |
| ar.S.L104 | -0.4231 | 0.048 | -8.855 | 0.000 | -0.517 | -0.329 |
| sigma2 | 80.8415 | 5.606 | 14.420 | 0.000 | 69.854 | 91.829 |

```

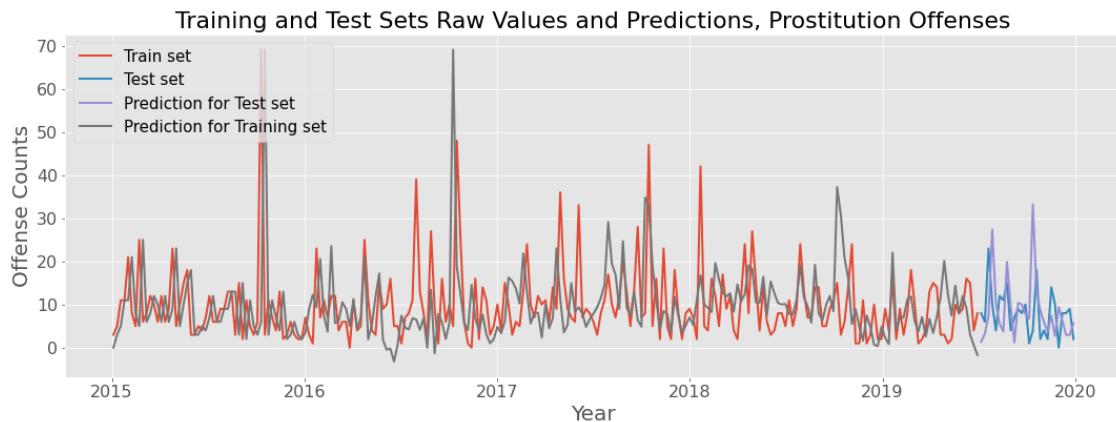
Ljung-Box (L1) (Q): 0.00 Jarque-Bera (JB): 319.98
Prob(Q): 0.96 Prob(JB): 0.00
Heteroskedasticity (H): 0.76 Skew: -0.80
Prob(H) (two-sided): 0.27 Kurtosis: 8.86
=====
```

Warnings:

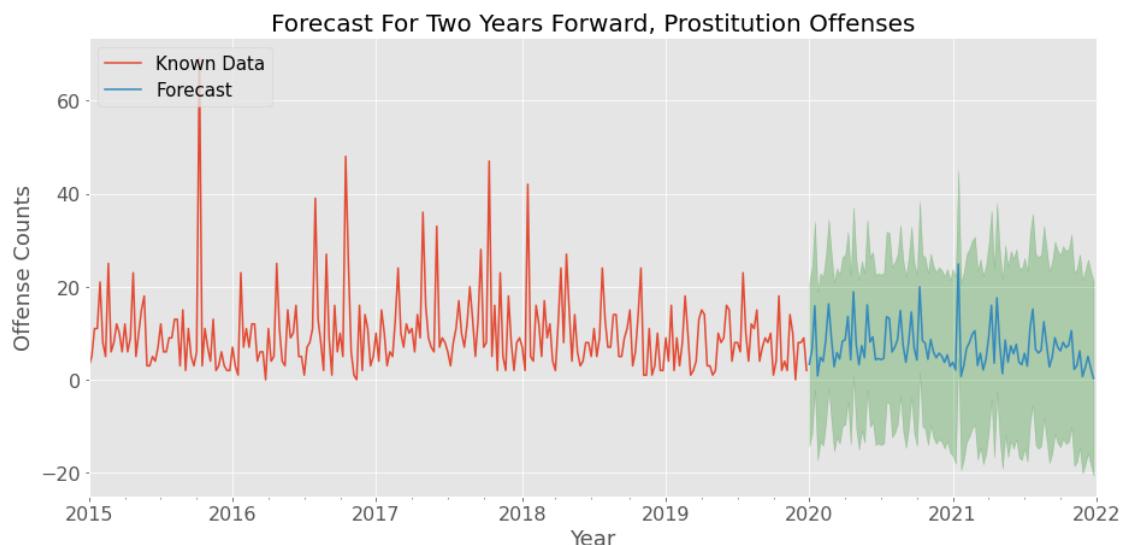
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
 """



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Bribery

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
```

====

SARIMAX Results

```
=====
Dep. Variable:                   Bribery    No. Observations:             262
Model: SARIMAX(1, 1, 0)x(2, 1, 0, 52) Log Likelihood:        -498.722
Date:   Fri, 16 Jul 2021          AIC:                  1005.443
Time:   12:24:10                BIC:                  1018.813
Sample: 01-04-2015 - 01-05-2020 HQIC:                  1010.849
Covariance Type: opg
=====
```

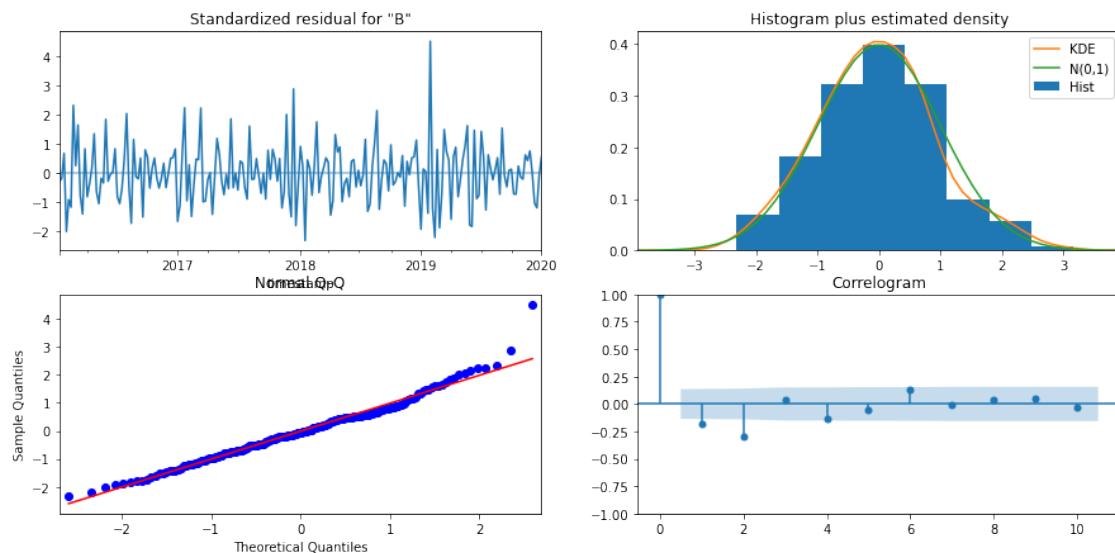
| | coef | std err | z | P> z | [0.025 | 0.975] |
|-----------|---------|---------|--------|-------|--------|--------|
| <hr/> | | | | | | |
| ar.L1 | -0.5278 | 0.055 | -9.665 | 0.000 | -0.635 | -0.421 |
| ar.S.L52 | -0.6081 | 0.087 | -6.976 | 0.000 | -0.779 | -0.437 |
| ar.S.L104 | -0.3715 | 0.084 | -4.423 | 0.000 | -0.536 | -0.207 |
| sigma2 | 6.0779 | 0.461 | 13.197 | 0.000 | 5.175 | 6.981 |
| <hr/> | | | | | | |

