Predictive Analytics – Session 2

Time Series Prediction – Considerations and Methods

Associate Professor Ole Maneesoonthorn

Associate Professor in Econometrics and Statistics

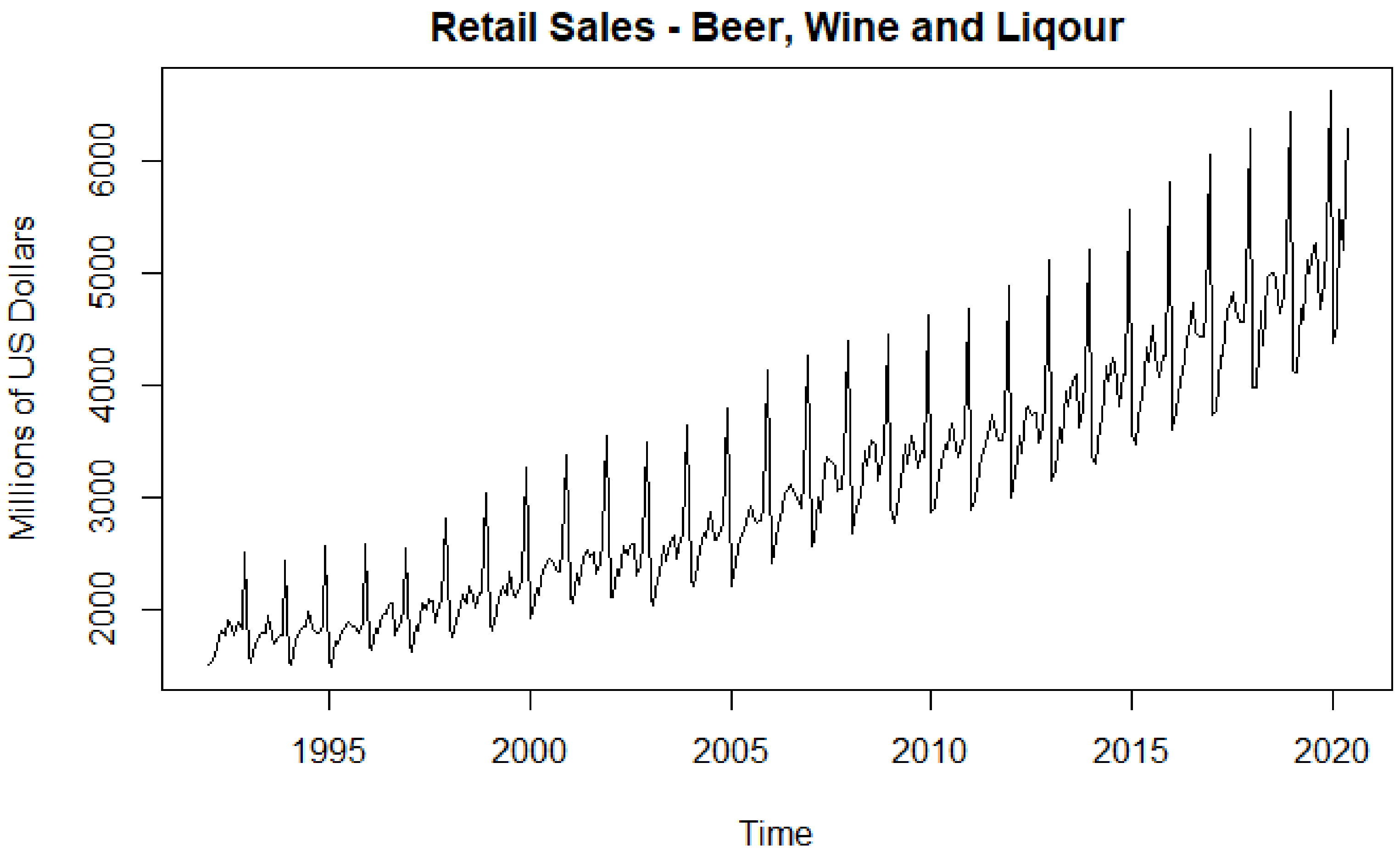
Melbourne Business School

O.Maneesoonthorn@mbs.edu

What is time series data?

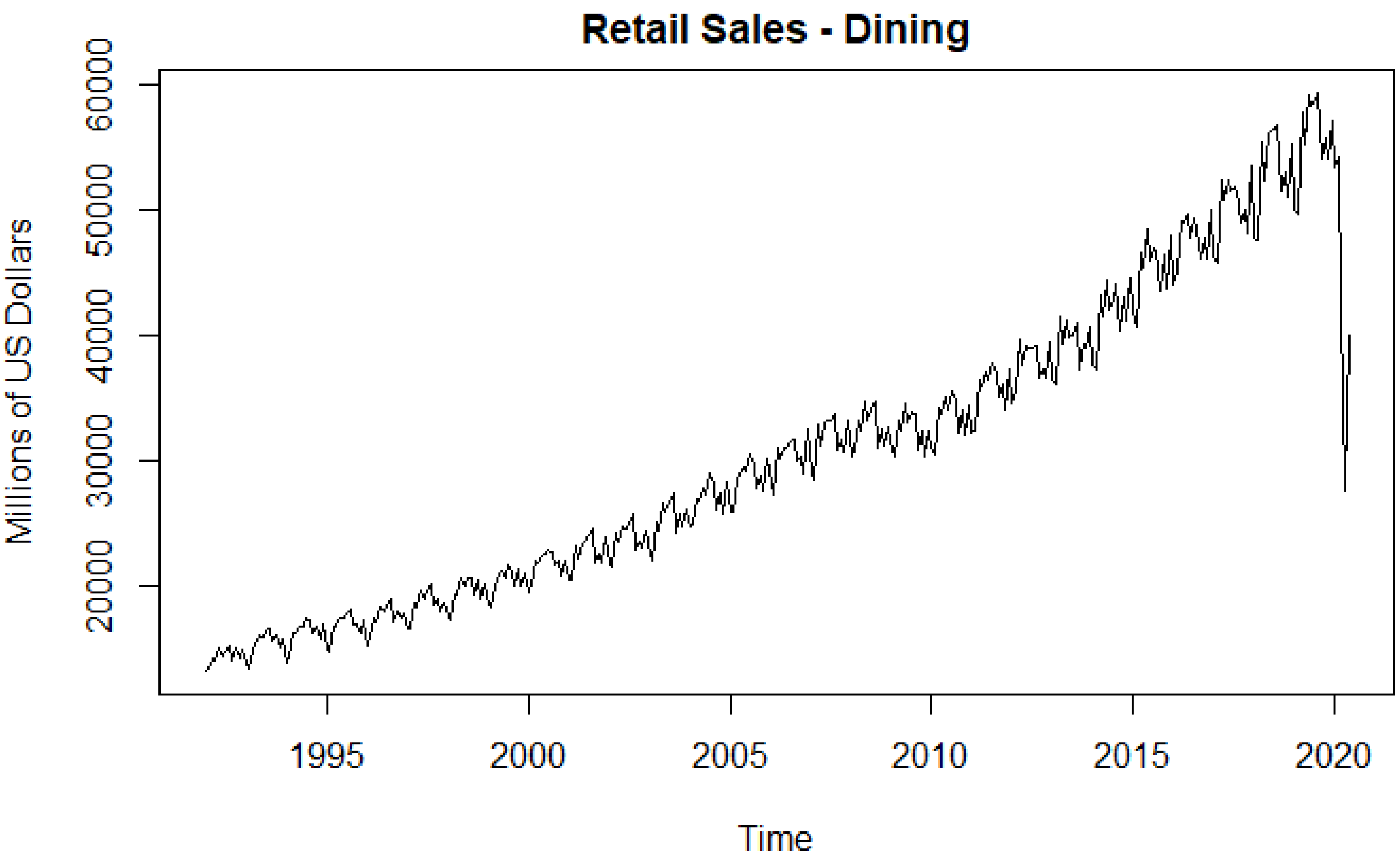
Time Series Characteristics

- Examples:



Time Series Characteristics

- Examples:



Time Series Characteristics

The main components:

1. **Trend** – long term general direction
2. **Cyclical** – shorter term movements that last more than one whole season
3. **Seasonal** – frequent and repeated patterns at regular frequency
4. **Error** – random noise

Time Series Characteristics

The time series components:

- **Trend and cyclical** components are typically not separable
- **Seasonal** components can be extracted using various methods
- Key question: Are you interested in the general trend only? Or are you interested in the whole time series, seasonal and all?

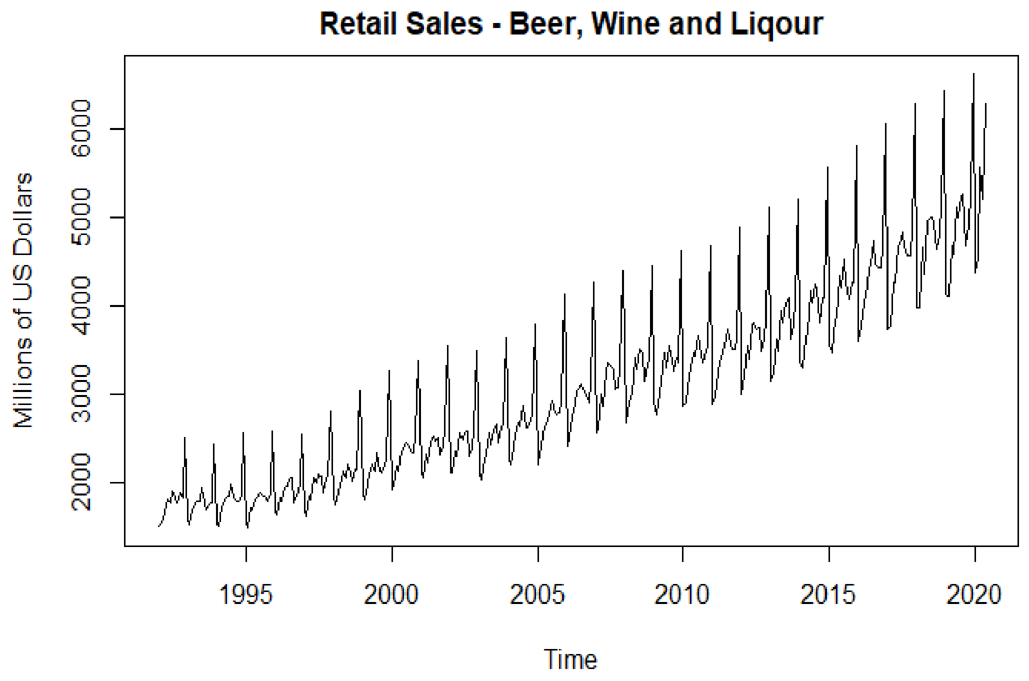
Considerations in Time Series Prediction

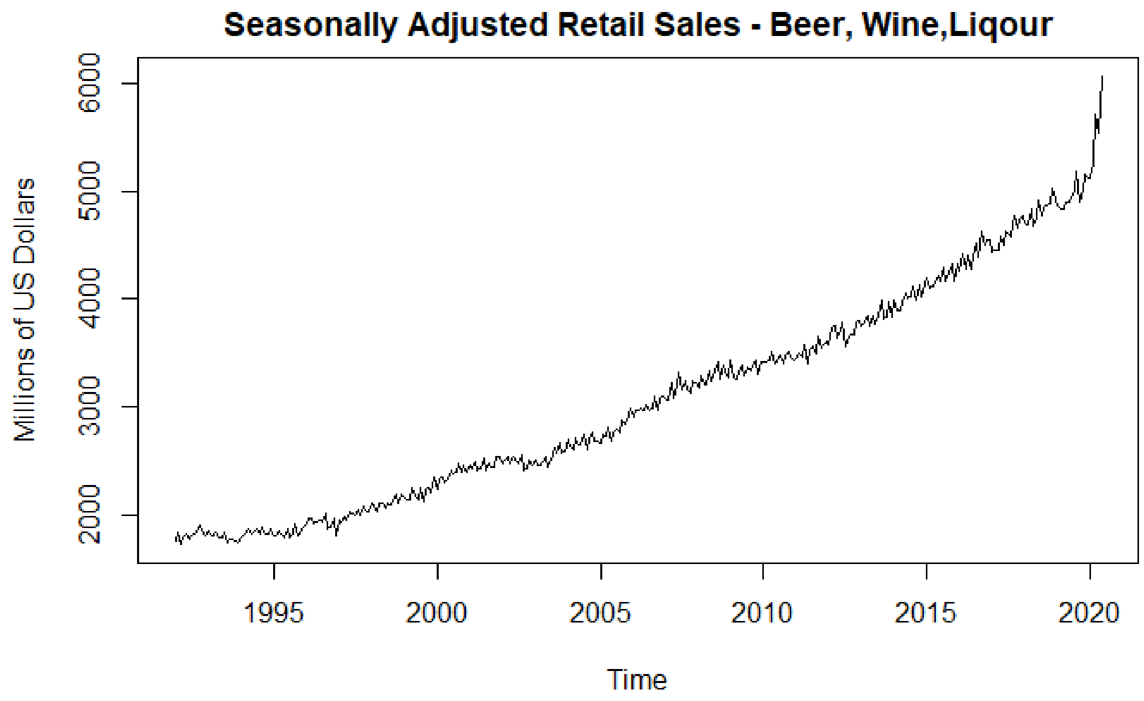
What components to focus on?

- Interested in the general trend only?
    - → **Remove** the seasonal component – **seasonally adjusted data**
    - More common when data is used for other purposes, e.g. as a regressor in another model
    - Removing seasonality can reveal a **better picture of trend/cyclical components**
    - → Decomposition of the time series data

Considerations in Time Series Prediction

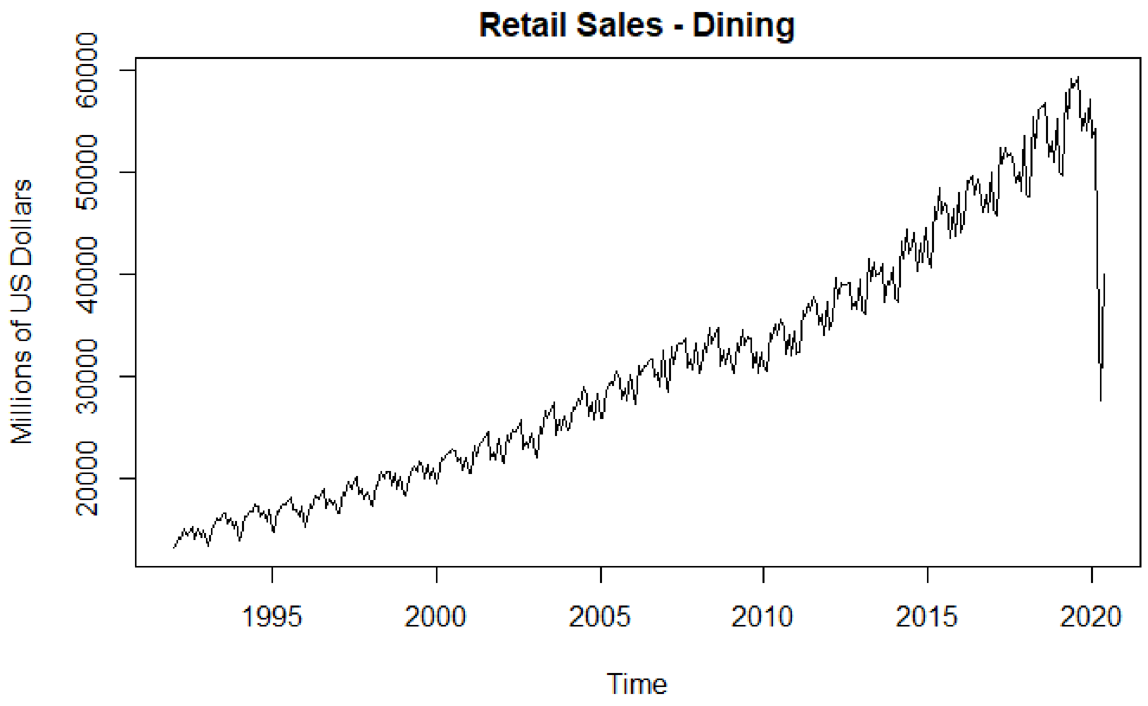
- Examples: **Seasonal adjustments**

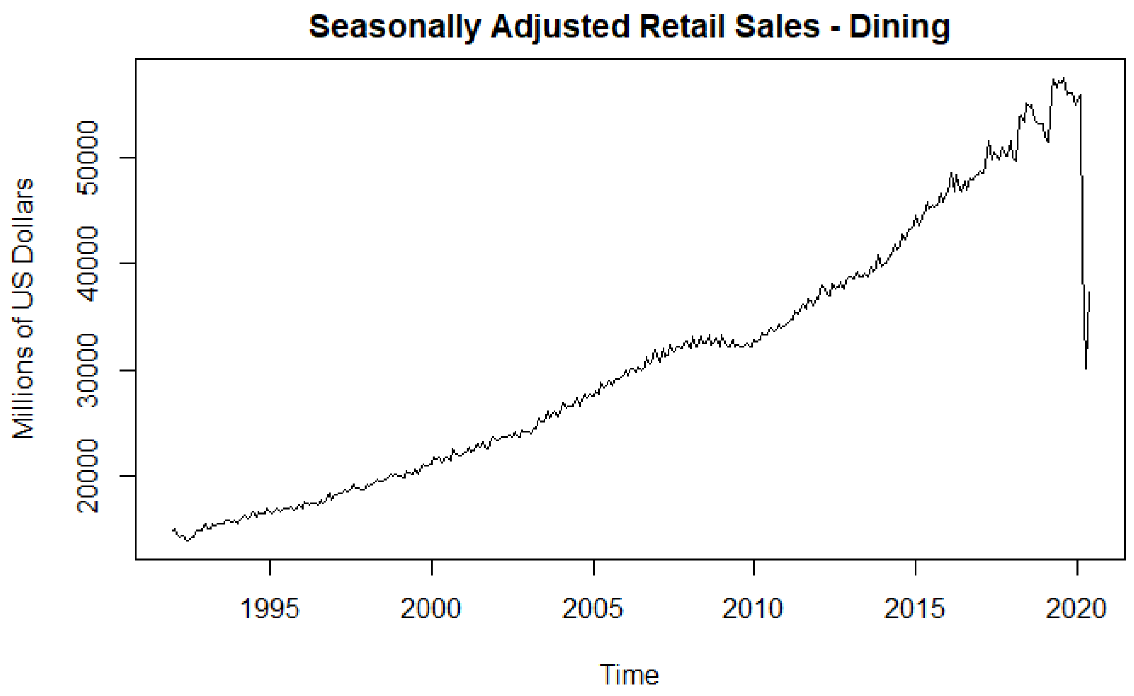




Considerations in Time Series Prediction

- Examples: **Seasonal adjustments**





Considerations in Time Series Prediction

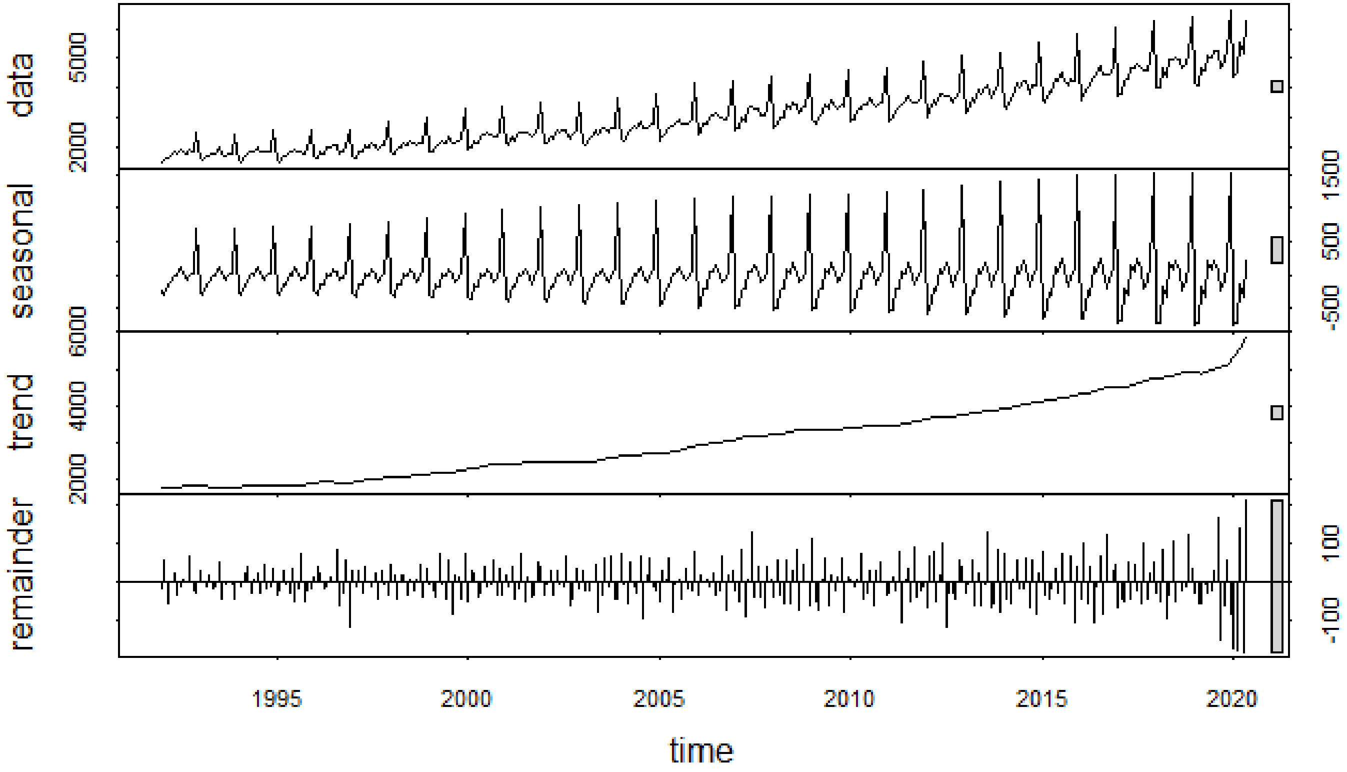
What components to focus on?

- Interested in the whole time series, seasonal and all?
    - → Predict the **raw data**
    - More common in business context
    - E.g. inventory planning, budgeting, rostering

Considerations in Time Series Prediction

What else can decomposition tell us?

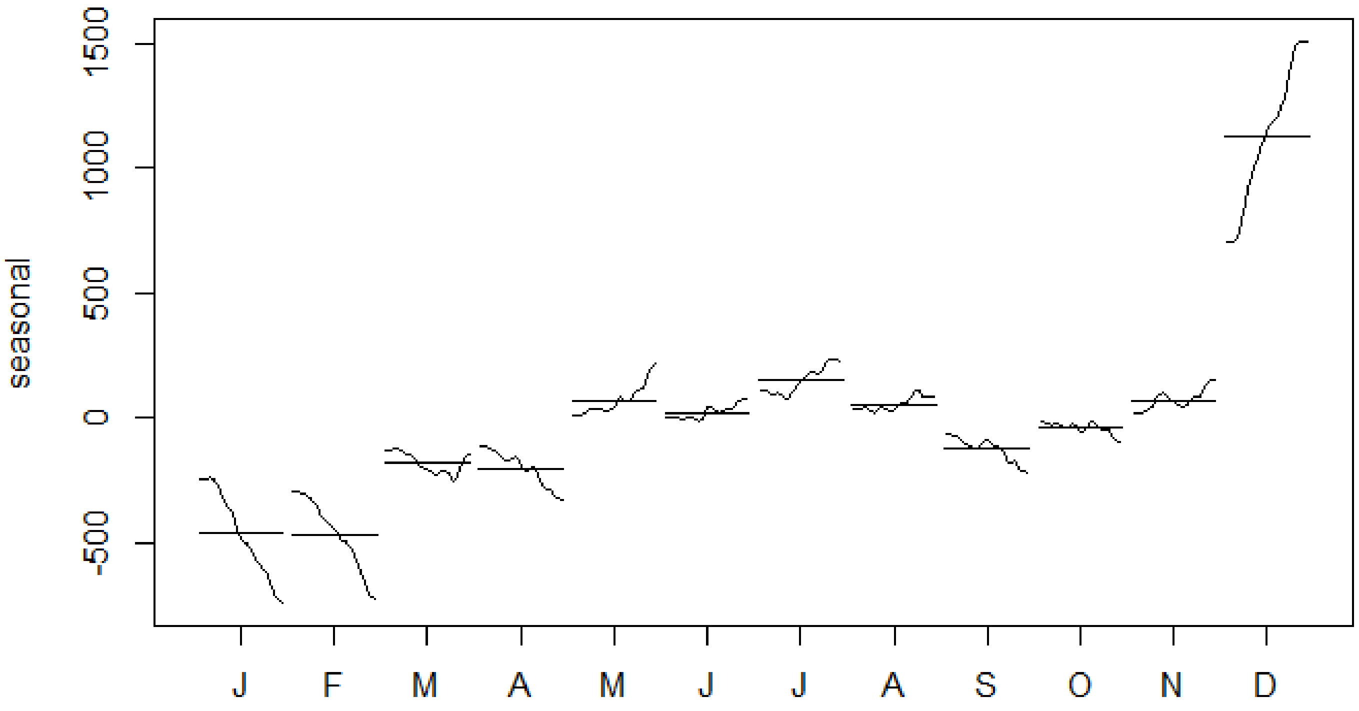
- Are there **changes in seasonality**?



Considerations in Time Series Prediction

What else can decomposition tell us?

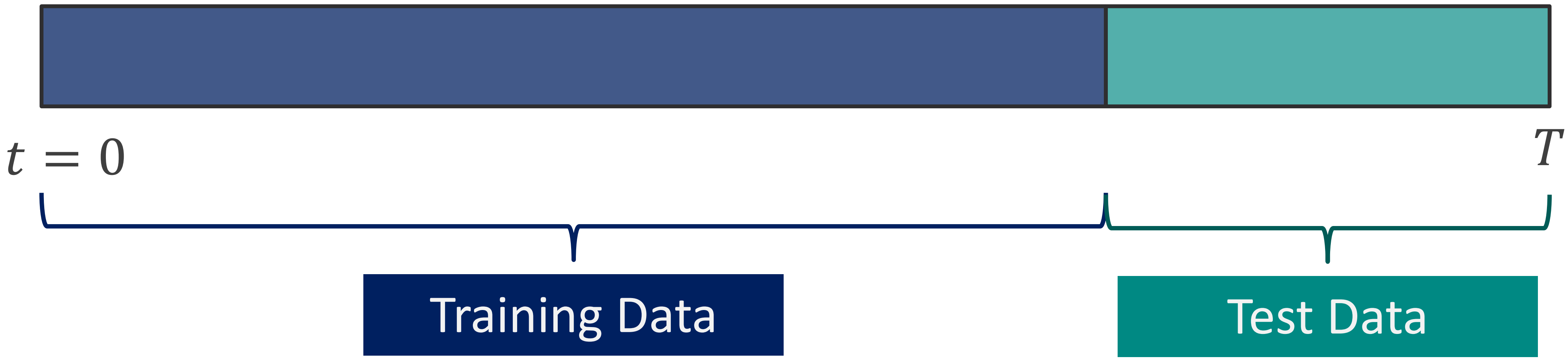
- Are there any trend in **seasonal movements**?



Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

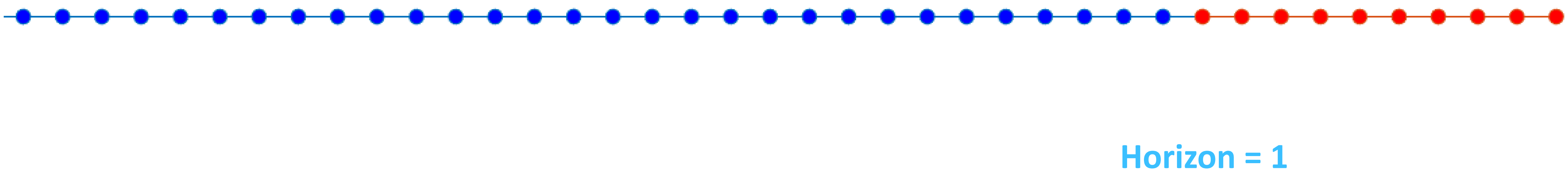
- Recall the predictive assessment in time series:



Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

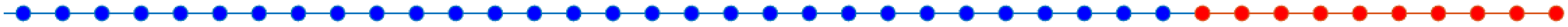
- Are we going to do **long-horizon** predictions?



Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

- Are we going to do **long-horizon** predictions?

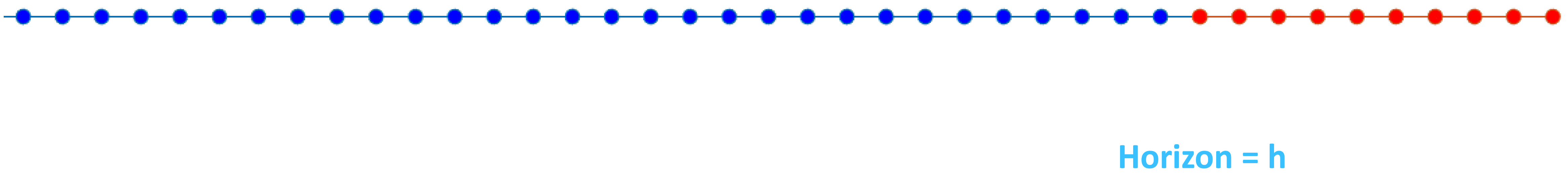


Horizon = 2

Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

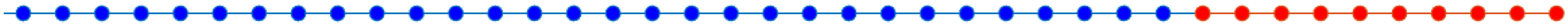
- Are we going to do **long-horizon** predictions?



Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

- Or are we going to do **real-time** predictions?



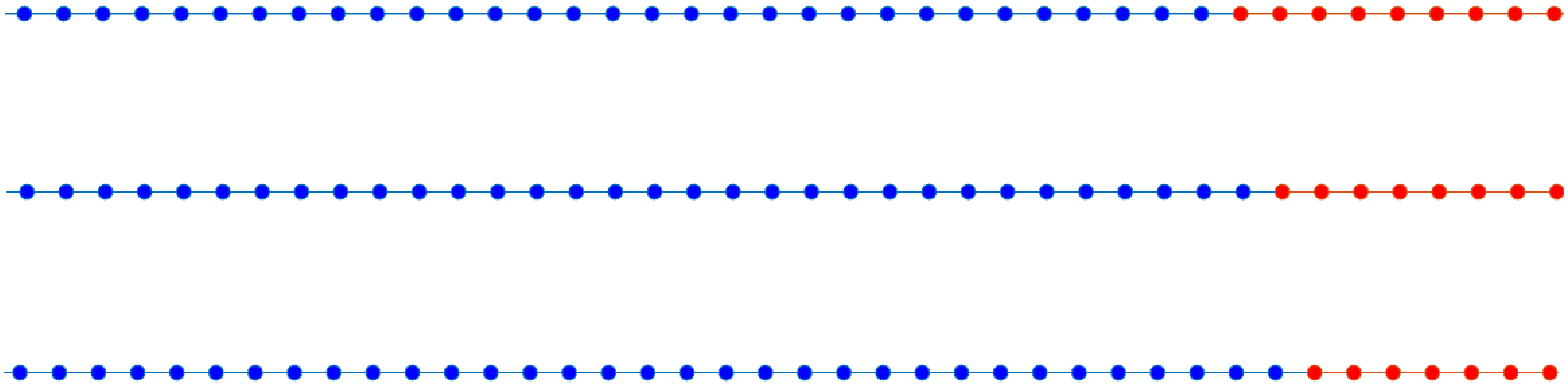
Horizon = 1

Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

- Or are we going to do **real-time** predictions?

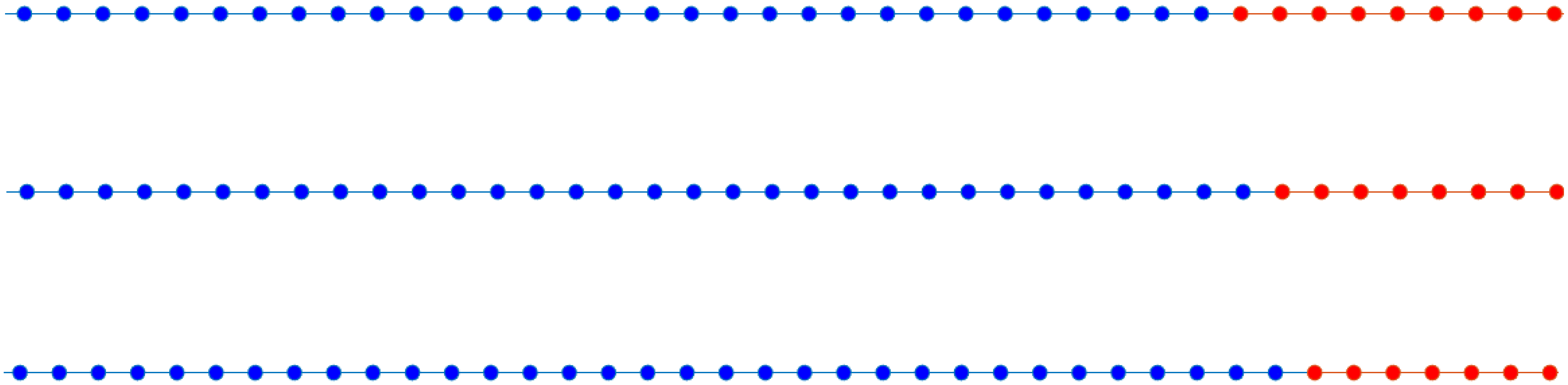
Horizon = 1 ALWAYS!



Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

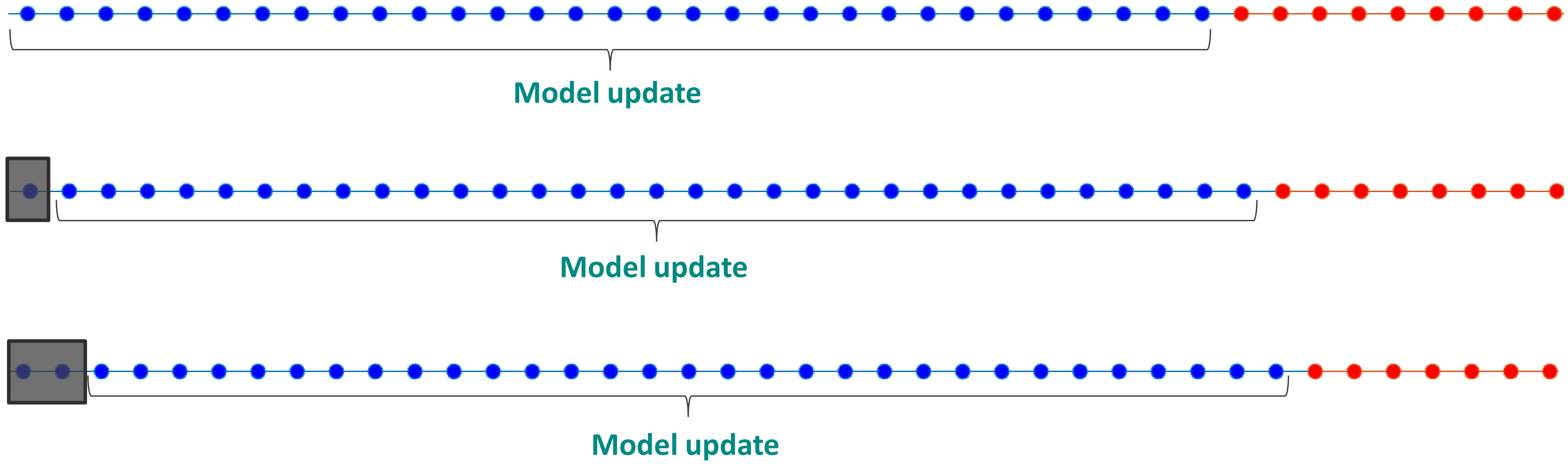
- Should we do **fixed window** or **expanding window** updates?



Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

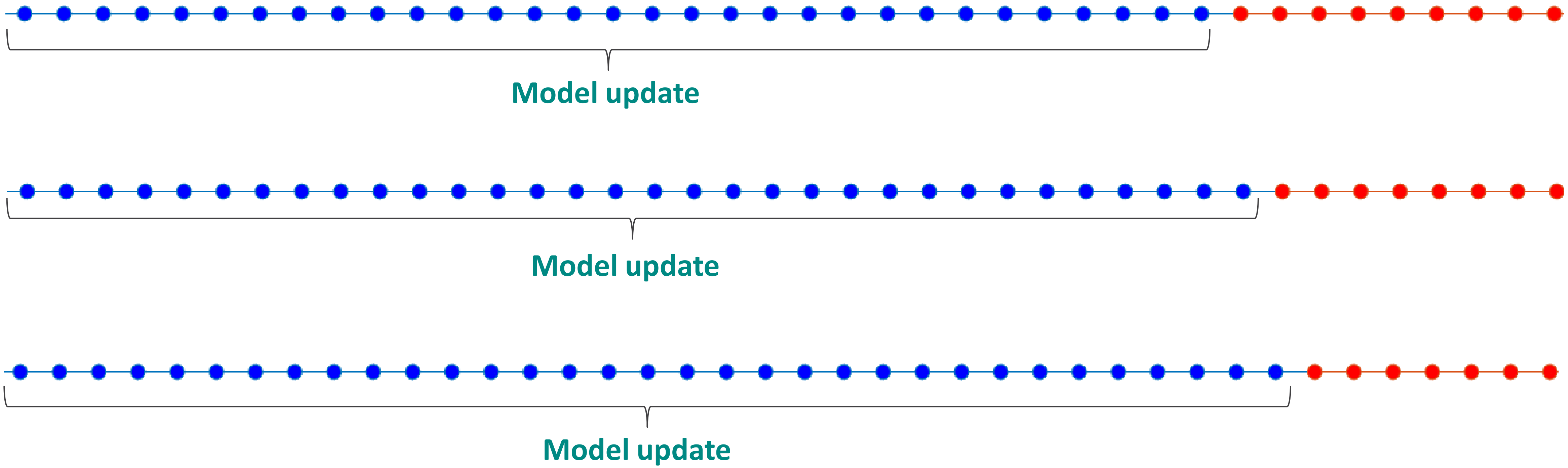
- **Fixed window** updates



Considerations in Time Series Prediction

What prediction scheme is suitable for my context?

- **Expanding window** updates



Considerations in Time Series Prediction

Metrics for time series predictive performance

- RMSE, MAE, MAPE used in the usual way
- **MASE** – benchmark naïve is now different
    - **Time series**: information becomes available as we move on in time
    - Naïve forecast: the **most recently observed value**
    - Interpretation in the usual way with benchmark MASE value of 1 indicating “as good as” naïve

Considerations in Time Series Prediction

Metrics for time series predictive performance

- **Additional metric**: Are the forecast errors truly random?
- That is, if we know what the forecast error is today, does that inform us what the forecast error is likely to be tomorrow?
- Measured by the correlation between the errors observed at two different time points
- Known as the **autocorrelation**, or ACF for short

$$ACF = Corr(e_t, e_{t-1})$$

- ACF should be **as close to zero as possible**

Considerations in Time Series Prediction

Metrics for time series predictive performance

- Again, think about what the predictive outcome is being used for
- **User-specific loss** can also be incorporated into the predictive assessment

Time Series Methods

Time series regression

- All variables now have time index t to indicate time
- Time series regression account for **trend** and **seasonality** by

$$y_t = b_0 + b_1 t + b_2 S_{2t} + \cdots + b_m S_{mt} + e_t$$

  - The variable t is a count of time periods → b_1 is average growth per time period
  - $S_{2t}, S_{3t}, \ldots, S_{mt}$ are dummy variables indicating which season the time period belongs to
  - $b_2, b_3, \ldots, b_m$ are estimates of seasonality relative to the first season
  - (Dummy variable: leave one category out)

Time Series Methods

Time series regression - variants

- Other variants of the time series regression may improve the predictive accuracy
- Nonlinear trends – quadratic, polynomial, spline
- **Log transform** of the variable of interest
    - Particularly useful for business variables with proportion growth context (e.g. sales)
    - Variable must be positive for this transformation
- Inclusion of other explanatory variables

Time Series Methods

Exponential Smoothing

- Best prediction = **running average** of the past values
- Use **all observations** from the past…
- But give **more weight to most recent** observations
- Weight function decays exponentially as we move into the historical past
- Based on time series decomposition
    - Trend + Seasonal + Error

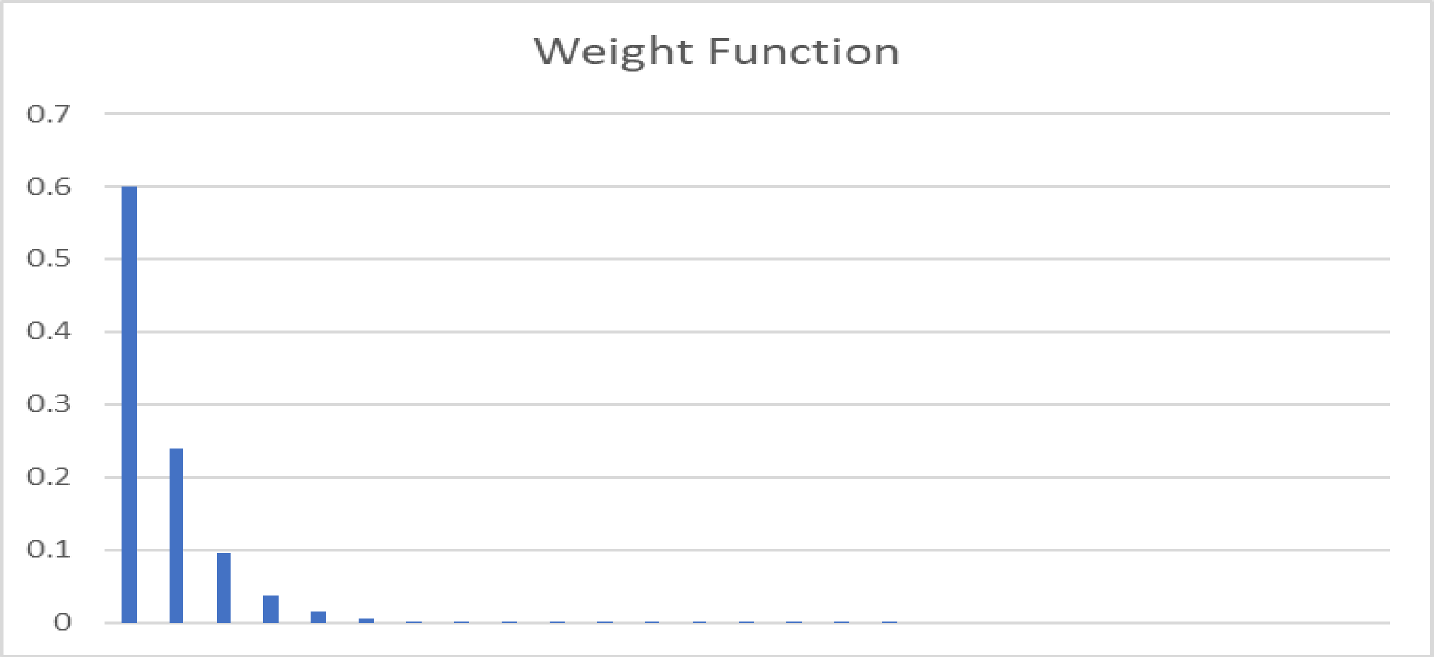
Time Series Methods

Exponential Smoothing

- Exponential smoothing moving average – the basic version

$$\hat{y}_t = wy_{t-1} + (1-w)wy_{t-2} + (1-w)^2wy_{t-3} + \cdots$$

- Weight function



Recent obs Older obs

Time Series Methods

Exponential Smoothing

- For time series with **trend**:
  - Smoothing applies to the trend component as well
  - Option: linear trend (absolute growth) or multiplicative trend (proportional growth)
  - Option to “dampen” trend so that predictions do not get explosive with long horizon
- For time series with **seasonality**:
  - Algorithm will “detrend” – i.e. remove the trend – the time series
  - Then apply a smoothing algorithm to predict seasonality component
  - Option: linear (absolute difference) or multiplicative (proportional difference)

Time Series Methods

Exponential Smoothing

- Additional feature in R – you can characterise the **error** component
    - **Usage**: Model selection statistics
    - **Usage**: Prediction interval can be computed
    - **Option**: Additive or multiplicative error

Time Series Methods

Exponential Smoothing

- Exponential smoothing in R: use the ets() function
- ETS stands for Error Trend and Seasonality
- E.g. ETS(A,A,A)
    - Additive error
    - Additive trend
    - Additive seasonality
- E.g. ETS(M,Ad,M)
    - Multiplicative error
    - Additive dampened trend
    - Multiplicative seasonality

Time Series Methods

Autoregressive Integrated Moving Average (ARIMA) Model

- A variation of the regression with some twist!
- Autoregressive (AR) model – based on **past values**:

$$y_t = b_0 + b_1 y_{t-1} + e_t$$

- Moving average (MA) model – based on **past errors**:

$$y_t = b_0 + c_1 e_{t-1} + e_t$$

  where e_{t-1} is the prediction error from the previous time point

- → **Learning** from past information and past mistakes!
- Can go back further than one period…

Time Series Methods

Autoregressive Integrated Moving Average (ARIMA) Model

- Integrated – taking a difference before modelling
    - Only required if the time series has a trend
- ARIMA(**p**,**d**,**q**) – a mixture of all three components
    - AR component = **p** past terms of the actual data
    - Integration order **d**=0 (model the raw data) or **d**=1 (model the change over time)
    - MA component = **q** terms of past prediction errors

Time Series Methods

Autoregressive Integrated Moving Average (ARIMA) Model

- Extension to seasonal ARIMA - specification

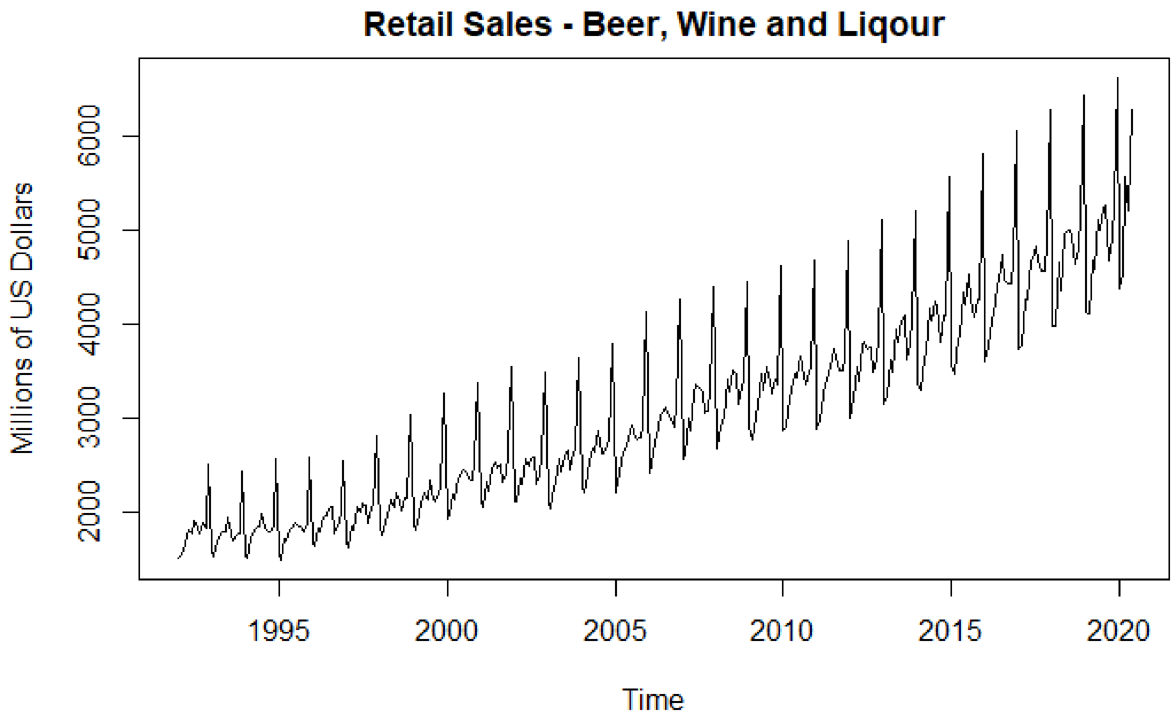
ARIMA(p,d,q)(**P**,**D**,**Q**)[**m**]

- m = seasonal frequency (e.g. monthly data m=12)
- **D** indicate seasonal differencing.
  - That is, do we model the raw data or model the season-on-season change?
  - If data is seasonal, set D=1
- **P** = number of seasonal lags of the past data
  - E.g. for P=2, we include y_{t-m} and y_{t-2m}
- **Q** = number of seasonal lags of the past prediction error
  - E.g. for Q=2, we include e_{t-m} and e_{t-2m}

Time Series Prediction

An example of time series predictive assessment

- Let us focus on the Retail Sales – Beer, Wine and Liquor
- What are the characteristics that we see?



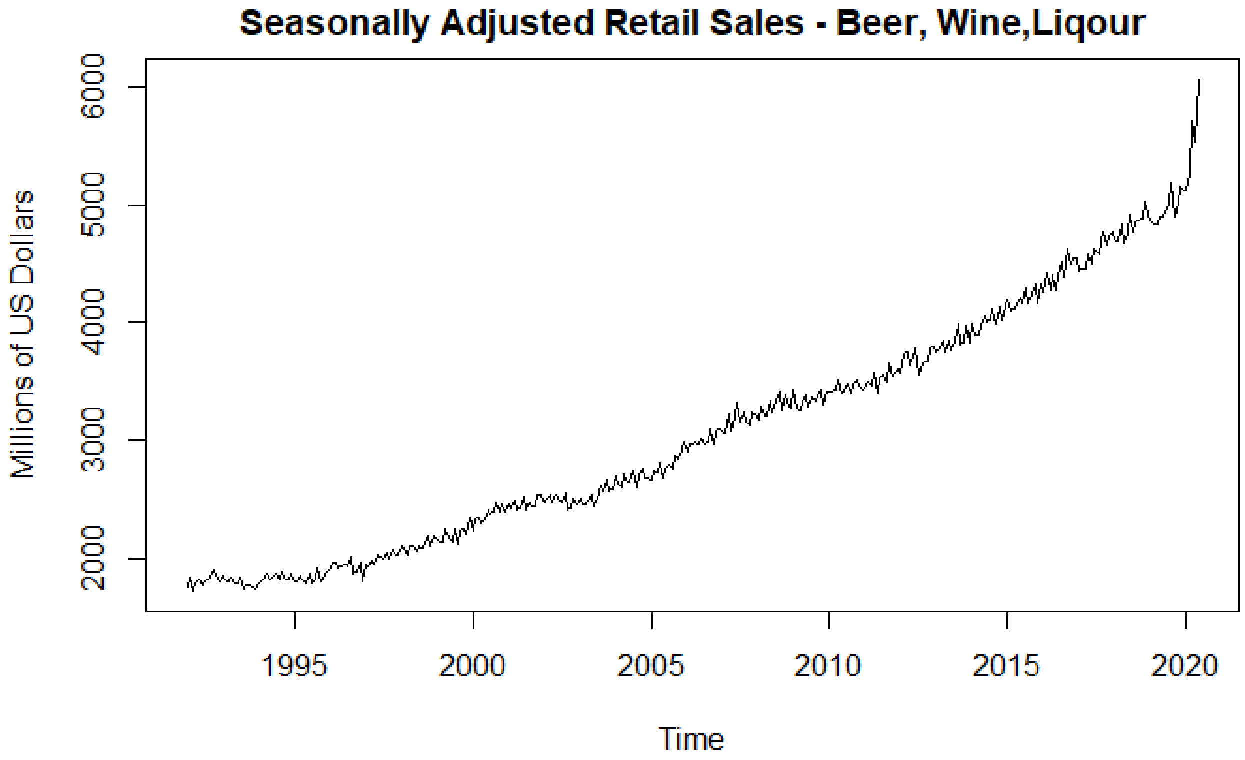
Raw data features:

- Strong seasonality
- Changing seasonal pattern
- Strong increasing trend

Time Series Prediction

An example of time series predictive assessment

- Let us focus on the Retail Sales – Beer, Wine and Liquor
- What are the characteristics that we see?



Seasonally adjusted data

- Steady trend up to 2020
- Accelerating trend post COVID lockdown

Time Series Prediction

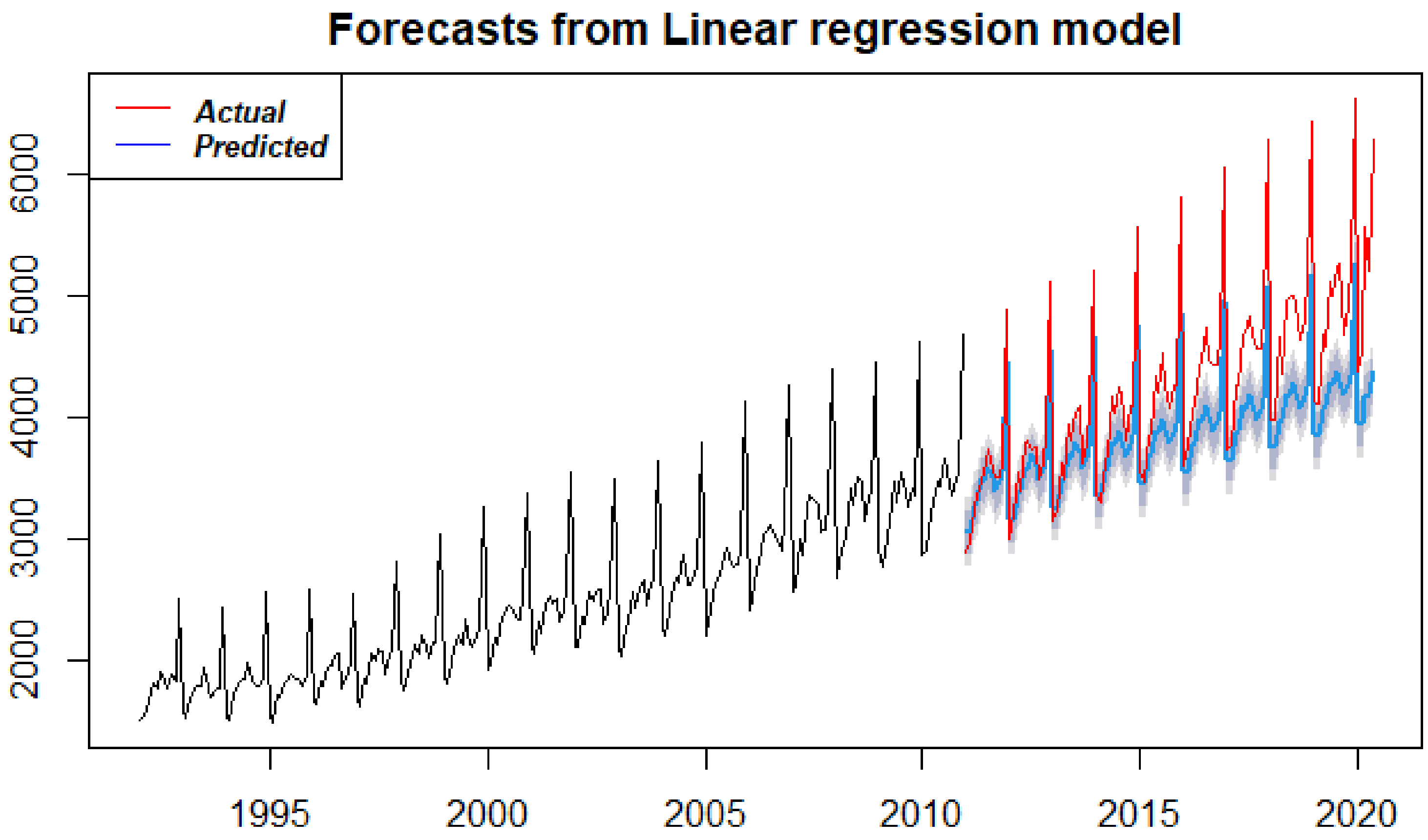
An example of time series predictive assessment

- Data split
    - Training set: January 1992 to December 2010
    - Test set: January 2011 to June 2020
- **Long horizon predictive assessment**
- **Real time predictive assessment**

Time Series Prediction

An example of time series predictive assessment

- Long horizon predictive assessment – time series regression



Red line is the test set data.

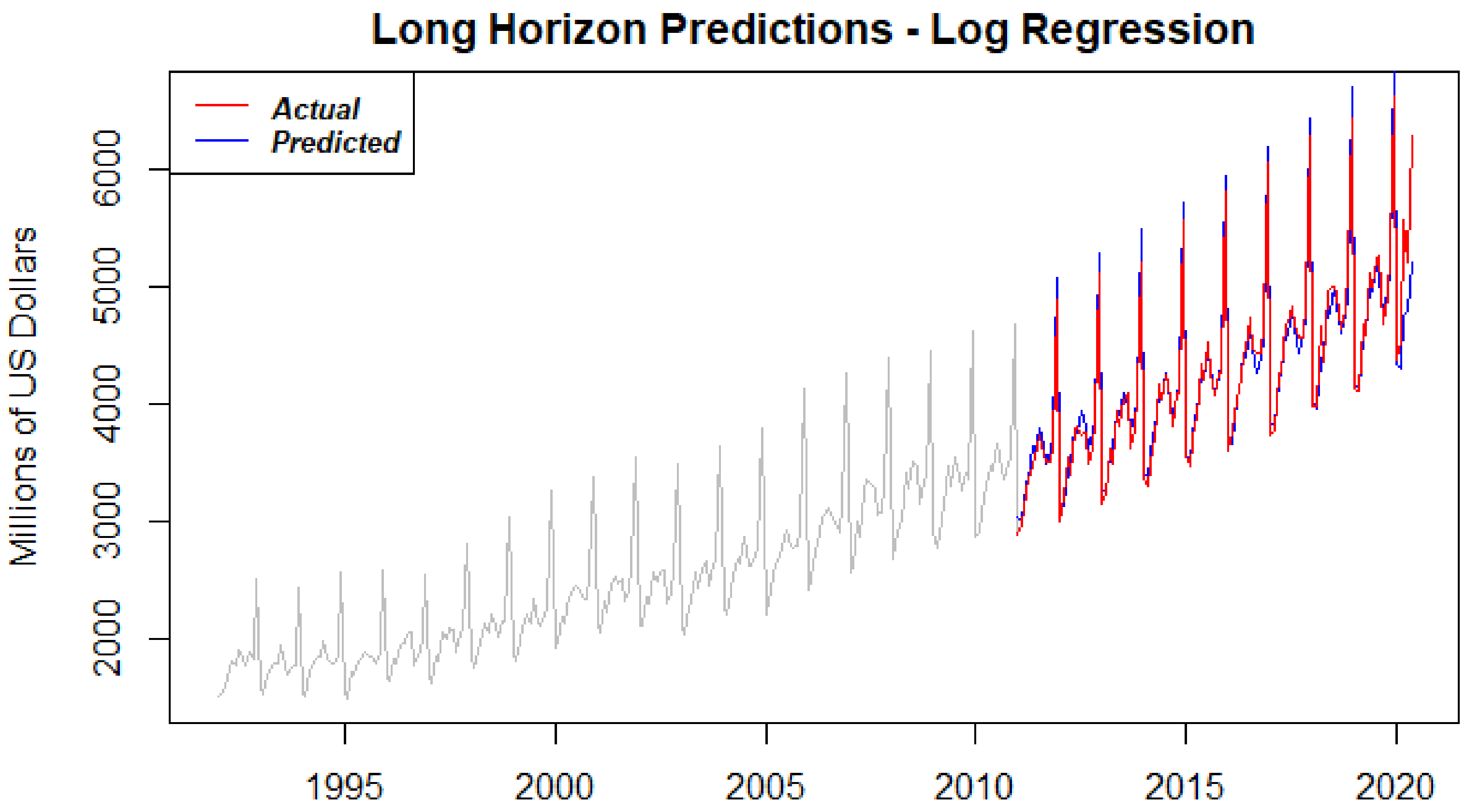
Blue is the predicted value.

Gray area indicate prediction interval.

Time Series Prediction

An example of time series predictive assessment

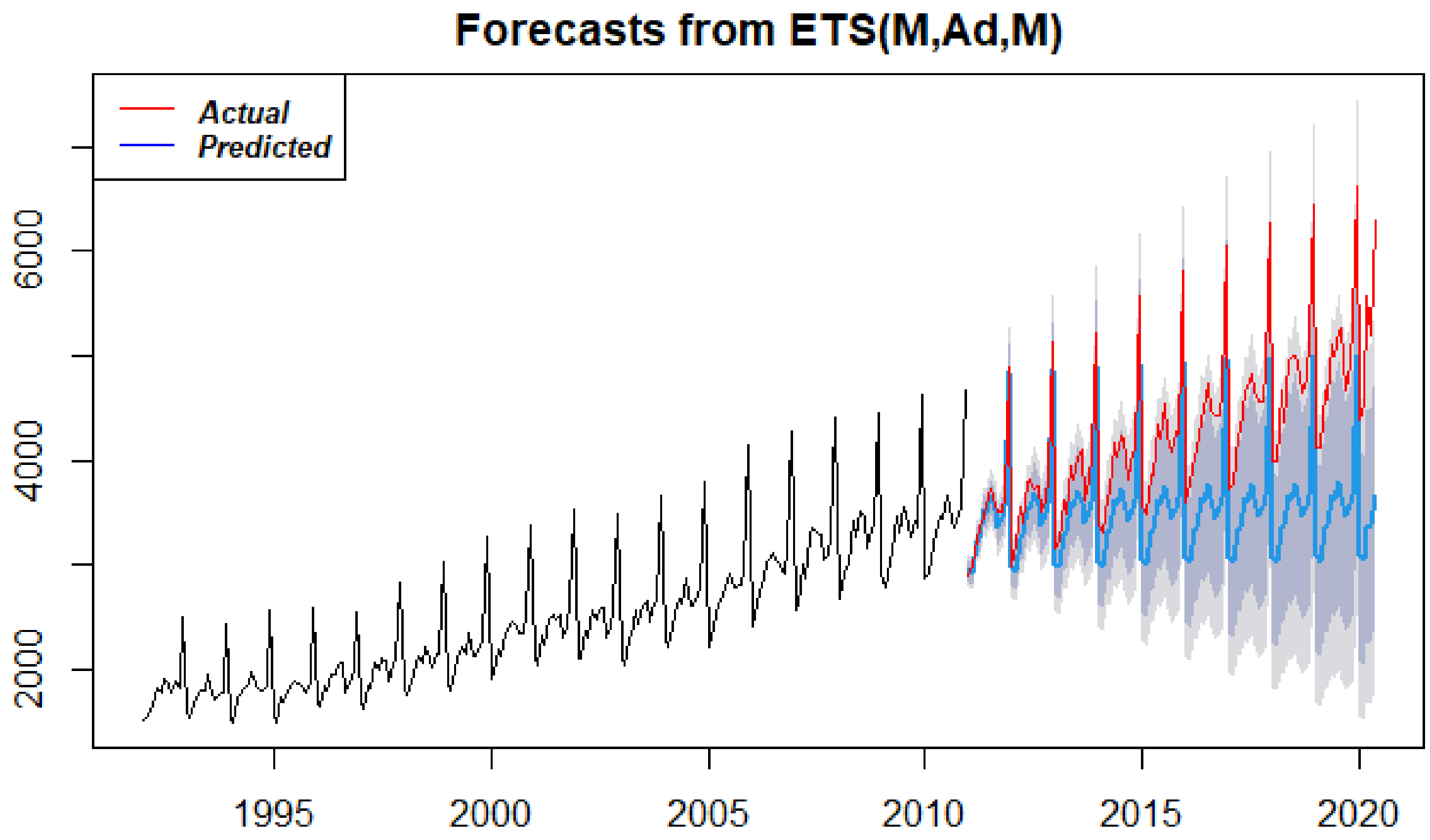
- **Long horizon** predictive assessment – time series log-linear regression



Time Series Prediction

An example of time series predictive assessment

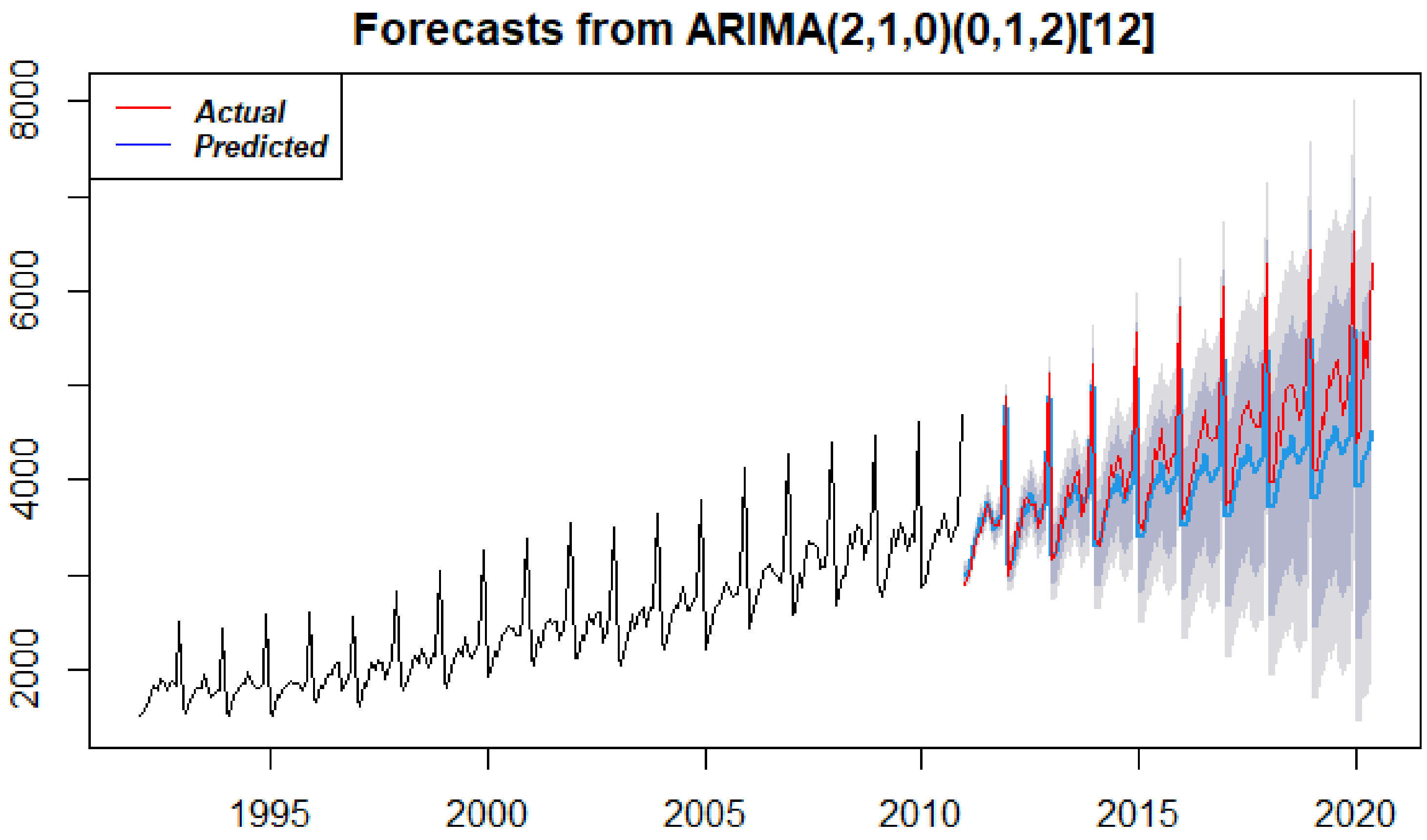
- Long horizon predictive assessment – exponential smoothing



Time Series Prediction

An example of time series predictive assessment

- **Long horizon** predictive assessment – ARIMA with seasonal components



Time Series Prediction

An example of time series predictive assessment

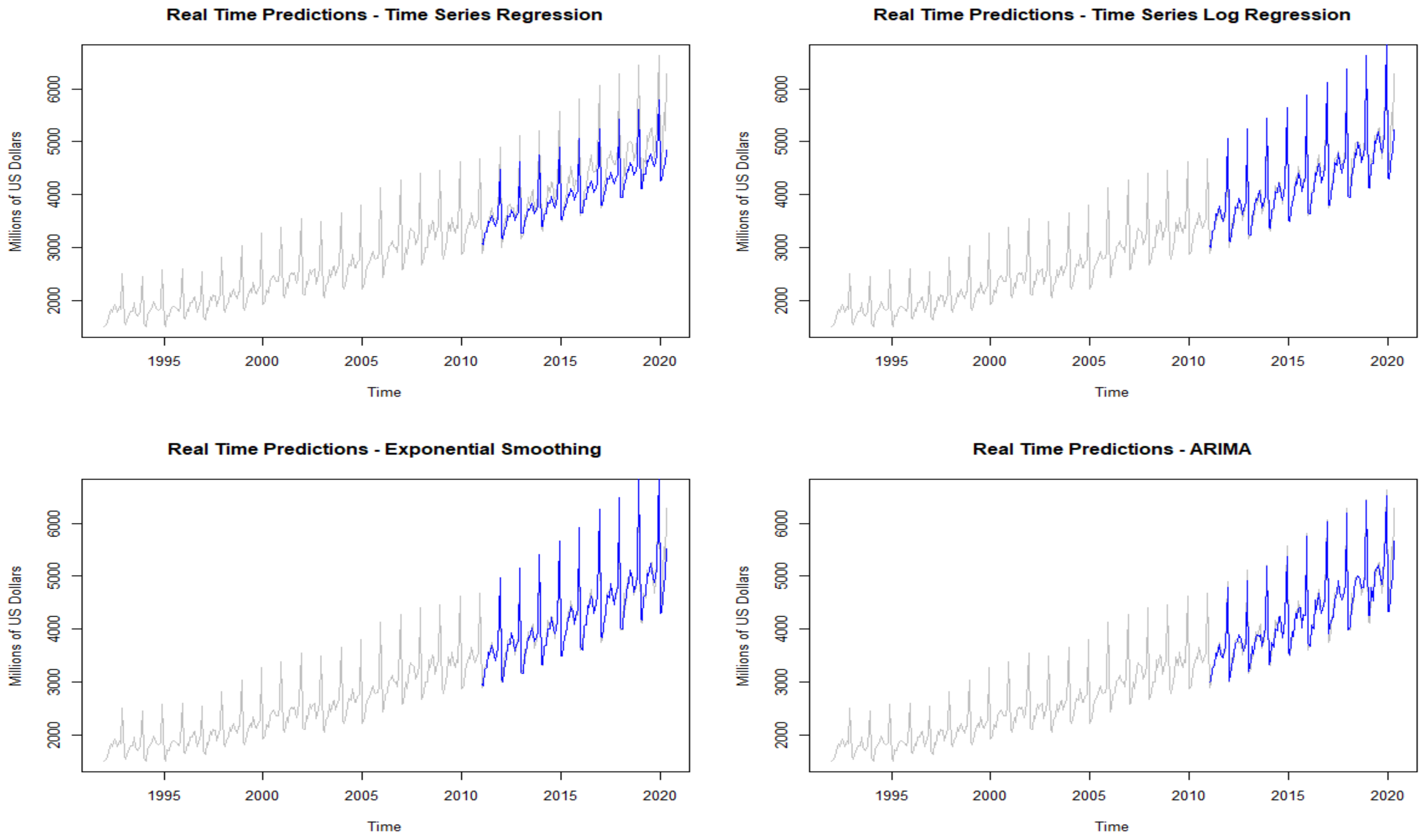
- **Long horizon predictive assessment**
- No use eyeballing the plots, let's check the accuracy measures

	RMSE	MAE	MAPE	MASE	ACF
Regression	541.188	411.912	8.642	2.010	0.509
Log Regression	**172.483**	**106.078**	**2.345**	**0.518**	**0.274**
Exp Smoothing	**898.006**	**731.770**	**15.914**	**3.570**	**0.865**
ARIMA	442.975	325.746	6.835	1.589	0.617

Time Series Prediction

An example of time series predictive assessment

- **Real time** predictive assessment



Time Series Prediction

An example of time series predictive assessment

- **Real time** predictive assessment – accuracy metrics

	RMSE	MAE	MAPE	MASE	ACF
Regression	**363.986**	**270.177**	**5.688**	**1.318**	0.282
Log Regression	169.069	103.585	2.289	0.505	**0.286**
Exp Smoothing	144.030	96.745	**2.150**	0.472	**-0.003**
ARIMA	**130.377**	**94.008**	2.183	**0.459**	0.187

Time Series Prediction

An example of time series predictive assessment

- A range of methods available
- Remember the assessment process must match the intended use
  - Long horizon vs real time
  - What may be accurate in long horizon may not be accurate for real time
- As usual, **user-specific loss** will bring in the business perspective