

**Seasonal Adjustment using  
X-13ARIMA-SEATS**



Revised: 10/9/2017



Summary .....	1
Data Input.....	3
Limitations .....	4
Analysis Options .....	5
Tables and Graphs.....	6
Analysis Summary .....	7
Data Table.....	9
Trend-Cycle Plot.....	10
Seasonality Plot.....	11
Seasonality Table .....	12
Seasonality Plot.....	13
Irregular Component.....	14
Seasonally Adjusted Data .....	15
Seasonal Subseries Plot.....	16
Annual Subseries Plot.....	17
References.....	18

**Summary**

The **X-13ARIMA-SEATS Seasonal Adjustment** performs a seasonal adjustment of time series data using the procedure currently employed by the United States Census Bureau. As part of the procedure, the time series is decomposed into 3 components:

1. a trend-cycle component
2. a seasonal component
3. an irregular component

Each component may be plotted separately or saved, together with the seasonally adjusted data.

The seasonal adjustment calculations are performed by the “seasonal” package in R. To run the procedure, R must be installed on your computer together with the “seasonal” package. For information on downloading and installing R, refer to the document titled “R – Installation and Configuration”.

### Sample StatFolio: *X13ARIMASEATS.sgp*

### Sample Data

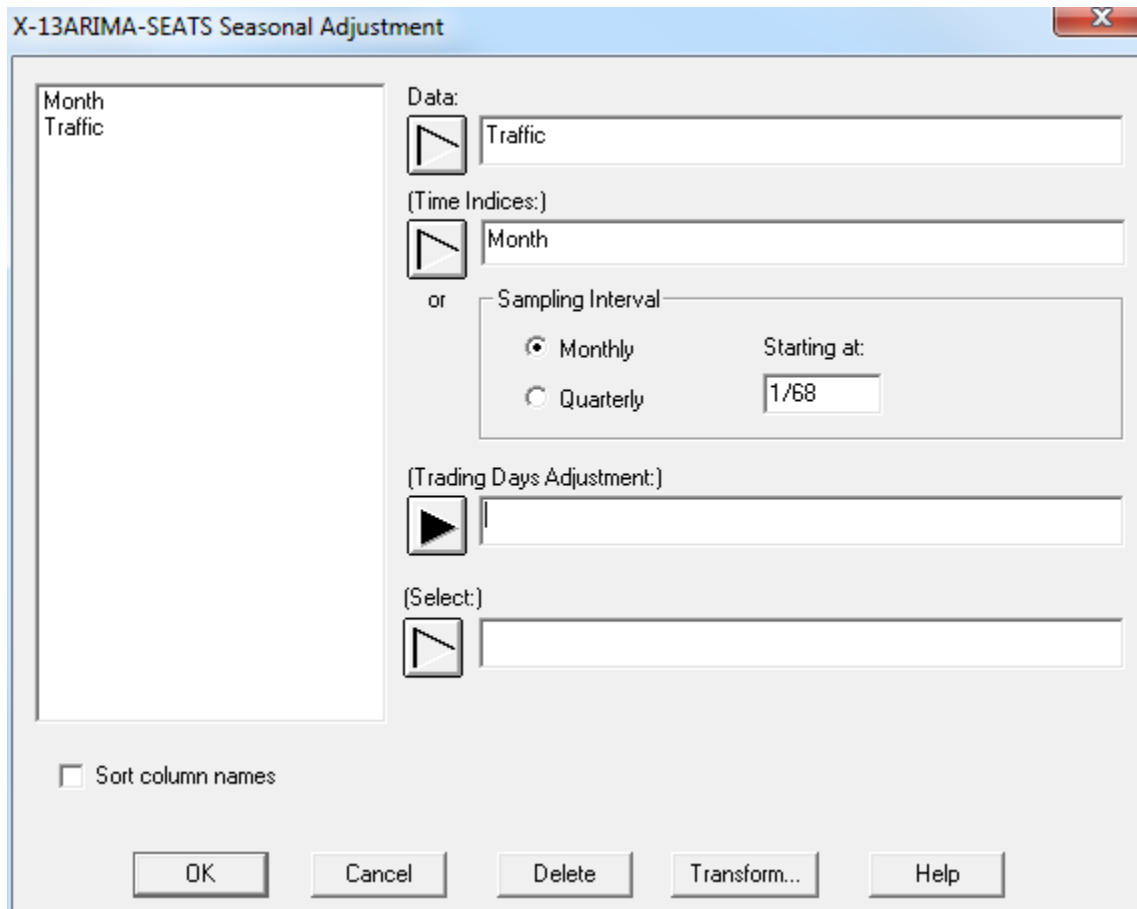
The file *golden gate.sgd* contains monthly traffic volumes on the Golden Gate Bridge in San Francisco for a period of  $n = 168$  months from January, 1968 through December, 1981. The table below shows a partial list of the data from that file:

<i>Month</i>	<i>Traffic</i>
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 and information about the time periods covered. This procedure only supports data recorded on a monthly or quarterly basis:



- **Data:** numeric column containing  $n$  equally spaced numeric observations.
- **Time indices:** data associated with each observation. Each value in this column must be unique and arranged in ascending order with no gaps. For monthly data, the column must have a type equal to 'month'. For quarterly data, the column must have a type equal to 'quarter'.
- **Sampling Interval:** If time indices are not provided, this defines the type of data (monthly or quarterly) and the time corresponding to the first row in the data file. For example, the data from the Golden Gate Bridge were collected once every month, beginning in 01/68. Note: this section is ignored if a column containing time indices is provided.
- **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. Note: if you wish to use the trading day feature contained within X-13ARIMA-SEATS, leave this field blank.

- **Select:** subset selection.

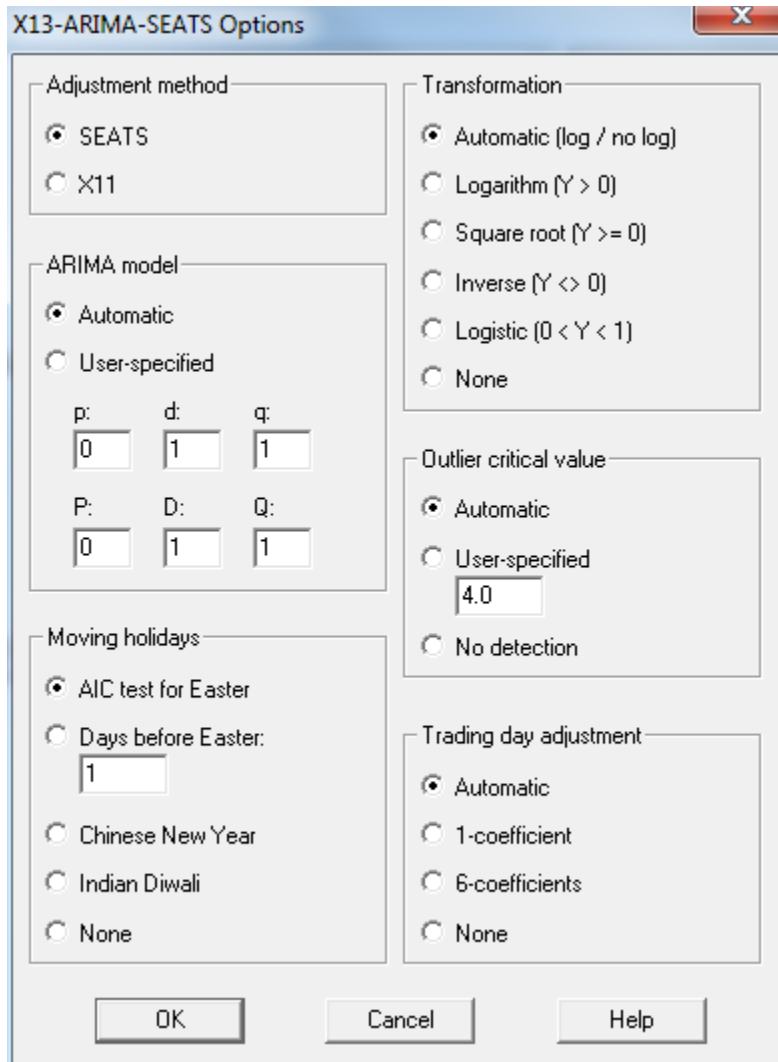
## Limitations

The X-13ARIMA-SEATS program has certain limitations. The most important are:

1. The data must be either monthly with a seasonality equal to 12, or quarterly with a seasonality equal to 4.
2. The maximum number of observations in the time series cannot exceed  $n = 780$ .
3. The automated trading day adjustment can only be applied if the starting year is 1776 or later.
4. The automated adjustment for Easter can only be applied if the starting year is 1901 or later.
5. A limited amount of missing data is permitted, provided that 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

The *Analysis Options* dialog box sets the options used by the seasonal adjustment procedure:

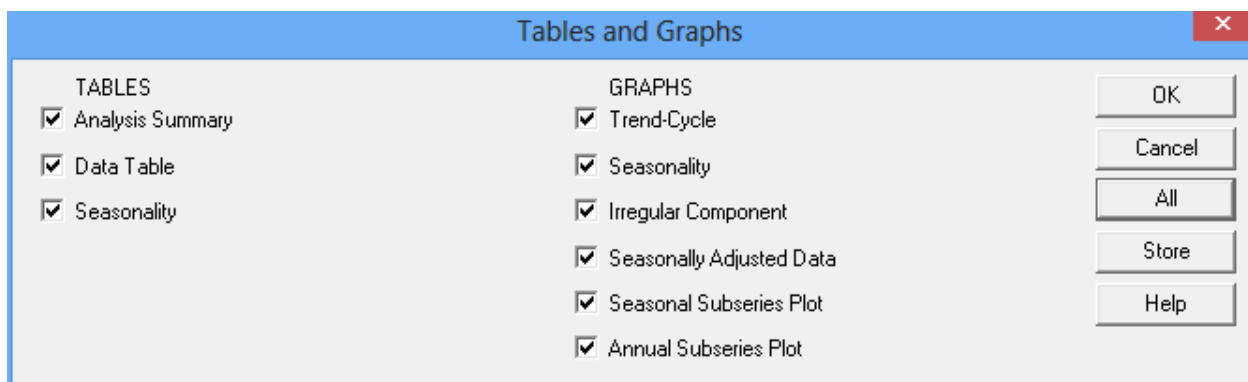


- **Adjustment method:** Specifies the method used to perform the adjustment. The default is X13-SEATS, but the earlier X-11 procedure can be selected instead.
- **Transformation:** Specifies the transformation to be applied to the time series before fitting the model. “Automatic” causes the program to select the best transformation internally. Other settings force particular transformations to be used.
- **ARIMA Model:** Specifies the type of ARIMA model used to model the time series. “Automatic” causes the program to select the best model internally. “User-specified” allows the user to select the orders of the nonseasonal part of the model (p,d,q) and the seasonal part of the model (P,D,Q), where  $p$  is the order of the AR term,  $d$  is the order of differencing, and  $q$  is the order of the MA term.

- **Outlier critical value:** The critical value of a Z test used to identify outliers (between 3 and 5). The lower the value, the more likely it is that a point will be declared to be an outlier. If set to “Automatic”, the critical value will be set based on the number of observations in the time series.
- **Movable holidays:** Adds effects to adjust for the specified holiday. “AIC test for Easter” automatically compares effects lasting different numbers of days and selects the best. “Days before Easter” specifies an effect covering the indicated days before Easter Sunday.
- **Trading day adjustment:** Adds effects to adjust for the number of trading days in a month or quarter. “Automatic” compares models with 1 and 6 coefficients and picks the better one.

## Tables and Graphs

The following tables and graphs may be created:



## Analysis Summary

The *Analysis Summary* displays the output generated by R:

### X-13ARIMA-SEATS Seasonal Adjustment - Traffic

Data variable: Traffic

Number of observations = 168

Time indices: Month

Length of seasonality = 12

#### Output From R

X-13ARIMA-SEATS Seasonal Adjustment

```

TSData <-
read.csv("C:\\Users\\NEIL~1.STA\\AppData\\Local\\Temp\\statgraphics_data.csv",dec=".")
'
  sep=",",stringsAsFactors=FALSE)
library("seasonal")
t=ts(TSData$x,start=c(1968,1),frequency=12)
s=seas(t)
summary(s)

##
## Call:
## seas(x = t)
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## Leap Year      -1.69562    0.49720  -3.410 0.000649 ***
## Weekday         0.02219    0.01995   1.112 0.265948
## Easter[15]      2.22602    0.34374   6.476 9.43e-11 ***
## AO1974.Mar     -9.16669    0.91052 -10.068 < 2e-16 ***
## LS1979.Apr     -5.22118    1.30117  -4.013 6.00e-05 ***
## AO1979.May     -7.19946    0.90444  -7.960 1.72e-15 ***
## MA-Seasonal-12  0.79681    0.05152  15.466 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## SEATS adj. ARIMA: (0 1 0)(0 1 1) Obs.: 168 Transform: none
## AICc: 552, BIC: 575.4 QS (no seasonality in final): 0
## Box-Ljung (no autocorr.): 17.31 Shapiro (normality): 0.9851 .

results=data.frame(s$data)
write.csv(results,"C:\\Users\\NEIL~1.STA\\AppData\\Local\\Temp\\statgraphics_results.csv",
  row.names=FALSE

```

The output of the analysis is located in the lines beginning with **##**. A full description of the possible output is provided in the X-13ARIMA-SEATS Reference Manual (2016). Among the most important results are:

- **Coefficients:** estimated coefficients and significance tests. In the above output, there is a coefficient to model the effect of a leap year, the effect of weekdays versus weekends, an Easter effect beginning 15 days before Easter Sunday, 2 additive outliers (AO) at March 1974 and May 1979, a level shift (LS) during April 1979, and a seasonal moving average

parameter at lag 12. Temporary change (TC) effects and seasonal outliers (SO) are also possible.