```
Ljung-Box (L1) (Q):            7.09    Jarque-Bera (JB):       32.28
Prob(Q):                      0.01    Prob(JB):                 0.00
Heteroskedasticity (H):        1.14    Skew:                     0.58
Prob(H) (two-sided):           0.59    Kurtosis:                 4.53
=====
```

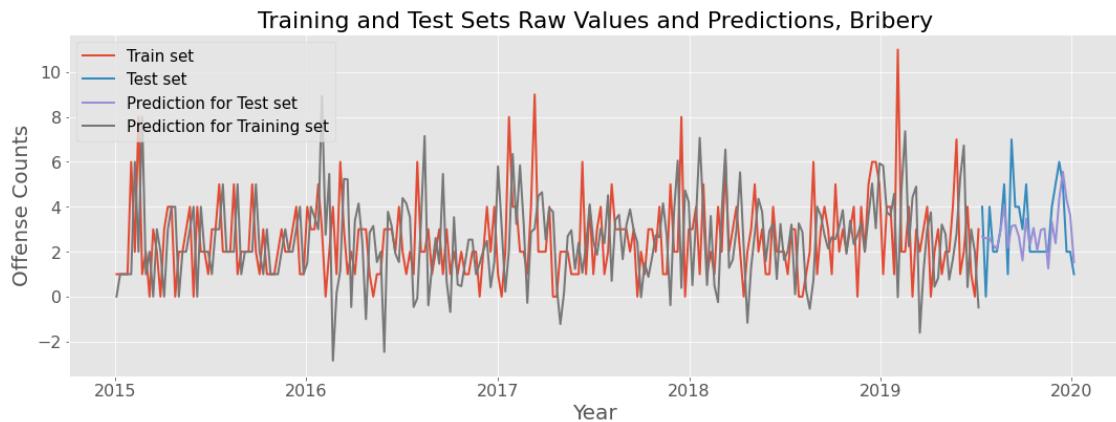
Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

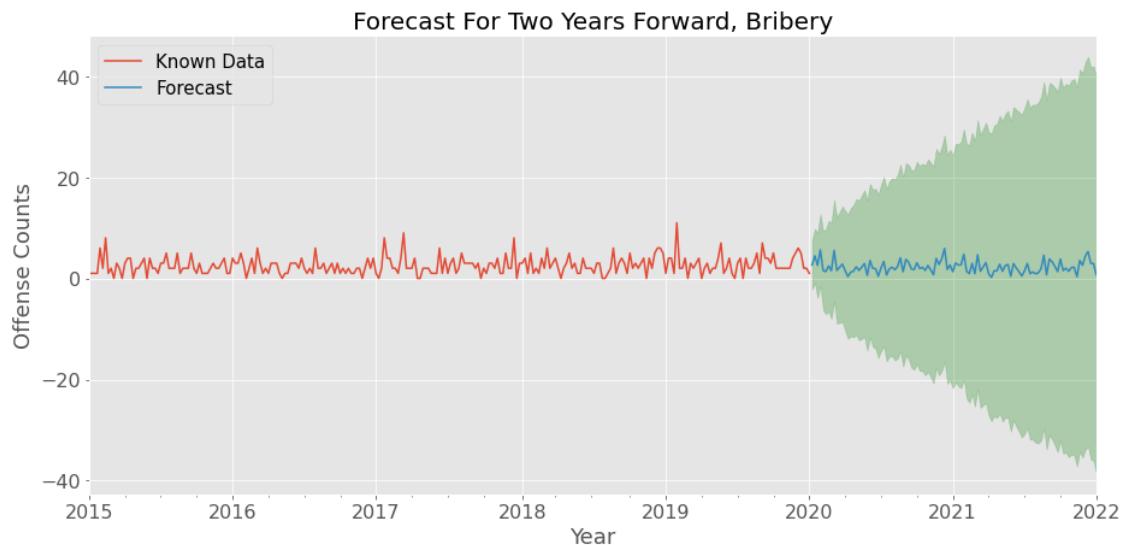
====



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Embezzlement

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
```

====

SARIMAX Results

```
=====
Dep. Variable: Embezzlement    No. Observations: 262
Model: SARIMAX(0, 1, 1)x(2, 1, [])  Log Likelihood -582.761
Date: Fri, 16 Jul 2021   AIC 1173.522
Time: 12:24:10   BIC 1186.892
Sample: 01-04-2015 - 01-05-2020   HQIC 1178.928
Covariance Type: opg
=====
```

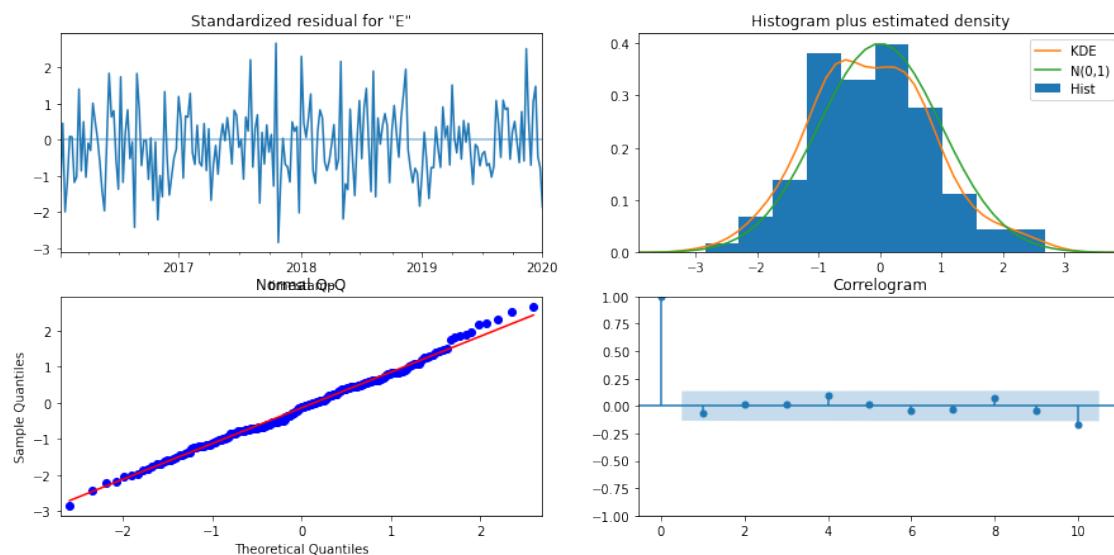
| | coef | std err | z | P> z | [0.025 | 0.975] |
|-----------|---------|---------|---------|-------|--------|--------|
| ma.L1 | -1.0684 | 0.036 | -29.657 | 0.000 | -1.139 | -0.998 |
| ar.S.L52 | -0.7387 | 0.077 | -9.546 | 0.000 | -0.890 | -0.587 |
| ar.S.L104 | -0.3269 | 0.088 | -3.707 | 0.000 | -0.500 | -0.154 |
| sigma2 | 11.5611 | 1.365 | 8.468 | 0.000 | 8.885 | 14.237 |

```
Ljung-Box (L1) (Q): 0.95 Jarque-Bera (JB): 1.28
Prob(Q): 0.33 Prob(JB): 0.53
Heteroskedasticity (H): 0.71 Skew: 0.19
Prob(H) (two-sided): 0.16 Kurtosis: 3.03
=====
```

