

Time Series Decomposition - STL



Recap

- **Decomposition**
 - braking down underlying pattern into **sub patterns** to identify the component factors
- Components
 - | A series = A combination of these components
 - Systematic
 - level, trend, seasonality
 - Non-Systematic
 - noise
- Classical Models
 - Additive: Time series data is a function of the **sum** of its components
 - Multiplicative: Time series data is a function of the **product** of its components

STL

| Seasonal and Trend decomposition using Loess

= statistical method of decomposing a Time Series data in to 3 components (seasonality, trend, residual(noise))

$$y_t = s_t + t_t + r_t$$

Components

- Trend (t_t)
 - general direction of the overall data
- Seasonality (s_t)
 - regular & predictable pattern (recur at a fixed interval of time)
- Residual(Noise; Randomness; r_t)
 - random fluctuation or unpredictable change

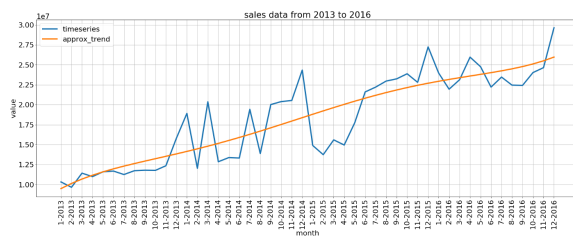


Loess

- nonparametric regression technique used local weighted regression to fit a smooth curve through points in a sequence
 - can reveal trend-cycle in data that might be difficult to model with parametric curve
 - e.g. non-linear, polynomial data

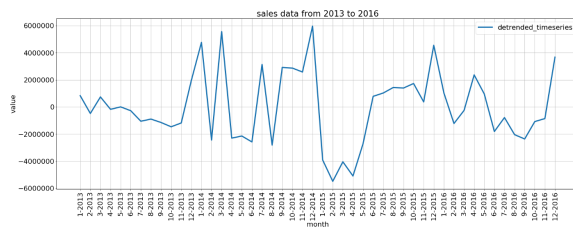
Step by Step Process

Step 1.



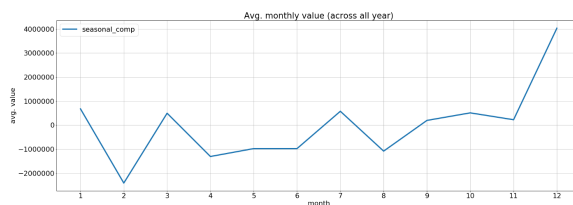
Find the **approximate trend line** which fits the Time Series data

Step 2.



Find **de-trended series** by subtracting Time Series data with approximate trend line

Step 3.

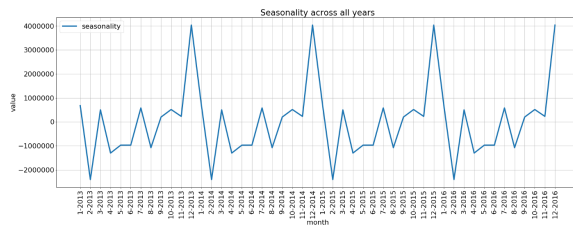


Find the **Seasonal** Component

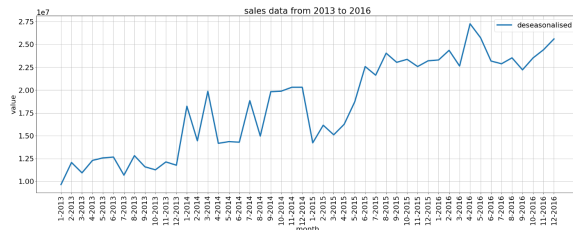
e.g. grouping de-trended series by month → find average value of each month

Step 4.

Repeat Seasonal component across the whole data

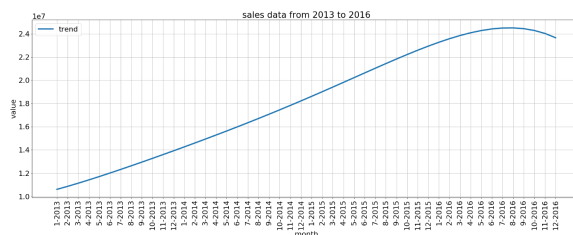


Step 5.



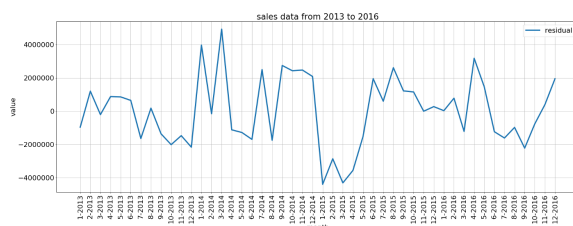
Find the **de-seasonalised series** by subtracting Time Series data with seasonal data

Step 6.



Find the **trend** by fitting the de-seasonalised data to a polynomial model

Step 7.



Find the residual by subtracting Time Series data with seasonal & trend

Pros & Cons

Pros

- Unlike SEATS, X11, STL can handle any type of seasonality (not only monthly, quarterly data)
- The seasonal component can be changed over time → the rate of change can be controlled by user
- The smoothness of the trend can also be controlled by the user
- Robust to outliers

Cons

- Does not handle calendar variation automatically (only provide facilities for additive decomposition)
 - e.g. In the case of stock prices, there is no data such as Saturday and Sunday. Therefore, before putting it into the STL, we have to transform data into daily one.

Reference

- Basic Concept of STL
 - <https://otexts.com/fpp2/stl.html>
 - <https://towardsdatascience.com/stl-decomposition-how-to-do-it-from-scratch-b686711986ec>
- statsmodels Library Application
 - https://www.statsmodels.org/v0.12.1/examples/notebooks/generated/stl_decomposition.html