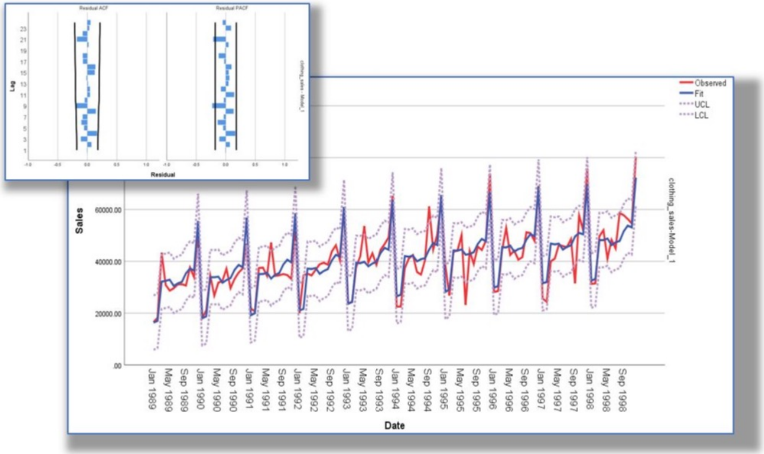


Forecasting made easy with SPSS Statistics

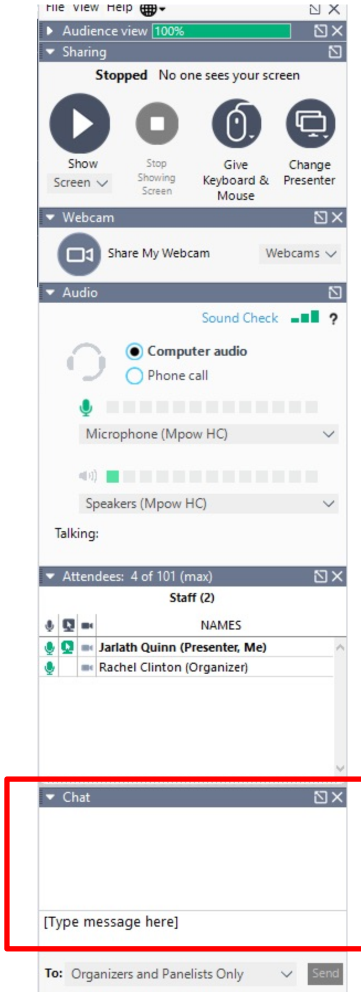
Jarlath Quinn

www.sv-europe.com

A SELECT INTERNATIONAL COMPANY

FAQ's

- Is this session being recorded? Yes
- Can I get a copy of the slides? Yes, we'll email a PDF copy to you after the session has ended.
- Can we arrange a re-run for colleagues? Yes, just ask us.
- How can I ask questions? All lines are muted so please use the chat facility – if we run out of time we will follow up with you.



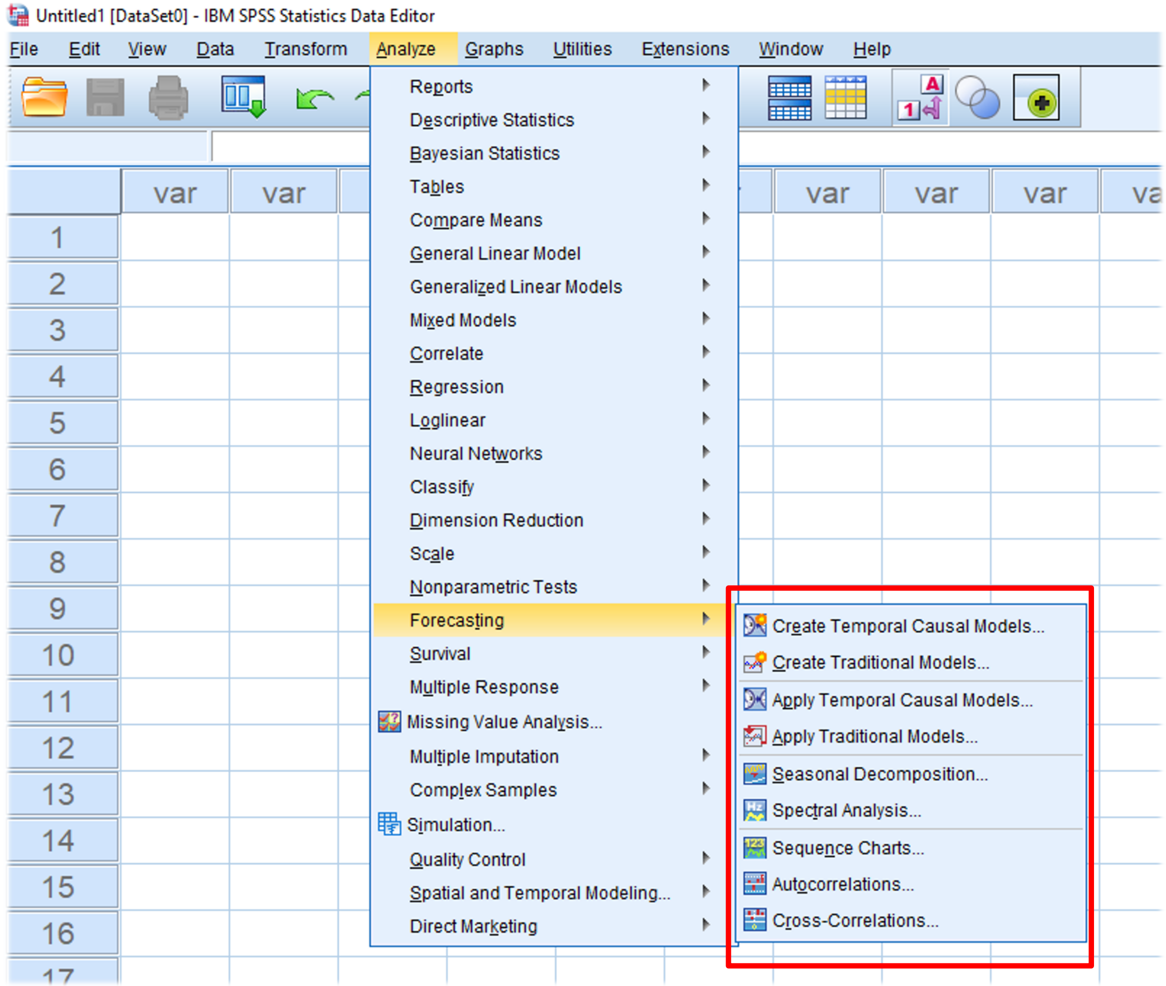
- Premier accredited partner to IBM and Predictive Solutions specialising in advanced analytics & big data technologies
- Work with open source technologies (R, Python, Spark etc.)
- Team each has 15 to 30 years of experience working in the advanced and predictive analytics industry

- Deep experience of applied advanced analytics applications across sectors
  - Retail
  - Gaming
  - Utilities
  - Insurance
  - Telecommunications
  - Media
  - FMCG

IBM SPSS Statistics v27 Base & Associated Modules

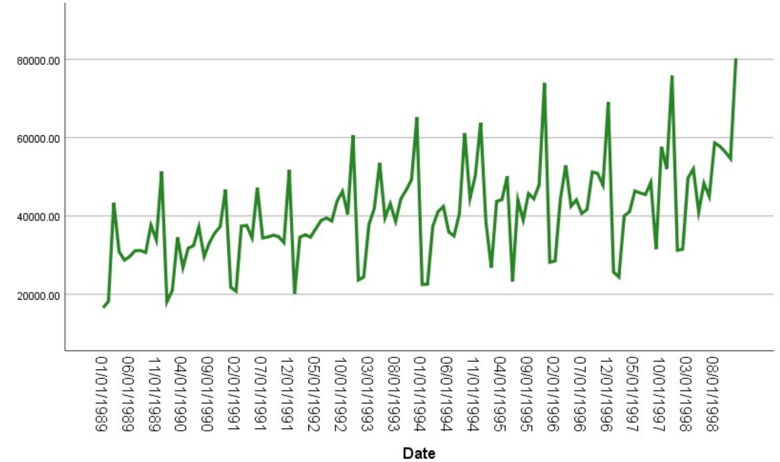
* You will need the SPSS Statistics Forecasting Module

- **If you can see sub-menu, then it is installed on your copy of SPSS Statistics**



Agenda

- The principles of Time Series forecasting
- Visualising time series
- Smoothing techniques
- Exponential smoothing methods
- Interpreting output and model fit
- Using predictor fields to improve accuracy
- Generating forecasts



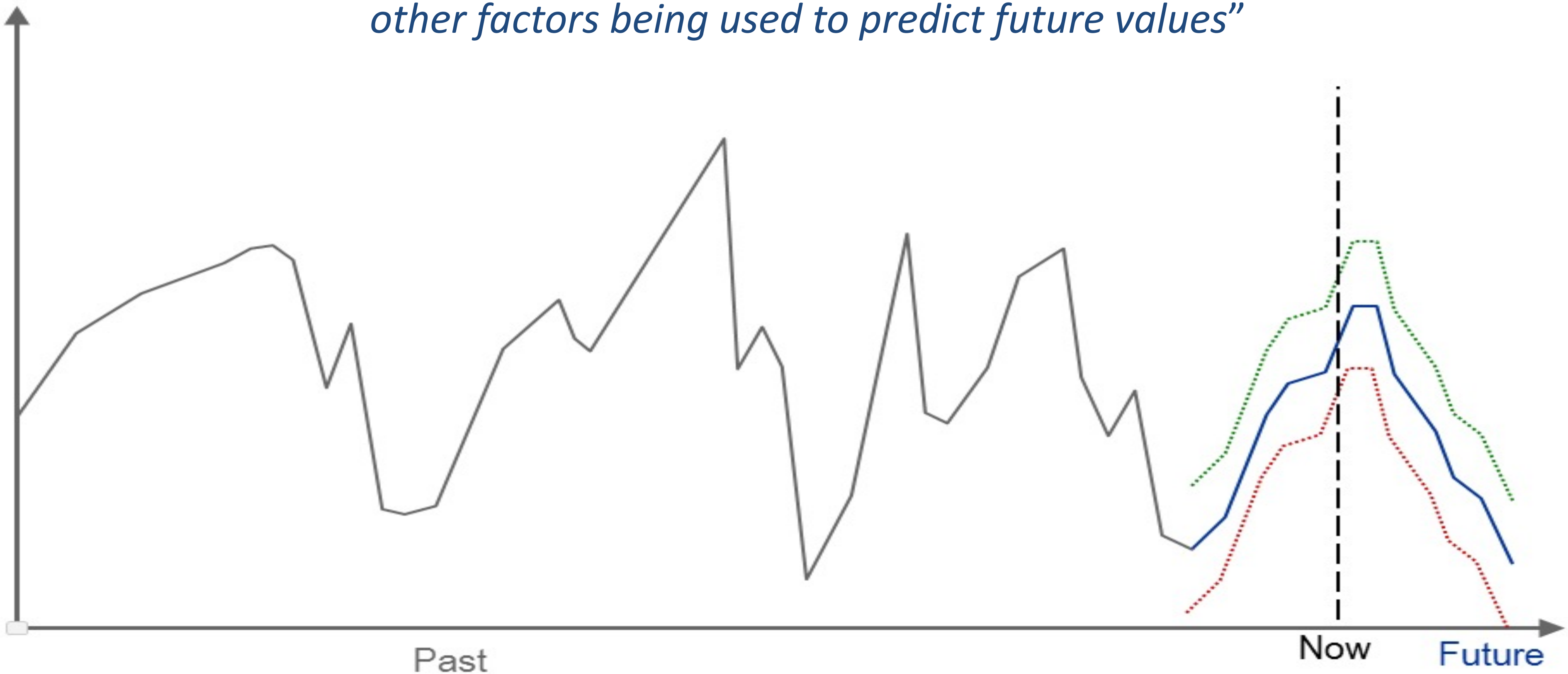
The Principles of Time Series Forecasting

What is Time Series?

- A 'Time Series' is simply a series of values of a quantity collected over a specific time period, often with equal intervals between them
- Examples of time series include:
  - Airline passenger numbers for a particular country over the last 40 years
  - Daily website hits during a three-month period
  - Hourly traffic volumes over the course of a week

Time Series Forecasting

" 'Things' that are observed repeatedly over time, with past values and other factors being used to predict future values"



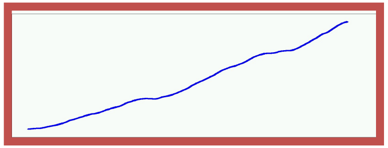
What is Time Series?

- Time series analysis is based on the principle that the past provides a model for the future
- Time series forecasting models often don’t require predictor/independent variables
- The goal of time series analysis is to separate the random variability (‘noise’) from the variability that can be explained
- A single time series may have several elements that enable effective forecasting

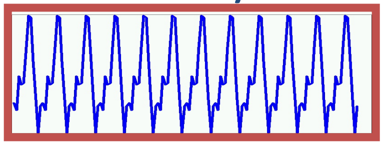
What's *in* a 'Time Series'?

Time Series

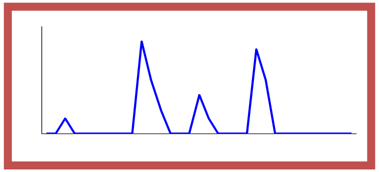
Trend



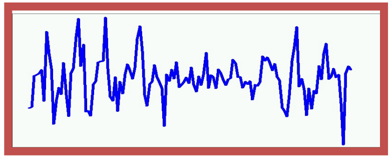
Seasonality



Events

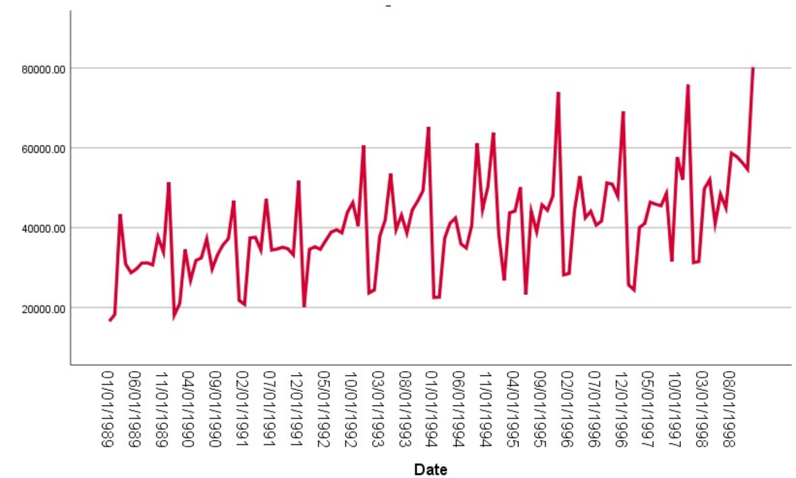


Noise



Considerations

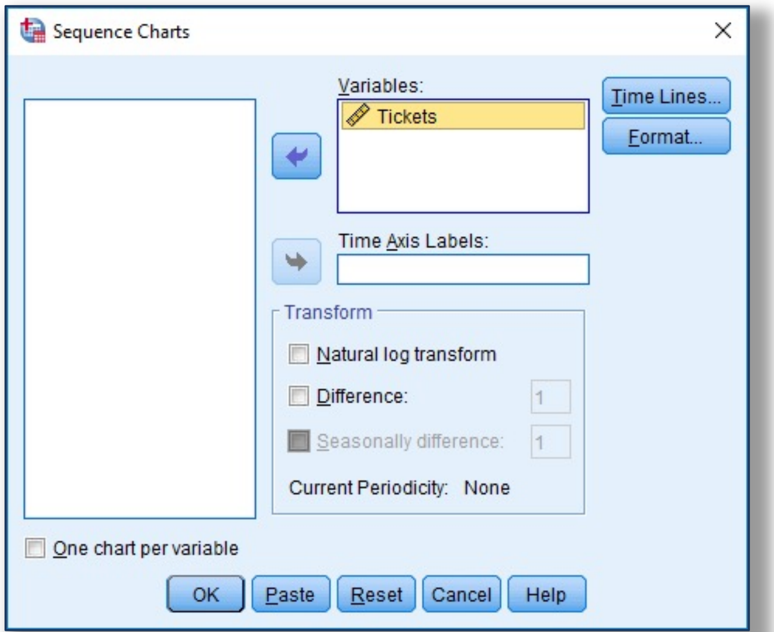
- Are the data points regularly spaced?
- How far can I forecast into the future?
- What is the periodicity?
- Should I exclude data?
- Do I have other predictor fields?
- Are there special events that need to be marked?
- Do I need to forecast more than one series?

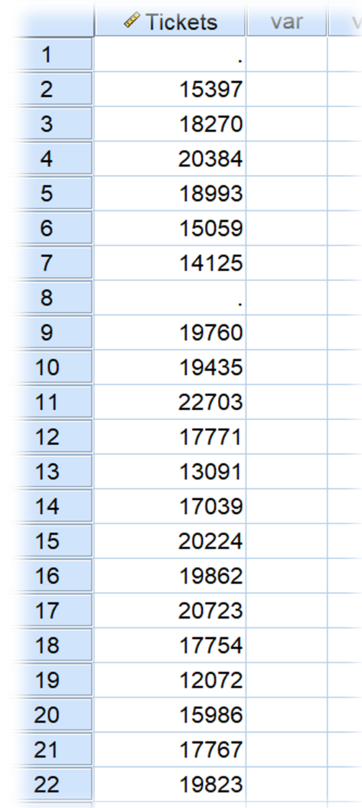


Visualising Time Series

Visualising Time Series

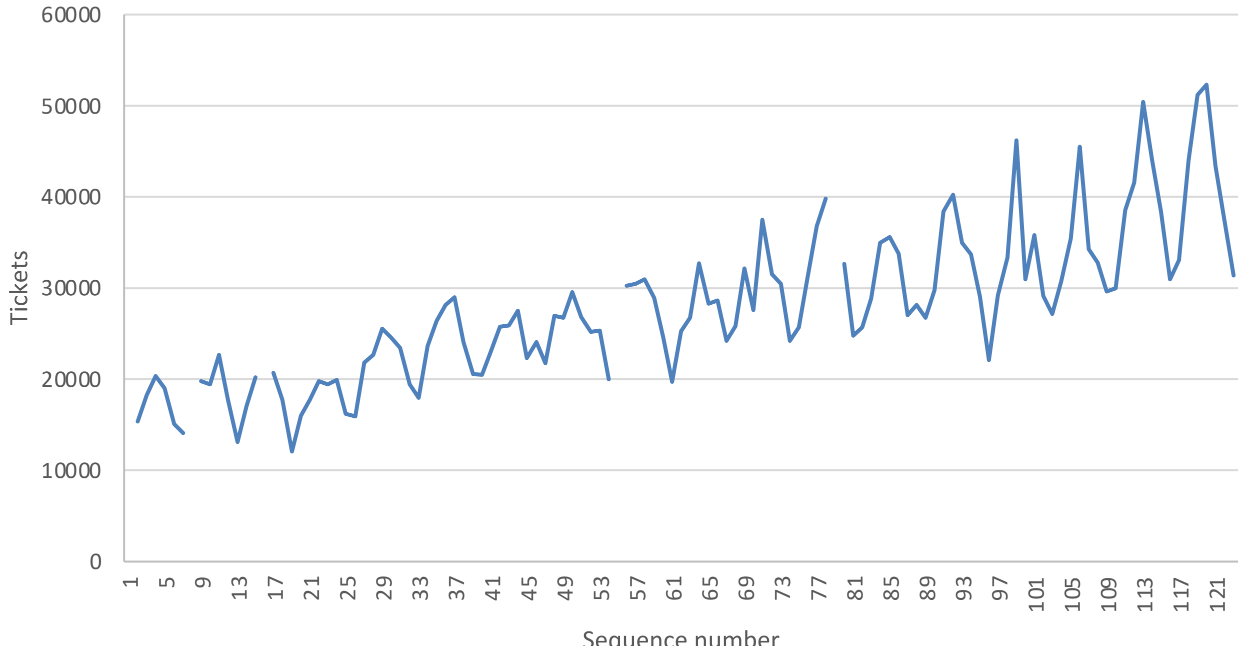
- Using the dataset **Ticket Sales Missing.sav**
- We can visualise the data series with **Sequence Charts**





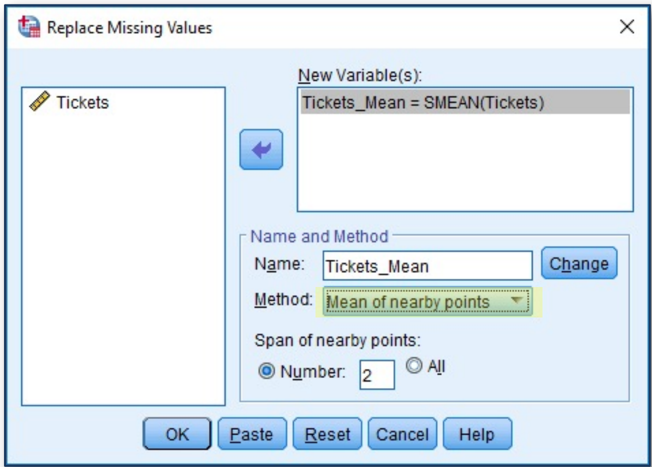
Visualising Time Series

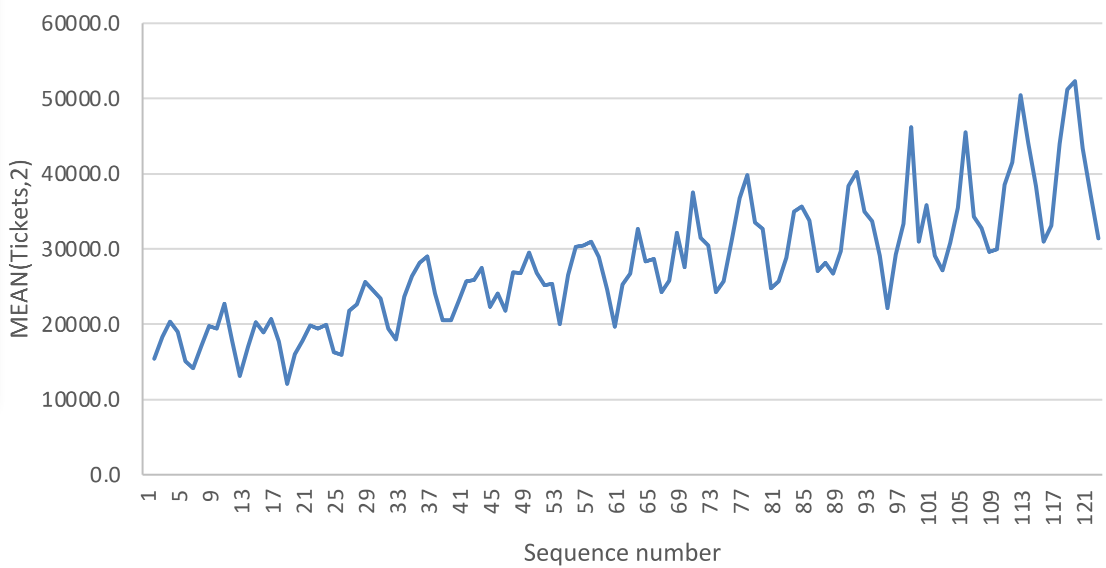
- A few things to note:
  - You don't *need* a variable for the time axis
  - There are gaps in the series i.e. missing data

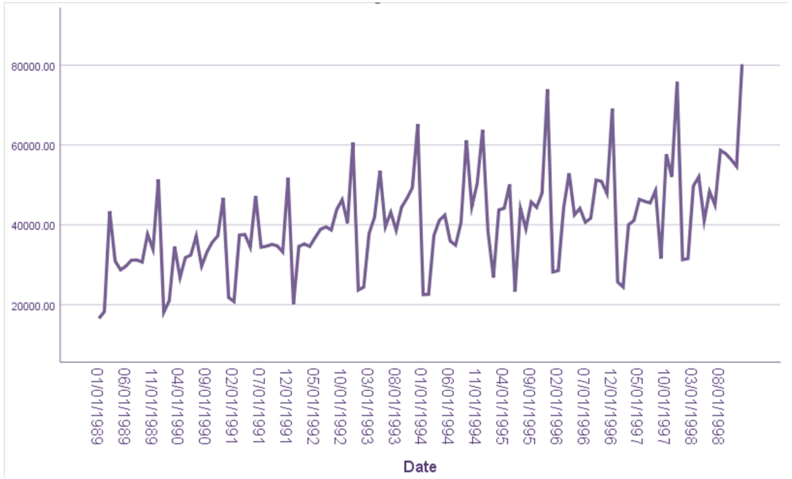


Visualising Time Series

- *One way* to deal with missing values in a sequence
- Using the method '**Mean of nearby points**'



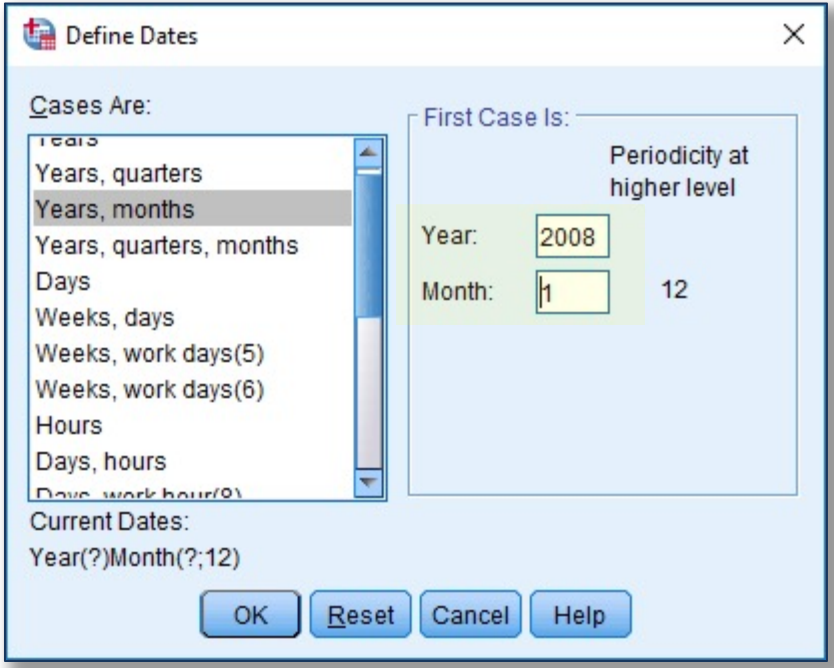




Assigning Periodicity

Assigning Periodicity

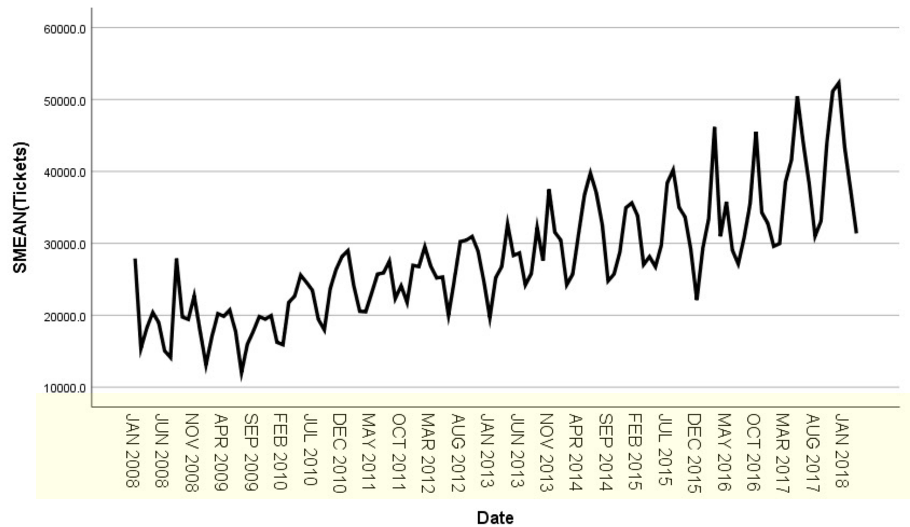
- SPSS Statistics can only see a sequence of numbers
- It doesn't know what separates each row of data
- They could be separated by minutes, weeks or years
- To address this, we explicitly **assign periodicity** to the dataset



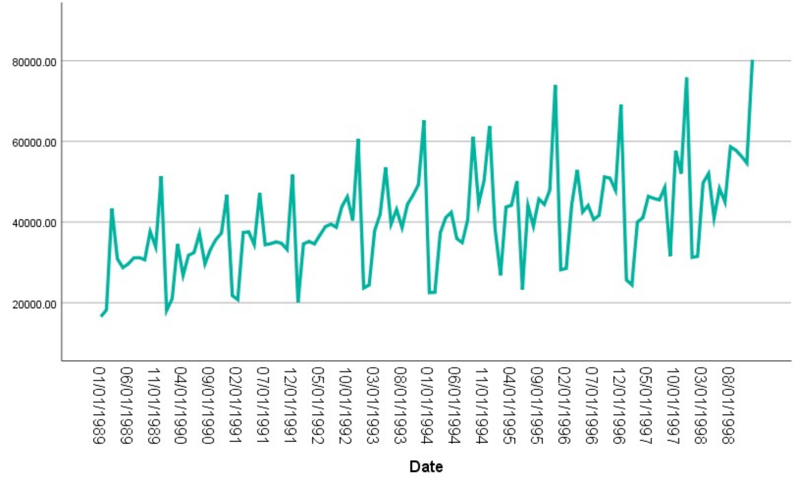
Assigning Periodicity

- SPSS Statistics now creates a series of fields that it can use in time series analysis to identify the periodic (and seasonal) separation between the sequence values

YEAR_	MONTH_	DATE_
2008	1	JAN 2008
2008	2	FEB 2008



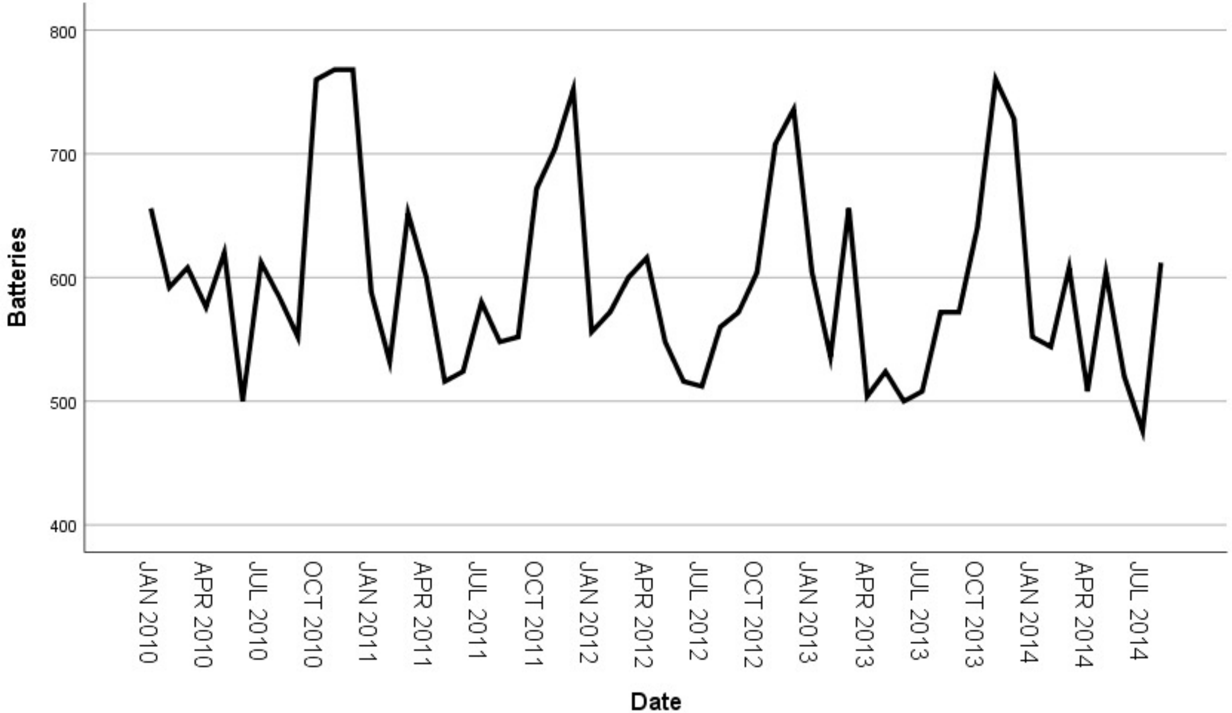
SMARTVISION Europe



Smoothing Series

Smoothing Series

- Creating a sequence plot of the file **Batteries sales.sav**
- Note the periodicity is already defined



Smoothing Series

- We can use SPSS to create a **smoothed** version of this series
- Smoothing is usually done to reveal a clearer picture of the series by simplifying it and removing some of the randomness.
- Smoothing a series is similar to the methods that **pure** Time Series techniques employ to create a model for forecasting
- The are *many different* ways that smoothing techniques can be applied.

