iOLAP

INTRO TO TIME SERIES ANALYSIS

Hana Rizvić

Advanced Analytics Team Lead

Renee Ahel

Freelance Data Scientist

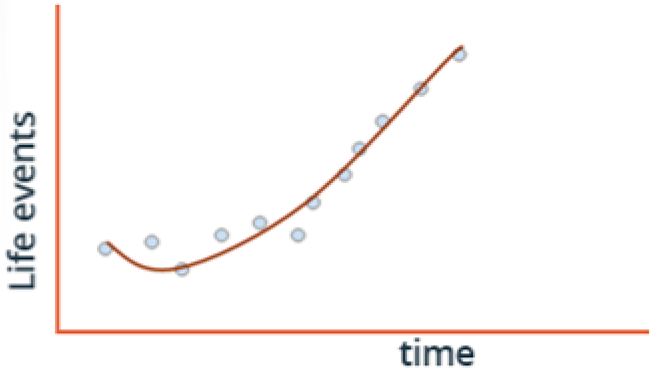
- ABOUT US
- MORE ABOUT IOLAP AND DSC
- FUTURE MEETUPS

AGENDA

- What are time series?
- Overview of usage
- Overview of methods
- Main components and properties of time series
- SARIMA explained

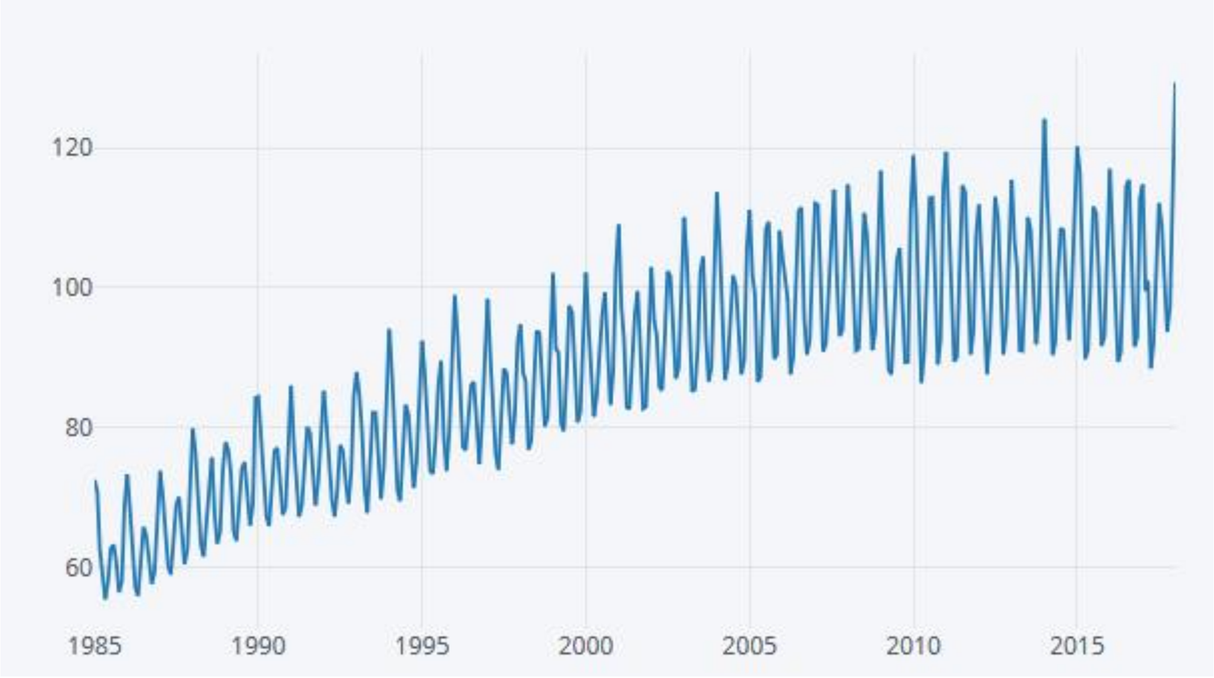
WHAT ARE TIME SERIES?

Notion of a real world
event as an abstraction of a sequence
of timely activities.



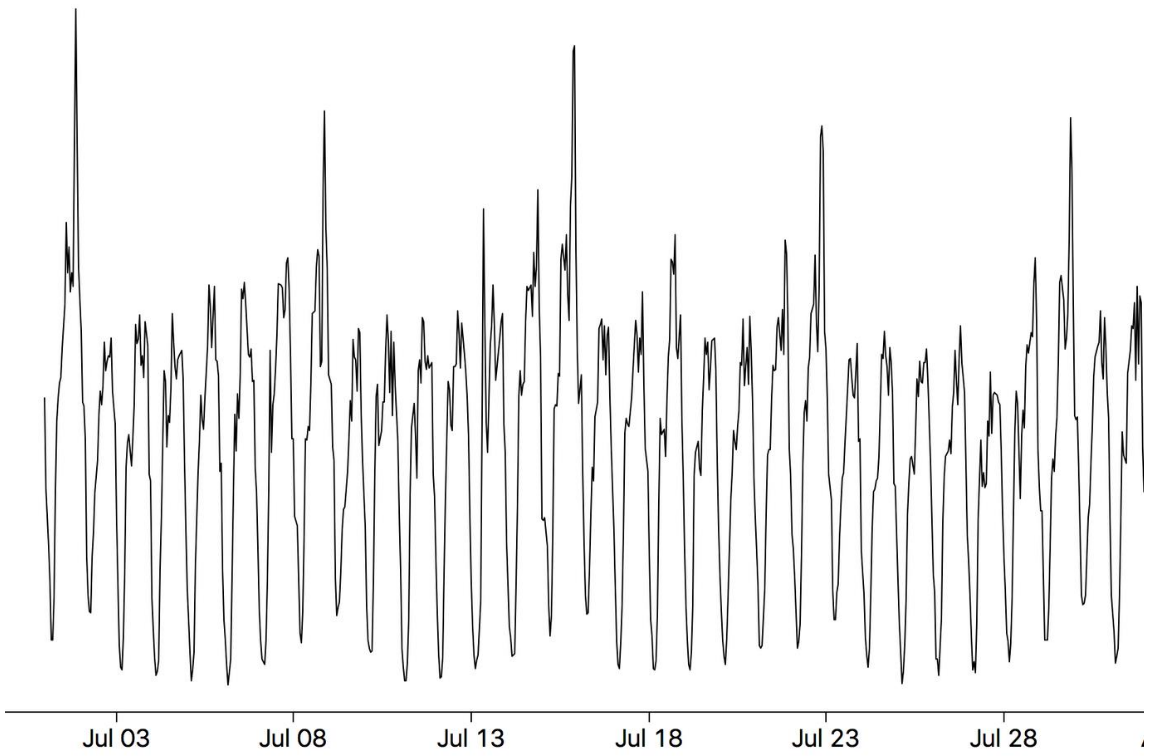
TIME SERIES EXAMPLES

Energy production Jan 1985 – Sep 2018



https://fred.stlouisfed.org/series/IPG2211A2N

Hourly sum of Uber trips in July 2017



FOR WHAT IS USED

In terms of statistics, econometrics, finance, meteorology...

- **Short – term predictions**
  - Planning and scheduling: equipment, personnel, financial, levels of inventory, personnel etc
- **Medium – term predictions**
  - Budgeting purposes mostly – require predictions of economic and industry variables
- **Long term predictions**
  - Needed for capital expansion plans, selecting R&D projects, launching new products, formulating long-term goals
- **Trend tracking**
  - Changes in trend due to newly introduced factors – marketing campaigns, environmental changes

OVERVIEW OF METHODS AND DIVISIONS

iOLAP

- Frequency domain
  - Spectral analysis
  - Wavelet analysis
- Time domain
  - Auto-correlation
  - Cross-correlation
- Assumptions
  - Parametric
  - Non-Parametric
- By look and count
  - Linear and non-linear
  - Uni- or Multivariate

ARIMA

AutoRegressive Integrated Moving Average

- **AR(p)** $$Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \cdots + b_p Y_{t-p} + e_t$$
- **I(d)** $$Y_t' = Y_t - Y_{t-1}$$
- **MA(q)** $$Y_t = b_0 + b_1 e_{t-1} + b_2 e_{t-2} + \cdots + b_q e_{t-q} + e_t$$
- **ARIMA(p,d,q) + Seasonal = ARIMA(p,d,q)(P,D,Q)**

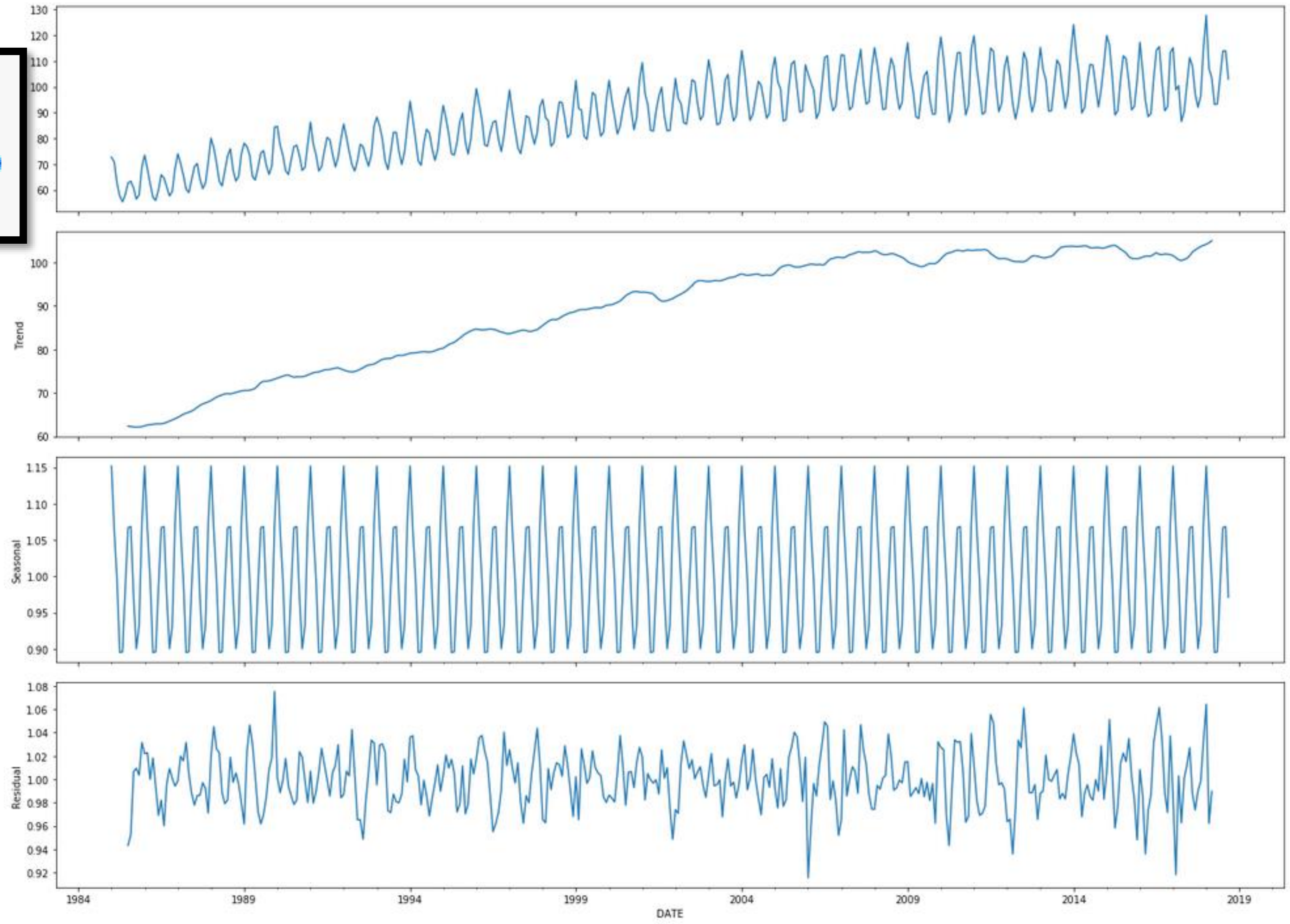
COMPONENTS

```
data.index = pd.to_datetime(data.index)
from statsmodels.tsa.seasonal import seasonal_decompose
result = seasonal_decompose(data, model='multiplicative')
fig = result.plot()
```

DATA = Trend X Seasonal X Error

DATA = Trend + Seasonal + Error

Additive or Multiplicative?



