

```
In [1]: #!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Fri May 7 15:37:04 2021

@author: nacho
"""

#%%
import statsmodels.api as sm
import pandas as pd
import numpy as np
import matplotlib
import seaborn as sns; sns.set()
from scipy import stats
from matplotlib import pyplot as plt
sns.set(color_codes=True)
import os, datetime
from collections import Counter
from matplotlib.offsetbox import AnchoredText
import matplotlib.ticker as ticker
from darts import TimeSeries
from darts.metrics import mape, mase
from darts.utils.statistics import check_seasonality, plot_acf, plot_
residuals_analysis
from darts.models import (
    NaiveSeasonal,
    NaiveDrift,
    Prophet,
    ExponentialSmoothing,
    ARIMA,
    AutoARIMA,
    StandardRegressionModel,
    Theta,
    FFT
)
import warnings
warnings.filterwarnings("ignore")
import logging
logging.disable(logging.CRITICAL)
#import movilidad
import statsmodels.api as sm
from itertools import product
import math
from sklearn.preprocessing import MinMaxScaler
matplotlib.rcParams['figure.figsize'] = [9.0, 5.0]
```

Importing plotly failed. Interactive plots will not work.

```
In [2]: path = "/home/nacho/Documents/coronavirus/Data_Mexico"
os.chdir(os.path.join(path))
delitos_df = pd.read_csv("data/IDEFC_NM_dic2020.csv", encoding='latin
-1',thousands=',')
#delitos_df = pd.read_csv("data/IDEFC_NM_mar2021.csv", encoding='lati
n-1',thousands=',')
#movilidad = movilidad.return_df()
df_movilidad = pd.read_csv("data/google_movilidad.csv", index_col='Da
te')
```

```
In [3]: delitos_df.head()
```

Out[3]:

	Año	Clave_Ent	Entidad	Bien jurídico afectado	Tipo de delito	Subtipo de delito	Modalidad	Enero	Febrero
0	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio doloso	Con arma de fuego	3	0
1	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio doloso	Con arma blanca	1	1
2	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio doloso	Con otro elemento	0	0
3	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio doloso	No especificado	2	0
4	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio culposo	Con arma de fuego	0	0

```
In [4]: df_movilidad.head()
```

Out[4]:

	Unnamed: 0	State	Residences
Date			
2020-02-16	2667462	NaN	-1.0
2020-02-17	2667463	NaN	-2.0
2020-02-18	2667464	NaN	-2.0
2020-02-19	2667465	NaN	-1.0
2020-02-20	2667466	NaN	-1.0

```
In [5]: #####
#####NACIONAL#####
#####
```

```
In [6]: ##hacer listas de valores unicos
delitos_list = delitos_df['Tipo de delito'].value_counts().index.tolist()
##Generar dict de delitos

delitos_sum_df = pd.DataFrame(index = pd.date_range('2015-01', '2021-01', freq='M'))
delitos_sum_df.index.name = 'date'
years = np.arange(2015,2021)
meses = ['Enero', 'Febrero', 'Marzo', 'Abril', 'Mayo', 'Junio', 'Julio', 'Agosto', 'Septiembre', 'Octubre', 'Noviembre', 'Diciembre']
#for i in range(len(delitos_sum_df)):
for index, row in delitos_sum_df.iterrows():
    for delito in delitos_list:
        values_delito = []
        for year in years:
            for mes in meses:
                value = delitos_df.loc[(delitos_df['Tipo de delito'] == delito) & (delitos_df['Año'] == year), mes].sum()
                values_delito.append(value)
            delitos_sum_df[delito] = values_delito
```

In [8]: delitos\_sum\_df

Out[8]:

	Robo	Homicidio	Lesiones	Secuestro	Feminicidio	Daño a la propiedad	Allanamiento de morada	Evasión de presos
date								
2015-01-31	55171	2548	15022	89	32	9468	908	9
2015-02-28	52765	2529	15953	108	30	9493	813	7
2015-03-31	56378	2511	17458	76	31	10313	1030	10
2015-04-30	52897	2563	18197	85	35	10007	1011	6
2015-05-31	54041	2773	18838	96	28	10347	1102	13
...	...	...	...	...	...	...	...	...
2020-08-31	49490	3779	15366	78	76	9888	1127	12
2020-09-30	50114	3452	15886	69	78	10158	1092	6
2020-10-31	52899	3697	16682	63	74	11307	1216	13
2020-11-30	49293	3601	15250	55	82	10065	987	7
2020-12-31	48674	3610	14874	57	77	10623	1006	7

72 rows × 40 columns

```
In [7]: ###generar df con cada delito con su corr y trend
df_dsct = pd.DataFrame(list(product(delitos_list)), columns=['delito'
])
df_dsct['corr'] = ""
df_dsct['trend'] = ""
###Generar df por delito por estado y su porcentaje de crecimiento res
pecto al año pasado
for delito in delitos_list:
    df_delito = delitos_sum_df[delito]
    trend = df_delito.resample('Y').sum().pct_change()[-1]
    if np.isinf(trend):
        trend = math.inf
    df_dsct.loc[df_dsct['delito'] == delito, ['trend']] = trend
```

```
In [10]: #%%corr con movilidad
df_corr_movilidad = df_movilidad[df_movilidad['State'].isnull()]

for delito in delitos_list:
    df_delito = delitos_sum_df[delito]
    movilidad_spec = df_corr_movilidad.loc[(df_corr_movilidad.index >
'2020-02-15') & (df_corr_movilidad.index < '2021-01-01')]
    movilidad_spec.index = pd.to_datetime(movilidad_spec.index)
    movilidad_spec = movilidad_spec['Residences'].resample('M').sum()
    df_delito = df_delito.loc['2020-02':]
    corr = stats.pearsonr(movilidad_spec, df_delito)[0]
    df_dsct.loc[df_dsct['delito'] == delito, ['corr']] = corr
```

```
In [11]: #%%limpiar df
df_dsct = df_dsct.sort_values(by=['corr'], ignore_index=True, ascending = False)
```

In [12]: df\_dsct

Out[12]:

	delito	corr	trend
0	Violencia de género en todas sus modalidades d...	0.253194	0.273585
1	Trata de personas	0.114857	0.0128676
2	Otros delitos que atentan contra la vida y la ...	-0.0445427	0.224678
3	Tráfico de menores	-0.0929915	-0.275862
4	Electorales	-0.137224	-0.228951
5	Homicidio	-0.294743	-0.0349346
6	Rapto	-0.322076	0.103448
7	Incesto	-0.33408	2
8	Secuestro	-0.376247	-0.38149
9	Violencia familiar	-0.415299	0.0469647
10	Despojo	-0.448734	0.00598255
11	Extorsión	-0.458547	-0.0888482
12	Otros delitos contra la familia	-0.464545	0.0650299
13	Contra el medio ambiente	-0.48649	-0.0937357
14	Aborto	-0.491871	-0.114685
15	Feminicidio	-0.510792	-0.00212314
16	Otros delitos que atentan contra la libertad y...	-0.51251	0.269723
17	Fraude	-0.52436	-0.0415755
18	Evasión de presos	-0.531599	-0.204724
19	Allanamiento de morada	-0.618816	-0.0975869
20	Violación equiparada	-0.660255	0.149973
21	Abuso de confianza	-0.669984	-0.0986362
22	Narcomenudeo	-0.702535	0.0923498
23	Otros delitos que atentan contra la libertad p...	-0.715252	-0.0834429
24	Otros delitos contra la sociedad	-0.726916	0.0370788
25	Otros delitos contra el patrimonio	-0.734469	-0.0791874
26	Violación simple	-0.743516	-0.0979789
27	Incumplimiento de obligaciones de asistencia f...	-0.760239	-0.275944
28	Amenazas	-0.787718	-0.051297
29	Acoso sexual	-0.798168	0.331589
30	Daño a la propiedad	-0.823498	-0.141287
31	Otros delitos del Fuero Común	-0.830544	-0.0525269
32	Lesiones	-0.837164	-0.135075
33	Corrupción de menores	-0.837453	0.117566
34	Falsedad	-0.850414	-0.250142

	delito	corr	trend
35	Abuso sexual	-0.852041	-0.0528254
36	Falsificación	-0.853735	-0.291809
37	Delitos cometidos por servidores públicos	-0.859945	0.0181463
38	Hostigamiento sexual	-0.866934	-0.0586022
39	Robo	-0.933648	-0.211102

```
In [13]: ###specific scaler
def spec_scaler(series, minv, maxv):
    series = series.values.reshape(-1, 1)
    scaler = MinMaxScaler(feature_range=(minv, maxv))
    scaler = scaler.fit(series)
    normalized_series = scaler.transform(series)
    return normalized_series

###prediction
def predict_series(df_delito_state, delito):
    df_pred = df_delito_state.copy()
    df_pred = pd.DataFrame(df_pred)

    df_pred['Year'] = pd.date_range('2015-01', '2021-01', freq='M')
    series = TimeSeries.from_dataframe(df_pred, 'Year', delito)
    #train, val = series.split_before(pd.Timestamp('20200201'))
    train, val = series.split_before(pd.Timestamp('20191230'))

    model = Prophet()
    #model = ExponentialSmoothing()
    model.fit(train)
    prediction = model.predict(len(val))
    prediction = prediction.pd_dataframe()
    prediction[prediction < 0] = 0
    return prediction
```



In [14]: `%%plot df_dsct`

```
for index in [0,1,2,-1,-2,-3,-4,-5]:

    index_delito_state = df_dsct.iloc[index,:]
    delito = index_delito_state.loc['delito']
    df_delito = delitos_sum_df[delito]

    decom = sm.tsa.seasonal_decompose(df_delito, model = 'additive')

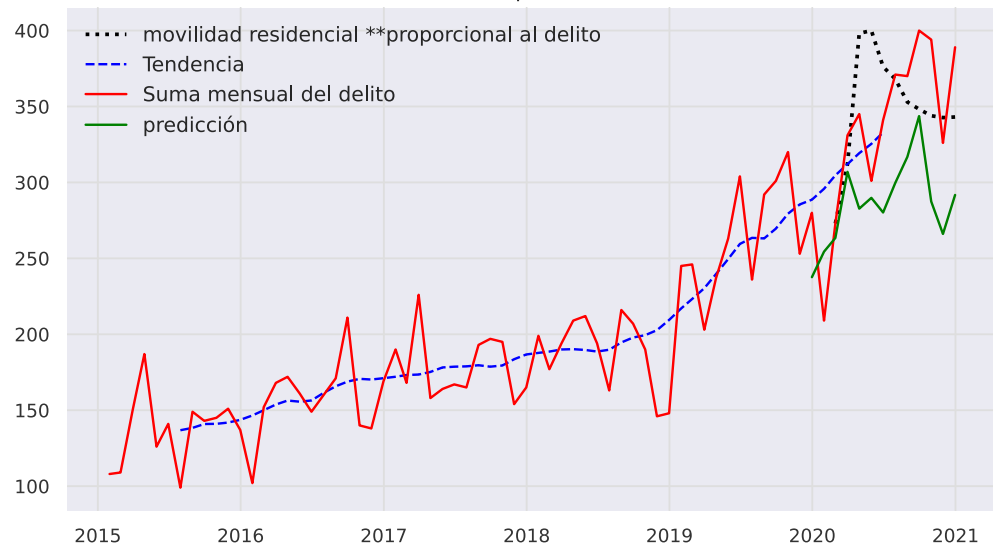
    movilidad_spec = df_corr_movilidad.loc[df_corr_movilidad.index <
'2021-01-01']
    movilidad_spec.index = pd.to_datetime(movilidad_spec.index)
    movilidad_spec = movilidad_spec['Residences'].resample('M').sum()
    #plot
    fig, ax = plt.subplots()
    ax.plot(movilidad_spec.index, spec_scaler(series = movilidad_spec
,
                                minv = df_delito.loc[df
_delito.index > '2020-02-15'].resample('M').sum().min(),
                                maxv = df_delito.loc[df
_delito.index > '2020-02-15'].resample('M').sum().max()),
            label = 'movilidad residencial **proporcional al delito',
            linestyle = ':',
            color = 'black',
            linewidth=2)
    ax.plot(decom.trend.index, decom.trend.values, label = 'Tendenci
a', linestyle='--', color='blue')
    ax.plot(decom.trend.index, df_delito.resample('M').sum(), label =
'Suma mensual del delito', color='red')
    ax.set_title(delito+' en la República Mexicana '+'
'\n'+ 'correlación del delito con movilidad: '+str(ro
und(df_dsct.iloc[index,:]['corr'], 2))+
'\n'+ 'tendencia del delito respecto al año anterior:
'+str(round(df_dsct.iloc[index,:]['trend'], 2)))
    df_pred = predict_series(df_delito, delito)
    ax.plot(df_pred.index,
            df_pred.values,
            label = 'predicción',
            color = 'green')

    ax.legend(loc='upper left')
    plt.plot()
```

## Violencia de género en todas sus modalidades distinta a la violencia familiar en la República Mexicana

correlación del delito con movilidad: 0.25

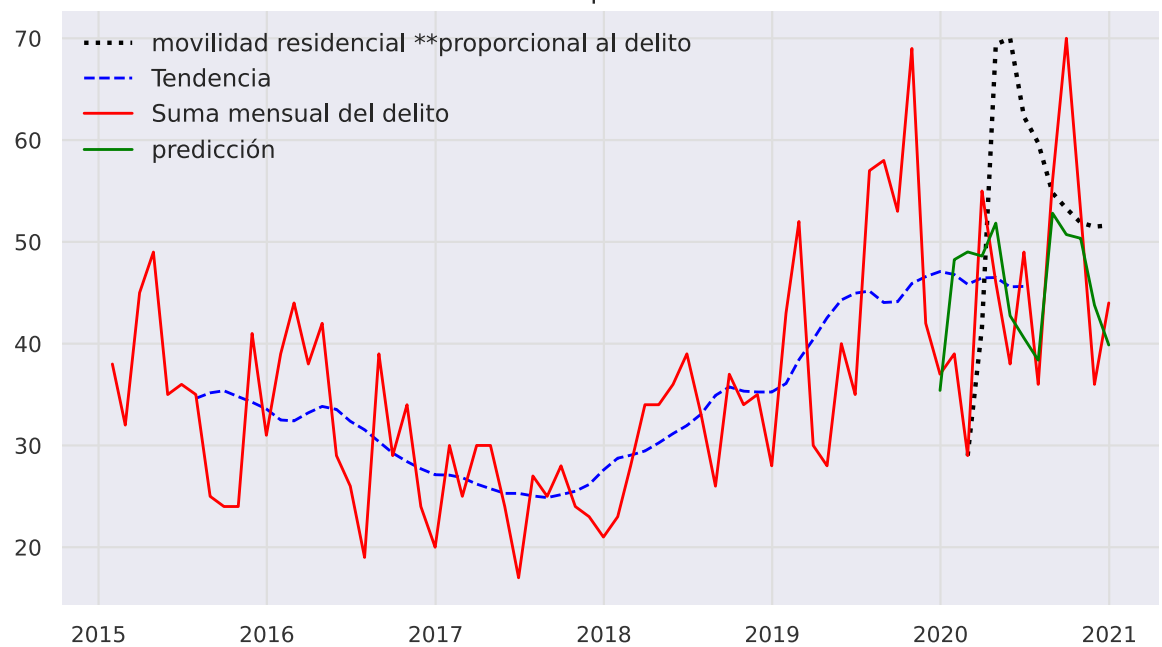
tendencia del delito respecto al año anterior: 0.27



## Trata de personas en la República Mexicana

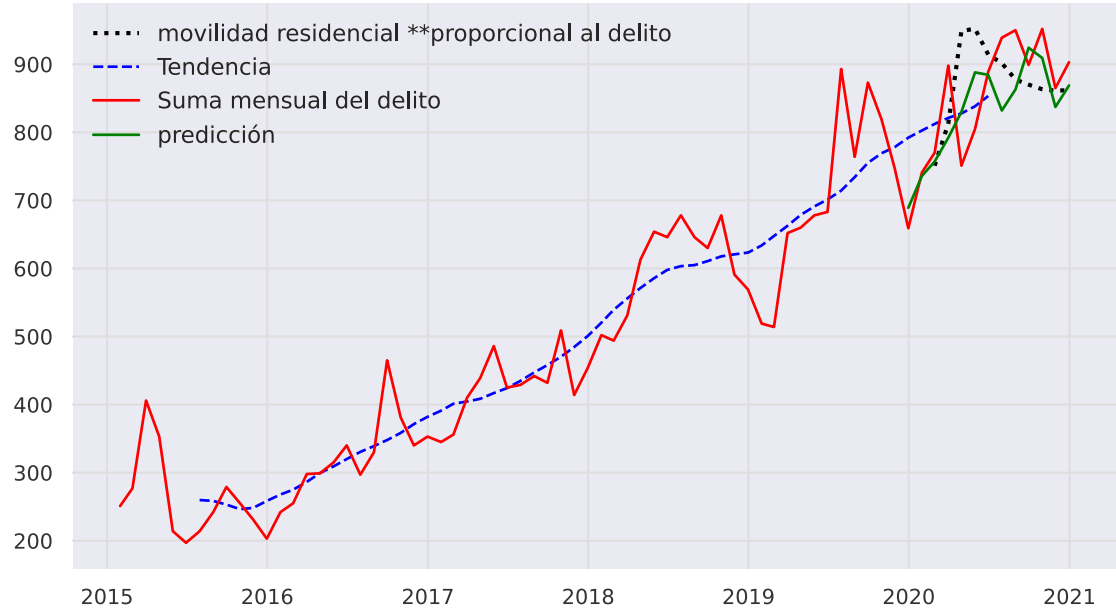
correlación del delito con movilidad: 0.11

tendencia del delito respecto al año anterior: 0.01



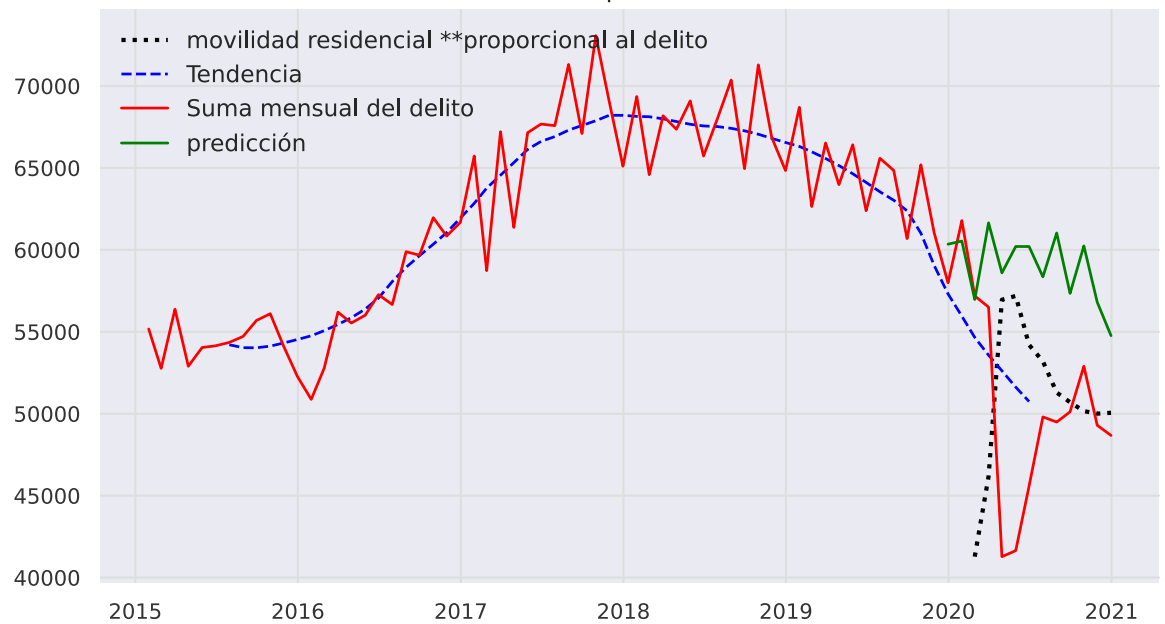
# Otros delitos que atentan contra la vida y la integridad corporal en la República Mexicana

correlación del delito con movilidad: -0.04  
tendencia del delito respecto al año anterior: 0.22

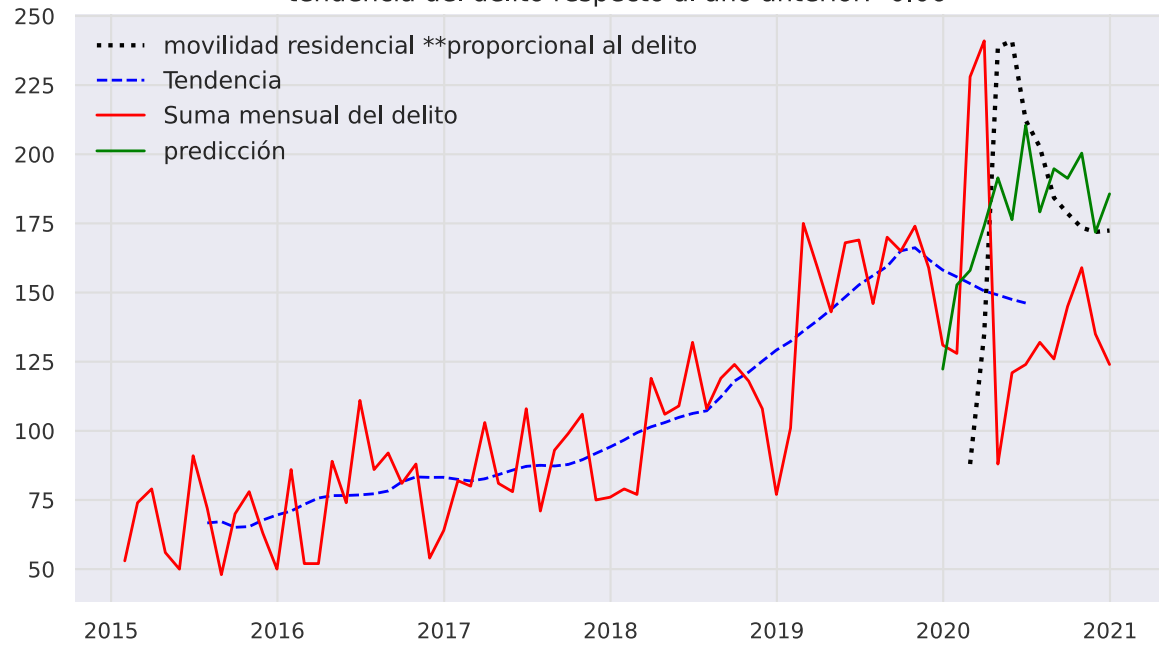


## Robo en la República Mexicana

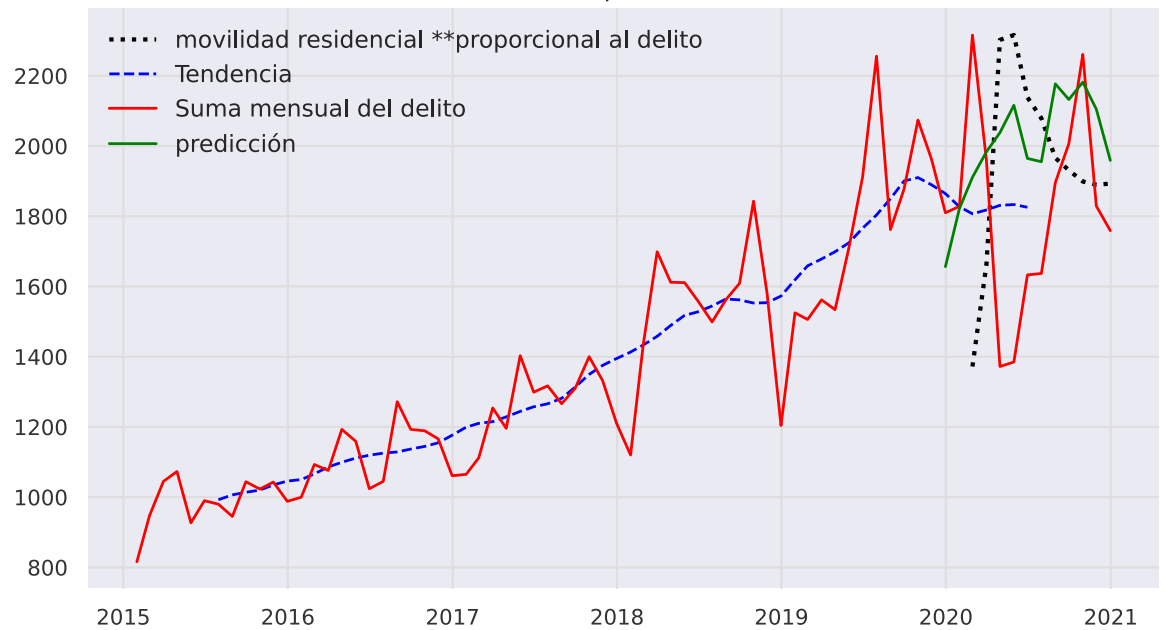
correlación del delito con movilidad: -0.93  
tendencia del delito respecto al año anterior: -0.21



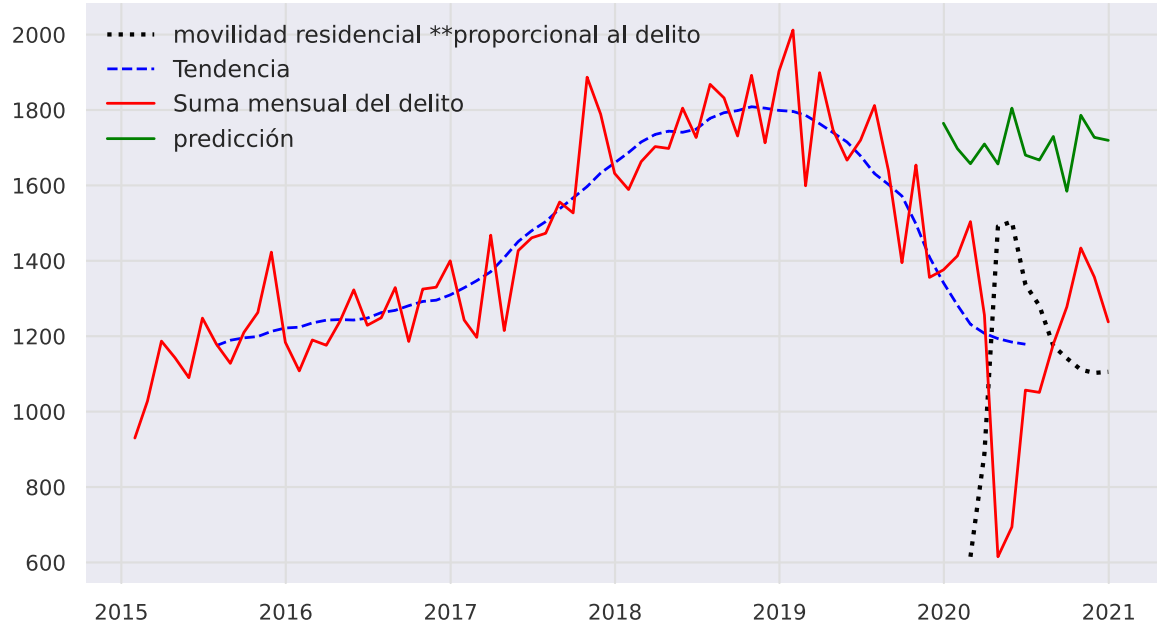
Hostigamiento sexual en la República Mexicana  
 correlación del delito con movilidad: -0.87  
 tendencia del delito respecto al año anterior: -0.06



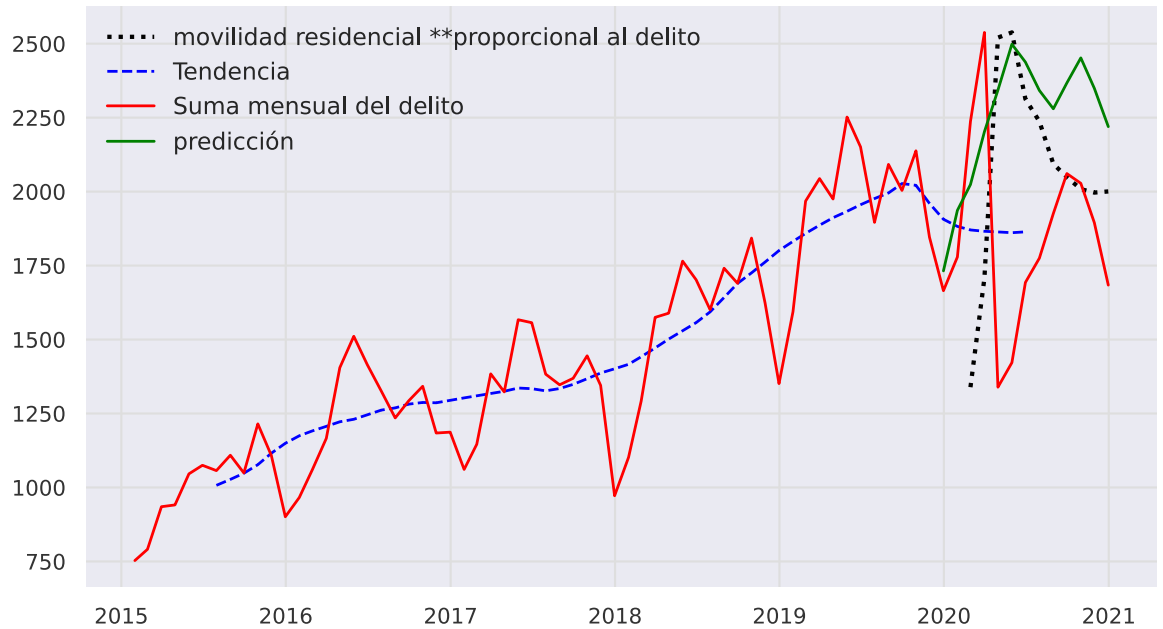
Delitos cometidos por servidores públicos en la República Mexicana  
 correlación del delito con movilidad: -0.86  
 tendencia del delito respecto al año anterior: 0.02



Falsificación en la República Mexicana  
 correlación del delito con movilidad: -0.85  
 tendencia del delito respecto al año anterior: -0.29



Abuso sexual en la República Mexicana  
 correlación del delito con movilidad: -0.85  
 tendencia del delito respecto al año anterior: -0.05



In [ ]: