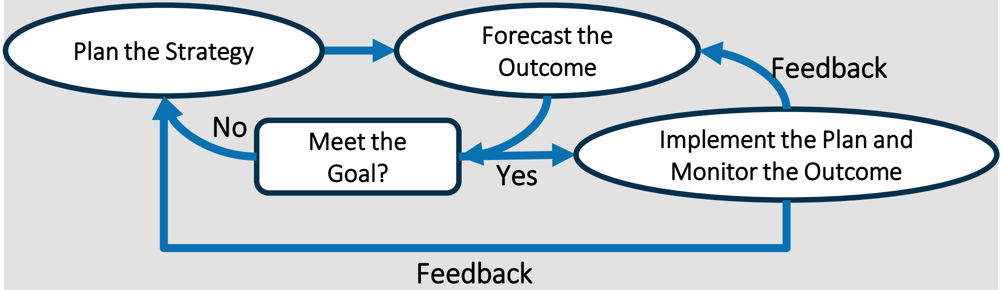
Time Series Forecasting Landscape

A quick guide to common forecasting practices

What Is Forecasting?

Forecasting is the process of predicting uncertain events, trends, or levels in the future on the basis of current information and historical data.

Forecasting is an integral part of the decision-making process.

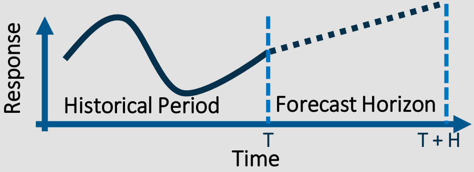


Forecasting vs. Predictive Modeling

Forecasting is considered a subfield in predictive modeling. A critical distinction between forecasting and other predictive modeling is the **time** component. Forecasts at a future time (horizon) depend on responses in the historical period.

Forecasting

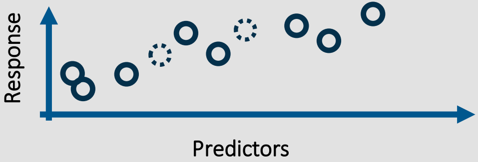
$$\hat{y}_{T+h} = f_h(y_T, y_{T-1}, \ldots, x_1, x_2, \ldots)$$



- Tourist visits next summer
- Product demand next month
- Website visits next week

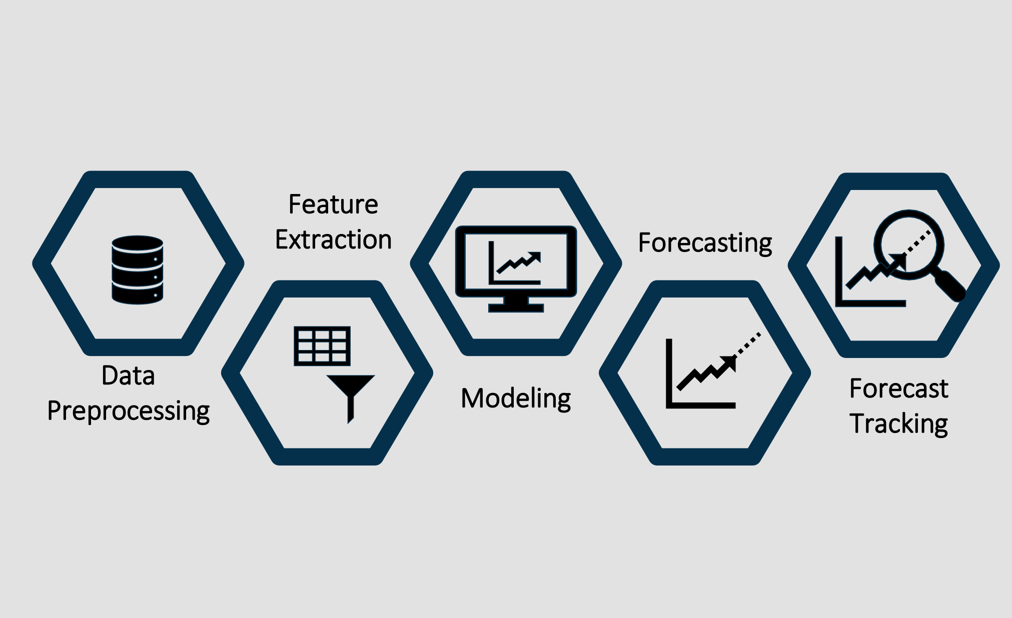
Predictive Modeling

$$\hat{y} = f(x_1, x_2, \ldots)$$



- MPG given car's characteristics
- Wine quality score given lab test results
- Customers' ratings given product information

Forecasting Analytic Process Flow



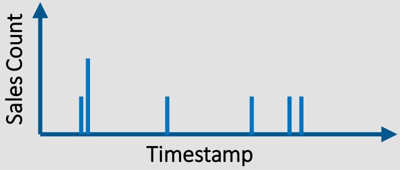
Data Preprocessing

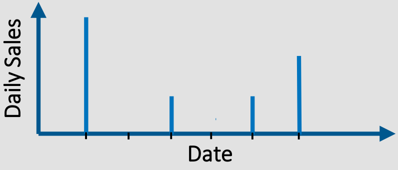
Transactional Data

Timestamp	Sales Count
Feb. 10, 2:00pm	1
Feb. 10, 3:00pm	2
Feb. 12, 9:00am	1
Feb. 14, 10:00am	1
Feb. 15, 1:00pm	1
Feb. 15, 4:00pm	1

Time Series Data

Date	Daily Sales
Feb. 10	3
Feb. 11	0
Feb. 12	1
Feb. 13	0
Feb. 14	1
Feb. 15	2





Notes

- Transactional data vs. time series data:
  - Transactional data are logged every time an event or transaction occurs, and often have no particular frequency.
  - For the purpose of analysis, transactional data need to be accumulated into time series data. Time series data are timestamped data that are collected over time at a fixed frequency.

- Typical examples:

Transactional Data		Time Series Data
Internet data	→	Website visits per hour
Point of sales (POS) data	→	Sales per month
Inventory data	→	Inventory draws per week
Trading data	→	Daily closing price

Feature Extraction

Feature extraction is the process of creating **new features** from an initial set of data. These features typically encapsulate the central properties of a data set.

Time Series Decomposition

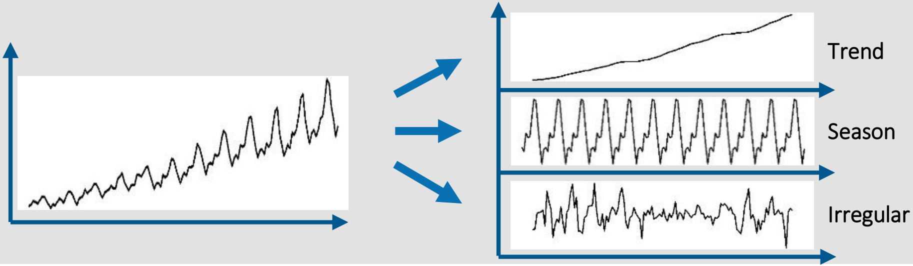
- Additive
- Multiplicative

Time Series Components

- Trend
- Cycle
- Seasonal
- Irregular
- Motifs
- SSA* group

Feature Extraction Methods

- Classical decomposition
- SSA* decomposition
- Spectral decomposition



*SSA (singular spectrum analysis)

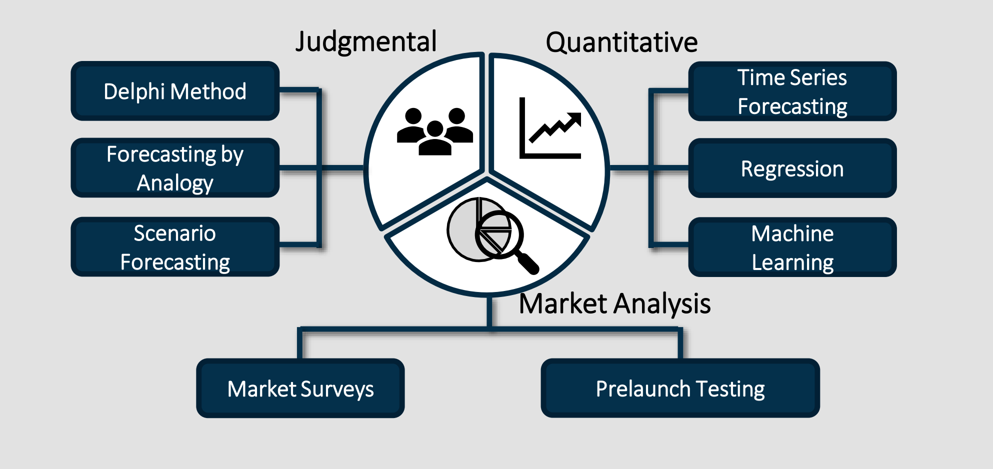
Notes

How can feature extraction help?

- **Exploratory studies:** Feature extraction identifies patterns (such as seasonal cycles and repetitive motifs) that help you understand your data.
- **Modeling and forecasting:** Feature extraction can generate new features to include in the model, which often helps improve forecast accuracy.
- **Interpretation:** Extracting, identifying, and associating features with your forecasting forecasting results help you interpret what causes the peaks and valleys of your forecast.

Modeling – A Brief Survey of Methods

Forecasting methods can be categorized into three types: judgmental, quantitative and market analysis. These methods often complement each other and can be combined to create more robust forecasts.

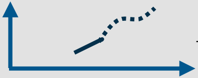


Modeling – Data-Dependent Methods

Different models should be used with different data types.

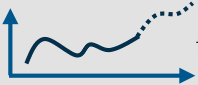
New time series

- Judgmental methods
- Market analysis methods



Short time series

- Judgmental methods
- Diffusion models
- Regression

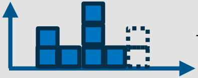


Long time series

- Time series forecasting
- Machine learning
- Regression

Intermittent time series

- Simple averages
- Croston model
- Pooling/clustering

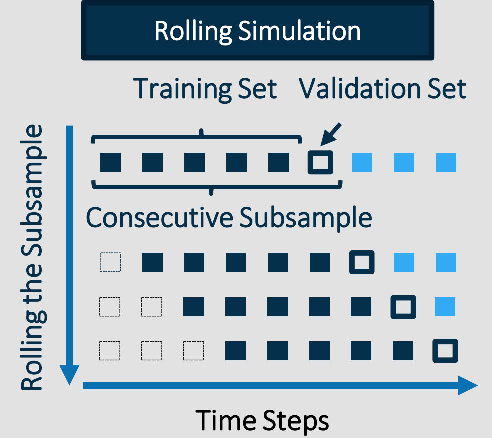


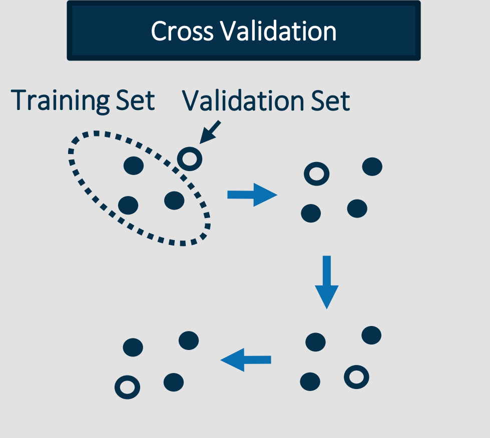
Count time series

- Count series forecasting
- Regression

Modeling – Validate Your Methods

Time series data require a specific validation scheme, called rolling simulation, which preserves the sequential structure. (This scheme is different from cross validation, which is often used in machine learning.) Average errors across all validation sets are used to select best model.





Notes

Can you apply machine learning models in forecasting?

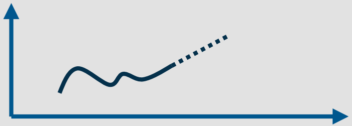
Machine learning algorithms have become popular and widely used in various application fields, including time series forecasting. The following points are important to consider when applying machine learning models to forecasting problems:

- Many machine learning models assume that the data do not have a particular order and can be shuffled, randomly sampled, or distributed. This is usually not true for time series data.
- Extracted features, such as lagged response variables and seasonal dummies, can help a machine learning algorithm to learn the dependencies on past values.
- Some machine learning models have large numbers of parameters, which can lead to overfitting for short time series. In these cases, time series models are often preferred.
- Neural network models, such as Long Short Term Memory (LSTM), or combinations of traditional time series models and neural networks, work well for problems that include independent variables and nonlinearities.

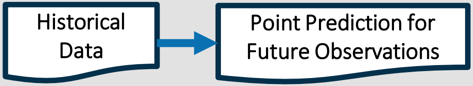
Forecasting

Making forecasts with uncertainty in mind.

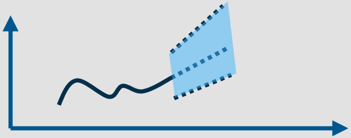
Point Forecast



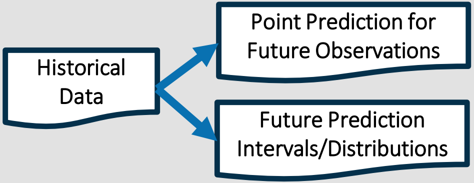
- Provide point prediction for future observations



Probabilistic Forecast



- Provide uncertainty estimation and distribution information for future observations



Notes

Point Forecast

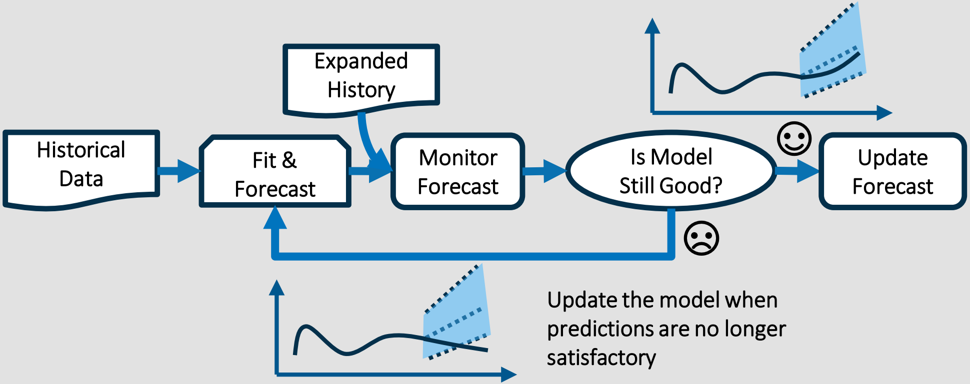
- Provides a single estimate for future observations at each time step
- Easier to estimate because a variety of tools are available for accurate mean estimation in the fields of time series forecasting, regression, and machine learning
- Shortcomings: forecast users and decision makers are often unaware of the uncertainty of a point forecast and use it to take high-risk actions

Probabilistic Forecast

- Provides more information about the forecast uncertainty than a point forecast
- Harder to obtain because many machine learning methods do not predict a full distribution
- Typically requires distributional assumptions or ability to generate additional samples that match characteristics of historical data
- Application areas: weather forecasting, energy forecasting, product demand in supply chain, financial modeling and other applications in which uncertainty and volatility play key roles in decision making

Forecast Tracking

Forecast tracking makes sure the model is adjusted in a timely manner.



Update the model when predictions are no longer satisfactory

Notes

How often should the model be updated?

- The most common practice for incremental forecast updating is to refresh all the models in a fixed cycle (e.g., every 2 months).
- However, this practice creates the potential for wasted effort because all models are re-examined even when they are performing well. Also, model changes could cause forecasts to be drastically different from what they were before.
- By monitoring forecast quality and updating models only as needed, the system creates a perfect balance between results stability and forecast quality.

Appendix

Ensemble Models

- Rather than attempt to specify a single best model to forecast your data, it is usually a good practice to create an ensemble of several appropriate models.
- Such combination models have generally proven to perform better than a single model.
- Taking a simple average of the models is usually sufficient, although complex weighting schemes may eke out slightly more accuracy with much greater effort.
- Avoid including clearly inappropriate (aka "poisonous") models in the ensemble, as these can degrade performance.

Appendix

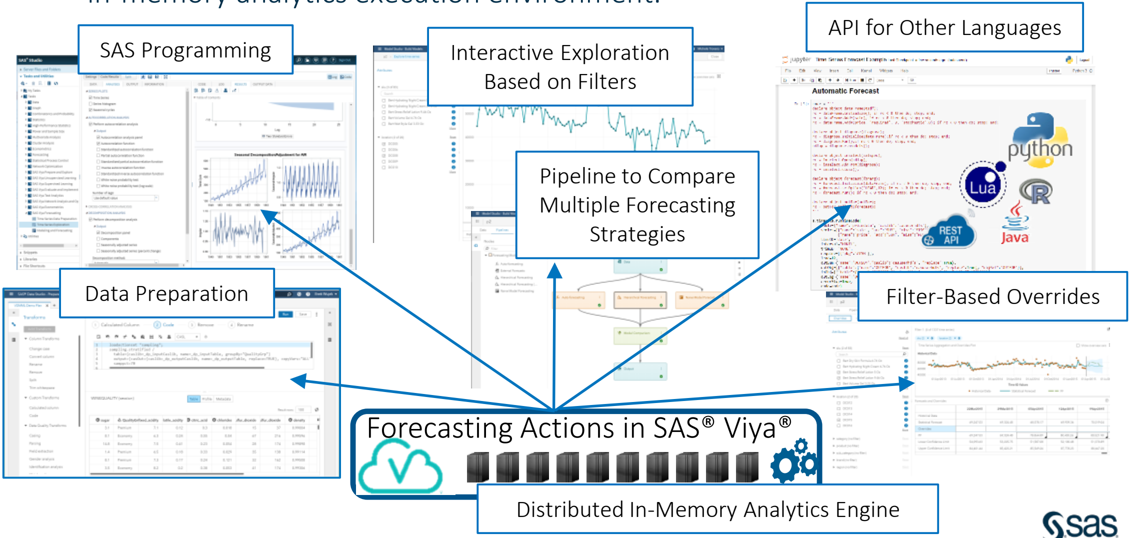
Forecasting Value Added (FVA)

- FVA is the change in a forecasting performance metric that can be attributed to a particular step or participant in the forecasting process.
- It is used to identify ***process waste*** – those activities that are failing to improve the forecast.
- The concept turns attention away from the end results (such as forecast accuracy) to focus on the overall efficiency and effectiveness of the forecasting process.
- FVA analysis is often used to compare an analyst's override of forecasts to the original statistical forecasts and to compare both to simple naïve forecasts.
- Positive FVA means the step in the process added value by making the forecast better. Negative FVA means the process step made the forecast worse.

SAS® Visual Forecasting

sas.com/vf

Access state-of-the-art forecasting algorithms in a powerful, distributed in-memory analytics execution environment.



Contributors

- Michele Trovero, Senior Manager, Forecasting R&D
- Michael Leonard, Director, Forecasting R&D
- Mahesh Joshi, Senior Manager & Esteemed Researcher, Forecasting R&D
- Youngjin Park, Forecasting and Machine Learning Specialist
- Steven Mills, Research Statistician Developer
- Iman Vasheghani Farahani, Research Statistician Developer
- Joe Katz, Principal Product Manager, Product Management
- Mike Gilliland, Principal Product Marketing Manager, Product Marketing
- Udo Sglavo, Vice President of Advanced Analytics
- Bruce Elsheimer, Senior Manager, Advanced Analytics Testing
- Jennifer Beeman, Senior Analytics Software Tester

sas.com