TIME SERIES FOR DATA SCIENCE

DR. ALINE QUADROS

DSR OCTOBER 2020

SEPTEMBER 2020

Contents

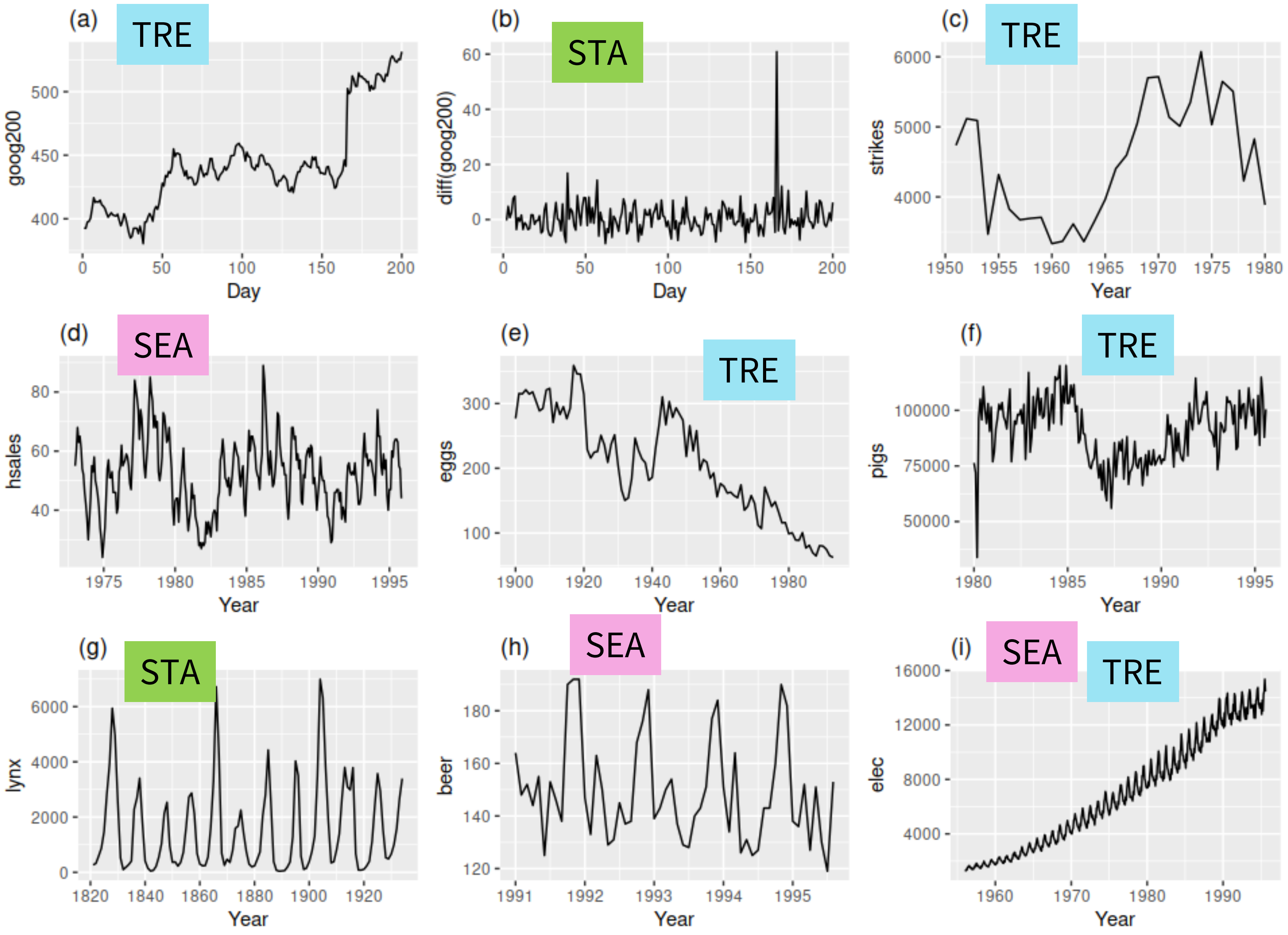
- What are time series?
- What kind of time series properties are there?
  - Trend, seasonality
  - Stationarity
  - Autocorrelation and Partial-autocorrelation
- Baselines
  - average, naïve, smoothing moving Average,
- Traditional Time Series Analysis:
  - Exponential smoothing (a.k.a ETS models)
  - Autoregressive models: ARIMA family (ARMA, ARIMA, SARIMA, SARIMAX)
- ML-based time-series:
  - Random forests and boosting trees
  - RNNs and LSTMs (in the Sequences class)

