Forecasting, 8. October

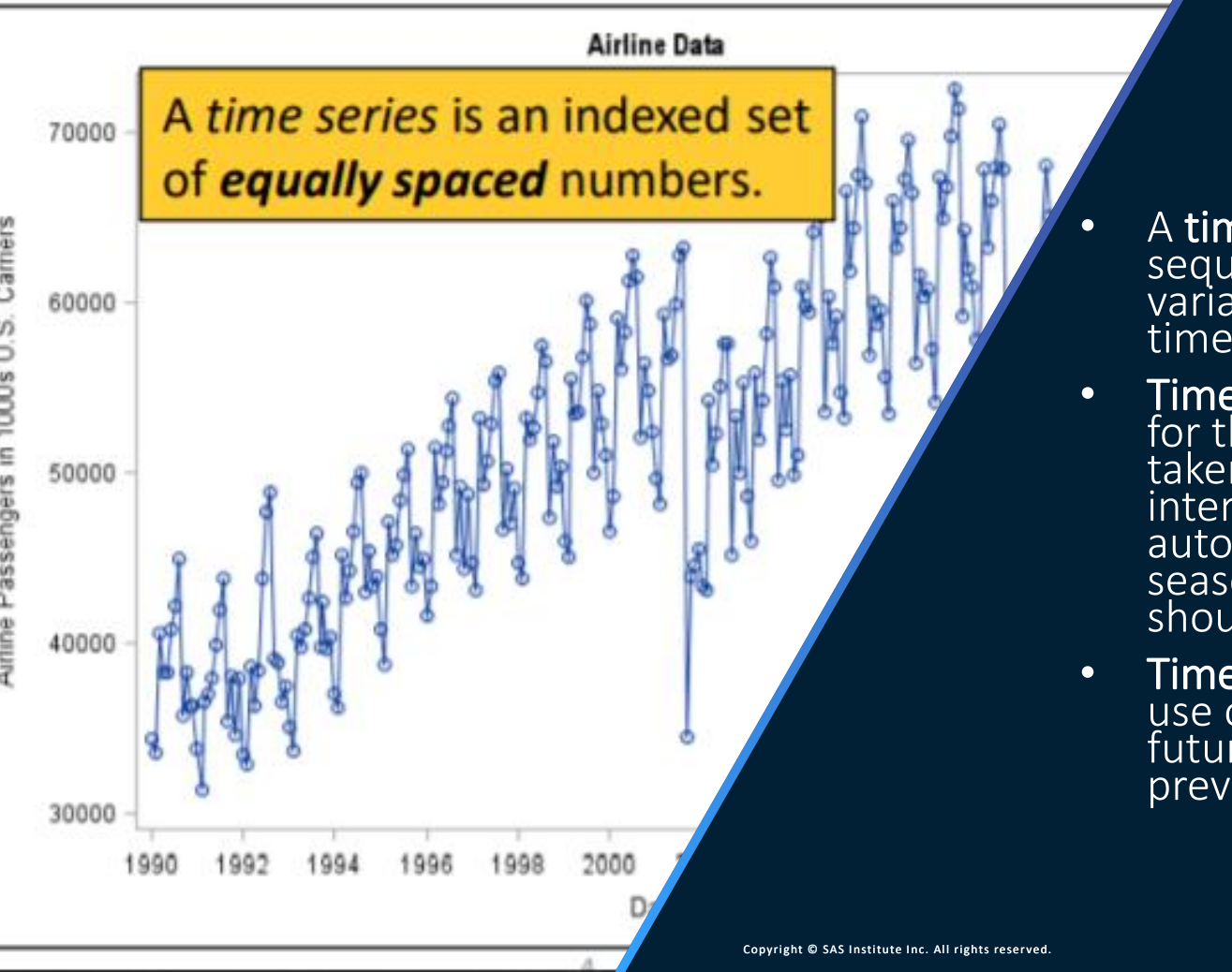
Pia Skare Rønnevik

fans
NORDIC SAS USER GROUP

sas

Agenda

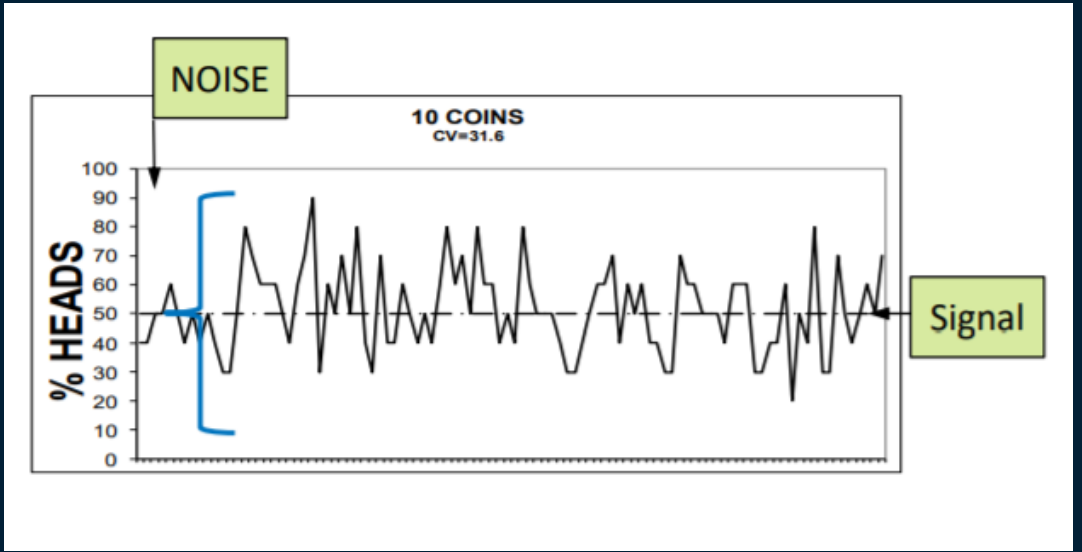
- What is time series ?
- Variation in time series data
- The different signal components
- Time series models
- Performace
- Measurements of model fit
- Simulating a Prospective Study
- Demo of how to make forecasting models in Model Studio



- A **time series** is an ordered sequence of values of a variable at equally spaced time intervals.
- **Time series analysis** accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for.
- **Time series *forecasting*** is the use of a model to predict future values based on previously observed values.

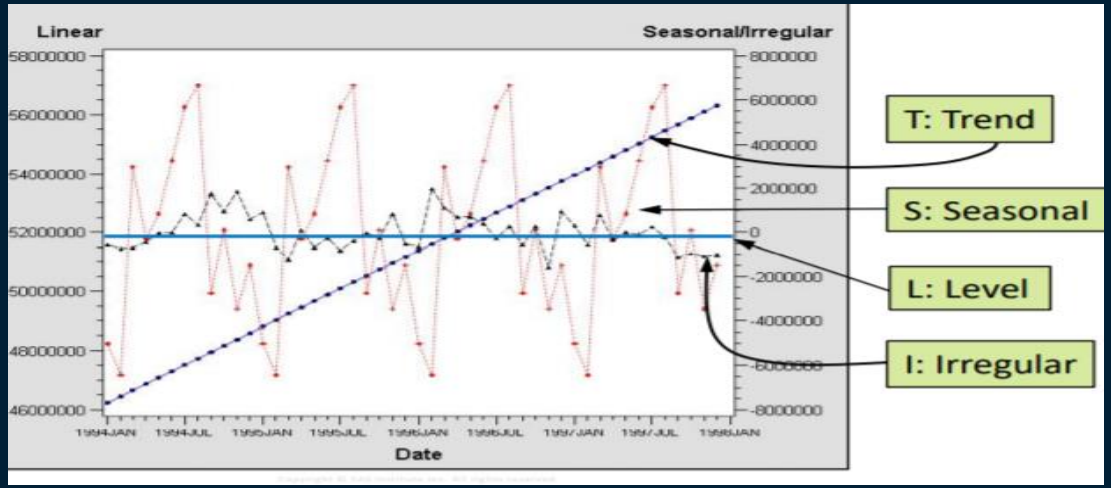
Variation in time series data

- signal
- noise



Signal Components

- level
- trend
- seasonality
- cycle
- exogenous (also known as explanatory variable effects)
- irregular



Autocorrelation

- The irregular part of a series contains both noise and signal.
- Autocorrelation is part of the signal in the irregular component.
- Autocorrelation simply means that current values in a time series are related with previous values.

Time series models

Exponential Smoothing Models (ESM)

- use weighted averages of past observations to forecast new values.

- gives more weights to the recent values than older observations. Thus, as observations get older (in time), the importance of these values get exponentially smaller.

- combines Error, Trend, and Seasonal components in a smoothing calculation (ETS). Each components can be combined either additively, multiplicatively, or be left out of the model.

- work well when the time series shows a clear trend and/or seasonal behavior.

Time series models

AutoRegressive Integrated Moving Average with eXogenous variables Models (ARIMAX)

- AR:Autoregressive → Time series is a function of its own past.

- I:Integrated → Differenced values between successive time points can be modeled, and after modeling returned to the undifferenced metric.

- MA: Moving Average → Time series is a function of past shocks (deviations, innovations, errors and so on)

- X: Exogenous → Time series is influenced by external factors.

Time series models

Unobserved Components Models (UCMs)

- also known as structural time series models.
- can be considered as a multiple regression model with time-varying coefficients.
- performs a time series decomposition into components:

$$Yt = Trend + Season + Cycle + Regressors$$

- the components in the model have their own models and its own source of error and forecast.

Time series models

Naive Models

- often used to compare your complex statistical models against:
  - Simple mean models
  - Simple random walk
  - Random walk drift

- the complex statistical models are only valuable if they perform better than Naive models.

Simple Regression Models

Time series models

- a good substitute if other forecasting models becomes problematic.
- time can be used as an input variable.
- seasonality can be modeled using dummy variables representing each season.
- predefined trend components: linear, quadratic, cubic, log-linear, exponential, and so on.

Performance

- **Simple models**: have no performance issues.

- **ESM**: can be constructed quickly and easily, so they always have good performance.

- **ARIMAX**: require many more computer cycles than simple or exponential smoothing models, so some approximations and shortcuts are used to speed performance.

- **UCM**: are very computer-intensive and should be tried only on small data sets or individual time series.

Measurements of model fit

Mean Absolute Percent Error (MAPE)

- MAPE is one of the most common accuracy measures in business forecasting.
- As a selection criterion, choose the model with the smallest value of MAPE.
- MAPE is the average of all of the individual absolute percent errors.

Mean Absolute Error (MAE)

- MAE is not commonly used as an accuracy measure in business forecasting.
- As a selection criterion, choose the model with the smallest value of MAE.
- The average of all of the individual absolute errors.

Measurements of model fit

Root Mean Square Error (RMSE)

- RMSE is the square root of the average of all of the individual squared errors, adjusted for the number of estimated model parameters.

- RMSE is commonly used as an accuracy measure in industrial, economic, and scientific forecasting.

- As a selection criterion, choose the model with the smallest value of RMSE.

fans
NORDIC SAS USER GROUP

sas

Simulating a Prospective Study

Divide the time series data into two segments:

- The fit sample is used to derive a forecast model.
- The holdout sample is used to evaluate forecast accuracy.

Full = Fit + Holdout data is used to fit a deployment model.

DEMO