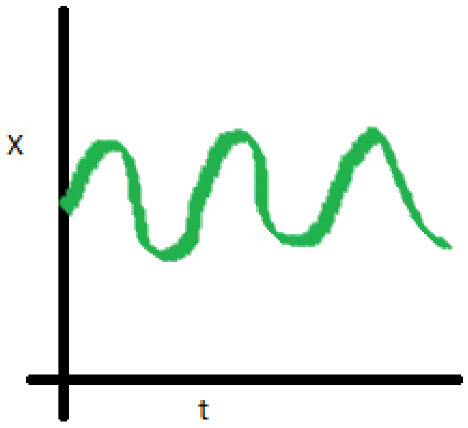
PROCESS FLOW

iOLAP

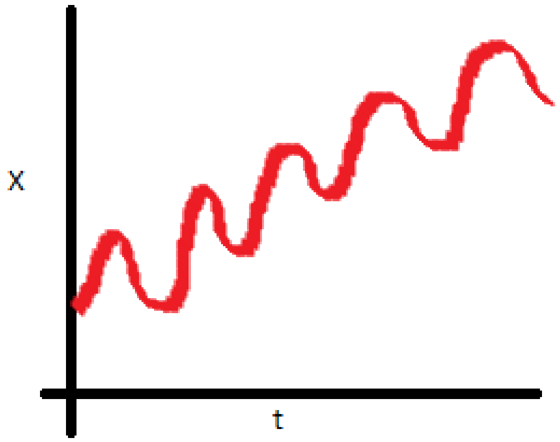
1. Visualize the time series
2. Stationarize the series
3. Plot ACF/PACF charts and find optimal parameters
4. Build the SARIMA model
5. Make predictions

STATIONARITY

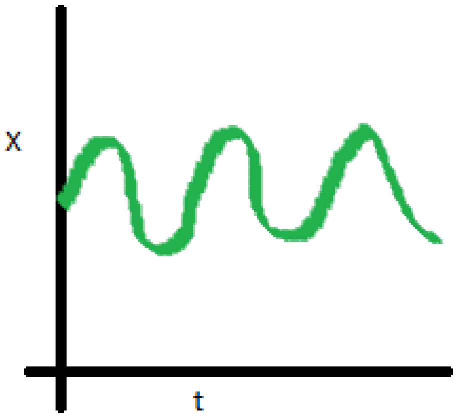
- Unit roots
- Augmented Dickey Fuller test



