Chapter 164

Lag Plots

Introduction

A *lag plot* is used to help evaluate whether the values in a dataset or time series are random. If the data are random, the lag plot will exhibit no identifiable pattern. If the data are not random, the lag plot will demonstrate a clearly identifiable pattern. The type of pattern can aid the user in identifying the non-random structure in the data. Lag plots can also help to identify outliers.

This procedure will generate all lag plots up to a user-specified maximum lag *k* in a single run.

Definition of Lag

For data values Y_1, Y_2, …, Y_N, the *k*-period (or *k*th) lag of the value Y_i is defined as the data point that occurred *k* time points before time *i*. That is

$$\text{Lag}_k(Y_i) = Y_{i-k}$$

For example, $\text{Lag}_1(Y_2) = Y_1$ and $\text{Lag}_3(Y_{10}) = Y_7$.

Lag Plot Patterns

The pattern in the lag plot can help you identify possibly hidden patterns in the data. The following sections give some examples of random and non-random data structures and their resulting lag plots.

Random Data

Random data gives rise to lag plots with no pattern. The points in the lag plot appear scattered from left to right and top to bottom.
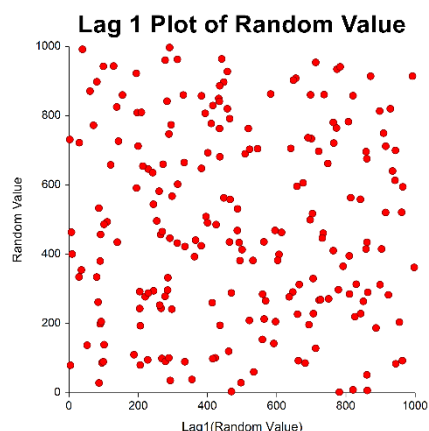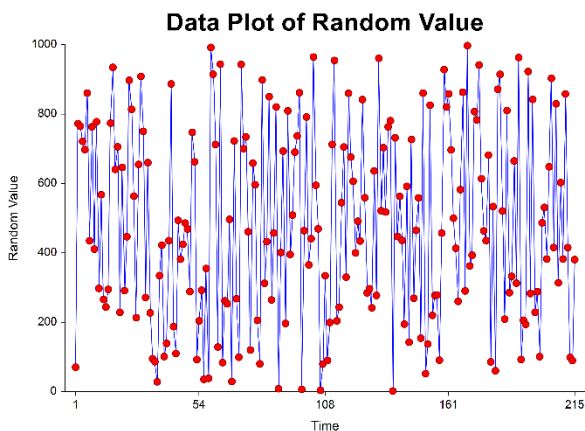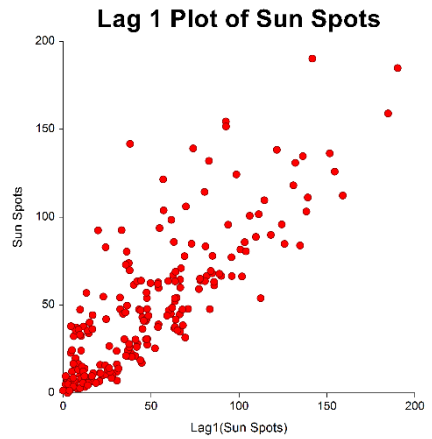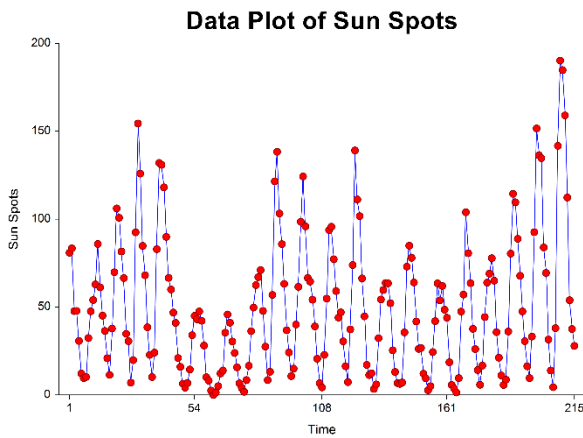
Data with Autocorrelation

Data with autocorrelation gives rise to lag plots with linear patterns that follow the diagonal. As the level of autocorrelation increases, the points cluster more tightly along the diagonal.
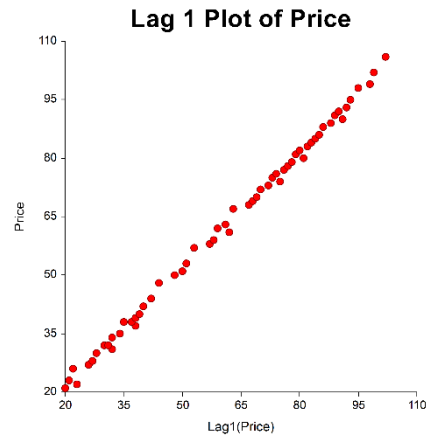
Data with Weak Autocorrelation
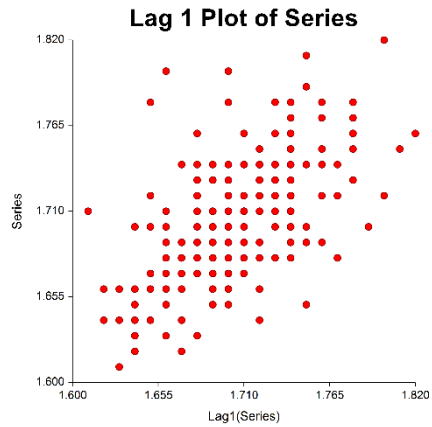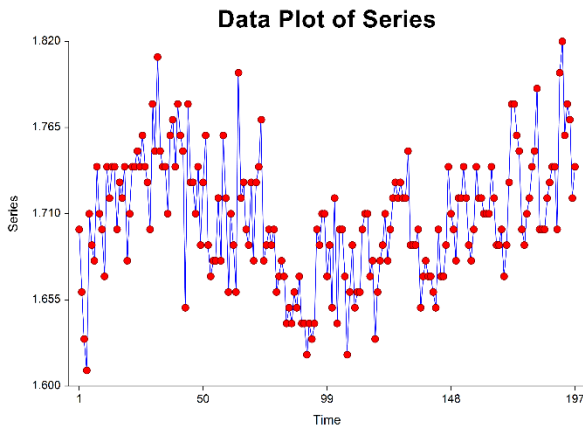
Data with Moderate Autocorrelation

Data with High Autocorrelation

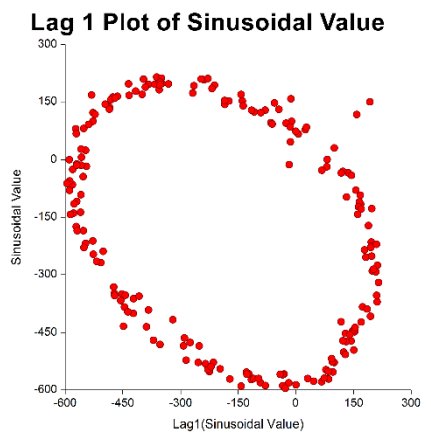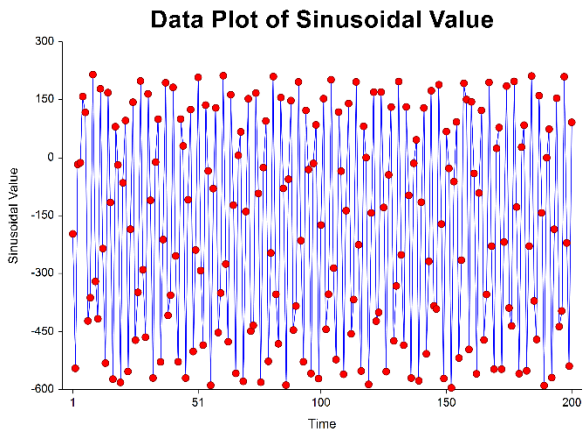## Sinusoidal Data

Single-cycle sinusoidal data gives rise to lag plots with circular or elliptical patterns. Values lying off the ellipse should be considered as potential outliers.

Data Structure

Each lag plot is constructed from a single numeric variable. If multiple variables are entered, a separate lag series is created for each variable. A data label variable can also be entered.

Example 1 – Creating Lag Plots

This section presents a tutorial on generating 1st and 2nd order lag plots. To run this example, take the following steps using the Sunspot dataset.

Setup

To run this example, complete the following steps:

1   Open the Sunspot example dataset
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Sunspot** and click **OK**.

2   Specify the Lag Plots procedure options
- Find and open the **Lag Plots** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab
_____

Variable(s) ....**Spots**
Maximum Lag Plotted **2**

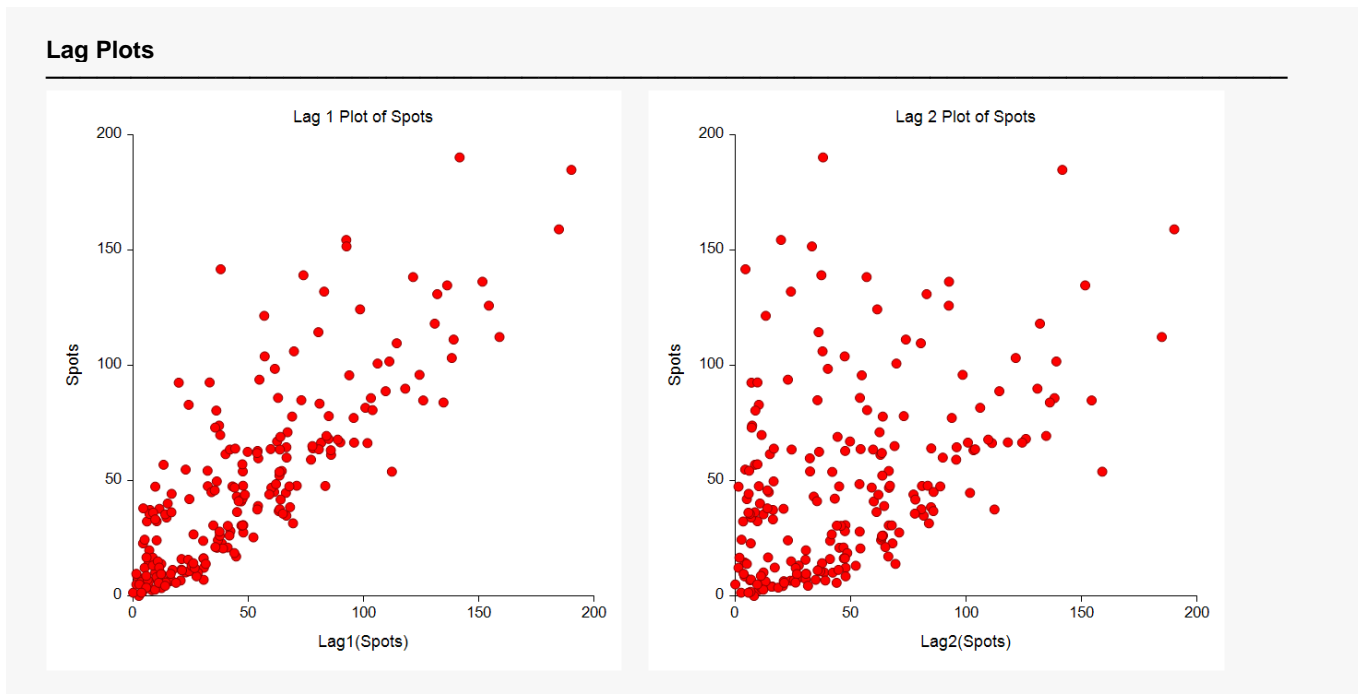**3   Run the procedure**
- Click the **Run** button to perform the calculations and generate the output.

Output

This section displays the lag plots. In the output screen you can double-click any of the plots to see it in full size.

Lag Plots
_____

The data variable "Spots" is plotted against the 1st and 2nd order lags in these two plots. The linear pattern in the 1st order lag plot indicates probable autocorrelation.

Example 2 – Combining Multiple Series on a Single Lag Plot

This section presents a tutorial on generating 1st and 2nd order lag series and combining the results onto a single plot. To run this example, take the following steps using the Sinusoidal dataset.

Setup

To run this example, complete the following steps:

1   Open the Sinusoidal example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Sinusoidal** and click **OK**.

2   Specify the Lag Plots procedure options

- Find and open the **Lag Plots** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.
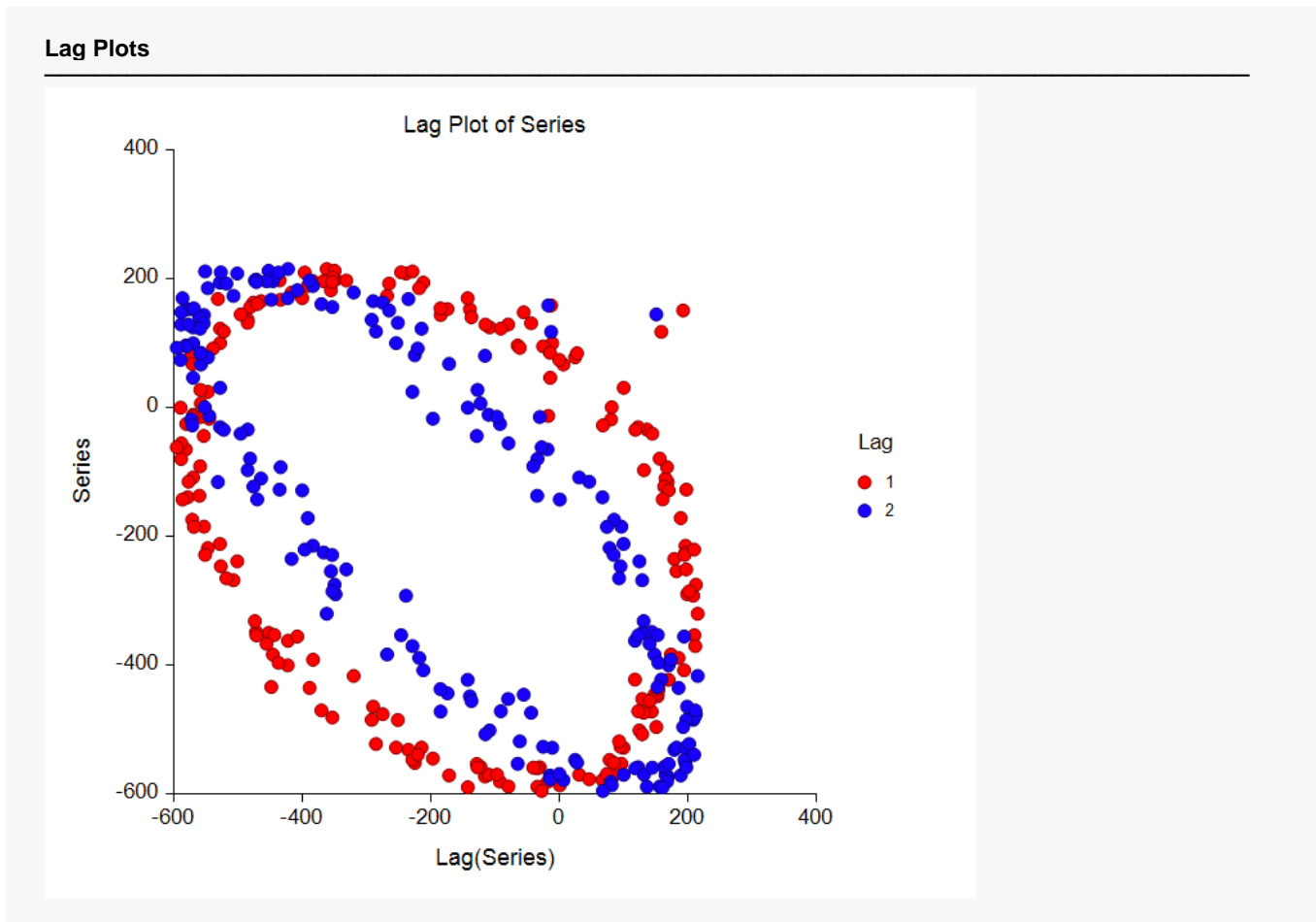
Variables Tab

Variable(s) ....**Series**
Maximum Lag Plotted **2**
Plot Layout...**All Lags Together on a Single Plot**

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

This section displays the lag plots. In the output screen you can double-click any of the plots to see it in full size.

Lag Plots
_____

The data variable "Series" is plotted against the 1st and 2nd order lags with the results combined in a single plot. The sinusoidal data exhibits the typical elliptical pattern in the lag plot.