- **SEATS adj:** the seasonal adjustment was performed by fitting an ARIMA model of order  $(0\ 1\ 0)\ (0\ 1\ 1)$ .
- **Obs:** there were  $n = 168$  observations in the time series.
- **Transform:** no transformation was performed on the data values before fitting the model.
- **AICc and BIC:** shows the values of the F-adjusted Achaike Information Criterion and the Bayes Information Criterion.
- **QS:** QS statistic used to test whether there is seasonality remaining in the final seasonally adjusted data. If  $Q$  is small, there is no remaining seasonality.
- **Box-Ljung:** value of the Box-Ljung statistic, used to test whether there is autocorrelation remaining in the residuals.
- **Shapiro:** P-value for the Shapiro-Wilk test for normality of the residuals. A small P-value would indicate that the residuals do not come from a normal distribution.



## Data Table

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

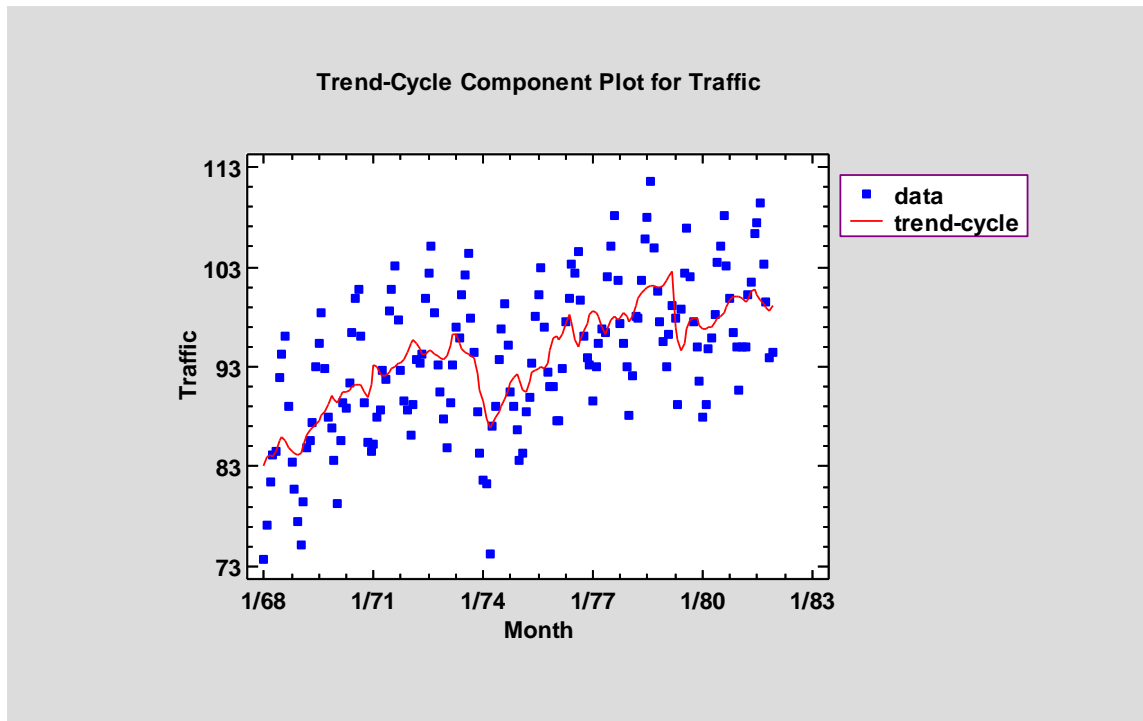
<i>Period</i>	<i>Data</i>	<i>Trend-Cycle</i>	<i>Seasonality</i>	<i>Irregular</i>	<i>Seasonally Adjusted</i>
1/68	73.637	83.0658	-9.18941	-0.305934	82.7598
2/68	77.136	83.8989	-5.76172	0.248347	84.1472
3/68	81.481	84.2564	-2.165	0.288595	84.545
4/68	84.127	84.0447	-0.520629	-0.251707	83.793
5/68	84.562	84.5076	0.453824	-0.465958	84.0416
6/68	91.959	85.5974	5.92172	0.550832	86.1482
7/68	94.174	86.0245	8.05086	0.0320708	86.0566
8/68	96.087	85.6312	10.2552	0.211774	85.8429
9/68	88.952	84.9513	4.20122	-0.167233	84.7841
10/68	83.479	84.5722	-0.978799	-0.180968	84.3912
11/68	80.814	84.412	-3.87296	0.308272	84.7202
12/68	77.466	84.1692	-6.33981	-0.352291	83.8169
1/69	75.225	84.3091	-9.16114	0.0104797	84.3196
2/69	79.418	85.1503	-5.78029	-0.375937	84.7744
3/69	84.813	86.1921	-2.2063	0.539043	86.7311
4/69	85.691	86.7114	-0.512279	-0.175605	86.5358
5/69	87.49	86.952	0.45623	0.0928854	87.0449
6/69	92.995	87.1825	5.8464	-0.000627182	87.1819
7/69	95.375	87.525	7.98855	-0.205106	87.3199
8/69	98.396	88.1007	10.2091	0.174878	88.2756
9/69	92.791	88.5568	4.21451	-0.024693	88.5321

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.
- **Seasonality:** the estimated seasonal component.
- **Irregular:** the estimated residual or irregular component.
- **Seasonally Adjusted Data:** the original data with only the seasonality removed.

## Trend-Cycle Plot

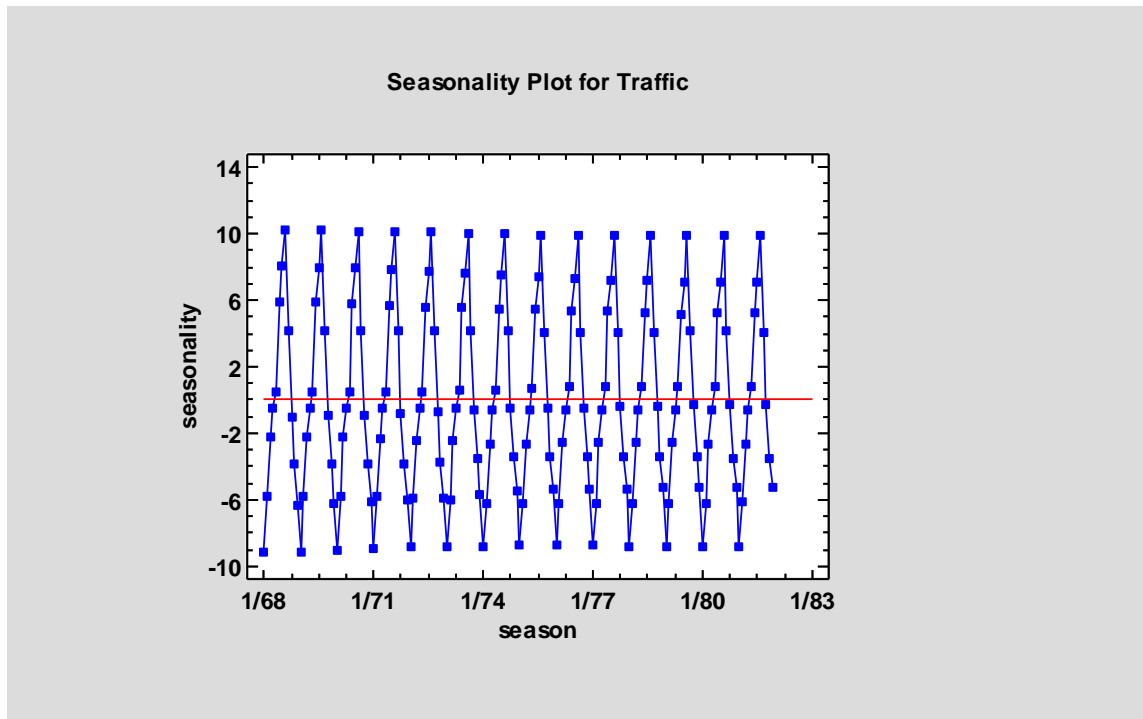
The *Trend-Cycle* plot shows the estimated trend-cycle component.



The trend-cycle component is estimated by smoothing the time series data. In the traffic data shown above, the general pattern is upward, although there were two major shocks to the system that had a significant impact on the basic trend.

## Seasonality Plot

The *Seasonality Plot* shows the estimated seasonality at each time period.



Seasonality represents the amount added to the trend-cycle component at each time period due to repeatable fluctuations at the seasonal frequency (once every 12 months for monthly data).

## Seasonality Table

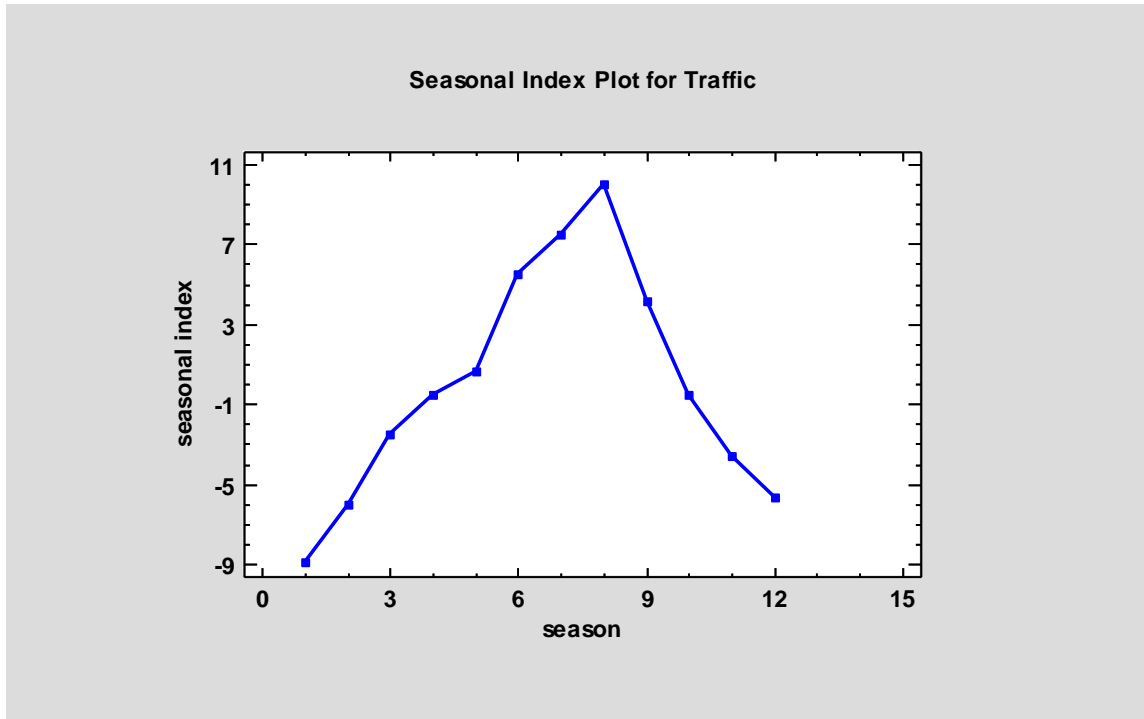
The *Seasonality* table displays the average effect of each season:

<b>Seasonal Indices for Traffic</b>	
Seasonal decomposition method: X-13ARIMA-SEATS	
<i>Season</i>	<i>Index</i>
1	-8.88135
2	-6.04812
3	-2.47301
4	-0.554707
5	0.644352
6	5.4973
7	7.50431
8	10.0089
9	4.14813
10	-0.562474
11	-3.58185
12	-5.63747

For example, the traffic index in August is approximately 10.0, meaning that traffic across the Golden Gate Bridge in August is 10.0 units higher than average.

## Seasonality Plot

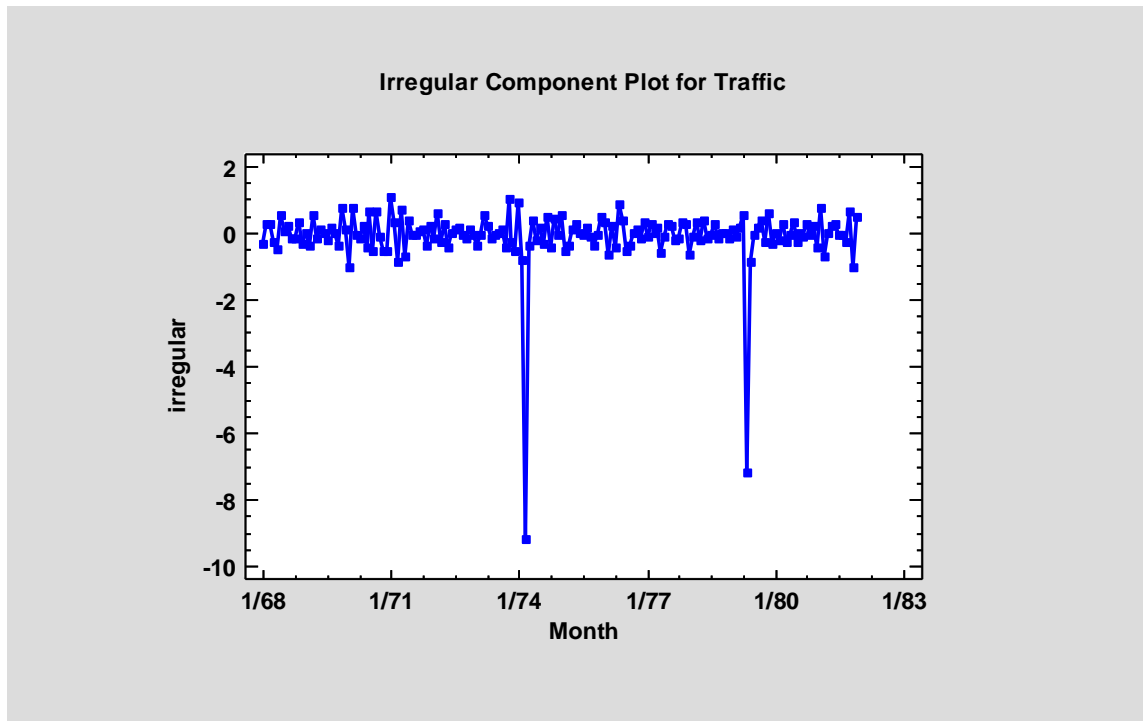
The *Seasonality Plot* displays the estimated indices:



Note the steady rise in traffic between January and August, followed by a decline from then until the end of the year.

## Irregular Component

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



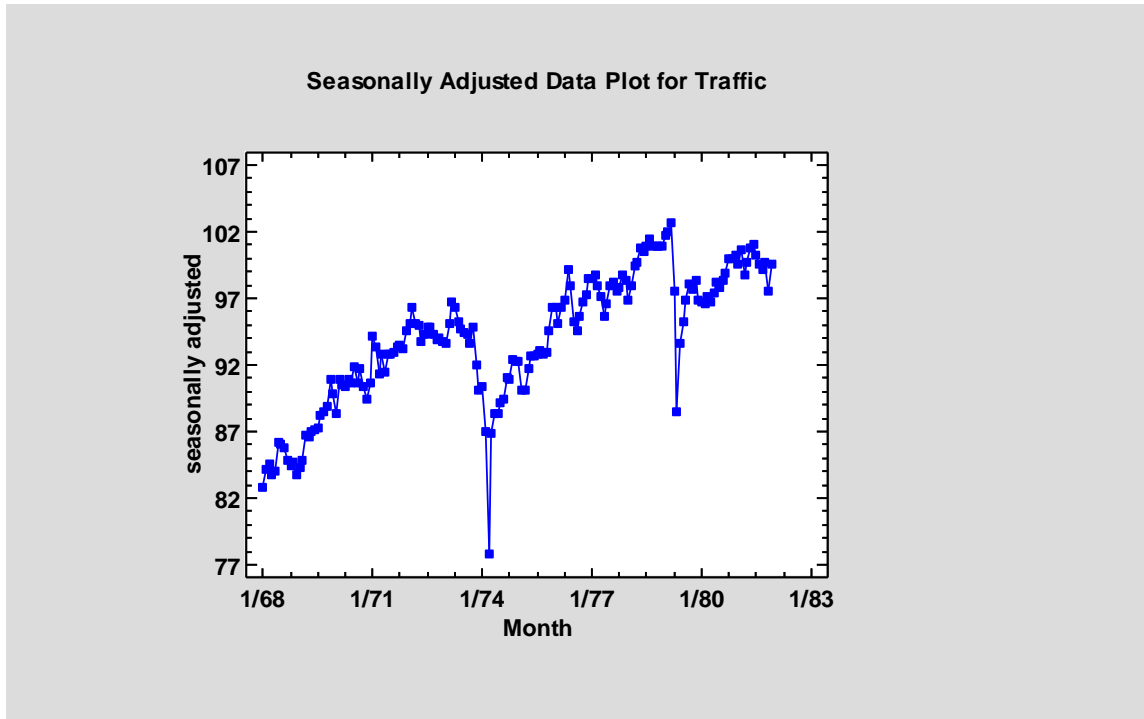
The irregular or residual component is that part of the time series that is unique to each time period, after removing the trend-cycle and seasonality.

In the sample data, note the large residual in March of 1974, equaling approximately -9.2. This indicates that traffic in that month was 9.2 units 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 only the seasonal effects, leaving both the trend-cycle and irregular components

The *Seasonally Adjusted Data* plot displays the results:

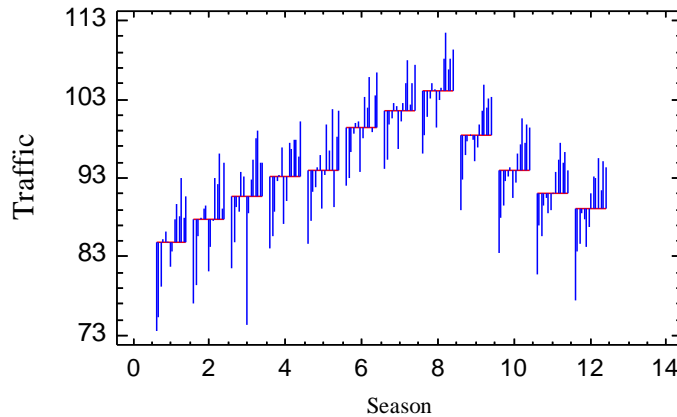


The effect of the gasoline shortages on the basic trend line is now quite apparent.

## Seasonal Subseries Plot

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

Seasonal Subseries Plot for Traffic



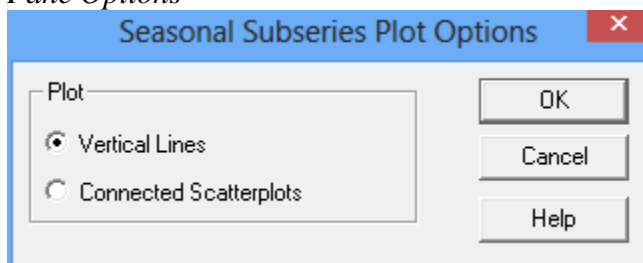
This plot is constructed as follows:

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

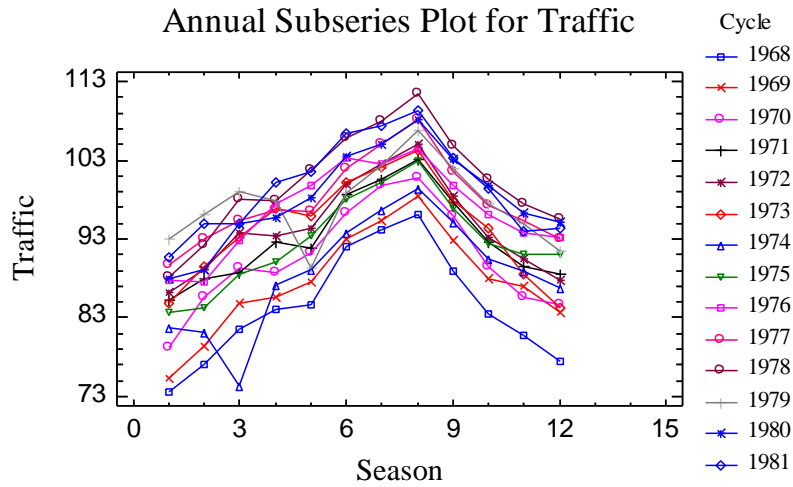


- **Vertical Lines** – draw a line from each observation to the average for its season.
- **Connected Scatterplots** – draw the data within each season as a connected X-Y plot.

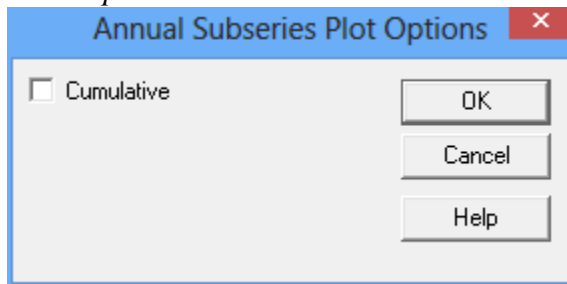


## Annual Subseries Plot

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

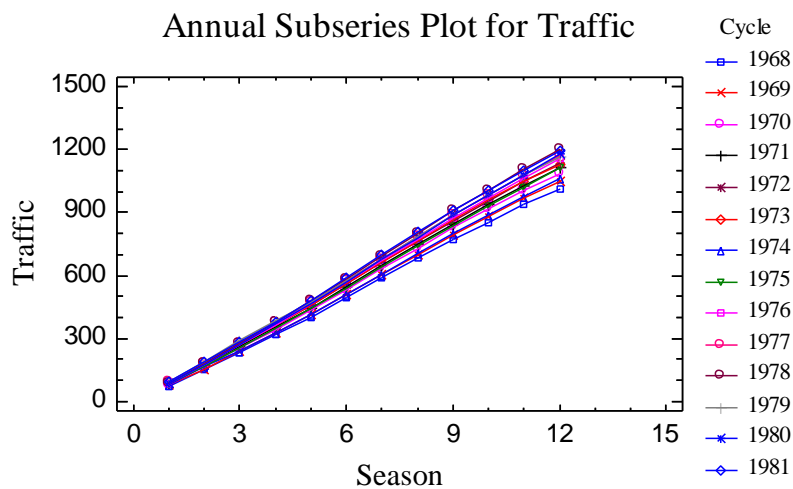


### Pane Options



- **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.

## References

X-13ARIMA-Seats Reference Manual, Version 1.1, Time Series Research Staff, Center for Statistical Research and Methodology, U.S. Census Bureau, 2016:  
<https://www.census.gov/ts/x13as/docX13AS.pdf>

Package ‘seasonal’, 2016: <https://cran.r-project.org/web/packages/seasonal/seasonal.pdf>

seasonal: R interface to X-13ARIMA-SEATS, Christoph Sax, 2016: <https://cran.r-project.org/web/packages/seasonal/vignettes/seas.pdf>