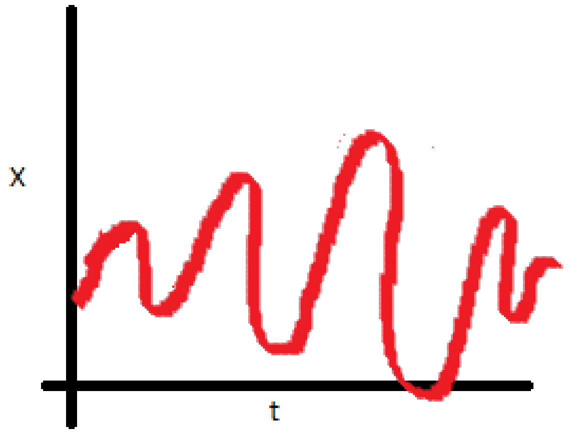
Stationary series



Non-Stationary series



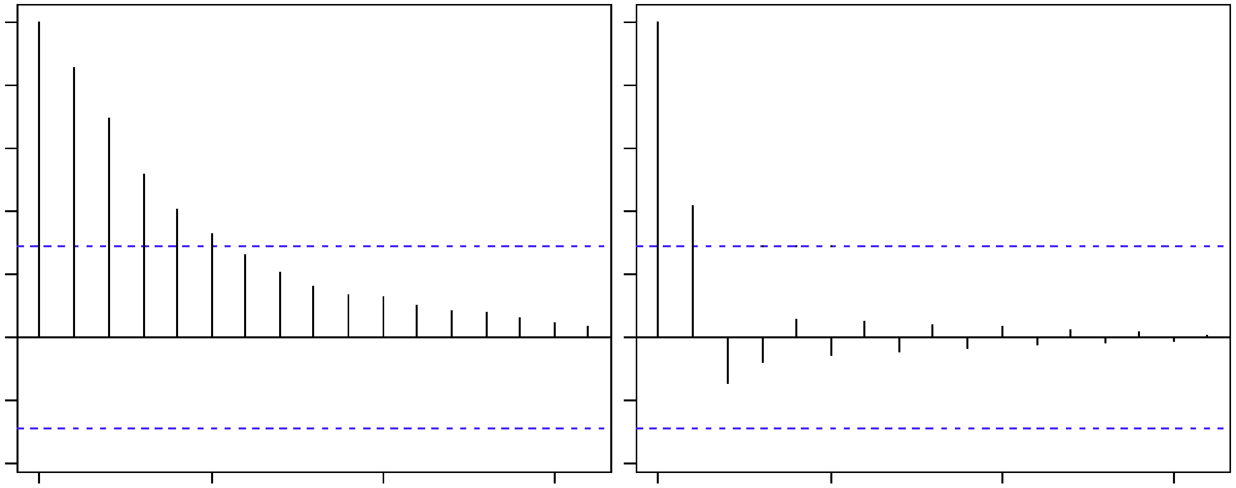
Stationary series



Non-Stationary series

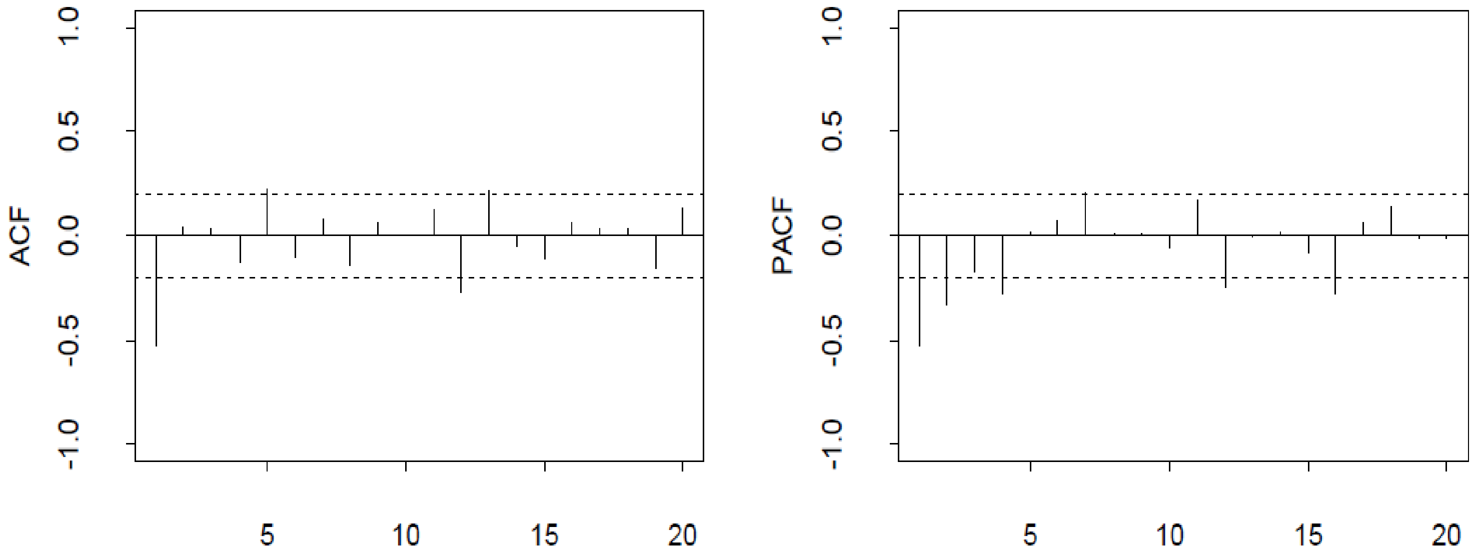
ACF AND PACF

ACF **PACF**



AR(1) process

MA(1) process



SARIMA - PARAMETERS

iOLAP

```
from pyramid.arima import auto_arima

train = data.loc['1985-01-01':'2016-12-01']
test = data.loc['2017-01-01':]

stepwise_model = auto_arima(train, start_p=1, start_q=1,
                           max_p=3, max_q=3, m=12,
                           start_P=0, seasonal=True,
                           d=1, D=1, trace=True,
                           error_action='ignore',
                           suppress_warnings=True,