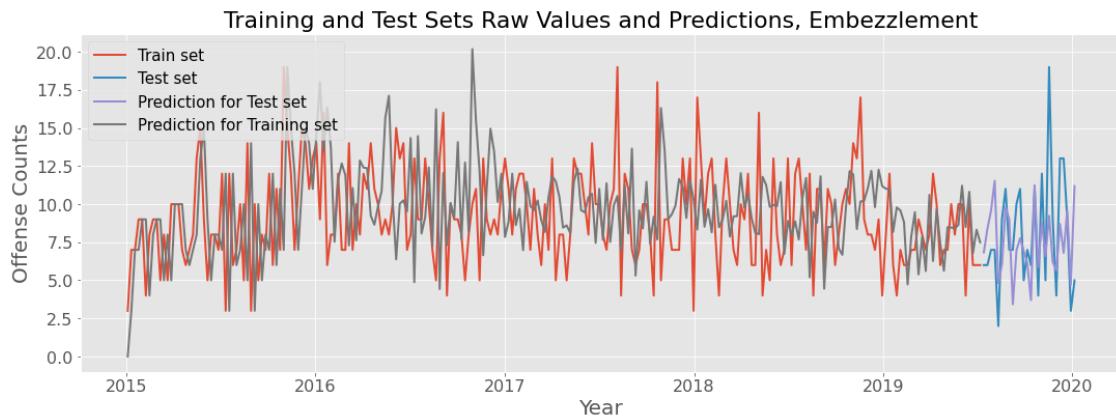
Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

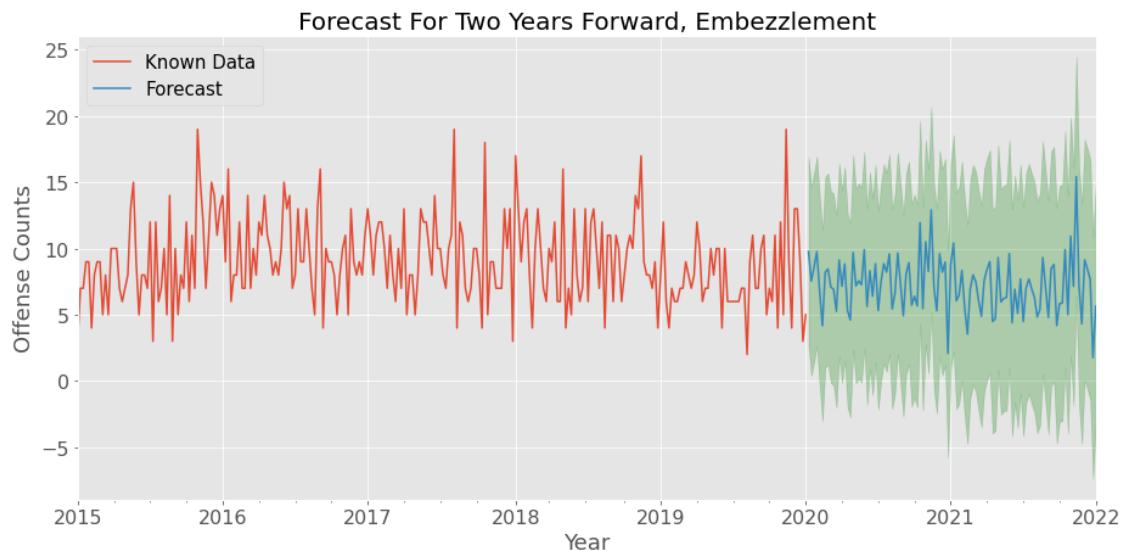
====



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Homicide Offenses

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
```

====

SARIMAX Results

```
=====
Dep. Variable: Homicide Offenses No. Observations: 262
Model: SARIMAX(0, 1, 1)x(2, 1, [], 52) Log Likelihood -488.621
Date: Fri, 16 Jul 2021 AIC 985.242
Time: 12:24:11 BIC 998.611
Sample: 01-04-2015 HQIC 990.647
          - 01-05-2020
Covariance Type: opg
=====
```

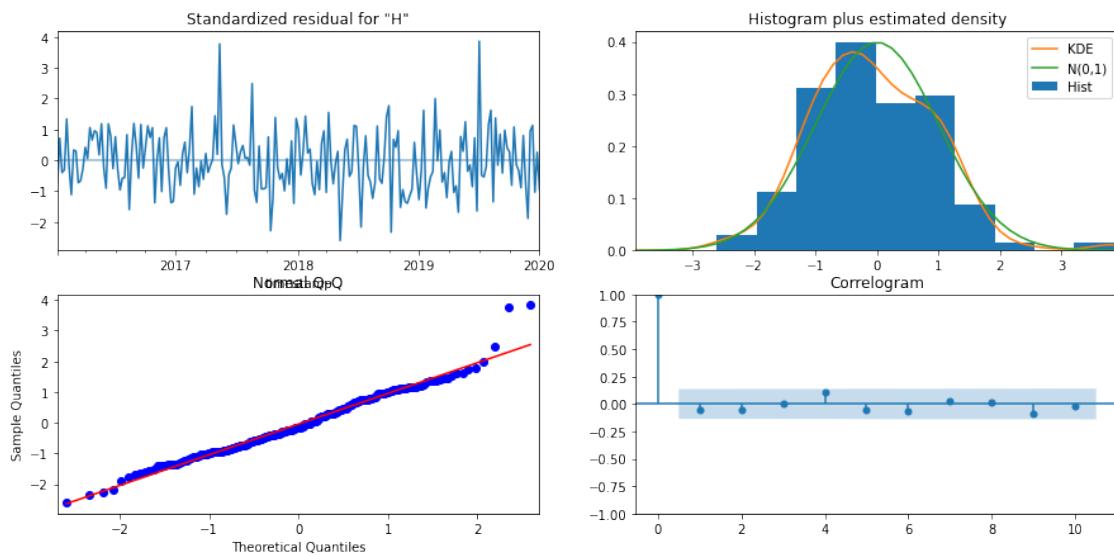
| | coef | std err | z | P> z | [0.025 | 0.975] |
|-----------|---------|---------|---------|-------|--------|--------|
| ma.L1 | -1.0235 | 0.020 | -49.946 | 0.000 | -1.064 | -0.983 |
| ar.S.L52 | -0.7096 | 0.063 | -11.268 | 0.000 | -0.833 | -0.586 |
| ar.S.L104 | -0.3780 | 0.092 | -4.130 | 0.000 | -0.557 | -0.199 |
| sigma2 | 5.0587 | 0.457 | 11.060 | 0.000 | 4.162 | 5.955 |

