

What's new in SAS Forecasting?

Sara Vidal
Senior Solution Specialist Analytics, SAS



Agenda

- What's new in SAS/ETS?
- What's new in Forecast Server?
- What's new in SAS Forecasting solutions?



WHAT IS SAS/ETS?

- ETS:

Collection of **E**conometric and **T**ime **S**eries procedures

Some tools are explanatory (Econometric: Why do things happen?) and some are predictive (Forecasting: What will happen tomorrow?)

- **Types of tools in SAS/ETS**

Data Aggregation

Time Series Analysis

Time Series Econometrics

Cross-section Econometrics

Panel Data Econometrics

- BUT: not meant for large-scale forecasting



- Binary\Censored\Truncated Regression with Endogenous Regressors
- Bayesian Multivariate Models
- Compound Distribution Models for Aggregate Loss Modeling
- Vector autoregression computational enhancements
- New features for State-Space Modeling
- Access to Federal Reserve Economic Data (FRED)



LIMITED DEPENDENT VARIABLE MODELS WITH ENDOGENOUS REGRESSORS

Endogeneity arises from:

- Simultaneity (feedback with Y and X)
- Measurement Error in the regressors
- Omitted Variables (Influence Y and correlated with included regressors)

QLIM now estimates:

- Probit\Logit with Endogenous Regressors
- Censored\Truncated Regression with Endogenous Regressors

– Endogenous regressors may be:

- Binary
- Continuous
- Censored or Truncated.

- Full Information Maximum Likelihood (FIML) estimates



PROBIT MODEL WITH A BINARY ENDOGENOUS EXPLANATORY VARIABLE

Model:

$$\begin{aligned}
 y_{1i}^* &= \alpha_1 y_{2i} + \beta_1 z_{1i} + \beta_2 z_{2i} + u_i \\
 y_{2i}^* &= \pi_1 z_{1i} + \pi_2 z_{2i} + \pi_3 z_{3i} + \pi_4 z_{4i} + \varepsilon_i \\
 y_{1i} &= \begin{cases} 1 & \text{if } y_{1i}^* > 0 \\ 0 & \text{if } y_{1i}^* \leq 0 \end{cases} \\
 y_{2i} &= \begin{cases} 1 & \text{if } y_{2i}^* > 0 \\ 0 & \text{if } y_{2i}^* \leq 0 \end{cases}
 \end{aligned}$$

Use case: Probability of loan default as a function of change in working status (firing)

Syntax:

```

proc qlim data=a;
  model y1 = y2 z1 z2 / discrete endotest(y2);
  model y2 = z1 z2 z3 z4 / discrete;
run;

```

Do we need to treat y2 as endogenous?

Endogenous Regressor

Instruments



BAYESIAN MULTIVARIATE REGRESSION

What if you only have a limited data set to estimate a price response?

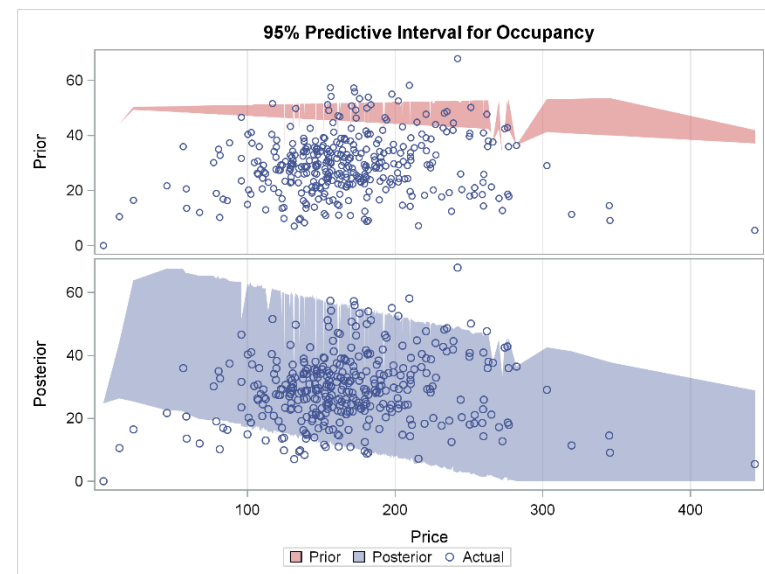
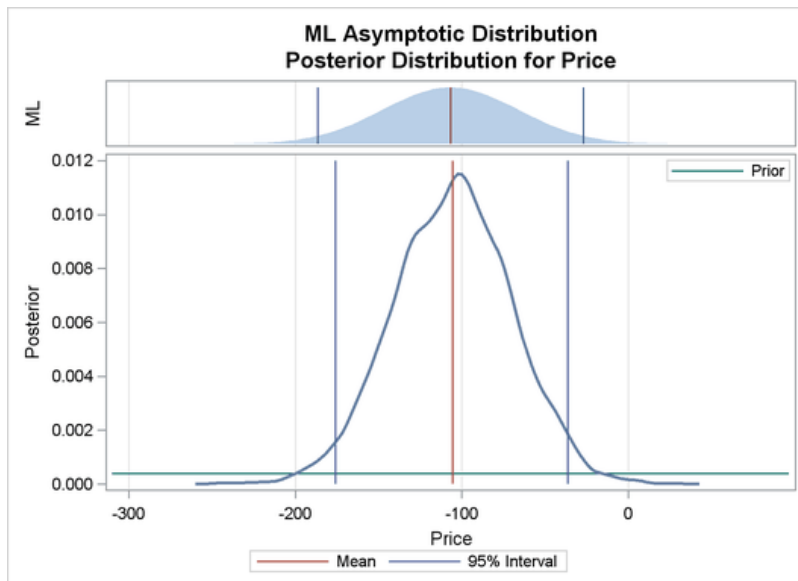
The QLIM procedure now enables you to supplement the data with information from other hotel properties:

Property A:	$E(\beta) = -0.08$	$SD(\beta) = 0.02$
Property C:	$E(\beta) = -0.014$	$SD(\beta) = 0.004$
Property D:	$E(\beta) = -0.018$	$SD(\beta) = 0.005 \rightarrow$
	mean = -0.037 variance 0.00018	

```
PROC QLIM data = PropertyB;
  MODEL Occupancy = Price Weather / censored(lb=0 ub=up);
  BAYES nmc=30000 nbi=10000 ntrds=1;
  PRIOR Occupancy.Price ~ normal(mean=-0.037, var=0.00018);
  MODEL Utilities = Weather Avg_Stay
RUN;
```



BAYESIAN DIAGNOSTIC TOOLS



Proc HPCDM

COMPOUND DISTRIBUTION MODELS FOR AGGREGATE LOSS MODELING: INSURANCE EXAMPLE

- Assume:
 - I have data on loss numbers per policy holder or line of business per year.
 - Let's say it follows a Poisson distribution.
 - I have data on magnitude of loss per policy holder or line of business.
 - Let's say these follow a gamma distribution.
- How do I combine this information to estimate the distribution of aggregate loss possibilities in a given year?



Proc HPCDM

COMPOUND DISTRIBUTION MODELS

- Let N represent the frequency random variable for the number of loss events that occur in the time period of interest. Let X represent the severity random variable for the magnitude of one loss event. Then, the aggregate loss S is defined as:

$$S = \sum_{j=1}^N X_j$$

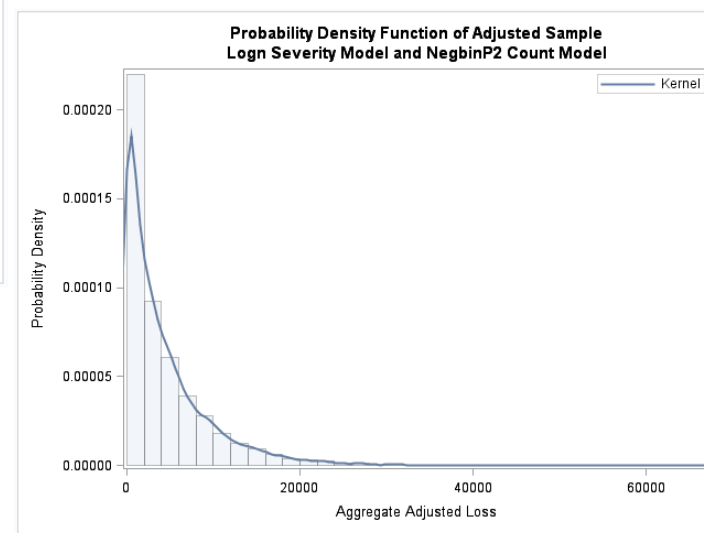
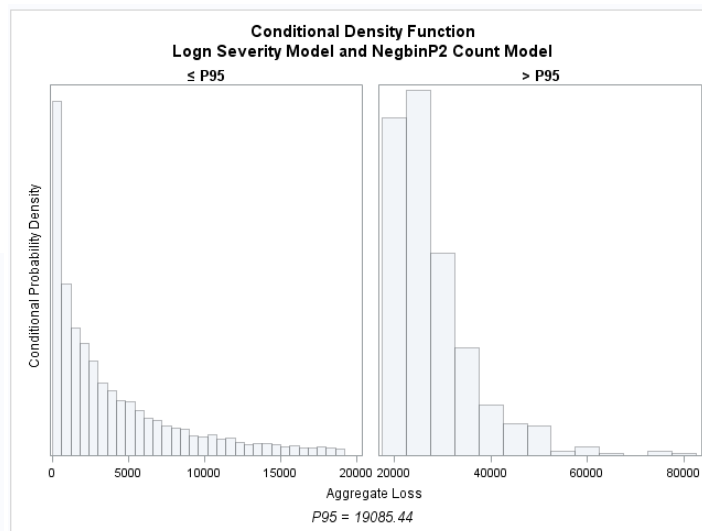
- PROC HPCDM accepts
 - the severity model of X as estimated by the SEVERITY procedure (with regression effects).
 - The frequency model of N as estimated by the COUNTREG procedure (with regression effects).
- Multithreaded Monte Carlo Simulation to form aggregate loss distribution



PROC HPCDM

GRAPHICS

Adjusted Sample Percentile Perturbation Analysis		
Percentile	Estimate	Standard Error
0	0	0
1	0	0
5	0	0
25	604.60948	126.17884
50	2568.4	270.29467
75	6126.0	579.35289
95	13865.7	1135.7
99	20869.0	896.03211
99.5	22880.5	1118.9
Number of Perturbed Samples = 50		
Size of Each Sample = 10000		



PROC VARMAX

ENHANCEMENTS

- Huge performance and scalability improvements for models estimated through the method of maximum likelihood. In some test cases, thousands of times faster!
- BOUND: specify inequality constraints on parameters
- INITIAL: initial values for parameters
- Multivariate DCC GARCH
- Four new forms of univariate GARCH models
EGARCH, PGARCH, QGARCH, TGARCH



- Traditional time series tools require time series to have the same start and end dates and not to have missing observations
- The SSM procedure in SAS/ETS allows for multivariate time series modeling of differing lengths
 - Common trends with diagnostic tools
 - Common seasons with diagnostic tools
 - Modeling language to create variables on the fly
- New in 13.1:
 - Break checks for any component (CHECKBREAK)
 - State transition equation now permits inclusion of regression effects



- UAH contains monthly satellite global temperature readings starting December 1978

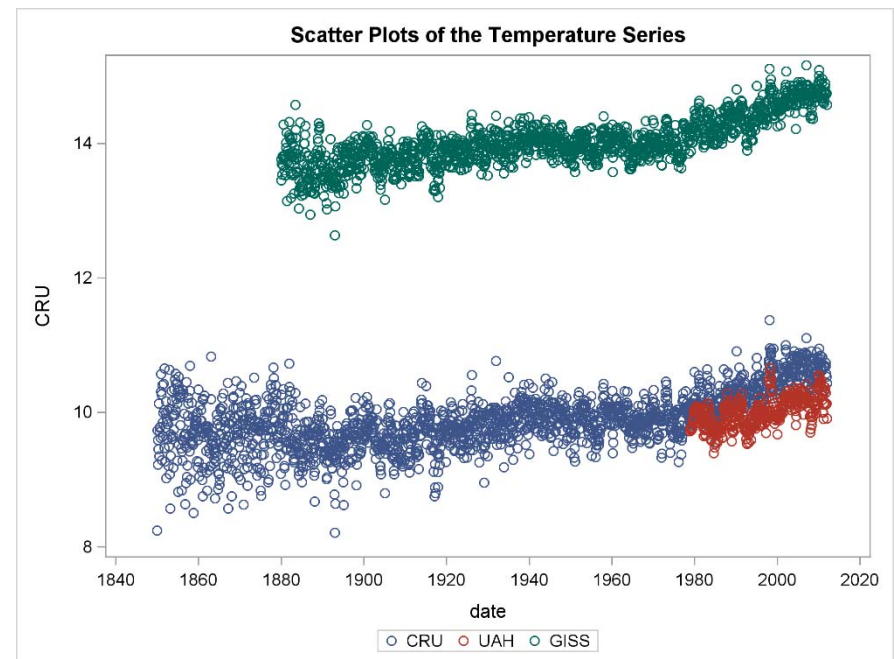
- CRU contains monthly temperature data starting January 1850

- Could be different parts of world, but all have seasons and *might* have monthly temperature
- GISS contains monthly temperature data starting January 1880

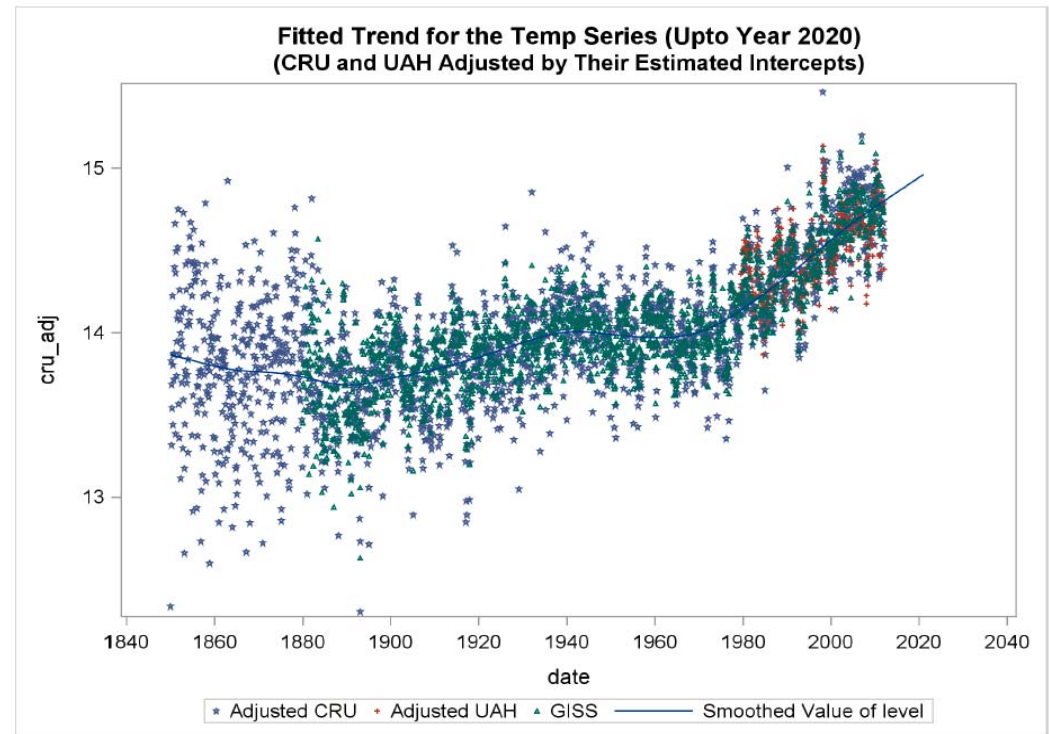
$$GISS_t = \mu_t + a \zeta_t + a r_1 \epsilon_{1t}$$

$$CRU_t = \beta_{cru} + \mu_t + a \zeta_t + a \epsilon_{2t}$$

$$UAH_t = \beta_{uah} + \mu_t + a \zeta_t + a r_3 \epsilon_{3t}$$

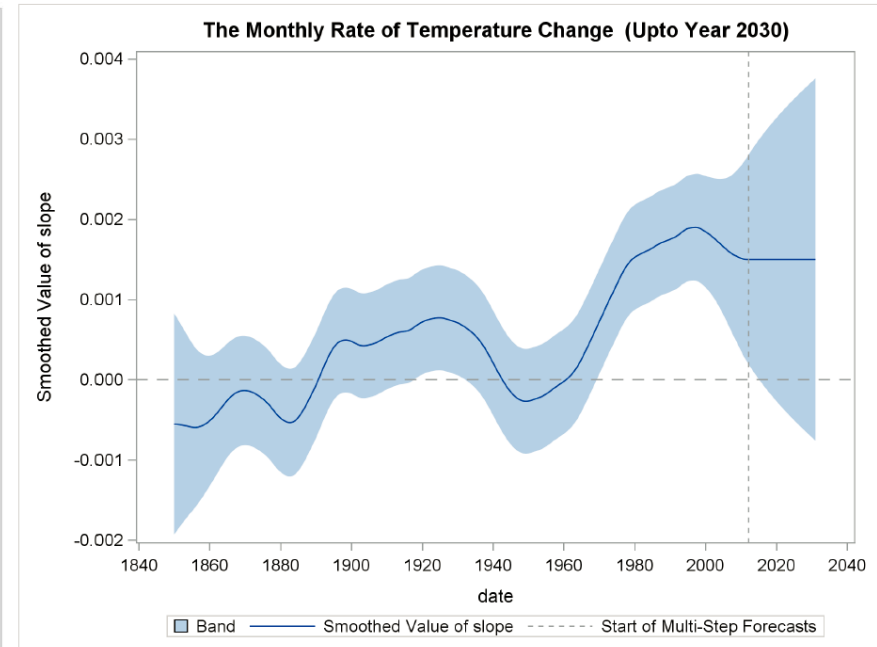
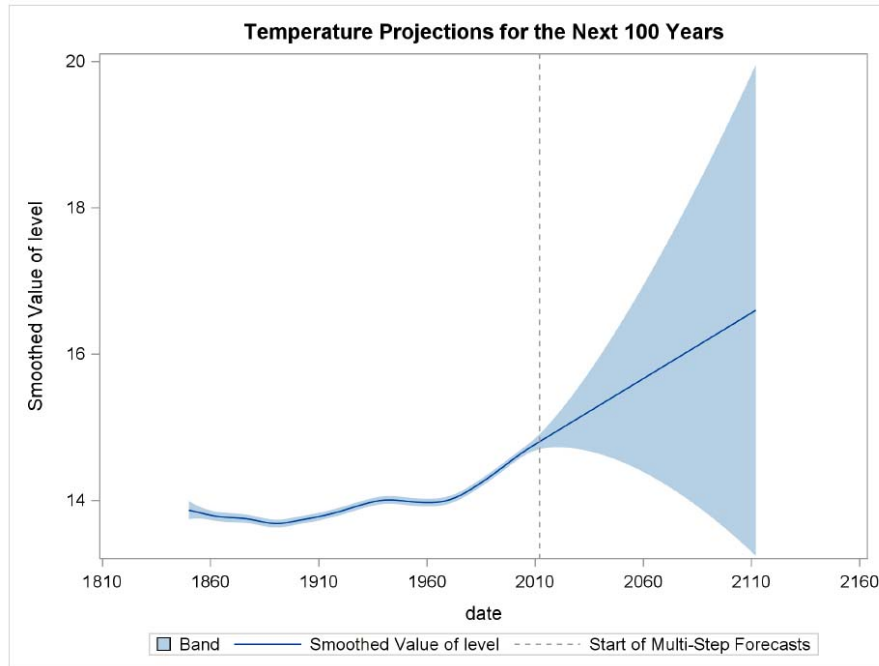


- In effect, we are using information from all three time series to more precisely estimate the long-term trend
- Removing seasonal effects, differing intercepts



New Econometrics

STATE-SPACE MODELS



Warming temperatures, but the growth has not accelerated over the past ten years

SASEFRED

ACCESS TO FEDERAL RESERVE ECONOMIC DATA

- Enables SAS users to retrieve economic data from the FRED website
- FRED databases contain more than 61,000 economic data time series from 48 national and international sources
- Time series are updated at annual, quarterly, monthly, weekly, and daily intervals
- FREE!!!

```
libname fred sasefred "%sysget(FRED)" OUTXML=fredex01 AUTOMAP=replace
MAPREF=MyMap XMLMAP="%sysget(FRED)fredex01.map"
APIKEY= 'XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX'
IDLIST= 'bopxgs,bopmgs'
START= '1997-01-01' END= '2011-01-01' FREQ= 'a'
OUTPUT=1
AGG= 'avg'
FORMAT=xml;
```



- AUTOREG: multiple structural change tests Bai and Perron (1998)
- COUNTREG supports STORE, RESTORE and SCORE
- New procedures: HPCOPULA, HPCDM, HPPANEL
- PANEL: system GMM estimator by Blundell and Bond (1998)
- SASEXFSD: Uses FactSet FASTFetch web service
- SEVERITY: Support for scoring functions
- UCM: bootstrap-based procedure for computing the standard errors of the series and component forecasts



Agenda

- What's new in SAS/ETS?
- What's new in Forecast Server?
- What's new in SAS Forecasting solutions?



What is SAS Forecast Server?

- **SAS Forecast Server** - client-server application that enables organizations to plan more effectively for the future by **generating large quantities of high-quality forecasts quickly and automatically**. **SAS Forecast Studio** is the client component that provides a graphical interface to the SAS Forecast Server Procedures.
- SAS Forecast Server 13.1 runs on the first maintenance release of SAS 9.4. SAS Forecast Server is now integrated with SAS Time Series Studio, enabling you analyze and structure your time-stamped data. You can export data from SAS Time Series Studio and import it into SAS Forecast Studio.
- SAS Forecast Server 13.1 supports integrated Windows authentication, support for clustered metadata servers, and support for clustered middle-tier servers.



- SAS® Forecast Server 13.1 provides many key enhancements and introduces new functionality such as multithreading
- Available with SAS 9.4M1, this release adds important new functionality to SAS forecasting software.
- SAS Forecast Server automatically diagnoses time series to find appropriate models, generates forecasts, and reconciles forecasts in a variety of ways



- Multithreading refers to concurrent execution of multiple threads.
- When multithreading is possible, it can produce greater performance gains than those of sequential (single-threaded) execution.
- Both the HPFENGINE procedure and the HPFDIAGNOSE procedure are now multithreaded. They can use multiple threads in certain phases of their computations within each BY group that they process as follows:



- PROC HPFDIAGNOSE can make use of threads in the following computational processes:
 - Model identification for each model family
 - Automatic input variable and calendar events selection
- PROC HPFENGINE can make use of threads in the following computational processes:
 - FORECAST statement variables
 - Automatic forecast model selection



SAS Forecast Server 13.1

GRID COMPUTING

- For organizations that face a vast number of time series or many different hierarchies, it is often desirable to model and forecast each hierarchical time series in parallel for computational efficiency.
- SAS Forecast Server 13.1 integrates seamlessly with SAS® Grid Manager, a flexible, centrally managed computing environment for workload balancing, high availability, and faster processing.
- SAS Grid Manager requires a separate license.



SAS Forecast Server 13.1

ADMINISTRATIVE ENHANCEMENTS

- SAS Forecast Server 13.1 now supports the following:
 - Integrated Windows authentication
 - Clustered metadata servers
 - Clustered middle-tier servers
 - SAS Workspace Servers in grid computing environments



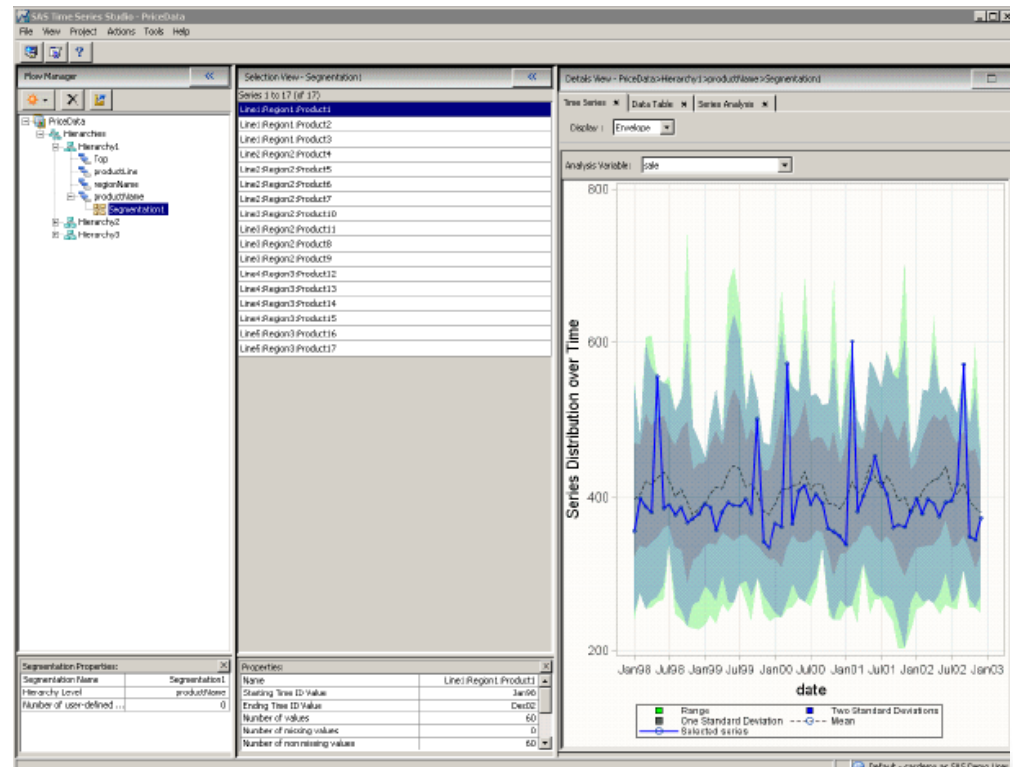
- Many organizations collect large amounts of transactional and time series data, such as sales histories, inventory histories, customer transactions, insurance claim histories, and Internet data.
- In working with these data, analysts often need to structure the time series data into hierarchical time series at particular frequencies to enhance their understanding of the data and to improve the accuracy of their analyses.
- Faced with large amounts of data, analysts often need to subset their time series data to support different analysis on different subsets of their data.



SAS Forecast Server 13.1

SAS TIME SERIES STUDIO

Using SAS Time Series Studio, you can subset the data by using hierarchical queries, graphical queries, parametric queries, or manual selection.



SAS Forecast Server 13.1

INTEGRATION WITH MICROSOFT EXCEL

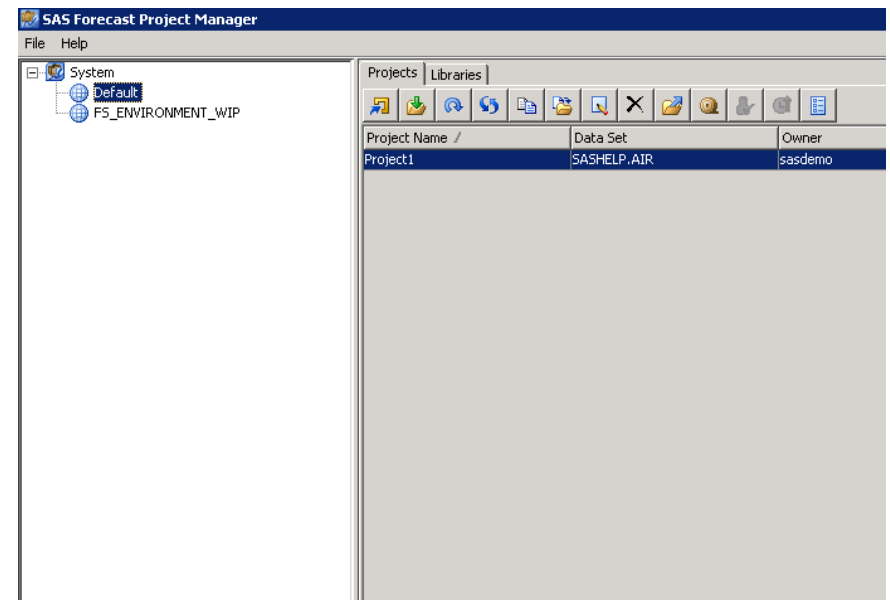
- SAS Forecast Server 13.1 includes access to specialized tasks that enable you to open and create SAS® Forecast Studio projects from Microsoft Excel spreadsheets.
- It also provides access to dedicated override tasks, enabling you to make changes to forecast values in SAS Forecast Studio projects from Microsoft Excel.



SAS Forecast Server 13.1

SAS FORECAST PROJECT MANAGER

- Use the SAS® Forecast Project Manager to manage your forecasting projects, forecasting environments, and environment libraries. It is a stand-alone management client designed to require only HTTP-based communication.
- The SAS Forecast Project Manager enables you to configure a SAS Forecast Server project. It provides all the management capabilities available in the SAS Forecast Server Plug-in for SAS® Management Console.



- Preparing the time series data is one of the most important activities in time series analysis, forecasting, and also data mining.
- Data preparation requires a great deal of flexibility to handle various issues of concern, and time series transformations and dimension reductions can be quite complex.
- The new TIMEDATA procedure deals with the challenges of iterative and recursive transformations and reductions, minimizing the number of passes through the data.
- The TIMEDATA procedure is used by both SAS Forecast Studio and SAS Time Series Studio and is now available for batch usage.



SAS Forecast Server 13.1

SAS® FORECAST SERVER PROCEDURES

- SAS Forecast Server Procedures is the new name for the product formerly called SAS® High-Performance Forecasting.
- This represents a change in name only and does not require any change in how you use the software.
- SAS High-Performance Forecasting was renamed SAS Forecast Server Procedures in order to distinguish its functionality from other SAS High-Performance Analytics products.



SAS Forecast Server 13.1

OTHER ENHANCEMENTS TO SAS FORECAST SERVER

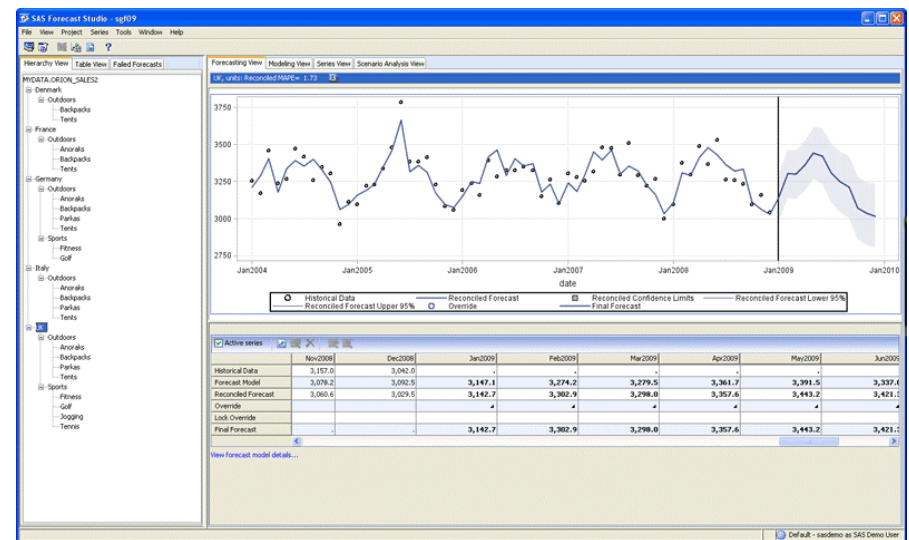
- Ability to change ODS format and execution server in running a report or stored process in SAS Forecast Studio
- Ability to create an output forecast data set for independent variables in SAS Forecast Studio
- Enhancements to %FSCREATE and %FSREPORT batch macros



SAS Forecast Server 13.1

SAS® FORECASTING FOR DESKTOP

- SAS Forecasting for Desktop is a desktop product that provides automation and productivity gains for smaller-scale forecasting tasks.
- No remote servers are required when you use SAS Forecasting for Desktop, so all work is performed and saved locally.
- SAS Forecasting for Desktop can be licensed separately.



Note: SAS Forecast Server is recommended if you have large amounts of data and large-scale forecasting projects.

Agenda

- What's new in SAS/ETS?
- What's new in Forecast Server?
- What's new in SAS Forecasting solutions?



SAS Forecasting solutions

(based on Forecast Server)

- SAS Demand-Driven Forecasting
- SAS Forecasting for SAP APO
- SAS Warranty Reserve Forecasting **NEW!**
- SAS Service Operations Optimization **NEW!**
- SAS Service Parts Optimization
- SAS Revenue Optimization Suite
- SAS Demand Forecasting for Retail
- SAS Retail Forecasting



SAS® Forum Switzerland 2014

make connections • share ideas • be inspired

Sara Vidal
sara.vidal@sas.com



Copyright © 2014, SAS Institute Inc. All rights reserved.

