analysing the patterns in the time series

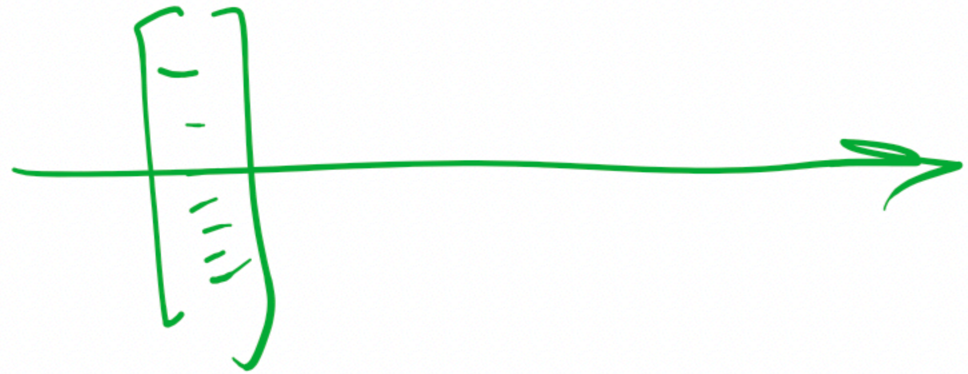
Time Series Analysis

& to make predictions.

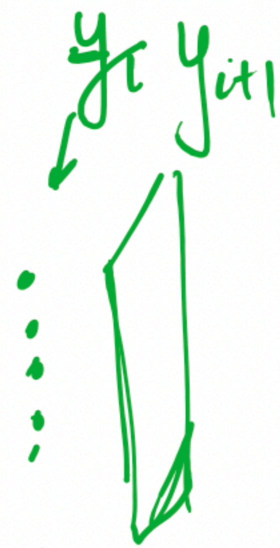
Time Series

- Time series analysis is the process of extracting meaningful non-trivial information and patterns from time series data.

- Time series forecasting is the process of predicting the future value of time series data based on past observations and other inputs
  - We make use of the historical information about a particular quantity to make forecasts about the value of the same quantity in the future.



ignored the past history.



time i

regressors.

dependent $\rightarrow$ $y = w_1 x_1 + w_2 x_2 \ldots + \ldots w_n x_n$

regressand

time i

regressors.

dependent $\rightarrow y = w_1 x_1 + w_2 x_2 + \dots + w_n x_n$

regressand

Input $\rightarrow$ output

x : predictors OR independent variables

Time Series Forecasting

Level. Trend

Time Series Forecasting

Decomposition based methods	Smoothing based methods	Regression based methods	Machine Learning based methods
Classical	Moving Averages	Regression with seasonality	Regression
STS, X11,..	Exponential Smoothing	ARIMA	Neural Network

't' is a regressor.

Fully Connected
LSTM
RNNs.
etc

Time Series Decomposition

- Time series decomposition is the process of deconstructing a time series into the number of constituent components with each representing an underlying phenomenon.
- Decomposition splits a time series into
  - Level
  - Trend component
  - Seasonal component
  - Noise
- The first 3 are the systematic components, while noise is the non-systematic component.

① Time Series Forecasting

$L_n \quad T_n \quad S_n$ noise.

$L_{n+1} \quad T_{n+1} \quad S_{n+1}$

- Forecasting with decomposition

The individual (decomposed) systematic components can be better forecasted using regression or similar techniques and combined together as an aggregated forecasted series.

Additive

Multiplicative

2

Time Series Forecasting

- Smoothing-based methods

window -based

These methods forecast future data by smoothing the past observations and projecting it to the future.

The future value of the time series is the weighted average of past observations.

Time Series Forecasting

- Regression-based methods

  This formulates forecasting as a supervised predictive model.

  However, the time t is an independent variable

  A linear regression model would look like $y_t = a \times t + b$

- A more sophisticated technique considers that data from adjacent time periods are correlated in a time series.

  One such technique is ARIMA (Auto Regressive Integrated Moving Average)

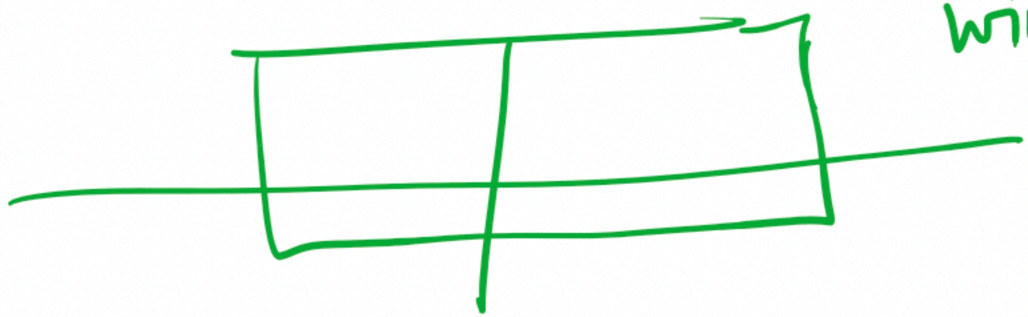
Time Series Forecasting

- Supervised Machine Learning Models for forecasting

We can extract a time window of input variables derived from the time series.

This kind of cross-sectional datasets can be used to make predictions.

window is also a cross section.



t=i

Decomposition based Methods

Time Series Decomposition

current value / Reference value

- Level: Average value of the time series variable
- Trend: Long-term tendency of data.

  - It represents change from one period to the next.

rise

t=n   t=n+1

- Seasonality: Repetitive behaviour during a cycle of time.

  Seasonality can be further split into hourly, daily, weekly, monthly, quarterly and yearly seasonality.

periodicity

- Cycle: Cyclic component represents longer-than-a-year pattern where there is no specific time frame between the cycles.

Time Series Decomposition

- Cycle:

  Examples: Cycle of stock market booms and crashes.

  Note that the length of the boom period, length of recession, time between subsequent booms and crashes is uncertain and random.

random white noise

- Noise: Anything that is not represented by level, trend, seasonality, or cyclic component is the noise in the time series.

Time Series Decomposition

- Time series decomposition can be classified into additive decomposition or multiplicative decomposition
- Additive decomposition: The decomposed components need to be added together to obtain the original time series

Time series = Trend + Seasonality + Noise

$$y_t = T_t + S_t + E_t$$

Level is not explicitly mentioned

- Multiplicative decomposition:

Time series = Trend × Seasonality × Noise

$$y_t = T_t \times S_t \times E_t$$

Time Series Decomposition

- How to decide whether to go for the additive or the multiplicative decomposition?
- Go for multiplicative decomposition if
  the magnitude or frequency of the seasonal fluctuation, or
  the variation in trend changes with the level of the time series (i.e. the trend changes non-linearly)

Time Series Decomposition

estimate T_t, S_t

- To estimate the additive decomposition
  - Estimate the trend T_t This can be the average of the last m data points.
  - Remove the trend (i.e. calculate the de-trended series)

    Compute $y_t - T_t$ for each data point in the series.
  - Estimate the seasonal component S_t

    Average $(y_t - T_t)$ for each m period. That is, calculate the average for all January values and repeat for all the months.
  - Calculate the noise component $E_t = y_t - T_t - S_t$ for each data point in the series.

Time Series Forecasting using Decomposed Data

Predict

- The idea is to forecast each of the decomposed components $\underbrace{\hat{S}_t \text{ and } \hat{T}_t}_{\text{future time}}$ and put them together for forecasted future time series values.

$$\hat{y}_t = \hat{S}_t + \hat{T}_t$$

- Forecast for the seasonal component:

  It is assumed that the seasonal component of the time series does not change. Hence, it is forecasted as is.

- The seasonality adjusted time series (i.e. the series without the seasonal component) can be forecasted by easier methods: regression, Holt's method or ARIMA.

Smoothing based Methods

Smoothing based Methods for Time Series Forecasting

Notation

- Time periods: $t = 1,2,3,\ldots,n$.

  Time periods can be seconds, days, weeks, months, or years depending on the problem

- Data series: $y_1, y_2, \ldots, y_n$ → true function

  We denote $Y(t)$ as a function which gives the observed value at time t

- Forecasts: F_{n+h} is the forecast for the h-th time period following n. Here h is called as the horizon.

  n: current h: horizon

- Forecast errors: $e_t = y_t - F_t$ for any given time t.

  ⇒ → predicted

Simple Forecasting Methods

n , $n+1$ $n+h$

y_n F_{n+1}

- Naive Method

  Here we simply assume that $F_{n+1} = y_n$

- Seasonal Naive Method

  The forecasted value is the previous value of the same season.

  For example, the next February revenue can be assumed as the last known February revenue.

$$F_{n+1} = y_{n-s}$$

Simple Forecasting Methods

- Average Method

  The next point is assumed as the average of all the data points in the series.

  This model calculates the forecasted value, F_{n+1}, as

$$F_{n+1} = \text{Average}(y_n, y_{n-1}, y_{n-2}, \ldots, y_1)$$

All previous values.

- Moving Average Smoothing

  Select a window of the last k periods to calculate the average, i.e.

  $y_n, y_{n-1}, y_{n-2}, \ldots, y_{n-k+1}$

Simple Forecasting Methods

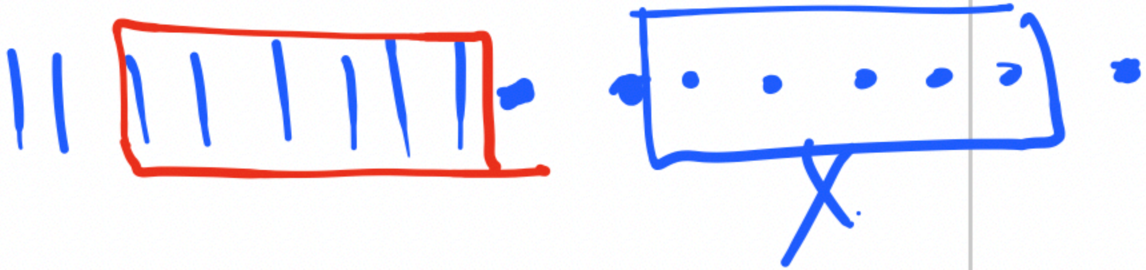
- Weighted Moving Average Smoothing

  For some cases, the most recent value could have more influence than some of the earlier values.

  Here the forecast for the next period is given by the model:

$$F_{n+1} = \frac{(a \times y_n + b \times y_{n-1} + c \times y_{n-2})}{(a + b + c)} \quad \text{for a window size of 3}$$

  where typically $a > b > c$



- These methods can make only one-step-ahead forecasts.

Exponential Smoothing based Forecasting

- Exponential Smoothing

  Exponential smoothing is the weighted average of the past data, with the recent data points given more weight than earlier data points.

- The weights decay exponentially towards the earlier data points

$$F_{n+1} = \alpha y_n + \alpha(1-\alpha)y_{n-1} + \alpha(1-\alpha)^2 y_{n-2} + \dots$$

beginning of the series.

- α is generally between 0 and 1. $\alpha = 1$ returns the naive forecast.
- This forecast equation can be re-written as

Smoothing

$$\boxed{F_{n+1} = \alpha y_n + (1-\alpha)F_n}$$

$\alpha y_n + \alpha[(1-\alpha)y_{n-1} + (1-\alpha)^2 y_{n-2} + \dots]$

$+ \alpha(1-\alpha)[y_{n-1} + (1-\alpha)y_{n-2} + \dots]$

$F_n = \alpha y_n + (1-\alpha)F_{n-1}$

- Generally α is chosen between 0.2 and 0.4

Exponential Smoothing based Forecasting

$F_n = \alpha y_n + (1-\alpha) F_{n-1}$

Disadvantages:

- The simple Exponential Smoothing model is suited only for time series without clear trend or seasonality.
- The forecast does not factor in any seasonality or trend.
- Also, we get the trivial result $F_{n+h} = F_{n+1}$.

  So forecasting over a longer horizon is not possible.

Holt's Two-Parameter Exponential Smoothing

α, β

Decomposition + smoothing

- The forecast uses two parameters α and β
- The forecast is expressed as a sum of the average value or level of the series L_n and the trend T_n

$\hat{L}_n \quad \hat{T}_n$

Level is current reference.

latest

$$F_{n+1} = \underline{L_n} + \underline{T_n}$$

where, level.

$$L_n = \alpha y_n + (1-\alpha) F_n$$

prediction info

exponential

$$T_n = \beta \underbrace{(L_n - L_{n-1})}_{\text{latest information about trend}} + (1-\beta) T_{n-1} \rightarrow \text{prediction about trend}$$

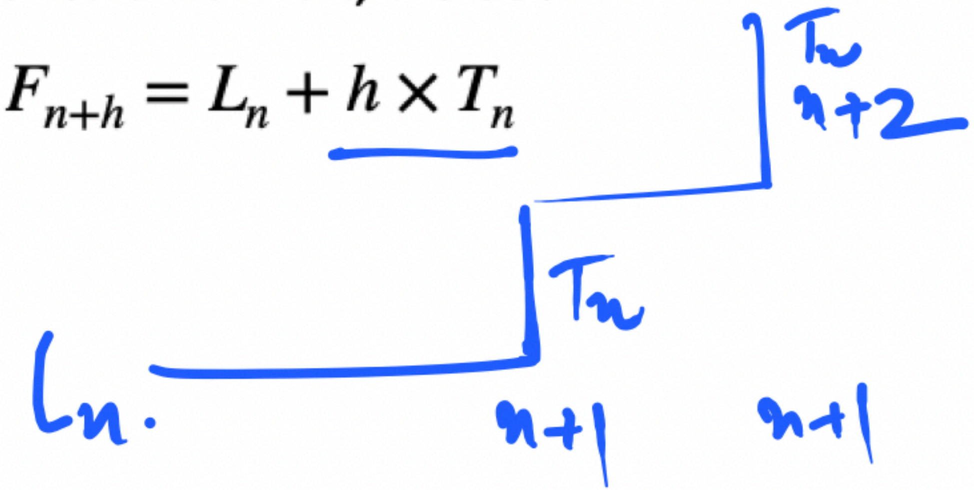
$$F_n = \underline{L_{n-1} + T_{n-1}}$$

$F_n \quad L_n \quad T_n.$

Holt's Two-Parameter Exponential Smoothing

- To make future forecast over a horizon, we use

$$F_{n+h} = L_n + \underline{h \times T_n}$$



Holt-Winter's Three-Parameter Exponential Smoothing

α, β, γ.

- **Additive Model:** To make future forecast over a horizon, we use

$$F_{n+h} = L_n + hT_n + S_{n+h-p}$$

actual info about the level adjusted with seas...

$$\alpha \rightarrow \quad L_n = \alpha(y_n - S_{n-p}) + (1-\alpha)(L_{n-1} + T_{n-1})$$

$$\beta \rightarrow \quad T_n = \beta(L_n - L_{n-1}) + (1-\beta)T_{n-1}$$

$$\gamma. \rightarrow \quad S_n = \gamma(y_n - L_{n-1} - T_{n-1}) + (1-\gamma)S_{n-p}$$

predicted level adjusted for seasonality

where p is the seasonality period.

- The parameters are estimated by fitting the smoothing equation on the training data.

Maximum likelihood estimation

Holt-Winter's Three-Parameter Exponential Smoothing

- **Multiplicative Model:** To make future forecast over a horizon, we use

$$F_{n+h} = (L_n + hT_n) \times S_{n+h-p}$$

$$L_n = \alpha y_n / S_{n-p} + (1-\alpha)(L_{n-1} + T_{n-1})$$

$$T_n = \beta(L_n - L_{n-1}) + (1-\beta)T_{n-1}$$

$$S_n = \gamma \frac{y_n}{L_{n-1} + T_{n-1}} + (1-\gamma)S_{n-p}$$

where p is the seasonality period.

- The parameters are estimated by fitting the smoothing equation on the training data.

Time Series Forecasting

Time Series Forecasting

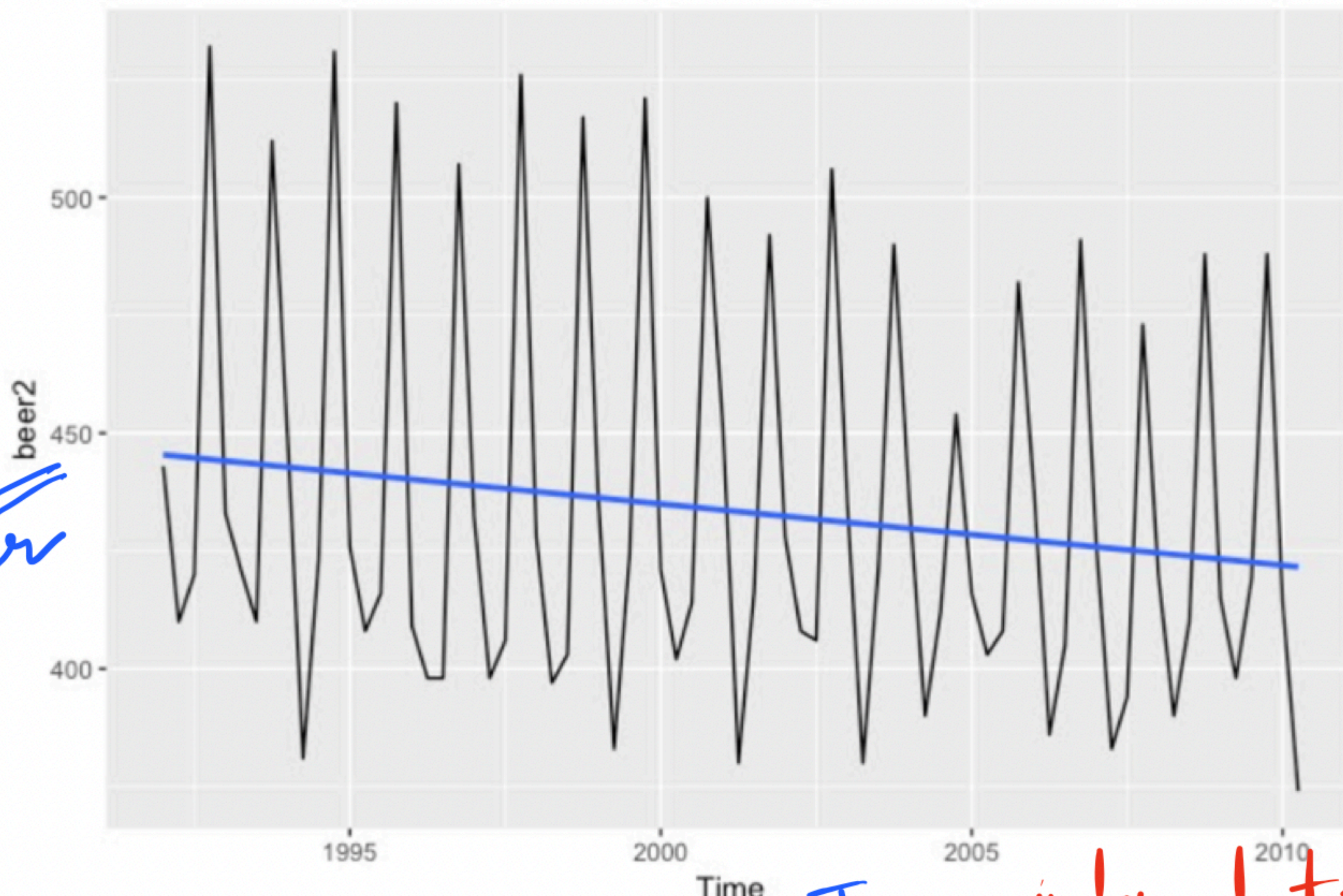
Decomposition based methods	Smoothing based methods	Regression based methods	Machine Learning based methods
Classical	Moving Averages	Regression with seasonality	Regression
STS, X11,..	Exponential Smoothing	ARIMA	Neural Network
		Autocorrelation	

Regression

- Time series value is predicted using time as the independent variable.
- Linear or polynomial functions can be used for regression.
- Forecast can be done for several steps ahead in the horizon

$n + h.$

Forecasting using Linear Regression



Beer

Time independent var

Regression with Seasonality

$Q_1 \quad Q_2 \quad Q_3 \quad Q_4$

- Seasonal dummy variables are introduced for each season
  - That is, each season triggers a 1 or 0 attribute value
  - Attribute of quarter Q1 is turned 1 if the quarter is Q1

$$\hat{y}_{n+1} = 44.189 - 1.132n - 27.268(Q1 = 1) - 16.363(Q2 = 1) + 91.99(Q3 = 1) + 23.131(Q4 = 1)$$

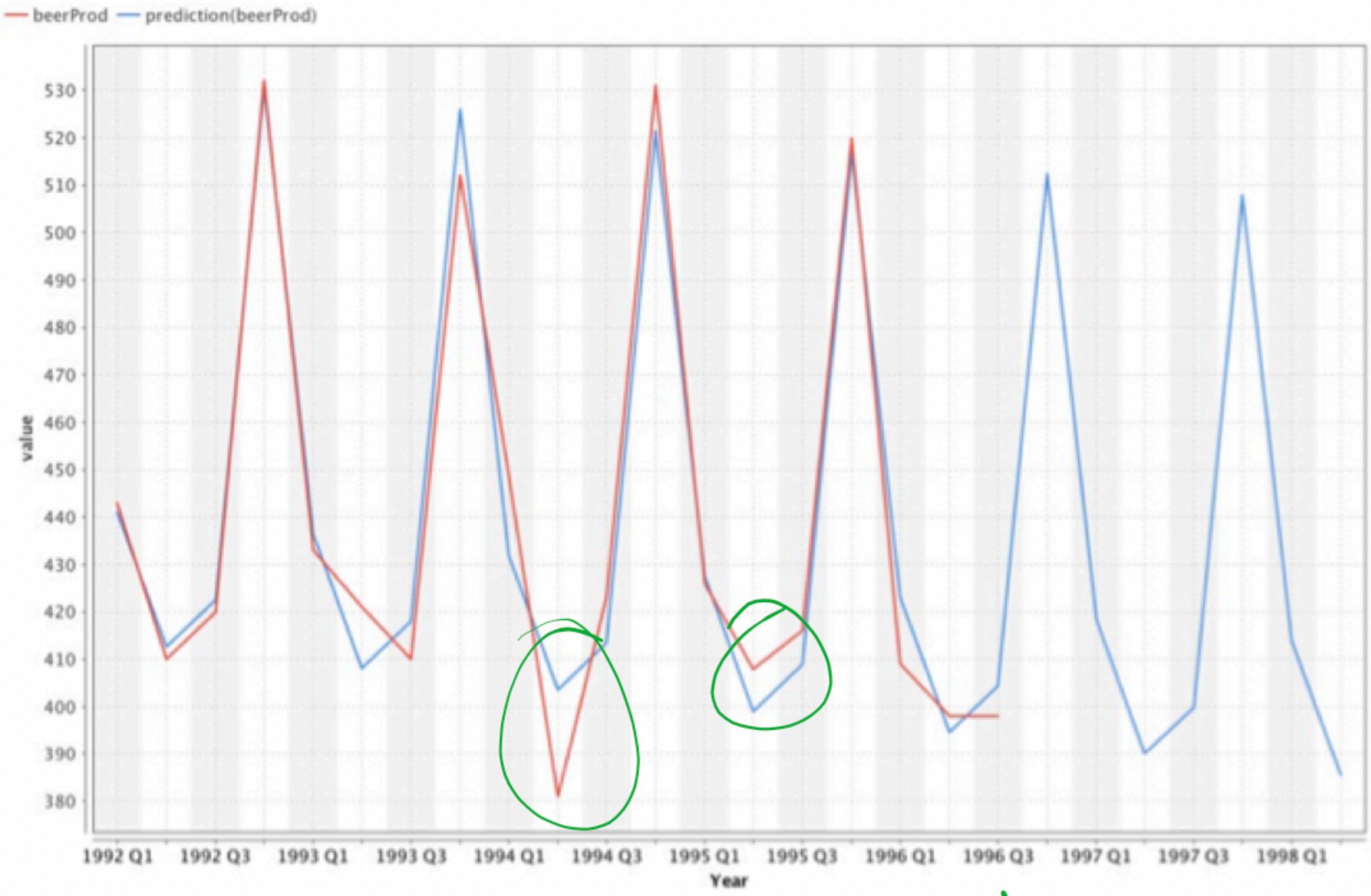
Seasonal Attributs

Dummy variables.

Year	time	Quarter = Q1	Quarter = Q2	Quarter = Q3	Quarter = Q4	beerProd
1992 Q1	1	1	0	0	0	443
1992 Q2	2	0	1	0	0	410
1992 Q3	3	0	0	1	0	420
1992 Q4	4	0	0	0	1	532
1993 Q1	5	1	0	0	0	433
1993 Q2	6	0	1	0	0	421
1993 Q3	7	0	0	1	0	410
1993 Q4	8	0	0	0	1	512
1994 Q1	9	1	0	0	0	449
1994 Q2	10	0	1	0	0	381

Forecasting with seasonal linear regression

red: actual
blue: predicted



horizon.

Autocorrelation

time is not a behavioural attribute

- Autocorrelations represent the correlations between adjacently located timestamps in a series.

- Typically, the behavioural attribute values at adjacently located timestamps are positively correlated.

- Autocorrelations in a time series are defined with respect to a particular value of the lag L

To capture seasonality.

component.

Lag Series

seasonal component

periodicity

original Lag=1 Lag=2 3 4 5 6

Year	prod	prod-1	prod-2	prod-3	prod-4	prod-5	prod-6
1992 Q1	443	?	?	?	?	?	?
1992 Q2	410	443	?	?	?	?	?
1992 Q3	420	410	443	?	?	?	?
1992 Q4	532	420	410	443	?	?	?
1993 Q1	433	532	420	410	443	?	?
1993 Q2	421	433	532	420	410	443	?
1993 Q3	410	421	433	532	420	410	443
1993 Q4	512	410	421	433	532	420	410
1994 Q1	449	512	410	421	433	532	420
1994 Q2	381	449	512	410	421	433	532
1994 Q3	423	381	449	512	410	421	433
1994 Q4	531	423	381	449	512	410	421
1995 Q1	426	531	423	381	449	512	410

Q1

Q1