Ask the Expert

Enhance Forecasting Accuracy With Time-Series Segmentation and Machine Learning

Spiros Potamitis, Senior Product Marketing Manager

Spiros Potamitis

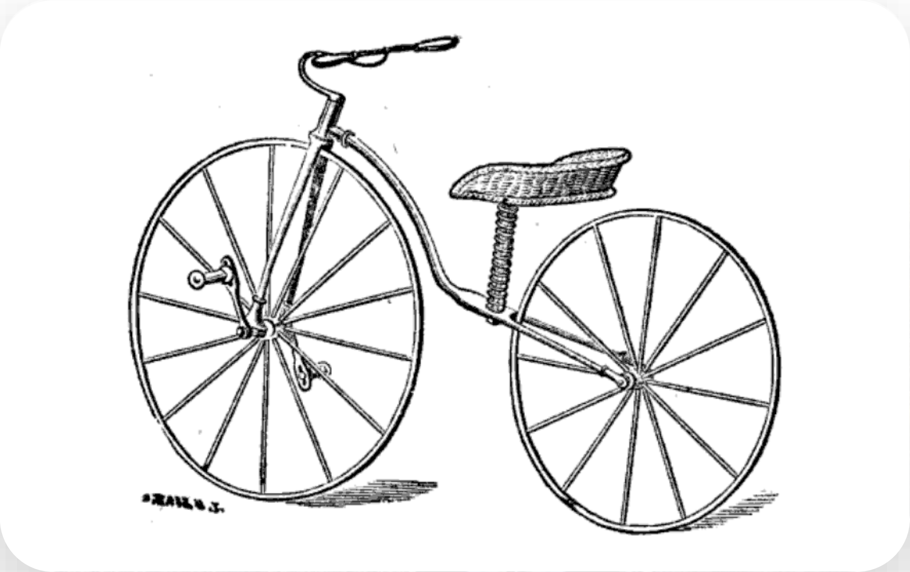
Senior Product Marketing Manager

Spiros is a data scientist and product marketing manager at SAS specializing in forecasting and machine learning. He works with R&D and product management to define the direction and vision of SAS products' while helping SAS customers apply advanced analytics and AI to drive business value. Spiros holds master's degrees in both computer engineering and information management from the University of Manchester.

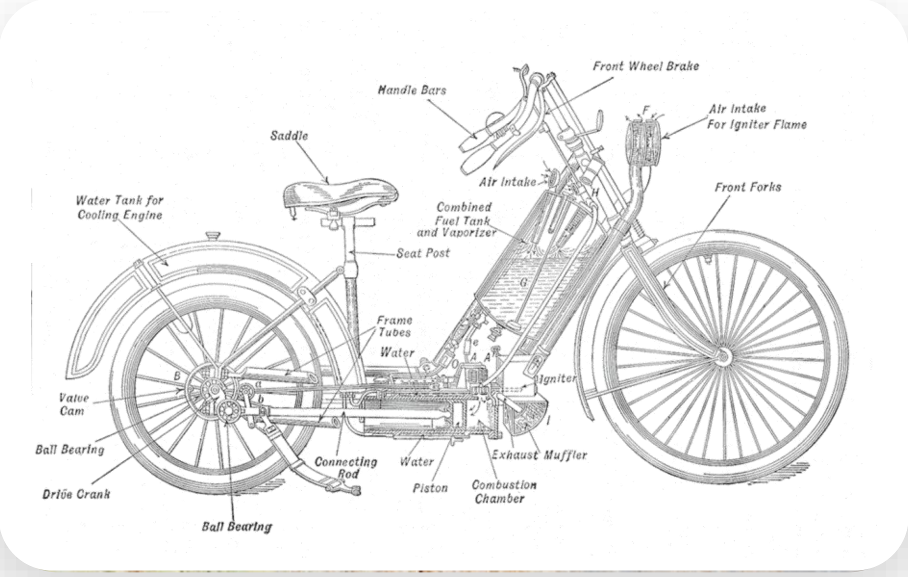
Enhance Forecasting Accuracy With Time-Series Segmentation and Machine Learning

Spiros Potamitis, SAS

Getting from A to B



- Flexible
- Cost efficient
- Great for short distances



- Fast and agile
- Easy to park
- Great for days with mild weather

- Comfortable
- Safe
- Great for families and carrying luggage

A Brief History of Forecasting Models

Statistical Models

- Exponential Smoothing
- Arima(X)
- UCM

Deep Learning

- RNN-based
- MLP-based
- CNN-based
- Transformer-based

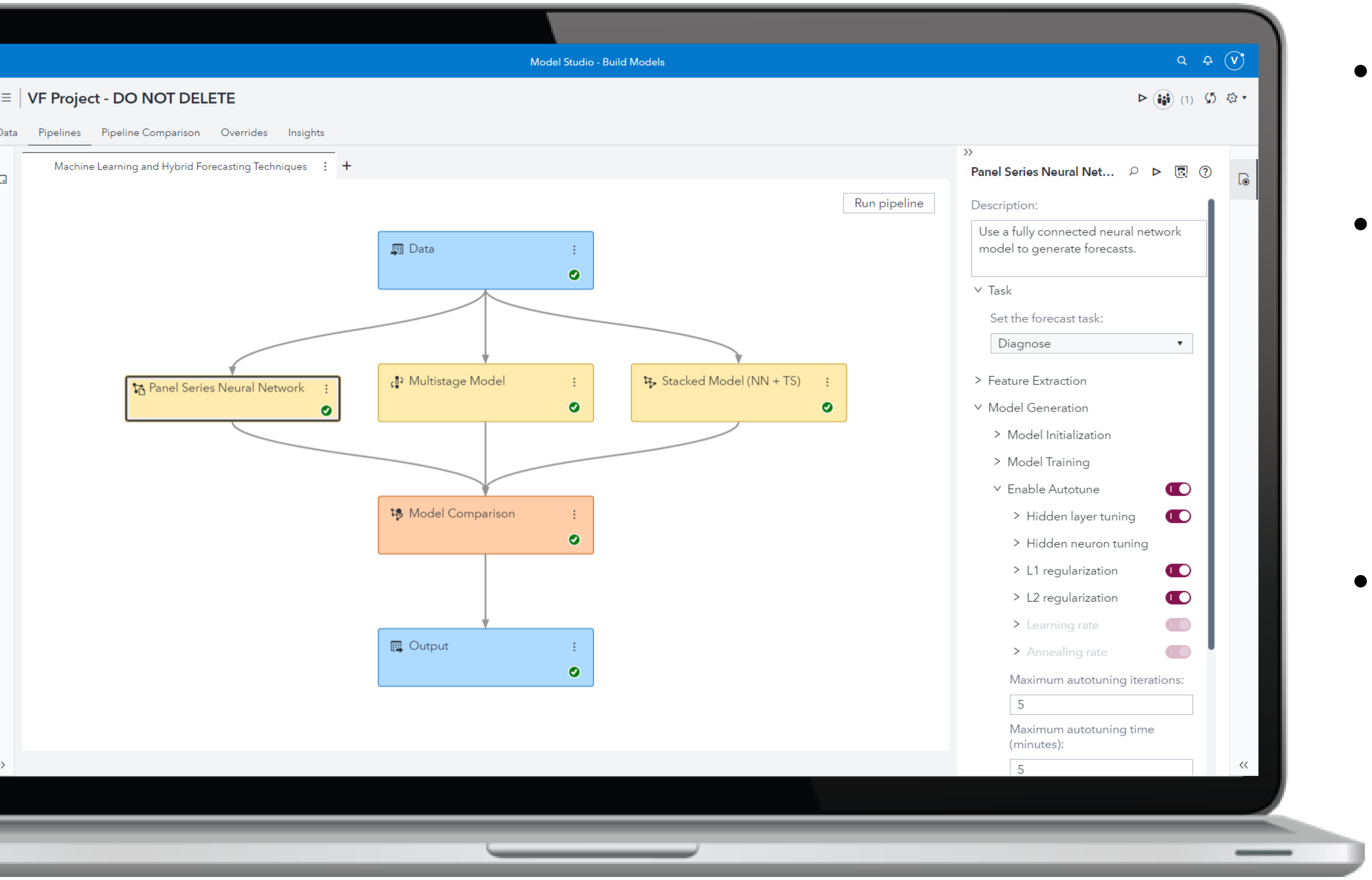
Machine Learning

- Random Forest
- LightGBM

GenAI

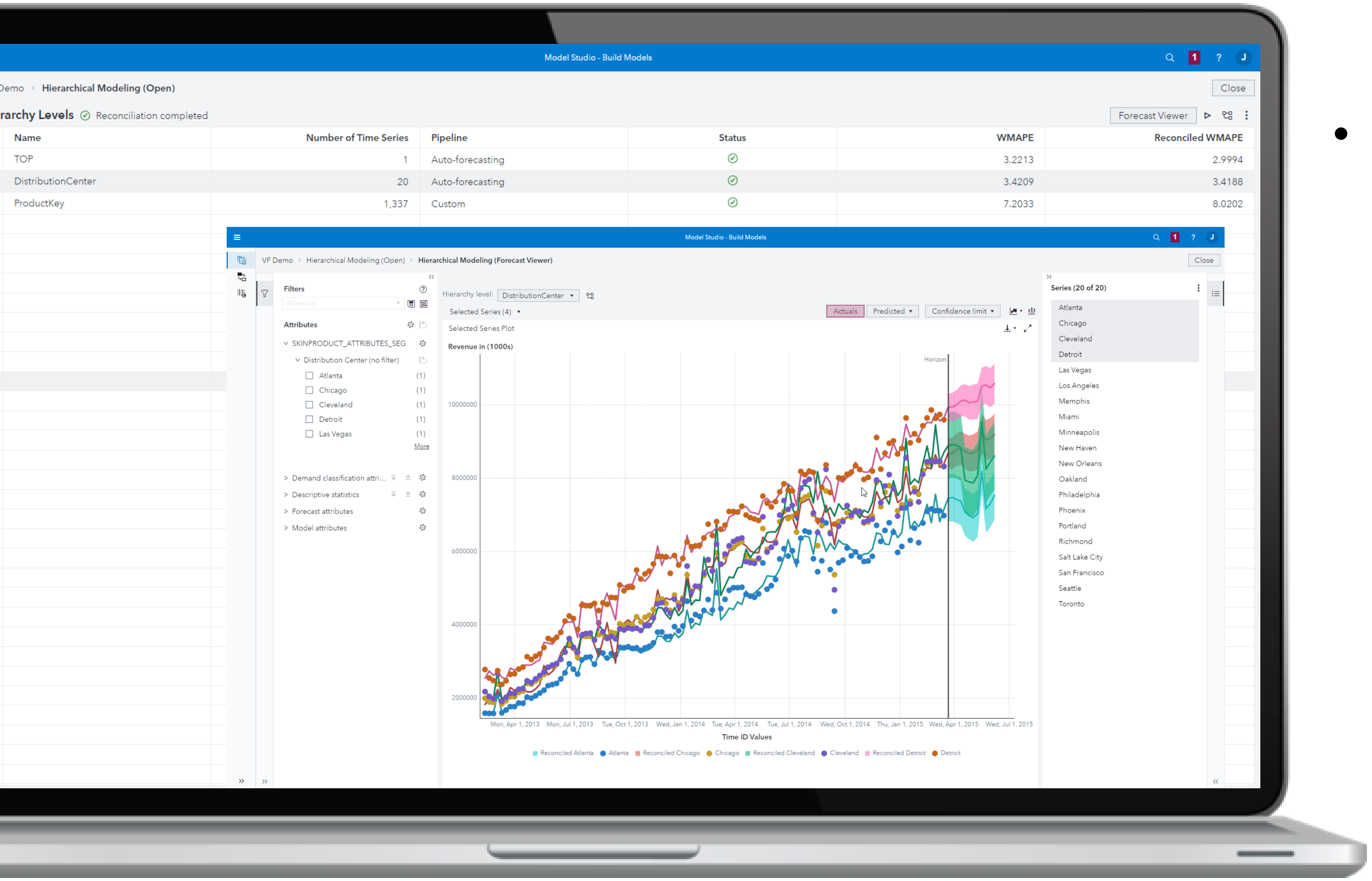
- Foundation Models

Advanced Forecasting Techniques



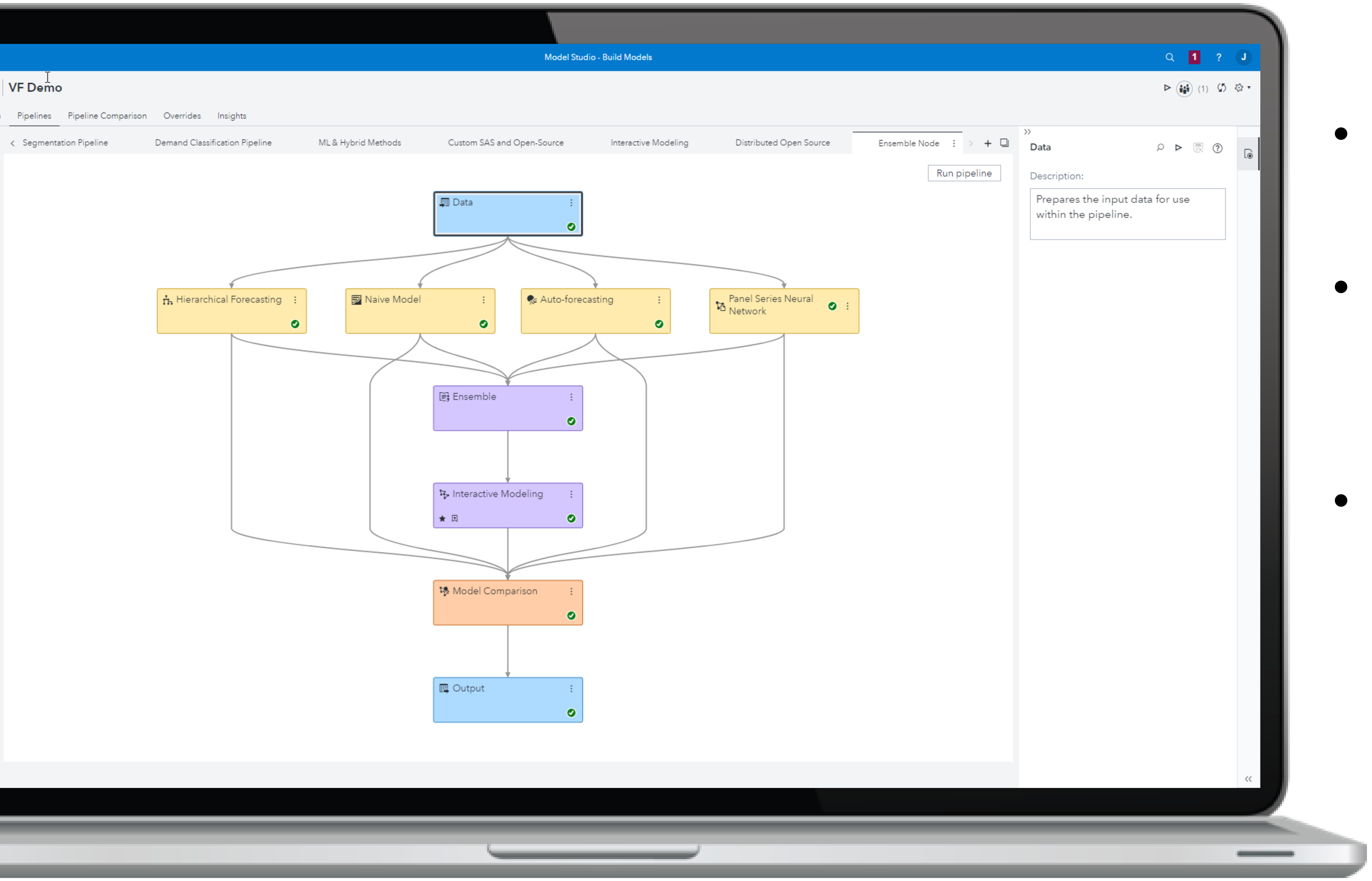
- Automatic data transformation and feature generation based on the specified technique
- Machine Learning and hybrid techniques
  - Panel Series Neural Networks (PSNN) with autotuning capabilities
  - Multistage Model (hybrid, incorporates a hierarchy)
  - Stacked Model (hybrid, NN + Time Series)
- Deep Learning techniques (RNNs, LSTMs, GRUs) with recursive strategy applied automatically for multistep forecasting

Advanced Forecasting Techniques



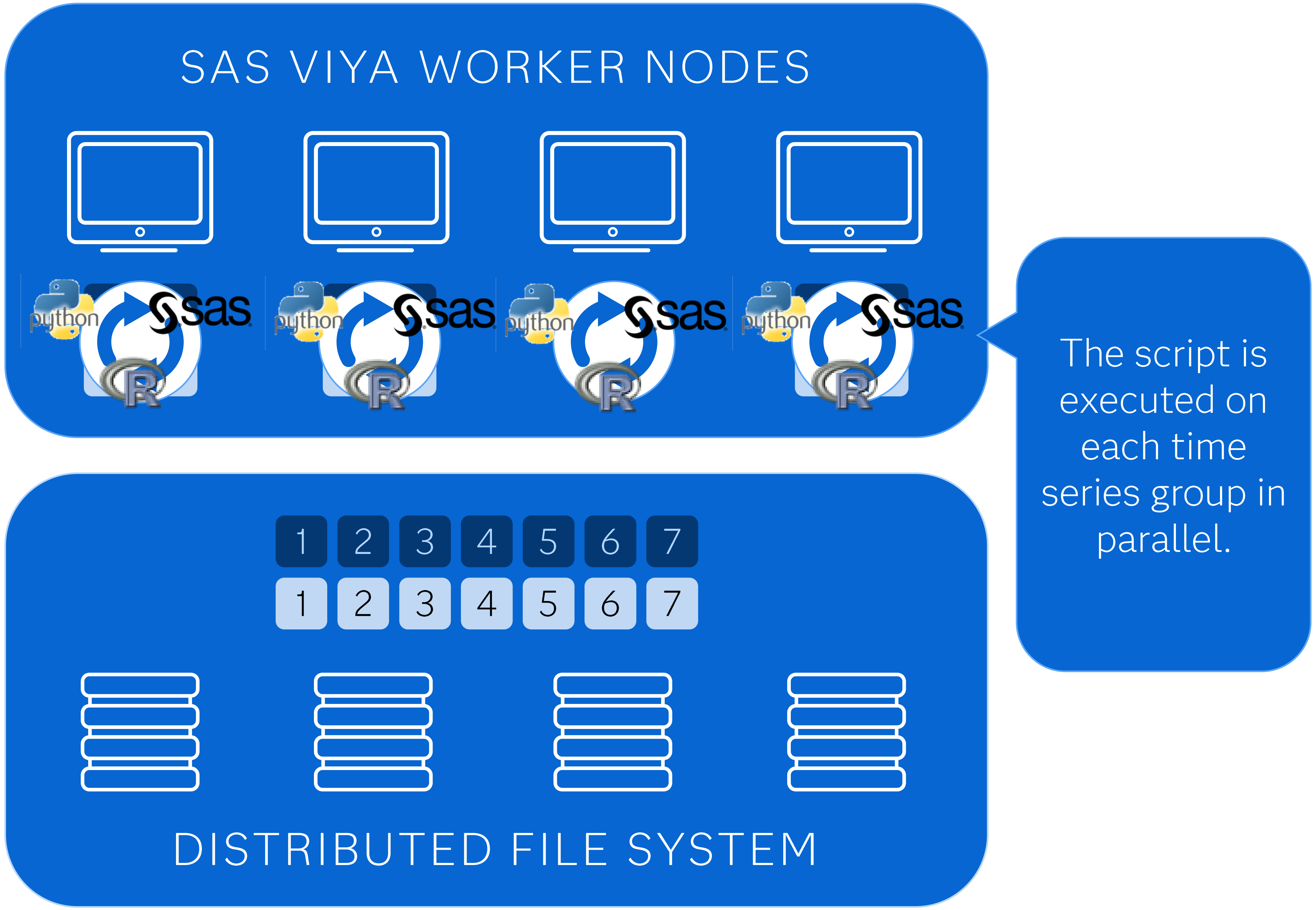
- Hierarchical Modeling node with customizable pipelines for each level of the hierarchy

Advanced Forecasting Techniques



- Hierarchical Modeling node with customizable pipelines for each level of the hierarchy
- Interactive modeling capabilities to assess and compare existing models as well as create new custom models for each series
- Ensemble Modeling to always select the best performing model for each series

Empower Open Source Users



- Distribute native Python and R code, along with SAS code, to run in parallel in the cloud
- Easily reuse open source forecasting algorithms in all business areas
- Scale open source algorithms for large volumes of time series
- Take advantage of all coding talent
- Apply, compare, and put into production the best performing algorithms from SAS and open source

Statistical Methods vs Machine Learning

- ✓ Simplicity
- ✓ Trend and seasonality
- ✗ Non-linear patterns
- ✗ Global Models

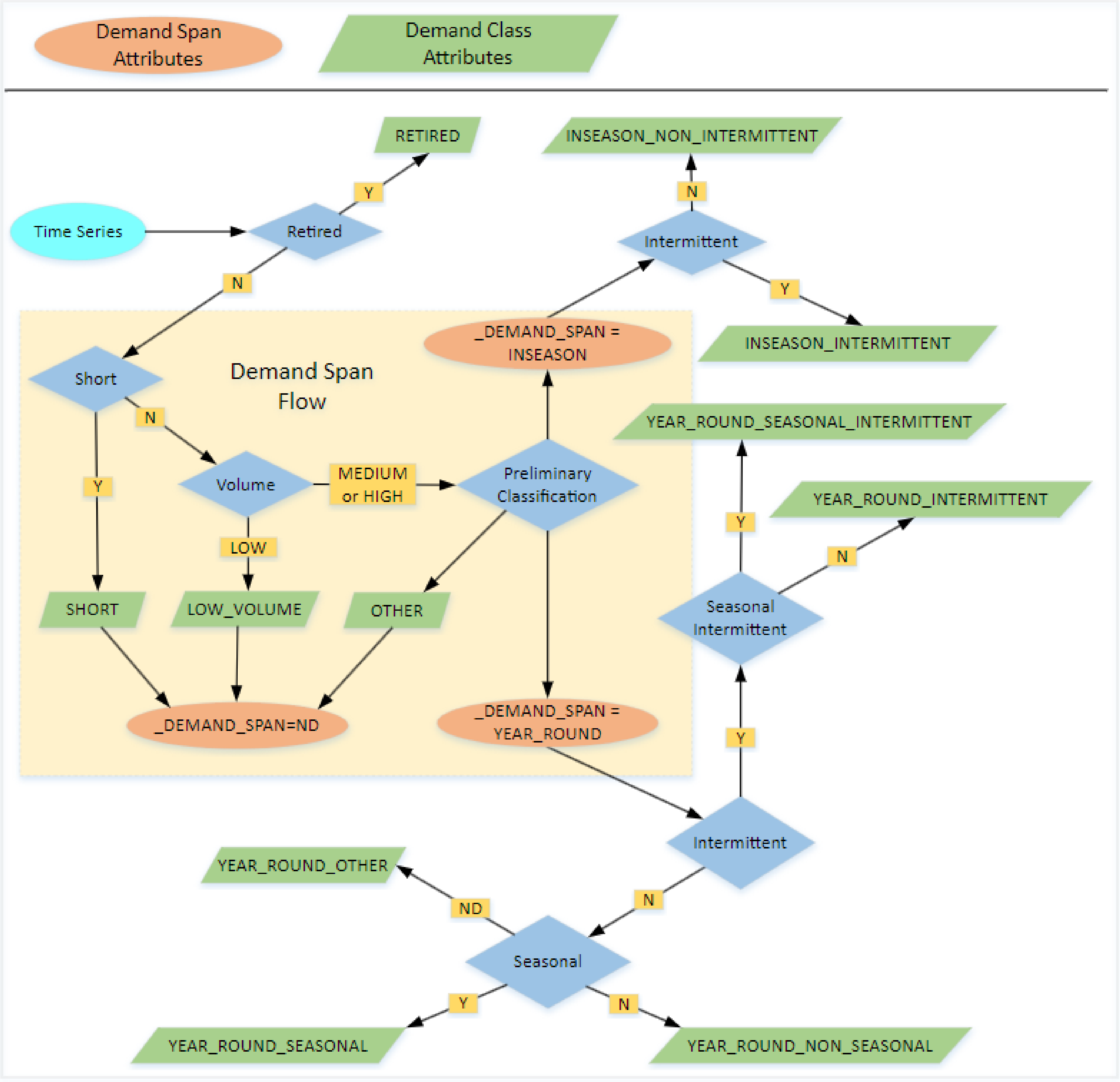
- ✓ Interrelated series
- ✓ Complex patterns
- ✗ Computationally intensive
- ✗ Expensive to scale in the cloud

Segmentation Methods

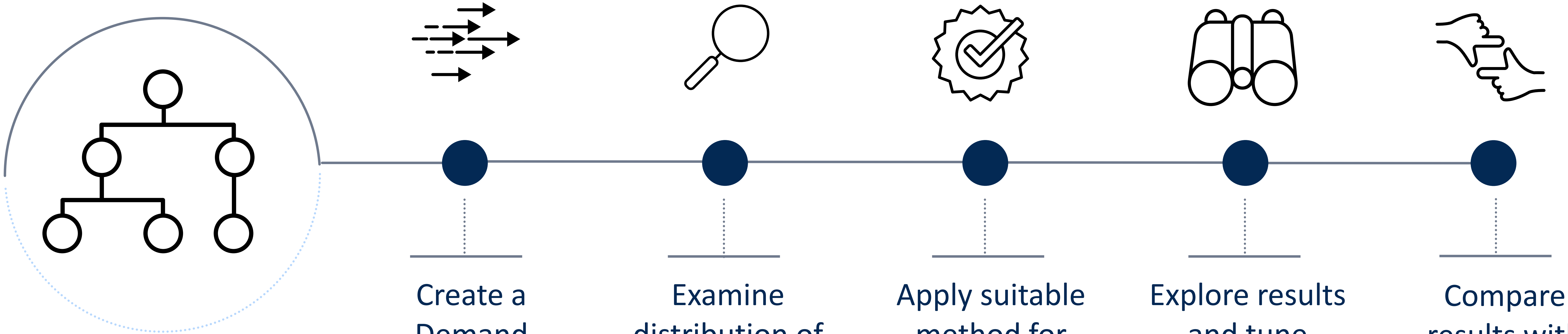
Optimize forecasting process in terms of accuracy and computational efficiency

SAS Visual Forecasting

- Externally Created Segments
- Demand Classification



Methodology Framework



Create a Demand Classification Pipeline

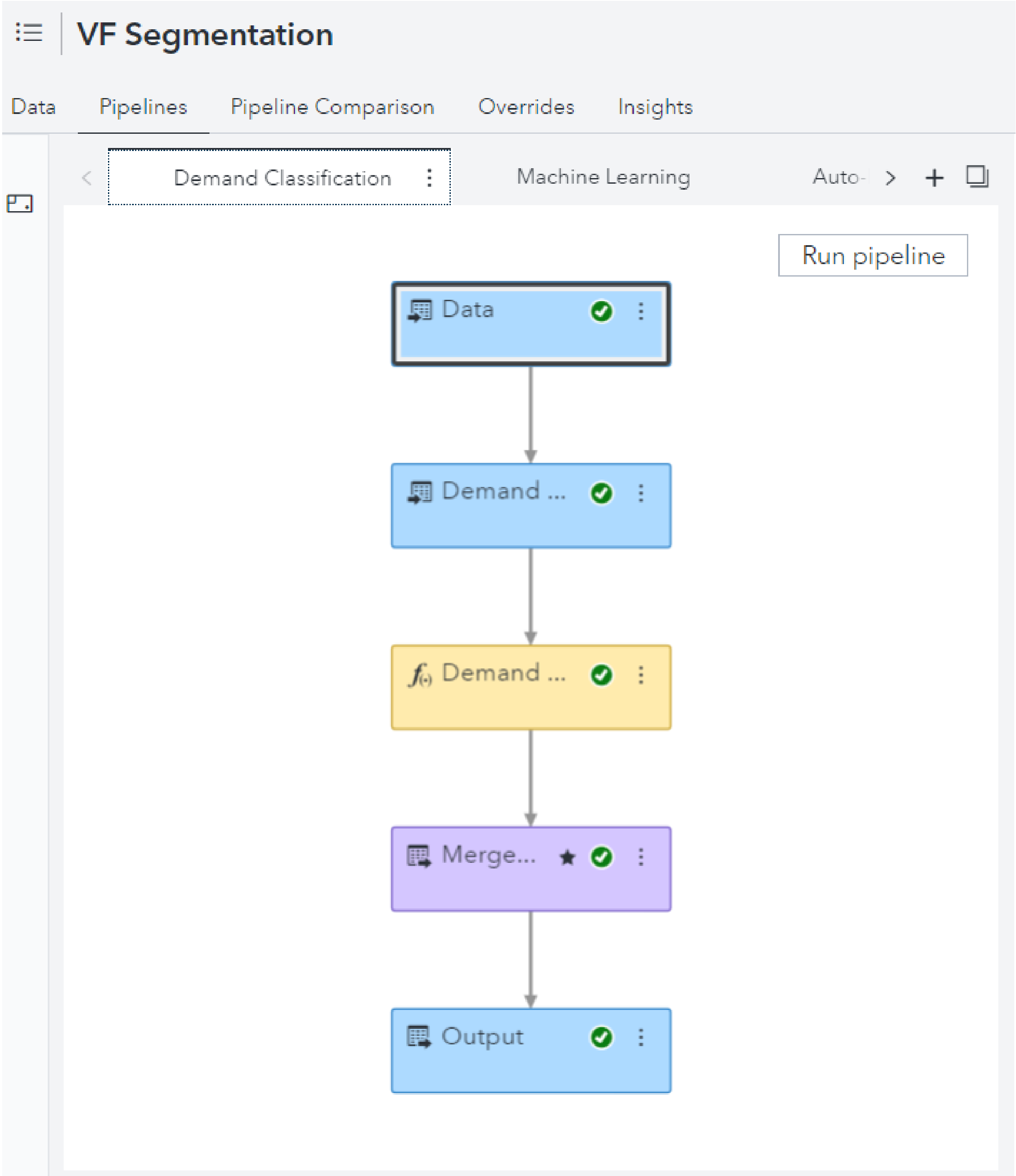
Examine distribution of time series to segments

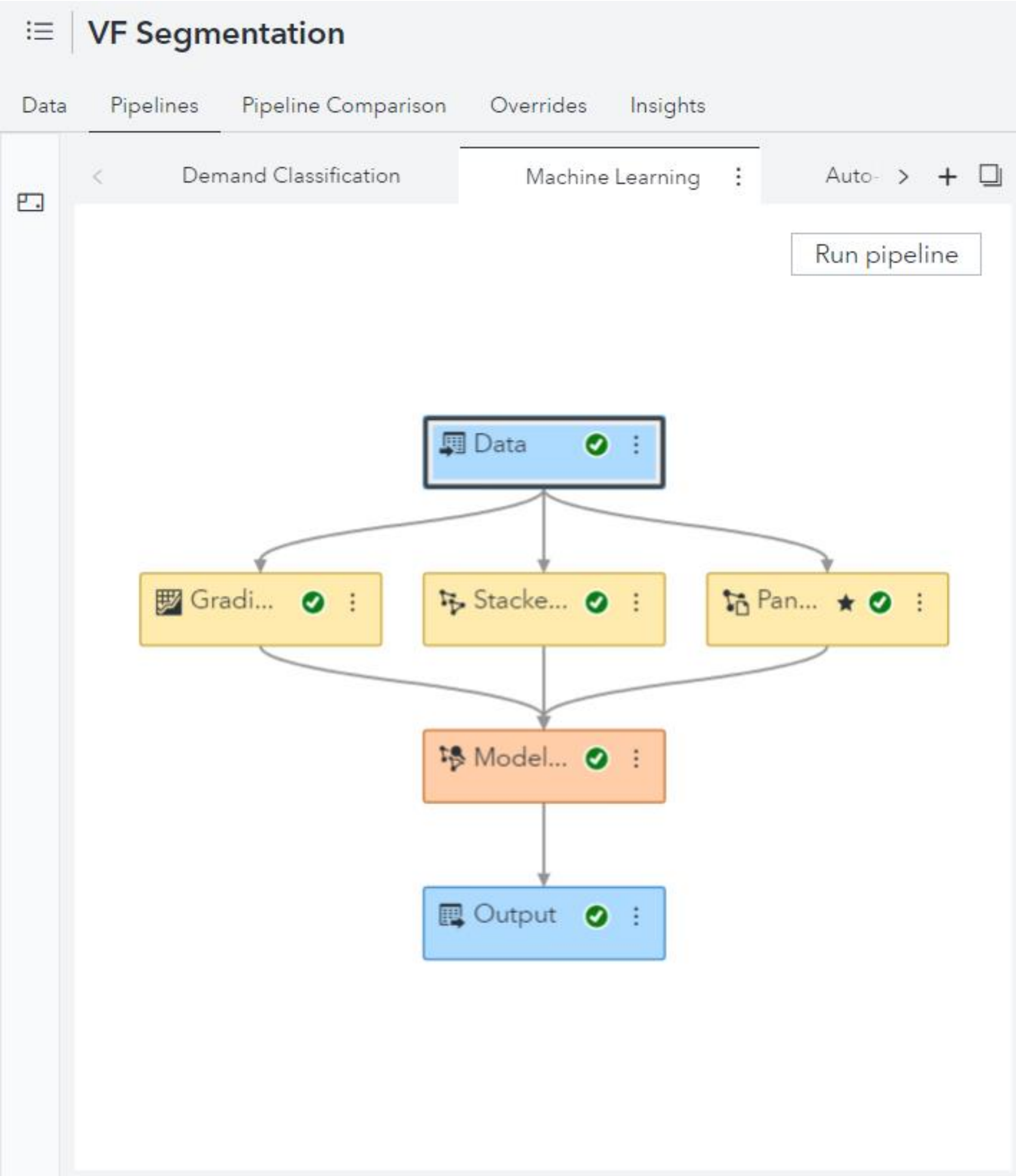
Apply suitable method for each segment

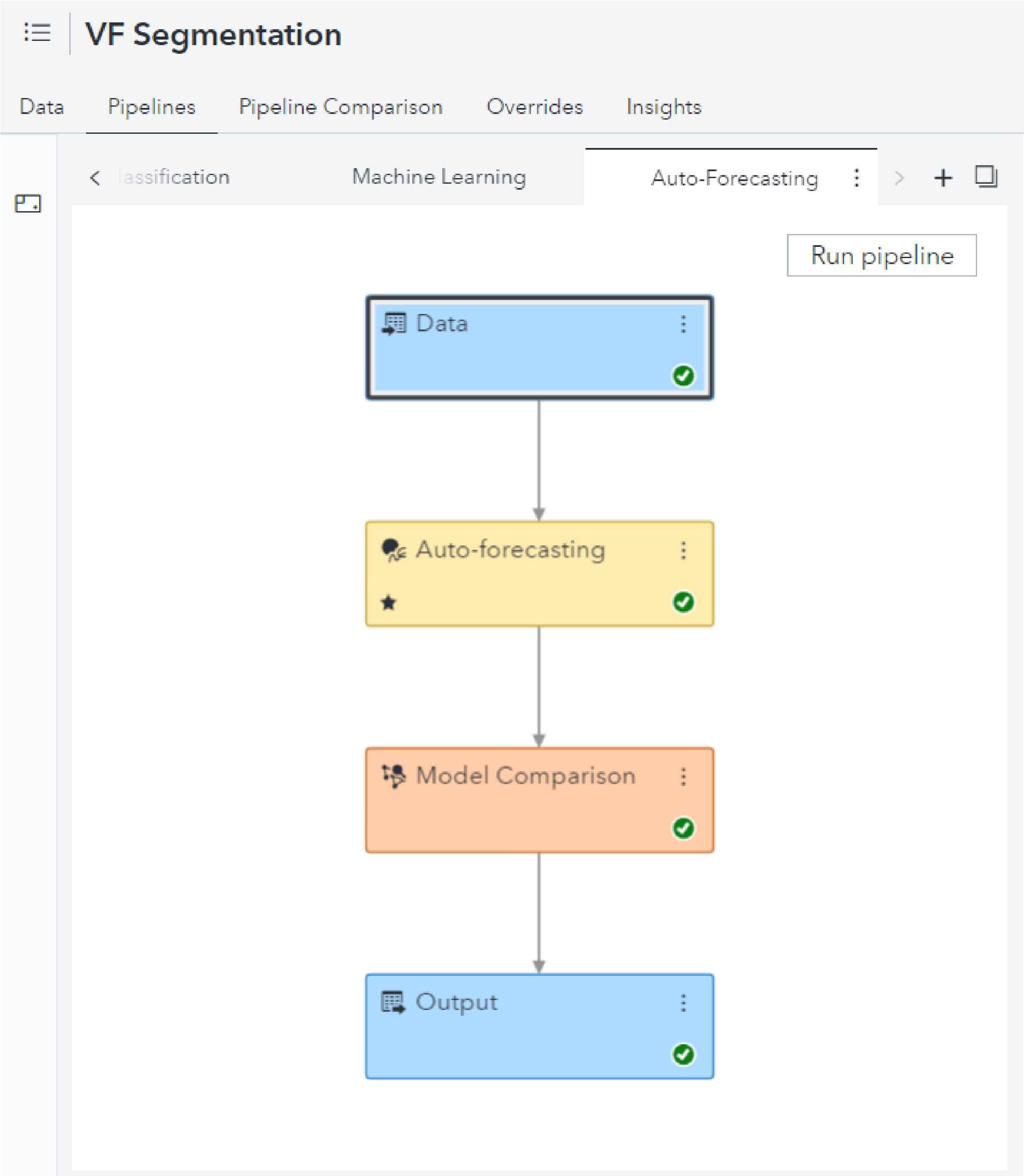
Explore results and tune algorithms

Compare results with competing pipelines

Pipelines’ Competition







Demand Classification Segments

Demand Classification Modeling

Segment Name	Number of Time Series	Pipeline	Status	WRMSE
INSEASON_INTERMITTENT	3	Regression Forecasting	✓	3.4770
INSEASON_NON_INTERMITTENT	0	Regression Forecasting	✓	0
LOW_VOLUME	0	Naive Forecasting	✓	0
OTHER	0	Naive (Moving Average) Forecasting	✓	0
RETIRED	25	Retired Forecasting	✓	0
SHORT	0	Naive (Moving Average) Forecasting	✓	0
YEAR_ROUND_INTERMITTENT	0	Auto-forecasting (Intermittent)	✓	0
YEAR_ROUND_NON_SEASONAL	519	Custom	✓	12,470.2467
YEAR_ROUND_OTHER	0	Naive (Moving Average) Forecasting	✓	0
YEAR_ROUND_SEASONAL	790	Custom	✓	12,228.2796
YEAR_ROUND_SEASONAL_INTERMITTENT	0	Temporal Aggregation Forecasting	✓	0

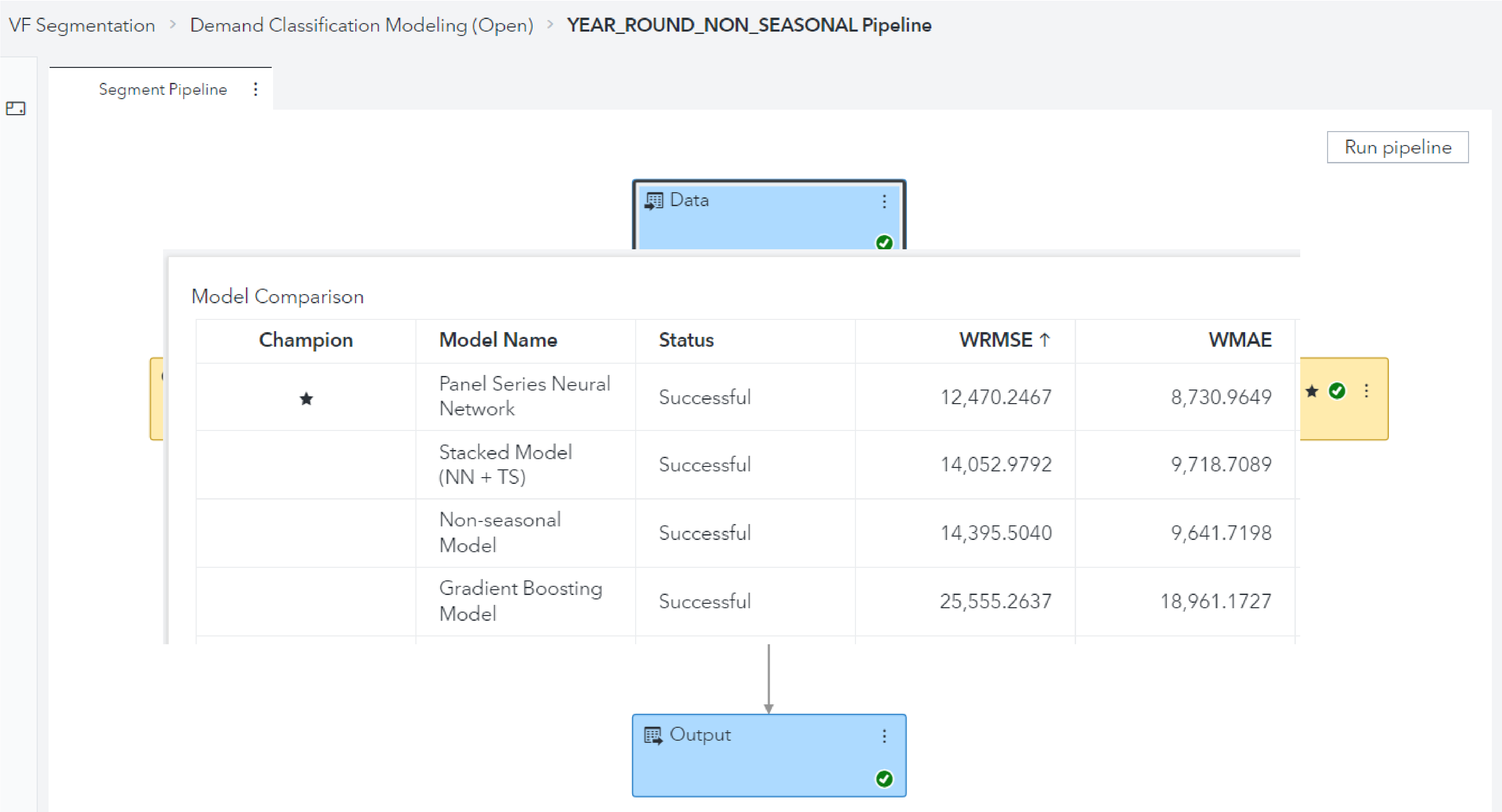
Nested Pipelines

Non-seasonal segment

13% accuracy improvement compared to statistical models

Seasonal segment

No significant difference in accuracy was observed



Champion	Model Name	Status	WRMSE ↑	WMAE
★	Panel Series Neural Network	Successful	12,470.2467	8,730.9649
	Stacked Model (NN + TS)	Successful	14,052.9792	9,718.7089
	Non-seasonal Model	Successful	14,395.5040	9,641.7198
	Gradient Boosting Model	Successful	25,555.2637	18,961.1727

Pipelines' Comparison

10% accuracy improvement vs ML only

20% accuracy improvement vs Statistical Methods only

VF Segmentation

Data | Pipelines | Pipeline Comparison | Overrides | Insights

Filter

Champion	Pipeline	WRMSE ↑	WMAE
★	Demand Classification	12,274.762	9,115.953
	Machine Learning	13,672.776	9,895.578
	Auto-Forecasting	15,376.808	11,011.275

Tips & Tricks

- Optimize for RMSE or MAE instead of MAPE
- The best model always depends on your data
- Stacked Model: experiment with number of hidden layers
- Panel Series Neural Networks: autotune is your frenemy
- ML methods are particularly useful when you’ve got external factors available

Summary

- Segmentation can give you the best of all worlds
- Forecasting process can be optimized both in terms of accuracy and computational efficiency to save on cloud costs
- You may achieve better results when applying ML methods to non-seasonal and complex data
- Traditional time-series models perform well when you've got seasonality and trend in the data

Resources

- Demand Classification Documentation
- How Will Generative AI Influence Forecasting Software
- Getting A Glimpse Into The Future Of Forecasting
- Neural Network–Based Forecasting Strategies in SAS® Viya®
- Measuring Forecasting Accuracy: Problems and Recommendations
- Forecast KPIs: RMSE, MAE, MAPE & Bias
- M5 accuracy competition: Results, findings, and conclusions

Thank you

Spiros.potamitis@sas.com