Moving Average Smoothing Example

- One of the most common forms of smoothing is using a **moving average**
- The graphic shows how a **moving average** is calculated with a span of 3 cases

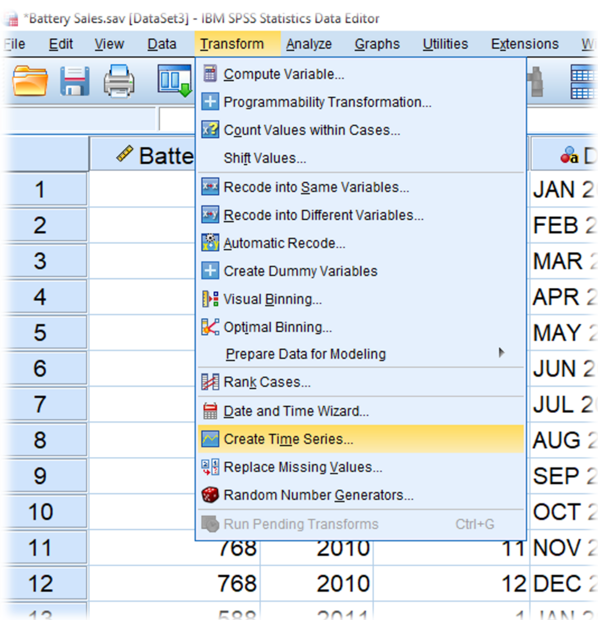
Case_ID	Variable	MA_3
1	1	
2	1	1.3
3	2	2
4	3	3.3
5	5	5.3
6	8	8.7
7	13	14
8	21	

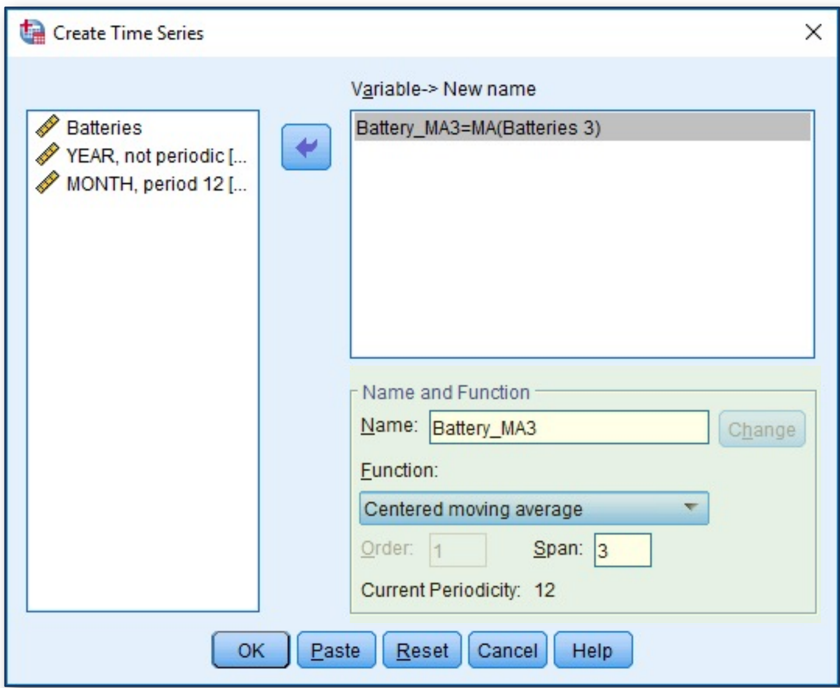
Case_ID	Variable	MA_3
1	1	
2	1	1.3
3	2	2
4	3	3.3
5	5	5.3
6	8	8.7
7	13	14
8	21	

Case_ID	Variable	MA_2
1	1	
2	1	1.3
3	2	2
4	3	3.3
5	5	5.3
6	8	8.7
7	13	14
8	21	

Smoothing Series

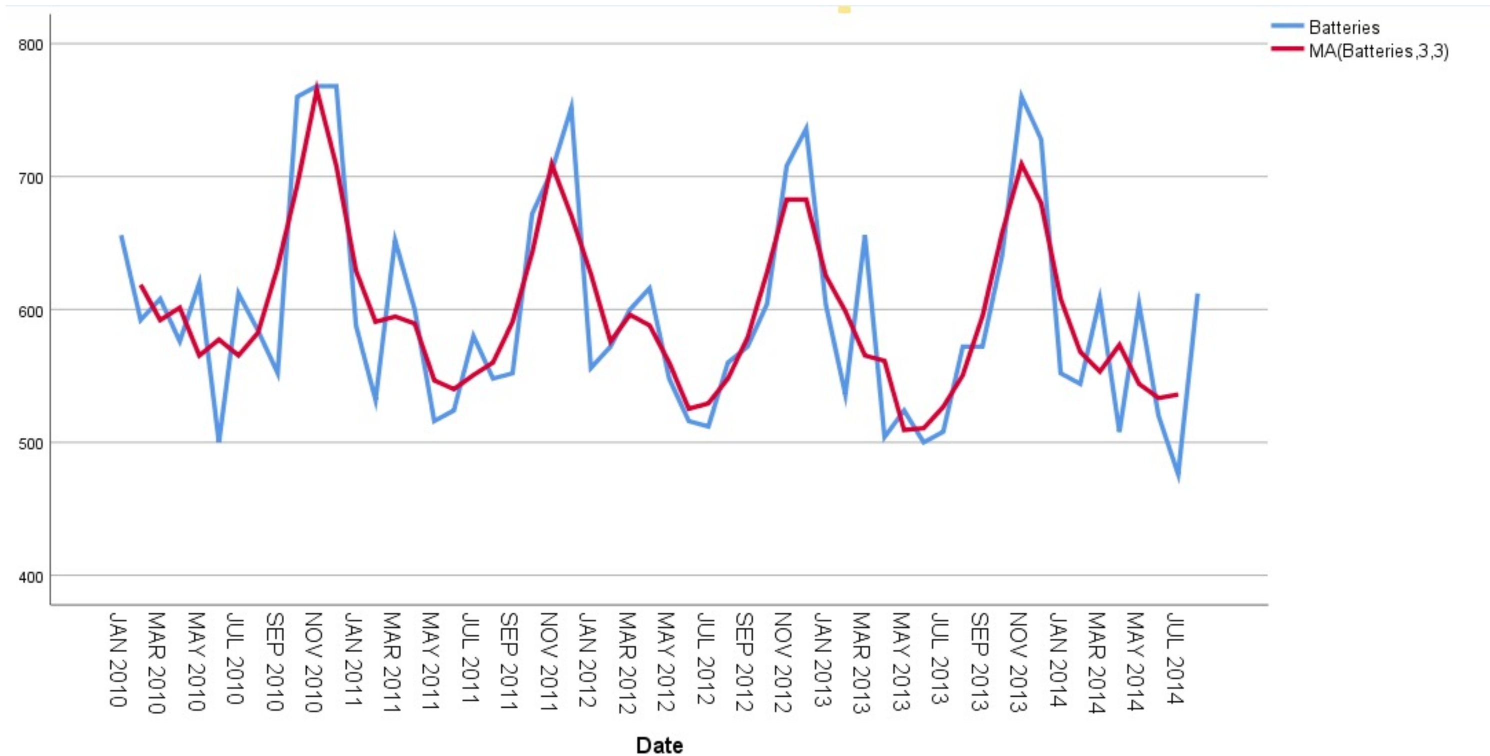
- The **Create Time Series** procedure allows us to create smoothed series



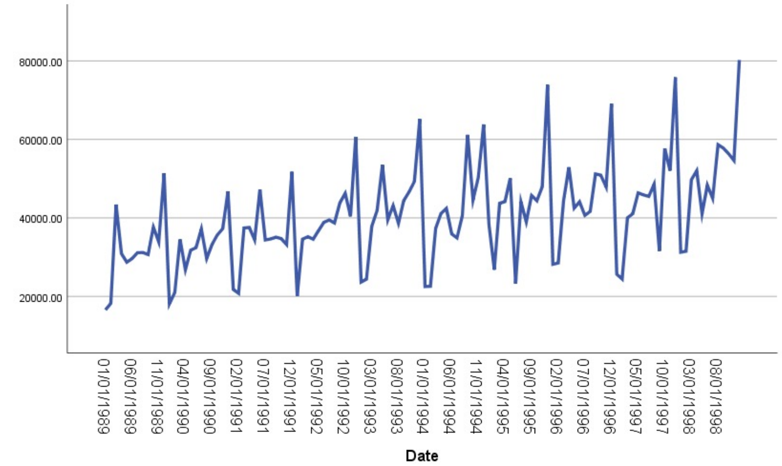


Smoothing Series

- The sequence chart clearly shows that the smoothed series moving average is a less noisy, albeit simplified version of the Batteries sales variable



SMARTVISION Europe



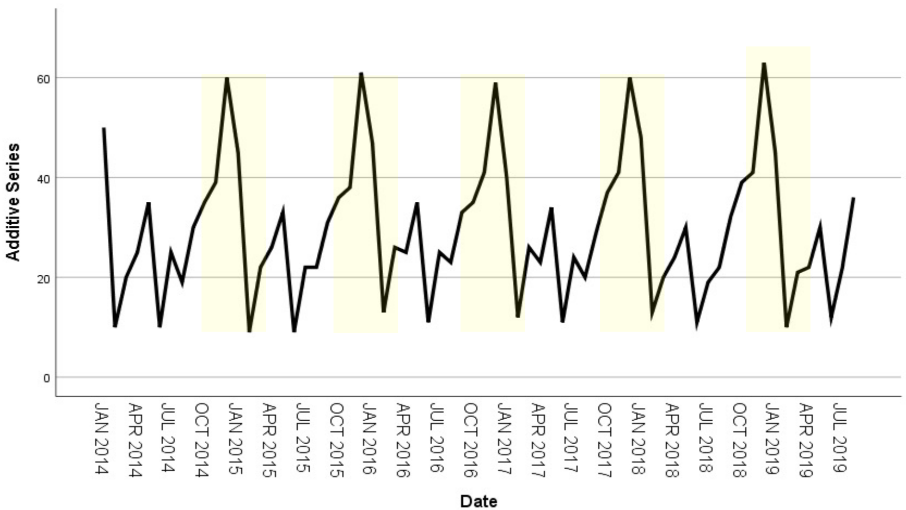
Exponential Smoothing Methods

Exponential Smoothing

- Using the simple moving average smoothing techniques that we saw earlier, each time point in the calculation has equal weight. For example, by using a span of say 5, the cases that are two time points away are treated as equally as those that are one time point away when the moving average is calculated.
- In ***Exponential* smoothing**, values that are *closer* (in time) are *given greater weight* than those that are further away.
- This approach can be employed in forecasting where the values that are *more recent* have a greater influence on estimating the future than those that are less recent.

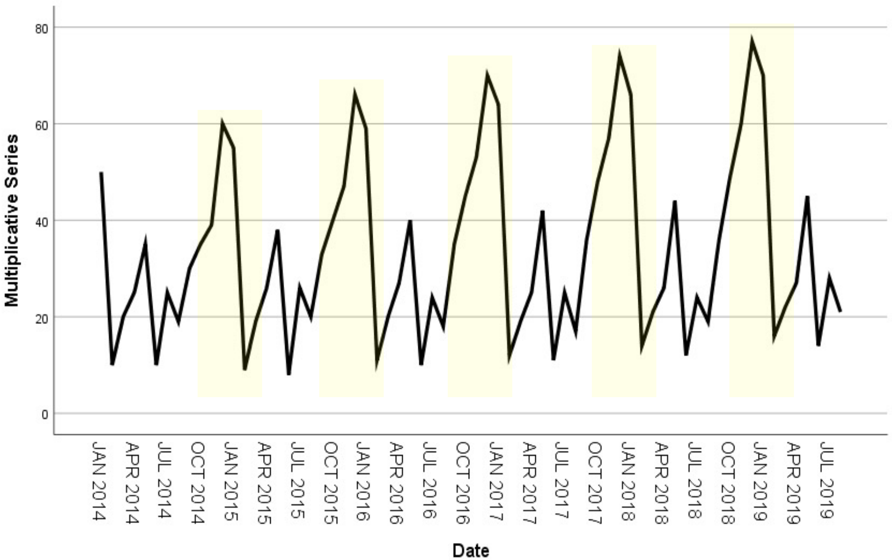
Additive vs Multiplicative Series

Additive Series



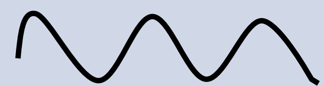
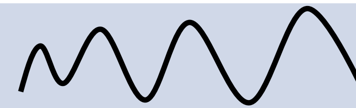
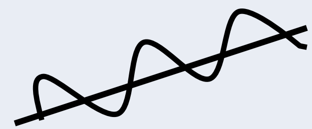
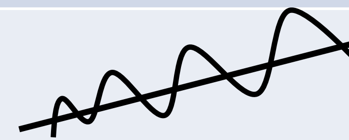
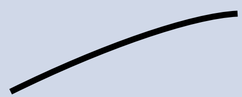
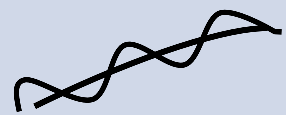
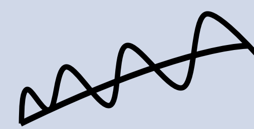
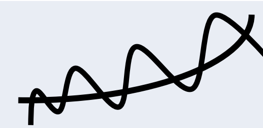
Seasonal Variation Constant

Multiplicative Series



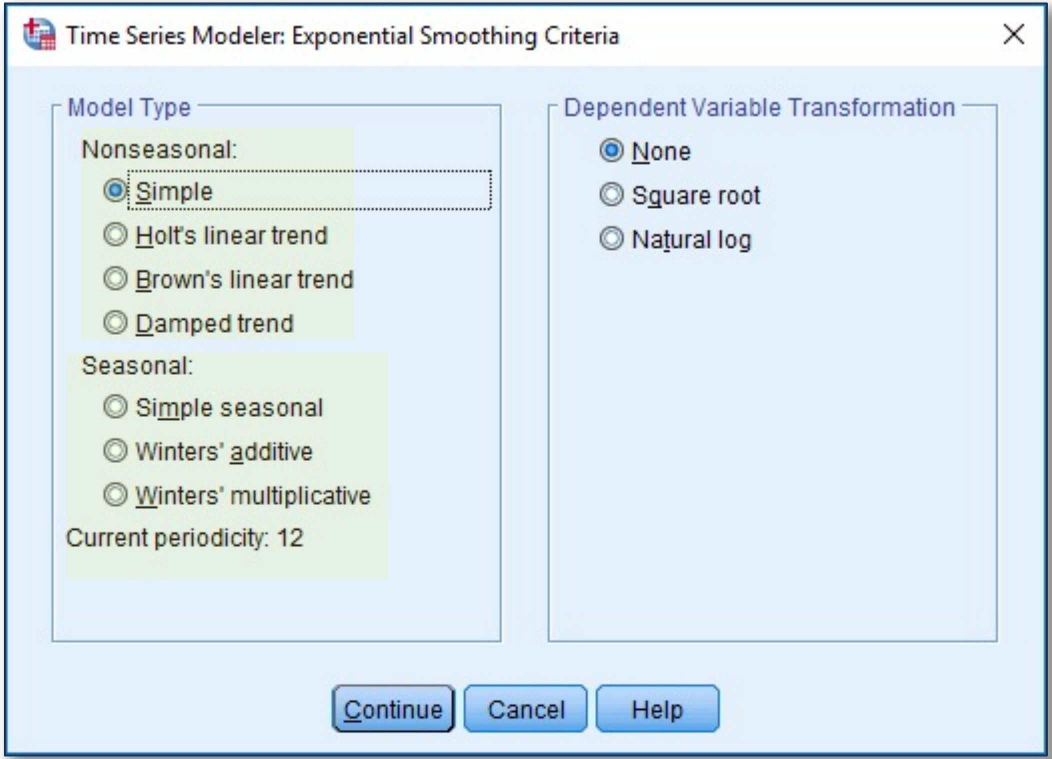
Seasonal Variation Changing

Basic Exponential Smoothing Methods

	Non Seasonal	Additive Seasonal	Multiplicative Seasonal
Constant Level			
Linear Trend			
Damped Trend			
Exponential Trend			

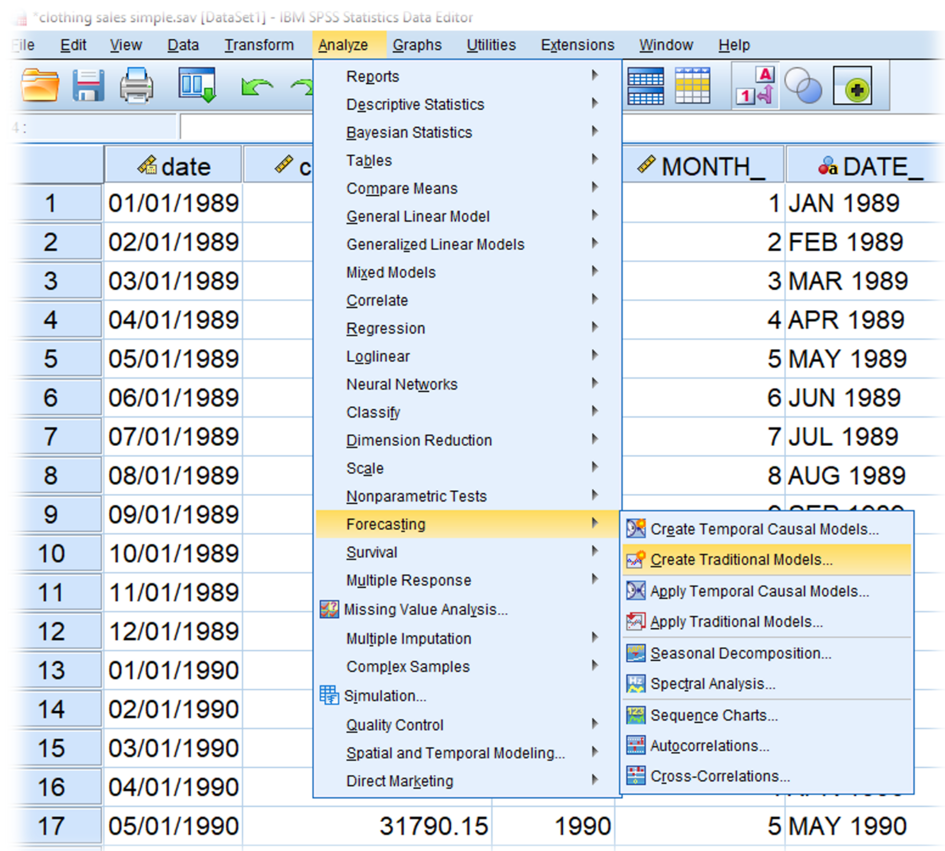
SPSS Statistics Exponential Smoothing Models

- SPSS Statistics has **7** standard Exponential smoothing models
- The standard models are divided into:
  - *4 non-seasonal* methods
  - *3 seasonal* methods
- However using the **Expert Modeler** method means that SPSS *automatically* chooses a model type for the series



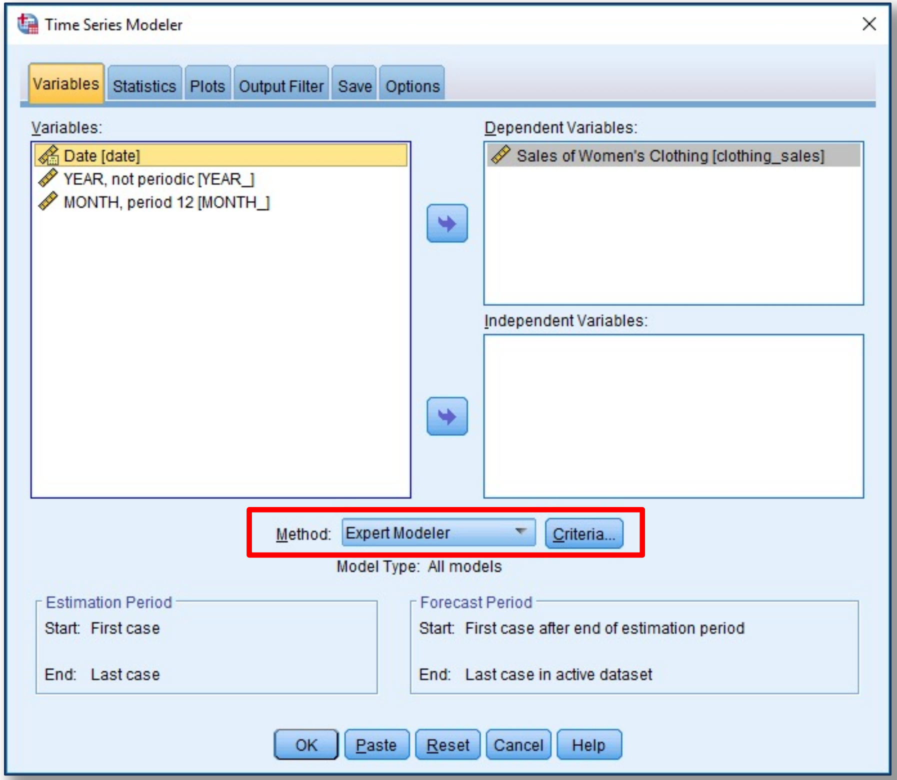
Forecasting with SPSS Expert Modeler

- To create a forecast using **Expert Modeler** from the main menu click:
  - **Analyze**
    - **Forecasting**
      - **Create Traditional Models**



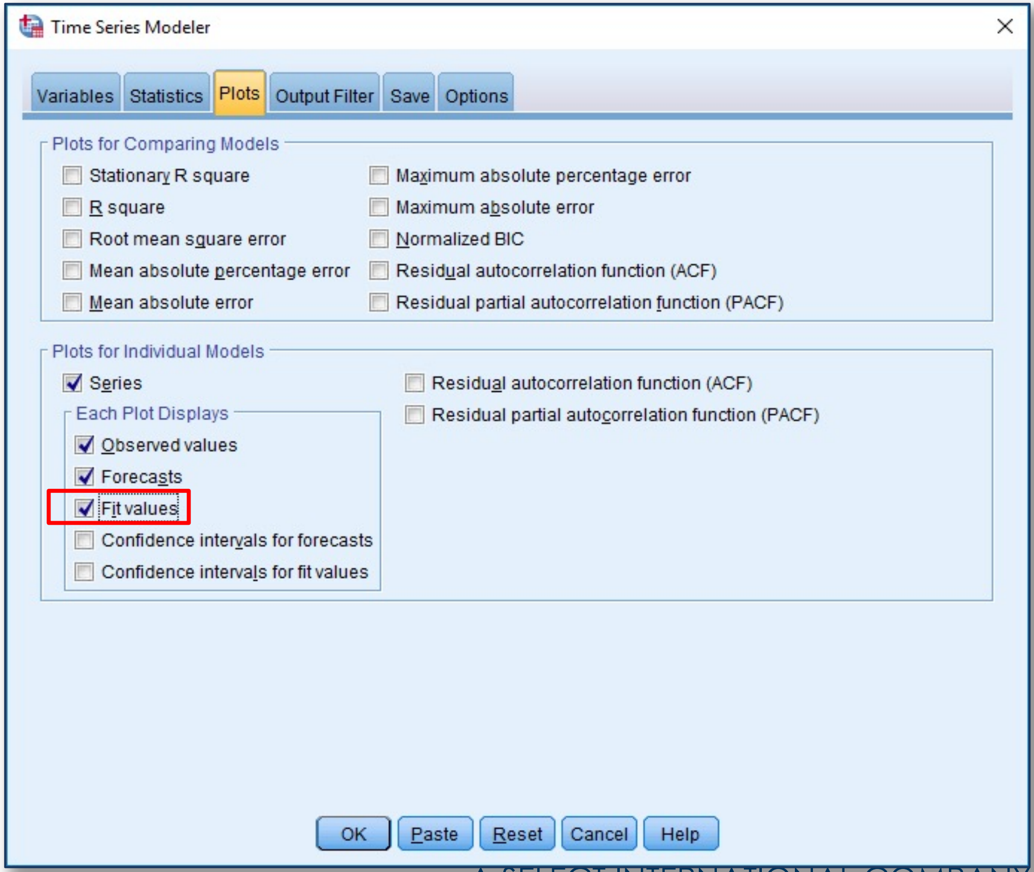
Forecasting with SPSS Expert Modeler

- Specify **Sales of Women's' Clothing** as the dependent variable
- Note that the default **Method** for model fitting is **Expert Modeler**
- Based on a measure of model fit, this method will try a number of model types and select the one with the best fit over the series



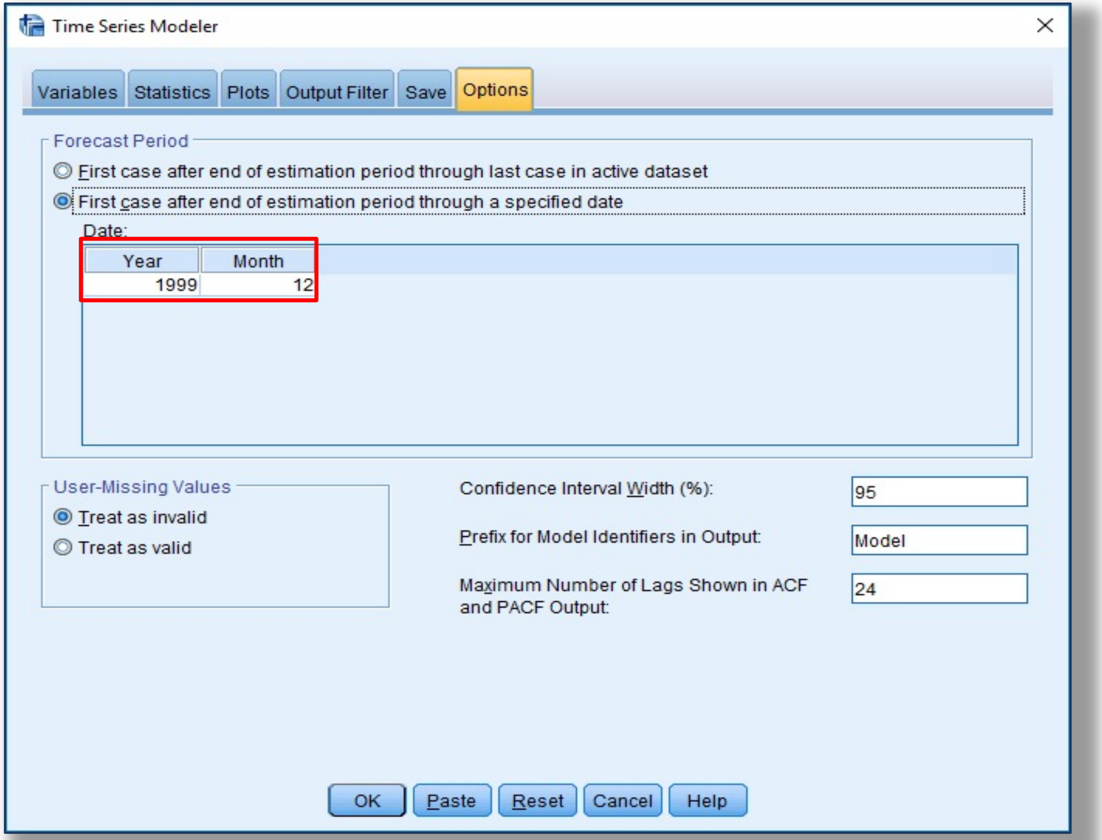
Forecasting with SPSS Expert Modeler

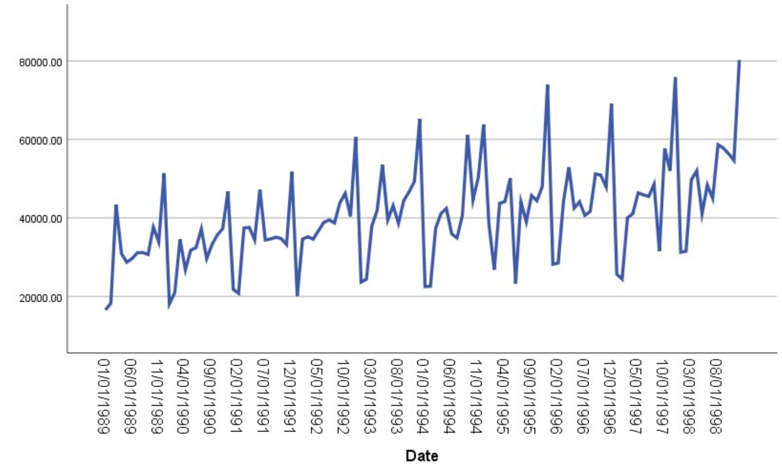
- We can request a plot to enable us to see how the model as been fitted to actual series



Forecasting with SPSS Expert Modeler

- We can specify how far we wish to forecast into the future





Interpreting Output and Model Fit

Forecasting with SPSS Expert Modeler

- Note that expert modeler chooses **Winters' Additive** as the best model type – this model is used with series showing a trend and seasonal effects

Model Description

			Model Type
Model ID	Sales of Women's Clothing	Model_1	Winters' Additive

Model Summary

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile 5	10	25	50	75	90	95
Stationary R-squared	.735	.	.735	.735	.735	.735	.735	.735	.735	.735	.735
R-squared	.815	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.815
RMSE	5296.069	.	5296.069	5296.069	5296.069	5296.069	5296.069	5296.069	5296.069	5296.069	5296.069
MAPE	9.412	.	9.412	9.412	9.412	9.412	9.412	9.412	9.412	9.412	9.412
MaxAPE	83.036	.	83.036	83.036	83.036	[illegible]	[illegible]	83.036	83.036	83.036	83.036
MAE	3624.699	.	3624.699	3624.699	3624.699	3624.699	3624.699	3624.699	3624.699	3624.699	3624.699
MaxAE	19302.778	.	19302.778	19302.778	19302.778	19302.778	19302.778	19302.778	19302.778	19302.778	19302.778
Normalized BIC	17.269	.	17.269	17.269	17.269	17.269	17.269	17.269	17.269	17.269	17.269

We can ignore these statistical values as they are used to compare the performance of *multiple models* so aren't relevant here

Model Fit Statistics

- **Stationary R square –** A *stationary* model is effectively one with the trend removed so that the values have the same variance and mean over time*. Larger values indicate better fit.*

- **Root mean square error (RSME) -** Squared errors are based on the square of the differences between *the fitted values and the observed values*. It's similar to a standard deviation value. *Smaller values indicate better fit.*

- **Mean absolute percentage error (MAPE) –** The average error values in percentage terms. Values such as 0.15 equate to an average of 15% error. *Smaller values indicate better fit.*
- **Mean absolute error -** Mean of the absolute values of the forecast errors. MAE is in the same units as the dependent series. MAE is appropriate when the cost of the forecast errors is proportional to the absolute size of the forecast error. *Smaller values indicate better fit.*

- **Maximum absolute percentage error -** The largest forecast error, expressed as a percentage. This measure gives a worst case scenario indication of model performance. It works best if there are no extremes to the data. *Smaller values indicate better fit.*
- **Maximum absolute error -**The largest forecast error. Expressed in the same units as the dependent series. This measure gives a worst case scenario indication of model performance. *Smaller values indicate better fit.*

- **Normalized BIC -** The Normalized *Bayesian Information Criterion* (BIC) fit measure enables you to compare different models for the same series. Normalized BIC "rewards" simpler models that fit better, while it "penalizes" models that use more parameters. It is based on a mean squared **This is the fit measure that Expert Modeler uses when comparing candidate models.** *Smaller values indicate better fit.*

Forecasting with SPSS Expert Modeler

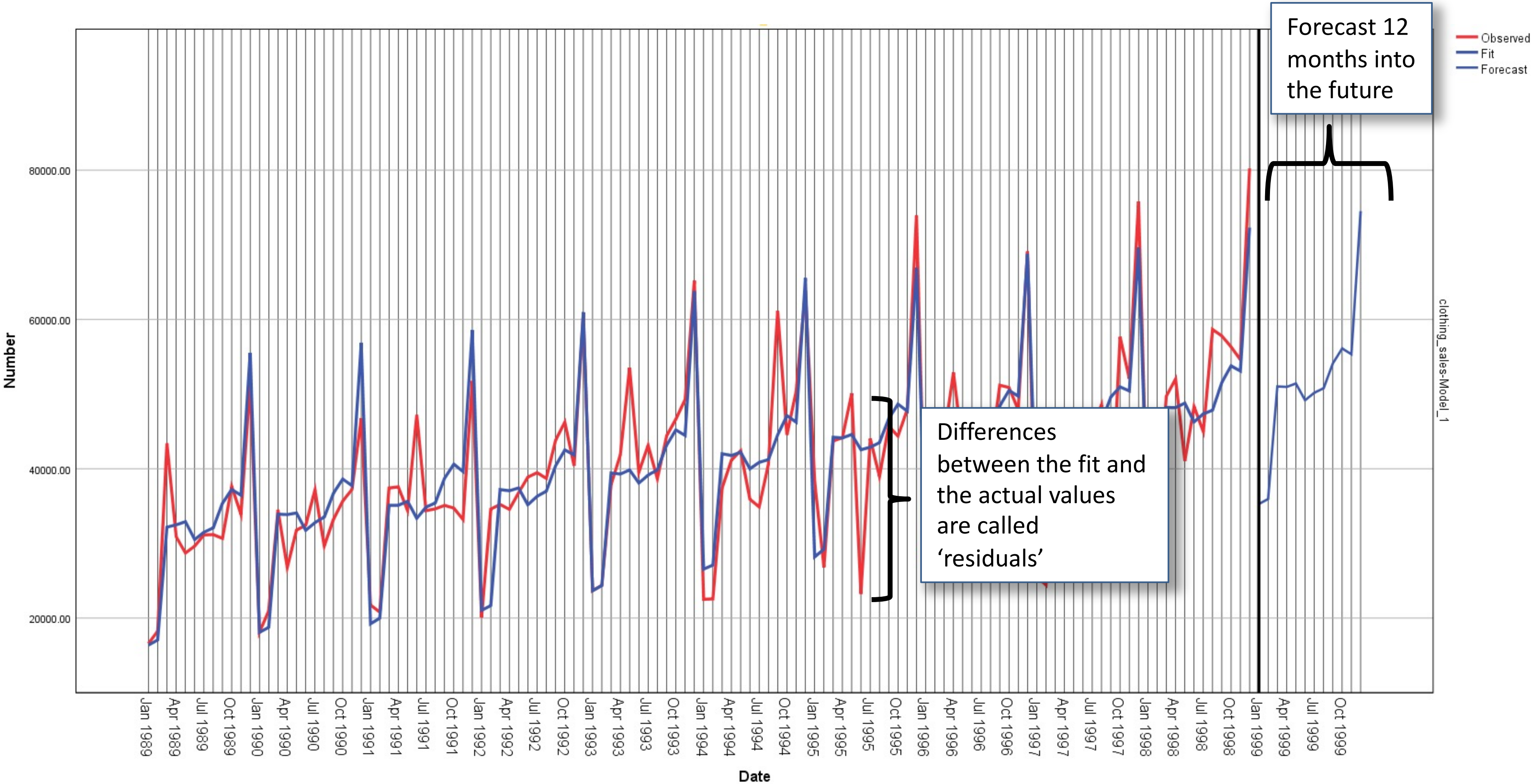
- **Ljung-Box Q: a lack of fit test to check that the model is correctly specified**

Model Statistics

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	DF	Sig.	
Sales of Women's Clothing-Model_1	0	.735	23.635	15	.072	0

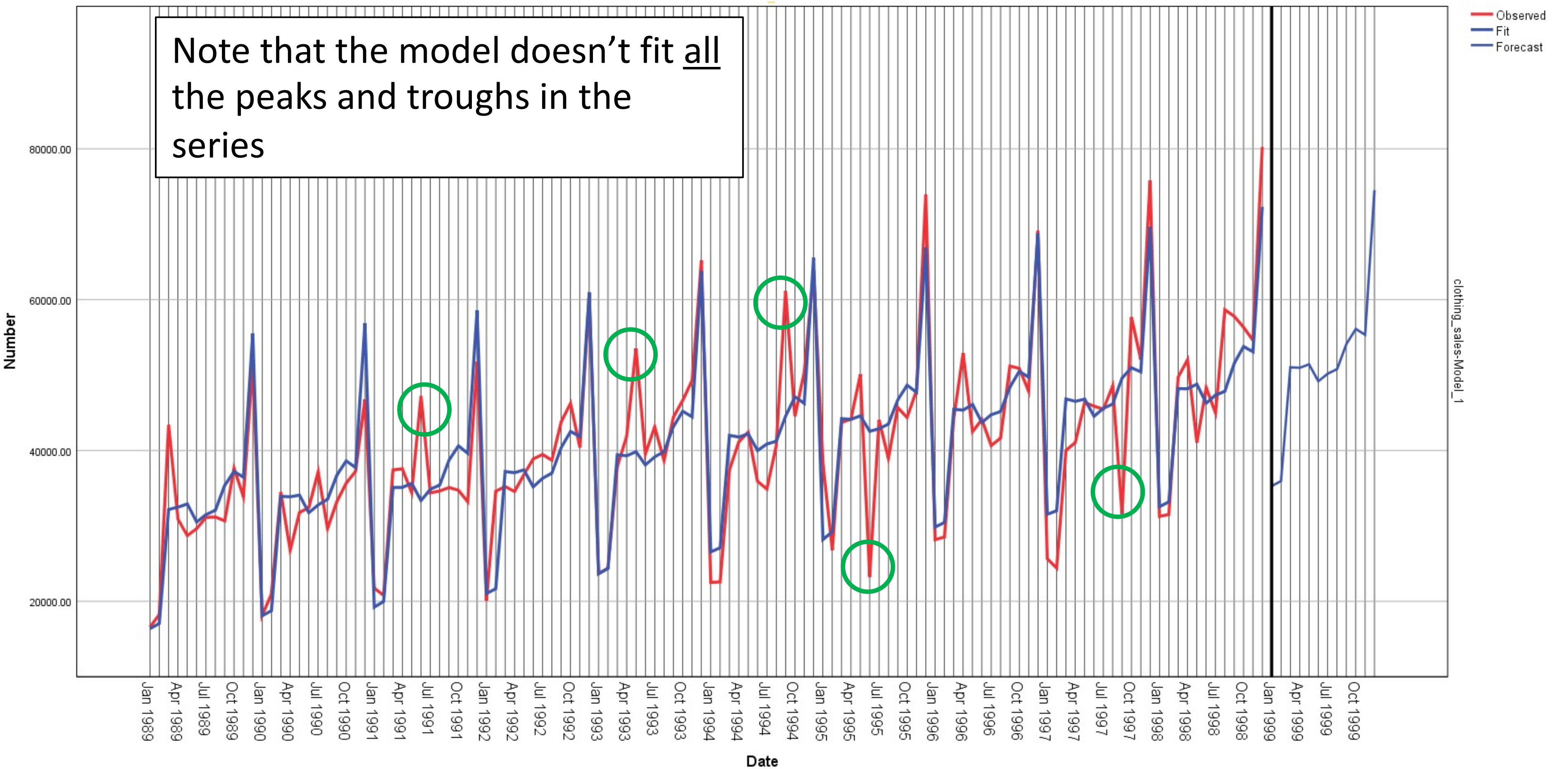
- Time series models are often evaluated by focussing on the *errors* in the model.
- These are also known as residuals and they represent the difference between the fit values and the actual values.
- Ideally when the model has been fitted, the correlation between the residuals in the sequence should be random (white noise).
- Values in the Sig. column above 0.05 indicate that the model doesn't leave any significant correlations after the model has been specified. So we *might* assume that it's doing a reasonable job (overall) of fitting the series.

Forecasting with SPSS Expert Modeler



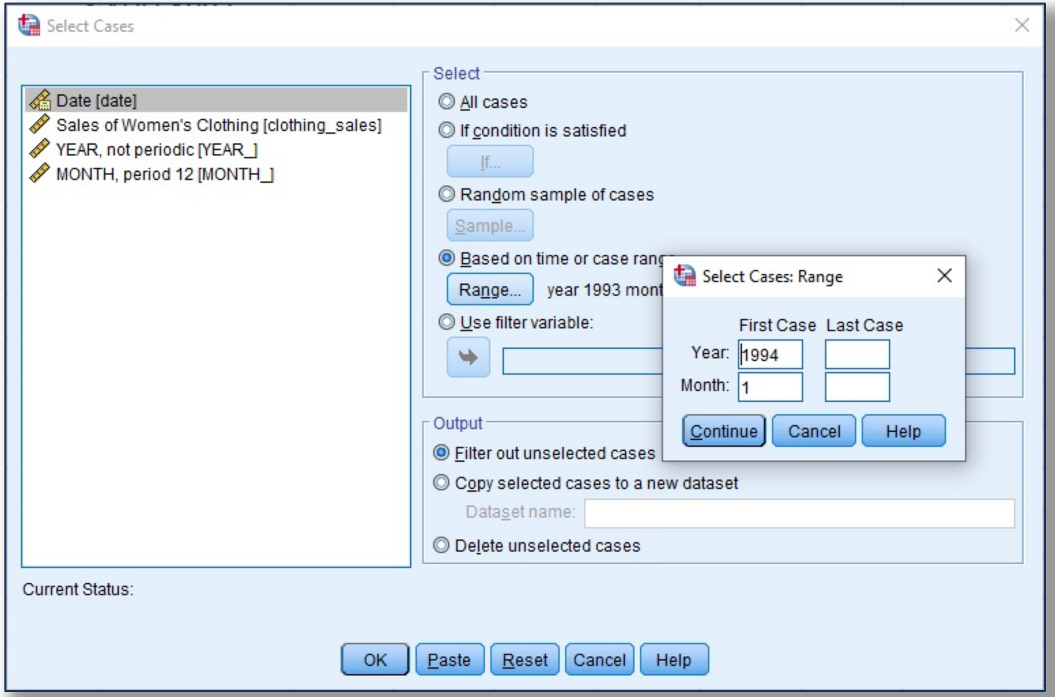
Forecasting with SPSS Expert Modeler

Note that the model doesn't fit <u>all</u> the peaks and troughs in the series



Forecasting with SPSS Expert Modeler

- Lets look at the effect of changing the *range of time* that we submit to the Expert Modeler algorithm.



Forecasting with SPSS Expert Modeler

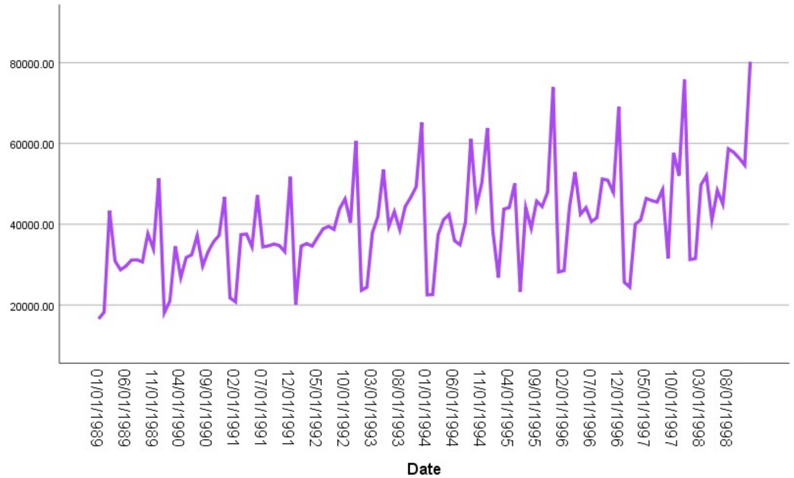
- Although the chosen Model Type is still Winter's Additive. The fit statistics are slightly different when using only the last 5 years of data

Fit Statistic	Mean
Stationary R-squared	.735
R-squared	.815
RMSE	5296.069
MAPE	9.412
MaxAPE	83.036
MAE	3624.699
MaxAE	19302.778
Normalized BIC	17.269

Model 1: 1989 - 1998

Fit Statistic	Mean
Stationary R-squared	.822
R-squared	.815
RMSE	5552.480
MAPE	9.634
MaxAPE	68.892
MAE	3897.216
MaxAE	19307.360
Normalized BIC	17.449

Model 2: 1994 - 1998



Using Predictor Fields with ARIMA

Using predictor fields with ARIMA

- As we have seen, the Expert Modeler in SPSS Forecasting attempts to automatically find the best-fitting model for each dependent series. So far we have only been working with Exponential Smoothing models However, by default the Expert Modeler considers two types of Time Series model: Exponential Smoothing and **ARIMA**.
- If we choose to work with any **independent** (predictor) variables, then the Expert Modeler will select an ARIMA model if any of these independent variables have a statistically significant relationship with the dependent series.
- The acronym ARIMA refers to the **three components** of this modelling approach.
- It stands for Autoregressive (**AR**) Integrated (**I**) Moving Average (**MA**)
- This is the **structure** of an ARIMA model. But not all ARIMA models use **all** these elements. In fact, most don't.

Using predictor fields with ARIMA

- The three components within the ARIMA model are usually shortened to ***p,d,q***
  - **p** refers to the **autoregression** component
  - **d** refers to the integration or ***differencing***
  - **q** refers to the **moving average** component

Using predictor fields with ARIMA

- **Autoregression (p)** refers to the correlation between a value in a series and the previous value(s). If I want to know the temperature today, is it useful to know the temperature yesterday? If so, then the autoregression ***p*** component is equal to **1** as this represents a lag of 1 day (a *first order* autoregression). If however the *day before yesterday* is a better component, then a value of **2** should be used to signify this (a *second order* autoregression).

Using predictor fields with ARIMA

- Integration or ***differencing (d)***. Many analysts prefer the term 'differencing' rather than 'integration' for the **d** component of ARIMA***.***
- Differencing is a technique to used to make a time series **stationary**. A stationary series is one where the mean, the variance and autocorrelation values are constant over time.
- You may recall that Time Series analysis attempts to break a series up into different components by isolating elements like the trend component. ARIMA models require a series to be stationary in order to estimate correctly and the differencing component deals with this.
- *Differencing* ***(d)*** *simply refers to the difference between one value in a sequence and another.* If we were looking at daily temperatures, A first order differencing would result in a sequence of values showing the difference between one day and the previous day. These values could of course be positive or negative.

Using predictor fields with ARIMA

- **Moving Average (q)**. We shouldn't confuse the term moving average here with the kind of moving average *smoothing* exercise that we undertook earlier. In fact the **q** component is another kind of autoregression. Except that this time it focusses on the *errors (*sometimes called 'shocks'*)* in the forecasting model.
- A first order **q** component (denoted as '1') in an ARIMA model indicates that the model's error in the immediate previous period is related to what the dependent variable will be now. For example, knowing that the model overestimated or underestimated by 20% in the previous period helps to predict the current value.

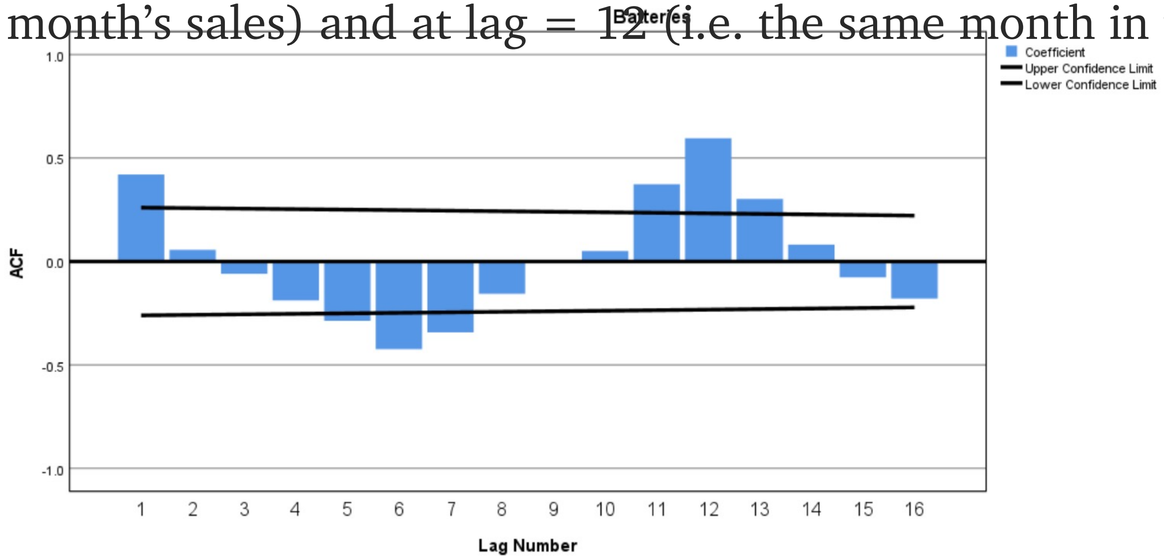
Using predictor fields with ARIMA

- Normally analysts thoroughly explore their time series data before attempting to fit a model
- In doing so they tend to pay a lot of attention to special charts known as correlograms
- These important charts have two forms:
  1. **ACF – Autocorrelation Function**
  2. **PACF – Partial Autocorrelation Function**

- These charts are especially important if one needs to *manually* specify an ARIMA model as they can be used to give a clues to the what the **p**, **d**, **q** values should be.

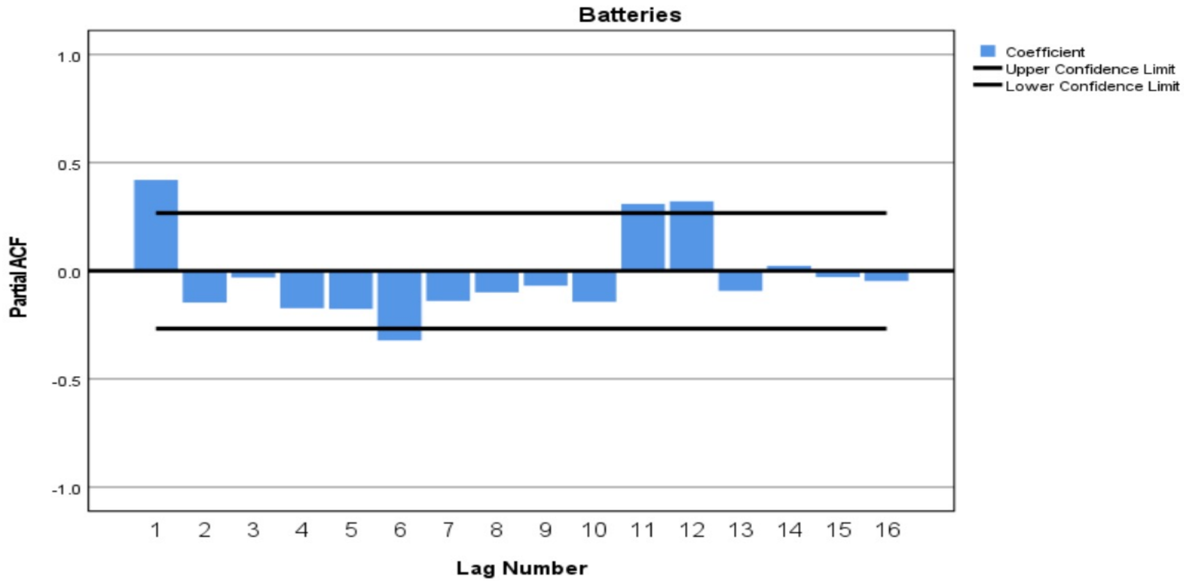
Using predictor fields with ARIMA

- **ACF – Autocorrelation Function**
- This displays how well the present value of the series is related with its past values
- Each bar shows the correlation with the present value at a given lag number
- In the batteries data we can see strong correlations at lag = 1 (i.e. the previous month's sales) and at lag = 12 (i.e. the same month in the previous year)



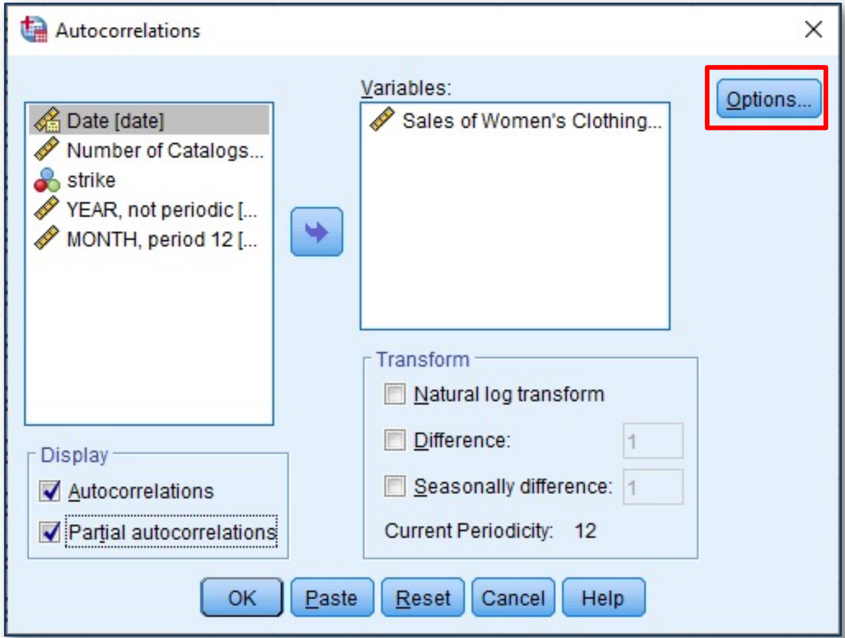
Using predictor fields with ARIMA

- **PACF – Partial Autocorrelation Function**
- Unlike the ACF it looks at the correlation of the residuals (the error remaining after removing the variation already explained by the earlier lags) with the current time point. It is a pure measure of correlation as it controls for all correlation values up to that time point.



Using predictor fields with ARIMA

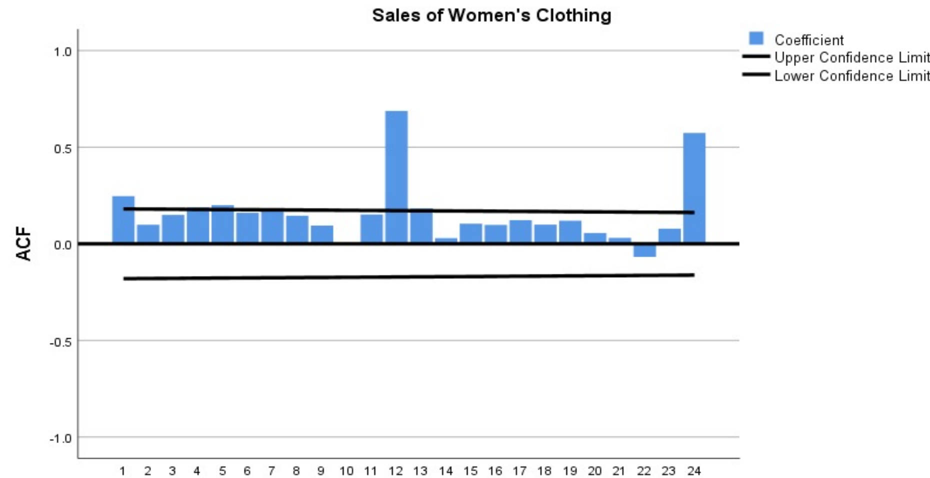
- Before building an ARIMA model, we can investigate the autocorrelation in the series.

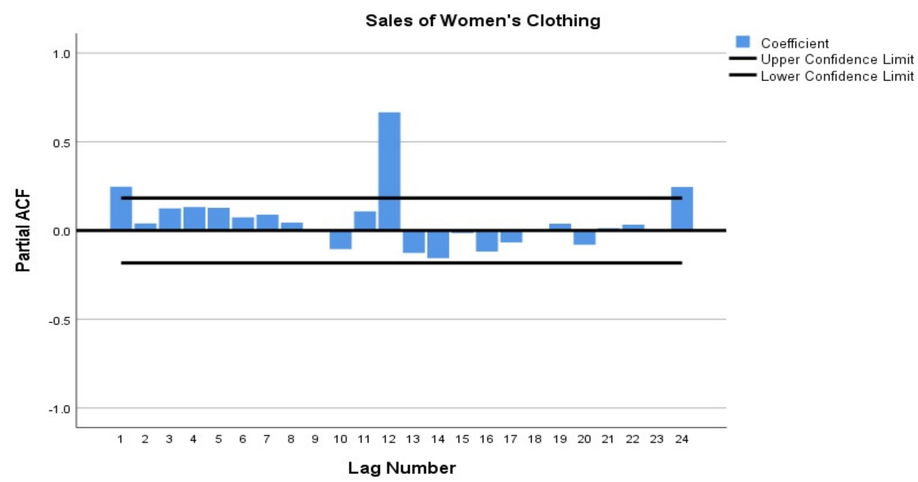


Requesting 24 lag maximum

Demo 4: Using predictor fields with ARIMA

- Both the ACF and PACF plots show strong seasonal correlations at the 12 and 24 month lag points.
- A weaker correlation is also shown at the non-seasonal 1st order (i.e. lag 1)
- Based on this, the ARIMA model should probably have at least this structure: (0,0,0) (1,0,0)

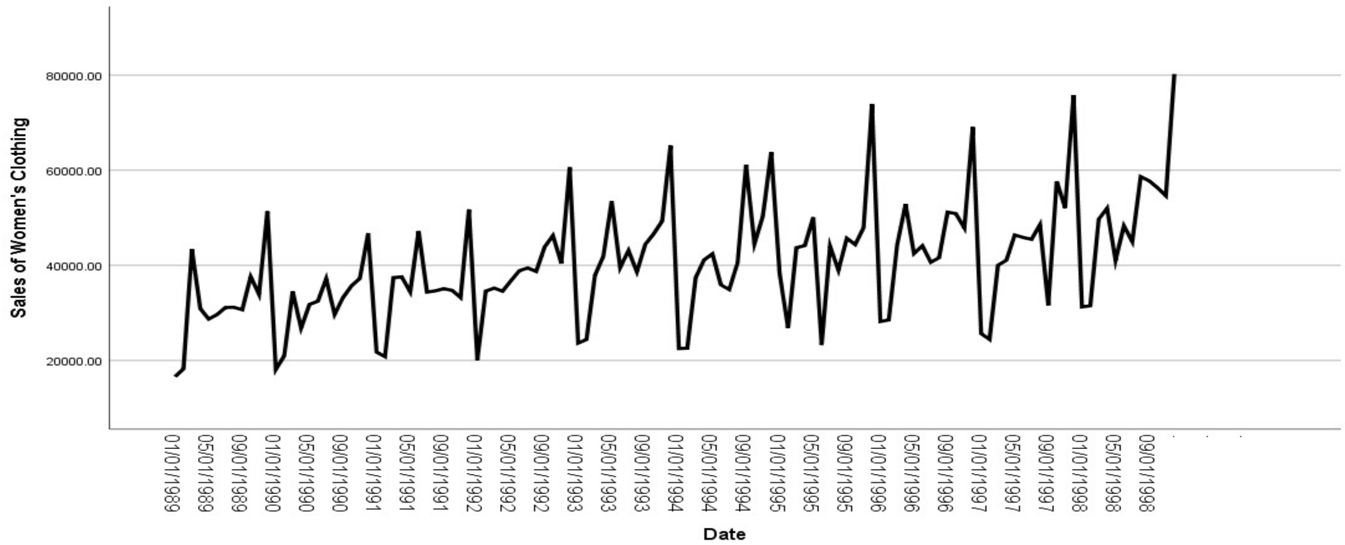




SMARTVISION Europe

Using predictor fields with ARIMA

- In this section we can use the clothing sales file but with two additional independent variables: **mail** and **strike**
- The variable **strike** has only two values: 1 and 0. This is called an **event** variable.
- Remember this is what the series looks like:

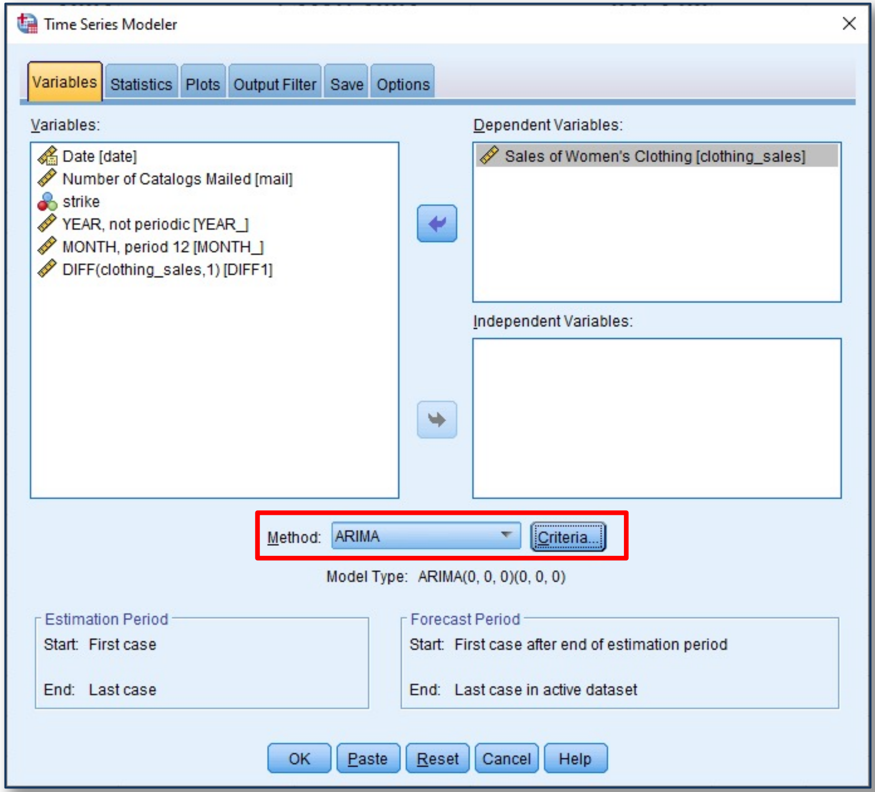


Demo 4: Using predictor fields with ARIMA

- As an experiment, let's return to the Time Series Modeler dialog an **manually** specify and ARIMA model.

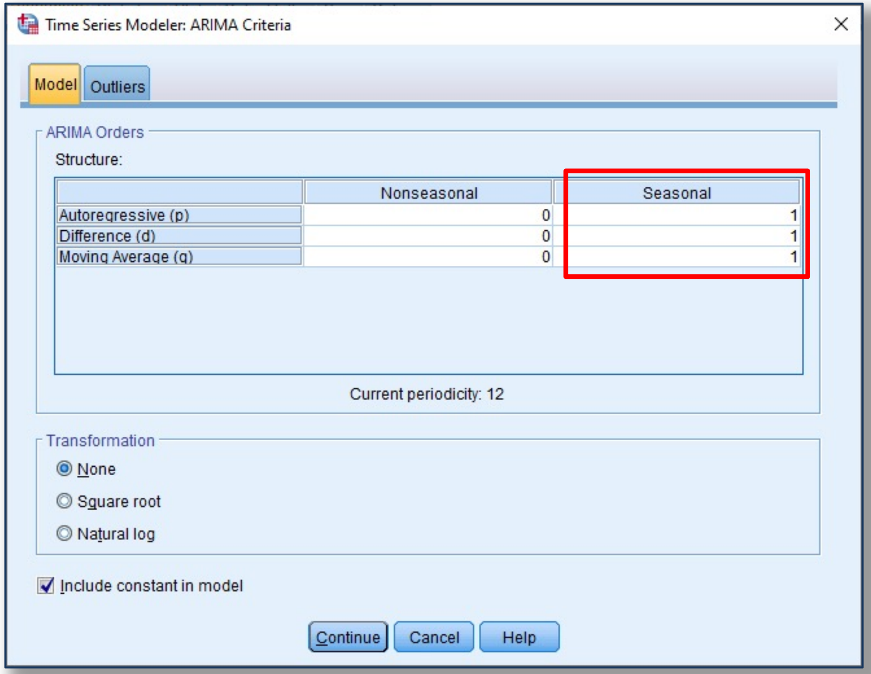
- To override **Expert Modeler** and specify our own **ARIMA** model, click the drop-down button marked **Method** and change it to **ARIMA**

- To specify the **p,d,q** values we click:
  - **Criteria**



Demo 4: Using predictor fields with ARIMA

- Further exploratory analyses indicate that not only is there strong seasonal correlation as shown in the ACF and PACF charts but that the data is not stationary so we should add a differencing term
- With this in mind let's specify the P,D,Q values in the seasonal column as (1,1,1).
- This means our ARIMA model is (0,0,0), (1,1,1)
- To see how the model performs click:
    - **Continue**
        - **OK**



Using predictor fields with ARIMA

Winter's Additive Exponential Smoothing

Fit Statistic	Mean
Stationary R-squared	.735
R-squared	.815
RMSE	5296.069
MAPE	9.412
MaxAPE	83.036
MAE	3624.699
MaxAE	19302.778
Normalized BIC	17.269

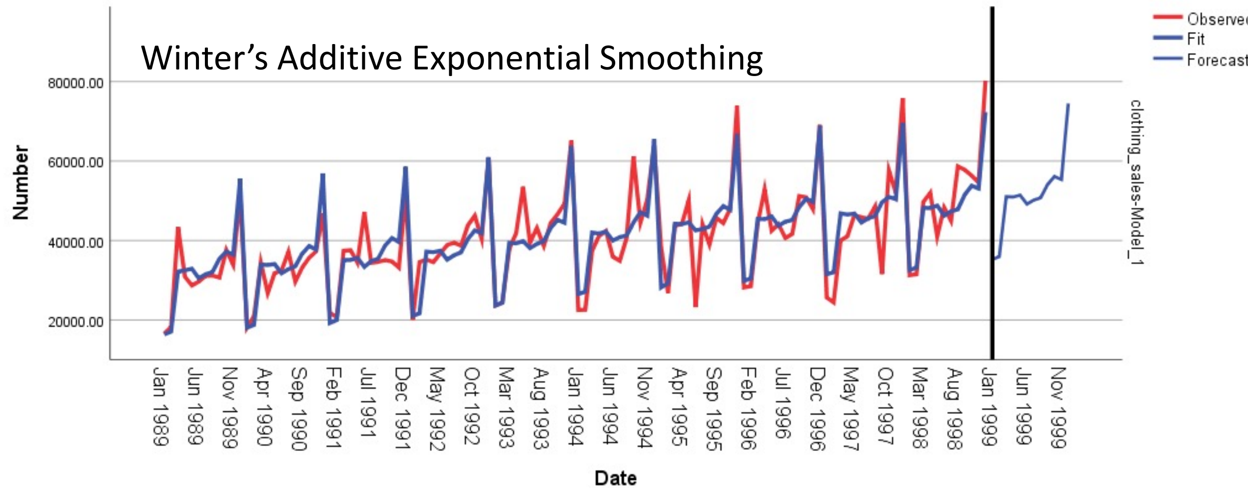
ARIMA (0,0,0) (1,1,1)

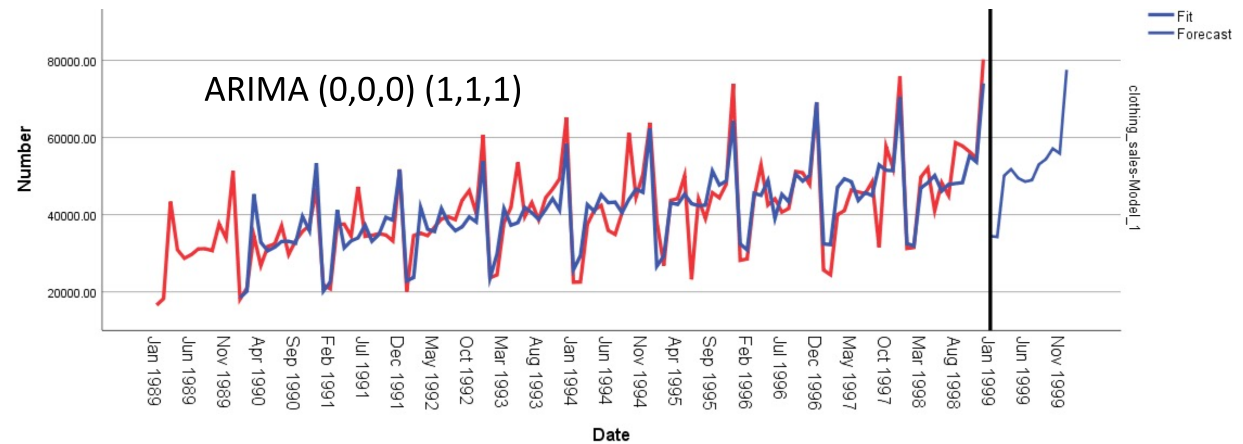
Fit Statistic	Mean
Stationary R-squared	.306
R-squared	.753
RMSE	6084.494
MAPE	11.374
MaxAPE	84.309
MAE	4409.612
MaxAE	21318.207
Normalized BIC	17.557

- The first thing we notice is that the Stationary R-Squared value is much smaller. This is the effect of adding the **D** (differencing component to the model). But apart from that there's not a **huge** difference between the fit values.

Using predictor fields with ARIMA

- The sequence charts for the Exponential Smoothing model vs our manually specified ARIMA model show that they look very similar.
- Note that the ARIMA model doesn't include a model fit line for the first 12 months of the data (as it would need the previous 12 months to do so)

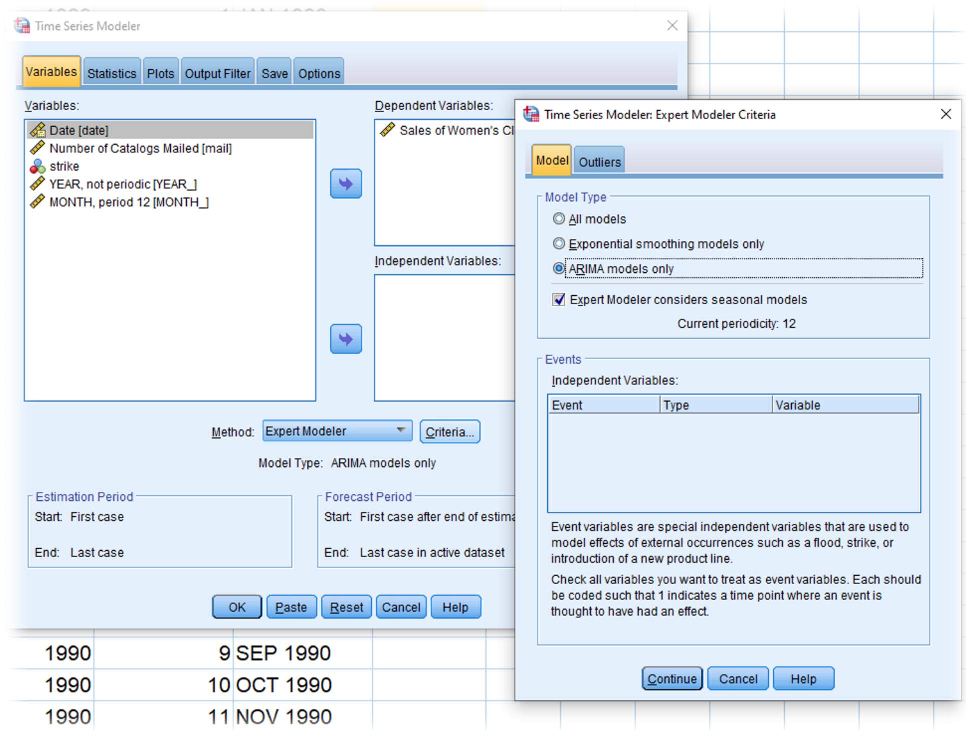




SMARTVISION Europe

Using predictor fields with ARIMA

- Returning to Time Series Modeler dialog we can use the **Expert Modeler** function to automatically specify a model for us
- Obviously if we just ran this procedure as we did at the start, it will select a **Winters Additive Exponential Smoothing** model again
- So we will force it to *only consider ARIMA models*



Using predictor fields with ARIMA

Model Description

			Model Type
Model ID	Sales of Women's Clothing	Model_1	ARIMA(0,0,0)(0,1,1)

- The results show that Expert Modeler has *dropped* the seasonal autoregression term (P). Remember, this method looked at the previous relevant time point and uses it to estimate the next timepoint based on a simple regression formula
- It has however kept the differencing function (D). This method is used to make the series stationary so the variance doesn’t change over time.
- It has also kept the moving average term (Q). This method fits a model by starting from an average value and constantly takes into consideration how wrong it was in the previous time point every time it makes a forecast one time point into the future.

Using predictor fields with ARIMA

Winter's Additive Exponential Smoothing

Fit Statistic	Mean
Stationary R-squared	.735
R-squared	.815
RMSE	5296.069
MAPE	9.412
MaxAPE	83.036
MAE	3624.699
MaxAE	19302.778
Normalized BIC	17.269

ARIMA (0,0,0) (1,1,1)

Fit Statistic	Mean
Stationary R-squared	.306
R-squared	.753
RMSE	6084.494
MAPE	11.374
MaxAPE	84.309
MAE	4409.612
MaxAE	21318.207
Normalized BIC	17.557

ARIMA (0,0,0) (0,1,1)

Fit Statistic	Mean
Stationary R-squared	.304
R-squared	.752
RMSE	6062.496
MAPE	11.364
MaxAPE	85.067
MAE	4400.095
MaxAE	21802.869
Normalized BIC	17.506

- Remember that the Expert Modeler attempts to find models with the smallest Normalised BIC value. That's why it chose Winter's Additive in our first example and the ARIMA (0,0,0) (0,1,1) when we forced it to only consider ARIMA models. Nevertheless the improvement in fit over our manually specified model is very small.

Using predictor fields with ARIMA

- The dataset also contains two predictor fields which can be specified as independent variables. In this context Exponential Smoothing algorithms don’t make use of predictor fields, *but ARIMA can*.

- The variable **mail** refers to the number of catalogues that were posted to customers. Note that if we wish to use this to forecast into the future we would need to specify the future values for each monthly mailing volume. You can see these anticipated future values at the bottom of the data file.

- The variable **strike** indicates whether or not a postal strike *event* had occurred that month.

mail	strike	YEA
11288	0	1
11096	0	1
11224	0	1
11483	0	1
11643	0	1
10893	0	1
11147	1	1
12260	0	1
11168	0	1
14370	0	1
11890	0	1
11722	0	1
11589	0	1
11633	0	1
11951	0	1
11706	0	1

Using predictor fields with ARIMA

Winter's Additive Exponential Smoothing

Fit Statistic	Mean
Stationary R-squared	.735
R-squared	.815
RMSE	5296.069
MAPE	9.412
MaxAPE	83.036
MAE	3624.699
MaxAE	19302.778
Normalized BIC	17.269

ARIMA (0,0,0) (0,1,1) without independent variables

Fit Statistic	Mean
Stationary R-squared	.304
R-squared	.752
RMSE	6062.496
MAPE	11.364
MaxAPE	85.067
MAE	4400.095
MaxAE	21802.869
Normalized BIC	17.506

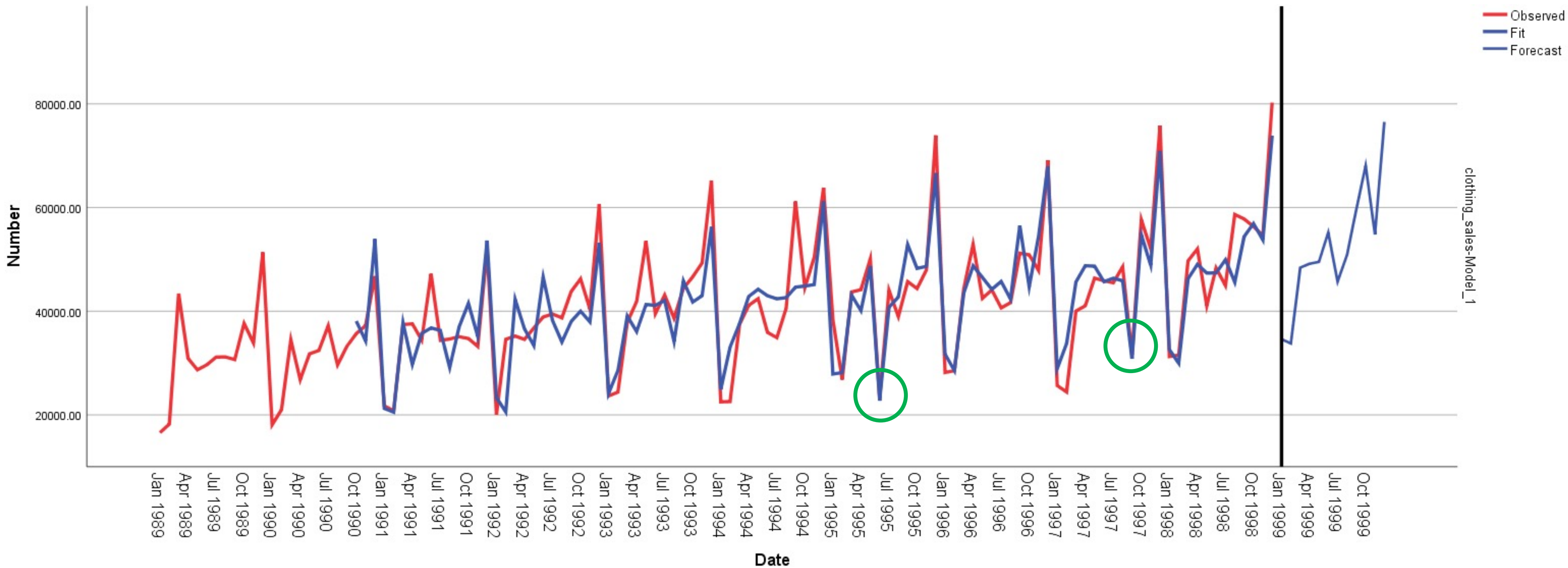
ARIMA (0,0,0) (0,1,1) with independent variables

Fit Statistic	Mean
Stationary R-squared	.492
R-squared	.803
RMSE	5349.490
MAPE	9.810
MaxAPE	46.514
MAE	4045.744
MaxAE	16525.282
Normalized BIC	17.309

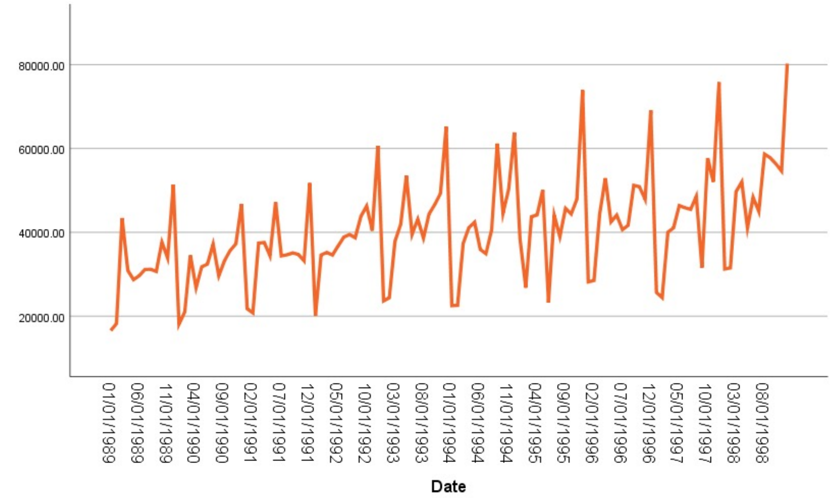
- The model fit seems to have improved slightly on the previous ARIMA model that did not include independent variables. Many of the statistics in the Exponential Smoothing model still seem to indicate a slightly better overall fit. But the biggest change is in the MaxAPE and MaxAE where the maximum absolute percentage error and maximum absolute error values are both quite smaller. This is almost certainly due to the inclusion of the **strike** variable.

Using predictor fields with ARIMA

- The immediate effect of adding an event variable can be seen in the sequence chart for the model. Postal strikes occurred in June 1996 and September 1997. As such, the event variable contains the value 1 at each of these time points. The model immediately picks this up so the fit line correctly reflects the downturn in revenue on these two occasions. Thus an event variable can be used to help explain anomalous values caused by external effects with the aim of providing more accurate estimates.



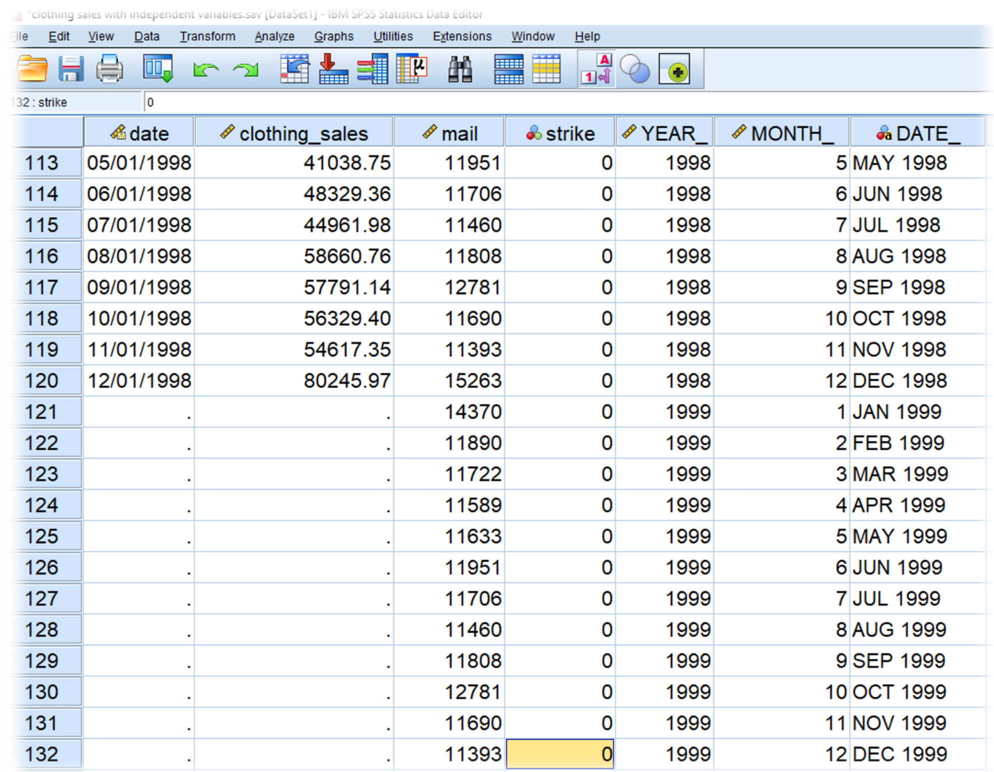
SMARTVISION Europe



Generating Forecasts

Generating Forecasts

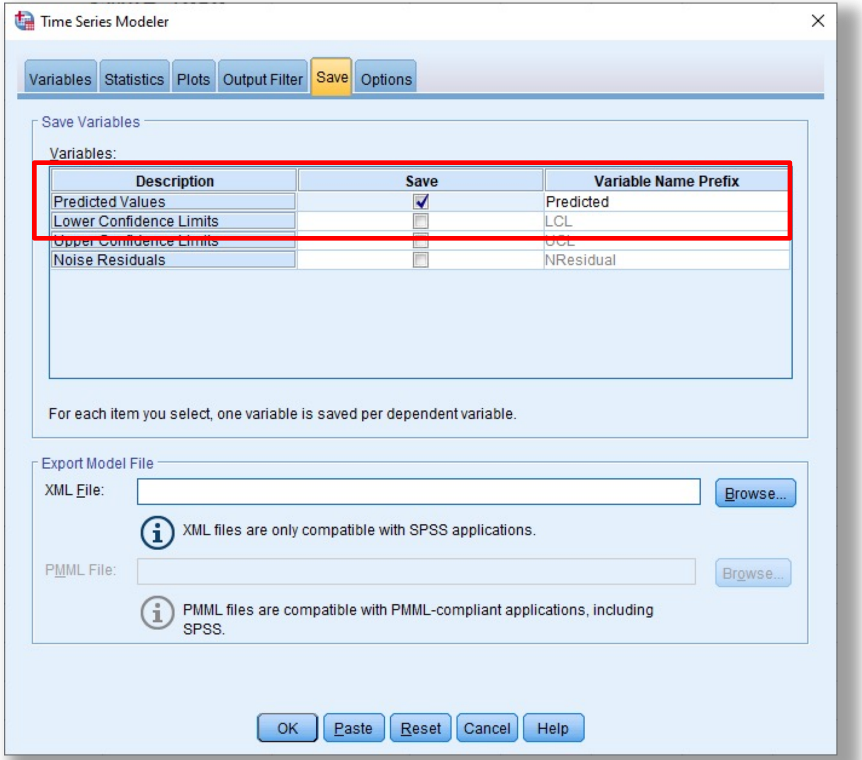
- One of the simplest ways to create new data *showing forecasted values* is to simply add the forecasts to the end of the data file itself.
- In our current example file, it's necessary to add values for the two independent variables **mail** and **strike** because they are part of the model and will therefore be needed to generate these new values.



	date	clothing_sales	mail	strike	YEAR_	MONTH_	DATE_
113	05/01/1998	41038.75	11951	0	1998	5	MAY 1998
114	06/01/1998	48329.36	11706	0	1998	6	JUN 1998
115	07/01/1998	44961.98	11460	0	1998	7	JUL 1998
116	08/01/1998	58660.76	11808	0	1998	8	AUG 1998
117	09/01/1998	57791.14	12781	0	1998	9	SEP 1998
118	10/01/1998	56329.40	11690	0	1998	10	OCT 1998
119	11/01/1998	54617.35	11393	0	1998	11	NOV 1998
120	12/01/1998	80245.97	15263	0	1998	12	DEC 1998
121	.	.	14370	0	1999	1	JAN 1999
122	.	.	11890	0	1999	2	FEB 1999
123	.	.	11722	0	1999	3	MAR 1999
124	.	.	11589	0	1999	4	APR 1999
125	.	.	11633	0	1999	5	MAY 1999
126	.	.	11951	0	1999	6	JUN 1999
127	.	.	11706	0	1999	7	JUL 1999
128	.	.	11460	0	1999	8	AUG 1999
129	.	.	11808	0	1999	9	SEP 1999
130	.	.	12781	0	1999	10	OCT 1999
131	.	.	11690	0	1999	11	NOV 1999
132	.	.	11393	0	1999	12	DEC 1999

Generating Forecasts

- Simply return to the Time Series dialog and click the tab marked:
  - **Save**
- Then check the appropriate options to request the predicted values



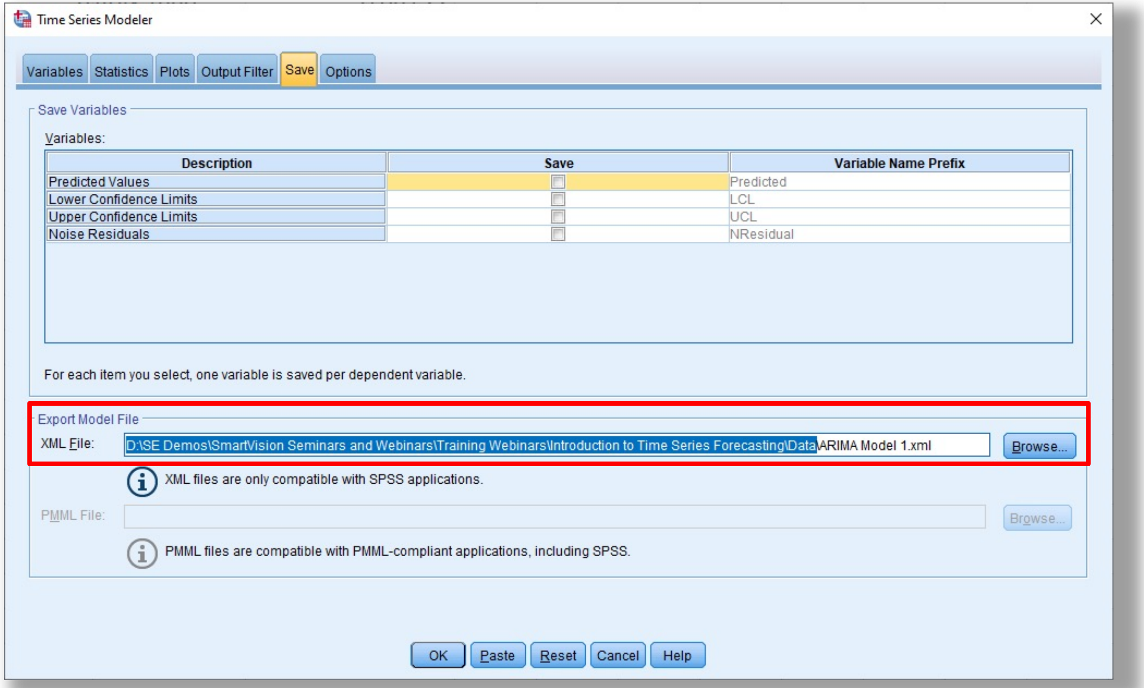
Generating Forecasts

- The forecasted clothing sales for the year 1999 are now shown in the dataset.

date	clothing_sales	mail	strike	YEAR_	MONTH_	DATE_	Predicted_clothing_sales_Model_1
05/01/1998	41038.75	11951	0	1998	5	MAY 1998	47359.96
06/01/1998	48329.36	11706	0	1998	6	JUN 1998	47391.22
07/01/1998	44961.98	11460	0	1998	7	JUL 1998	49934.48
08/01/1998	58660.76	11808	0	1998	8	AUG 1998	45590.14
09/01/1998	57791.14	12781	0	1998	9	SEP 1998	54405.31
10/01/1998	56329.40	11690	0	1998	10	OCT 1998	56927.51
11/01/1998	54617.35	11393	0	1998	11	NOV 1998	53767.85
12/01/1998	80245.97	15263	0	1998	12	DEC 1998	73884.71
.	.	14370	0	1999	1	JAN 1999	34621.09
.	.	11890	0	1999	2	FEB 1999	33784.23
.	.	11722	0	1999	3	MAR 1999	48400.32
.	.	11589	0	1999	4	APR 1999	49167.51
.	.	11633	0	1999	5	MAY 1999	49522.23
.	.	11951	0	1999	6	JUN 1999	55169.71
.	.	11706	0	1999	7	JUL 1999	45732.82
.	.	11460	0	1999	8	AUG 1999	50807.40
.	.	11808	0	1999	9	SEP 1999	59576.83
.	.	12781	0	1999	10	OCT 1999	68085.74
.	.	11690	0	1999	11	NOV 1999	54808.51
.	.	11393	0	1999	12	DEC 1999	76528.76

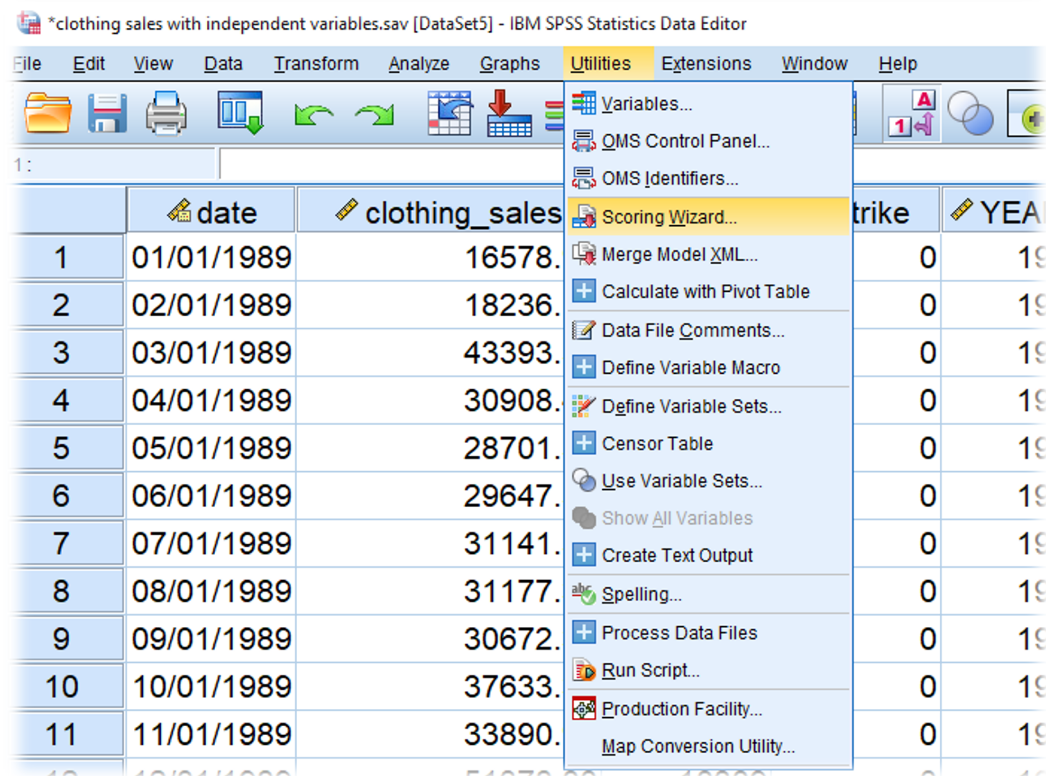
Generating Forecasts

- An alternative approach is to save the model as an XML file.



Generating Forecasts

- We can apply the saved xml model file to a new dataset by using the SPSS **Scoring Wizard**



For a deeper dive, try our online course

Introduction to Time Series Forecasting with IBM SPSS Statistics

Home / Online courses / Introduction to Time Series Forecasting with IBM SPSS Statistics

176 students

Overview Curriculum Instructor

Introduction to Time Series Forecasting

Course materials

£75.00

Buy this course

Download our e-book for free

The insider's guide to predictive analytics

£0.00

- 1 +

Add to basket

Category: books

SMARTVISION Europe

A SELECT INTERNATIONAL COMPANY

Working with Smart Vision Europe Ltd.

- **Sourcing Software**
  - You can buy your analytical software from us often with discounts
  - Assist with selection, pilot, implementation & support of analytical tools
  - http://www.sv-europe.com/buy-spss-online/
- **Training and Consulting Services**
  - Guided consulting & training to develop in house skills
  - Delivery of classroom training courses / side by side training support
  - Identification & recruitment of analytical skills into your organisation
- **Advice and Support**
  - offer 'no strings attached' technical and business advice relating to analytical activities
  - Technical support services around SPSS

Contact us:

+44 (0)207 786 3568
info@sv-europe.com
Twitter: @sveurope
Follow us on Linked In
Sign up for our Newsletter

Thank you