                           stepwise=True)

print(stepwise_model.aic())
```

ARIMA(p,d,q)(P,D,Q)

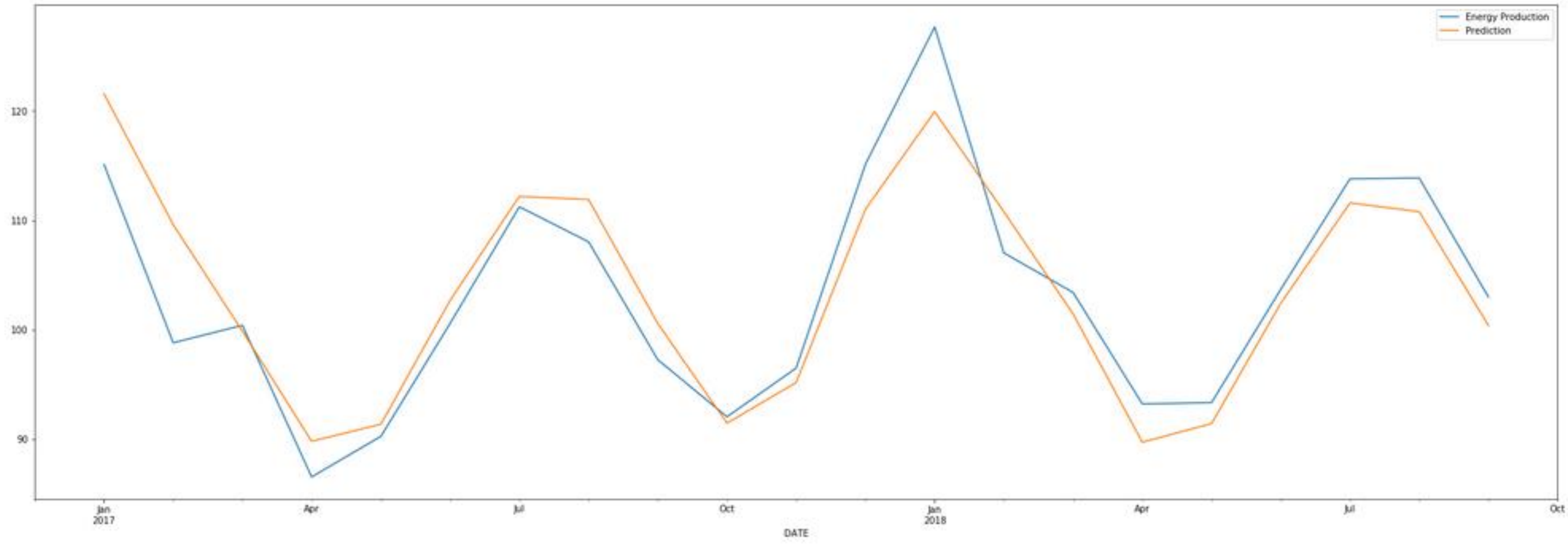
```
Fit ARIMA: order=(0, 1, 2) seasonal_order=(2, 1, 1, 12); AIC=1696.141, BIC=1723.555, Fit time=3.047 seconds
Fit ARIMA: order=(2, 1, 2) seasonal_order=(2, 1, 1, 12); AIC=1689.560, BIC=1724.806, Fit time=8.076 seconds
Fit ARIMA: order=(1, 1, 1) seasonal_order=(2, 1, 1, 12); AIC=1686.438, BIC=1713.852, Fit time=5.175 seconds
Fit ARIMA: order=(1, 1, 3) seasonal_order=(2, 1, 1, 12); AIC=1689.153, BIC=1724.399, Fit time=8.269 seconds
```

```
print(str('ARIMA')+str(stepwise_model.order)+str(stepwise_model.seasonal_order))

ARIMA(1, 1, 2)(2, 1, 1, 12)
```

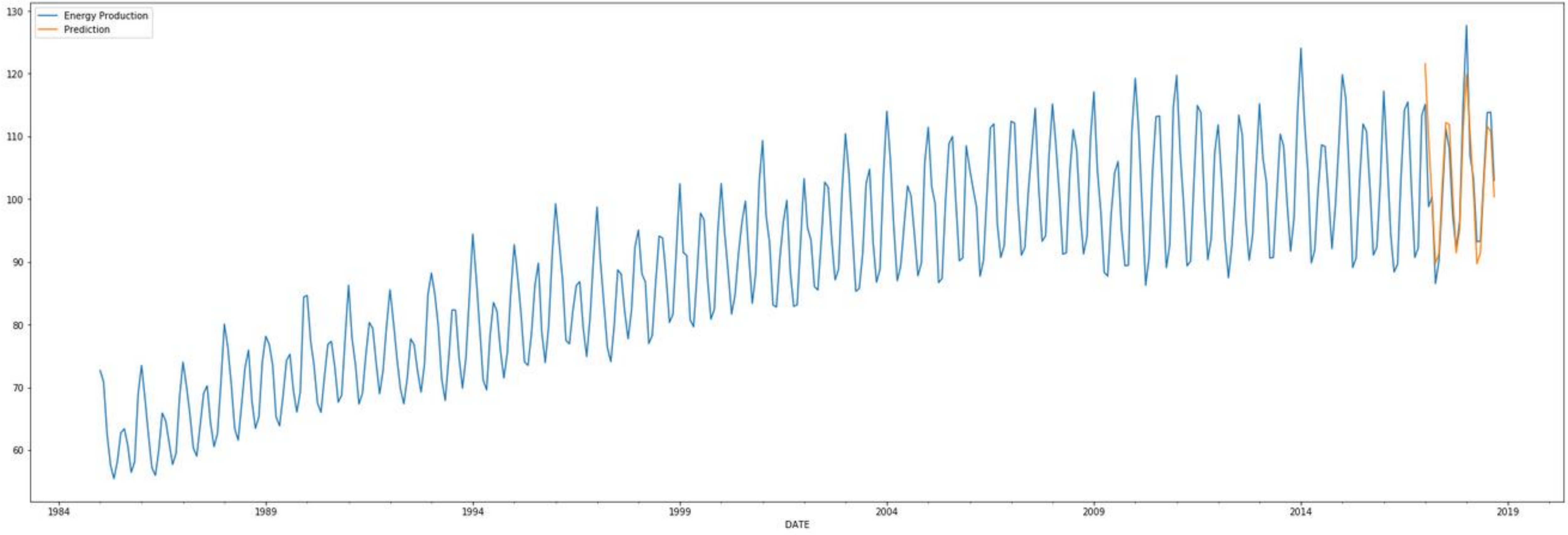
SARIMA - PREDICTION

```
stepwise_model.fit(train)
future_forecast = stepwise_model.predict(n_periods=len(test))
future_forecast = pd.DataFrame(future_forecast,index = test.index,columns=['Prediction'])
pd.concat([test,future_forecast],axis=1).plot()
```



SARIMA - FINAL

iOLAP



RECOMMENDED READING

- **Forecasting : methods and applications**
  - Spyros Makridakis, Steven C. Wheelwright, Victor E. McGee
- **Practical Time Series Analysis**
  - By Avishek and Prakash
- **Time Series Analysis: Forecasting and Control (3rd Edition)**
  - Box et al.

iOLAP

THANK YOU!
Q&A