Seasonal Decomposition

Summary
The **Seasonal Decomposition** procedure divides a time series into three components:

1. trend-cycle
2. seasonality
3. irregular

Each component may be separately plotted or saved. In addition, the decomposition can be used to create a seasonally adjusted version of the original time series. Seasonal subseries and annual subseries plot may also be created.

Sample StatFolio: *tsdecomp.sgp*

Sample Data:
The file *golden gate.sgd* contains monthly traffic volumes on the Golden Gate Bridge in San Francisco for a period of $n = 168$ months from January, 1968 through December, 1981. The table below shows a partial list of the data from that file:

Month	Traffic
1/68	73.637
2/68	77.136
3/68	81.481
4/68	84.127
5/68	84.562
6/68	91.959
7/68	94.174
8/68	96.087
9/68	88.952
10/68	83.479
11/68	80.814
12/68	77.466
1/69	75.225
…	…

The data were obtained from a publication of the Golden Gate Bridge.

Data Input

The data input dialog box requests the name of the column containing the time series data:

- **Data:** numeric column containing *n* equally spaced numeric observations.

- **Time indices:** time, date or other index associated with each observation. Each value in this column must be unique and arranged in ascending order.

- **Sampling Interval:** If time indices are not provided, this defines the interval between successive observations. For example, the data from the Golden Gate Bridge were collected once every *month*, beginning in 01/68.

- **Seasonality**: the length of seasonality *s*, or number of observations in a full cycle of the seasonal pattern. For example, monthly data such as traffic on the Golden Gate Bridge have a seasonality of $s = 12$. Hourly data that repeat every day have a seasonality of $s = 24$.

- **Trading Days Adjustment**: a numeric variable with n observations used to normalize the original observations, such as the number of working days in a month. The observations in the *Data* column will be divided by these values before being plotted or analyzed.

- **Select:** subset selection.

Statistical Model

In analyzing time series data, it is common to view the data as consisting of several components:

1. *Trend* (T) – a general long-term pattern observed over the entire data set. For example, many economic time series tend to show an increasing trend when viewed over many years.

2. *Cycle (C)* – cyclical variations around the trend line. Unlike seasonal effects, these cycles do not have a fixed frequency. General up and downs of the general world economy is a typical example.

3. *Seasonality (S)* – cyclical variations with a fixed frequency, such as yearly cycles in the sales of lawnmowers. Seasonal effects repeat on a regular and predictable basis.

4. *Random or Irregular (R)* – the residual component left behind after the other three components are accounted for.

There are two basic models upon which a decomposition of a time series into its component parts may be based: a *multiplicative* model and an *additive* model. The multiplicative model assumes that the data at time *t* may be represented as the product of the four components according to:

$$Y_t = T_t C_t S_t R_t \qquad (1)$$

The additive model assumes that the components add:

$$Y_t = T_t + C_t + S_t + R_t \qquad (2)$$

The goal of the *Seasonal Decomposition* procedure is to divide an observed time series into its component parts. In particular, the procedure derives:

1. *Seasonal indices* representing the effect of each season. Knowing the effect of different seasons is often quite important.

2. An estimated of the combined *trend-cycle* component. No attempt is made to separate these two components, however, since both represent relatively long-term effects.

3. The residuals (irregular component).

Analysis Summary

The *Analysis Summary* displays the number of observations in the time series, the length of seasonality, and the decomposition method selected.

Seasonal Decomposition - Traffic
Data variable: Traffic (Golden Gate Bridge Traffic Volume)

Number of observations = 168
Start index = 1/68
Sampling interval = 1.0 month(s)
Length of seasonality = 12

Seasonal Decomposition
Method: Multiplicative

Note: a limited amount of missing data is permitted, providing there are not too many missing values close together. Missing values are replaced by interpolated values according to the method outlined in the *Calculations* section of the *Time Series – Descriptive Methods* documentation.

Analysis Options

Analysis Options permits the data to be transformed before being plotted or analyzed:
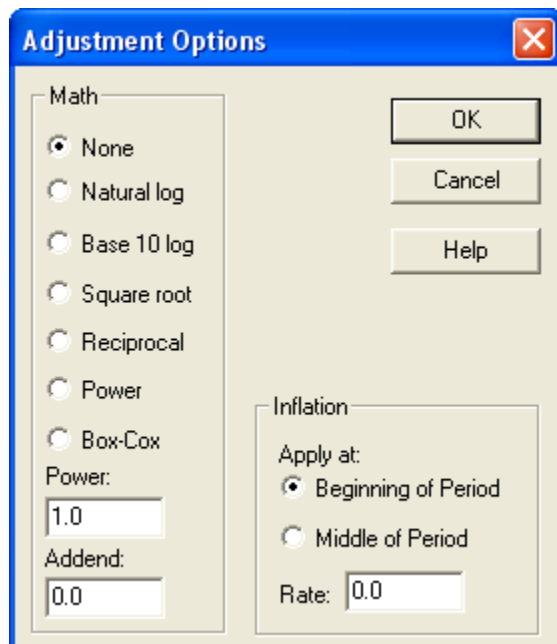
- **Math**: transforms the data by performing the indicated mathematical operation.

- **Inflation**: adjusts the data for inflation using the specified inflation rate

If more than one transformation is requested, they are applied in the following order:

1. trading day adjustment
2. inflation adjustment
3. math adjustment

For a detailed discussion of the transformation options, refer to the *Time Series – Descriptive Methods* documentation.

Data Table

The *Data Table* displays the results of the decomposition:

Data Table for Traffic
Seasonal decomposition method: Multiplicative

Period	Data	Trend-Cycle	Seasonality	Irregular	Seasonally Adjusted
1/68	73.637				81.0636
2/68	77.136				82.2507
3/68	81.481				84.2308
4/68	84.127				84.6577
5/68	84.562				84.4944
6/68	91.959				87.0608
7/68	94.174	84.5557	111.375	103.054	87.1383
8/68	96.087	84.7169	113.421	102.394	86.7454
9/68	88.952	84.9508	104.71	100.136	85.0664
10/68	83.479	85.1548	98.032	98.5183	83.8931
11/68	80.814	85.342	94.6943	98.1288	83.7451
12/68	77.466	85.5072	90.5959	96.2271	82.2811
1/69	75.225	85.6004	87.8793	96.7423	82.8118
2/69	79.418	85.7466	92.6194	98.7608	84.684
3/69	84.813	86.0028	98.6166	101.945	87.6753
4/69	85.691	86.3519	99.2347	99.8607	86.2316
5/69	87.49	86.7945	100.801	100.721	87.42
6/69	92.995	87.3052	106.517	100.844	88.0417
7/69	95.375	87.7297	108.715	100.593	88.2496
8/69	98.396	88.1522	111.621	100.769	88.83
9/69	92.791	88.5946	104.737	100.161	88.7377

Included in the table are:

- **Data**: the original time series Y_t, including any replacement values that have been calculated for missing data.

- **Trend-Cycle**: an estimate of the combined trend-cycle component ($T_t C_t$ for a multiplicative decomposition and $T_t + C_t$ for an additive decomposition*).*

- **Seasonality**: the estimated seasonal component S_t.

- **Irregular**: the estimated residual or irregular component R_t.

- **Seasonally Adjusted Data**: the original data with only the seasonality removed.
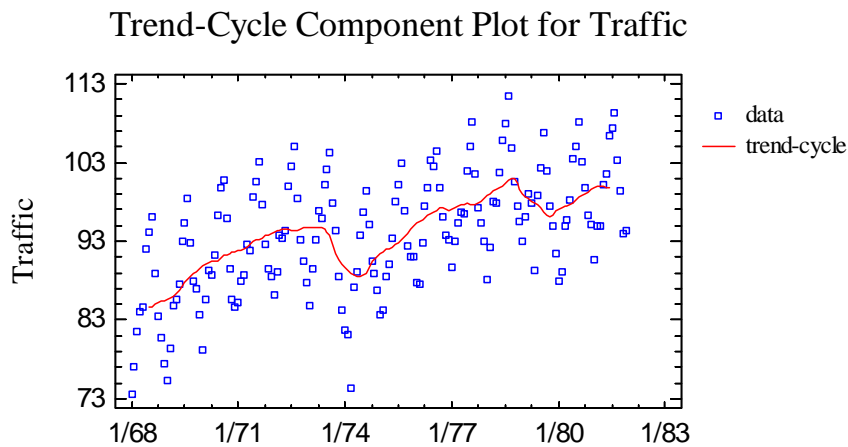
Pane Options
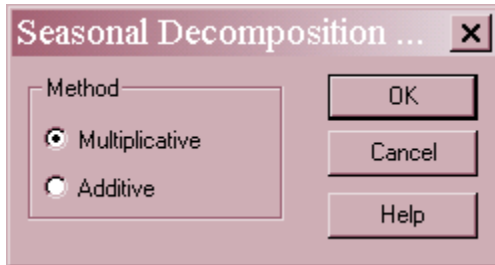
- **Method**: the type of decomposition performed.

Trend-Cycle Plot

The *Trend-Cycle* plot shows the estimated trend-cycle component.

The trend-cycle component is estimated by smoothing the time series data using a simple moving average with span *k* equal to the length of seasonality *s*. In the traffic data shown above, the general pattern is upward, although there were two major shocks to the system that had a major impact on the basic trend.

Seasonal Indices

Once the trend-cycle has been estimated, it can be removed from the data. For a multiplicative model, this is done by dividing the original data by the estimated component (called a 'ratio-to-moving average"), leaving:

$$\hat{S}_t \hat{R}_t = \frac{Y_t}{\hat{T}\hat{C}_t} \tag{3}$$

For an additive model, the trend-cycle is subtracted from the original data, leaving:

$$\hat{S}_t \hat{R}_t = Y_t - \hat{T}\hat{C}_t \tag{4}$$

The resulting estimates of the seasonal-irregular component are then averaged across all observations within each season to remove the irregular component, resulting in an estimate of the seasonal component. The seasonal components are then adjusted so that an average season has a value of 1.0 if using the multiplicative method and 0 if using the additive method.

The *Seasonal Indices* table displays the results:

Seasonal Indices for Traffic
Seasonal decomposition method: Multiplicative

Season	Index
1	90.8385
2	93.7815
3	96.7354
4	99.3731
5	100.08
6	105.626
7	108.074
8	110.769
9	104.568
10	99.5064
11	96.5
12	94.148

This table shows the estimated seasonal component S_t in the traditional format of a percentage, such that an average season would have an index equal to 100. For example, the traffic index in August is approximately 110.8, meaning that traffic across the Golden Gate Bridge in August is 10.8% higher than average.
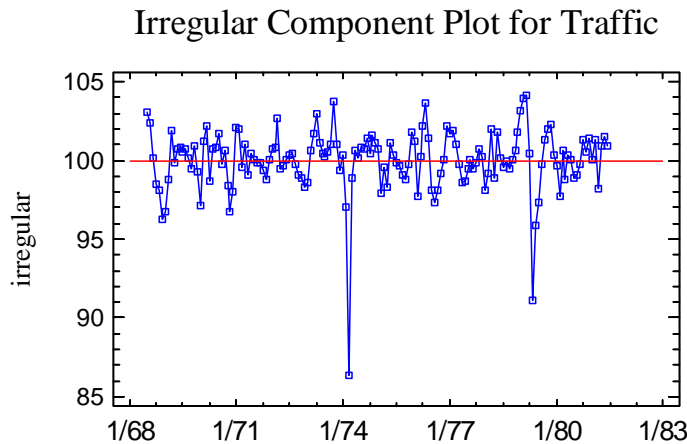
Seasonal Indices Plot

The *Seasonal Indices* plot displays the estimated indices:

Seasonal Index Plot for Traffic
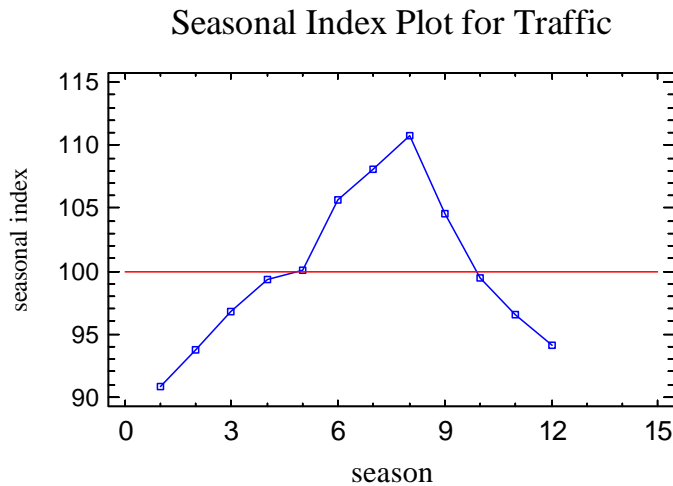
Irregular Component

The *Irregular Component* plot displays the residual or irregular component R_t at each time period:

Irregular Component Plot for Traffic

For a multiplicative model, the estimated irregular component is obtained by dividing the observations by the estimated trend-cycle and seasonal components:

$$\hat{R}_t = \frac{Y_t}{\hat{T}\hat{C}_t\hat{S}_t} \tag{5}$$

and then normalized so that the average residual equal 1.0 (corresponding to an index of 100). For an additive model, the irregular component is obtained by subtraction:

$$\hat{R}_t = Y_t - \hat{T}\hat{C}_t - \hat{S}_t \tag{6}$$

and then normalized so that the average residual equals 0.

In the sample data, note the large residual in March of 1974, equaling approximately 86.4. This indicates that traffic in that month was 13.6% below what would have been expected, given the estimated trend-cycle and seasonal effects.

Seasonally Adjusted Data

Once all of the components have been estimated, the original time series can be seasonally adjusted by removing from it only the seasonal effects, leaving both the trend-cycle and irregular components. For a multiplicative decomposition, the seasonally adjusted data is given by

$$Y_t^{adj} = \frac{Y_t}{\hat{S}_t} \tag{7}$$

For an additive decomposition, the seasonally adjusted data is given by

$$Y_t^{adj} = Y_t - \hat{S}_t \tag{8}$$

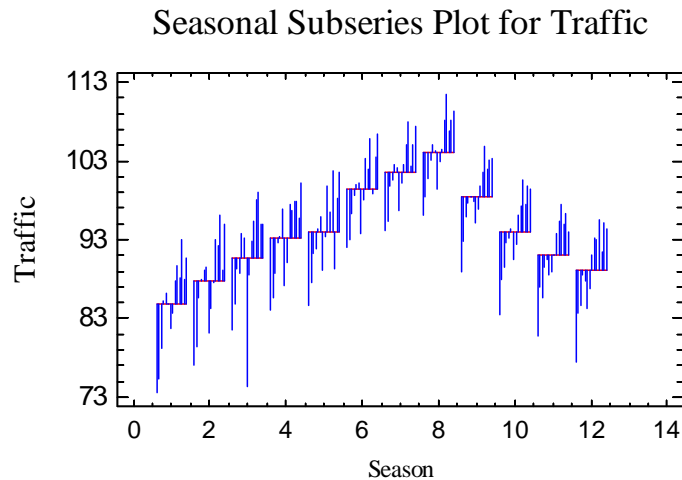The *Seasonally Adjusted Data* plot displays the results:

Seasonally Adjusted Data Plot for Traffic
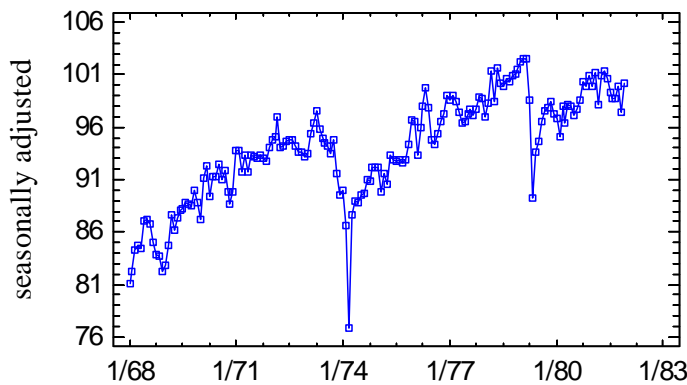
The effect of the gasoline shortages on the basic trend line is now quite apparent.

Seasonal Subseries Plot

An alternative way to plot seasonal data is through a *Seasonal Subseries Plot*:

Seasonal Subseries Plot for Traffic

This plot is constructed as follows:

1. The observations corresponding to each season are collected and horizontal lines drawn at the average for the season.

2. Vertical lines are drawn from each observation to the average of the season it corresponds to.

In such a plot, one can see all of the time series components:

(i) The seasonal pattern is observable by looking at the differences between the averages of each season.
(ii) The trend-cycle is observable by looking at patterns *within each season*.
(iii) The irregular component is observable by looking at the length of the vertical lines contrasted with the trend within each season.
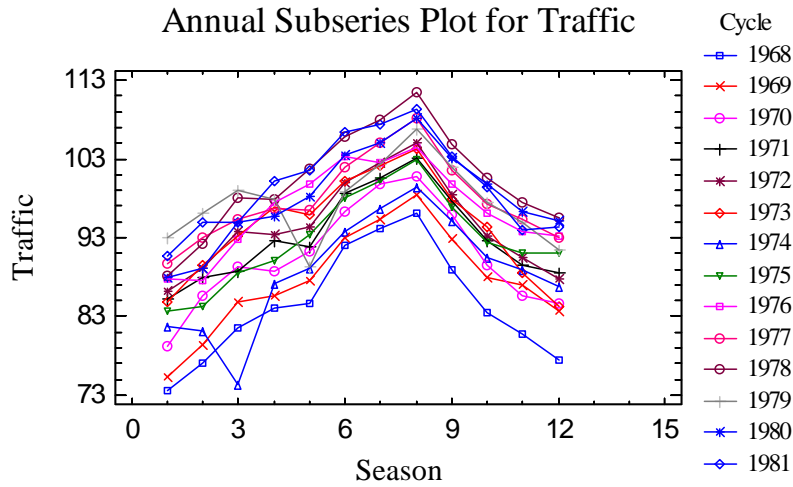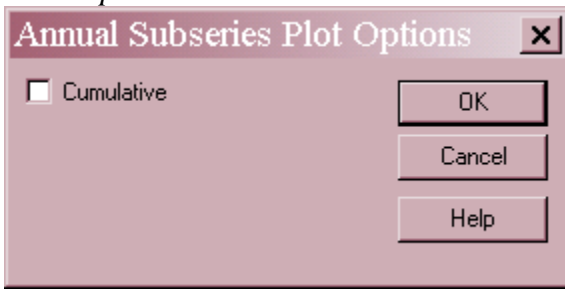
Pane Options
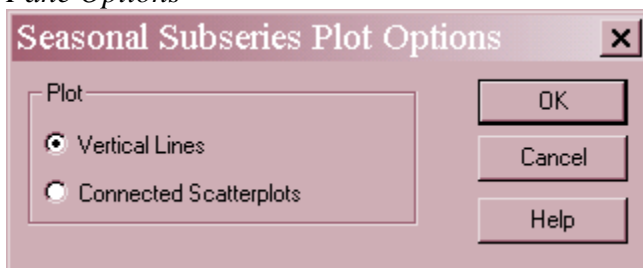
- **Vertical Lines** – draw a line from each observation to the average for its season.
- **Connected Scatterplots** – draw the data within each season as a connected X-Y plot.

Annual Subseries Plot

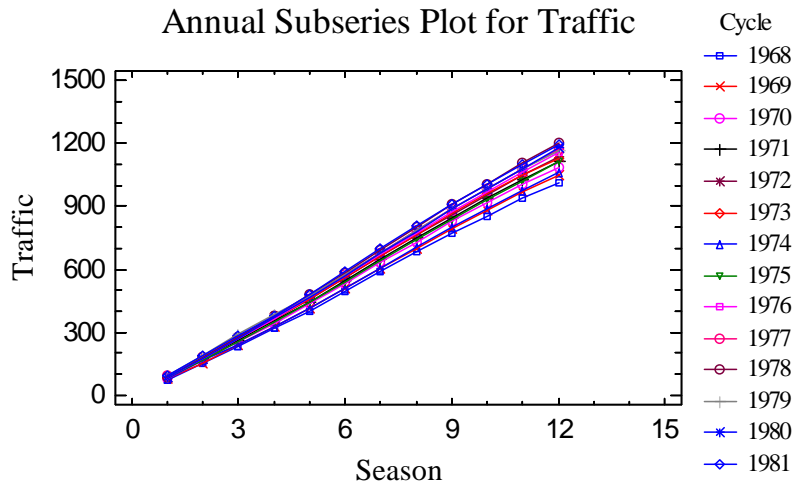The *Annual Subseries Plot* displays each cycle as a separate plot:

Pane Options

- **Cumulative**: if checked, the value plotted on the vertical axis is the cumulative sum of the observations during the cycle.

Example – Cumulative Plot

Save Results

The following results may be saved to the datasheet:

1. *Data* – the original time series data, after replacement of any missing values.

2. *Trend-cycle* – the estimated trend-cycle component.

3. *Seasonal indices* – the estimated seasonal indices.

4. *Irregular* – the irregular component.

5. *Seasonally adjusted data* – the seasonally adjusted data.

Note: if a multiplicative model was used, the *Seasonal indices* and *Irregular* components will be normalized so that the average value equals 100. For an additive model, the *Seasonal indices* will be normalized so that the average value equals 0.