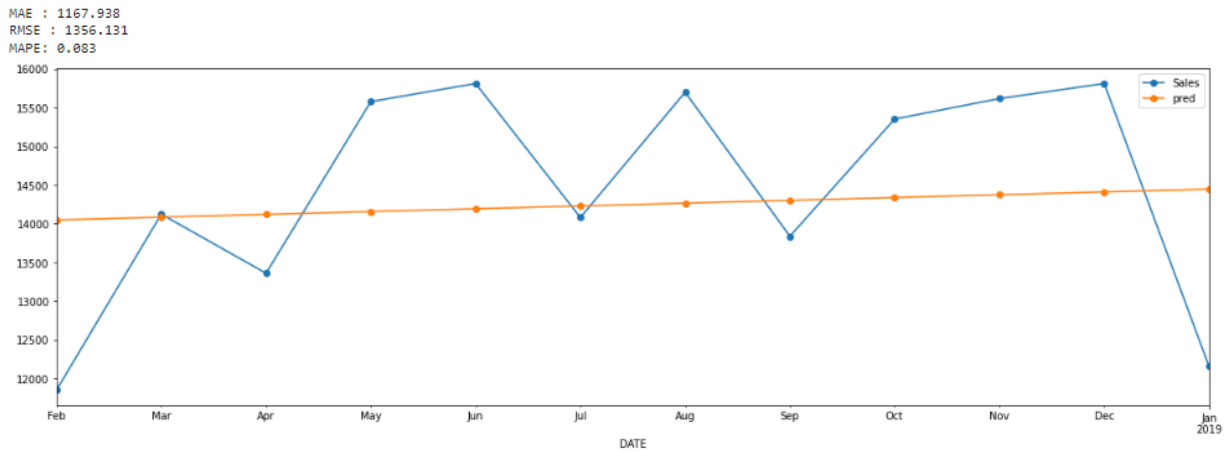
Time Series Analysis Lecture — 2

2. Double Exponential Smoothing (DES):



- **Purpose:** Addresses SES's lack of trend capture by incorporating trend into forecasts.
- **Components**:
  - **Level:** Short-term average value.
  - **Trend:** Direction and rate of data movement over time.
- **Formulation:** Applies exponential smoothing to both level and trend. The formulation of DES is as follows:

  $$\hat{y}_{t+h} = l_t + hb_t$$

  where

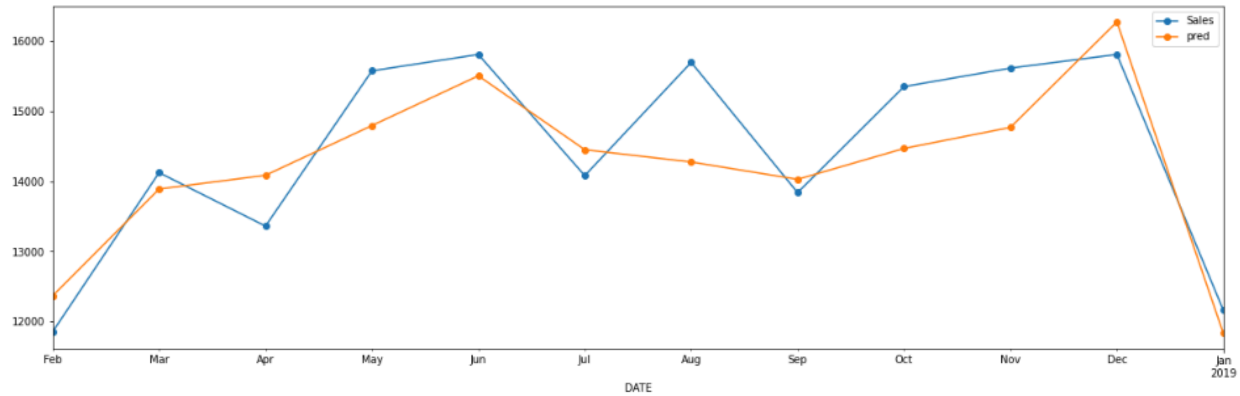
  $$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$$

  l_t is called the **level** of time series at time t.

- **Smoothing Parameters:**
  - α: Corresponds to level series.
  - β: Corresponds to trend series; requires tuning.
- **Performance:** Better fit than SES, with a lower error rate (8.3% vs. 10% for SES).
- **Limitations:** Does not account for seasonality.

3. Triple Exponential Smoothing (TES):

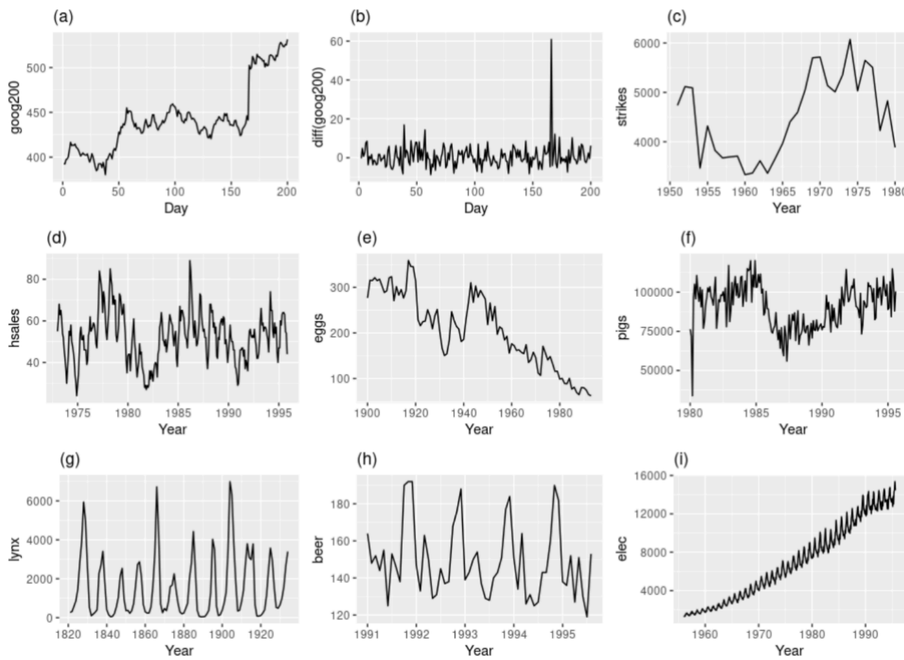
MAE : 588.607
RMSE : 680.844
MAPE: 0.04



- **Purpose:** Extends DES by adding support for seasonality.
- **Components**:
    - Level: Short-term average value.
    - Trend: Rate of data movement over time.
- **Seasonal:** Seasonal variations.
- **Formulation**: Incorporates level, trend, and seasonality in forecasting.
- **Smoothing Parameters:**
    - α: Corresponds to level.
    - β: Corresponds to trend.
    - γ: Corresponds to seasonality.
- **Performance:** Captures more information than DES, significant reduction in MAPE error to 4%.
- **Model Selection:** Additive vs. multiplicative models chosen based on performance; require testing.

Stationarity:

- **Definition:** Time series with constant mean, variance, amplitude, and frequency over time.
- **Non-Stationarity Indicators:** Presence of trends, seasonality.
- **Example:** Stationary heartbeat with consistent mean and standard deviation.
- **Characteristics:** No long-term predictable patterns.
- **Importance:** Many models require stationarity for accurate results.
- **Conversion:** Non-stationary series often transformed to stationary.
- **Assessment:** Stationarity is determined visually or with statistical tests like Dickey-Fuller.



- **Non-Stationary:** Plots a, c, e, f (due to trend or changing mean), d, h (due to seasonality), i (due to trend, seasonality, unstable variance).
- **Stationary:** Plot b (despite one outlier), plot g (assumed for modeling, irregular cyclic pattern).

Dickey-Fuller Test:

- **Purpose:** Tests stationarity in time series.
- **Hypotheses:**
  - H0: Time series is non-stationary.
  - H1: Time series is stationary.
- **Implementation:** sm.tsa.stattools.adfuller() in statmodels library.
- **Interpretation:** p-value < 0.05 indicates stationarity.

Converting Non-stationary to Stationary Time Series:

Detrending:

- **Method:** Differencing the series (value(t) = observation(t) - observation(t-1)).
- **Library Function**: diff() in pandas.
- **Non-linear Trends:** This may require multiple differencing steps.

Deseasonalizing:

- **m-Differencing:** Subtracting observation at the current timestep from the one at the last seasonal period (value(t) = observation(t) - observation(t-m)).
- **Seasonality Period (m):** Determined by the data's seasonal cycle.

Process:

- **Detrend**: First, use differencing to remove the trend.
- **Deseasonalize:** Then, remove seasonality, potentially using m-differencing.

Autocorrelation and Seasonality Detection:

- **Autocorrelation:** Correlation of a time series with its lagged version; identifies optimal lag (m) where series overlap.
- **Autocorrelation Function (ACF):** Shows direct and indirect correlation impacts; useful for spotting random series.
- **Partial Autocorrelation Function (PACF)**: Shows unique correlation by removing indirect effects; helps in identifying direct relationships and seasonality.
- **ACF and PACF Plots:** Reveal significant lags with correlations outside confidence intervals, indicating potential seasonality.
- **Usage:** ACF applied to stationary series; PACF to original series for direct impacts.

Correlation vs. Causation:

- **Correlation:** Relationship between two variables without implying cause.
- **Causation:** One variable directly affects another.
- **Example:** Ice cream sales and drownings correlate due to temperature, not causation.
- **Confounding Variable:** A third element influencing both correlated variables.
- **Forecasting Use:** Correlation useful even without causality.
- **Misinterpretation Risk:** Assuming causality from correlation can lead to incorrect conclusions.
- **Model Improvement:** Understanding causality helps identify better predictive features.