```
Ljung-Box (L1) (Q): 0.68 Jarque-Bera (JB): 17.33
Prob(Q): 0.41 Prob(JB): 0.00
Heteroskedasticity (H): 1.97 Skew: 0.47
Prob(H) (two-sided): 0.01 Kurtosis: 4.05
=====
```

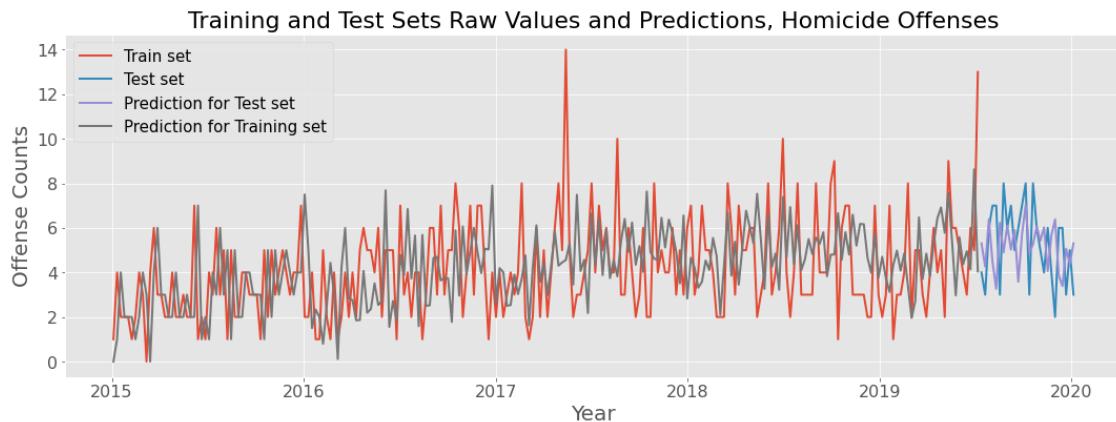
Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

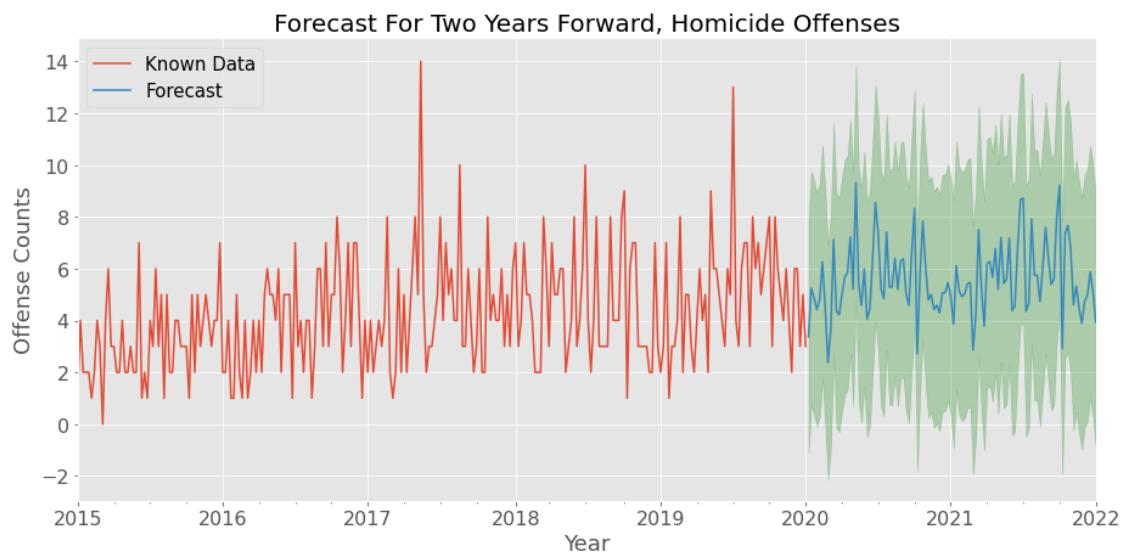
====



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Human Trafficking

THE FINAL MODEL SUMMARY:

```
<class 'statsmodels.iolib.summary.Summary'>
```

"""

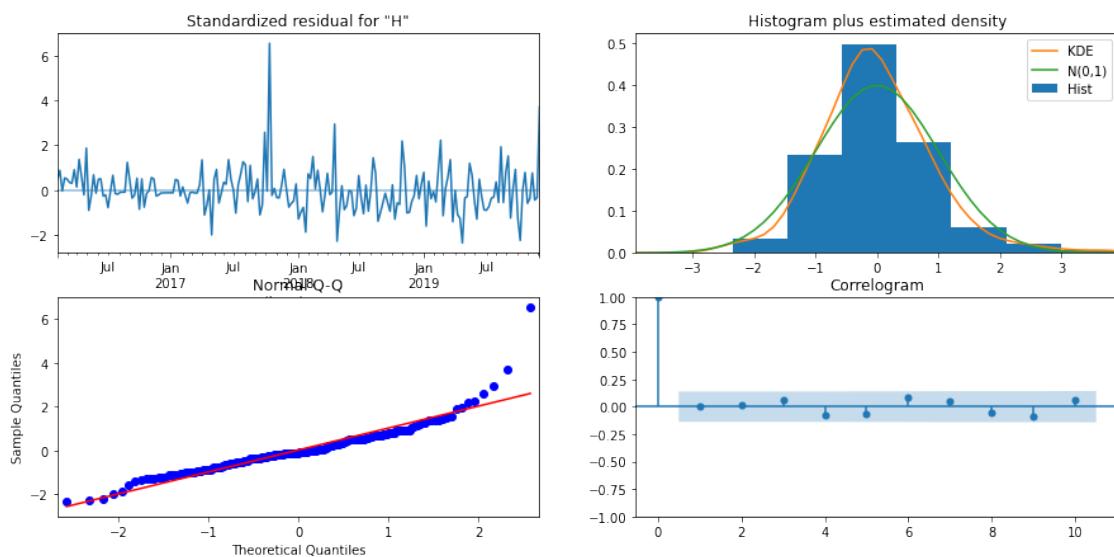
SARIMAX Results

| Dep. Variable: | Human Trafficking | No. Observations: | 254 | | | |
|-------------------------|--------------------------------|-------------------|----------|-------|--------|--------|
| Model: | SARIMAX(2, 1, 1)x(2, 1, 0, 52) | Log Likelihood | -329.007 | | | |
| Date: | Fri, 16 Jul 2021 | AIC | 670.015 | | | |
| Time: | 12:24:12 | BIC | 689.834 | | | |
| Sample: | 02-01-2015 - 12-08-2019 | HQIC | 678.035 | | | |
| Covariance Type: | opg | | | | | |
| | | | ===== | | | |
| | coef | std err | z | P> z | [0.025 | 0.975] |
| ar.L1 | 0.0605 | 0.089 | 0.680 | 0.496 | -0.114 | 0.235 |
| ar.L2 | 0.1227 | 0.089 | 1.378 | 0.168 | -0.052 | 0.297 |
| ma.L1 | -0.9674 | 0.030 | -32.027 | 0.000 | -1.027 | -0.908 |
| ar.S.L52 | -0.7577 | 0.050 | -15.241 | 0.000 | -0.855 | -0.660 |
| ar.S.L104 | -0.5915 | 0.048 | -12.239 | 0.000 | -0.686 | -0.497 |
| sigma2 | 1.1424 | 0.107 | 10.668 | 0.000 | 0.933 | 1.352 |
| Ljung-Box (L1) (Q): | 0.00 | Jarque-Bera (JB): | 778.67 | | | |
| Prob(Q): | 0.99 | Prob(JB): | 0.00 | | | |
| Heteroskedasticity (H): | 2.96 | Skew: | 1.72 | | | |
| Prob(H) (two-sided): | 0.00 | Kurtosis: | 12.01 | | | |
| | | | ===== | | | |

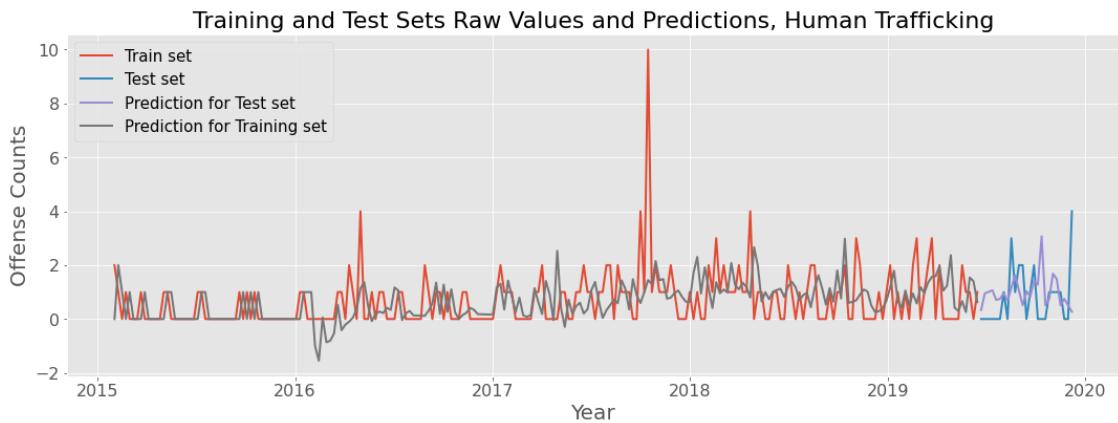
Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

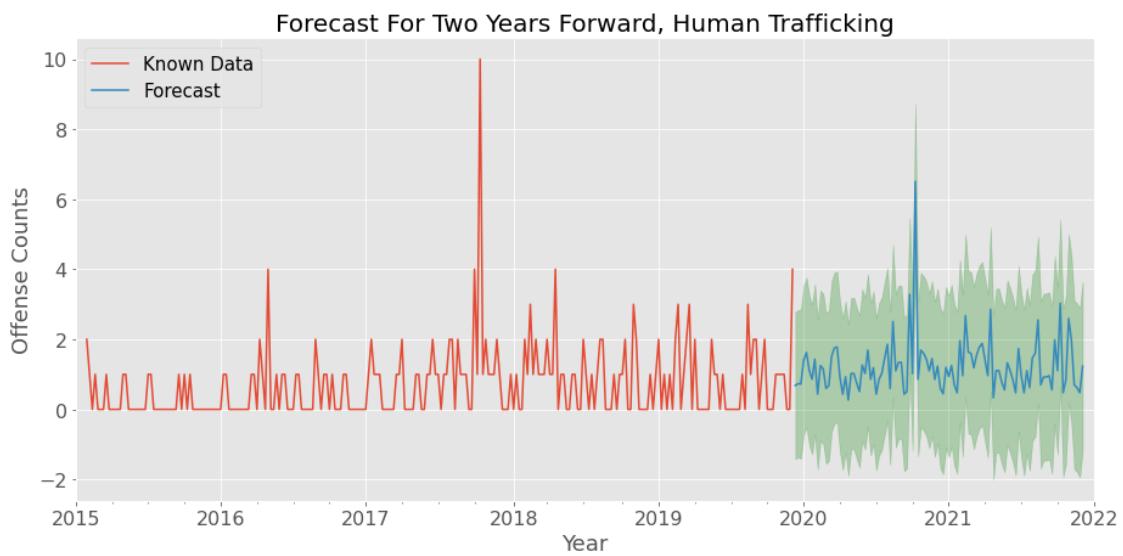
"""



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Gambling Offenses

THE FINAL MODEL SUMMARY:

```

<class 'statsmodels.iolib.summary.Summary'>
"""
=====
              SARIMAX Results
=====

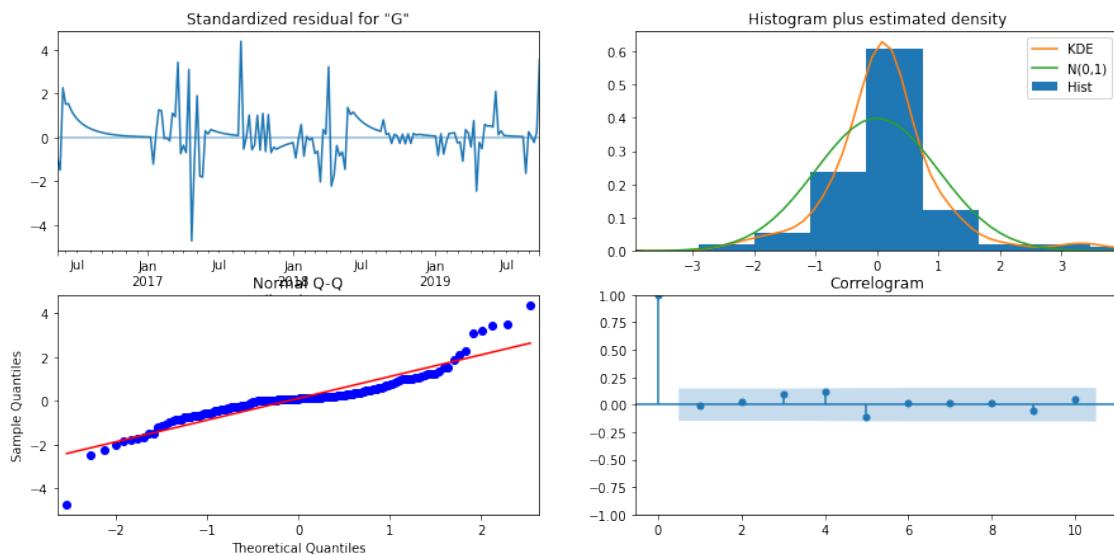
Dep. Variable:          Gambling Offenses    No. Observations:                  230
Model:                 SARIMAX(2, 1, 1)x(2, 1, 0, 52)    Log Likelihood:                -184.117
Date:                   Fri, 16 Jul 2021      AIC:                            380.234
Time:                     12:24:13        BIC:                            399.291
Sample:                 05-10-2015      HQIC:                           387.962
                           - 09-29-2019
Covariance Type:             opg
=====

            coef      std err           z      P>|z|      [0.025      0.975]
-----
ar.L1       0.0104      0.070     0.149      0.881     -0.126      0.147
ar.L2      -0.1026      0.072    -1.418      0.156     -0.244      0.039
ma.L1      -0.8722      0.037   -23.807      0.000     -0.944     -0.800
ar.S.L52    -0.7778      0.071   -10.981      0.000     -0.917     -0.639
ar.S.L104   -0.5218      0.048   -10.976      0.000     -0.615     -0.429
sigma2      0.3527      0.029    12.315      0.000      0.297      0.409
=====

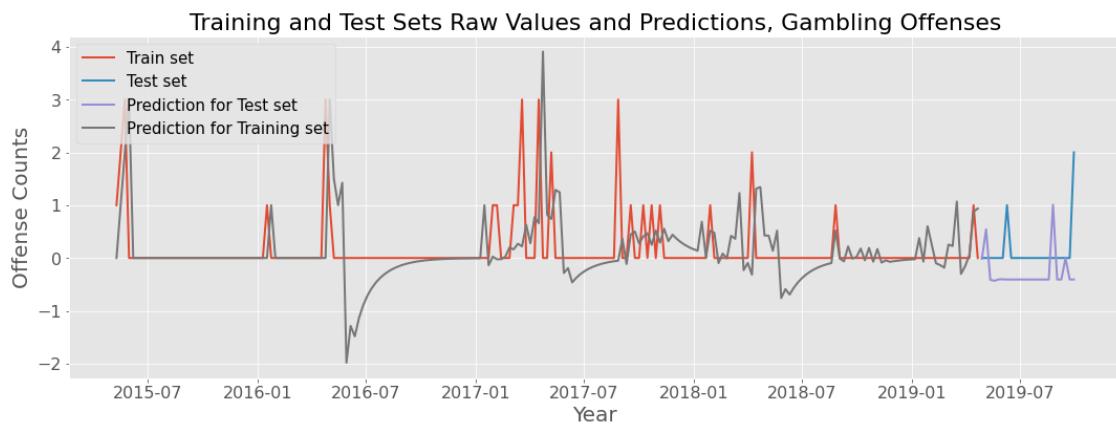
Ljung-Box (L1) (Q):      0.01    Jarque-Bera (JB):                  266.41
Prob(Q):                  0.93    Prob(JB):                      0.00
Heteroskedasticity (H):      0.39    Skew:                          0.28
Prob(H) (two-sided):      0.00    Kurtosis:                     8.98
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

