File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Markdown

Time-Series Forecasting FBProphet

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import missingno as msno
%matplotlib inline
```

STEP #1: IMPORTING DATA

In [2]:
```
# !pip install fbprophet
```

In [3]:
```
# from fbprophet import Prophet
```

In [4]:
```
%%time
chicago_df_1 = pd.read_csv('Chicago_Crimes_2005_to_2007.csv', error_bad_lines=False)
chicago_df_2 = pd.read_csv('Chicago_Crimes_2008_to_2011.csv', error_bad_lines=False)
chicago_df_3 = pd.read_csv('Chicago_Crimes_2012_to_2017.csv', error_bad_lines=False)
```

```
C:\Users\Perry\anaconda3\lib\site-packages\IPython\core\magics\execution.py:1321: FutureWarning: The error_bad_lines argument has been deprecated and will be removed in a future version.


  exec(code, glob, local_ns)
b'Skipping line 533719: expected 23 fields, saw 24\n'
b'Skipping line 1149094: expected 23 fields, saw 41\n'
Wall time: 25.3 s
```

In [5]:
```
chicago_df_1.shape
```

Out[5]: (1872343, 23)

In [6]:
```
chicago_df_2.shape
```

Out[6]: (2688710, 23)

In [7]:
```
chicago_df_3.shape
```

Out[7]: (1456714, 23)

In [8]:
```
df = pd.concat(objs=[chicago_df_1, chicago_df_2, chicago_df_3], axis=0)
df.shape
```

Out[8]: (6017767, 23)

In [9]:
```
1872343 + 2688710 + 1456714
```

Out[9]: 6017767

STEP #2: EXPLORING THE DATASET [...]

STEP #3: PREPARING THE DATA

In [21]:
```
df_prophet = df.resample('M').size().reset_index()
df_prophet.columns = ['Date', 'Crime Count']
df_prophet.head()
```

Out[21]:

	Date	Crime Count
0	2005-01-31	33983
1	2005-02-28	32042
2	2005-03-31	36970
3	2005-04-30	38963
4	2005-05-31	40572

In [22]:
```
# Rename the columns names
df_prophet.rename(mapper={"Date": "ds", "Crime Count": "y"}, axis=1)
#
df_prophet.rename(columns={"Date": "ds", "Crime Count": "y"})
#
df_prophet.columns = ["ds", "y"]
```

In [23]:
```
df_prophet
```

Out[23]:

	ds	y
0	2005-01-31	33983
1	2005-02-28	32042
2	2005-03-31	36970
3	2005-04-30	38963
4	2005-05-31	40572
...	...	...
140	2016-09-30	23235
141	2016-10-31	23314

142	2016-11-30	21140
143	2016-12-31	19580
144	2017-01-31	11357

145 rows × 2 columns

In [24]:
```
for row in df.to_dict(orient="records"):
    print(row)
    break
```

```
{'ID': 4673626, 'Date': Timestamp('2006-04-02 13:00:00'), 'Block': '055XX N MANGO AVE', 'Primary Type': 'OTHER OFFENSE', 'De
scription': 'HARASSMENT BY TELEPHONE', 'Location Description': 'RESIDENCE', 'Arrest': False, 'Domestic': False}
```

STEP #4: MAKE PREDICTIONS

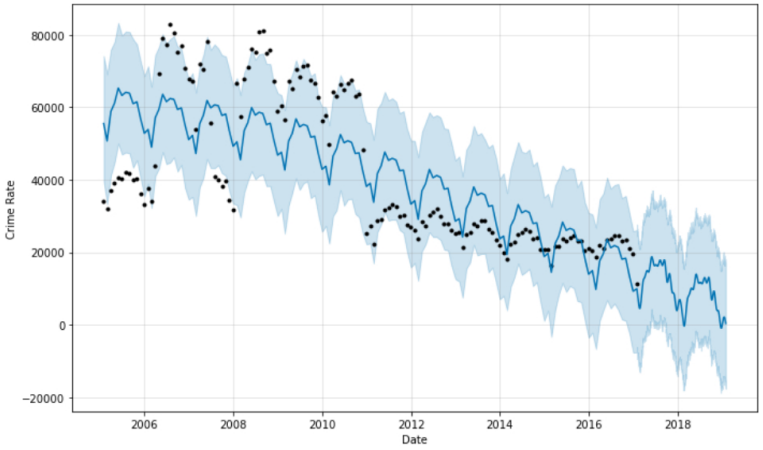
In [25]:
```
from fbprophet import Prophet

model = Prophet()
model.fit(df_prophet)
```

```
INFO:fbprophet:Disabling weekly seasonality. Run prophet with weekly_seasonality=True to override this.
INFO:fbprophet:Disabling daily seasonality. Run prophet with daily_seasonality=True to override this.
```

Out[25]: <fbprophet.forecaster.Prophet at 0x1d5837563a0>

In [26]:
```
# Forcasting into the future
future = model.make_future_dataframe(periods=730)
forecast = model.predict(future)
```

In [27]:
```
figure = model.plot(forecast, xlabel='Date', ylabel='Crime Rate')
```



In [28]:
```
figure = model.plot_components(forecast)
```