Resources

- Free excellent book
  - Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. www.OTexts.com/fpp2
  - *Most of the plots used in this presentation come from this book*
- Bastian Kubsch’s repository: https://github.com/bkubsch/time_series
- Online class = https://online.stat.psu.edu/stat510/lesson/1

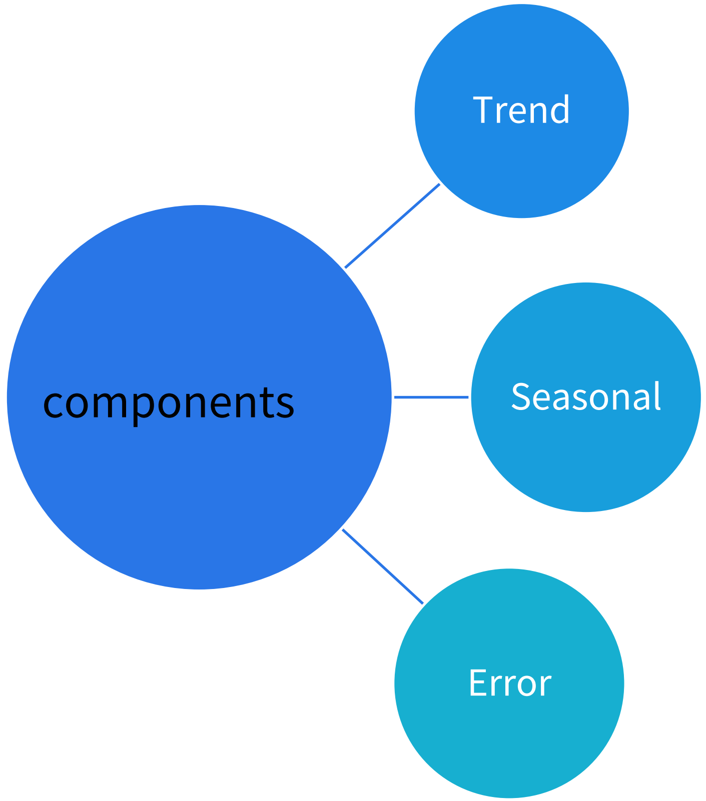
PROPERTIES

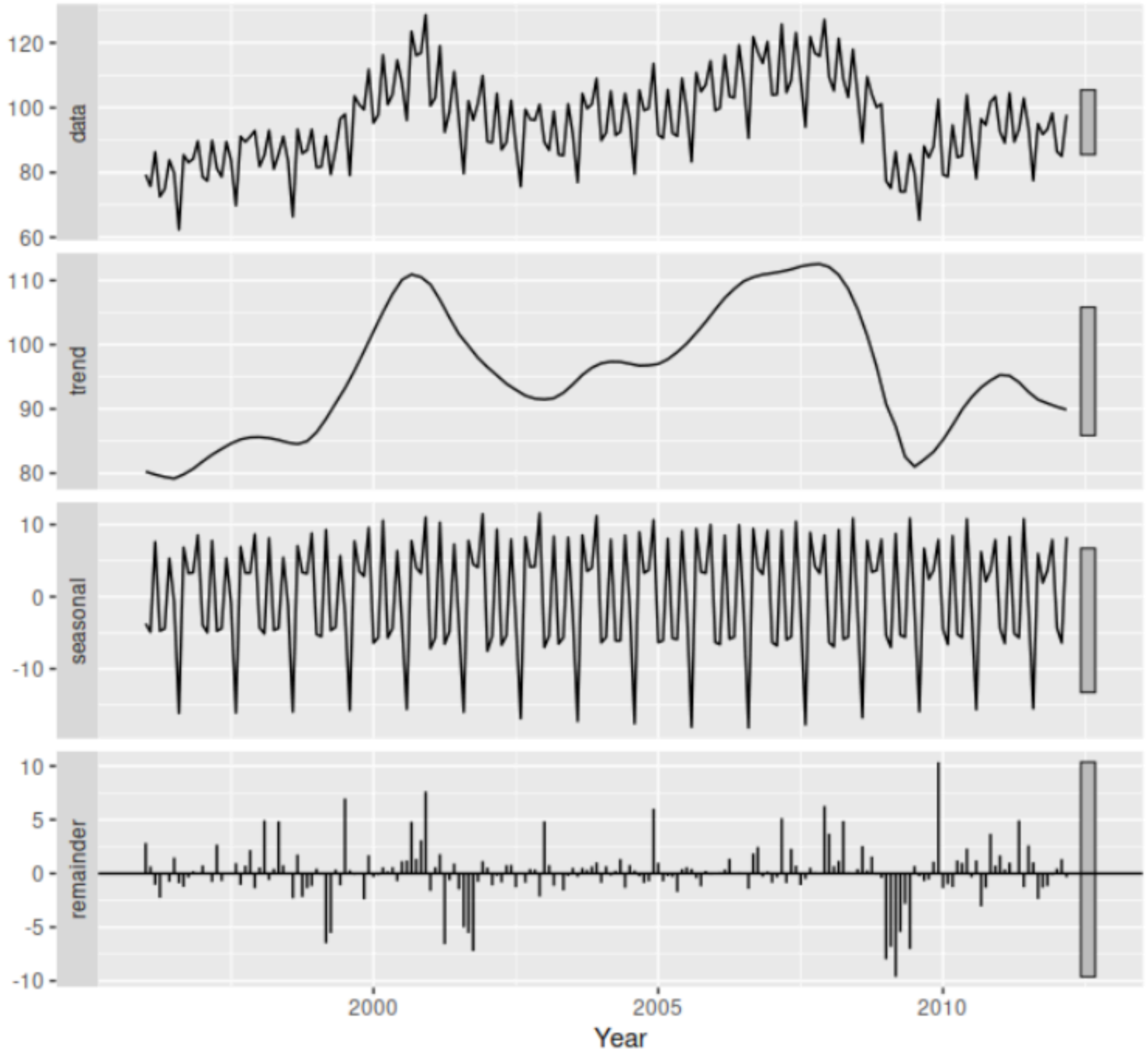
SEPTEMBER 2020



Trend
Seasonality
Stationarity

Time series decomposition





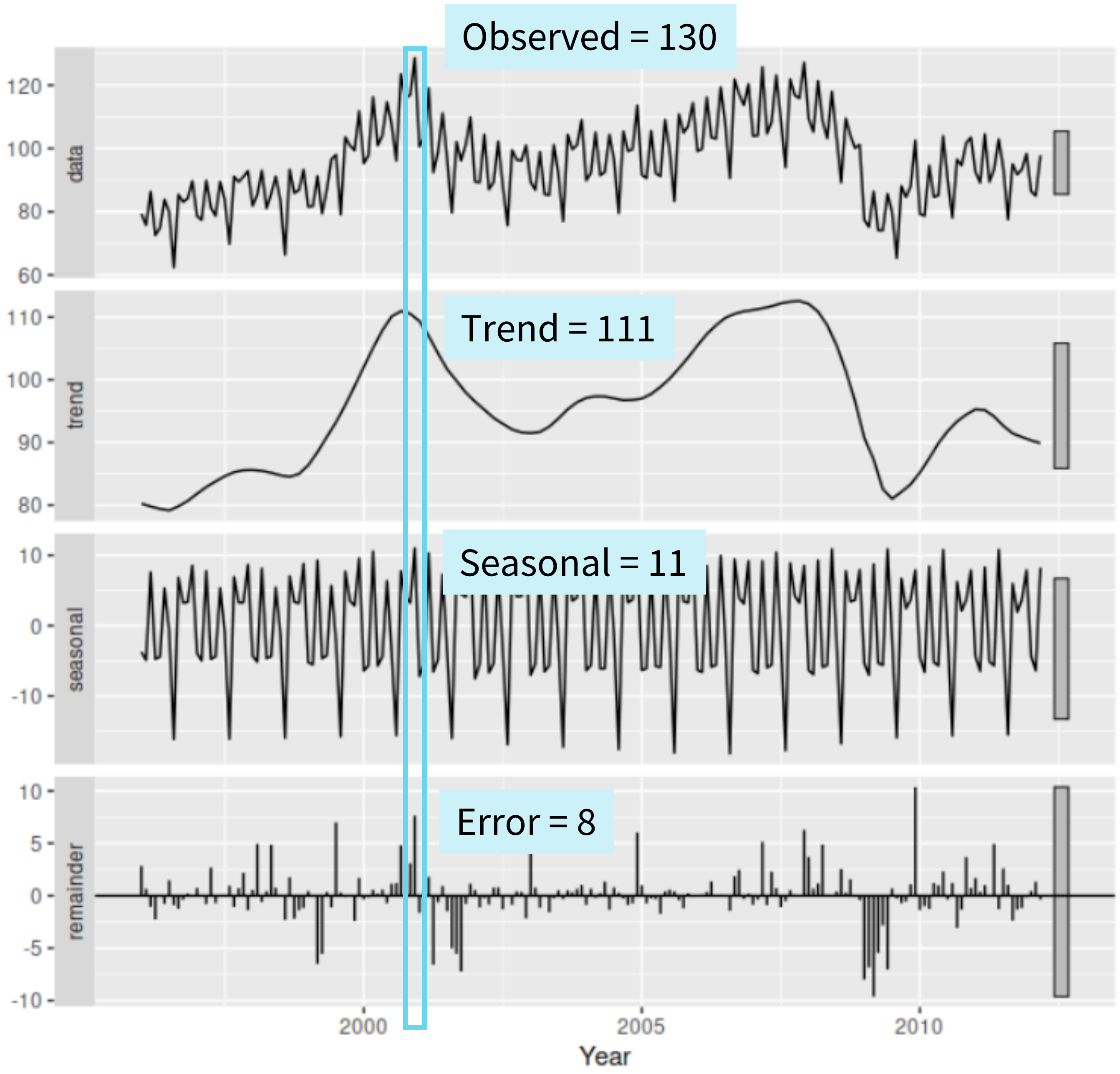
Additive = the components add to each other
Multiplicative= the effects are multiplied

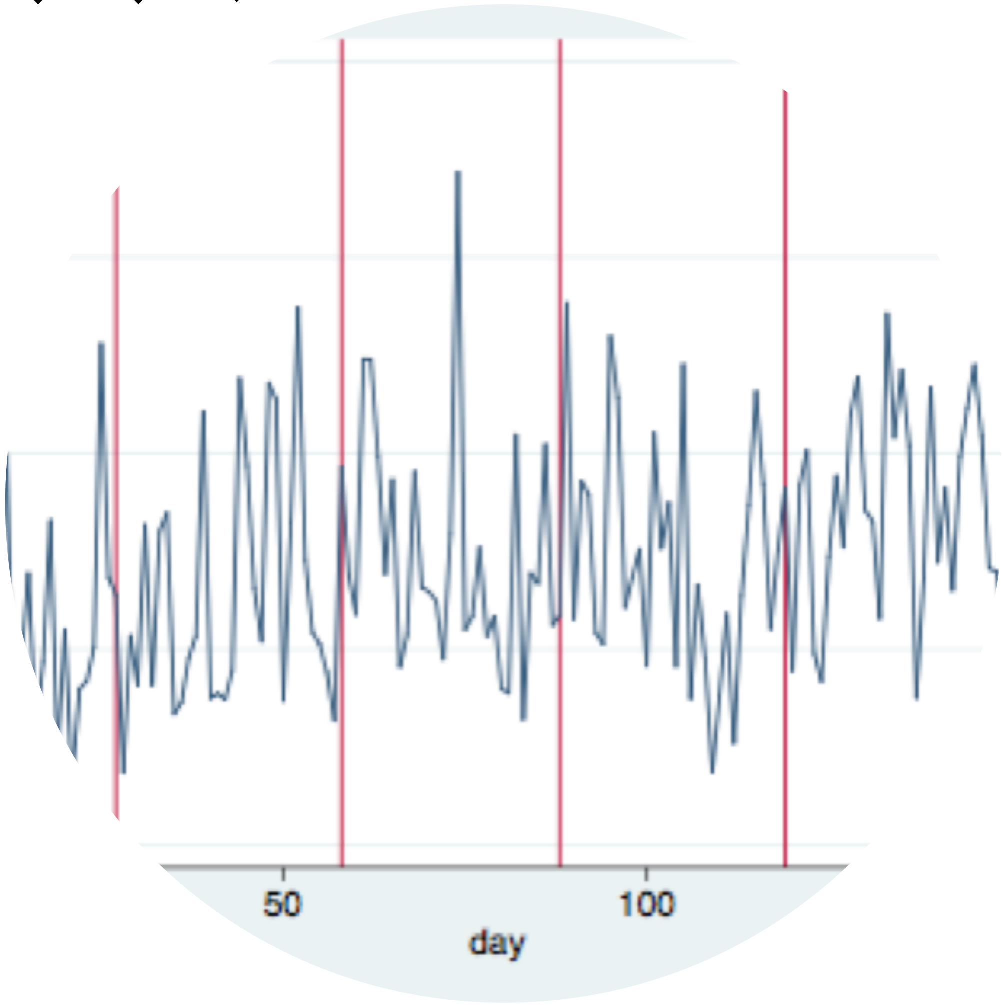
Time series decomposition

Example of additive decomposition

Look at the value at time 2001:

The observed value (130) is decomposed as
130 = 111 + 11 + 8





Stationarity

- **Even though the data has a time axis, or was observed over time, it’s values do not depend on the time component itself**
- In other words, stationary data is data without trends or seasonality

BASELINES

SEPTEMBER 2020

Simplest possible forecasts

Simplest representation of trend, seasonality and autocorrelation

In practice, they will be used as a reference to compare our candidate models

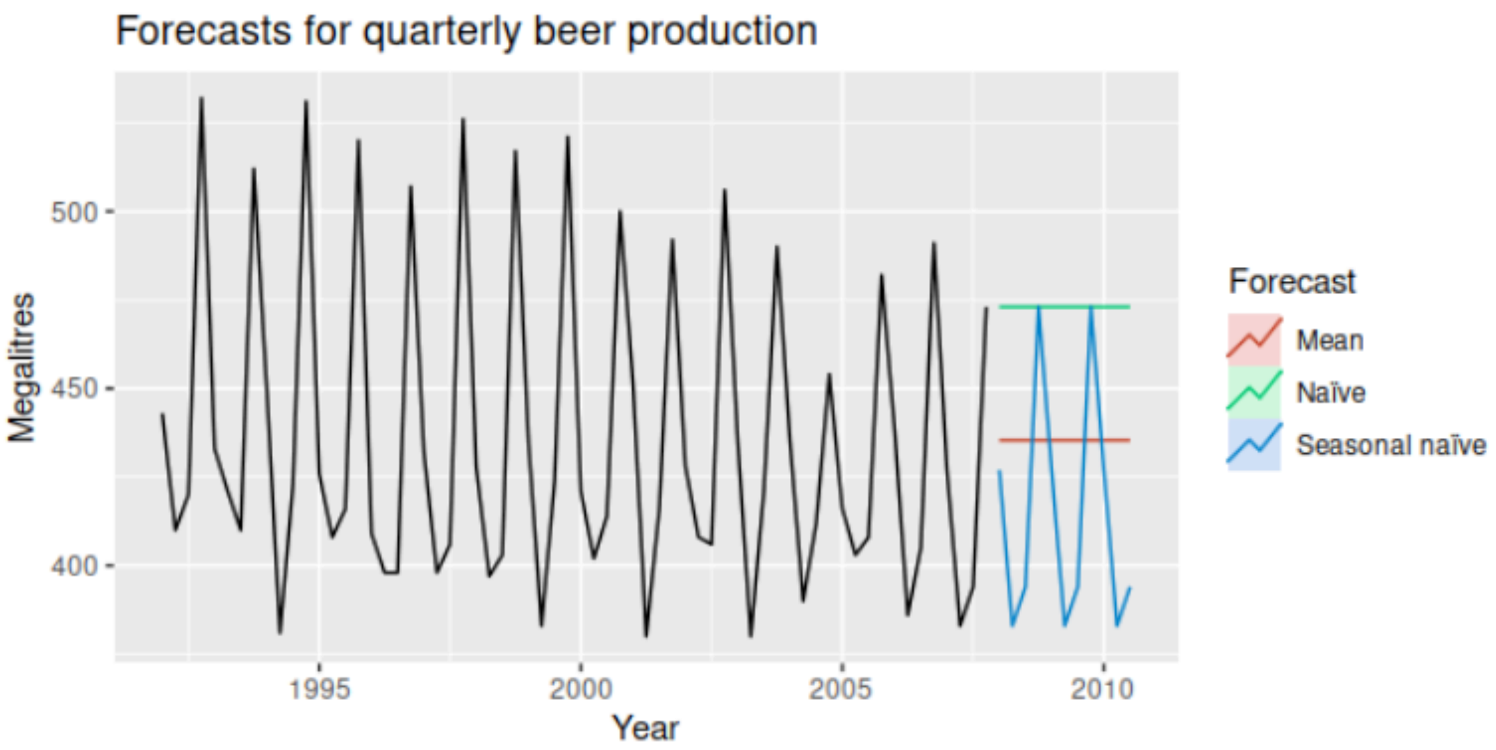


Figure 3.1: Forecasts of Australian quarterly beer production.

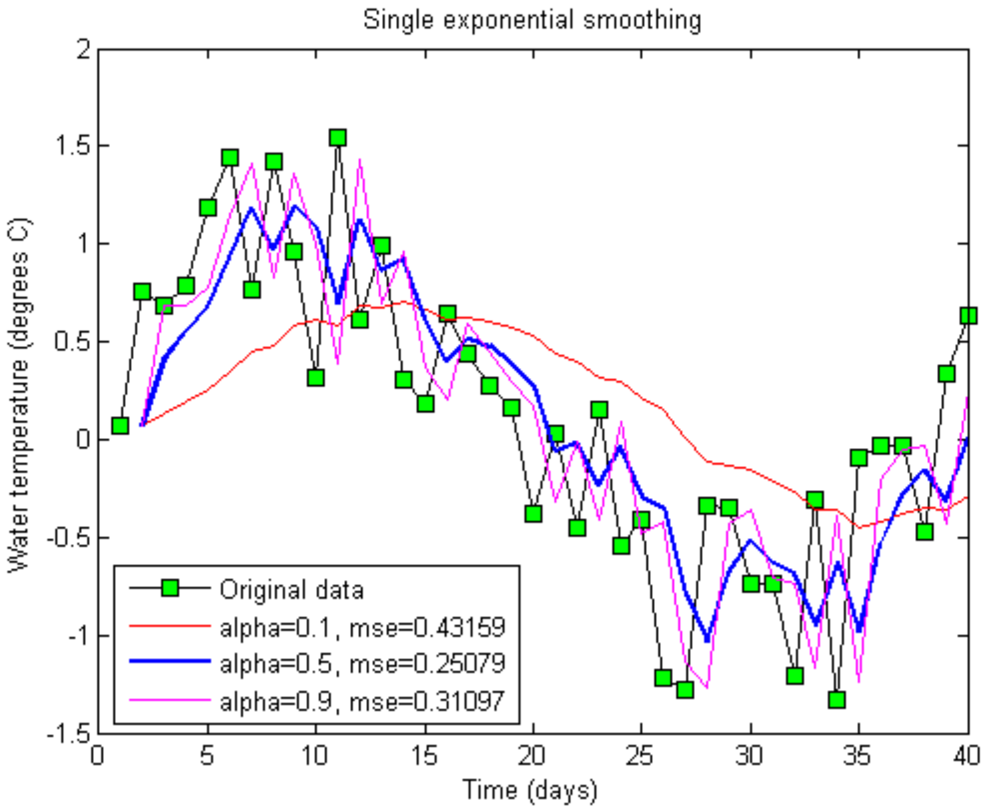
EXPONENTIAL SMOOTHING

SEPTEMBER 2020

Exponential Smoothing

Exponential smoothing methods are weighted averages of past observations,

The weights decaying exponentially as the observations get older, and the degree is given by parameter alpha.



Exponential Smoothing

The way the weights are distributed along time is determined by the smoothing parameter alpha (α)

Small α	Shifts weight toward the past
Large α	Shifts the weights toward the most recent observation
If α =1?	Same as naive fcst

Alpha

The way the weights are distributed along time is determined by the smoothing parameter alpha (α)

	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$
y_T	0.2000	0.4000	0.6000	0.8000
y_{T-1}	0.1600	0.2400	0.2400	0.1600
y_{T-2}	0.1280	0.1440	0.0960	0.0320
y_{T-3}	0.1024	0.0864	0.0384	0.0064
y_{T-4}	0.0819	0.0518	0.0154	0.0013
y_{T-5}	0.0655	0.0311	0.0061	0.0003

Methods

- Simple Exponential Smoothing
  - For time series with no clear trend or seasonality
- Holt-Winter’s family of methods
  - For time series with trend and/or seasonality

Time-series scenarios (and appropriate models)

		SEASONALITY		
		None	Constant (Additive)	Increasing (Multiplicative)
TREND	None	(N,N) **Simple exponential smoothing**	(N, A)	(N, M)
	Linear (Additive)	(A, N) **Holt's linear method**	(A, A) **Additive Holt-Winter's**	(A, M) **Multiplicative Holt-Winter's**
	Exponential (Additive damped)	(Ad, N) **Additive damped trend method**	(Ad, A)	(Ad, M) **Holt-Winter's damped**

AUTO REGRESSIVE MODELS

SEPTEMBER 2020

Autoregressive models

- Classic and robust methods. Can be applied in many different situations
- ARIMA = Autoregressive Integrated Moving Average
- Exponential smoothing models are based on trend and seasonality in the data
- ARIMA models describe the autocorrelations in the data

The notation **(p, q, d)**

ARIMA

AUTOREGRESSIVE component
a.k.a. **p**

Linear combination of past values; Regression against itself

DIFFERENCING component
a.k.a. **d**

Helps to stabilise the mean, reducing trend and seasonality

MOVING AVERAGE component
a.k.a. **q**

Weighted moving average of past errors

VARIATIONS:

NON-SEASONAL ARIMA
ARIMA(p, d, q)

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

SARIMAX
+ features

Understanding (p, d, q) and m

- p = is the AR term. Is the number of lags of Y to use as predictors
- q = is the MA term. Is the number of lags to use to get the forecast errors
- d determines how many periods to lag before calculating the differences
  - For an Array[10, 4, 2, 9, 34] => d =1 results in [-6, -2, 7, 25]
  - If the data is already stationary, d = 0
- The m indicates the number of observations per seasonal cycle, will be used in SARIMAX and SARIMA.

WHAT'S NEXT?

ML MODELS

SEPTEMBER 2020

ML-based time series forecasting

- Tree models (Random forests, boosted trees)
- LSTMs and RNN (next classes)
- AWS SAGEMAKER DEEPAR
  - Based on RNN
  - https://docs.aws.amazon.com/sagemaker/latest/dg/deepar.html

ML and TIME SERIES basics

- Never shuffle the data: order matters!

- Always create a baseline model (naïve, average, moving average) to have a clearer understanding of how well your candidate models are performing
  - They should at least be better than the average

- Use as validation and test sets a period as long as the one you are being asked to predict:
  - Example: If you need to predict sales of next 3 months, use the last 3 months of data as test set, and the 3 months before that for cross-validation (for gridsearch, for example)