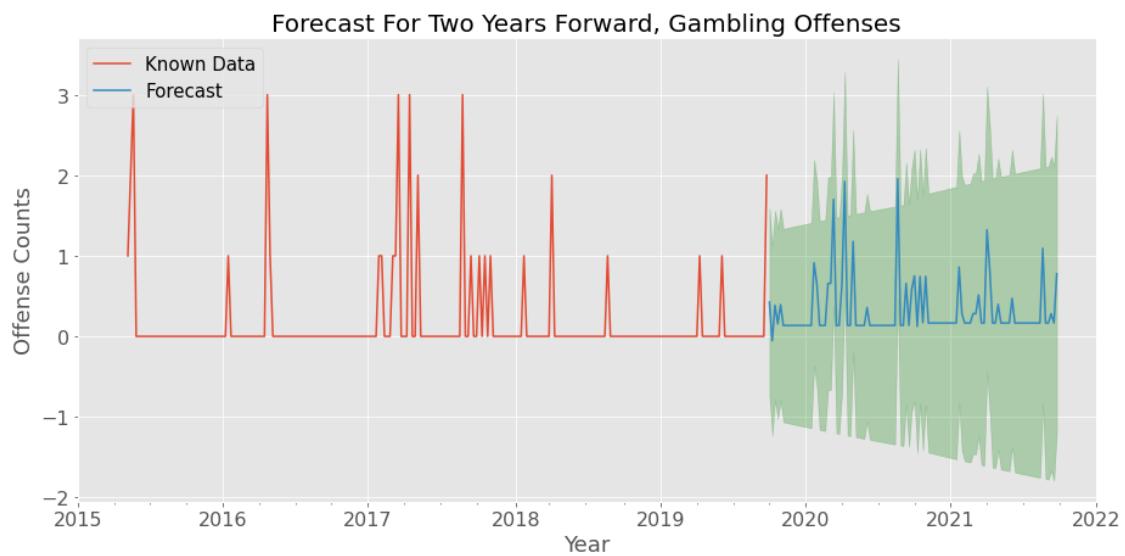
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Animal Cruelty

THE FINAL MODEL SUMMARY:

```

<class 'statsmodels.iolib.summary.Summary'>
"""
=====
SARIMAX Results
=====
Dep. Variable: Animal Cruelty No. Observations: 196
Model: SARIMAX(0, 1, 1)x(2, 1, [], 52) Log Likelihood -450.562
Date: Fri, 16 Jul 2021 AIC 909.124
Time: 12:24:14 BIC 920.976
Sample: 04-10-2016 HQIC 913.940
- 01-05-2020
Covariance Type: opg
=====

```

| | coef | std err | z | P> z | [0.025 | 0.975] |
|-----------|---------|---------|---------|-------|--------|--------|
| ma.L1 | -0.6934 | 0.063 | -10.953 | 0.000 | -0.817 | -0.569 |
| ar.S.L52 | -0.8118 | 0.130 | -6.230 | 0.000 | -1.067 | -0.556 |
| ar.S.L104 | -0.3824 | 0.147 | -2.597 | 0.009 | -0.671 | -0.094 |
| sigma2 | 24.2980 | 3.551 | 6.843 | 0.000 | 17.339 | 31.257 |

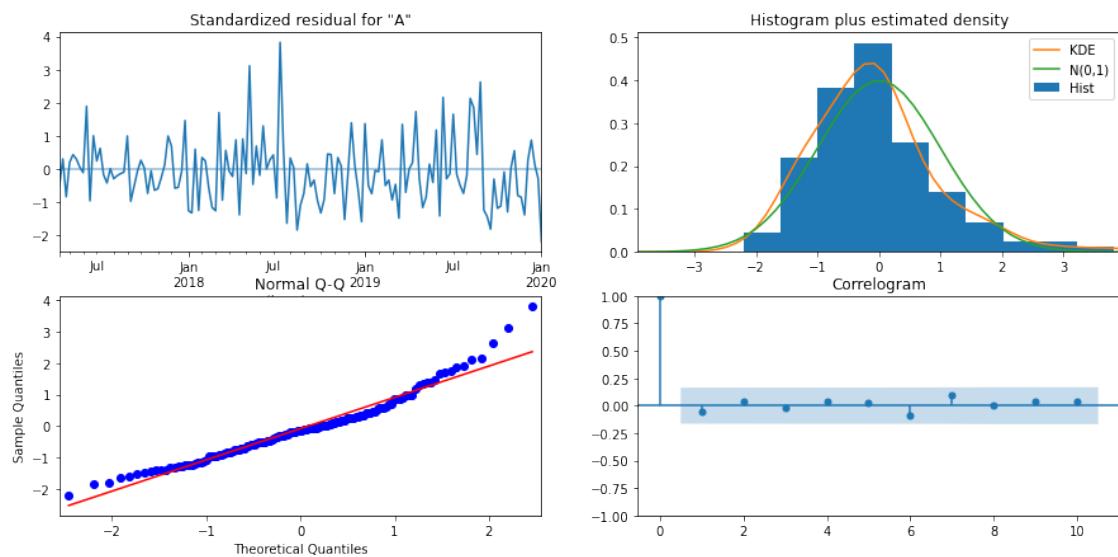
```

Ljung-Box (L1) (Q): 0.45 Jarque-Bera (JB): 32.54
Prob(Q): 0.50 Prob(JB): 0.00
Heteroskedasticity (H): 2.19 Skew: 0.89
Prob(H) (two-sided): 0.01 Kurtosis: 4.52
=====
```

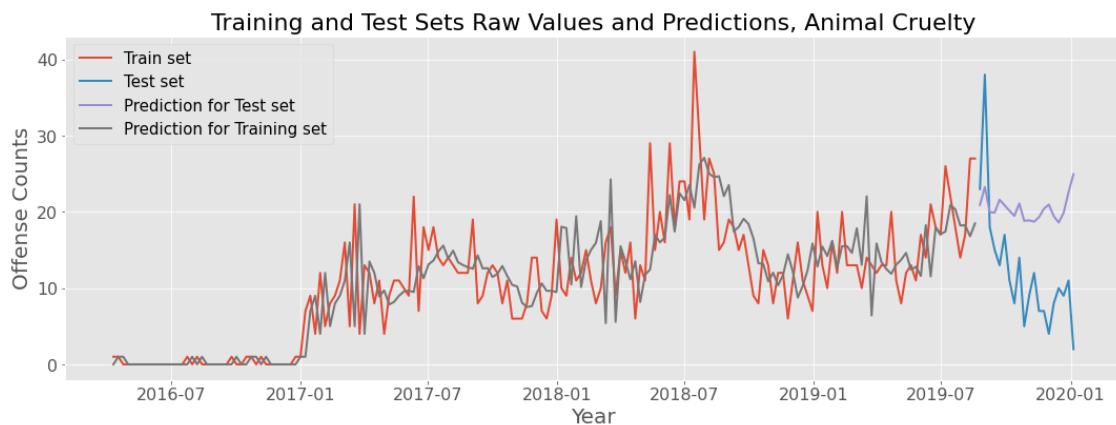
Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

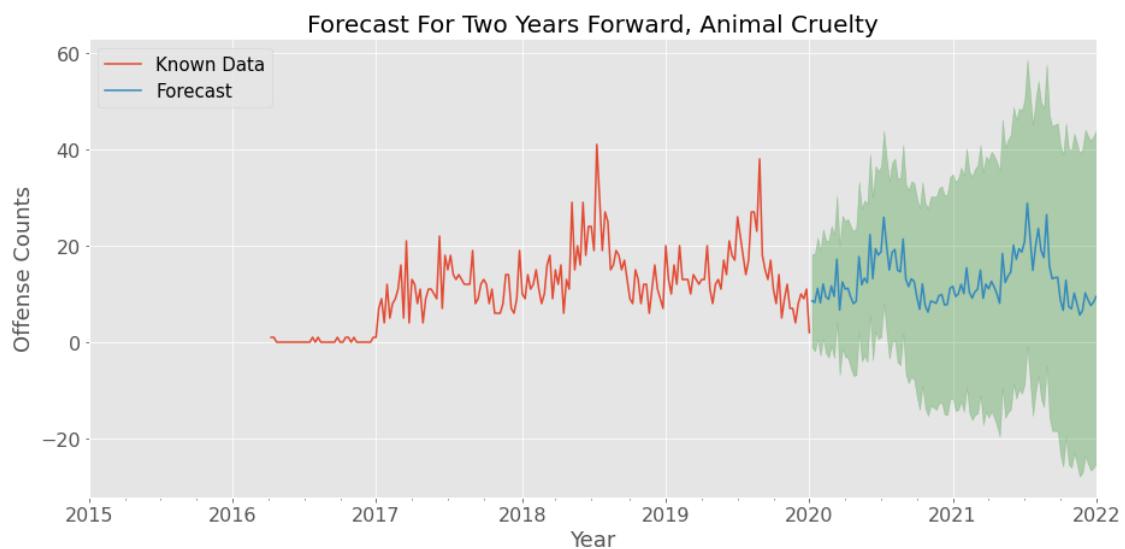
"""



