

# **STATGRAPHICS *Online***

## Statistical Analysis and Data Visualization System

Revised 6/21/2012

Copyright 2012 by StatPoint Technologies, Inc.

All rights reserved.

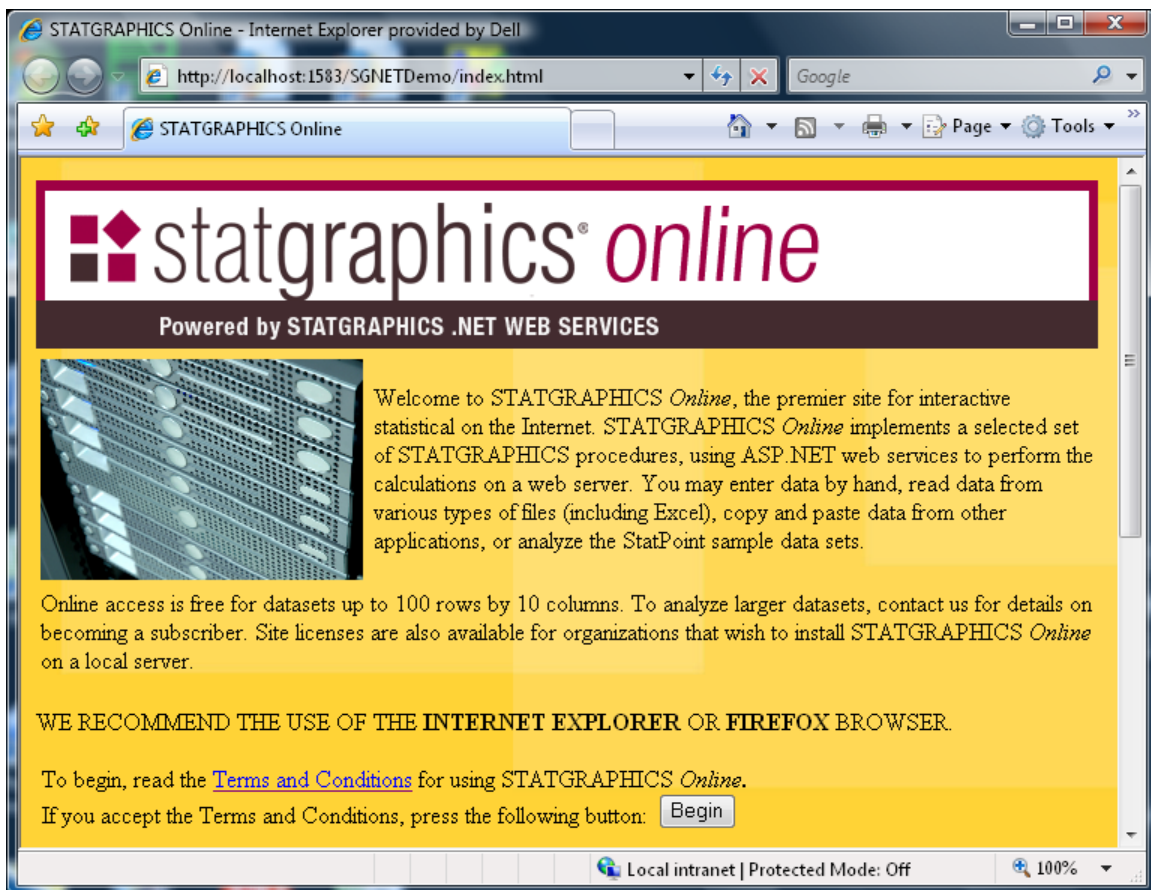
# Table of Contents

<b>Introduction</b> .....	<b>1</b>
<b>Chapter 1: Getting Started</b> .....	<b>2</b>
1.1 Creating a User Account.....	2
<b>Chapter 2: Entering and Manipulating Data</b> .....	<b>4</b>
2.1 Entering Data into the Data Editor.....	4
2.2 Opening a Sample Dataset .....	9
2.3 Opening a User File .....	9
2.4 Formatting Data .....	11
2.5 Managing Data Files .....	11
<b>Chapter 3: Selecting System Preferences</b> .....	<b>13</b>
3.1 General Settings .....	13
3.2 Internationalization .....	15
3.3 Fonts.....	16
3.4 Graphics .....	17
3.5 Fills .....	18
3.6 Lines.....	19
3.7 Points.....	20
<b>Chapter 4 – Running Statistical Procedures</b> .....	<b>21</b>
4.1 Data Input.....	21
4.2 Analysis Options .....	24
4.3 Tables and Graphs.....	25
4.4 Results to Save.....	28
4.5 Calculate .....	28
4.6 Preferences .....	30
4.7 Save Script .....	30
<b>Chapter 5 – Working with Scripts</b> .....	<b>31</b>
<b>Chapter 6 – STATLETS</b> .....	<b>33</b>
<b>Index</b> .....	<b>37</b>

## Introduction

STATGRAPHICS *Online* is a statistical package that allows you to perform a wide variety of statistical analyses from within your web browser. There is no need to download any software, since all calculations are done on a remote server. Using ASP.NET, data is submitted to a STATGRAPHICS XML web service running on a remote computer, which analyzes the data and returns both tabular and graphical results to the web browser. Data may reside anywhere, including your local computer.

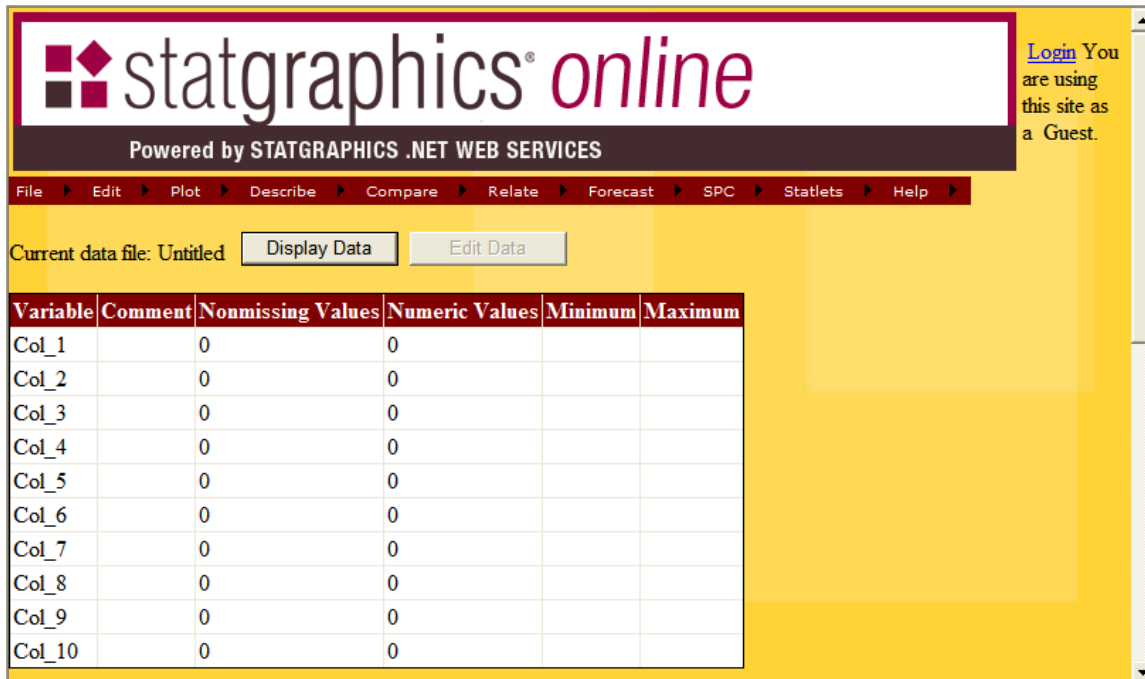
To use STATGRAPHICS *Online*, go to [www.statgraphicsonline.com](http://www.statgraphicsonline.com). The introductory page gives several hints about using the program:



Click on the *Terms and Conditions* link to read the conditions under which the software is provided. If you accept the terms and conditions, check the indicated box and press the *Begin* button to start analyzing data.

## Chapter 1: Getting Started

STATGRAPHICS *Online* can analyze data residing in various types of files, including Excel workbooks, XML files, or delimited text files. The main page, which is displayed when you start the program, shows the name of the current data file, together with a summary of the variables that it contains:



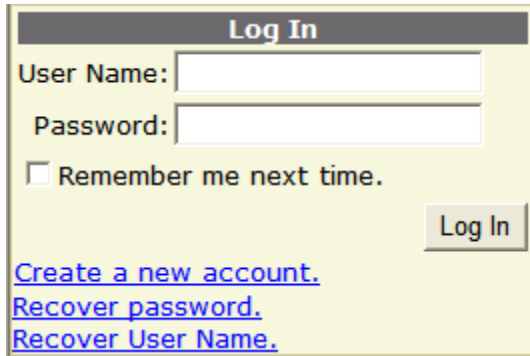
The screenshot shows the main interface of Statgraphics Online. At the top, there is a logo for "statgraphics® online" with the text "Powered by STATGRAPHICS .NET WEB SERVICES" below it. A navigation menu includes "File", "Edit", "Plot", "Describe", "Compare", "Relate", "Forecast", "SPC", "Statlets", and "Help". On the right side, there is a "Login" link and a message: "You are using this site as a Guest." Below the navigation menu, the current data file is identified as "Untitled", with "Display Data" and "Edit Data" buttons. A table summarizes the variables in the data file:

Variable	Comment	Nonmissing Values	Numeric Values	Minimum	Maximum
Col_1		0	0		
Col_2		0	0		
Col_3		0	0		
Col_4		0	0		
Col_5		0	0		
Col_6		0	0		
Col_7		0	0		
Col_8		0	0		
Col_9		0	0		
Col_10		0	0		

Initially, an empty data file named *Untitled* is created, containing 10 columns and 100 rows. If you wish to try out the statistical procedures using sample datasets provided by StatPoint Technologies, you can select any menu item from the main item. If you wish to analyze your own data, you must first login using the link in the upper right corner of the page.

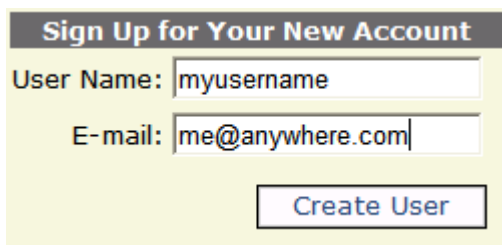
### 1.1 Creating a User Account

In order to analyze your own data, you must first create a user account. When you press the *Login* link, you will see the following input fields:



The image shows a 'Log In' form with a dark header bar containing the text 'Log In'. Below the header, there are two input fields: 'User Name:' and 'Password:'. A checkbox labeled 'Remember me next time.' is positioned below the password field. To the right of the checkbox is a 'Log In' button. At the bottom of the form, there are three blue underlined links: 'Create a new account.', 'Recover password.', and 'Recover User Name.'

If you have already created a user account, enter your user name and password. Otherwise, click on *Create a new account*:



The image shows a 'Sign Up for Your New Account' form with a dark header bar containing the text 'Sign Up for Your New Account'. Below the header, there are two input fields: 'User Name:' with the text 'myusername' and 'E-mail:' with the text 'me@anywhere.com'. Below the email field is a 'Create User' button.

User names contain 1 to 16 characters and are used to identify data files belonging to an individual user. Both your user name and e-mail address cannot already be in use. If you have forgotten your user name and/or password, use the links on the first login screen.

When you press the *Create User* button, a password for your account will be automatically generated and sent to the e-mail address that you enter. Once your password is received, return to the first *Login* screen and enter it. After you login, you will be returned to the main page. You can change your password by selecting *My Account* under *Edit* on the main menu.

## Chapter 2: Entering and Manipulating Data

To enter data into STATGRAPHICS Online, you have several choices:

1. Press the *Edit Data* button on the main page to load the data editor. This will display a datasheet of rows and columns into which you may type data. You can also use your system's clipboard to copy and paste data from other applications.
2. Select *File – Open Sample Dataset* to load one of StatPoint's sample datasets.
3. Select *File – Open User Data File – From Client* to load a data file residing on your computer.
4. Select *File – Open User Data File – From Server* to load a data file that you have previously saved on the server.
5. Leave the data file as *Untitled* and select a statistical procedure from the main menu. A sample dataset with data appropriate for the procedure you selected will then automatically be loaded.

### 2.1 Entering Data into the Data Editor

When you press the *Edit Data* menu on the main page, the *Data Editor* page shown below will be displayed:

statgraphics® online  
Powered by STATGRAPHICS .NET WEB SERVICES

File Save As... Rows Columns Cells Cancel

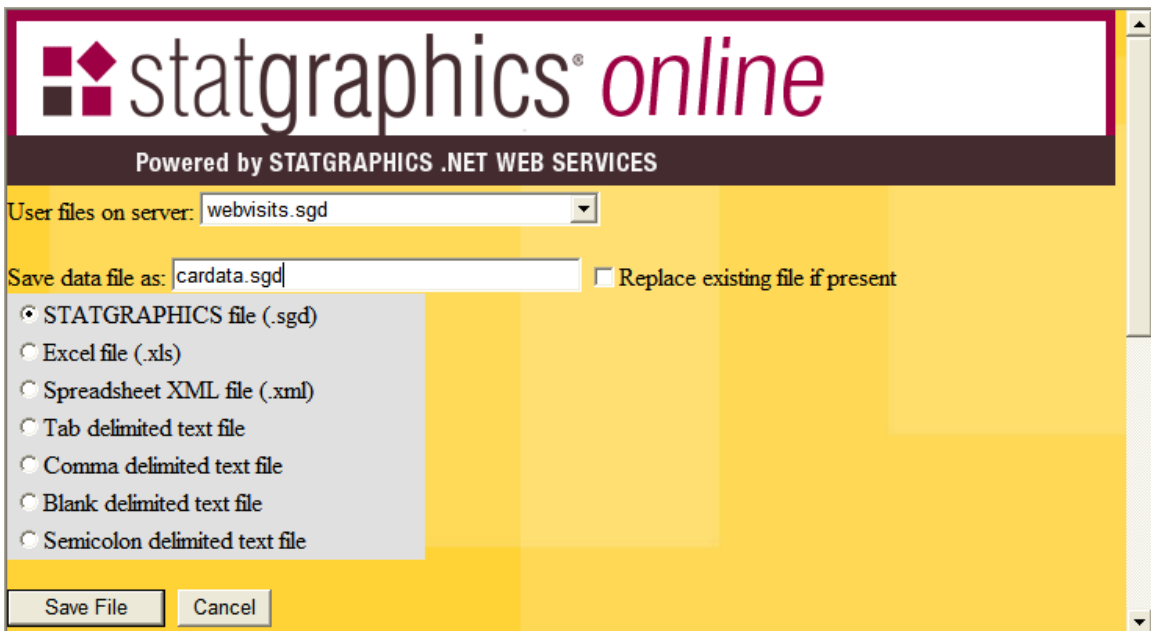
Name	A	B	C	D	E	F
Comment	Col_1	Col_2	Col_3	Col_4	Col_5	Col_6
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						

Input Results ✓ ↶ ✂ 📄 📁 ⏪ ⏩ 🖨

The Data Editor contains a data book with two sheets: one sheet for *Input* data that you wish to analyze, and a second sheet into which calculated *Results* will be placed. You can switch between the two sheets using the toolbar along the bottom of the page. To enter data, simply type it into the cells of the table. You should also assign meaningful names to the columns of the table. Column names are used to select data while in the statistical procedures and are displayed on tables and graphs. You can also use the row beneath to column names to indicate additional information about each column, such as the unit of measurement. A typical completed data sheet is shown below:

	A	B	C	D	E	F
Name	Make	Model	Type	Min Price	Mid Price	Max Price
Comment				price for basic version in \$1,000	average of min and max prices	price for premium version in \$1,000
1	Acura	Integra	Small	12.9	15.9	18.8
2	Acura	Legend	Midsize	29.2	33.9	38.7
3	Audi	90	Compact	25.9	29.1	32.3
4	Audi	100	Midsize	30.8	37.7	44.6
5	BMW	535i	Midsize	23.7	30	36.2
6	Buick	Century	Midsize	14.2	15.7	17.3
7	Buick	LeSabre	Large	19.9	20.8	21.7
8	Buick	Roadmaster	Large	22.6	23.7	24.9
9	Buick	Riviera	Midsize	26.3	26.3	26.3
10	Cadillac	DeVille	Large	33	34.7	36.3
11	Cadillac	Seville	Midsize	37.5	40.1	42.7

The menu across the top of the *Data Editor* page allows you to perform several important operations. To save data you have entered, select *Save File As*, which will display the following page:



You may save data in any of the indicated formats. If you plan to use the data only with STATGRAPHICS *Online*, we recommend that you save it as a STATGRAPHICS *.sgd* file. SGD files are XML files designed to contain information in a format that can be easily manipulated by the statistical procedures. A typical example is shown below:



```

<?xml version="1.0"?>
<statgraphics_data>
  <FORMAT>
    <DateOrder>MDY</DateOrder>
    <DateSeparator>/</DateSeparator>
    <DecimalSymbol>.</DecimalSymbol>
  </FORMAT>
  <COLUMNS>
    <C1 name="Make" type="C"/>
    <C2 name="Model" type="C"/>
    <C3 name="Type" type="C"/>
    <C4 name="Min Price" type="N" comment="price for basic version in $1,000"/>
    <C5 name="Mid Price" type="N" comment="average of min and max prices"/>
  </COLUMNS>
  <ROW>
    <C1>Acura</C1>
    <C2>Integra</C2>
    <C3>Small</C3>
    <C4>12.9</C4>
    <C5>15.9</C5>
  </ROW>
  <ROW>
    <C1>Acura</C1>
    <C2>Legend</C2>
    <C3>Midsize</C3>
    <C4>29.2</C4>
    <C5>33.9</C5>
  </ROW>
</statgraphics_data>

```

It contains information about how the data is formatted, column names and comments, and the raw data. You can also edit SGD files in any text editor, provided you maintain the same basic format.

Other operations that you can perform in the Data Editor include:

1. **Pasting columns of data from other applications** – First copy data from the other applications to the system clipboard. In the STATGRAPHICS data sheet, highlight the location at which you wish to paste the data. Then click on the Copy button in the lower toolbar.
2. **Paging within large datasets** – To reduce the amount of information that must be sent between your browser and the server, large datasets are displayed in blocks of 100 rows each. You can page between blocks using the arrows on the lower toolbar.

3. **Printing data** – Click on the printer icon on the lower toolbar to send a copy of the data to your printer.
4. **Cancel entries** – As you make changes to your data, they are changed within the browser but not sent to the server until:
  - a. You perform an operation such as selecting something from the menu that causes a new page to be posted.
  - b. You click on *Update* on the lower toolbar.
 To undo all input since the last post to the server, click on *Cancel* on the lower toolbar. NOTE: changes posted to the server are not permanent until you select *File Save* from the top menu.
5. **Delete rows, columns or cells** – Highlight the data to be deleted and select *Delete* from the top menu.
6. **Insert rows, columns or cells** – Highlight the data after which space is to be added and select *Insert* from the top menu.
7. **Sort data** – Highlight a single column and select *Rows – Sort* to sort data according to the values in the highlighted column.

After you enter and save the data, press *Cancel* to return to the main page. The name of the saved file and a summary of its contents will be displayed:

The screenshot shows the 'statgraphics online' web interface. At the top, there is a navigation menu with options: File, Edit, Plot, Describe, Compare, Relate, Forecast, SPC, Statlets, and Help. Below the menu, the current data file is identified as '93cars.sgd', with buttons for 'Display Data' and 'Edit Data'. A table displays the following data summary:

Variable	Comment	Nonmissing Values	Numeric Values	Minimum	Maximum
Make		93	0	Acura	Volvo
Model		93	7	90	900
Type		93	0	Compact	Midsize
Min Price	price for basic version in \$1,000	93	93	6.7	45.4
Mid Price	average of min and max prices	93	93	7.4	61.9
Max Price	price for premium version in \$1,000	93	93	7.9	80
MPG City	miles per gallon in city driving	93	93	15	46
MPG Highway	miles per gallon in highway driving	93	93	20	50
Air Bags	0=none, 1=driver only, 2=driver and passenger	93	93	0	2
Drive Train		93	0	all	front
Cylinders		92	92	3	8
Engine Size	liters	93	93	1	5.7
Horsepower	maximum	93	93	55	300

The summary shows:

1. *Nonmissing values* – the number of rows in the column for which an entry has been made. Empty cells will be treated as missing values by the statistical procedures.
2. *Numeric values* – the number of rows in the column that have valid numeric entries. When performing a statistical operation that requires numeric values, only these cells will be used in the analysis.
3. *Minimum and maximum* – the smallest and largest values in the column. For columns with no numeric values, this is based on an alphanumeric comparison.

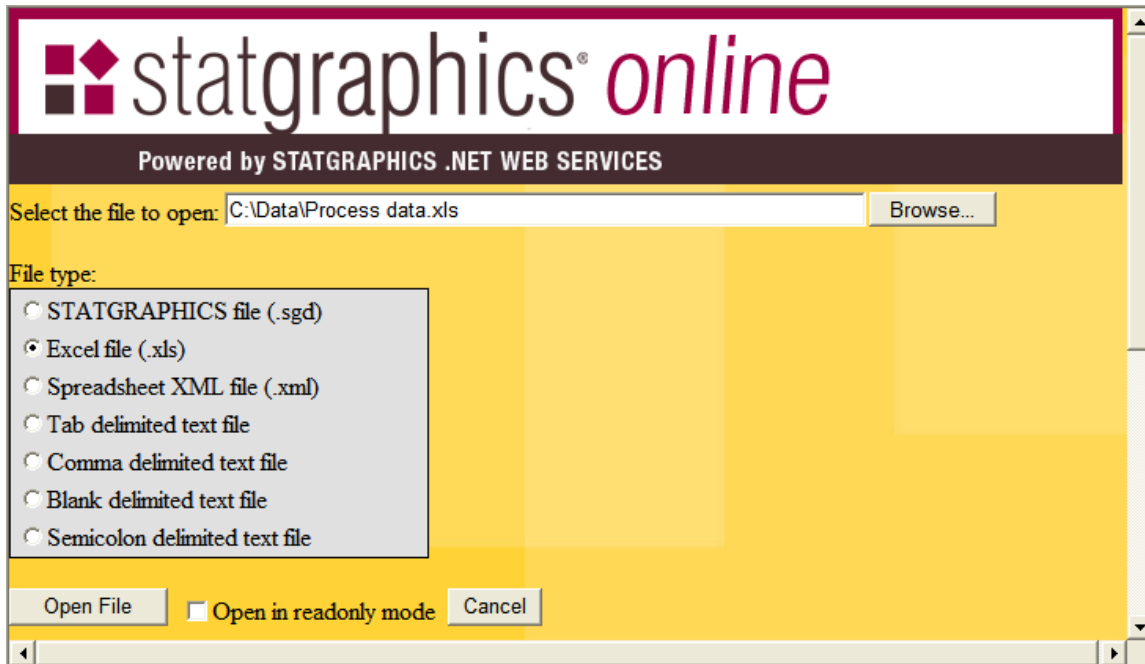
## 2.2 Opening a Sample Dataset

StatPoint Technologies has provided a collection of sample datasets that may be loaded by selecting *File – Open Sample Dataset*. Choose the file to be loaded and press *Open Dataset*.



## 2.3 Opening a User File

Users may also open data files from their client machine by selecting *File – Open User Data File – From Client*. This will display the page shown below:



Press the *Browse* button to search your computer for the file that you wish to load. Indicate the type of file it is and press *Open File*. This will:

1. Upload a copy of the file to the STATGRAPHICS *Online* server.
2. Make the selected data file the active file for access by the statistical procedures.

You will then be returned to the main page.

#### NOTES:

1. If you wish to restrict the program (and yourself) from making changes to the input file, select *Open in readonly mode*. You will still be able to display the data in the data editor, but you will not be able to make changes to it.
2. If you read an Excel file, only the data in the first sheet of the Excel workbook will be copied to the server and loaded. If you make changes to the file and resave it on the server, please remember that it contains only that one sheet.
3. The first two rows of Excel files and text files are assumed to contain column names and comments. If they do not, you may have to make adjustments to the data in the data editor by inserting additional blank rows at the top of the file.

## 2.4 Formatting Data

When you read data from external files (non-STATGRAPHICS files), the data is assumed to follow the formatting conventions of your current locale. This is also true when you type data into the Data Editor. The settings that affect how data are read are:

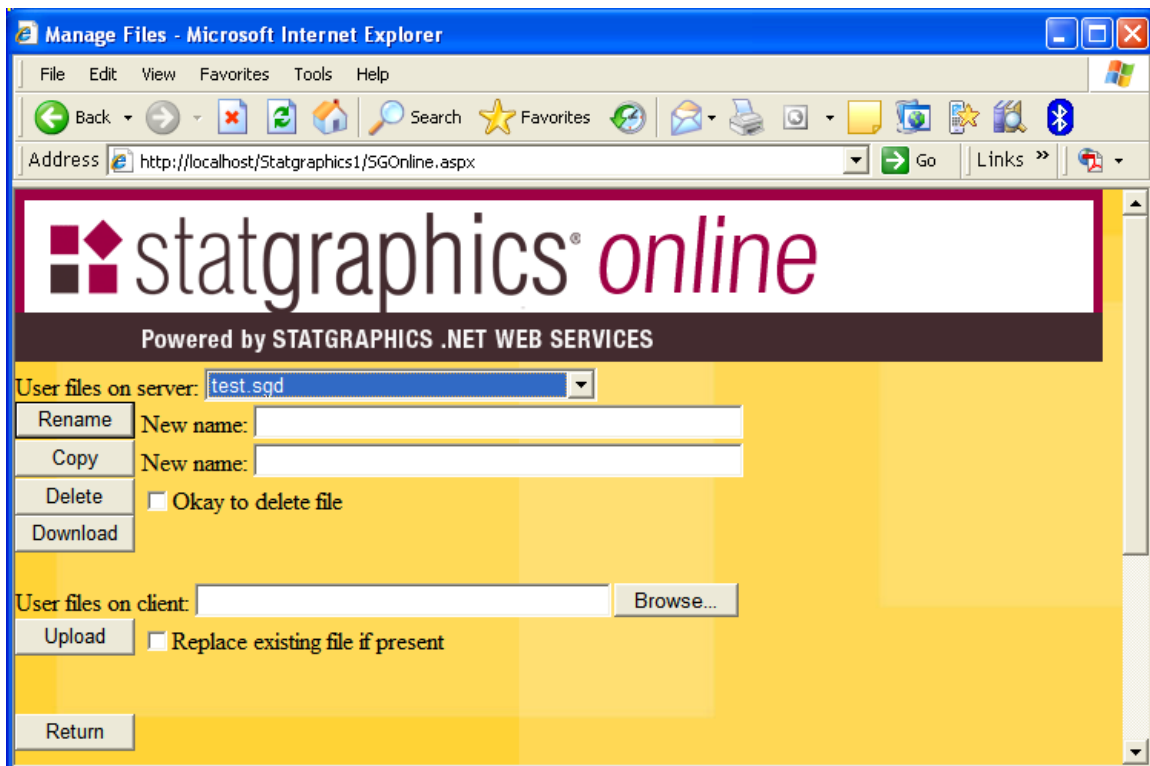
1. *Decimal separator* – the symbol placed to the left of the first decimal digit (either a period or a comma).
2. *Date delimiter* – usually a slash (/), period (.), or dash (-).
3. *Date order* – usually Month-Day-Year or Day-Month-Year.

You should be sure that the data in your files follows the conventions of your locale.

NOTE: you can specify a different format for the statistical output if desired by changing the setting under *Edit – Preferences* on the main menu.

## 2.5 Managing Data Files

The *File* menu contains a selection titled *Manage Data Files* that allows you to perform common operations on user data files that have been saved on the server. It displays the page shown below:



The buttons perform the following operations:

**Rename** – To rename a file, select the file you wish to rename from the drop down list. Then enter a new name for the file. NOTE: since this option does not change the contents of the file, the file extension should usually not be changed.

**Copy** – To create a copy of a file, select the file you wish to copy from the drop down list. Then enter a name for the new file. NOTE: since this option does not change the contents of the file, the file extension of the new file should usually be the same as the original file.

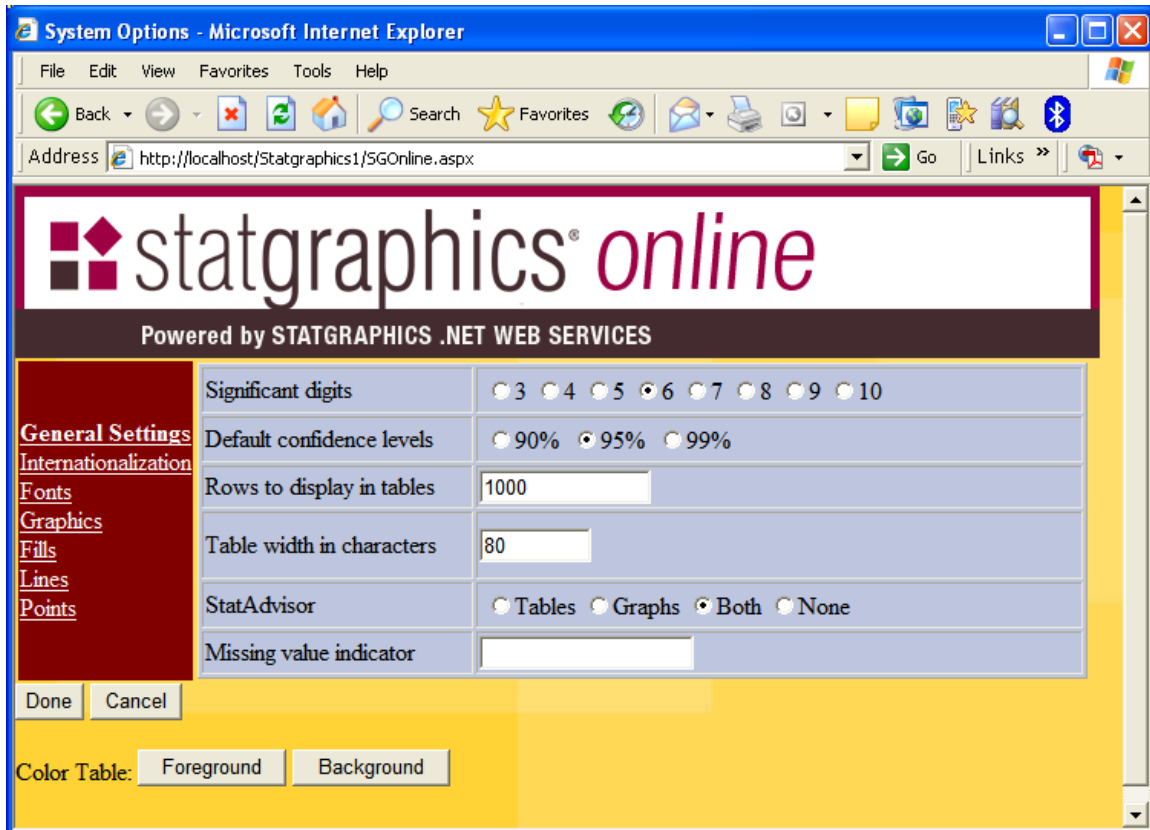
**Delete** – To delete a file, select the file you wish to copy from the drop down list. Check the *Okay to delete file* box and press *Delete*.

**Download** – This option is designed to allow you to download a file from the server to your client computer. Select the file you wish to copy from the drop down list. When you press download, a link will be added to the page specifying the URL necessary to access that file. You can then use your browser's capabilities to download the file to your computer.

**Upload** - This option is designed to allow you to upload a file from your client computer to the server. Use the *Browse* button to locate the file you wish to upload and then press *Upload*.

## Chapter 3: Selecting System Preferences

The *Edit* menu contains a selection titled *Set Preferences* that allows you to change the default settings used by STATGRAPHICS *Online*. It consists of multiple pages, each similar to the page shown below:



### 3.1 General Settings

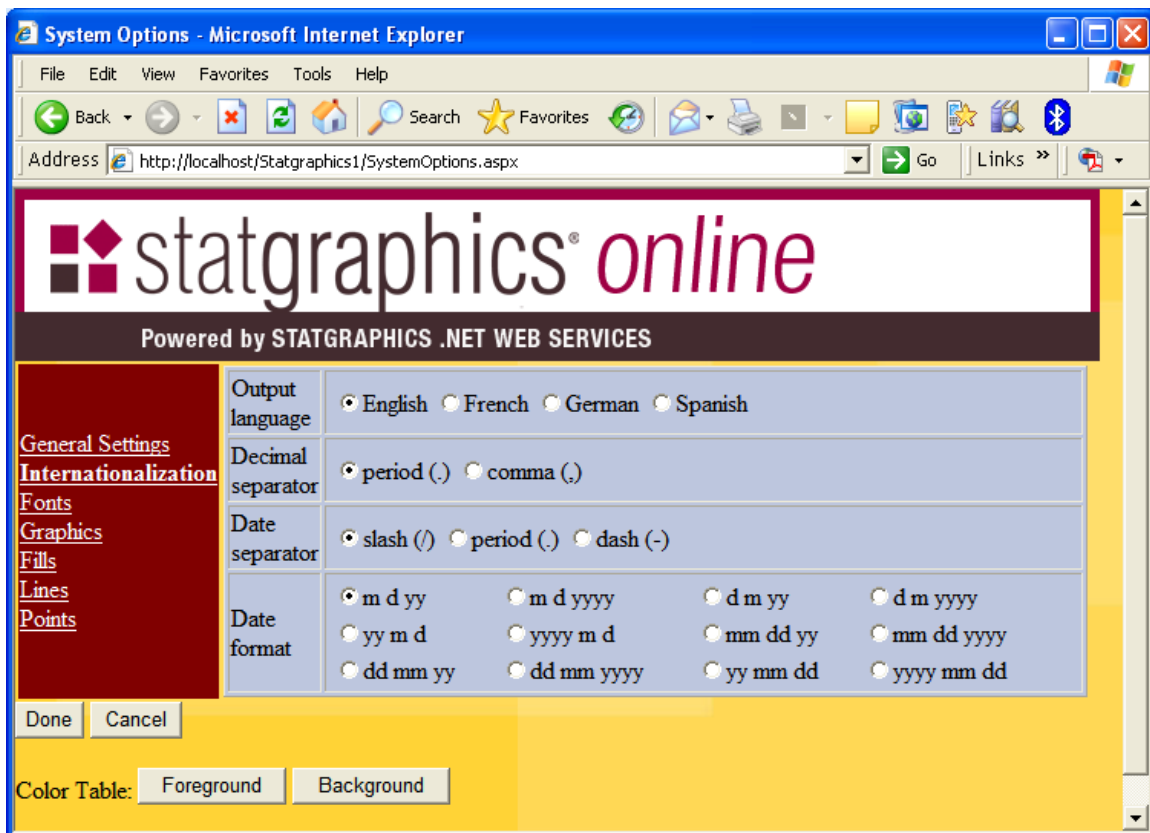
The *General Settings* page allows you to specify the following settings:

- *Significant digits* – the number of significant digits shown when displaying numeric values.
- *Default confidence level* – the confidence level used by default when creating interval estimates and hypothesis tests, and that used by the StatAdvisor when interpreting statistical results.
- *Rows to display in tables* – the maximum number of rows displayed when a table is created. This prevents tables from becoming excessively large when analyzing large datasets.

- *Table width in characters* – the maximum number of characters displayed across the page when outputting a table. Tables exceeding this width will be displayed in multiple segments.
- *StatAdvisor* – controls whether StatAdvisor information is displayed after tables and graphs. The StatAdvisor interprets the statistical output.
- *Missing value indicator* – a special indicator used in user data files to indicate missing data, if any. Empty cells are always interpreted as containing missing values.

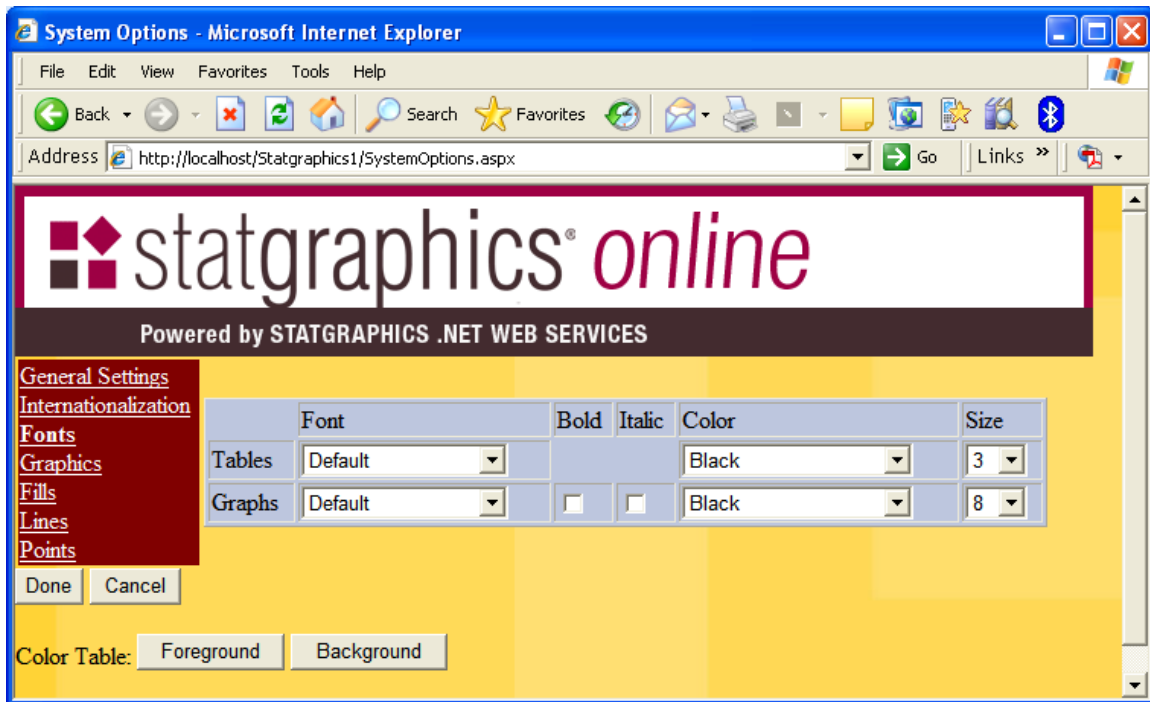


### 3.2 Internationalization



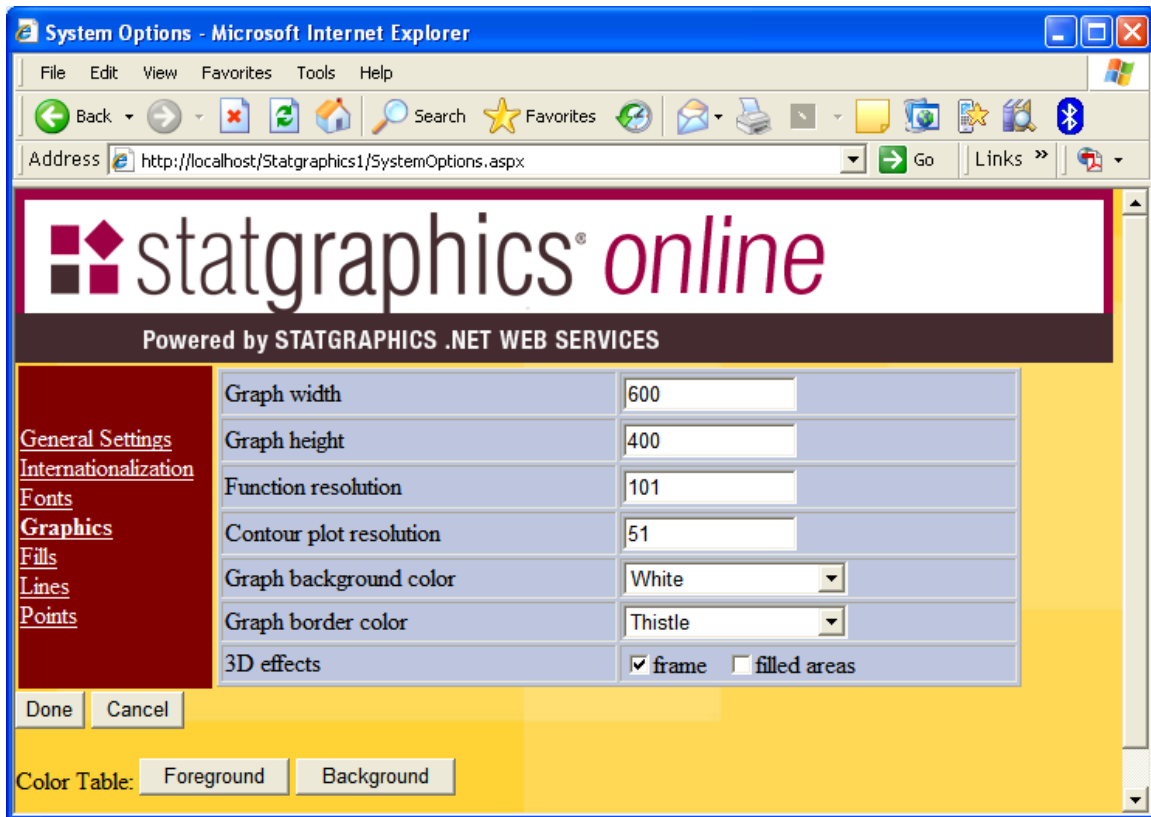
- *Output language* – the language used in output tables and graphs. Currently, input pages are only available in English.
- *Decimal separator* – the decimal separator to be used in displaying *output*. All numeric values on *input* pages and in user data files are assumed to follow the conventions of the current locale setting for your computer.
- *Date separator* - the date separator to be used in displaying *output*. All date and date-time values on *input* pages and in user data files are assumed to follow the conventions of the current locale setting for your computer.
- *Date format* - the format for months, days and years to be used in displaying *output*. A double letter such as *mm* requests that numbers below 10 be written with a leading zero. All date and date-time values on *input* pages and in user data files are assumed to follow the conventions of the current locale setting for your computer.

### 3.3 Fonts



This page controls the fonts used when outputting tables and graphs. Press the *Foreground* and *Background* buttons to display the available colors that may be selected in the dropdown lists.

### 3.4 Graphics



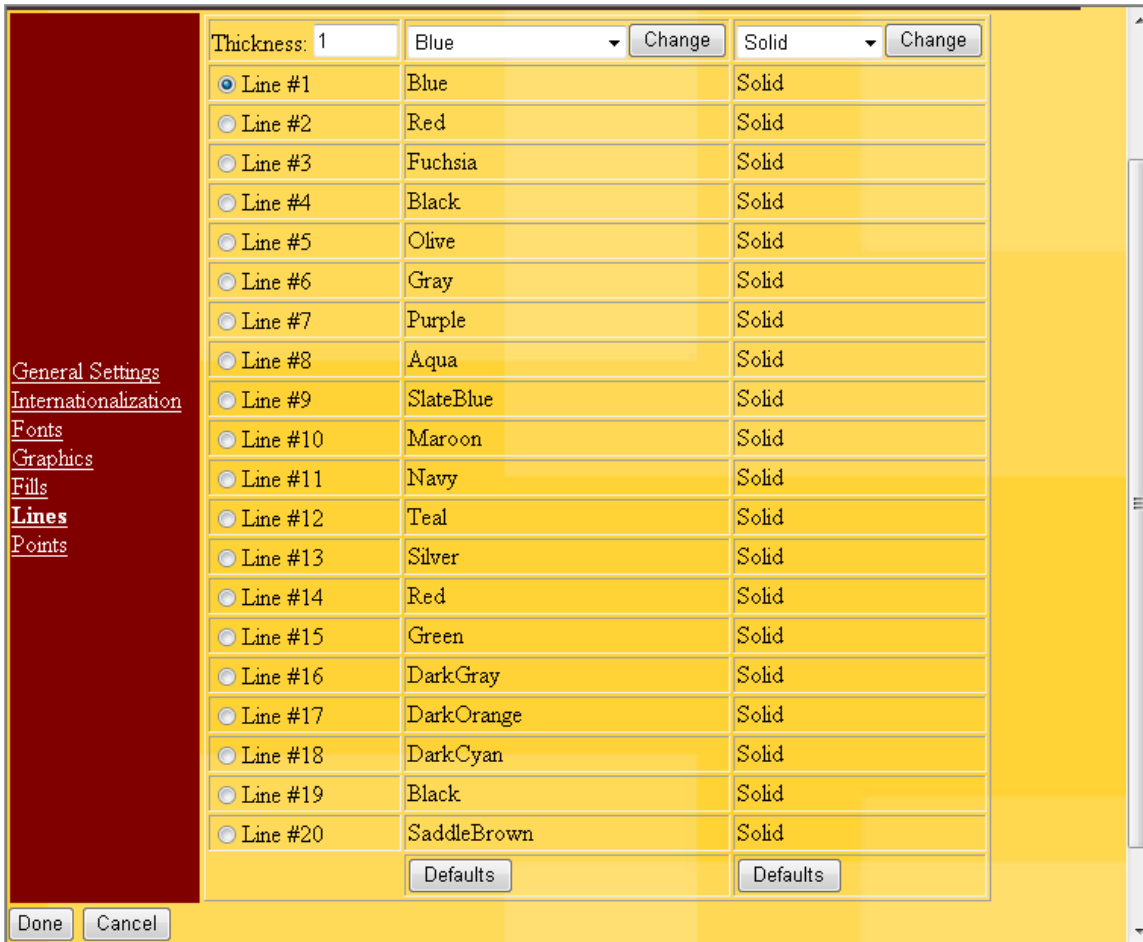
- *Graph width and height* – the size of graphs in pixels.
- *Function resolution* – the number of X values at which functions will be evaluated when drawn on a graph. Higher resolution results in smoother curves.
- *Contour plot resolution* – the number of X and Y values at which a function will be evaluated when contour plots are created. Higher resolution results in smoother divisions between the contour levels.
- *Graph background color* – the color displayed within the axes of a graph.
- *Graph border color* – the color displayed outside the axes of a graph.
- *3D effects* – whether 3D effects are used around the frame of a graph and when displaying filled areas such as bars and pie slices.

### 3.5 Fills



Graphs may contain up to 20 different types of filled areas. You may change the color and/or type of fill used in each area.

### 3.6 Lines



Graphs may contain up to 20 different line types. You may change the color and/or type of each line.

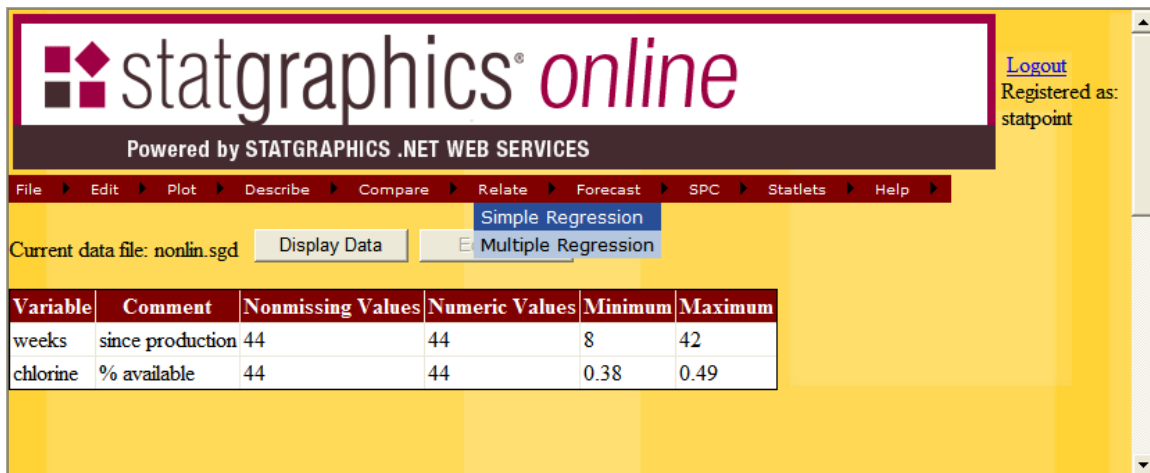
### 3.7 Points



Graphs may contain up to 20 different point types. You may change the size, color and/or type of each point set.

## Chapter 4 – Running Statistical Procedures

The main menu provides access to a large collection of statistical procedures. After opening the file containing the data you wish to analyze, use the menu to select the desired procedure. For example, suppose you wished to analyze the data contained in the sample dataset titled “Shelf Life Study”. This dataset contains 44 samples of different ages. For each sample, the number of weeks since it was produced is recorded, as is the concentration of chlorine. Select *File – Open Sample Dataset* to open the desired file. Then select *Relate – Simple Regression* from the main menu, as shown below:

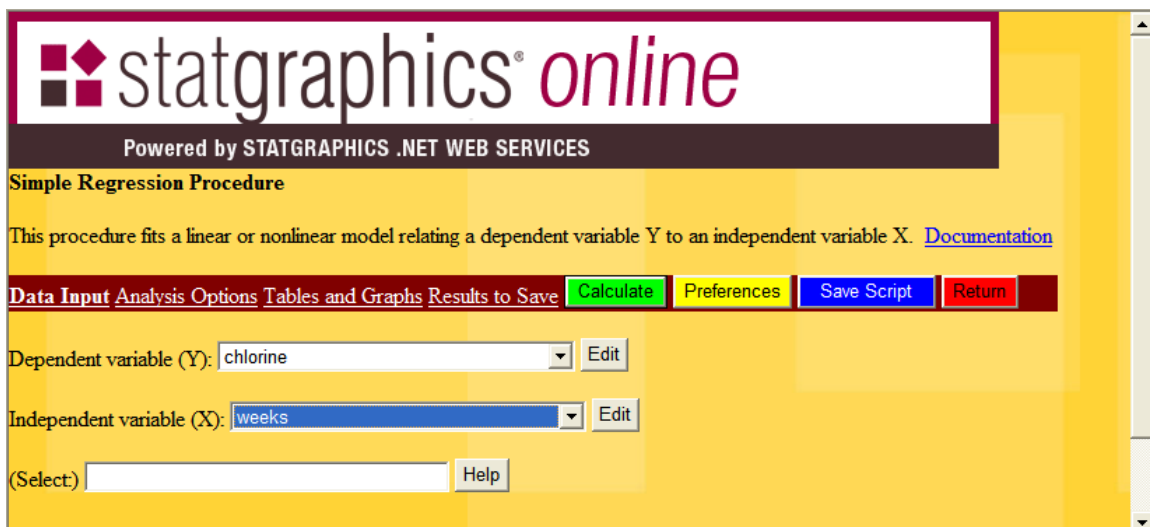


The screenshot shows the main menu of the statgraphics online application. The logo "statgraphics online" is at the top left, with "Powered by STATGRAPHICS .NET WEB SERVICES" below it. A navigation menu includes File, Edit, Plot, Describe, Compare, Relate, Forecast, SPC, Statlets, and Help. The "Relate" menu is open, showing "Simple Regression" selected and "Multiple Regression" as an alternative. Below the menu, the current data file is "nonlin.sgd" and there is a "Display Data" button. A table displays the data summary:

Variable	Comment	Nonmissing Values	Numeric Values	Minimum	Maximum
weeks	since production	44	44	8	42
chlorine	% available	44	44	0.38	0.49

### 4.1 Data Input

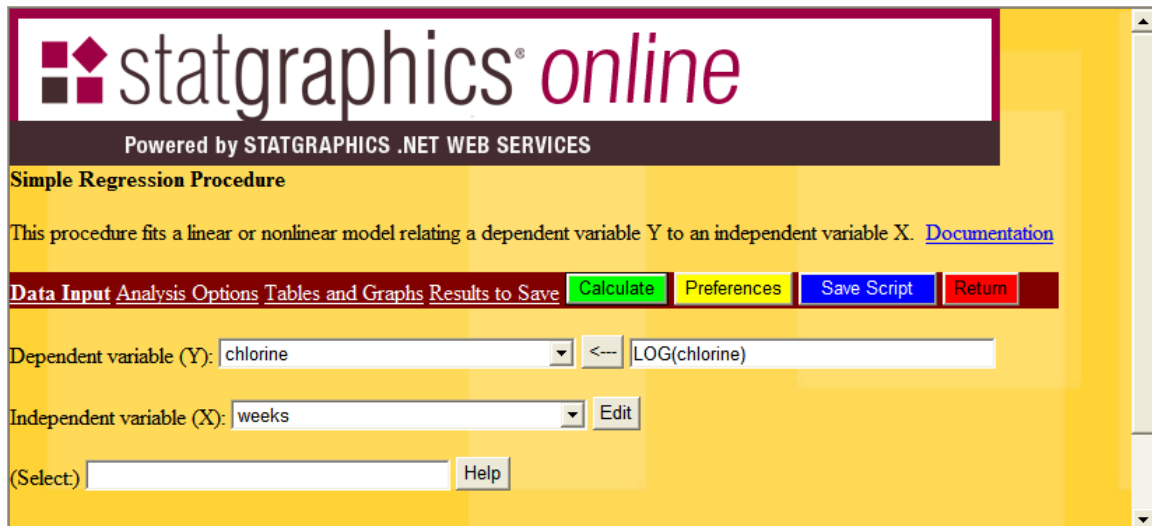
Whenever you select a statistical procedure from the main menu, a page will be displayed on which to specify the columns containing the data that you wish to analyze. The initial page for the *Simple Regression* procedure is shown below:



The screenshot shows the configuration page for the Simple Regression procedure. The title is "Simple Regression Procedure". Below the title, it states: "This procedure fits a linear or nonlinear model relating a dependent variable Y to an independent variable X. [Documentation](#)". A navigation bar includes "Data Input", "Analysis Options", "Tables and Graphs", "Results to Save", "Calculate", "Preferences", "Save Script", and "Return". The "Data Input" section has two dropdown menus: "Dependent variable (Y):" set to "chlorine" and "Independent variable (X):" set to "weeks". Both have "Edit" buttons. There is also a "(Select)" dropdown menu with a "Help" button.

You can specify the name of a variable in either of two ways:

1. Press the arrow to dropdown the list of variables and select the desired column.
2. Press the *Edit* button to the right of the column to display an edit field to the right of the button, as shown below:



statgraphics® online  
Powered by STATGRAPHICS .NET WEB SERVICES

**Simple Regression Procedure**

This procedure fits a linear or nonlinear model relating a dependent variable Y to an independent variable X. [Documentation](#)

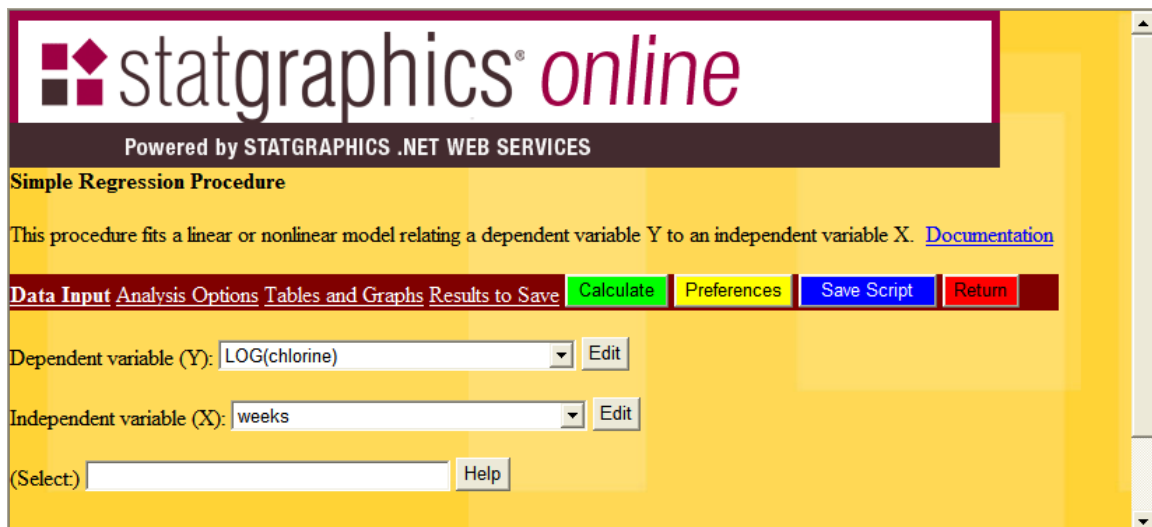
**Data Input** Analysis Options Tables and Graphs Results to Save Calculate Preferences Save Script Return

Dependent variable (Y): chlorine <-- LOG(chlorine)

Independent variable (X): weeks Edit

(Select) Help

You can type any valid STATGRAPHICS expression into the edit field, such as the expression above which will calculate the natural logarithms of *chlorine*. Then press the arrow button to put the edit field contents into the dropdown list:



statgraphics® online  
Powered by STATGRAPHICS .NET WEB SERVICES

**Simple Regression Procedure**

This procedure fits a linear or nonlinear model relating a dependent variable Y to an independent variable X. [Documentation](#)

**Data Input** Analysis Options Tables and Graphs Results to Save Calculate Preferences Save Script Return

Dependent variable (Y): LOG(chlorine) Edit

Independent variable (X): weeks Edit

(Select) Help

STATGRAPHICS expressions are described in detail in the document [www.statpoint.com/STATGRAPHICS Operators.pdf](http://www.statpoint.com/STATGRAPHICS_Operators.pdf).

Each *Data Input* page also contains a *Select* field, which can be used to select a subset of the rows in the file. Typical entries include:



<i>Entry</i>	<i>Use</i>	<i>Example</i>
FIRST (k)	Selects the first <i>k</i> rows.	FIRST(25)
LAST (k)	Selects the last <i>k</i> rows.	LAST(25)
ROWS (start,end)	Selects rows between <i>start</i> and <i>end</i> , inclusive.	ROWS(21,30)
RANDOM (k)	Selects a random set of <i>k</i> rows.	RANDOM(25)
EXCLUDE(k)	Selects all rows except row <i>k</i> .	EXCLUDE(25)
column = value	Selects only rows for which <i>column</i> equals <i>value</i> .	weeks = 30
column <> value	Selects only rows for which <i>column</i> does not equal <i>value</i> .	weeks <> 30
column > value	Selects only rows for which <i>column</i> is greater than <i>value</i> .	weeks > 30
column < value	Selects only rows for which <i>column</i> is less than <i>value</i> .	weeks < 30
column >= value	Selects only rows for which <i>column</i> is greater than or equal to <i>value</i> .	weeks >= 30
column <= value	Selects only rows for which <i>column</i> is less than or equal to <i>value</i> .	weeks <= 30
condition1 & condition2	Selects only rows that meet both conditions.	weeks >= 20 & weeks <= 40
condition1   condition2	Selects only rows that meet at least one of the conditions.	weeks = 30   weeks = 40

If you push the *Help* button next to the *Select* field, it will display a list showing typical structures:

(Select)

- All rows
- First \_\_\_ rows
- Last \_\_\_ rows
- Rows between \_\_\_ and \_\_\_
- Random \_\_\_ rows
- All rows except \_\_\_
- Rows for which \_\_\_ equals \_\_\_
- Rows for which \_\_\_ does not equal \_\_\_
- Rows for which \_\_\_ > \_\_\_
- Rows for which \_\_\_ < \_\_\_
- Rows for which \_\_\_ >= \_\_\_
- Rows for which \_\_\_ <= \_\_\_

If you select a radio button and then press *Help* again, an entry will be placed in the *Select* field as shown below:

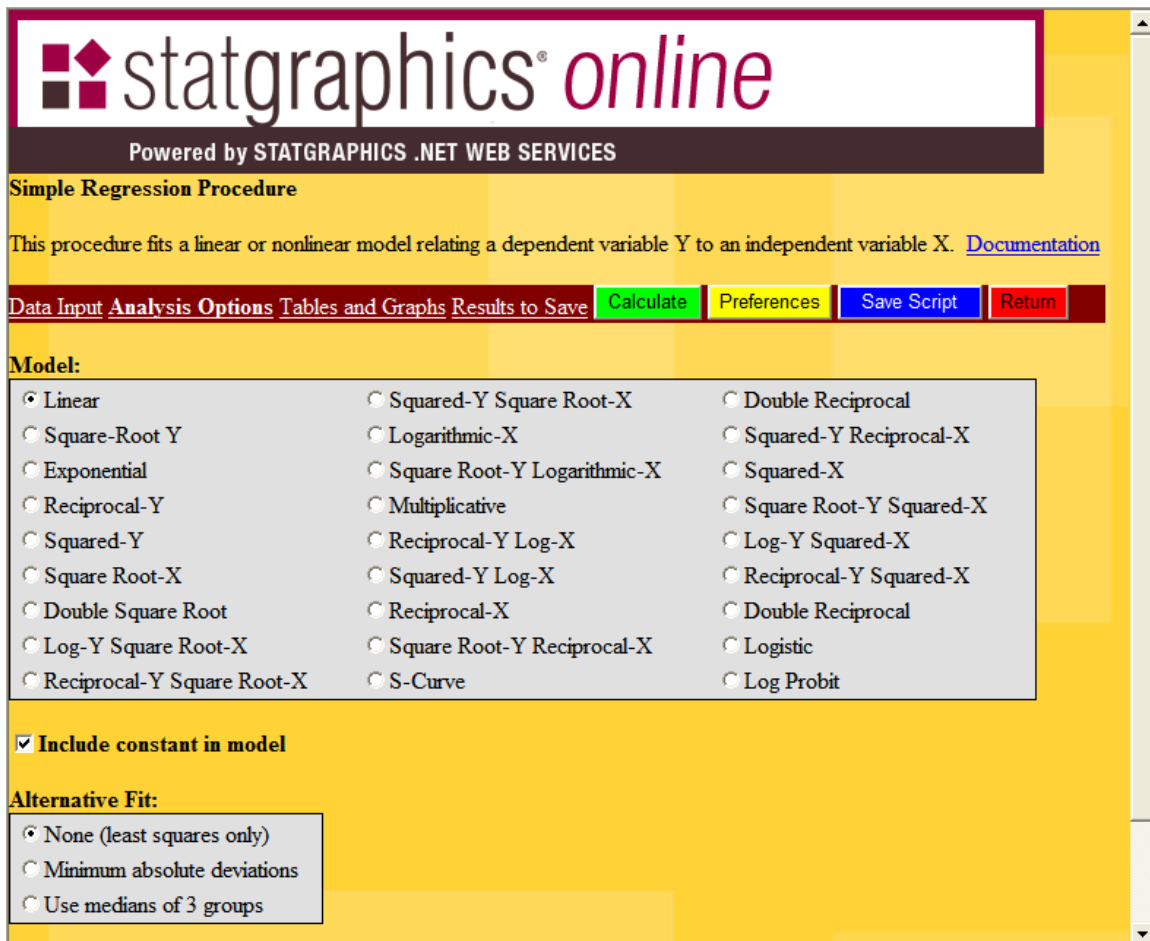


(Select) ROWS(???,???) Help

Replace any ???'s with your desired entries.

## 4.2 Analysis Options

After specifying the columns containing the data to be analyzed, click on *Analysis Options* to display a second page containing overall options for the selected statistical procedure:



statgraphics<sup>®</sup> online  
Powered by STATGRAPHICS .NET WEB SERVICES

### Simple Regression Procedure

This procedure fits a linear or nonlinear model relating a dependent variable Y to an independent variable X. [Documentation](#)

Data Input **Analysis Options** Tables and Graphs Results to Save Calculate Preferences Save Script Return

**Model:**

<input checked="" type="radio"/> Linear	<input type="radio"/> Squared-Y Square Root-X	<input type="radio"/> Double Reciprocal
<input type="radio"/> Square-Root Y	<input type="radio"/> Logarithmic-X	<input type="radio"/> Squared-Y Reciprocal-X
<input type="radio"/> Exponential	<input type="radio"/> Square Root-Y Logarithmic-X	<input type="radio"/> Squared-X
<input type="radio"/> Reciprocal-Y	<input type="radio"/> Multiplicative	<input type="radio"/> Square Root-Y Squared-X
<input type="radio"/> Squared-Y	<input type="radio"/> Reciprocal-Y Log-X	<input type="radio"/> Log-Y Squared-X
<input type="radio"/> Square Root-X	<input type="radio"/> Squared-Y Log-X	<input type="radio"/> Reciprocal-Y Squared-X
<input type="radio"/> Double Square Root	<input type="radio"/> Reciprocal-X	<input type="radio"/> Double Reciprocal
<input type="radio"/> Log-Y Square Root-X	<input type="radio"/> Square Root-Y Reciprocal-X	<input type="radio"/> Logistic
<input type="radio"/> Reciprocal-Y Square Root-X	<input type="radio"/> S-Curve	<input type="radio"/> Log Probit

Include constant in model

**Alternative Fit:**

<input checked="" type="radio"/> None (least squares only)
<input type="radio"/> Minimum absolute deviations
<input type="radio"/> Use medians of 3 groups

You can stick with the default settings or make any desired changes.

## 4.3 Tables and Graphs

The third page lists all of the tables and graphs that can be generated by the selected statistical procedure, together with any options:

The screenshot shows the Statgraphics Online interface for the Simple Regression Procedure. The header includes the Statgraphics Online logo and the text "Powered by STATGRAPHICS .NET WEB SERVICES". Below the header, the procedure name "Simple Regression Procedure" is displayed, followed by a brief description: "This procedure fits a linear or nonlinear model relating a dependent variable Y to an independent variable X." A navigation bar contains buttons for "Data Input", "Analysis Options", "Tables and Graphs" (which is highlighted), "Results to Save", "Calculate", "Preferences", "Save Script", and "Return".

The main configuration area is divided into two columns: "Tables/Graphs" and "Options".

**Tables/Graphs:**

- 
- 
- Analysis Summary
- Plot of Fitted Model
  -
- Forecasts

**Options:**

Plot:  
 All models  Least squares fit only  Alternative fit only

X-axis resolution:

Include:  
 Prediction limits  Confidence limits

Confidence level:  %

Two-sided interval  Lower bound  Upper bound

Forecast at X:  

<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
----------------------	----------------------	----------------------	----------------------

Press *All* to select all of the available output, or check only the tables and graphs that you want.

For each graph, there is a button labeled *Titles and Scaling* that allows you to override the default selections. These buttons display the page shown below:

Graphics Options for  (leave blank for defaults)

Item	Text	Scale From	Scale To	Scale By	Skip Increment	Options
Title	<input type="text"/>					
Subtitle	<input type="text"/>					
X-axis	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text" value="0"/>	<input type="checkbox"/> Log scale <input type="checkbox"/> No power <input type="checkbox"/> Skip repeats <input type="checkbox"/> Rotate ticmarks
Y-axis	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text" value="0"/>	<input type="checkbox"/> Log scale <input type="checkbox"/> No power <input type="checkbox"/> Skip repeats
Z-axis	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text" value="0"/>	<input type="checkbox"/> Log scale <input type="checkbox"/> No power <input type="checkbox"/> Skip repeats
Right-axis	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text" value="0"/>	<input type="checkbox"/> Log scale <input type="checkbox"/> No power <input type="checkbox"/> Skip repeats
Legend title	<input type="text"/>	X offset: <input type="text" value="1.1"/>	Y offset: <input type="text" value="1.0"/>			

**Grid direction:**

None  Horizontal  Vertical  Bidirectional

**Margins:**

Top:	Bottom:	Left:	Right:
<input type="text" value="22.0"/> %	<input type="text" value="22.0"/> %	<input type="text" value="22.0"/> %	<input type="text" value="22.0"/> %

**3D Viewpoint:**

Horizontal angle:	Vertical angle:
<input type="text" value="20"/>	<input type="text" value="20"/>

To override the title or scaling for an axis, enter the desired information. Otherwise, default settings will be applied. Some of the special options available on this page are:

- *Log scale* – Check this box to draw an axis using a logarithmic scale. Note that graphs on which logarithmic scaling does not make sense will ignore this setting.
- *No power* – Check this box to suppress scaling axes using a notation such as (*X 1000*) for very large or very small numbers.
- *Skip repeats* – Check this box to indicate that consecutive axis labels that are identical should not be drawn. This applies only to non-numeric axes.
- *X and Y offset* – This indicates the desired position of the legend block (if any) with respect to the axes. For offset purposes, the axes are considered to extend from (0, 0) at the bottom left corner to (1, 1) at the upper right corner. A value such as 1.1 for the X offset indicates that the legend block should be offset by 10% to the right of the axes. If a legend block is partially offscreen, use this setting and the page margins to adjust its position.
- *Margins* – This indicates the offset of each axis from the edge of the graphics drawing area.
- *3D viewpoint* – These angles specify the location of the viewer with respect to the origin of a 3D graph. Increasing the angles moves the viewer further to the right and further above the graph.

## 4.4 Results to Save

The fourth page indicates which results, if any, should be saved for later access:

Result to Save	Column name
<input checked="" type="checkbox"/> Predicted values	PREDICTED
<input checked="" type="checkbox"/> Lower limits for predictions	LOWERPLIMS
<input checked="" type="checkbox"/> Upper limits for predictions	UPPERPLIMS
<input type="checkbox"/> Lower limits for forecast means	LOWERCLIMS
<input type="checkbox"/> Upper limits for forecast means	UPPERCLIMS
<input type="checkbox"/> Residuals	RESIDUALS
<input type="checkbox"/> Studentized residuals	SRESIDUALS
<input type="checkbox"/> Leverages	LEVERAGES
<input type="checkbox"/> Predictions from resistant model	ALTPREDICTED
<input type="checkbox"/> Residuals from resistant model	ALTRESIDUALS
<input type="checkbox"/> Model statistics	MODELSTATS

When results are saved, they are placed in an XML file and a link to that file is added to the bottom of the output page. In addition, the results are loaded into the second page of the *Data Editor*.

## 4.5 Calculate

Once all desired options have been selected, press the *Calculate* button to begin the statistical calculations. An XML script is automatically created and sent to the server. The server reads the indicated data, performs the calculations, and sends back an HTML page to the web browser for display. The HTML page contains tabular output and imbedded images, as in the example below:

**Simple Regression Procedure**

This procedure fits a linear or nonlinear model relating a dependent variable Y to an independent variable X. [Documentation](#)

Data Input Analysis Options Tables and Graphs Results to Save Calculate Preferences Save Script Return

STATGRAPHICS Web Services  
Generated: 11/7/2007 5:02:05 PM

Data file: nonlin.sgd

**Simple Regression - chlorine vs. weeks**

Dependent variable: chlorine (% available)  
Independent variable: weeks (since production)  
Linear model:  $Y = a + b \cdot X$

**Coefficients**

	Least Squares Estimate	Standard Error	T	P-Value
Intercept	0.48551	0.00589066	82.4204	0.0000
Slope	-0.00271679	0.000243115	-11.1749	0.0000

**Analysis of Variance**

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	0.0295587	1	0.0295587	124.88	0.0000
Residual	0.00994133	42	0.000236698		
Total (Corr.)	0.0395	43			

Correlation Coefficient = -0.865055  
R-squared = 74.8321 percent  
R-squared (adjusted for d.f.) = 74.2328 percent  
Standard Error of Est. = 0.015385  
Mean absolute error = 0.012834  
Durbin-Watson statistic = 0.992081 (P= 0.0001 )  
Lag 1 residual autocorrelation = 0.451981

**The StatAdvisor**

The output shows the results of fitting a linear model to describe the relationship between chlorine and weeks. The equation of the fitted model is

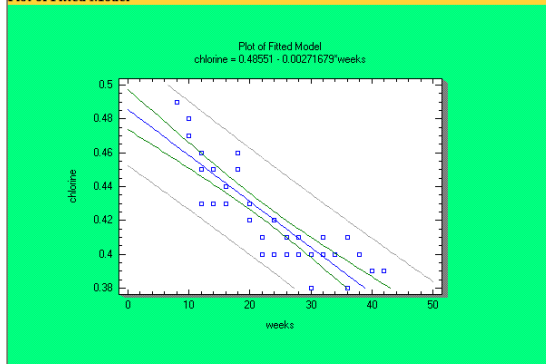
$$\text{chlorine} = 0.48551 - 0.00271679 \cdot \text{weeks}$$

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between chlorine and weeks at the 95% confidence level.

The R-Squared statistic indicates that the model as fitted explains 74.8321% of the variability in chlorine. The correlation coefficient equals -0.865055, indicating a moderately strong relationship between the variables. The standard error of the estimate shows the standard deviation of the residuals to be 0.015385.

The mean absolute error (MAE) of 0.012834 is the average value of the residuals. The Durbin-Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the P-value is less than 0.05, there is an indication of possible serial correlation at the 95% confidence level. Plot the residuals versus row order to see if there is any pattern that can be seen.

**Plot of Fitted Model**



**The StatAdvisor**

The output shows the results of fitting a linear model to describe the relationship between chlorine and weeks. The equation of the fitted model, shown as a solid line, is

$$\text{chlorine} = 0.48551 - 0.00271679 \cdot \text{weeks}$$

The inner bounds show 95% confidence limits for the mean chlorine of many observations at given values of weeks. The outer bounds show 95% prediction limits for new observations.

Results have been saved to a file named:  
[http://localhost/statgraphics1/temp/rwxp5qgxp5imebtyp5tiew\\_results.xml](http://localhost/statgraphics1/temp/rwxp5qgxp5imebtyp5tiew_results.xml)

Once in the browser, you can scroll through the output. You can also use the browser to print the output, copy it to your system's clipboard, or save the graphics images on your computer.

## 4.6 Preferences

At any time while in a statistical procedure, you can press the *Preferences* button to change system preferences. This accesses each of the pages described earlier in Chapter 3, which include settings such as the output language. If the *Calculate* button has already been pressed, it will be executed again using the new settings.

## 4.7 Save Script

If you wish to view or save the XML script generated by the statistical procedure, press the *Save Script* button. This will display the following page:



Scripts are normally saved with the extension *.sgs*. Once a script has been saved, it can be executed again at a later time by selecting *File - Execute Script* from the main menu.



## Chapter 5 – Working with Scripts

The instructions sent to the server when the *Calculate* button is pressed within a statistical procedure are contained in an XML script. The script tells the .NET Web Services where to locate the data to be analyzed, which procedure to execute, and what options are desired. As outlined in Chapter 4, these scripts may be saved for later reexecution.

A typical script contains several sections. The top section indicates the identity of the script and specifies the values of any global settings:

```
<?xml version="1.0" ?>
<statgraphics>
  <globals>
    <BorderColor Value="SpringGreen" />
    <InputDateFormat Value="M/D/YYYY" />
    <OutputDateFormat Value="M/D/YYYY" />
    <ThreeDFrameEffects Value="Yes" />
    <FillColor1 Value="SteelBlue" />
    <PointType1 Value="Plus" />
  </globals>
```

Each global settings is followed by its desired value.

The second section of the script indicates the name of the data source and information about how the data is formatted:

```
<data Alias="nonlin.xml" Source="temp/nonlin.sgd">
  <DecimalSeparator Value="." />
  <DateFormat Value="M/D/YYYY" />
  <MissingValue Value="" />
</data>
```

The *Alias* is the name of the original data file. *Source* specifies the relative location where the data is stored on the server. Data not contained in STATGRAPHICS .SGD files on the server is copied to a STATGRAPHICS data file when it is first loaded.

The last part of the script specifies information about the procedure to be run:

```
<proc name="SREG">
  <input>
    <Y Value="chlorine" />
    <X Value="weeks" />
  </input>
  <options>
    <Model Value="SquaredYReciprocalX" />
  </options>
```

```

<output>
  <table Name="Summary" />
  <graph Name="Model">
    <Resolution Value="101" />
  </graph>
  <table Name="Forecasts">
    <X Value="10" />
    <X Value="20" />
    <X Value="30" />
    <X Value="40" />
  </table>
  <graph Name="Observed" />
  <table Name="LackOfFit" />
  <table Name="Comparison" />
  <table Name="Residuals" />
  <graph Name="ResidsVsX" />
  <graph Name="ResidsVsPred" />
  <graph Name="ResidsVsRow" />
  <table Name="Influential" />
</output>
<results>
  <Residuals Name="RESIDS" />
</results>
</proc>
</statgraphics>

```

The *Input* section specifies the columns containing the data to be analyzed. The *Options* section sets the values of any analysis options. The *Output* section identifies the tables and graphs to be created. The *Results* section indicates which numerical results should be saved for later use.

Currently, there are two procedures on the STATGRAPHICS Online menu that deal with scripts:

1. *Execute script* – This option reexecutes a saved script. When the script is executed, it will look for the original data file specified in the *Alias* tag. If the data has changed, the results will be different.
2. *Manage Scripts* – The option allows scripts to be copied, renamed, deleted, or downloaded to your client machine.

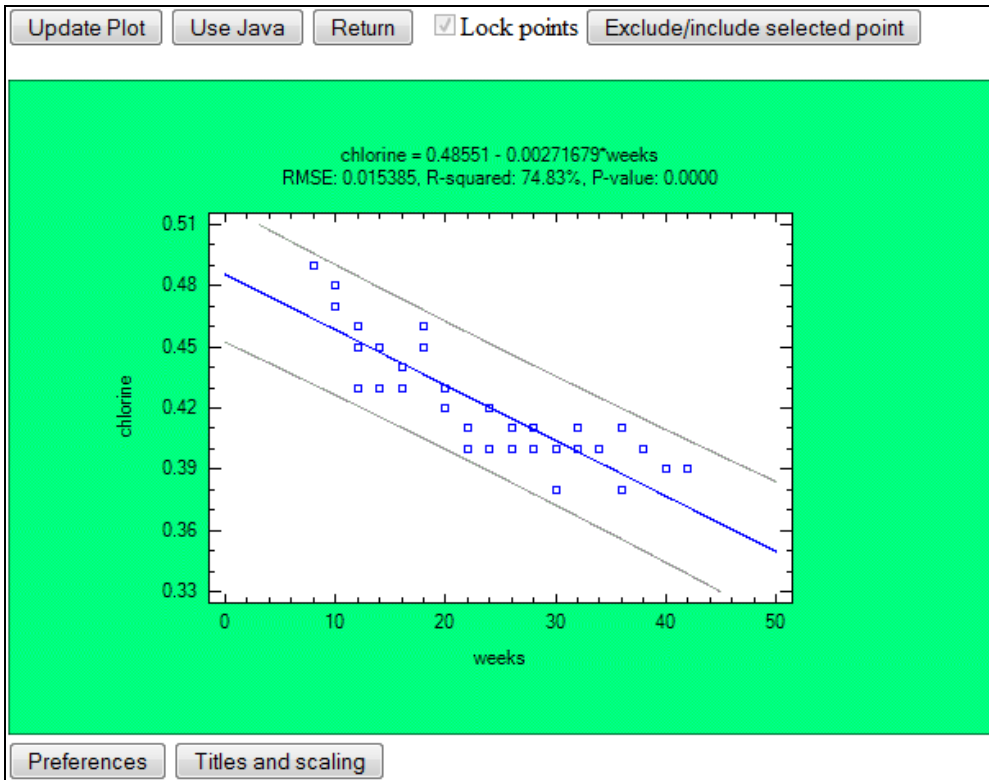
Beginning with Version XVI of STATGRAPHICS Centurion, these scripts will also be executable by the desktop Windows program.

## Chapter 6 – STATLETS

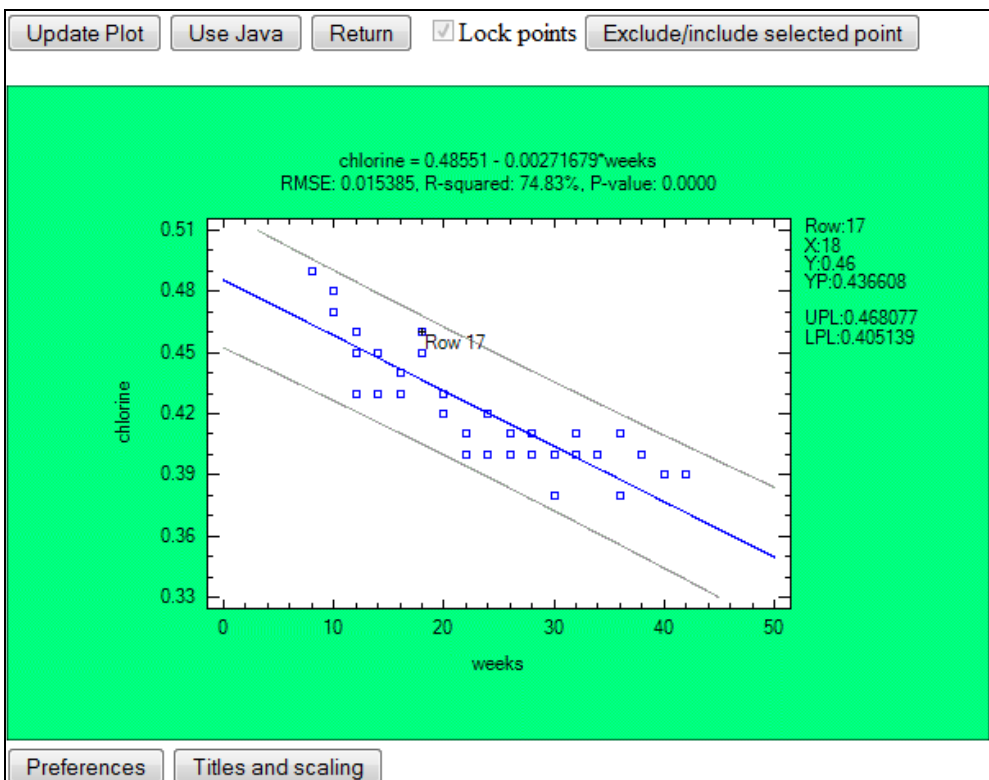
The main menu contains a special section labeled Statlets that provides access to special procedures that allow you to interact with your data. For example, select *File - Open Sample Dataset* from the main menu and open the sample dataset for the *Shelf life study*. Then select *Statlets - Regression Analysis - Curve Fitting* from the menu. This will display the page shown below:

The screenshot shows the 'Curve-Fitting Statlet' interface. At the top, there is a logo for 'statgraphics® online' and a banner that says 'Powered by STATGRAPHICS WEB SERVICES'. Below this, the title 'Curve-Fitting Statlet' is followed by a descriptive paragraph: 'This applet fits linear and nonlinear models involving a dependent variable Y and a single independent variable X. To remove a individual point, select it and then press the +/- button. If your browser is Java-enabled, you can also use your mouse to drag points to new locations. If the linear model is not adequate, you can try transforming X or Y or both. [Documentation](#)'. The interface includes several input fields and buttons: 'Y: chlorine' with an 'Edit' button; 'X: weeks' with an 'Edit' button; '(Labels):' with an 'Edit' button; '(Select):' with a 'Help' button; a checked checkbox for 'Add regression curve' with 'Confidence limits: None' and '95%' dropdowns, and 'Prediction limits: Two-sided' and '95%' dropdowns; 'X power: 1' and 'Y power: 1' input fields with 'Opt X', 'Opt Y', 'Opt Both', and 'Reset' buttons; an unchecked checkbox for 'Add LOWESS Smooth' with a 'Window: 50%' dropdown; and a row of buttons: 'Update Plot', 'Use Java', 'Return', a checked checkbox for 'Lock points', and 'Exclude/include selected point'.

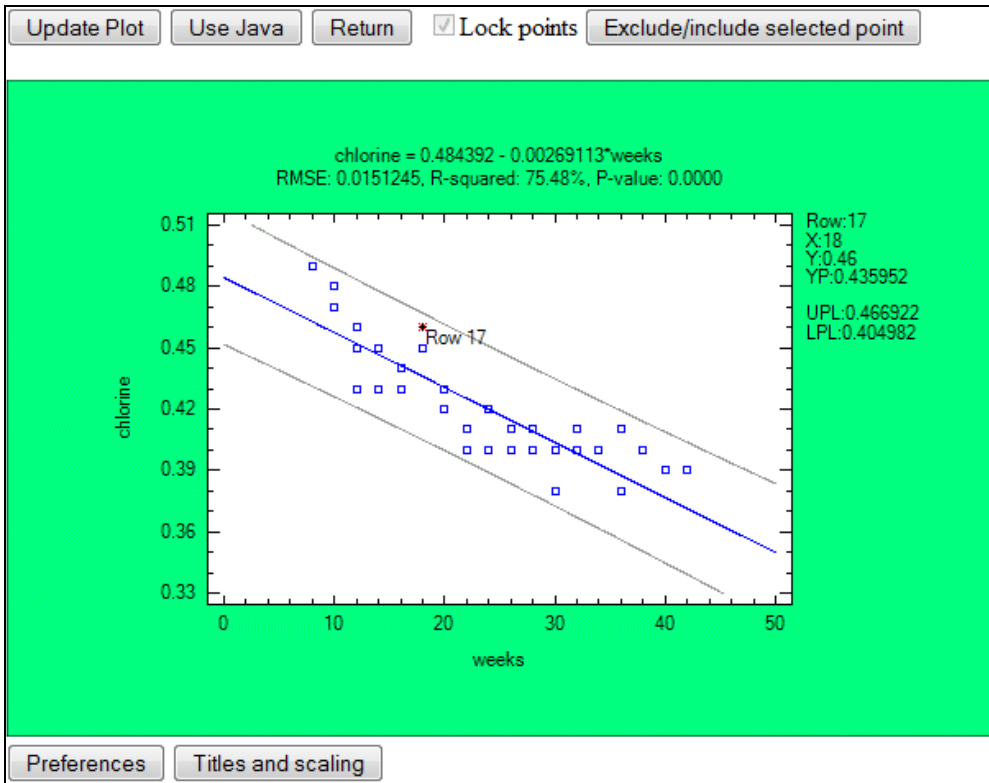
Select the *Y* and *X* variables and then press *Update Plot*. This will add the graph shown below to the page:



If your web browser supports HTML5 canvases, you may click on or touch any point on the graph to display its row number and coordinates:

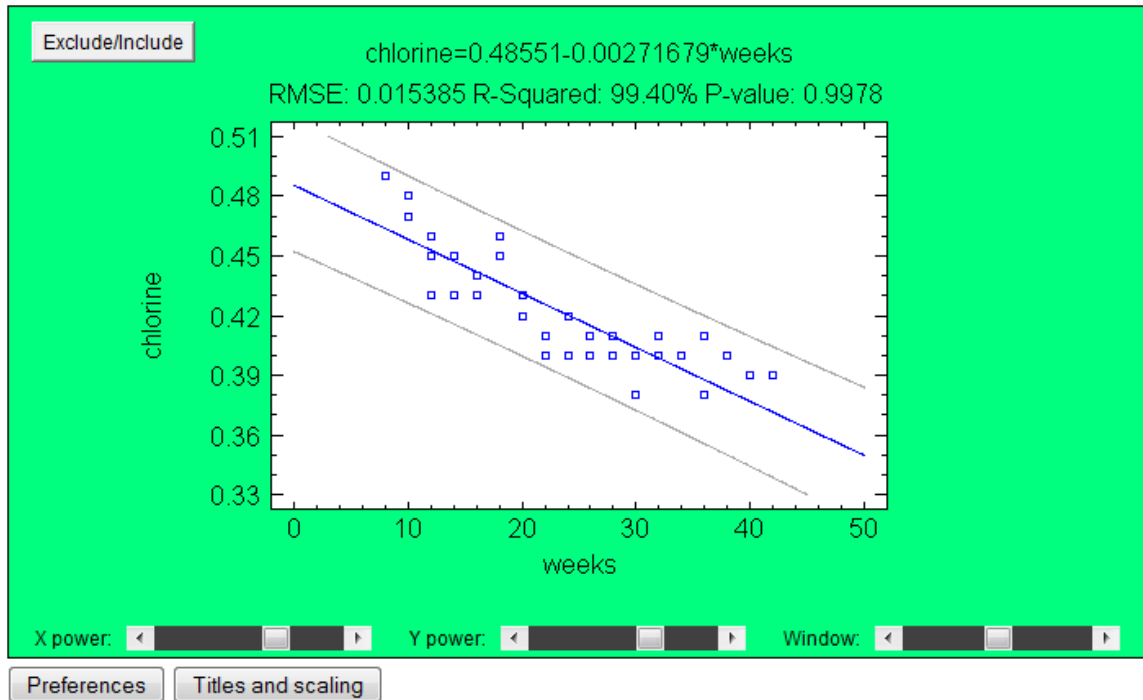


After clicking on a point, you may then press the *Exclude/include selected point* button to remove it from the fit:



Excluded points are displayed in red.

If your web browser supports Java, you may press the *Use Java* button to load a Java applet in place of the canvas:



Java applets allow changes to be made to the graph without posting the page back to the server. With the *Curve Fitting* applet, you may:

1. Click on a point to display its row number or label. The coordinates of the point and its predicted value will be displayed in the right margin of the graph.
2. Drag a point to a new location with your mouse and observe the effect on the fitted line. (You must first uncheck *Lock points*).
3. Click on a point and press the *Exclude/Include* button to remove that point completely from the fit.
4. Click on any location of the graph to display a vertical cursor at the selected value of X. The predicted response and limits at that X will be displayed in the right margin of the graph.
5. Use the *X power* and *Y power* sliders to fit a nonlinear regression model involving powers of the variables.
6. Use the *Window* slider to change the width of the LOWESS smoothing window (if *Add LOWESS smooth* has been selected).

By interacting with the applet, you can often see interesting features of your data.

# Index

column names .....	5	titles and scaling.....	25
confidence level .....	13	width and height.....	17
data		language setting .....	15
copy and paste.....	7	main page .....	2
delete .....	8	missing values .....	14
editor .....	4	passwords .....	3
entry .....	4	preferences .....	13
formatting.....	11	sample datasets.....	9
insert.....	8	scripts .....	31
paging.....	7	select field	
printing.....	8	Boolean expressions.....	23
saving .....	6	EXCLUDE.....	23
sorting .....	8	FIRST.....	23
data files		LAST.....	23
copying.....	12	RANDOM.....	23
deleting.....	12	ROWS .....	23
downloading.....	12	significant digits.....	13
renaming .....	12	StatAdvisor .....	14
SGD files.....	7	STATGRAPHICS expressions .....	22
uploading.....	12	statistical procedures	
date format .....	15	analysis options.....	24
decimal separator .....	15	calculate .....	28
Excel files.....	10	data input.....	21
fonts.....	16	preferences .....	30
graphs		results to save.....	28
3D effects .....	17	save script.....	30
3D viewpoint.....	27	tables and graphs.....	25
background color .....	17	Statlets.....	33
fills .....	18	tables	
function resolution .....	17	maximum rows.....	13
legend block offset.....	27	maximum width .....	14
lines .....	19	terms and conditions .....	1
margins.....	27	user accounts .....	2
points.....	20	user data files .....	9