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Sex Offenses

THE FINAL MODEL SUMMARY:

```

<class 'statsmodels.iolib.summary.Summary'>
"""
=====
              SARIMAX Results
=====

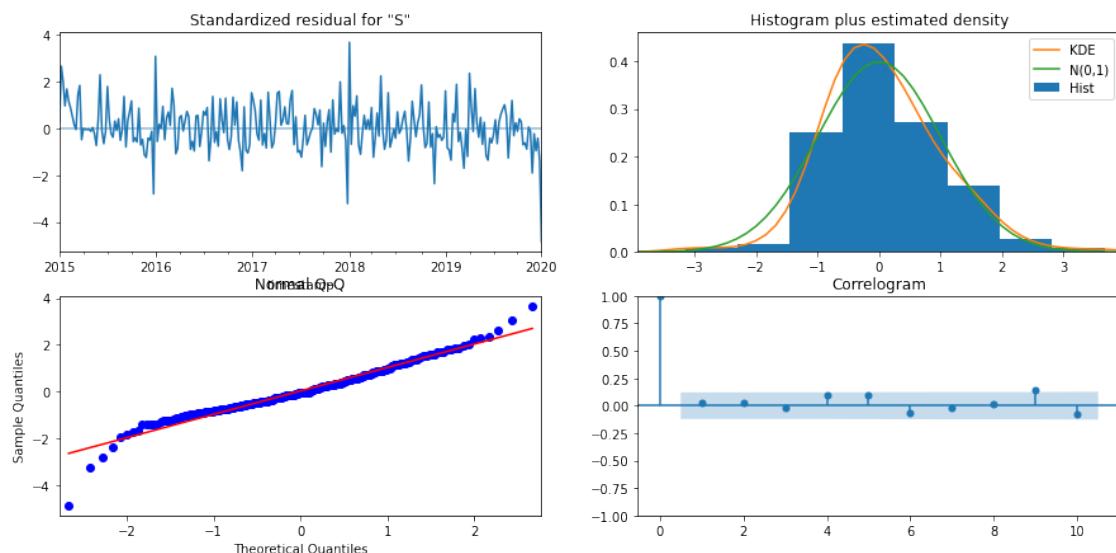
Dep. Variable:                  Sex Offenses    No. Observations:                   262
Model: SARIMAX(2, 0, 1)x(1, 0, [], 52)    Log Likelihood:                -1150.300
Date: Fri, 16 Jul 2021            AIC:                            2310.599
Time: 12:24:14                    BIC:                            2328.441
Sample: 01-04-2015 - 01-05-2020   HQIC:                           2317.770
Covariance Type:                  opg
=====

            coef      std err       z     P>|z|      [0.025      0.975]
-----
ar.L1      0.8545      0.083    10.310      0.000      0.692      1.017
ar.L2      0.1441      0.082     1.752      0.080     -0.017      0.305
ma.L1     -0.8308      0.052   -16.010      0.000     -0.932     -0.729
ar.S.L52    0.3267      0.049     6.613      0.000      0.230      0.424
sigma2    366.4911     25.233    14.524      0.000    317.035    415.947
=====

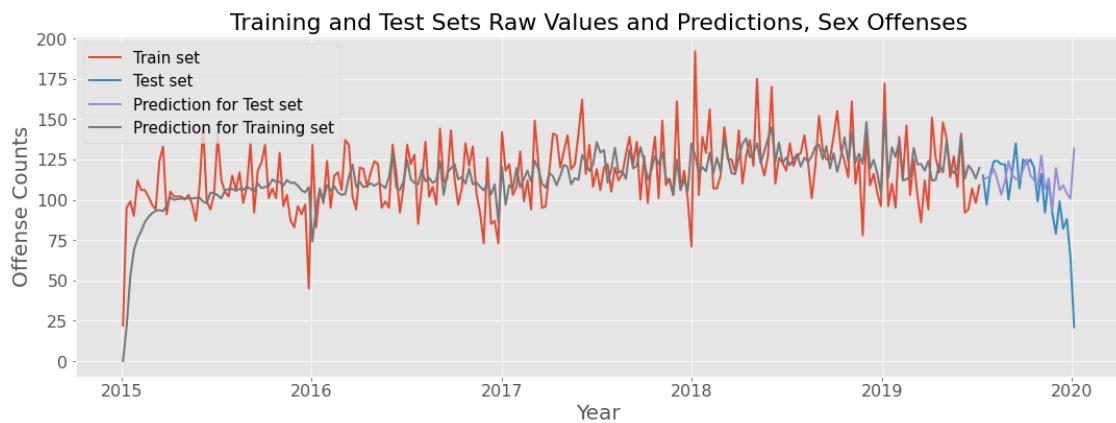
Ljung-Box (L1) (Q):             0.15    Jarque-Bera (JB):                 72.75
Prob(Q):                      0.70    Prob(JB):                     0.00
Heteroskedasticity (H):        1.14    Skew:                         -0.10
Prob(H) (two-sided):           0.55    Kurtosis:                     5.57
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

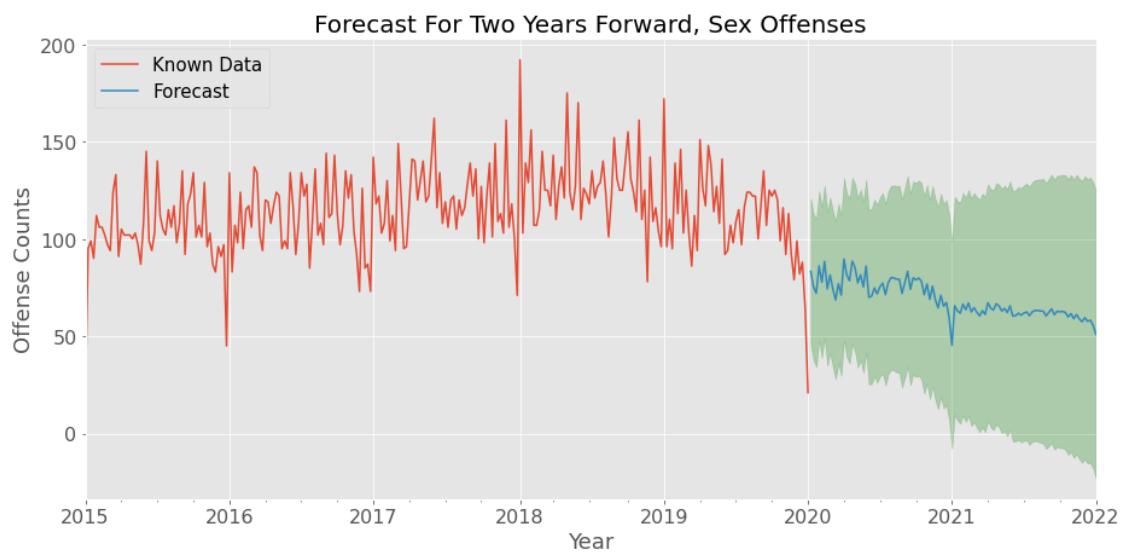
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



OFFENSE CATEGORY: Weapon Law Violations

THE FINAL MODEL SUMMARY:

```

<class 'statsmodels.iolib.summary.Summary'>
"""
=====
          SARIMAX Results
=====

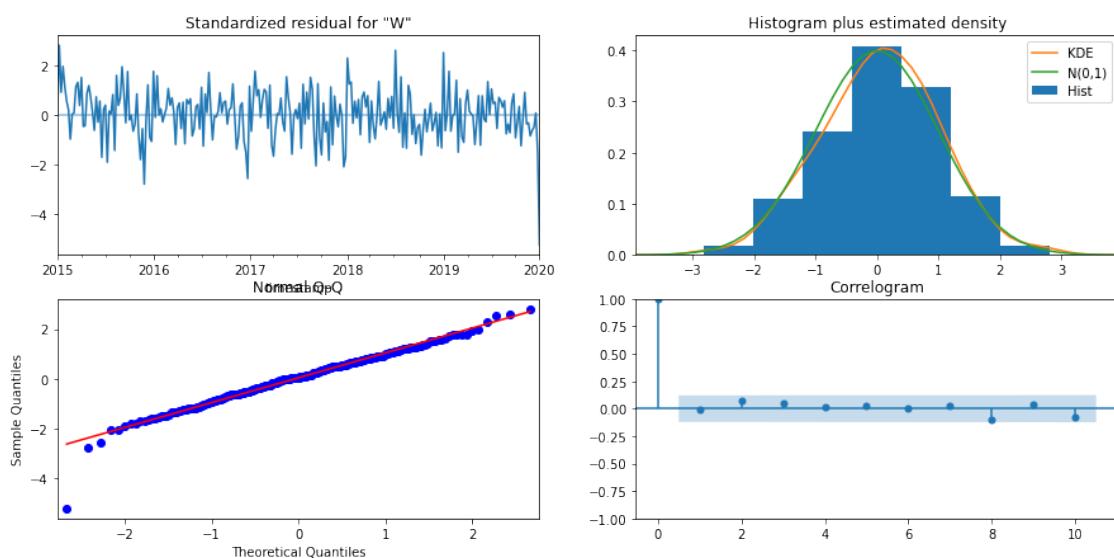
Dep. Variable:           Weapon Law Violations   No. Observations:                  262
Model:                 SARIMAX(1, 0, 1)x(0, 0, 1, 52)   Log Likelihood:                -1066.105
Date:                 Fri, 16 Jul 2021             AIC:                            2140.210
Time:                     12:24:15               BIC:                            2154.483
Sample:                01-04-2015 - 01-05-2020   HQIC:                           2145.946
Covariance Type:            opg

=====

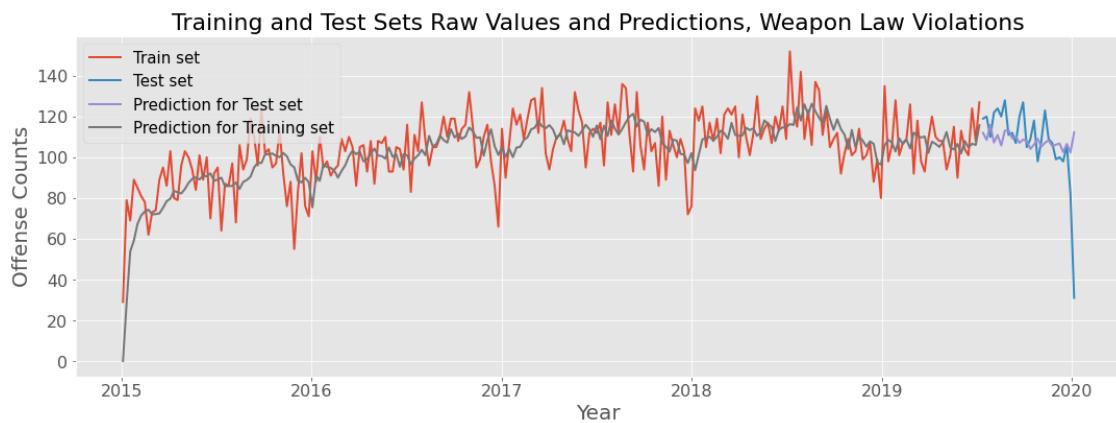
              coef    std err          z      P>|z|      [0.025      0.975]
-----
ar.L1      1.0000   9.32e-05   1.07e+04      0.000      1.000      1.000
ma.L1     -0.7818      0.041     -18.994      0.000     -0.863     -0.701
ma.S.L52    0.1099      0.058       1.880      0.060     -0.005      0.224
sigma2     194.3384     13.072      14.867      0.000    168.718    219.959
Ljung-Box (L1) (Q):      0.02 Jarque-Bera (JB):        75.04
Prob(Q):            0.90 Prob(JB):            0.00
Heteroskedasticity (H):  1.09 Skew:            -0.54
Prob(H) (two-sided):    0.70 Kurtosis:           5.39
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""

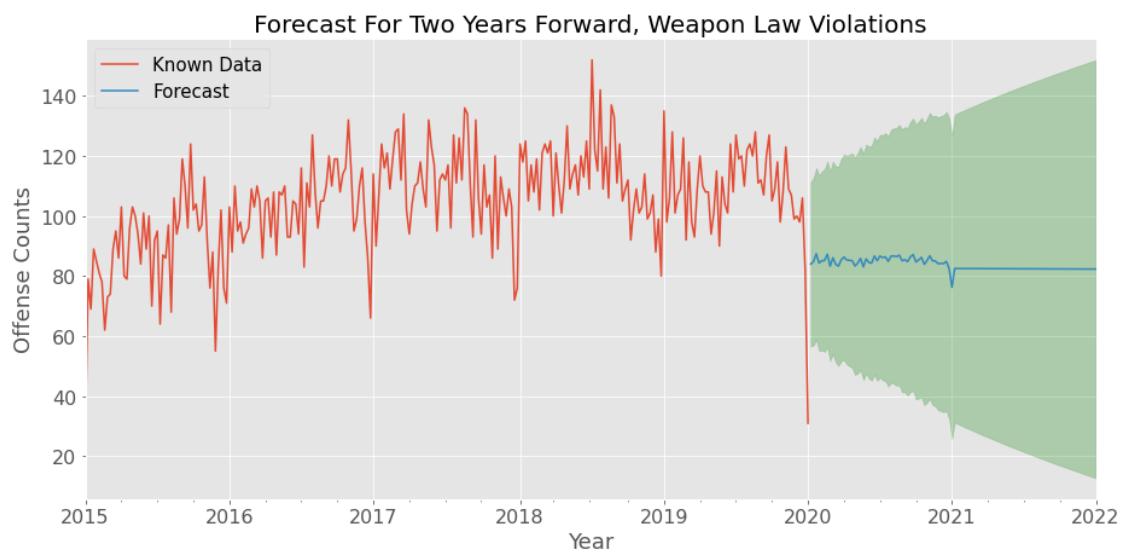
```



PREDICTION FOR TRAIN AND TEST sets:



FORECAST:



7 iNTERPRET

8 CONCLUSIONS & RECOMMENDATIONS

Summarize your conclusions and bullet-point your list of recommendations, which are based on your modeling results.

9 TO DO/FUTURE WORK

-

[]: