



USER'S MANUAL



Statistical Tool for Agricultural Research

Version 2.0.1

January 2014

A NOTE TO THE READER:

An electronic copy of this user's manual comes with the STAR installer. The STAR User's Manual may be printed/copied and distributed to any number of users. STAR is a freeware developed for non-profit use. Hence, selling of either the software or the user's manual is prohibited.

Biometrics and Breeding Informatics
Plant Breeding, Genetics and Biotechnology Division
INTERNATIONAL RICE RESEARCH INSTITUTE

Contents

INTRODUCTION	1
FEATURED MODULES	1
NEW FEATURES IN STAR 2.0.1	2
ABOUT THE USER'S MANUAL.....	2
CITING THE SOFTWARE	2
GETTING STARTED	3
INSTALLATION.....	3
LAUNCHING STAR.....	10
PROJECT EXPLORER PANEL	11
EDITOR PANEL	12
MENU BAR	13
DATA VIEWER.....	15
RESULT VIEWER	16
PROJECT AND FILE MANAGEMENT	19
CREATING NEW PROJECT	19
OPEN EXISTING PROJECT.....	21
RENAMING PROJECT	21
EXPORTING PROJECT	23
DELETING PROJECT	24
IMPORTING PROJECT.....	26
IMPORTING DATA FROM EXTERNAL SOURCE	27
IMPORTING DATA FROM OTHER PROJECT.....	29
IMPORTING DATA FROM PACKAGE.....	31
CREATING NEW FOLDER	32
COPYING FILE/FOLDER	34
RENAMING FILE/FOLDER	36
MOVING FILE/FOLDER	38
DELETING FILE/FOLDER	40
SWITCHING WORKSPACE	41
DATA HANDLING AND MANIPULATION	45
REPRESENTING MISSING VALUE(S)	45
INSERT ROW(S).....	45
DELETE ROW(S)	47
INSERTING COLUMN(S).....	48
DELETING COLUMN(S).....	50
CREATE NEW VARIABLES.....	51
EDITING VARIABLE INFORMATION.....	55
SORTING.....	57

MERGING DATASETS	60
APPENDING DATASETS	63
CREATING DATA SUBSET	65
RESHAPING	68
<i>Reshaping Data from Long (Serial) to Wide (Parallel) Format</i>	68
<i>Reshaping Data from Wide (Parallel) to Long (Serial) Format</i>	71
AGGREGATING DATA.....	75
RANDOMIZATION AND LAYOUT	81
COMPLETELY RANDOMIZED DESIGN	81
RANDOMIZED COMPLETE BLOCK DESIGN	85
LATIN SQUARE DESIGN	91
SPLIT PLOT DESIGN	94
STRIP PLOT DESIGN.....	102
SPLIT-SPLIT PLOT DESIGN.....	106
STRIP-SPLIT PLOT DESIGN	112
SPLIT-SPLIT-SPLIT PLOT DESIGN	117
STRIP-SPLIT-SPLIT PLOT DESIGN	122
BALANCED INCOMPLETE BLOCK DESIGN.....	127
GRAPHS	133
BAR GRAPH.....	133
BOX PLOT	140
HISTOGRAM	146
LINE GRAPH	151
PIE CHART	157
SCATTER PLOT	162
DESCRIPTIVE STATISTICS	169
CROSS TABULATION.....	175
NORMALITY TEST.....	179
HOMOGENEITY OF VARIANCES.....	187
FITTING DISTRIBUTIONS	195
T-TEST	201
ONE SAMPLE T-TEST	201
PAIRED SAMPLE T-TESTS	205
INDEPENDENT SAMPLE T-TESTS	209
CHI-SQUARE TEST	215
GOODNESS OF FIT TEST.....	215
TEST OF INDEPENDENCE.....	217
TEST ON PROPORTION.....	223
ANALYSIS OF VARIANCE.....	227

COMPLETELY RANDOMIZED DESIGN	228
RANDOMIZED COMPLETE BLOCK DESIGN	235
LATIN SQUARE DESIGN	240
SPLIT PLOT DESIGN	246
STRIP PLOT DESIGN.....	253
SPLIT-SPLIT PLOT DESIGN.....	260
STRIP-SPLIT PLOT DESIGN	267
SPLIT-SPLIT-SPLIT PLOT DESIGN	274
STRIP-SPLIT-SPLIT PLOT DESIGN	276
BALANCED INCOMPLETE BLOCK DESIGN.....	278
PAIRWISE MEAN COMPARISON	284
PARTITIONING SUM OF SQUARES.....	290
<i>Group Comparison</i>	290
<i>Trend Comparison</i>	294
CORRELATION ANALYSIS.....	297
REGRESSION ANALYSIS	305
LINEAR REGRESSION.....	305
NON-LINEAR REGRESSION.....	312
NONPARAMETRIC TESTS.....	321
ONE SAMPLE	321
TWO RELATED SAMPLES	324
TWO INDEPENDENT SAMPLES	326
TEST FOR SEVERAL RELATED SAMPLES	328
TEST FOR SEVERAL INDEPENDENT SAMPLES	331
MULTIVARIATE ANALYSIS	336
PRINCIPAL COMPONENT ANALYSIS.....	336
CLUSTER ANALYSIS.....	343
MULTIVARIATE ANALYSIS OF VARIANCE	351
MULTIDIMENSIONAL SCALING.....	357
DISCRIMINANT ANALYSIS	362
APPENDIX	370
DATA FROM STAR PACKAGE	370
<i>AgroData1</i>	370
<i>AgroData2</i>	370
<i>CarbonStock</i>	371
<i>AugLS</i>	371
<i>AugRCB</i>	372
<i>CochranBIB</i>	372
<i>Corr</i>	373
<i>CropStatRCBDOverSite</i>	373
<i>Dule_AlphaGxE</i>	374
<i>Dune</i>	375

<i>Fertilizer</i>	376
<i>fishData</i>	377
<i>GeneticDistMatrix</i>	377
<i>Gerua</i>	377
<i>GomezMaizeCross</i>	378
<i>GomezMaizeCrossSummary</i>	379
<i>GomezSplitRCBD</i>	380
<i>GomezSplit2RCBD</i>	380
<i>GomezStrip</i>	381
<i>GomezStripSplit</i>	382
<i>GomezTenureStatus</i>	383
<i>GomezTwoCropSeasonRCB</i>	383
<i>grp_pss</i>	384
<i>IASRICorrReg</i>	384
<i>IASRILSD</i>	385
<i>IASRITomar</i>	386
<i>IASRITomar2</i>	387
<i>InjuryProfile</i>	387
<i>Inoculation</i>	388
<i>Insecticides</i>	389
<i>JohnAlpha</i>	389
<i>KemptonRowCol</i>	390
<i>KijojiDS2010Cluster</i>	391
<i>Maize</i>	391
<i>ManovaData1</i>	392
<i>Nurul2012</i>	392
<i>OP_pss</i>	393
<i>PGTYPIC</i>	394
<i>SeedingRate</i>	394
<i>Sensors</i>	395
<i>Sevilla2012Cluster</i>	395
<i>soil_redox</i>	396
<i>soil_var</i>	397
<i>Survival</i>	397
<i>tf_rcb</i>	398
<i>upland</i>	398
<i>White_heads</i>	399
<i>Year</i>	399
<i>Yield94</i>	399

Introduction

Statistical Tool for Agricultural Research (STAR) is a computer program for data management and basic statistical analysis of experimental data. It has a user-friendly graphical interface where items are accessible via drop-down menus. Its graphical interface was created using the Eclipse Rich Client Platform (RCP) and uses the R language and environment for statistical computing and graphics. The programs uses functions in R that are specifically written for the development of this computer program and those that are available in the following R packages: ade4, agricolae, dae, DiGGER, doBy, fields, fitdstrplus, gdata, gplots, Hmisc, leaps, lme4, lmerTest, MASS, multcomp, mvtnorm, numDeriv, pastecs, permute, plyr, qtl, R.methodsS3, R.oo, rJava, spam, stringr and vegan.

STAR has been developed primarily for the analysis of data from agricultural field trials, but many of the features can be used for analysis of data from other sources.

Featured Modules

STAR has the following modules:

- Data management with a spreadsheet view
- Randomization for commonly used experimental design
- Graphics
- Descriptive statistics
- Analysis of variance for experimental designs
- Correlation and linear regression analysis
- Multivariate Analysis
- Non-parametric data analysis

New Features in STAR 2.0.1

The following are the revisions made in the latest version of STAR:

- A comma-separated-value file will be created as an additional output for the Descriptive Statistics
- Residual plot is displayed when testing the assumption for the reduced model in Multiple Regression
- Results of the test of Heteroscedasticity are displayed when Breush Pagan and Goldfeldt Quandt are requested for the Multiple Regression
- Randomized Complete Block Design as the default design for the MANOVA
- Default probabilities can be changed for the Discriminant Analysis
- Results are displayed if the divisive method is chosen for the Cluster Analysis
- PCA biplot are displayed for all possible combination of PC1 to PC3 for the Principal Component Analysis

About the User's Manual

The STAR user's manual provides step-by-step instructions on how to perform certain tasks of interest to users. Tasks are discussed using sample datasets that are included when STAR is installed. In most tasks, dialog boxes are displayed and the user has to specify required fields and additional options as desired. Descriptions and limitations of these fields and options are also provided. Screen images are also included for additional references.

This manual is written using the following format:

- Menu items, names of dialogs and form controls are in **bold** letters
- Project names, filenames, variable names and directories are *italicized*
- Menus appear in the form **Project | New Project**, which means “choose **New Project** from the **Project** menu”

Citing the Software

The suggested citation of STAR in publication is as follows:

STAR, version 2.0.1 2014. Biometrics and Breeding Informatics, PBGB Division, International Rice Research Institute, Los Baños, Laguna.

Getting Started

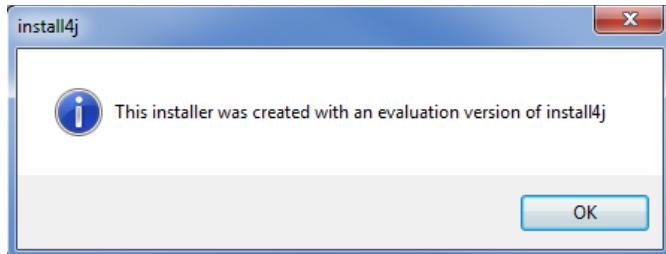
Installation

The installer of STAR 2.0.1 and its required R packages can be downloaded from bbi.irri.org. To download the installers, click the links labeled as *STAR 2.0.1* and *R-Packages 1.3* at the right-hand side of the page.

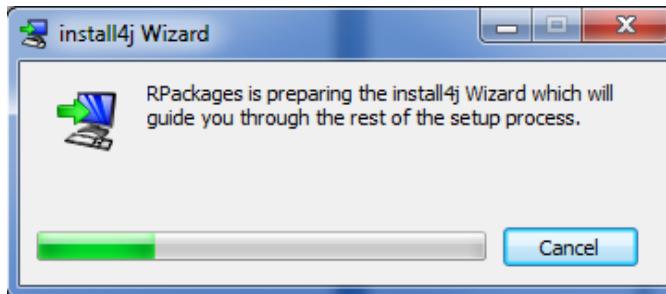
The screenshot shows the IRRI Biometrics and Breeding Informatics website. The main navigation bar includes links for Home, About Us (which is highlighted), Products, Trainings, Research, and Contact Us. The 'About Us' page content discusses the group's research support, consultation, and training services. A sidebar on the right is titled 'Products Download' and lists several links: 'Statistical Tool for Agricultural Research STAR 2.0.1 Installation Procedure', 'Plant Breeding Tools PBTools 1.2 Installation Procedure', 'STAR and PBTools required R Packages R-Packages1.3', and 'Android Tablet Application for Data Collection FieldLab 2.8'. Below this is a section titled 'Interact with us on' with links for facebook, twitter, and scribd. A red box highlights the 'Products Download' section, and another red box highlights the 'R-Packages1.3' link.

User need to install *R-Packages* first before installing *STAR*. The steps to install the required R packages are as follows:

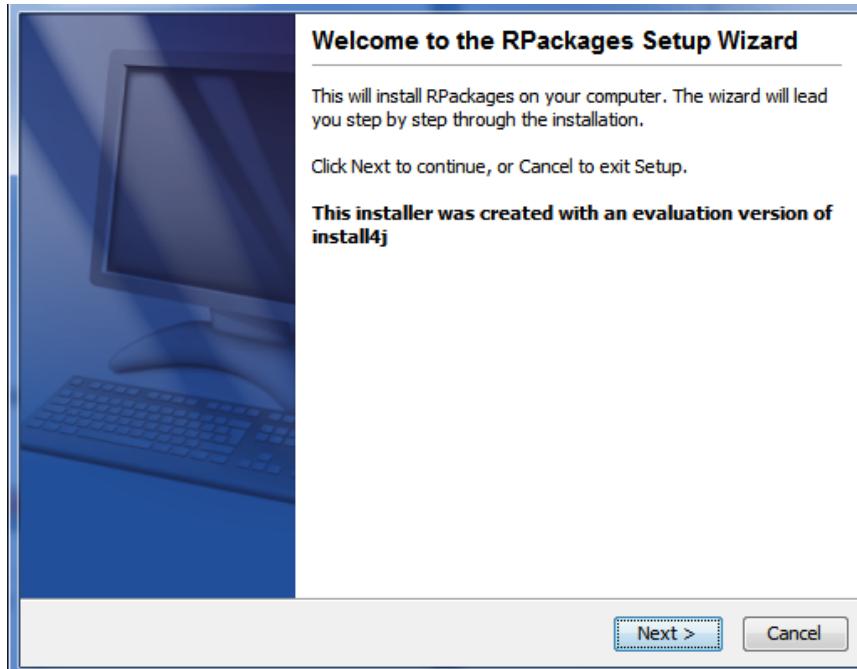
- Double-click the installer icon  **RPackages-1.2-windows-setup-32bit** to launch the setup. A dialog as shown below is displayed. Click the **OK** button.



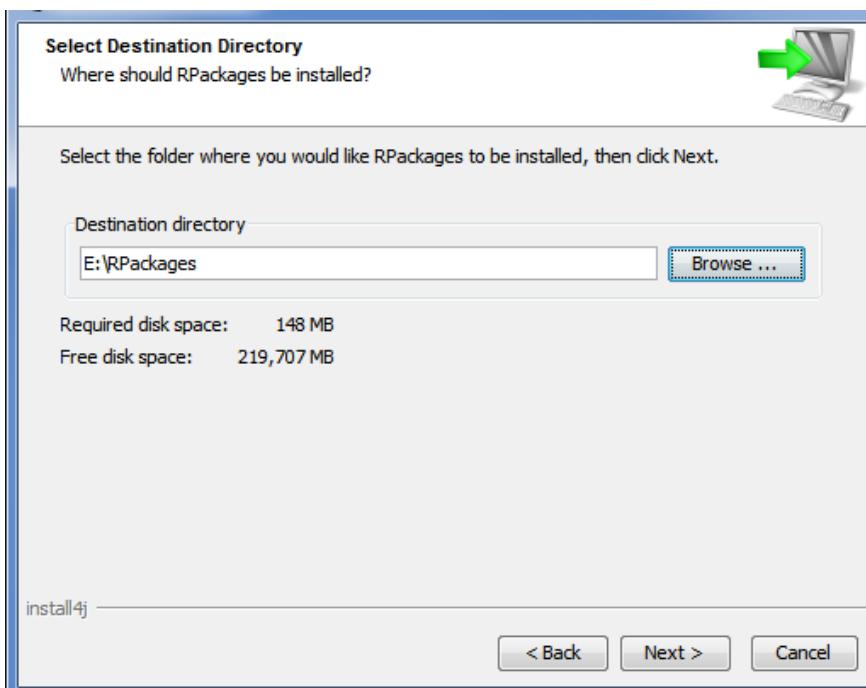
- The next step is the preparation of the wizard. A dialog with progress bar is displayed. Wait until the preparation is completed.



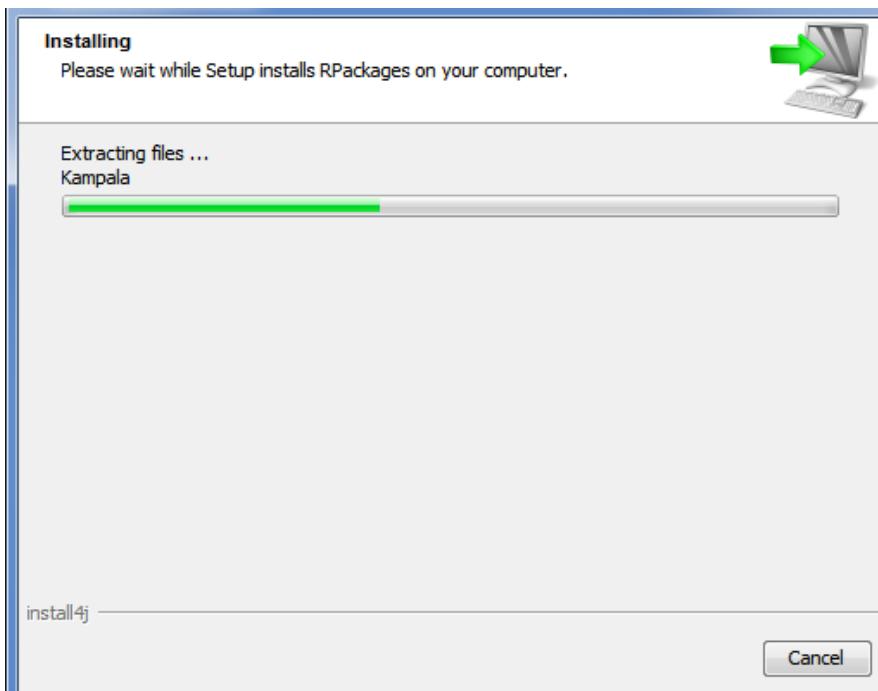
- When installing R Packages for the first time, the **Welcome** dialog as shown below is displayed. Click the **Next** button.



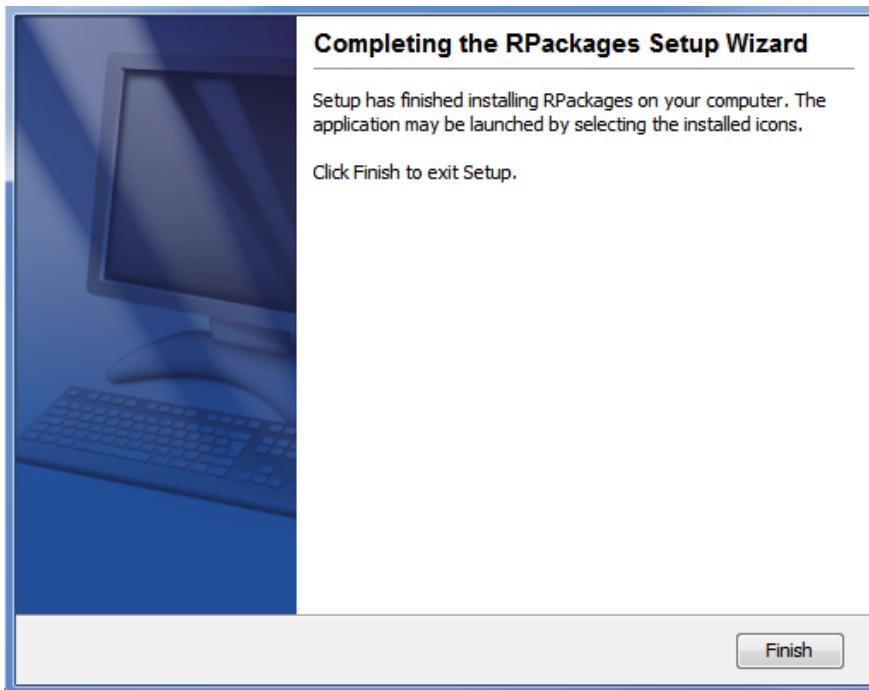
- Select the destination folder, preferably not in **C:\Program Files** as problems were encountered when RPackages is installed in this location. Click the **Next** button.



- Wait until the installation is completed.

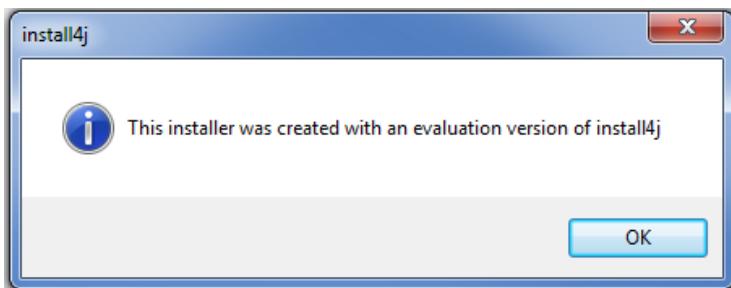


- Click the **Finish** button to exit Setup.

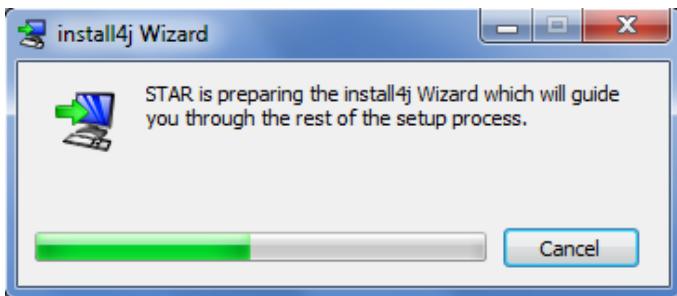


The steps to install STAR are as follows:

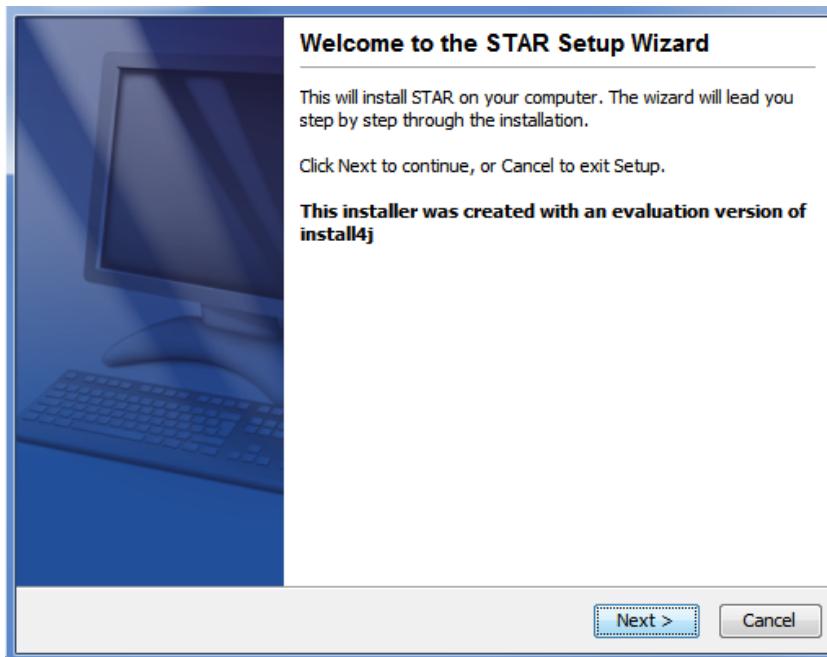
- Double-click the installer icon  STAR-2.0.1-windows-setup-32bit to launch the setup. A dialog as shown below is displayed. Click the **OK** button.



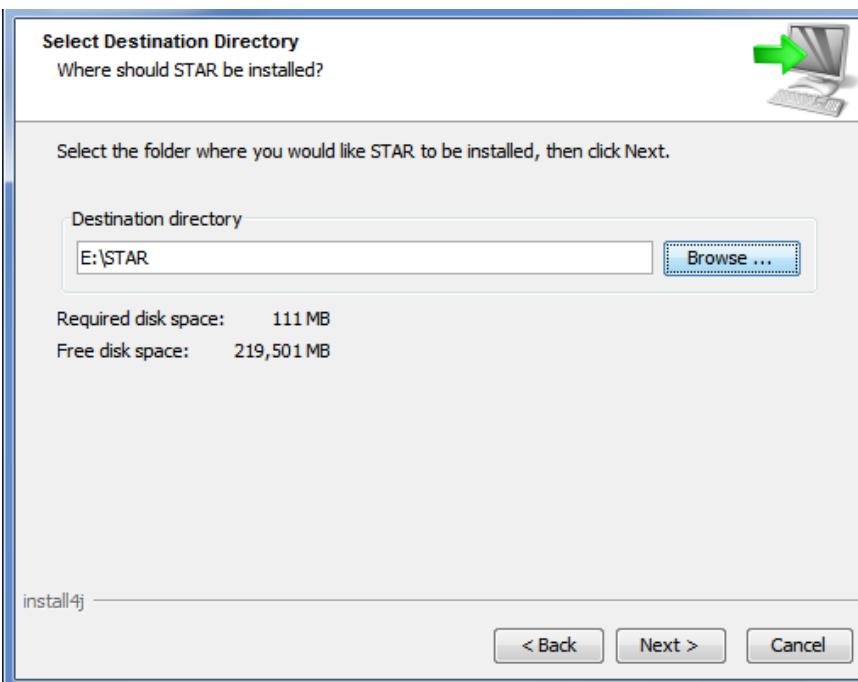
- The next step is the preparation of the wizard. A dialog with progress bar is displayed. Wait until the preparation is completed.



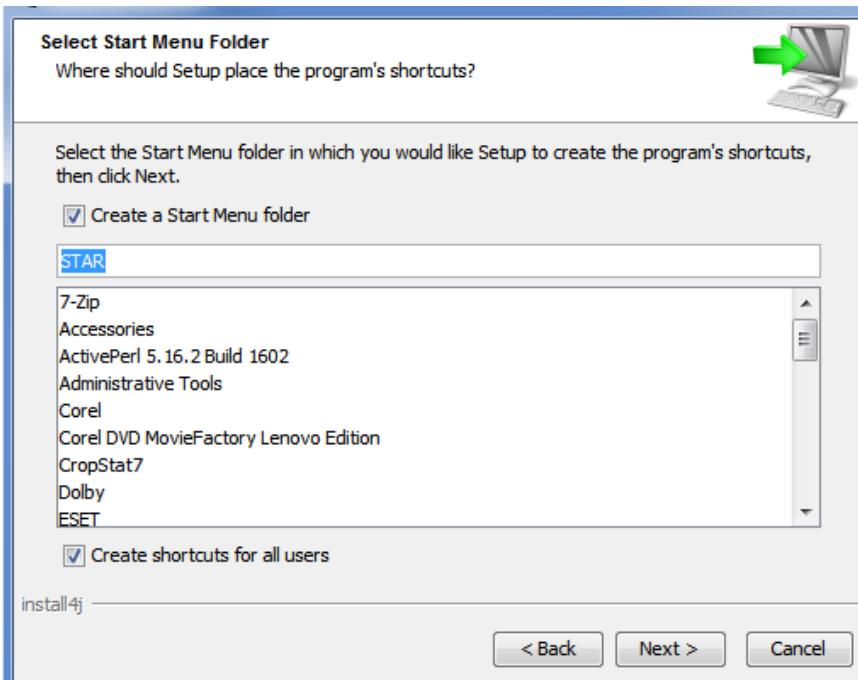
- When installing STAR for the first time, the **Welcome** dialog as shown below is displayed. Click the **Next** button.



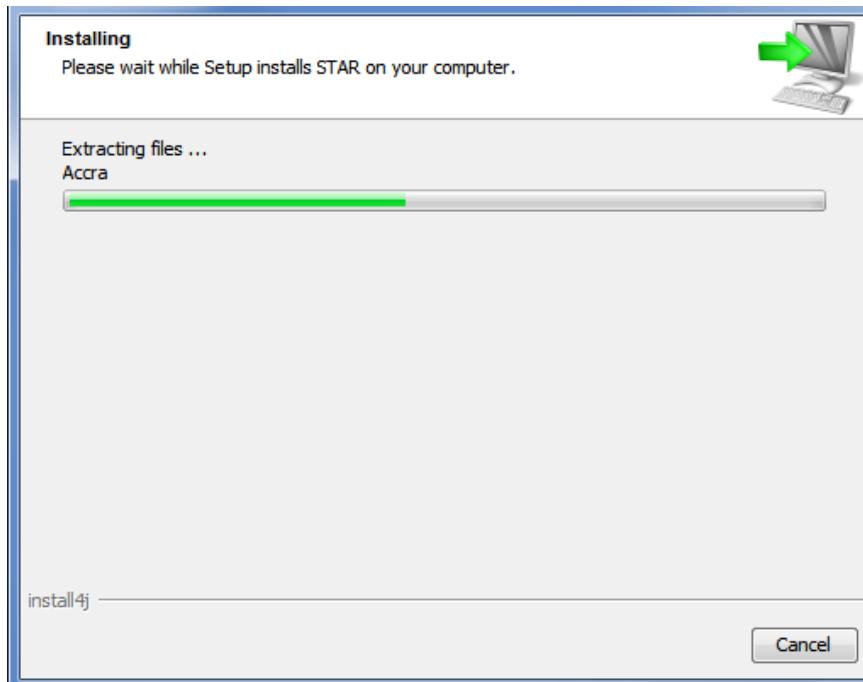
- Select the destination folder, preferably not in **C:\Program Files** as problems were encountered when STAR is installed in this location. Click the **Next** button.



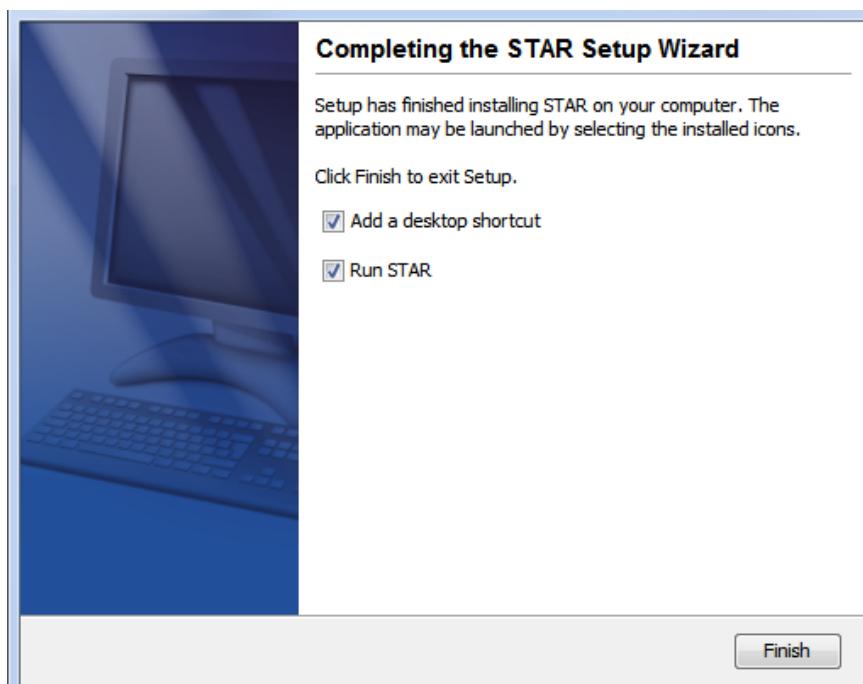
- Select which start menu folder you want to place the STAR icons in and click the **Next** button.



- Wait until the installation is completed.



- Finally, specify if you want to create a desktop icon and run STAR. Click the **Finish** button.



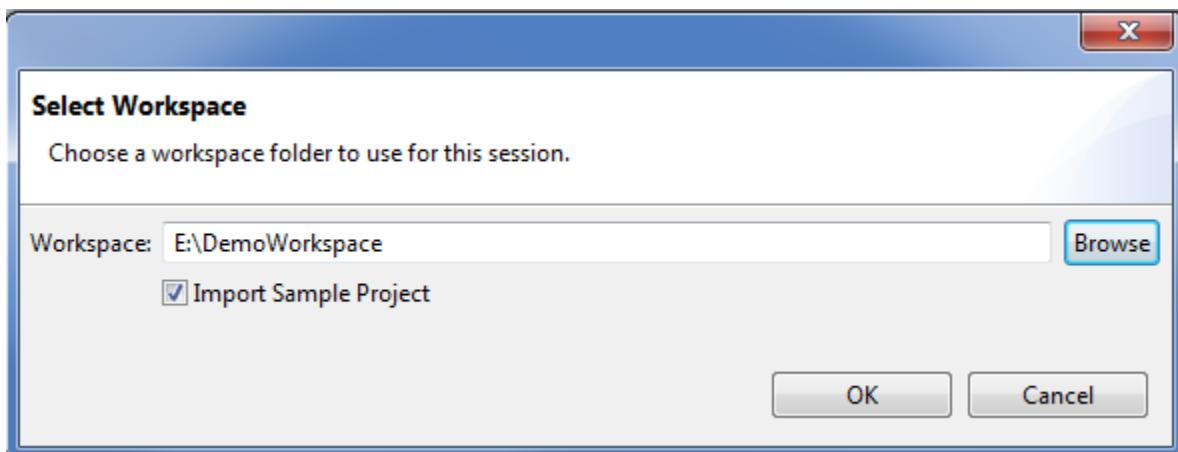
Launching STAR

To launch STAR, double-click the STAR shortcut icon on the desktop (if there is one) or click **Start** on the Windows task bar, choose **All Programs**, then click on the **STAR** folder and click the **STAR** icon.

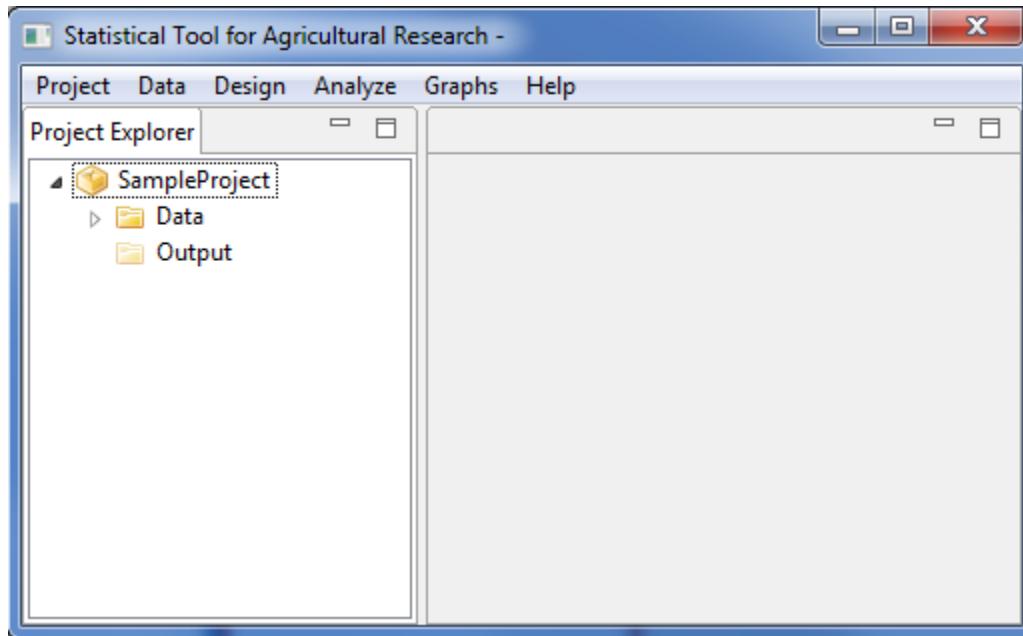
Upon launching, the splash image will appear.



When STAR is launched for the first time, the **Select Workspace** dialog is displayed. A default workspace folder is indicated in the **Workspace** entry box. To indicate a different workspace folder, click the **Browse** button and select the new workspace folder. By default, the checkbox **Import Sample Project** is selected. Then click the **OK** button.



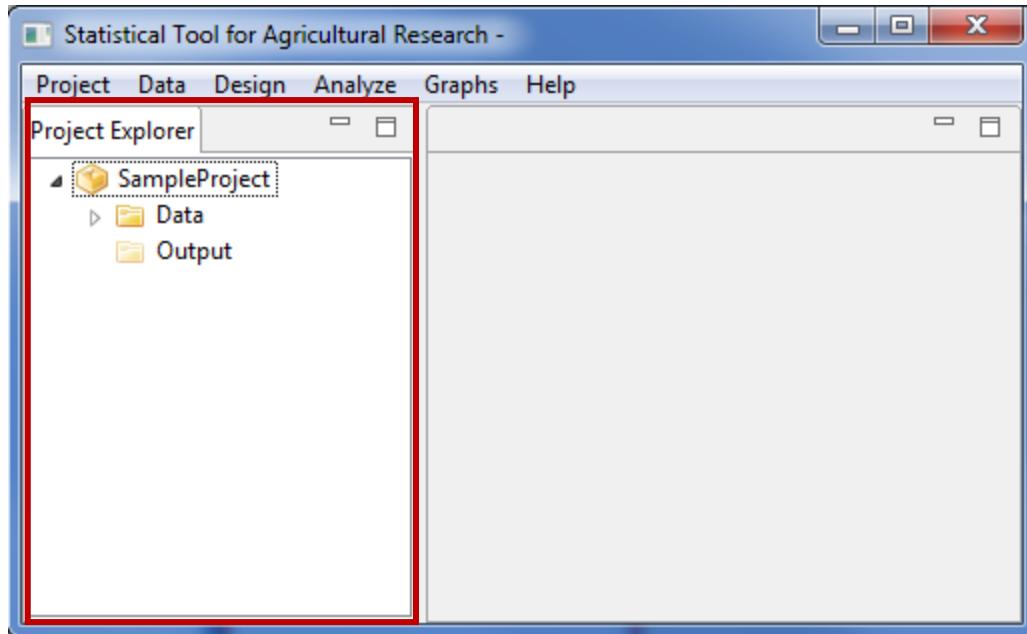
The main window of STAR will appear.



STAR main window has a menu bar and is divided into two panels, the **Project Explorer** panel located on the left-hand side and the **Editor Panel** on the right-hand side.

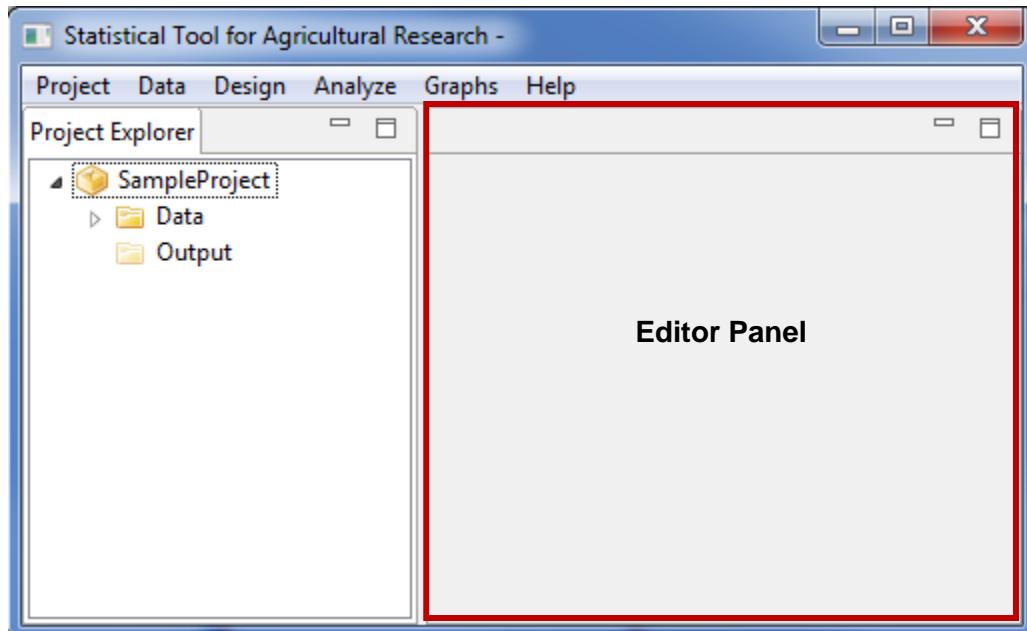
Project Explorer Panel

Project Explorer can be seen at the left-hand side of the main window. It serves as a file manager of the active project, where data and results of the analysis are displayed in a tree. It displays the last opened project from the previous STAR session. When the “Import Sample Project” option is selected in the **Select Workspace** dialog, a default project named *SampleProject* is displayed with *Data* folder and *Output* folder inside it. The *Data* folder contains some sample datasets that will be used in this manual. The *Output* folder, on the other hand, will contain all output files that will be created when an analysis is performed.



Editor Panel

Editor panel is located at the right-hand side of the main window. It consists of **Data Viewer** tab and the **Result Viewer** tab. By default, the editor panel is empty.

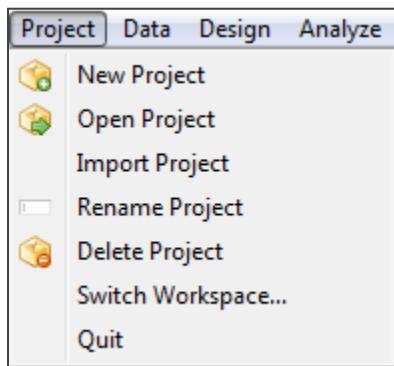


Menu Bar

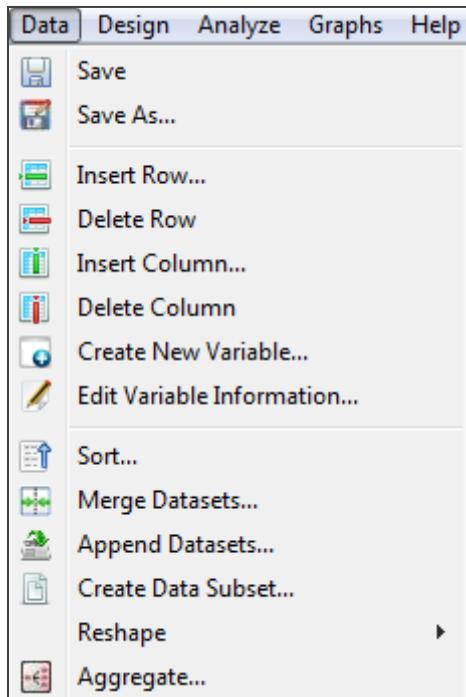
Most of the tasks that you perform in STAR start with menu selections. There are six drop-down menus available and this are:

Project Data Design Analyze Graphs Help

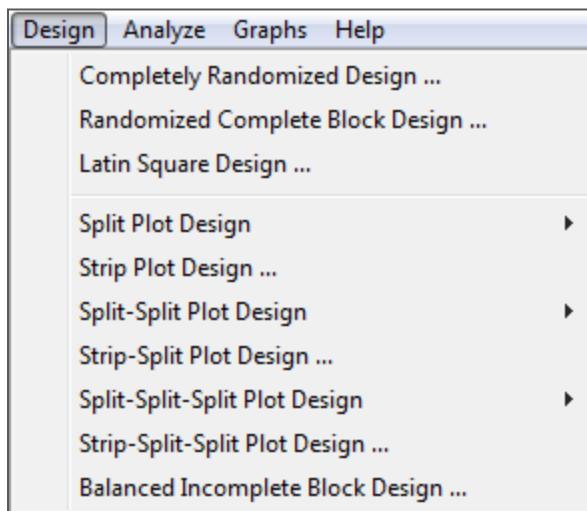
Project Menu contains functions for creating and managing project. This menu can also be used to quit or terminate sessions in STAR. It contains the following submenu items:



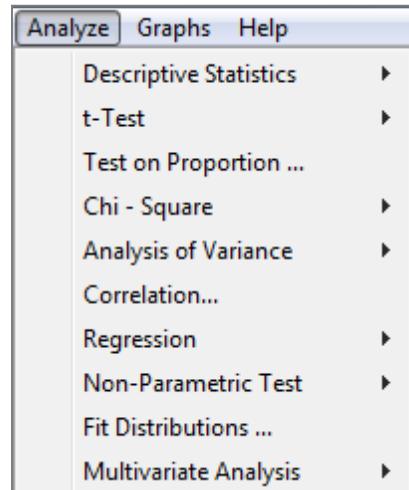
Data Menu contains functions for reading, managing and manipulating datasets. Items under this menu can be used only when a data in CSV format is displayed in the active tab of the editor panel. It contains the following submenu items:



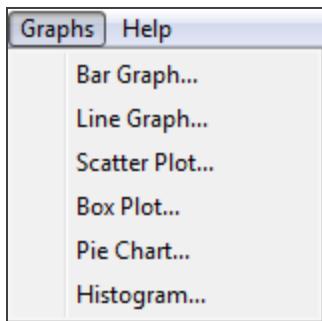
Design Menu contains functions for generating random assignment of factor levels for commonly used experimental design. It contains the following submenu items:



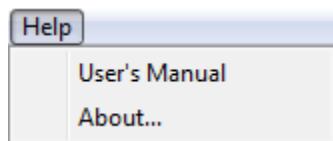
Analyze Menu contains functions to perform statistical analysis. Items under this menu can be used only when a data in CSV format is displayed in the active tab of the editor panel. It contains the following submenu items:



Graph Menu contains functions for creating the basic graphs. It contains the following submenu items:

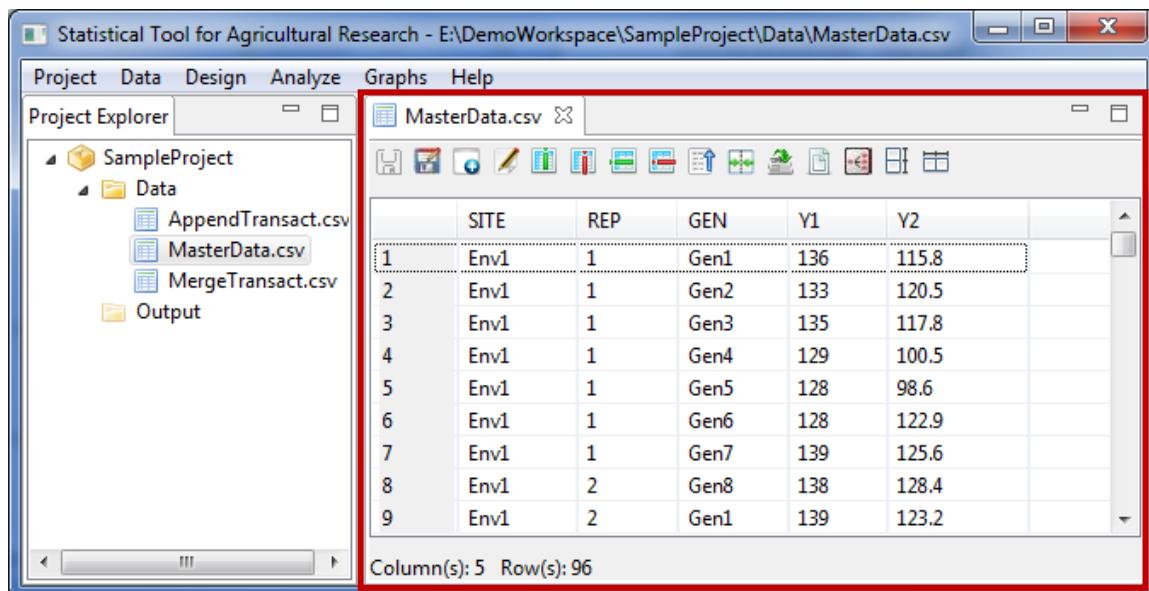


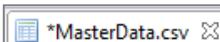
Help Menu used to access STAR's Help System or User's Manual. It contains the following submenu items:



Data Viewer

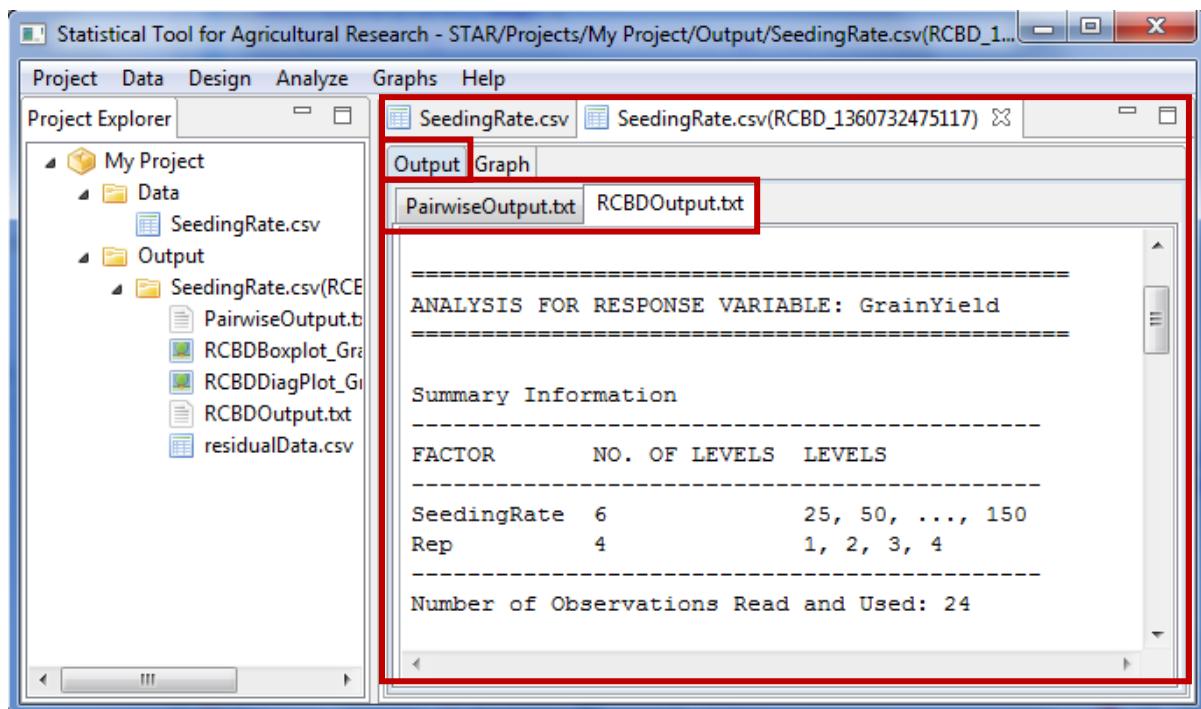
The **Data Viewer** tab is displayed in the editor panel when a data in CSV format is double-clicked from the **Project Explorer** tree. It is used for viewing data in spread-sheet format, editing data values and performing data management. Several **Data Viewer** tabs can be opened simultaneously inside the **Editor Panel** but only one is considered to be the active tab. Each **Data Viewer** tab contains toolbars which can be used for managing the data. The toolbars have the same functionalities as the options under the **Data** menu.



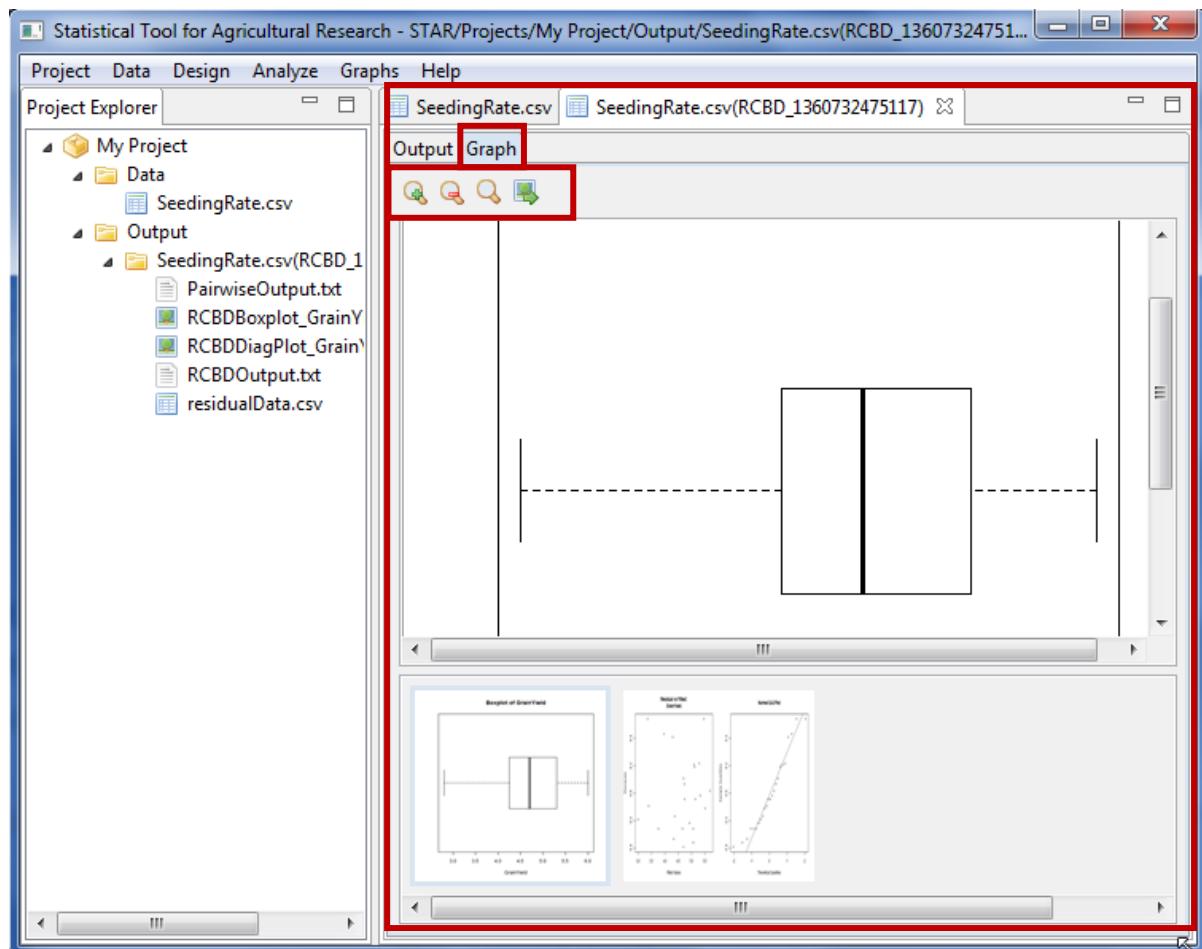
The name of the tab indicates the name of the dataset. When an asterisk appears in the tab name *, it indicates that there has been modification made in the data. If data manipulation or analysis is performed on the data on the active tab, STAR captures all information in the active tab (even unsaved modifications) and submit it for manipulation, graph creation or analysis. If a **Data Viewer** with unsaved changes is closed, the user will be prompted to save the changes.

Result Viewer

The **Results Viewer** tab is displayed in the editor panel when an analysis is performed or by double-clicking a results folder inside the *Output* folder. Depending on the contents of the results folder, the Result Viewer may have an *Output* page and/or the *Graph* page.



All outputs in *txt* format are displayed in the *Output* page while all graphical outputs in *png* format are displayed in the *Graph* page. The *Graph* page has toolbars that can be used to minimize, maximize, and view the graph to normal size and to export graph to external sources. To view other graphs in the page, the left and right arrow keys can be used.

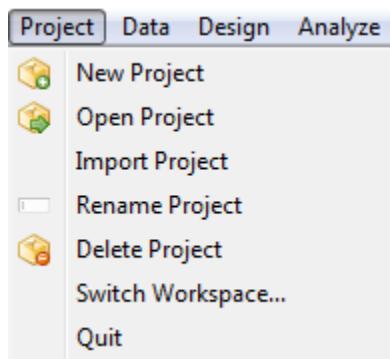


Project and File Management

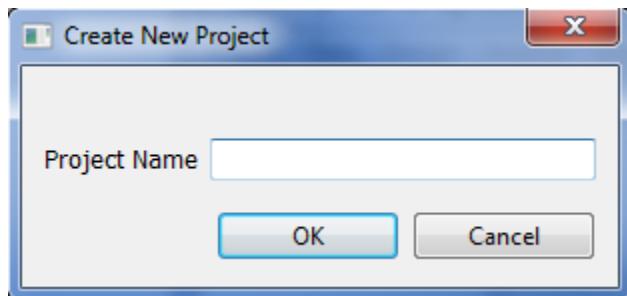
Creating New Project

Several projects can be created in STAR. However, only one project can be considered as the active project and is visible in the Project Explorer panel. The steps for creating new project are listed below.

- Select **Project | New Project** from the main menu.



- The **Create New Project** dialog box will appear. In the **Project Name** field, specify a name of the new project.

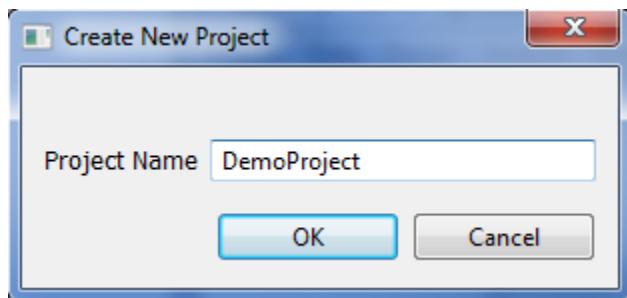


In naming a project, the following rules apply:

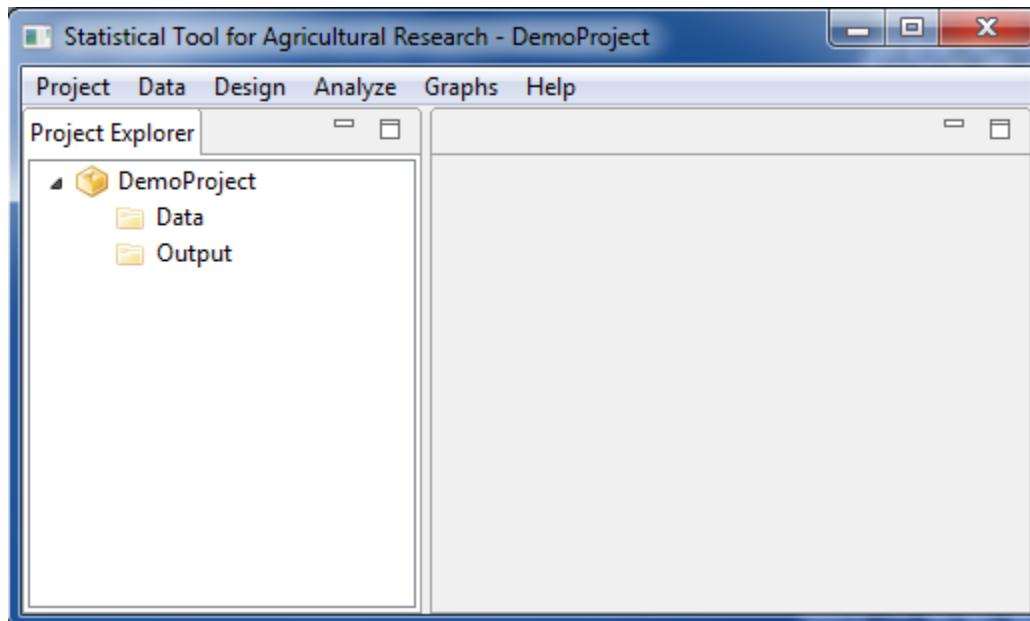
- The name must start with a letter. The remaining characters can be any letter, any digit, a period, underscore, blank or dash.

- The last character of the name should not be a period, underscore, blank or dash.
- The length of the name should not be less than 4 characters.
- The name must be different from the existing project names.
- The name is not case sensitive.

For the example, type *DemoProject*. The completed dialog box will appear as shown below:



- Click the **OK** button. The new project named *DemoProject*, will now be displayed in the **Project Explorer** panel. By default, two empty folders under the project name will be automatically created, the *Data* and *Output* folders. Note that these folders cannot be deleted, rename or move to another location.

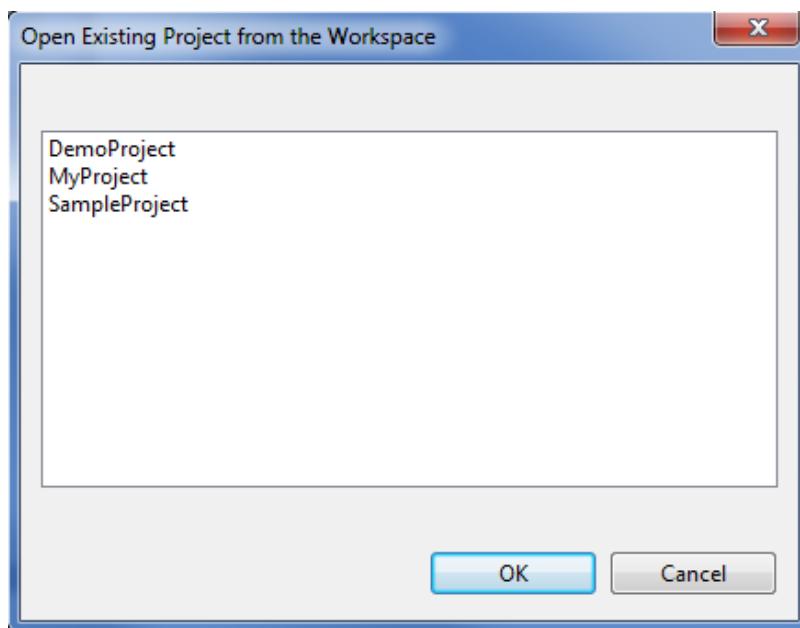


Note: As part of the examples, create another folder named *MyProject*.

Open Existing Project

The steps for opening existing project inside the STAR workspace are listed below.

- Select **Project | Open Project** from the main menu.
- The **Open Existing Project from the Workspace** dialog box will appear. The list of all existing projects in the workspace can be seen.

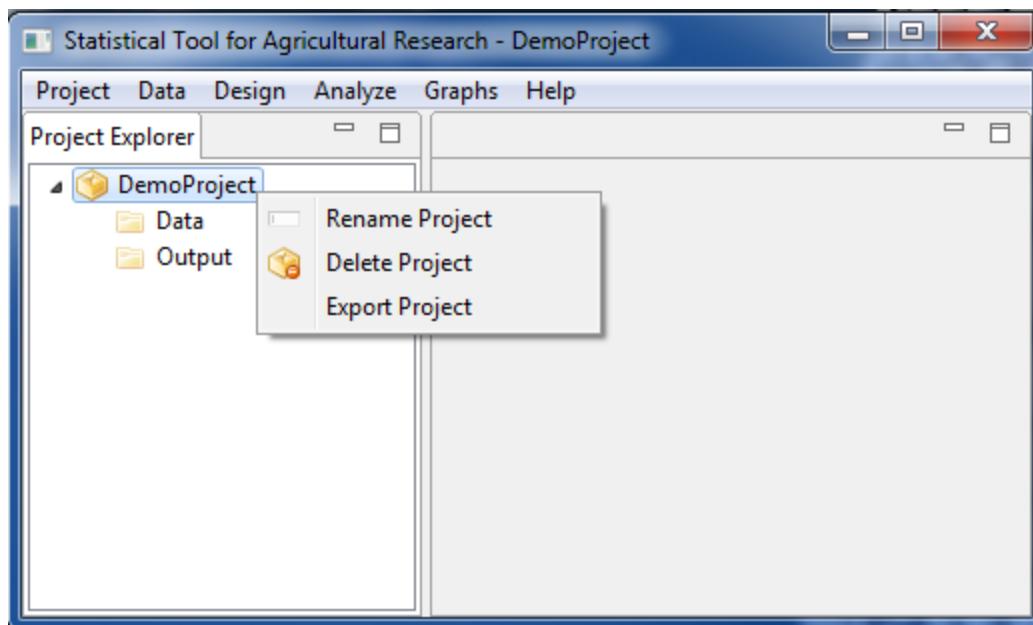


For the example, select the project named *DemoProject*. Click the **OK** button. *Demo Project* should now be the active project in the **Project Explorer**.

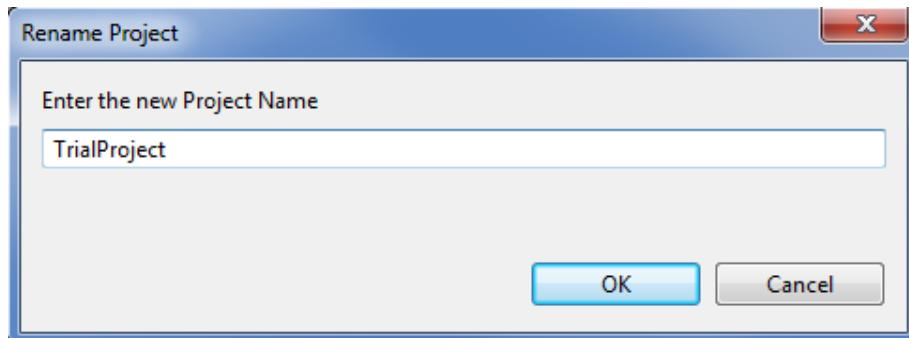
Renaming Project

The steps for renaming the active project are listed below.

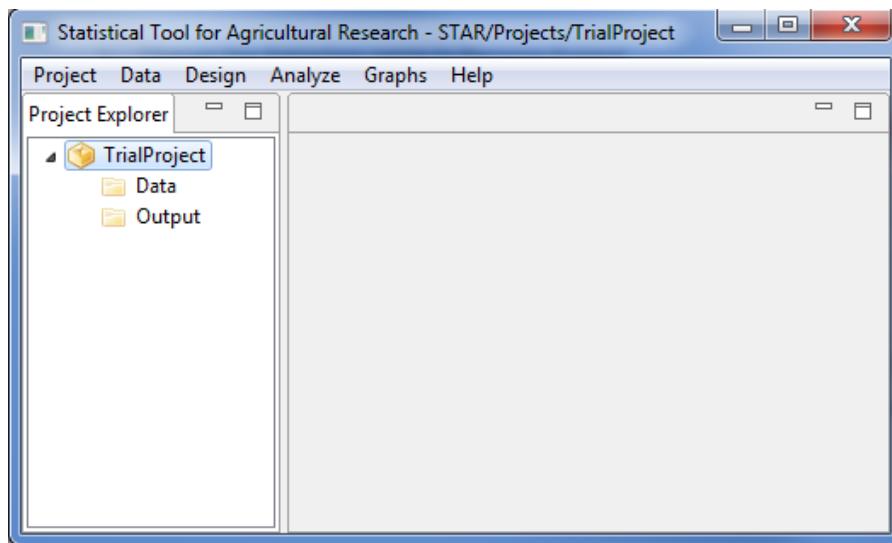
- Right-click on the active project, say *DemoProject*, then select **Rename Project** or select **Project | Rename Project** from the main menu.



- The **Rename Project** dialog box will appear. In the **New Project Name** field, specify a new project name, say *Trial Project*. Click the **OK** button.



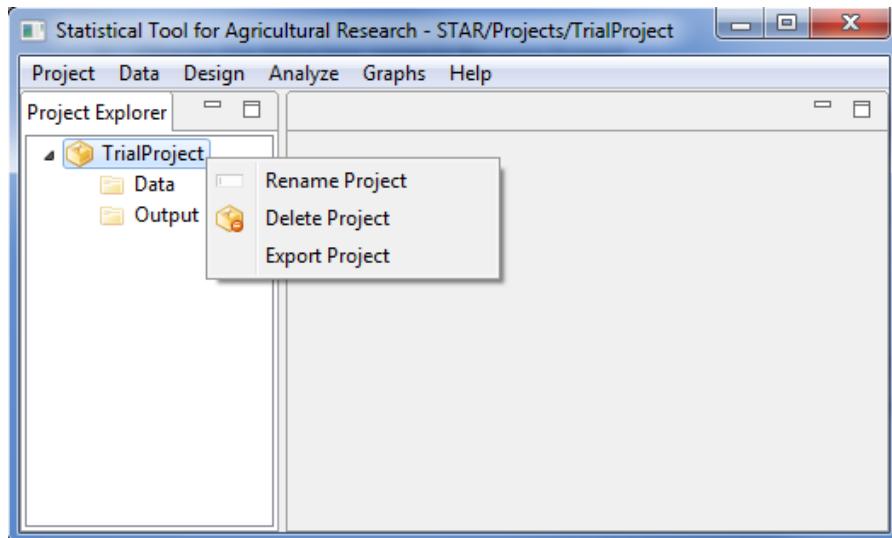
- The project should be displayed in the **Project Explorer** with the new name.



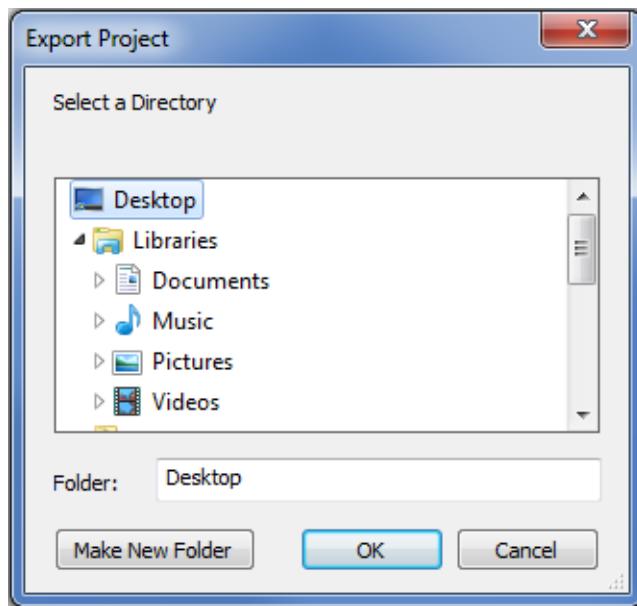
Exporting Project

If the user wants to save the active project to a different directory, the export project feature can be used. The steps are listed below.

- Right-click on the active project, say *TrialProject*, then select **Export Project**.



- The **Export Project** dialog box will appear. Select the directory where the active project will be saved. For the example, choose *Desktop*. Click the **OK** button.

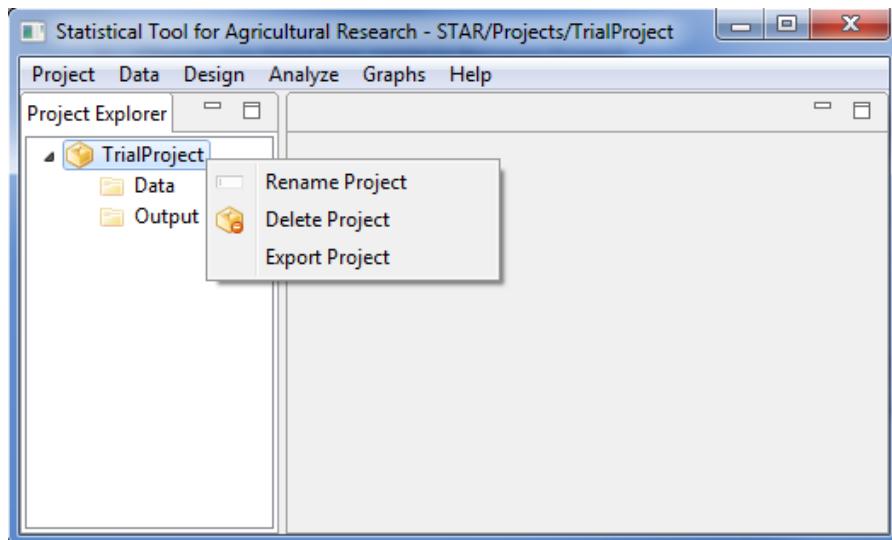


- The *TrialProject* folder is now on the desktop.

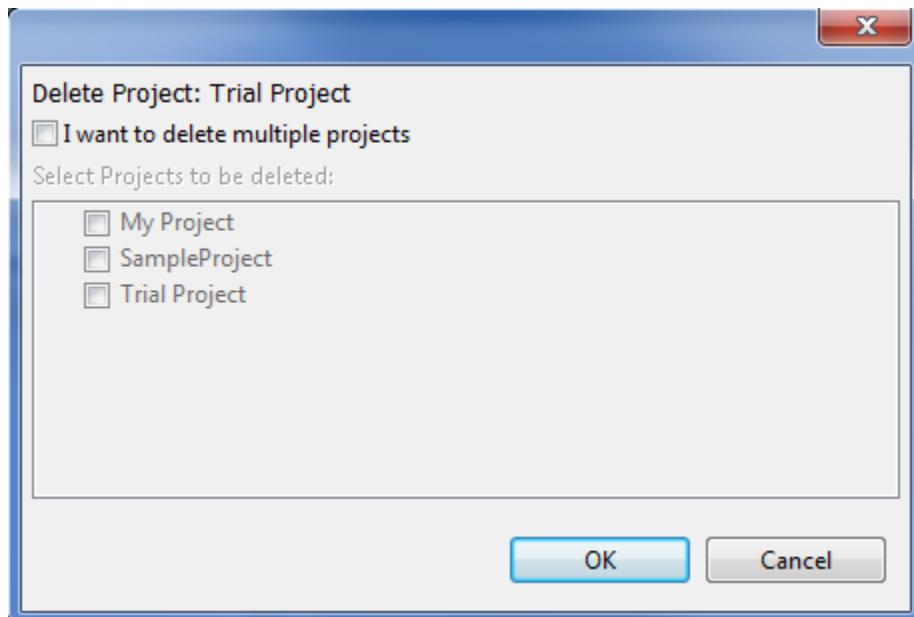
Deleting Project

The steps for deleting an existing project inside the STAR workspace are listed below.

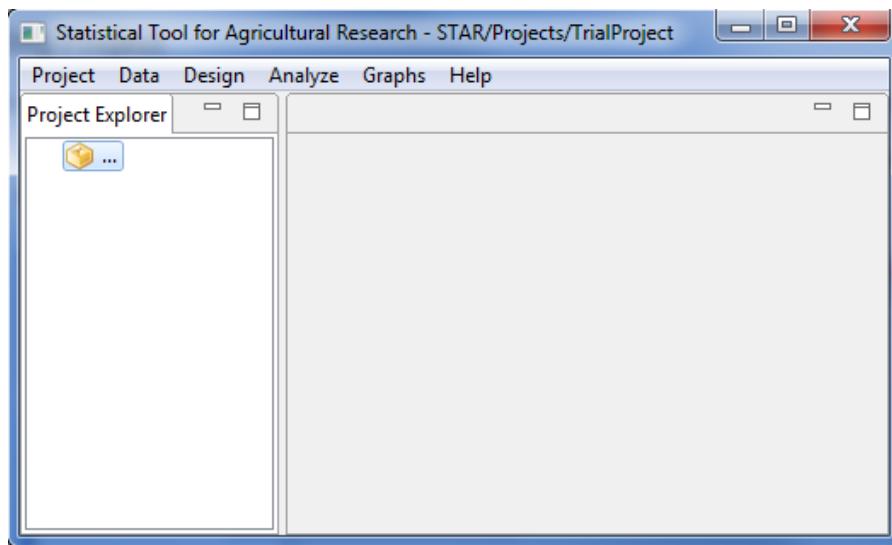
- Right-click on the active project, say *TrialProject* then select **Delete Project** or select **Project | Delete Project** from the main menu.



- Another dialog will be displayed. Click on the check box labeled **I want to delete multiple projects** if multiple projects are to be deleted. Then select from the list the projects to be deleted. For the example, delete only *TrialProject* by clicking the **OK** button.



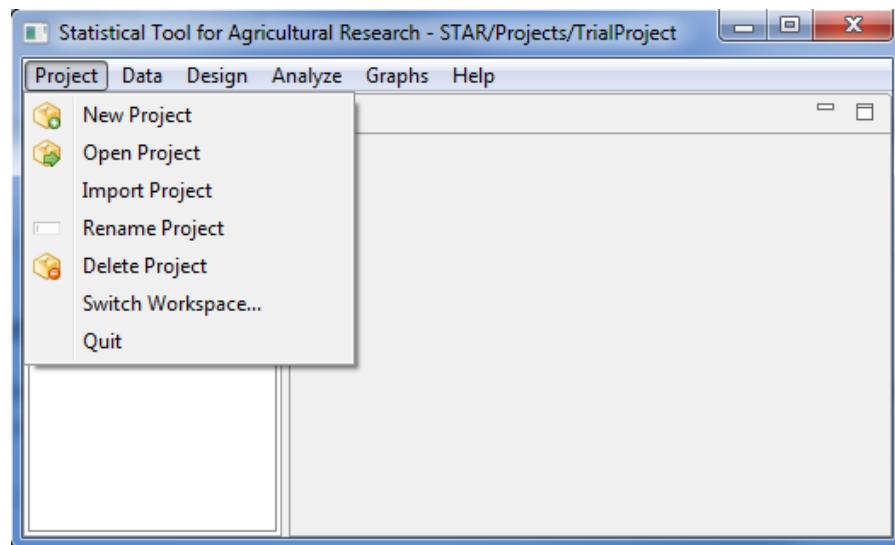
- Since *TrialProject* is now deleted from the STAR workspace, there is no active project displayed in the Project Explorer as shown below.



Importing Project

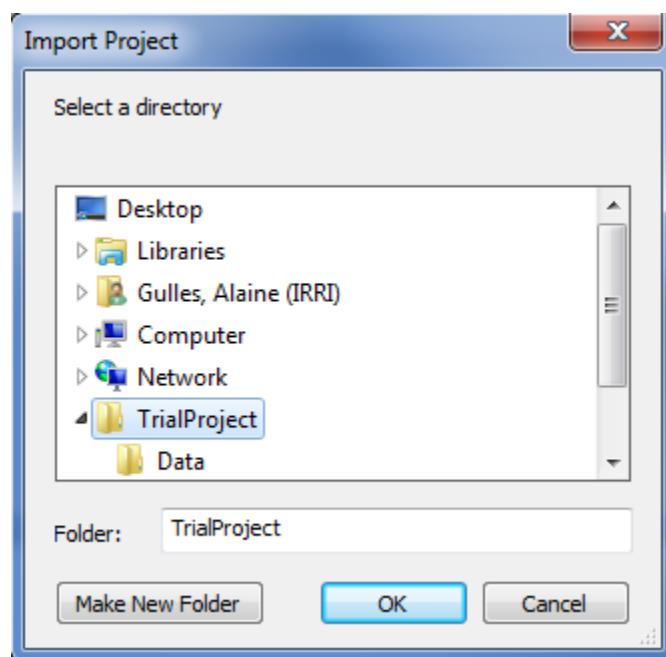
Projects, as long as it is a valid STAR project, can be saved inside the STAR workspace by using the import project feature. The steps are listed below.

- Click **Project | Import Project** from the main menu.



- Select the project folder that will be imported.

For the example, select *TrialProject* folder on the desktop. Click the **OK** button.



- The **TrialProject** is now the active project in the **Project Explorer**.

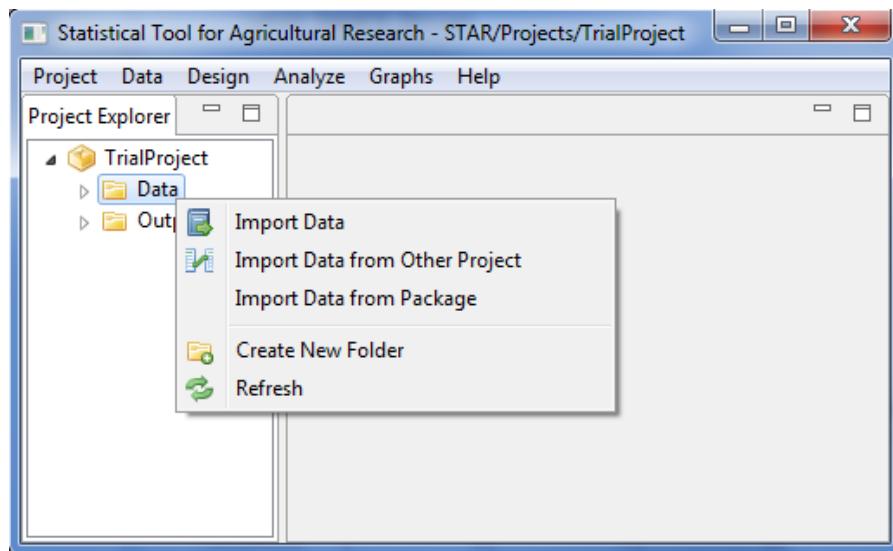
Importing Data from External Source

There are three file formats that STAR accepts to import from external source namely:

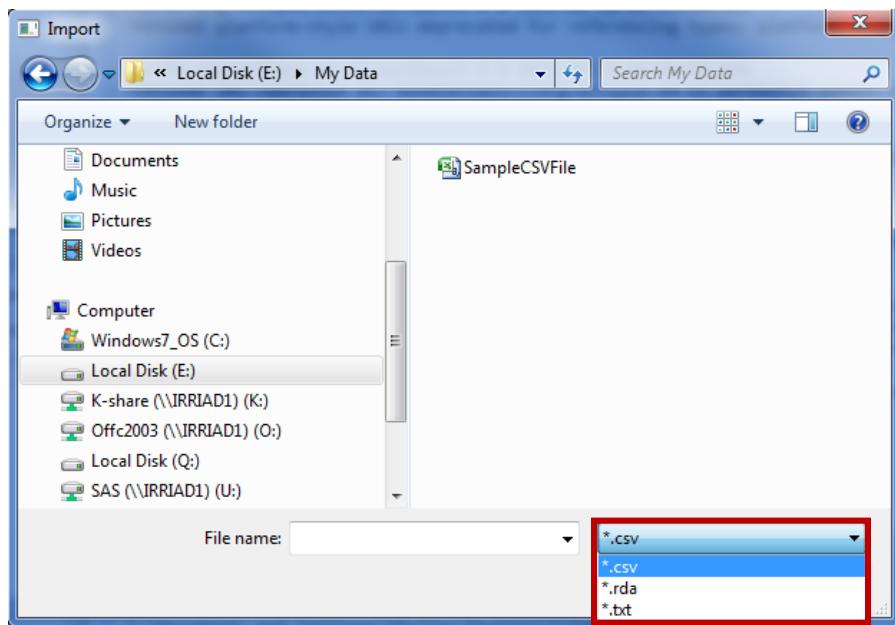
- i. R datasets. These are datasets that were saved in the R environment and have an extension .rda, .Rda, .RDA, or .Rdata;
- ii. text files. These are datasets with extension .txt whose file delimiter can be a tab, space, comma or a semi-colon; and
- iii. comma-separated value files. These are datasets with extension .csv.

The general steps for importing data from external source are listed below:

- On the **Project Explorer** panel, right-click the *Data* folder of the project named *TrialProject*. A pop-up menu will appear. Choose **Import Data**.

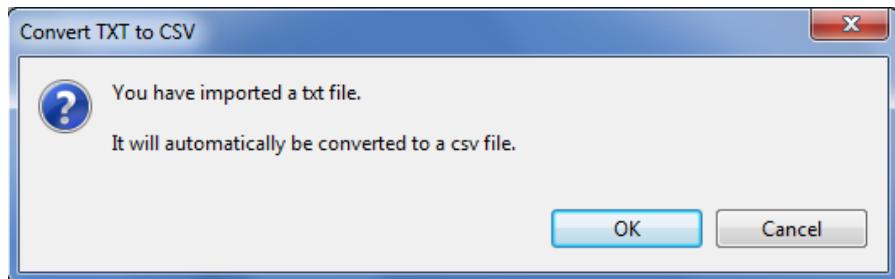


- The **Import** dialog box will appear. Choose the appropriate file type to be imported in the File of Types drop-down box. The default file type is ***.csv**. Go to the directory where the file to be imported is located. Select the file or type the file name on the **File name** text box.

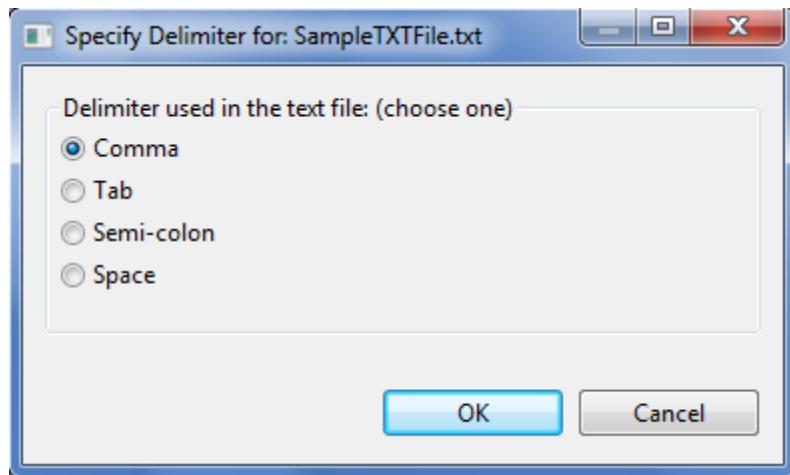


- Click the **Open** button.

If the ***.txt** or the ***.rda** file type is chosen, a message dialog box will appear indicating that the text or r data files will be automatically converted to a csv file. Click the **OK** button.



Further, if the ***.txt** file type is chosen, the **Specify Delimiter** dialog box will appear. Choose the appropriate delimiter of the text file to be open then click the **OK** button.

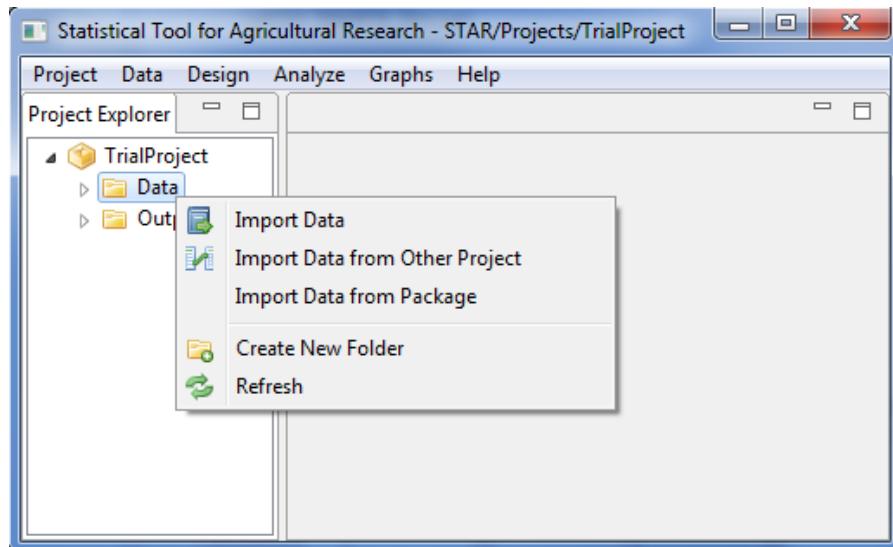


- The newly imported data will now be displayed inside the *Data* folder of the active project in the **Project Explorer**. To view the data in a spreadsheet format, double-click the icon of the data file.

Importing Data from Other Project

The steps for importing data from other project are listed below:

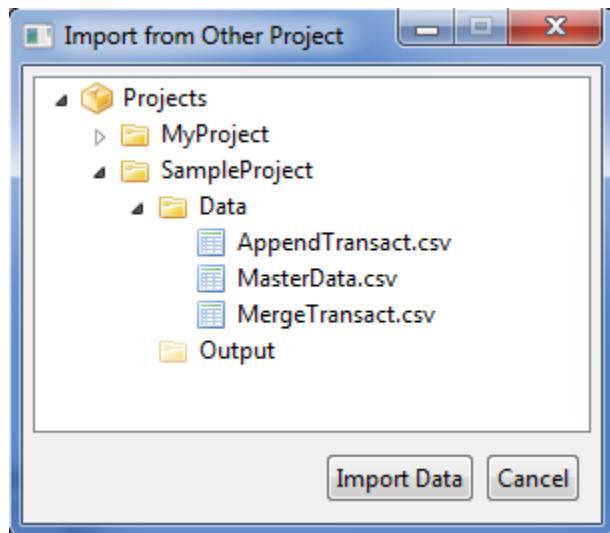
- On the **Project Explorer** panel, right-click the *Data* folder of the project named *TrialProject*. A pop-up menu will appear. Choose **Import Data from Other Project**.



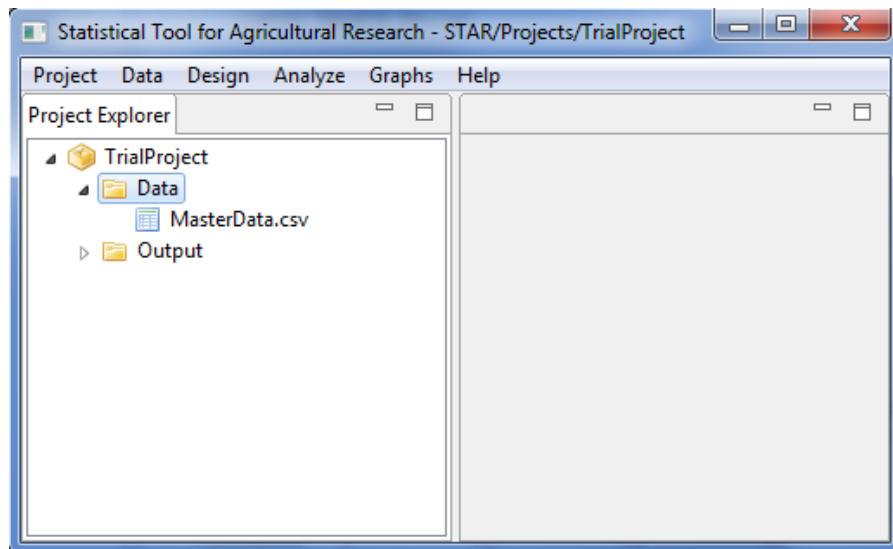
- The **Import from Other Project** dialog box will appear. This dialog box contains all existing projects in STAR. Select one or several project folders, one or several

folders within projects or one or several data files. Choose the data files to be imported.

For the example, choose *MasterData.csv* from *Projects | SampleProject | Data*.



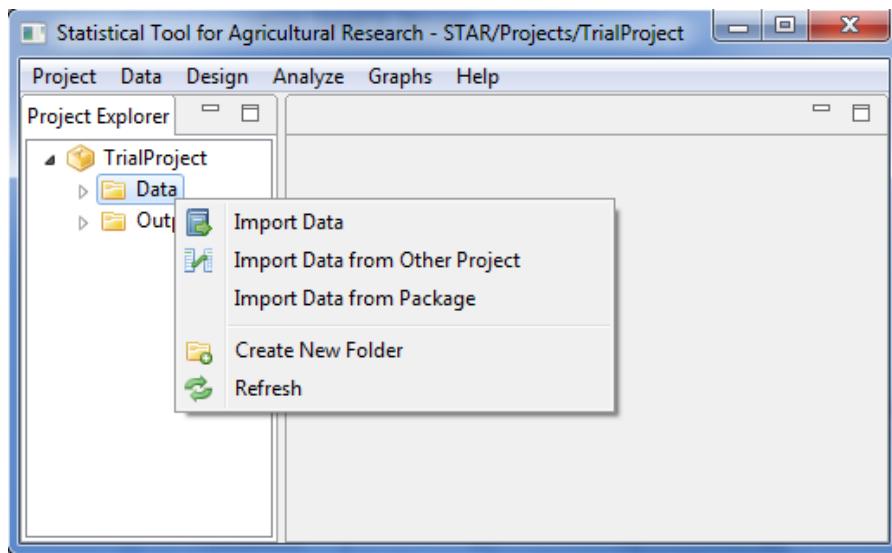
- Click the **Import Data** button. The new imported data file will now be displayed inside the *Data* folder of the active project in the **Project Explorer** panel. To view the data in a spreadsheet format, double-click the icon of the data file.



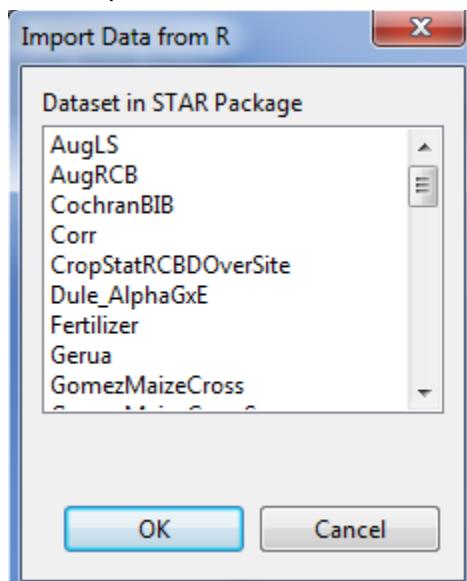
Importing Data from Package

Datasets in STAR package of R can be imported. The steps to import data from R package are listed below:

- On the **Project Explorer** panel, right-click the *Data* folder of the project named *TrialProject*. A pop-up menu will appear. Choose **Import Data from Package**.

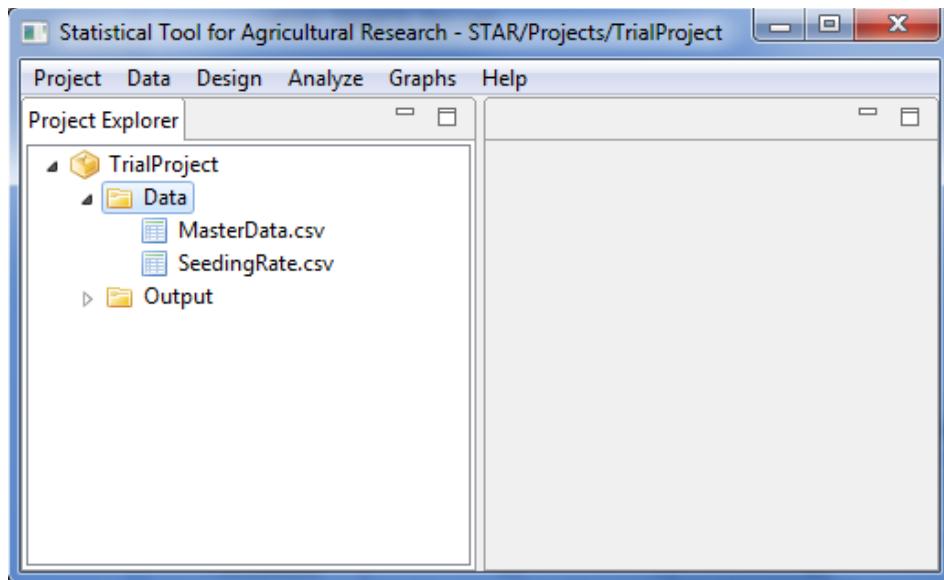


- The **Import Data from R** dialog box will appear. This dialog box contains all existing data file in STAR R package. Select one or several data files. Choose the data files to be imported.



For the example, choose *SeedingRate*.

- Click the **OK** button. The new imported data file will now be displayed inside the *Data* folder of the active project in the **Project Explorer** panel. To view the data in a spreadsheet format, double-click the icon of the data file.



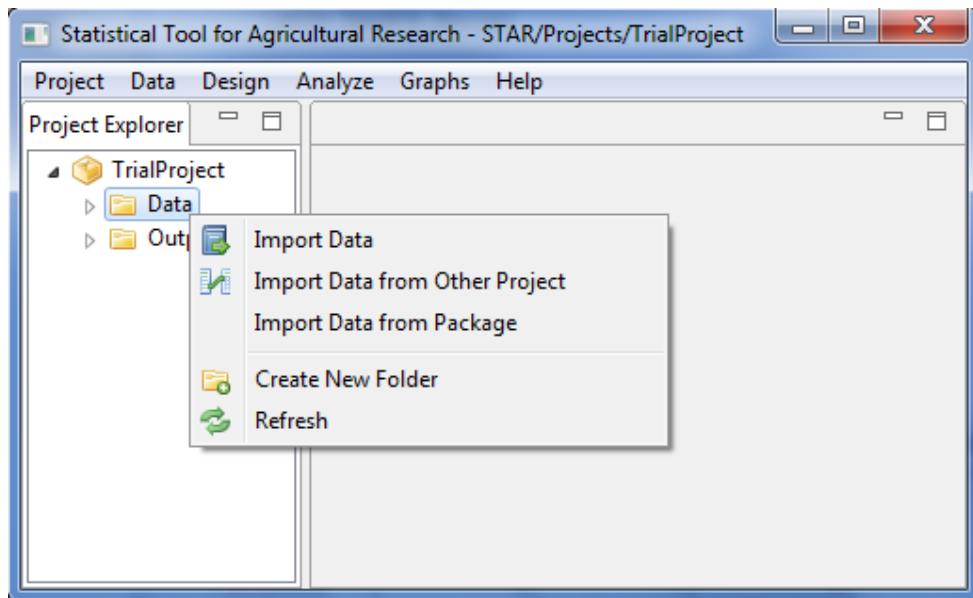
Creating New Folder

To organize the data files imported in a project or results of an analysis, user may want to create a sub folder inside the *Data* folder or inside the *Output* folder.

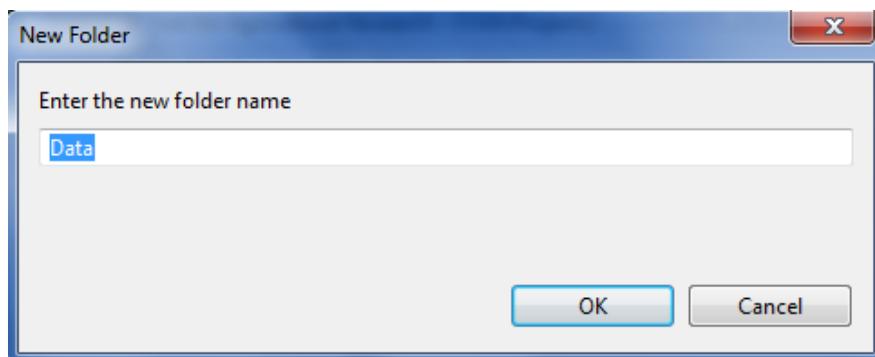
The steps in creating a sub folder are listed below:

- On the **Project Explorer** panel, right-click the folder where you want to create the sub folder. A pop-up menu will appear. Choose **Create New Folder**.

For the example, right-click the *Data* folder of the project named *TrialProject*.



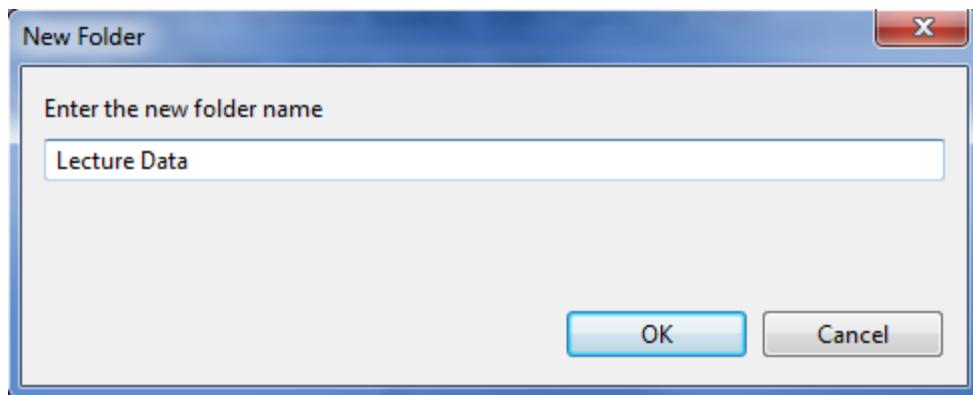
- The **New Folder** dialog box will appear. In the **Enter the new folder name** field, specify the name for the new folder. If the sub folder will be created within the *Data* folder, the default folder name is *Data*. If the sub folder will be created within the *Output* folder, the default folder name is *Output*.



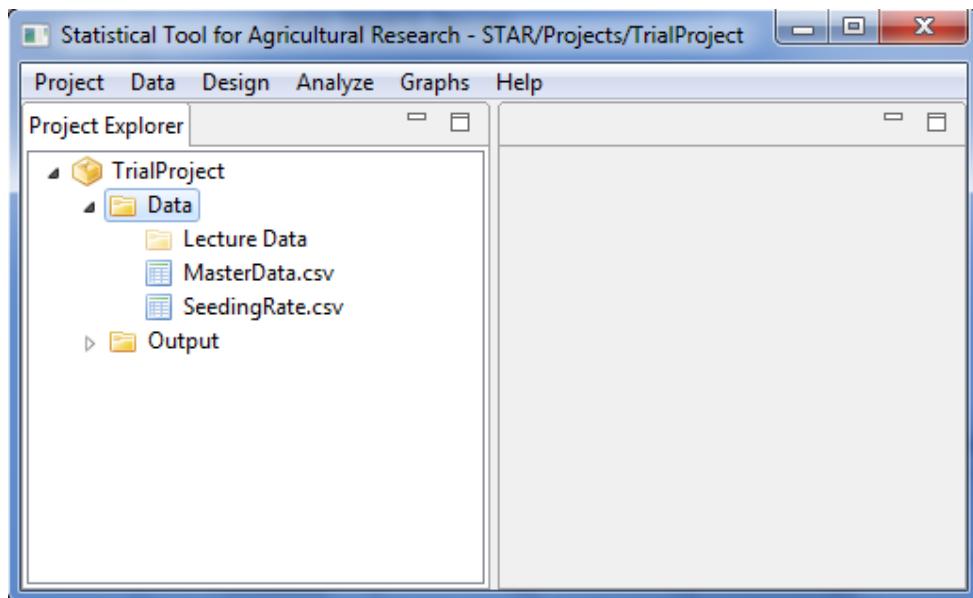
In naming a folder, the following rules apply:

- The name must start with a letter. The remaining characters can be any letter, any digit, a period, underscore, blank or dash.
- The last character of the name should not be a period, underscore, blank or dash.
- The length of the name should not be less than 4 characters.
- The name must be different from the existing folder names inside a parent folder.
- The name is not case sensitive.

For the example, type *Lecture Data* as the new folder name.



- A new sub-folder inside *Data* folder is now created and can be seen in the Project Explorer.



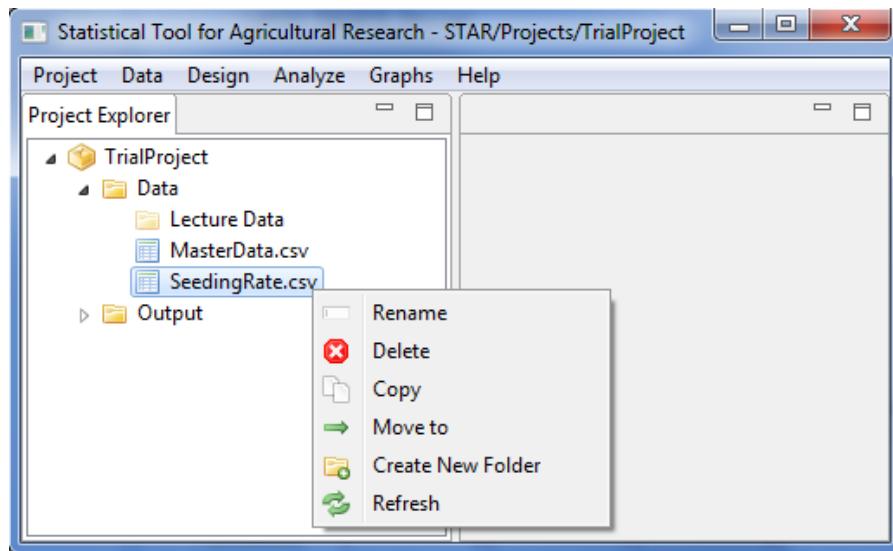
The user can create several sub folders within the *Data* folder or *Output* folder. If the sub folders are created within the *Data* folder, user can import data set to the sub folder, rename, copy, delete and move the sub folder.

Copying File/Folder

The steps to copy data files to other location are listed below:

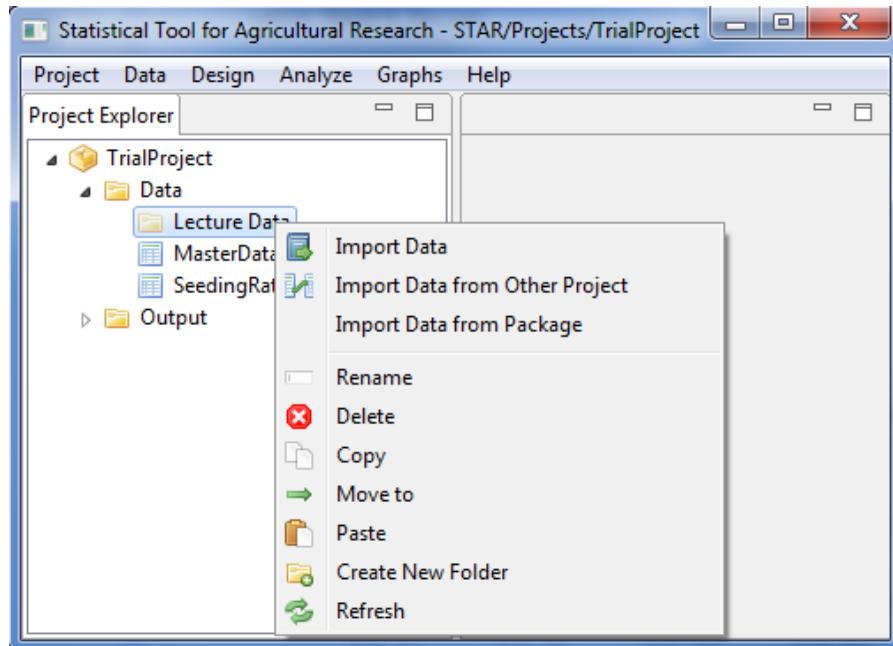
- On the **Project Explorer** panel, right-click the file you want to copy. A pop-up menu will appear. Choose **Copy**.

For the example, right-click *SeedingRate.csv* inside the *Data* folder of the project named *TrialProject*.

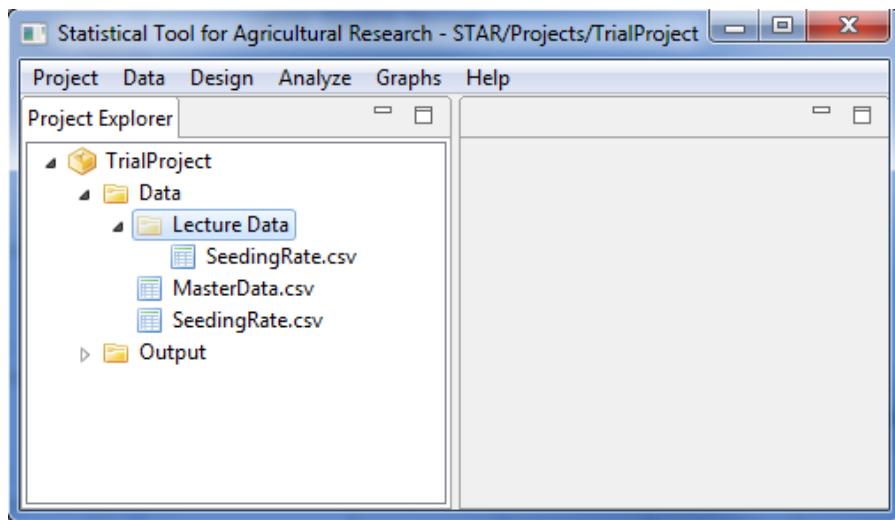


- Go to the desired destination folder, which can be within or outside STAR workspace, then right-click then choose **Paste**.

For the example, right-click on the *Lecture Data* folder then chose **Paste**.



- The file is now copied inside the destination folder.



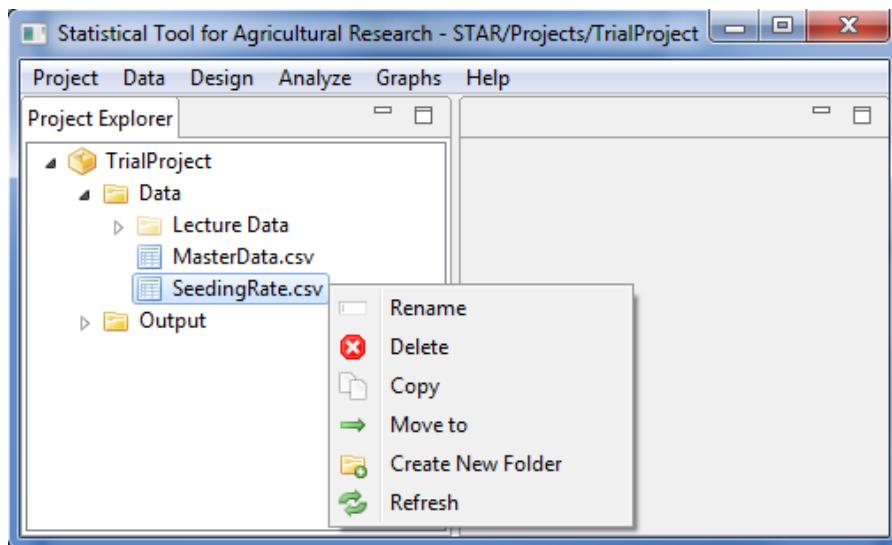
These steps also apply when copying the sub- folders within the *Data* and *Output* folders.

Renaming File/Folder

The steps to rename files are listed below:

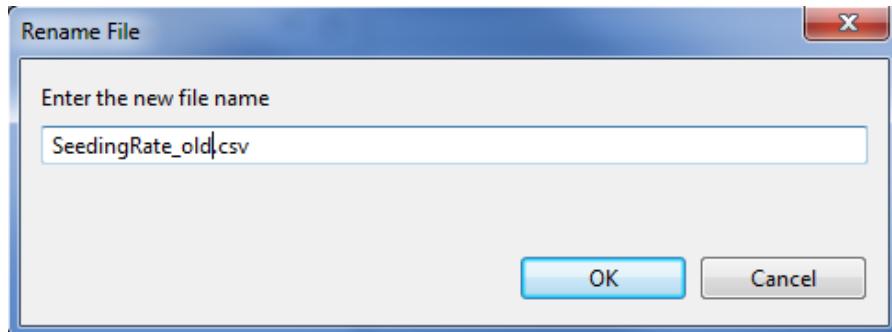
- On the **Project Explorer** panel, right click the file to rename. A pop-up menu will appear. Choose **Rename**.

For the example, right-click *SeedingRate.csv* inside the *Data* folder of the project named *TrialProject*.

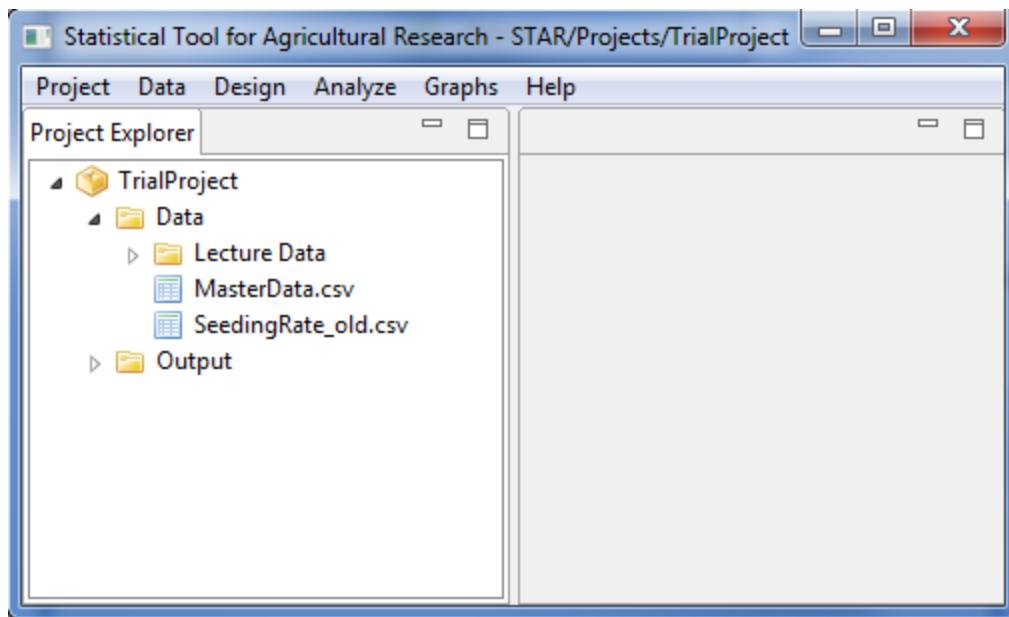


- The **Rename File** dialog box will appear. In the **Enter the new file name** field, specify the new filename.

For the example, set the new filename to *SeedingRate_old.csv*. The completed dialog box should appear as shown below:



- Click the **OK** button. The renamed data file should appear in the **Project Explorer** panel.



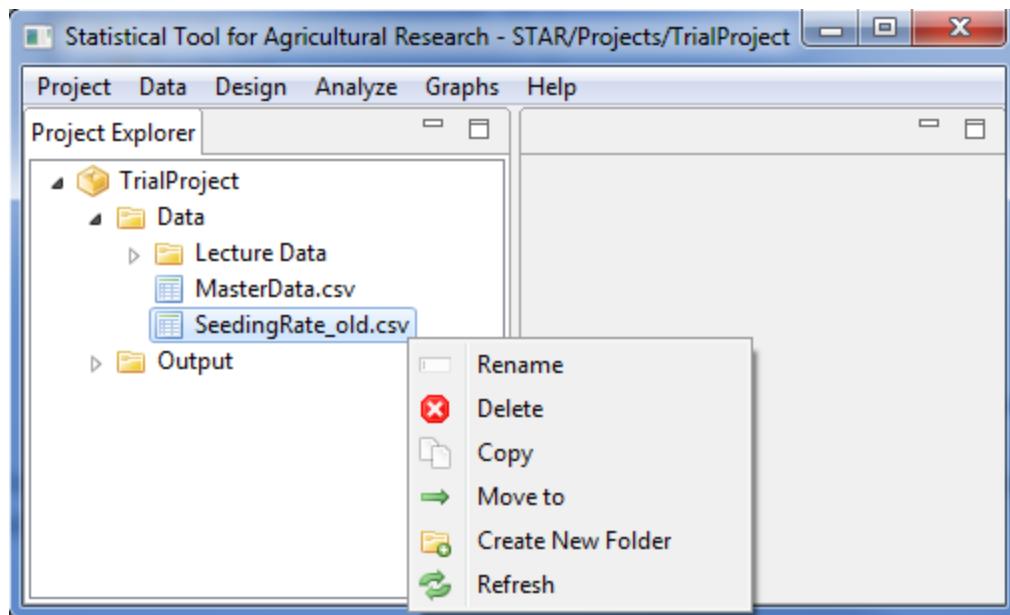
These steps also apply when renaming the sub- folders within the *Data* and *Output* folders.

Moving File/Folder

The steps to move data files to other location in the STAR workspace are listed below:

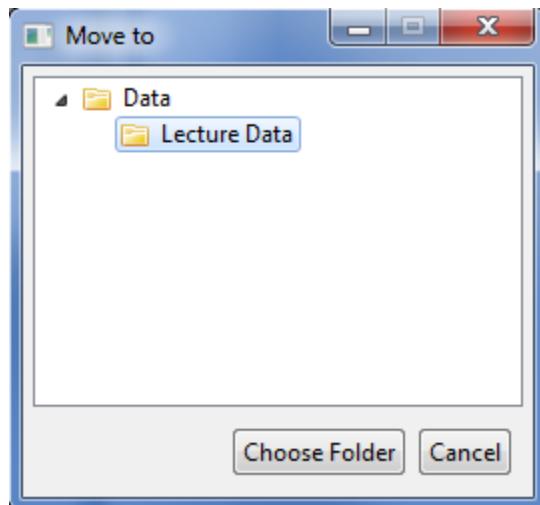
- On the **Project Explorer** panel, right click the file to you want to move. A pop-up menu will appear. Choose **Move to**.

For the example, right-click *SeedingRate_old.csv* inside the *Data* folder of the project named *TrialProject*.

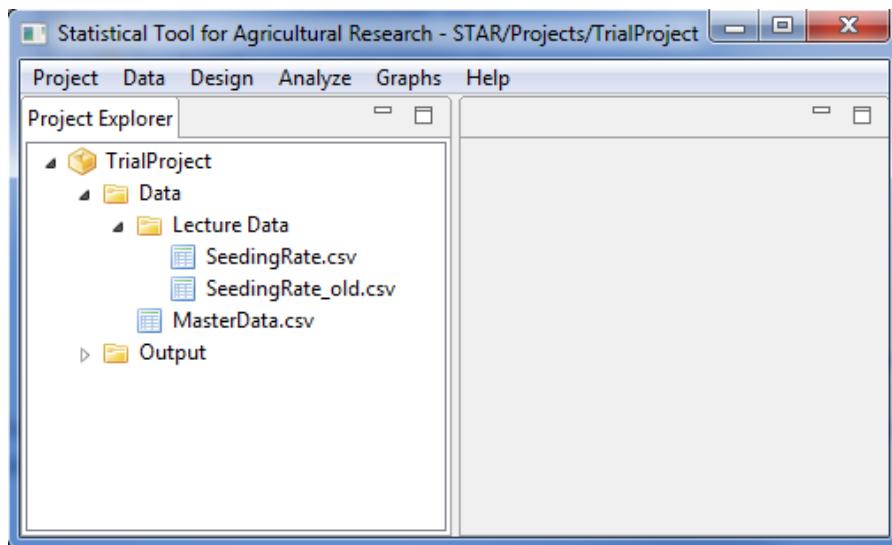


- The **Move to** dialog box will appear. Select the destination folder.

For the example, choose the *Lecture Data* folder inside the *Data* folder.



- Click the **Choose Folder** button. The file is now moved to the destination folder.



These steps also apply when moving the sub-folders within the *Data* and *Output* folder.

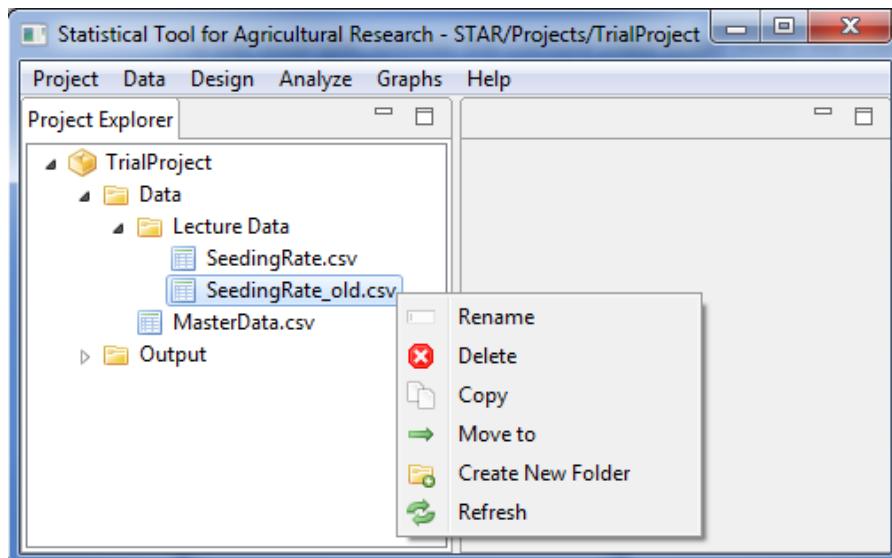
Deleting File/Folder

Files to be deleted must be closed before following the steps listed below. However, if a user wants to delete a folder, all items inside the folder to be deleted must be closed before deleting the folder. Note that the default folders created (*Data* folder and *Output* folder) cannot be deleted.

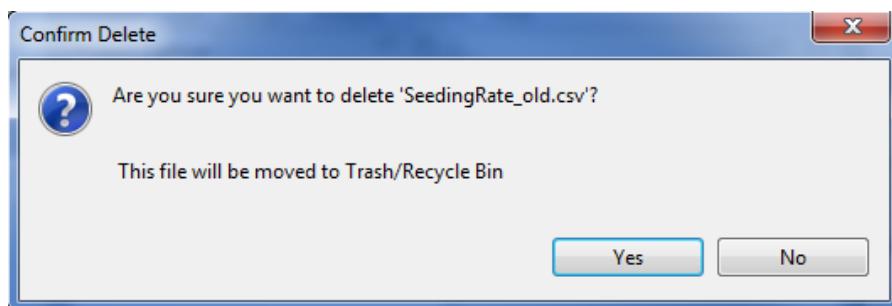
The steps to delete data files are listed below:

- On the **Project Explorer** panel, right click the file to you want to delete. A pop-up menu will appear. Choose **Delete**.

For the example, right-click *SeedingRate_old.csv* inside the *Lecture Data* of the *Data* folder of the project named *TrialProject*.



- The **Confirm Delete** message box will appear.



- Click the **Yes** button to delete.

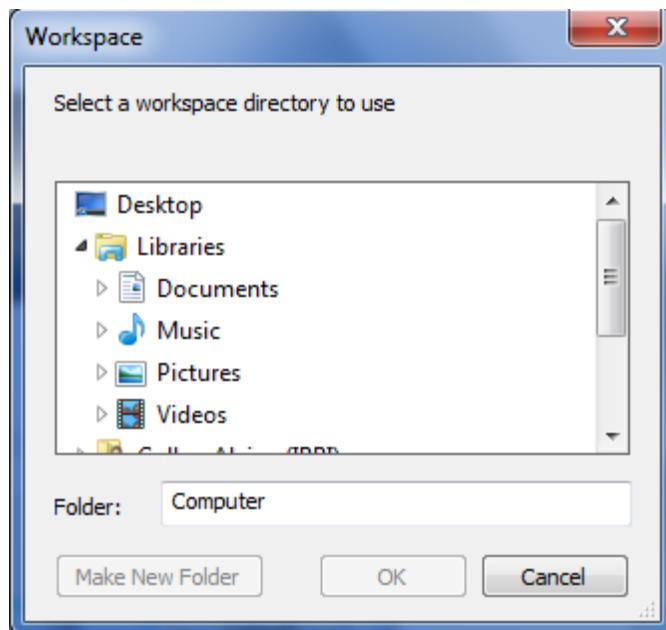
These steps also apply when deleting the sub-folders within the *Data* and *Output* folder.

Switching Workspace

Workspace is a location where projects, data and output during the session will be saved.

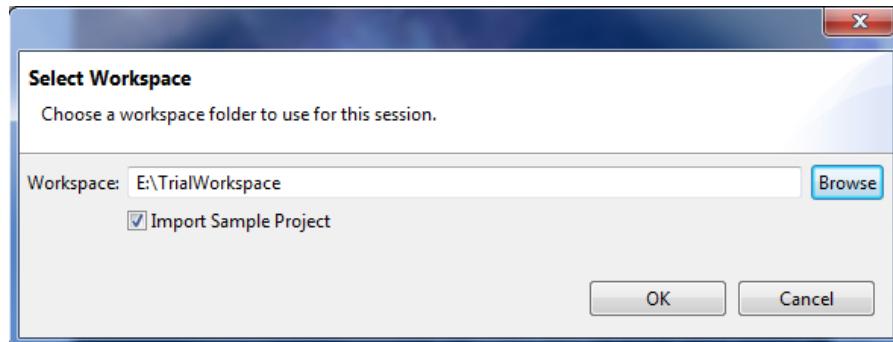
The steps to switch to a different workspace are listed below:

- Click **Project | Switch Workspace....**. The **Select Workspace** dialog box will appear. Click the **Browse** button to change the current workspace. The **Workspace** dialog box will appear.

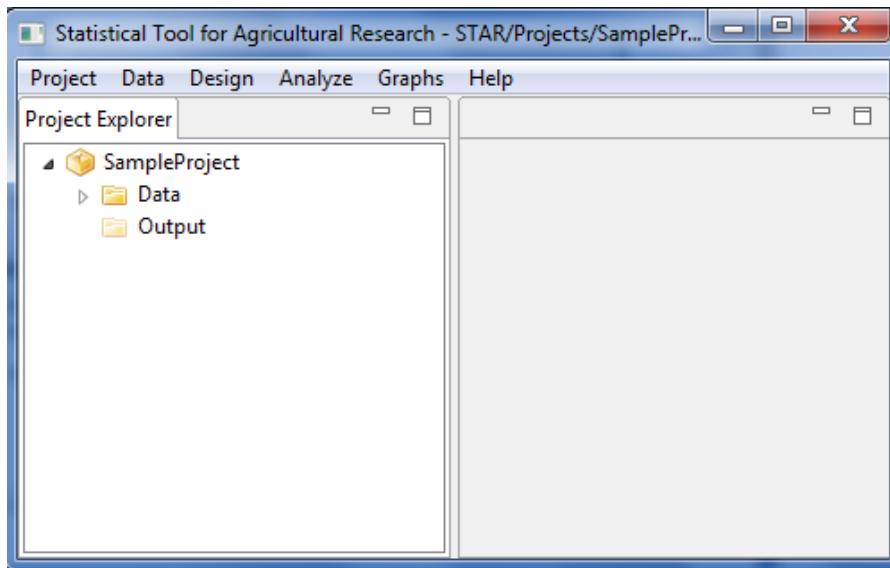


- Choose the new workspace or create a new workspace by clicking the **Make New Folder** button. Then click the **OK** button.

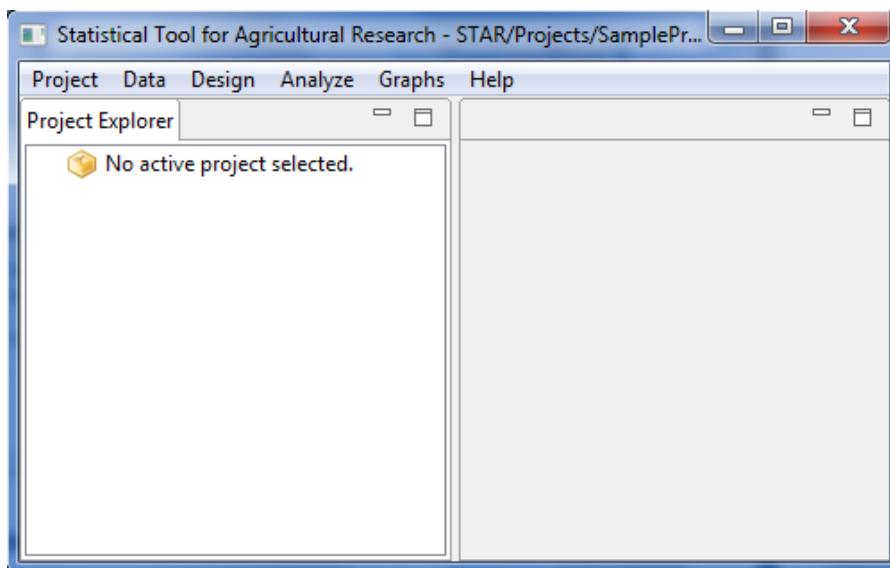
For the example, the completed dialog box should appear as shown below:



If the **Import Sample Project** is selected, the main window appears with *SampleProject* as the active project.



Otherwise, the main window appears with no active project.



Data Handling and Manipulation

Representing Missing Value(s)

In STAR, missing values can be represented by using any of the following: *NA*, space, period or blank. Using other symbol, say asterisk (*), to represent missing values will affect how STAR identifies the attributes of the columns of the dataset. A column with numeric values but contains asterisks will be classified as a factor column. This misclassification may cause problems later when the data is subjected to manipulation or analysis.

Insert Row(s)

The steps to insert row(s) are listed below:

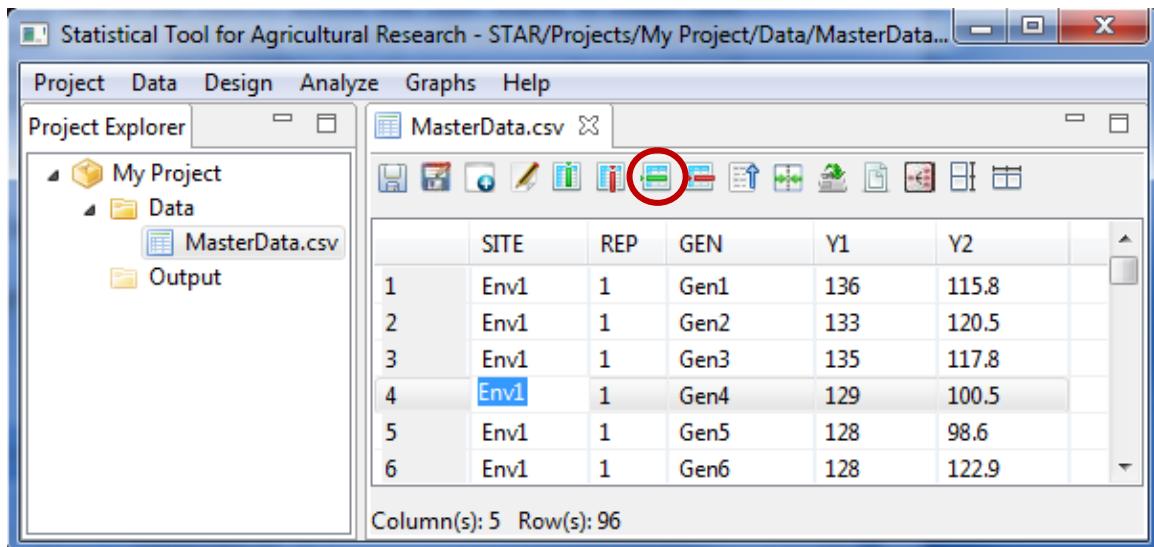
- On the **Project Explorer** panel, double-click the file where new row(s) will be inserted to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

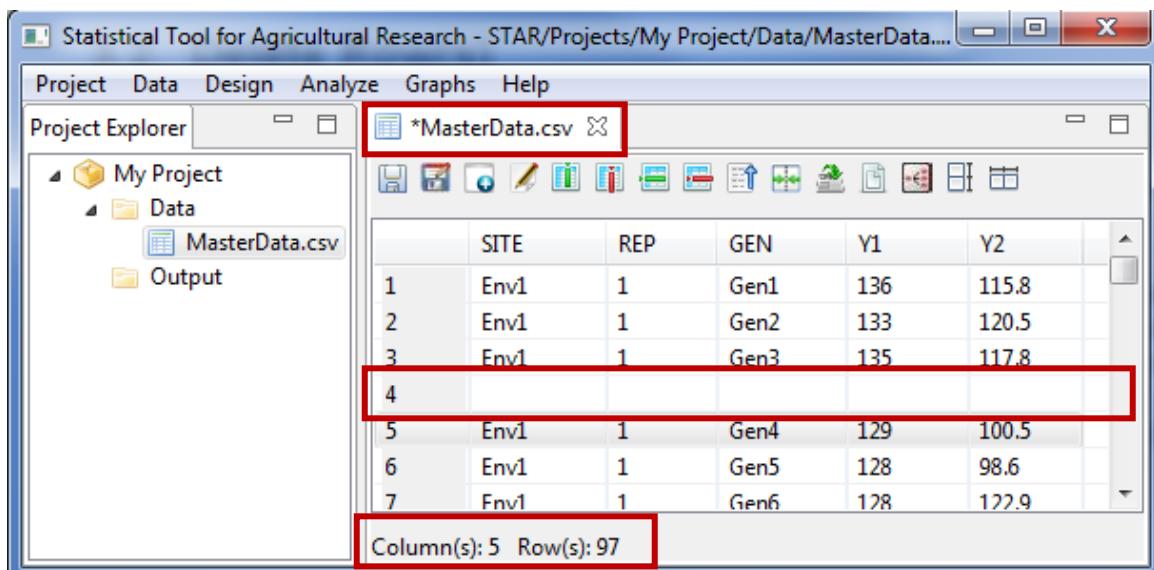
- Select any row(s) or any cell in the row(s) where new row(s) are to be inserted above it. To select several cells/rows, click any cell/row, then hold the **Ctrl** key and click on another cell/row. The number of selected rows is the number of row that will be inserted.

For the example, select the 4th row.

- Click **Data | Insert Row...** from the menu bar of the main window or click on the  Insert Row icon in the **Data Viewer** toolbar.



- The data now contains the newly inserted row.



If there is no row selected before executing the Insert Row function, the user will be prompted if the user wants to insert a row after the last row of the data or not.

Delete Row(s)

The steps to delete row(s) are listed below:

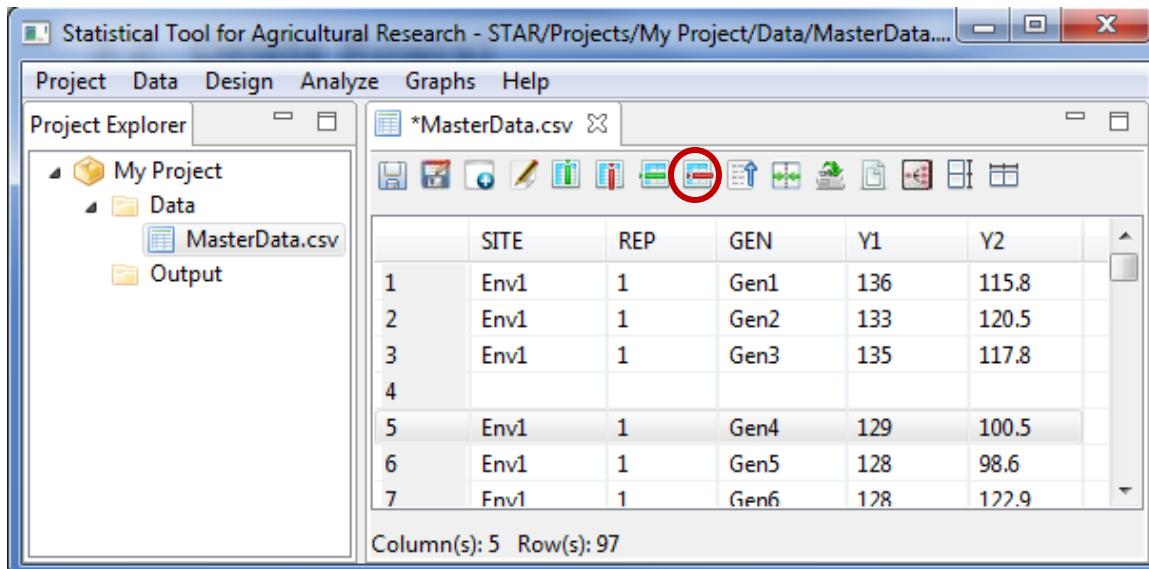
- On the **Project Explorer** panel, double-click the file where rows will be deleted to open it and view it in the Data Viewer tab.

For the example, use the modified and unsaved *MasterData.csv* in the project named *MyProject* created from the previous section that is displayed in the Data Viewer.

- Select any row(s) or any cell in the row(s) that will be deleted. To select several cells/rows, click any cell/row, then hold the Ctrl key and click on another cell/row. The number of selected rows is the number of row that will be deleted.

For the example, select the 4th row.

- Click **Data | Delete Row** from the menu bar of the main window or click on the  Delete Row icon in the Data Viewer tool bar.



- The **Delete Row** dialog box will appear. Confirm that the row deletion is desired by clicking the **Yes** button.
- The changes made should be reflected in the Data Viewer tab.

For the example, close the modified *MasterData.csv* and do not save the changes.

Inserting Column(s)

The steps to delete row(s) are listed below:

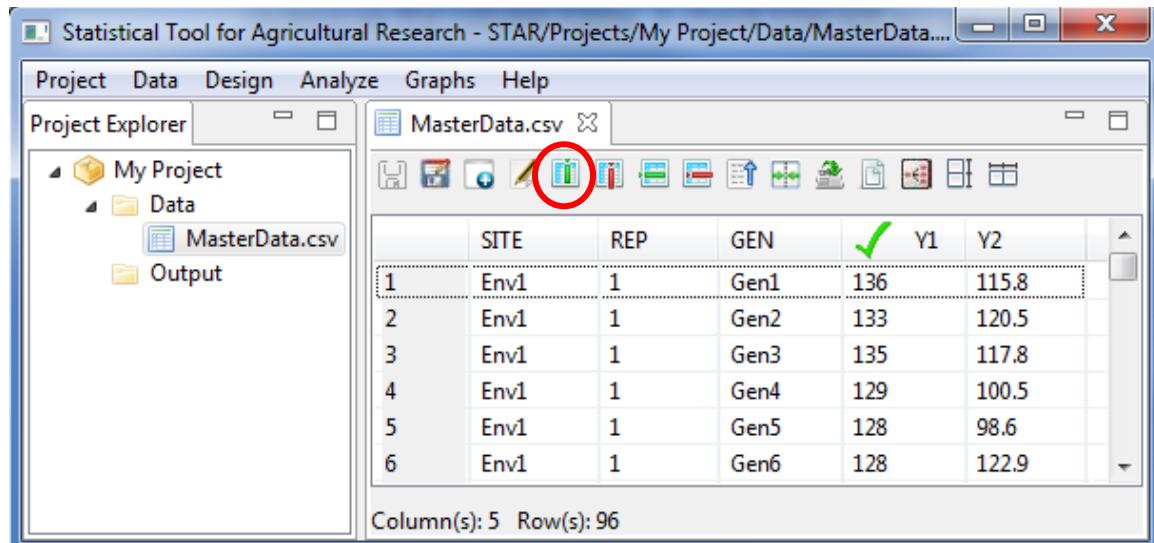
- On the **Project Explorer** panel, double-click the file where new columns will be inserted to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

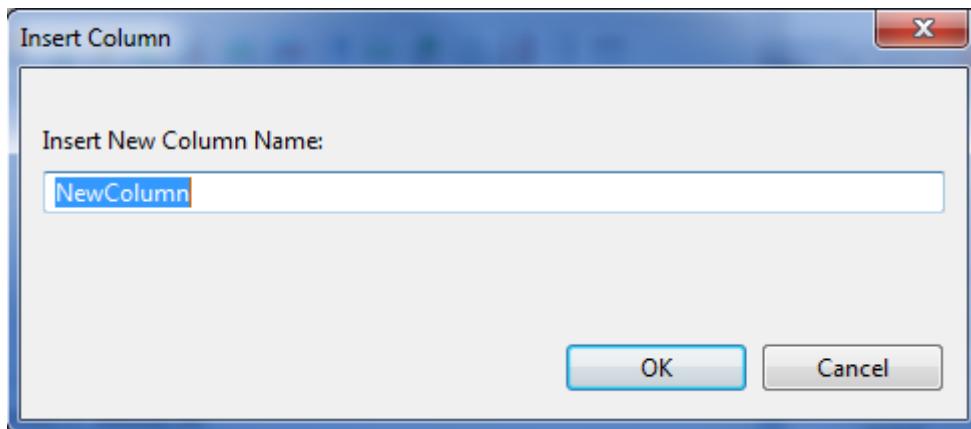
- Select a column where the new column will be inserted before it by clicking on the column name. A check icon will appear on the column header.

For the example, select the column named *Y1*.

- Click **Data | Insert Column...** from the menu bar of the main window or click on the  Insert Column icon in the Data Viewer tool bar.



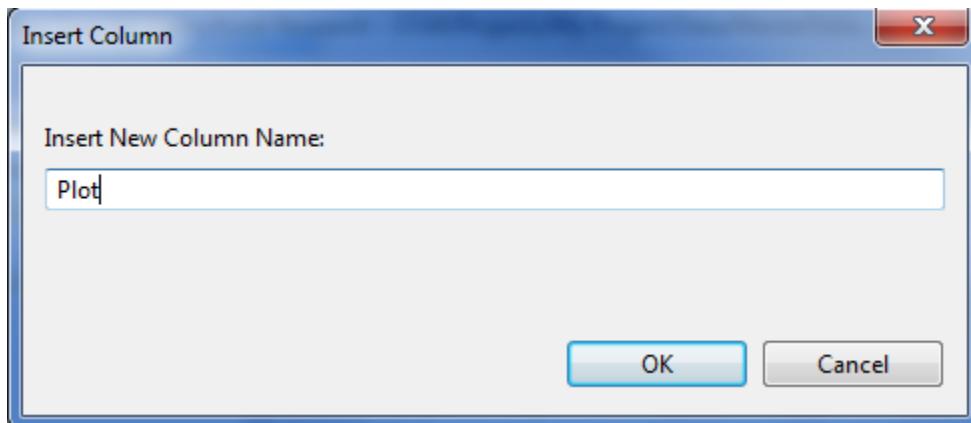
- The **Insert Column** dialog box will appear. In the **Insert New Column Name** field, specify a new name for the inserted column. The default new column name is *NewColumn*.



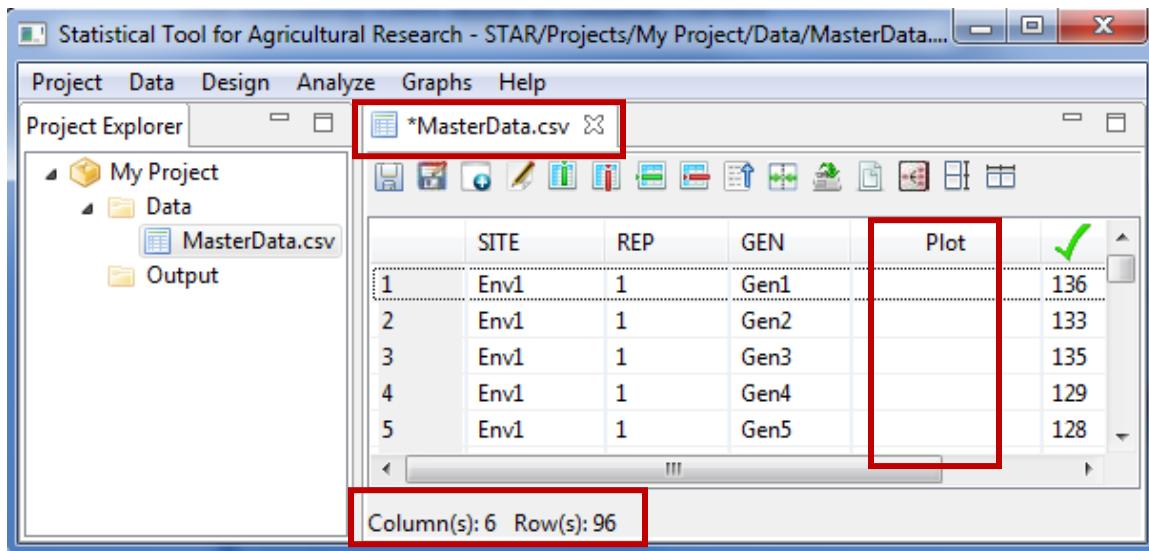
In naming the column, the following rules apply:

- The name must start with a letter while the remaining characters can be any letter, any digit, a period or underscore.
- The name must be different from the existing column names in the data.
- The name is case sensitive.

For the example, type *Plot*. The completed dialog box will appear as shown below:



- Click the **OK** button.



- The data now contains the newly inserted column named *Plot*.

Deleting Column(s)

The steps to delete row(s) are listed below:

- On the **Project Explorer** panel, double-click the file where columns will be deleted to open it and view it in the Data Viewer tab.

For the example, use the modified and unsaved *MasterData.csv* in the project named *MyProject* created from the previous section that is displayed in the Data Viewer tab.

- Select the column or one of the columns to be deleted by clicking on the column name. A check icon will appear on the column header.

For the example, select the column named *Plot*.

- Click **Data | Delete Column...** from the main window or click on the Insert Column icon  in the Data Viewer tool bar.
- Click the **OK** button.

Create New Variables

New variables can be created in the active data set by transforming existing variables or collapsing categories of two or more existing variables.

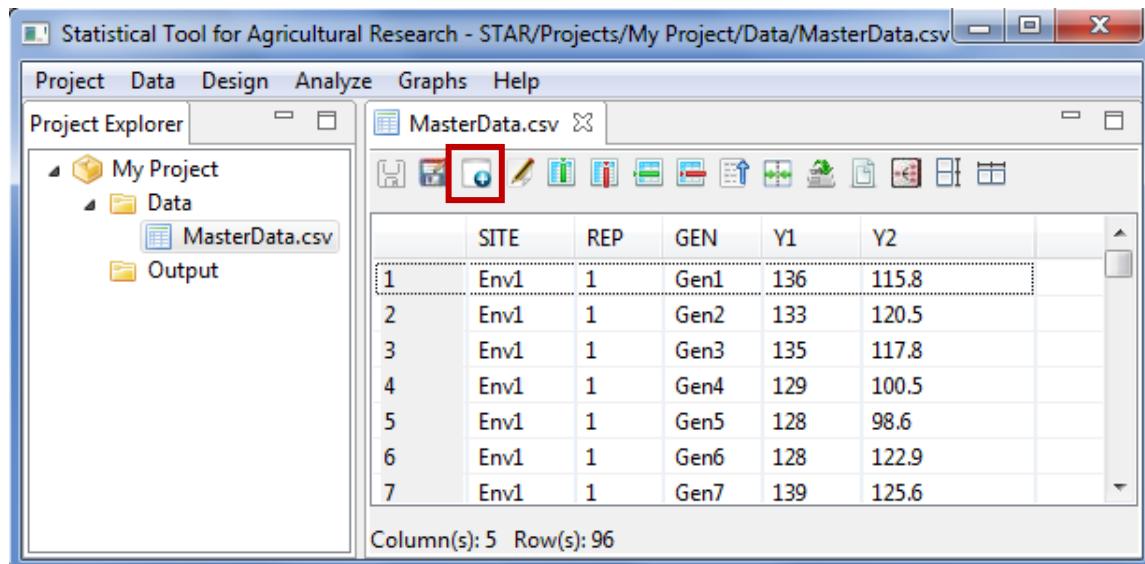
For transformations, there are six available functions, namely: logarithm, natural logarithm, square root, power, exponential and standardized. For logarithm, natural logarithm, square root and power transformation, the variables to be transformed should not contain negative values. For logarithm and natural logarithm transformation, values of the variable to be transformed will be incremented by 1 if it contains values equal to 0. For square root and power transformation, values of the variable will be incremented by 0.5 if it contains values equal to 0.

The steps to create new variables are listed below:

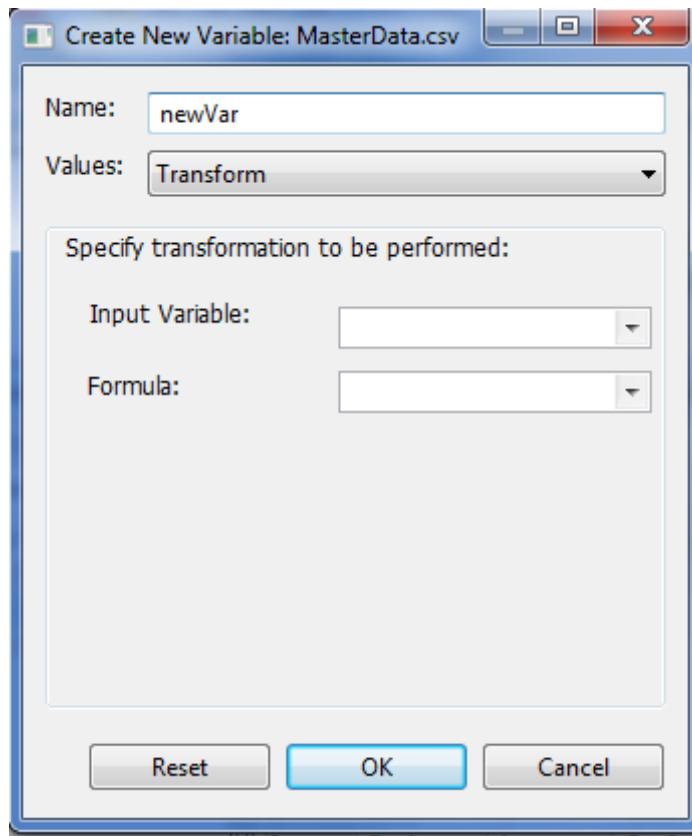
- On the **Project Explorer** panel, double-click the file to be used to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

- To create new variables, click **Data | Create New Variable** from the main window or click on the Create New Variable icon  in the Data Viewer toolbar.



The **Create New Variable** dialog box will appear.



- Specify the required fields and appropriate options.

Name

This is the name of the new variable to be created. The default variable name is *newVar*. In naming the column, the following rules apply:

- The name must start with a letter while the remaining characters can be any letter, any digit, a period or underscore.
- Each column name must be unique; duplication is not allowed.
- Column name is case sensitive

Values

There are two options to determine the values of the new variable. The user can either transform any existing numeric variables or concatenate the values of any variables in the data set. The default is *Transform*.

Input Variable

If *transformation* is selected in the **Value** drop-down box, this drop-down box will be visible. This box pertains to the variable where the transformation will be performed. All numeric variables of the active data will be displayed in the drop-down box.

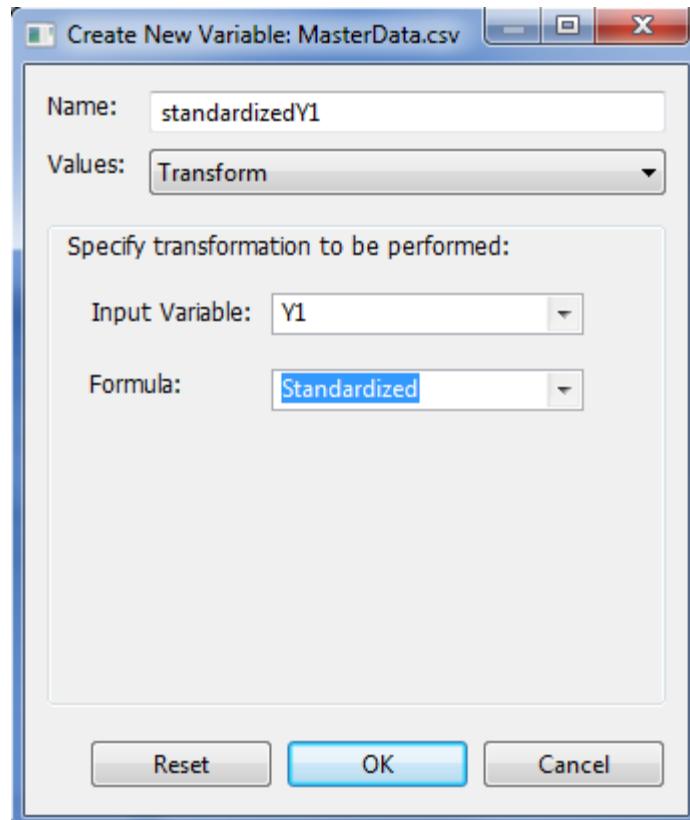
Formula

If *transformation* is selected in the **Value** drop-down box, this drop-down box will be visible. This box pertains to the formula that will be used when the value in the **Input Variable** is transformed. Transformation options available are logarithm, natural logarithm, square root, power, exponential and standardized.

Checkbox Table

If *Concatenate Columns* is selected in the **Value** drop-down box, this table will be visible. This table lets the user specify the columns to concatenate by ticking at least two column names. All columns of the active data will be displayed in the table.

For the example, suppose we want to standardize the variable Y1. The completed dialog box should appear as shown below:

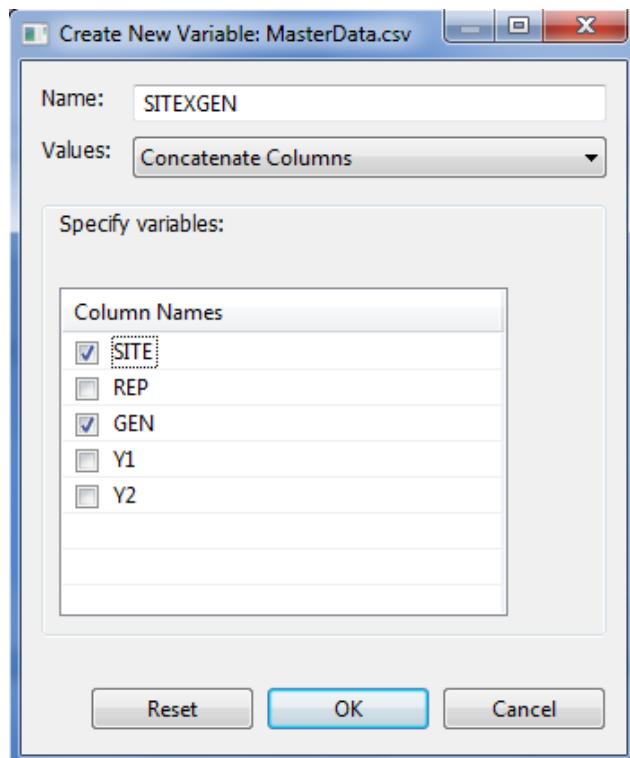


- Click the **OK** button. The data set with the transformed variable is displayed in the Data Viewer tab.

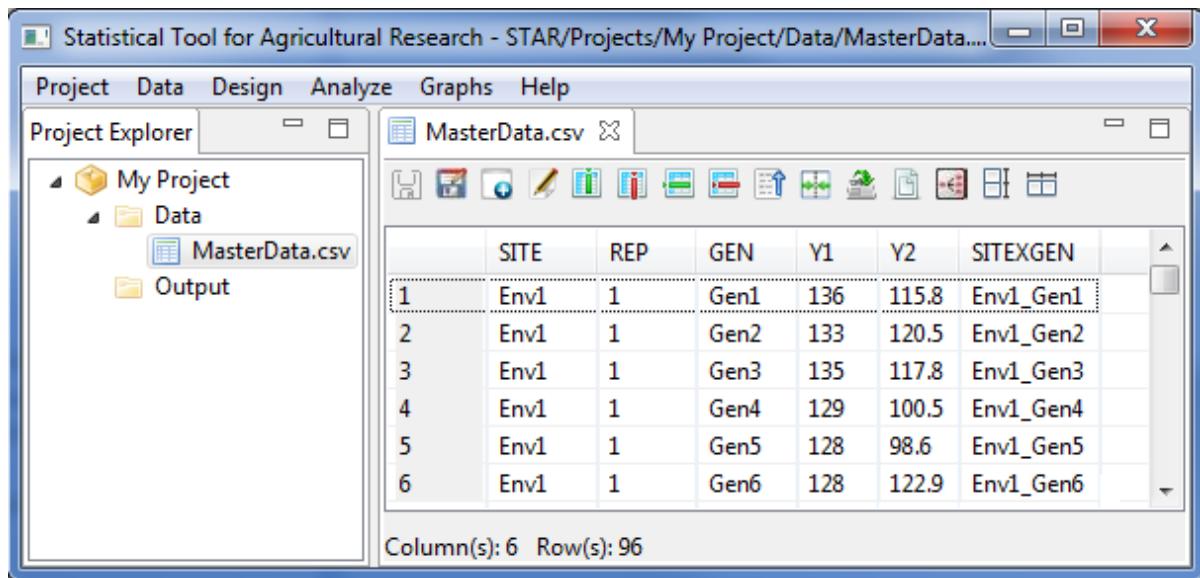
The screenshot shows the main window of the Statistical Tool for Agricultural Research. The menu bar includes Project, Data, Design, Analyze, Graphs, and Help. The Project Explorer tab is selected, showing a project named "My Project" with a "Data" folder containing "MasterData.csv". The Data Viewer tab is also selected, displaying the contents of "MasterData.csv". The table in the Data Viewer shows six columns: SITE, REP, GEN, Y1, Y2, and standardizedY1. The data rows show values for Env1 across six generations (Gen1 to Gen6) with corresponding Y1 and Y2 values and their standardized version. The status bar at the bottom indicates "Column(s): 6 Row(s): 96".

	SITE	REP	GEN	Y1	Y2	standardizedY1
1	Env1	1	Gen1	136	115.8	0.3472559098263
2	Env1	1	Gen2	133	120.5	0.1603218873403
3	Env1	1	Gen3	135	117.8	0.2849445689977
4	Env1	1	Gen4	129	100.5	-0.088923475974
5	Env1	1	Gen5	128	98.6	-0.151234816802
6	Env1	1	Gen6	128	122.9	-0.151234816802

Suppose we want to create a new variable by concatenating the values of two variables. The completed dialog box should appear as shown below:



- Click the **OK** button. The transform variable is displayed in the Data Viewer tab.



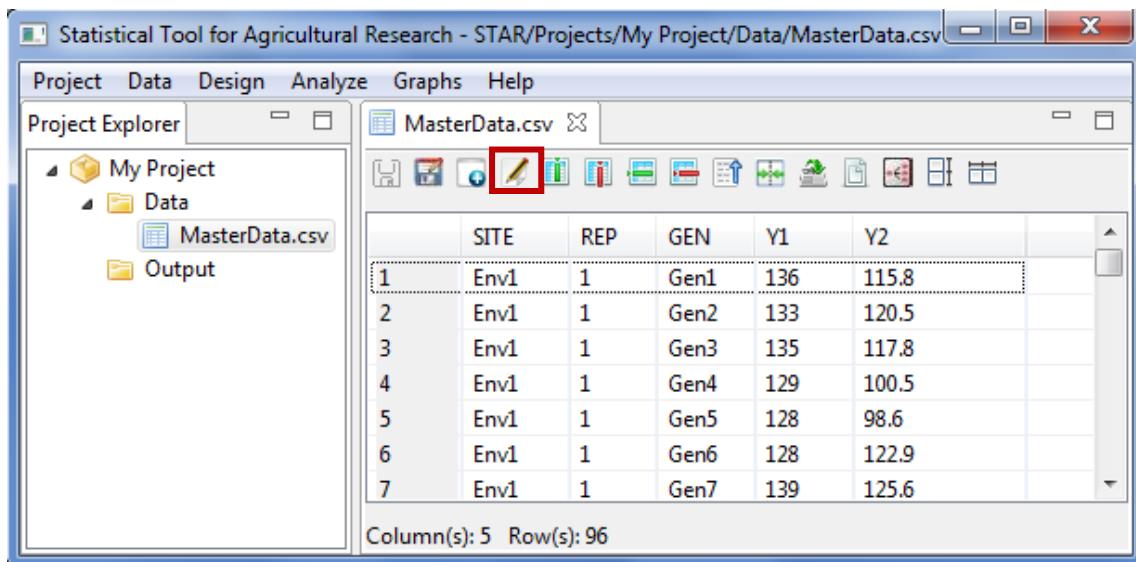
Editing Variable Information

The steps to edit the variable information are listed below:

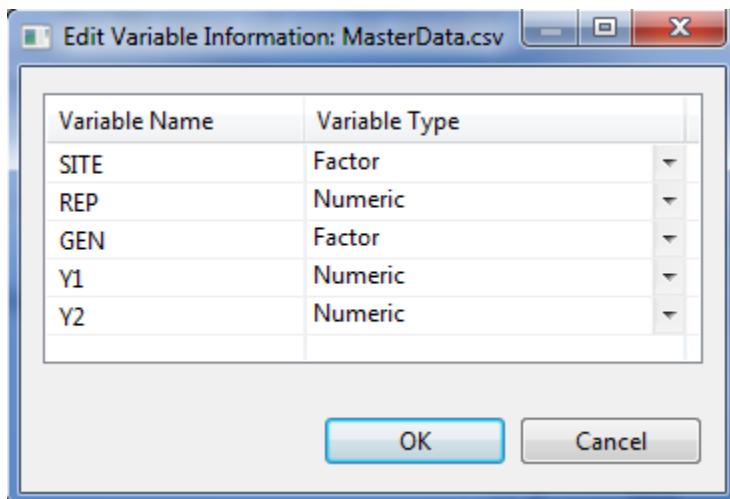
- On the **Project Explorer** panel, double-click the file to be used to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

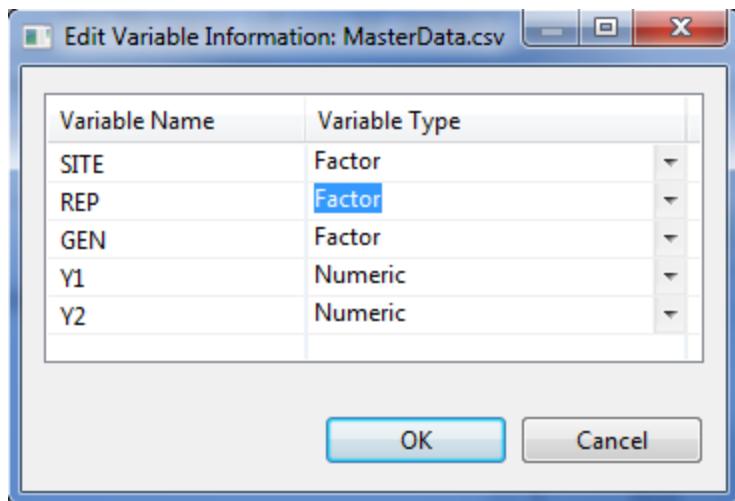
- To edit variable information, click **Data | Edit Variable Information** from the main window of STAR or click on the Edit Variable Information icon  in the Data Viewer toolbar.



- The **Edit Variable Information** dialog box will appear. The variables in the active data set are classified as either a factor or numeric variable. The user can modify the variable type of the variables in the data set.



For the example, change the Variable Type of *REP* from Numeric to Factor. The complete dialog box should appear as shown below:



- Click the **OK** button.

Sorting

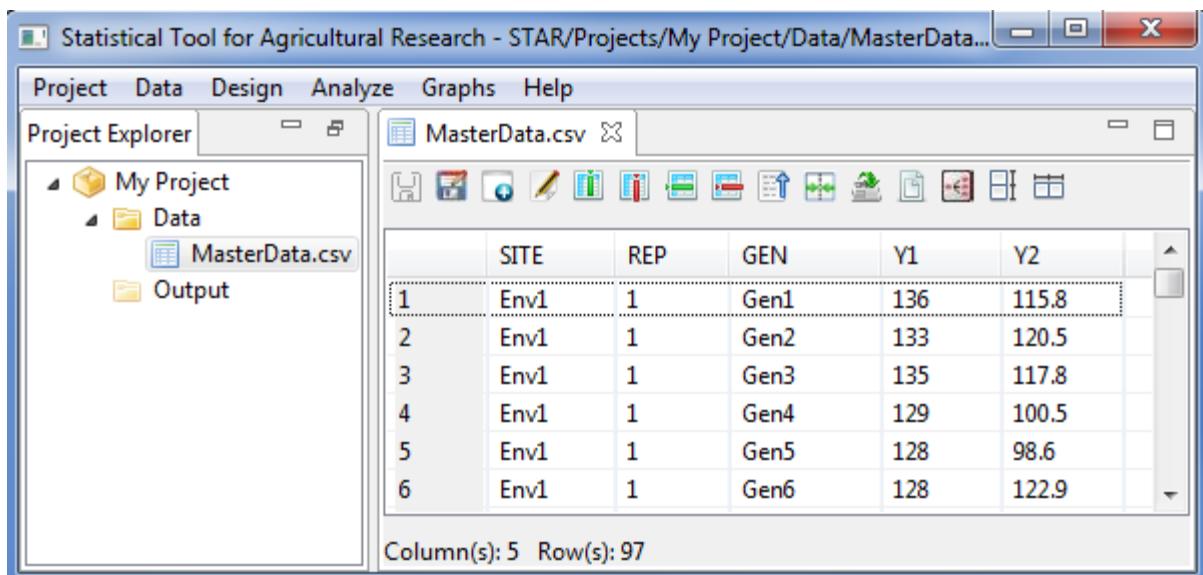
Re-arranging or sorting the rows or cases of the data file is often useful and sometimes necessary for certain types of analysis. The **Sort Cases Menu** can be used to sort rows based on the value of one or more sorting variables.

The steps to re-arrange the data set based on the value of one or more sorting variables are listed below:

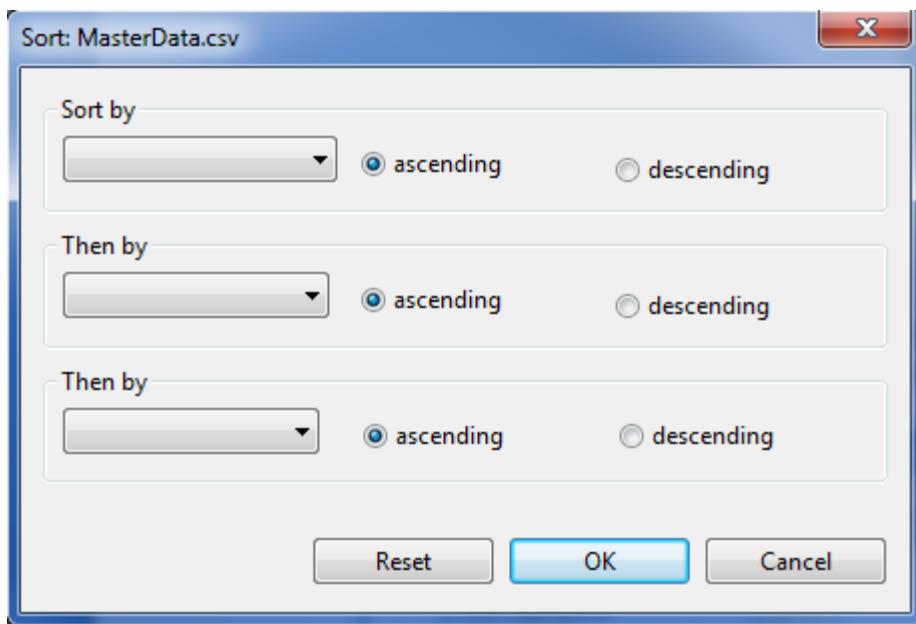
- On the **Project Explorer** panel, double-click the file to be used to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. The data file is arranged by *SITE*, *REP* then by *GEN*.

To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

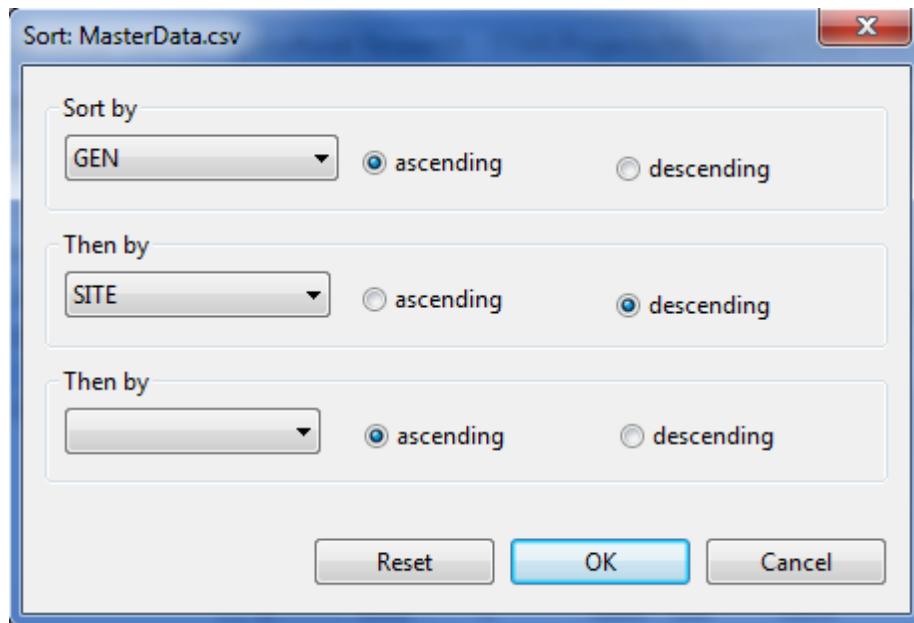


- Choose **Data | Sort...** from the main window or click on the Sort icon  in the Data Viewer tool bar. The **Sort** dialog box will appear.



- Specify up to three variables as the basis for sorting and the order of sorting. Click the drop-down list box to identify the sorting variables. The rows can be re-arranged or sorted in ascending or descending order, with ascending order as the default option. If two or three variables are selected, rows are sorted for each variable within categories of the preceding variables. For character variables, uppercase letters precede their lowercase counterparts.

For the example, suppose we want to re-arrange the content of the data *MasterData.csv*, by ascending order of the variable *GEN* then by descending order of the variable *SITE*. The completed dialog box should appear as shown below:



- Click the **OK** button to sort the active data. The **Sort** dialog box is closed and the sorted data is saved in the parent folder of the active data and displayed in the Data Viewer tab.

A screenshot of the Statistical Tool for Agricultural Research interface. The Project Explorer shows a project named 'My Project' with a 'Data' folder containing 'MasterData.csv' and 'MasterData_sorted.csv'. The Data Viewer tab is active, displaying a table with columns SITE, REP, GEN, Y1, and Y2. The data rows are:

	SITE	REP	GEN	Y1	Y2
1	Env3	1	Gen1	143	112.8
2	Env3	2	Gen1	134	115.5
3	Env3	3	Gen1	113	154.2
4	Env3	4	Gen1	121	163.5
5	Env2	1	Gen1	136	143.5
6	Env2	2	Gen1	144	130.6

Column(s): 5 Row(s): 96

Merging Datasets

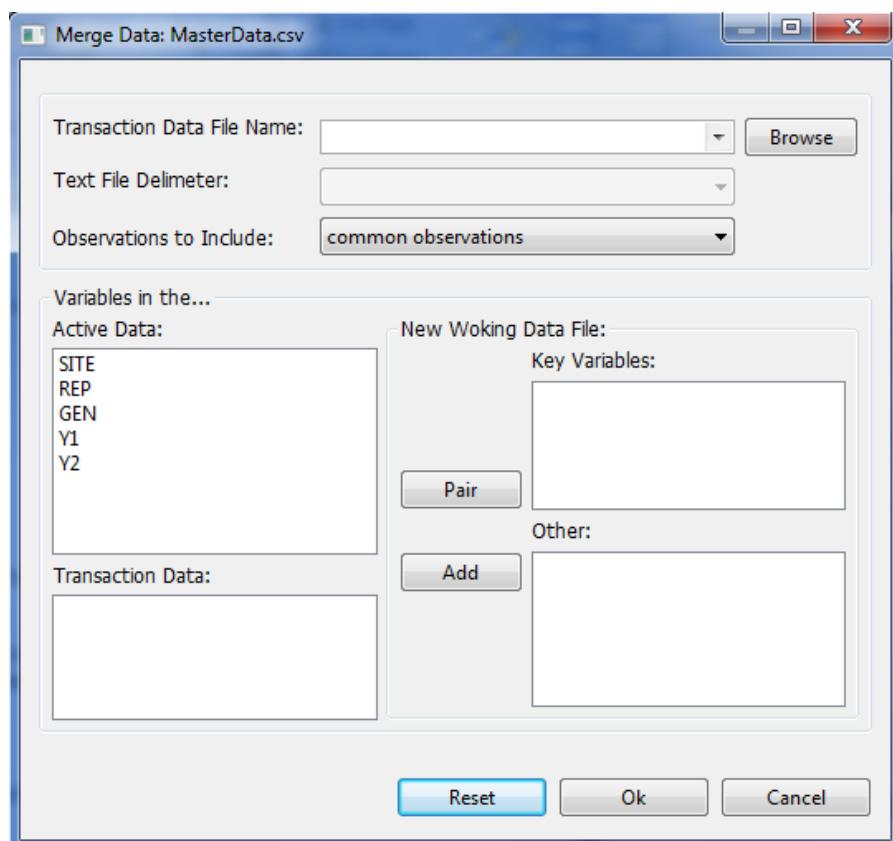
The **Merge** submenu can be used to combine the active data file (known as the master data) with another data file (referred as the transaction data) that contains the same cases or rows but different variables.

The steps to merge datasets are listed below:

- On the **Project Explorer** panel, using the project named *MyProject*, import the data files *MasterData.csv* and *MergeTransact.csv* from the *Data* folder of the project named *SampleProject*. Double-click the file *MasterData.csv* to open it and view it in the Data Viewer.

To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

- Choose **Data | Merge Datasets...** or click on the Merge Datasets icon  in the Data Viewer toolbar. The **Merge Data** dialog box will appear.



- Specify the required fields and appropriate options.

Transaction File Name

Specify this file by selecting a file using the drop-down combo box or by locating it using the **Browse** button. The files included in the drop-down combo box are files inside the *Data* folder not in its sub-folder. Only *.txt* and *.csv* data file format can be selected as the transaction file.

Text File Delimeter

If the transaction file selected is a text file, the delimeter should be specified. Four delimeters are available, namely: comma, space, tab and semi-colon.

Observations to Include

This option pertains to how rows or cases will be included in the new merged data file. The options available are to include common observations (default value), all observations in the active data, all observations in the transaction data and all observations.

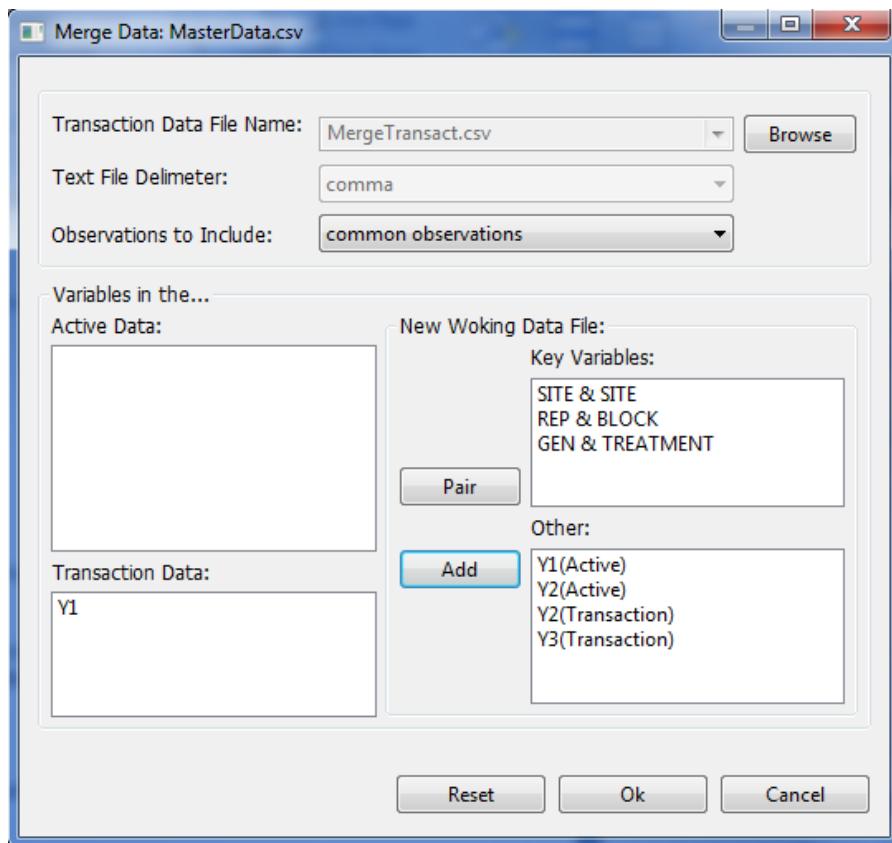
Key Variables

These are pair(s) of variables that are present in the active and transaction files and will be the basis for the merging. Variables can be paired by clicking one variable in the active data and clicking one variable in the transaction data then clicking the **Pair** button. If the pair of variables has different column names, the new merged data file that will be created will use the column name of the active data.

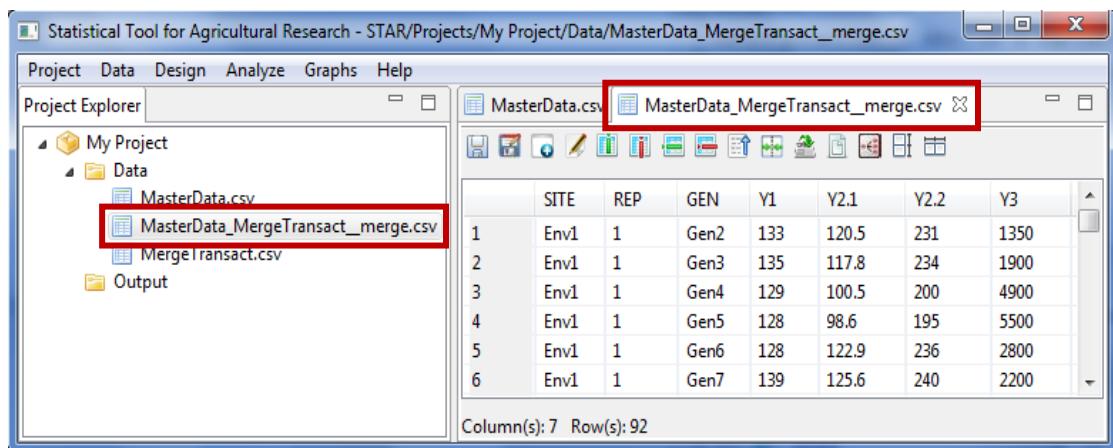
Other

These are the other variables the user wants to include in the merged data file. Variables can be added by selecting at least one variable in the active data and/or in the transaction data then clicking the **Add** button. If columns in this list have the same name, **.1** will be appended to the column name coming from the active data and **.2** to the one from the transaction data.

For the example, select *MergeTransact.csv* file as the transaction data file from the drop-down combo box. The completed dialog box should appear as shown below:



- Click the **OK** button. The merged data is saved in the parent folder of the active data and displayed in the Data Viewer tab. The default filename of the merged data set follows the format *<activeDataFilename>_<TransactionDataFilename>.merge.csv*.



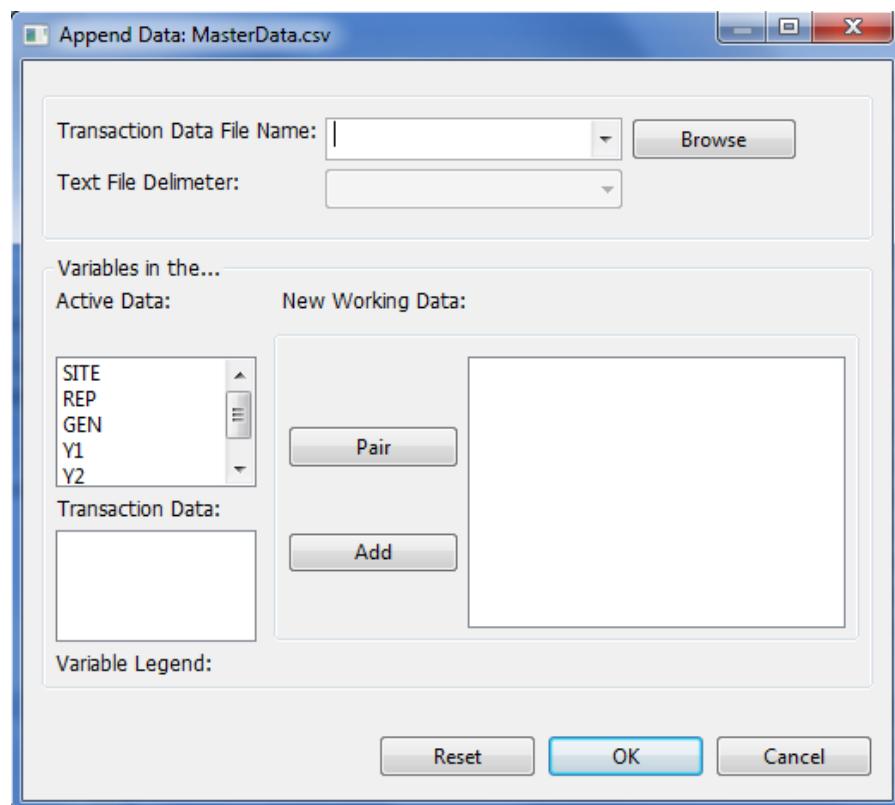
Appending Datasets

The Append submenu can be used combine two data files with the same variables but different cases.

- On the **Project Explorer** panel, using the project named *MyProject*, import the data files *MasterData.csv* and *AppendTransact.csv* from the *Data* folder of the project named *SampleProject*. Double-click the file *MasterData.csv* to open it and view it in the Data Viewer.

To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

- Choose **Data | Append Datasets...** or click on the Append Datasets icon  in the Data Viewer toolbar. The Append Data dialog box will appear.



- Specify the required fields and appropriate options.

Transaction File Name

Specify this file by selecting a file using the drop-down combo box or by locating it using the **Browse** button. The files included in the drop-down combo box are files

inside the *Data* folder not in its sub-folder. Only .txt and .csv data file format can be selected as the transaction file.

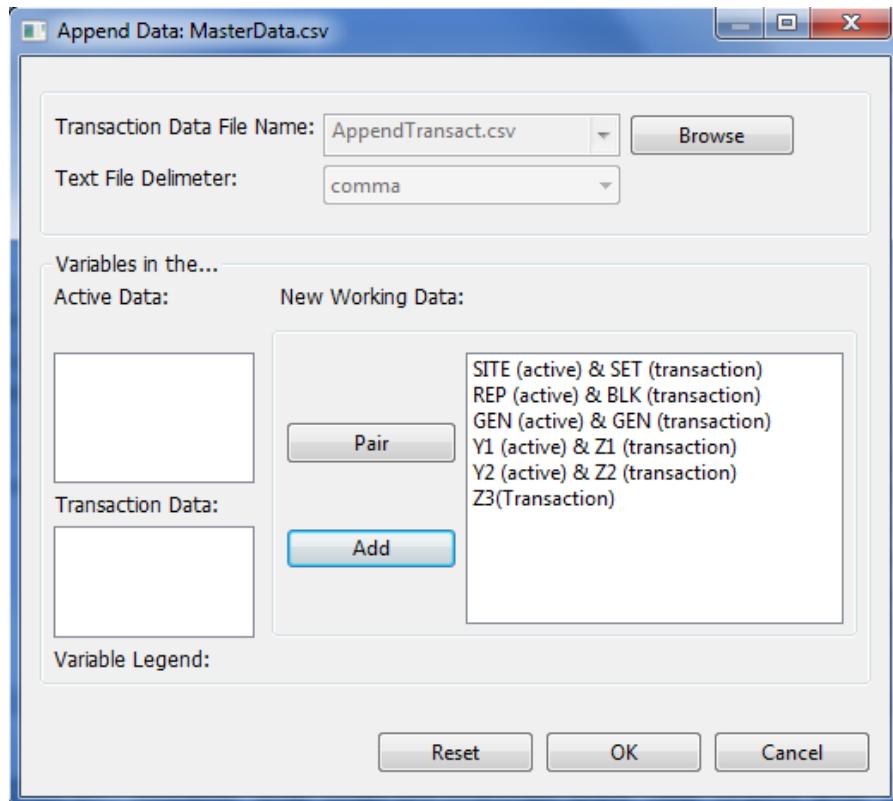
Text File Delimiter

If the transaction file selected is a text file, the delimiter should be specified. Four delimiters are available, namely: comma, space, tab and semi-colon

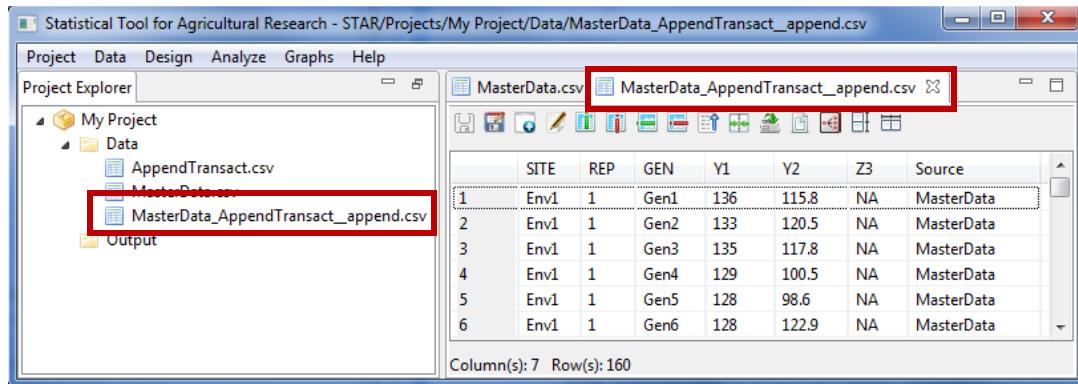
New Working Data

This is the list of variables or paired variables that will be included in the new data set. Variables can be paired by clicking one variable in the active data and clicking one variable in the transaction data then clicking the **Pair** button. Other variables can be added to this list by clicking the variable then clicking the **Add** button. For paired variables in this list, the appended data file that will be created will use the column name of the active data. For non-paired variables, .1 will be appended to the column name coming from the active data and .2 to the one from the transaction data.

For the example, select *AppendTransact.csv* file as the transaction data file from the drop-down combo box. The completed dialog box should appear as illustrated below:



- Click the **OK** button. The appended data is saved in the parent folder of the active data and displayed in the Data Viewer tab. The appended data will contain an additional column named **Source** which indicates where the data came from. The default filename of the appended data set follows the format `<activeDataFilename>_<TransactionDataFilename>_merge.csv`.



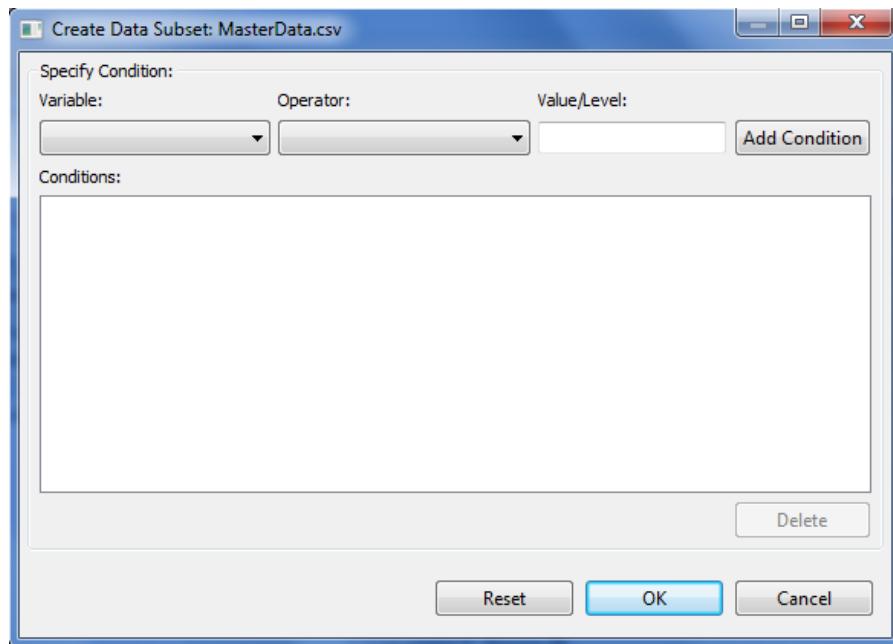
Creating Data Subset

The steps in creating a subset of a dataset are listed below:

- On the **Project Explorer** panel, using the project named *My Project*, import the data file *MasterData.csv* from the *Data* folder of the project named *SampleProject*. Double-click the file to open it and view it in the Data Viewer.

To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user manual.

- Choose **Data | Create Data Subset...** or click on the Create Data Subset icon in the Data Viewer toolbar. The **Create Data Subset** dialog box will appear.



- Specify the required fields and appropriate options.

Variable

This drop-down list box lists all the variables in the active data set. User can specify the variable by clicking the drop-down list box and selecting the variable of interest.

Operator

This drop-down list box lists the available relational operators that can be used. The seven available operators that can be used are: equal (==), not equal (!=), at least (\geq), at most (\leq), greater than (>), less than (<) and within the range. User can specify the variable by clicking the drop-down list box and selecting the operator of interest.

Value/Level

Specify the level or value of the variable specified. If the operator specified is *within* range, user must specify two values separated by comma. Otherwise, user should specify only one value/level.

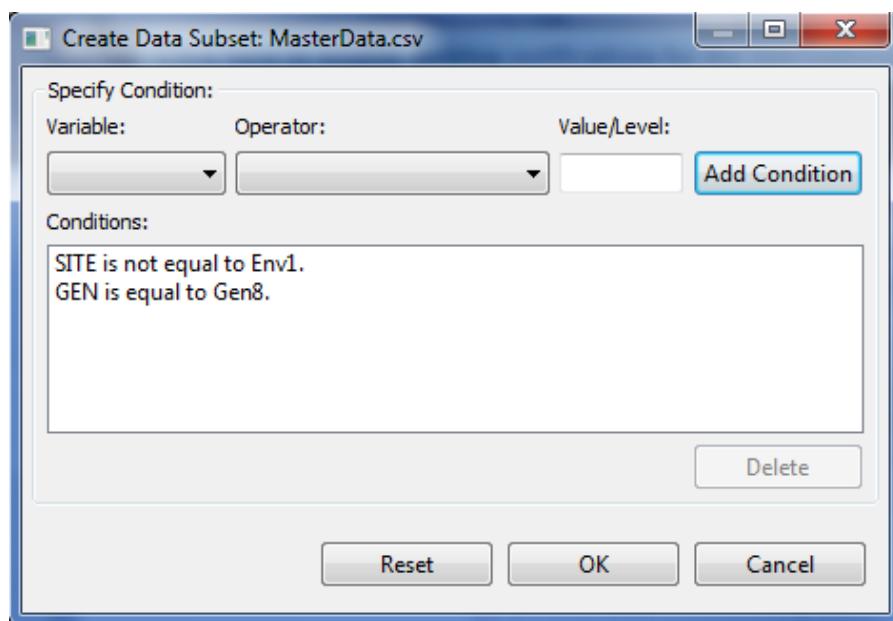
Add Condition

After specifying the conditional expression, user needs to click this button for the expression to be included in the **Conditions** list box.

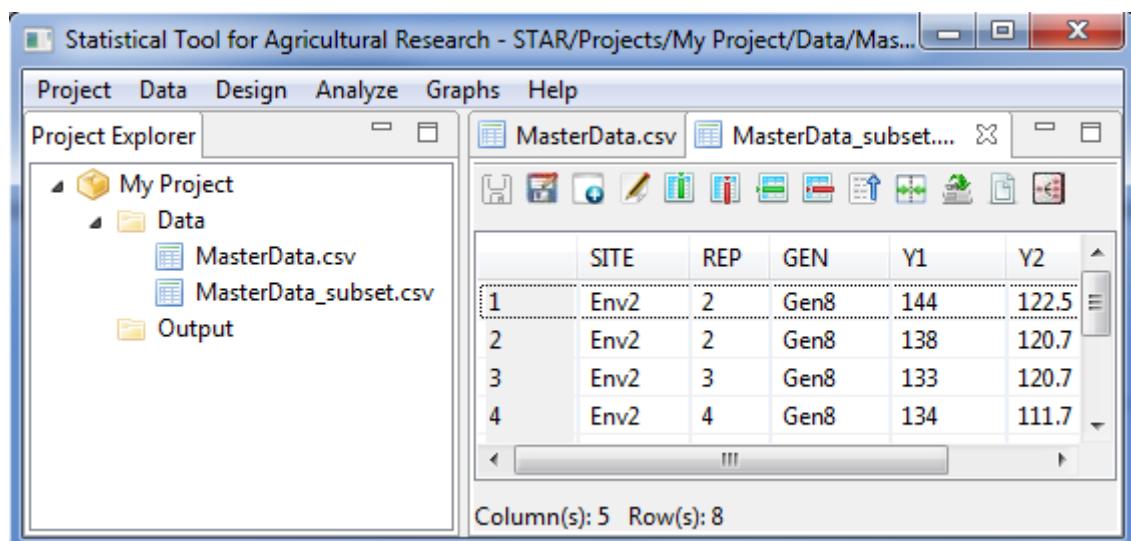
Conditions

All the conditional expressions specified in this list box will be used when sub setting the active data set.

For the example, the completed dialog box should appear as shown below:



- Click the **OK** button. The new data is saved in the parent folder of the active data and displayed in the Data Viewer tab. The default filename of the new data set follows the format <activeDataFilename>_subset.csv



Reshaping

The **Reshape** submenu can be used to restructure or reshape your active data set to follow the structure of the data set appropriate for the STAR procedure you want to use.

Reshaping Data from Long (Serial) to Wide (Parallel) Format

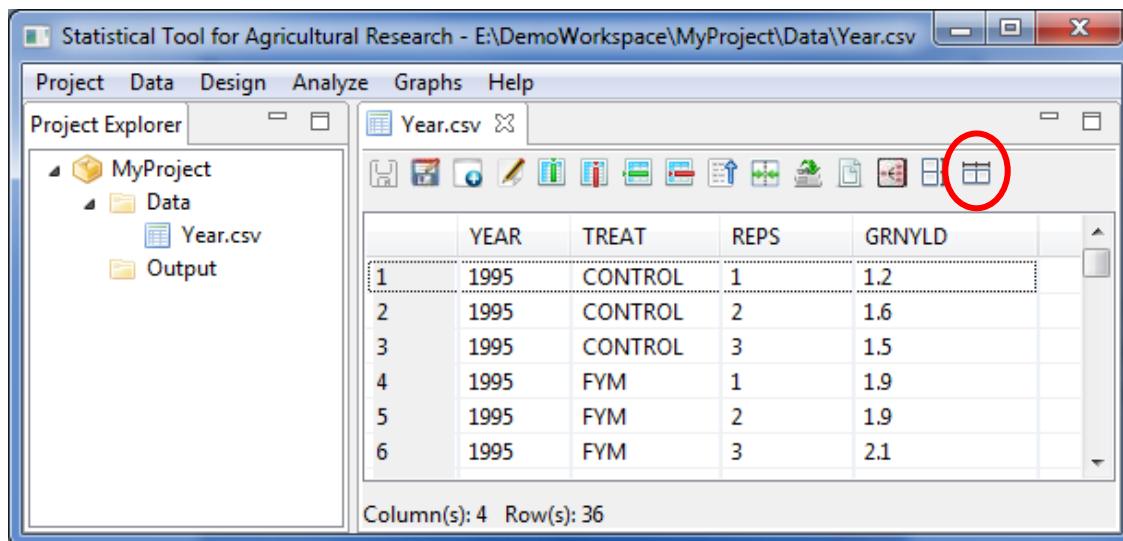
Reshaping data from long (serial) to wide (parallel) involves re-arranging a data file, such that, repeated measurements are in separate columns.

The steps for reshaping the data from long to wide format are listed below:

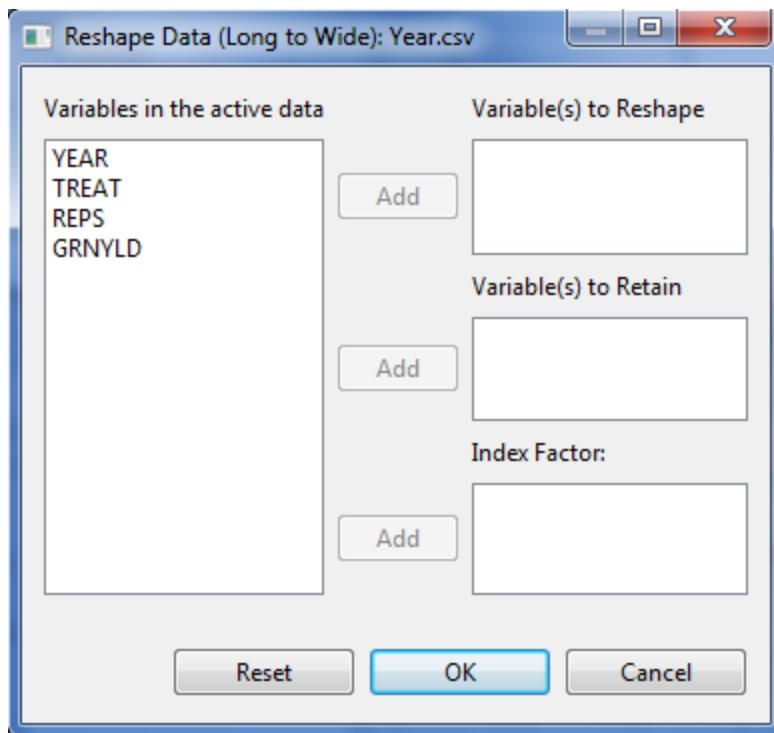
- On the **Project Explorer**, double-click the file to be reshaped to open and view it in the Data Viewer tab.

For the example, import the data file *Year* from the package. To import data from the package, see *Importing Data from Package* section of this user manual.

- Click **Data | Reshape | Long (Serial) to Wide (Parallel) ...** from the main window or click on the Reshape to Wide icon  in the Data Viewer tool bar.



- The **Reshape Data (Long to Wide)** dialog box will automatically appear.



- Specify the required fields and appropriate options.

Variables to be Reshape

This is the list of variable(s) whose values are to be divided into two or more columns and saved to different variables in terms of the levels of the index factor.

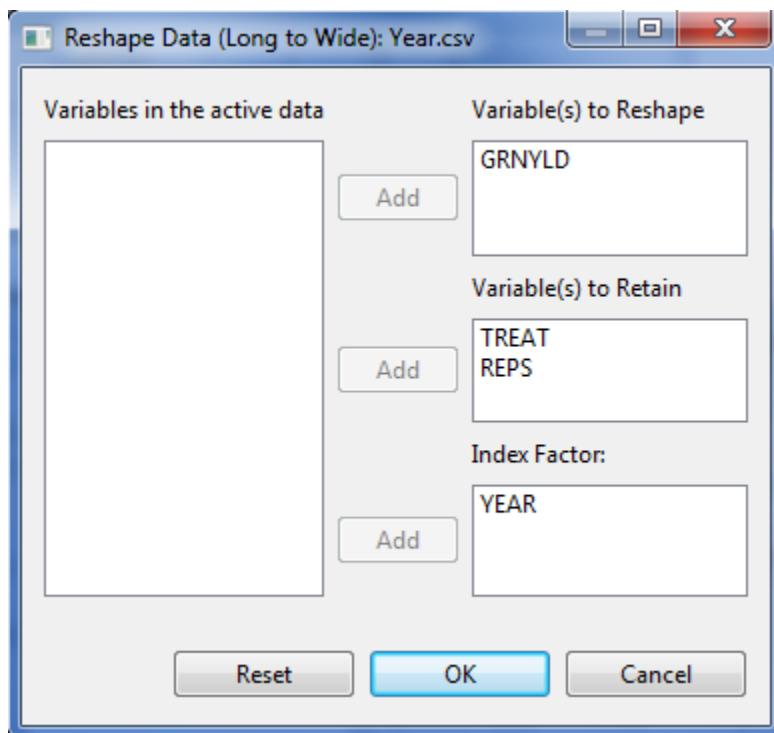
Variables to be Retained

This is the list of variable(s) which will be retained to describe the individual cases or observations.

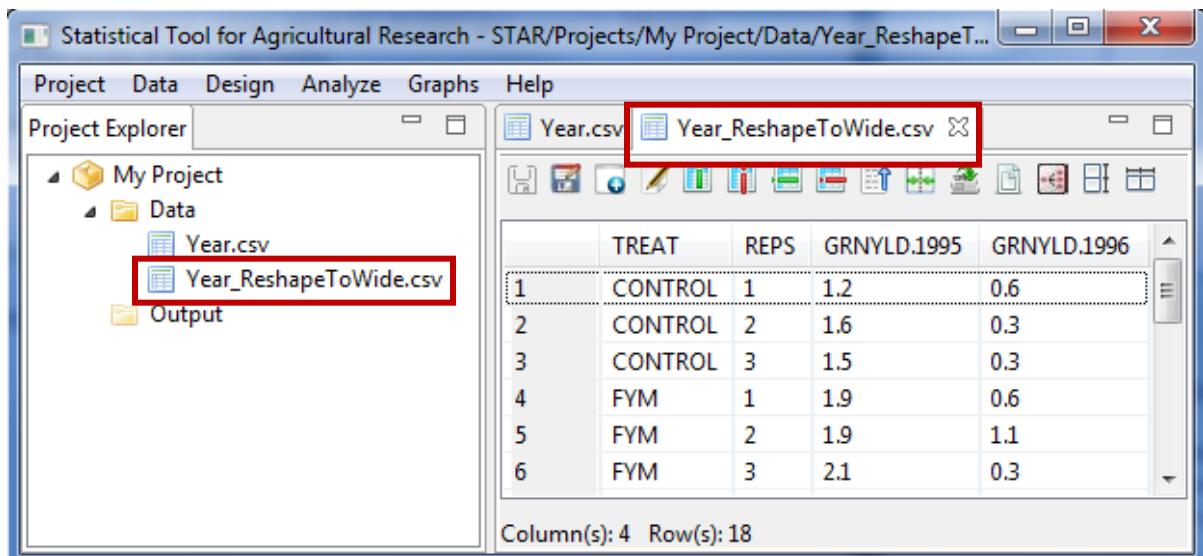
Index Factor

This pertains to variable(s) which will determine the groupings of the values of the variable(s) to be transposed

For the example, the completed **Reshape Data (Long to Wide)** dialog box should appear as illustrated below:



- Click the **Ok** button. The restructured data set is saved in the parent folder of the active data and displayed in the **Data Viewer** tab. The default filename of the restructured data set follows the format *<activeDataFilename>_ReshapeToWide.csv*.



Reshaping Data from Wide (Parallel) to Long (Serial) Format

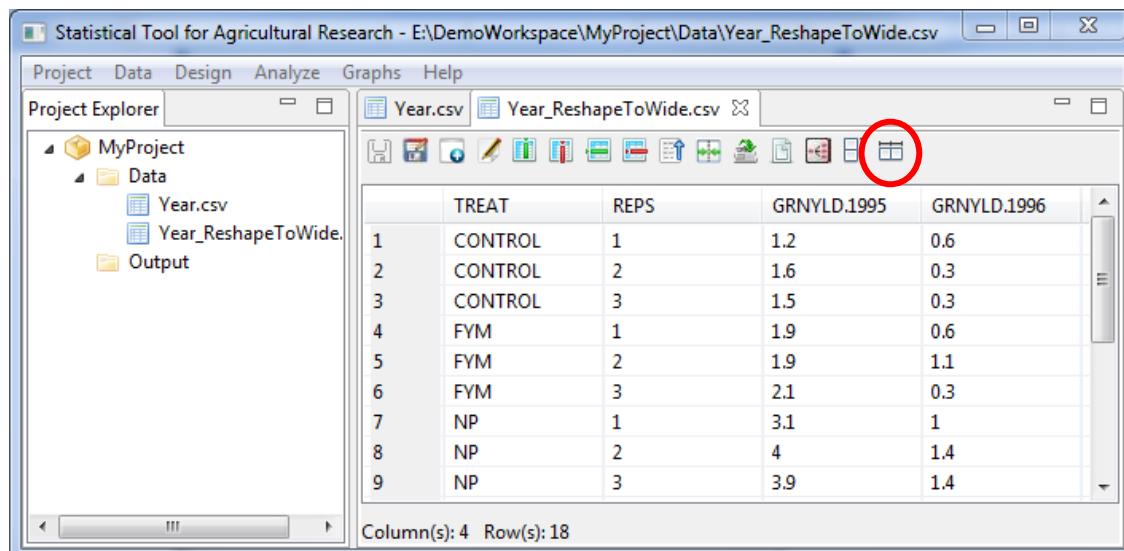
Manipulating data from wide (parallel) to long (serial) refers to re-arranging a multivariate into a univariate data. For instance, different columns representing measurements taken for a response variable over a period of time can be combined to form a single column, with the time variable used as an additional identifier variable.

The steps for reshaping the data from wide to long format are listed below:

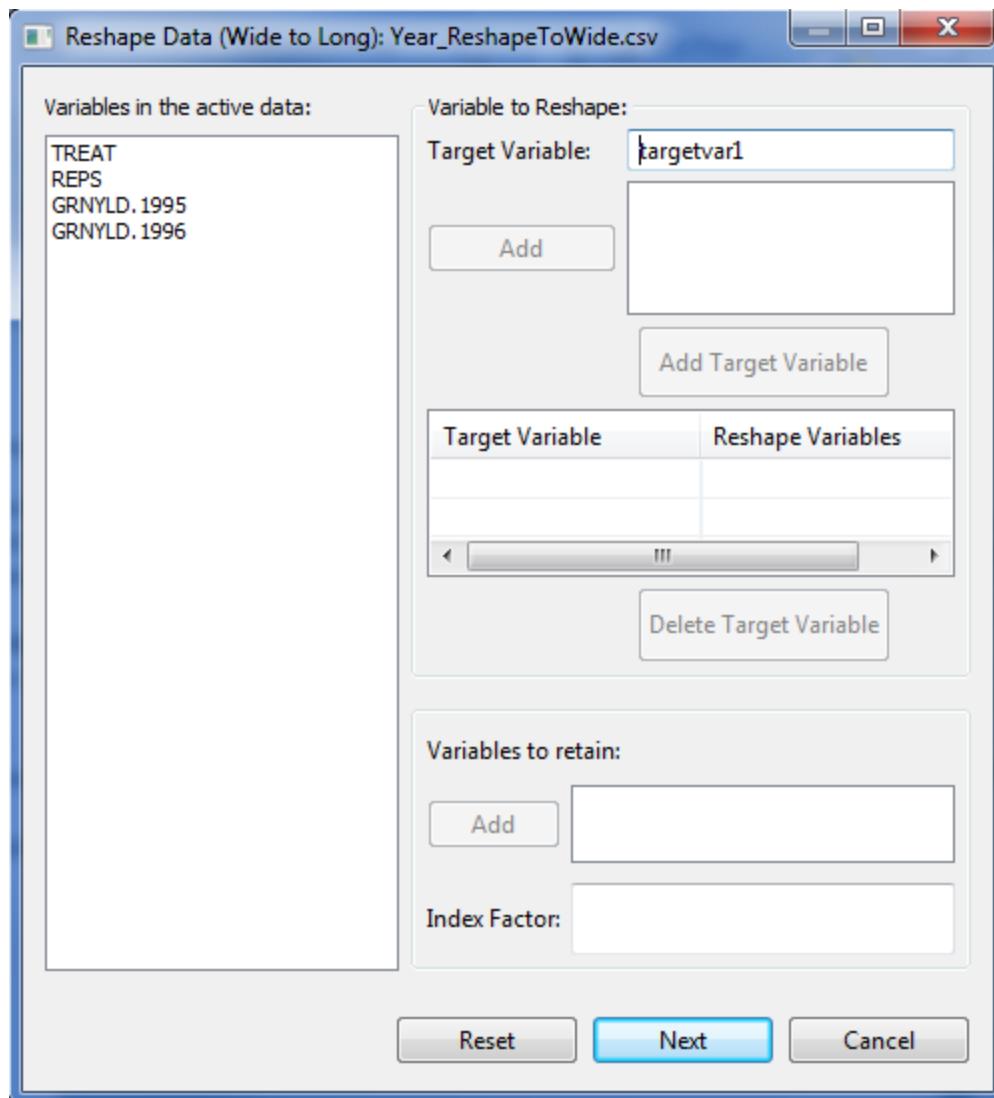
- On the **Project Explorer**, double-click the file to be reshaped to open and view it in the Data Viewer tab.

For the example, locate the *Year_ReshapeToWide.csv* file from the *Data* folder of the project named *MyProject*. This file was generated from the section *Reshape Data from Long to Wide Format* of this user's manual.

- Click **Data | Reshape | From Wide (Parallel) to Long (Serial) ...** from the main window or click on the Reshape to Long icon  in the Data Viewer tool bar.



- The **Reshape Data (Wide to Long)** dialog box will appear.



- Specify the required fields and appropriate options:

Variable to be Reshape

These are the variables that will be combined to form one variable. Click the variables in the list of variables in the active data while holding the Ctrl key then click the **Add** button. Moreover, the user should specify as the target variable the name of the variable to be created. The default target variable name is *targetvar1*. Then click the **Add Target Variable**. In specifying the name of the target variable, the following rules apply:

- The name must start with a letter. The remaining characters can be any letter, any digit, a period or underscore.
- Each column name must be unique; duplication is not allowed.
- Column name is case sensitive

User can create at least one target variable. This target variable should be unique and the length of the *Reshape Variables* should be the equal for all target variables to be created.

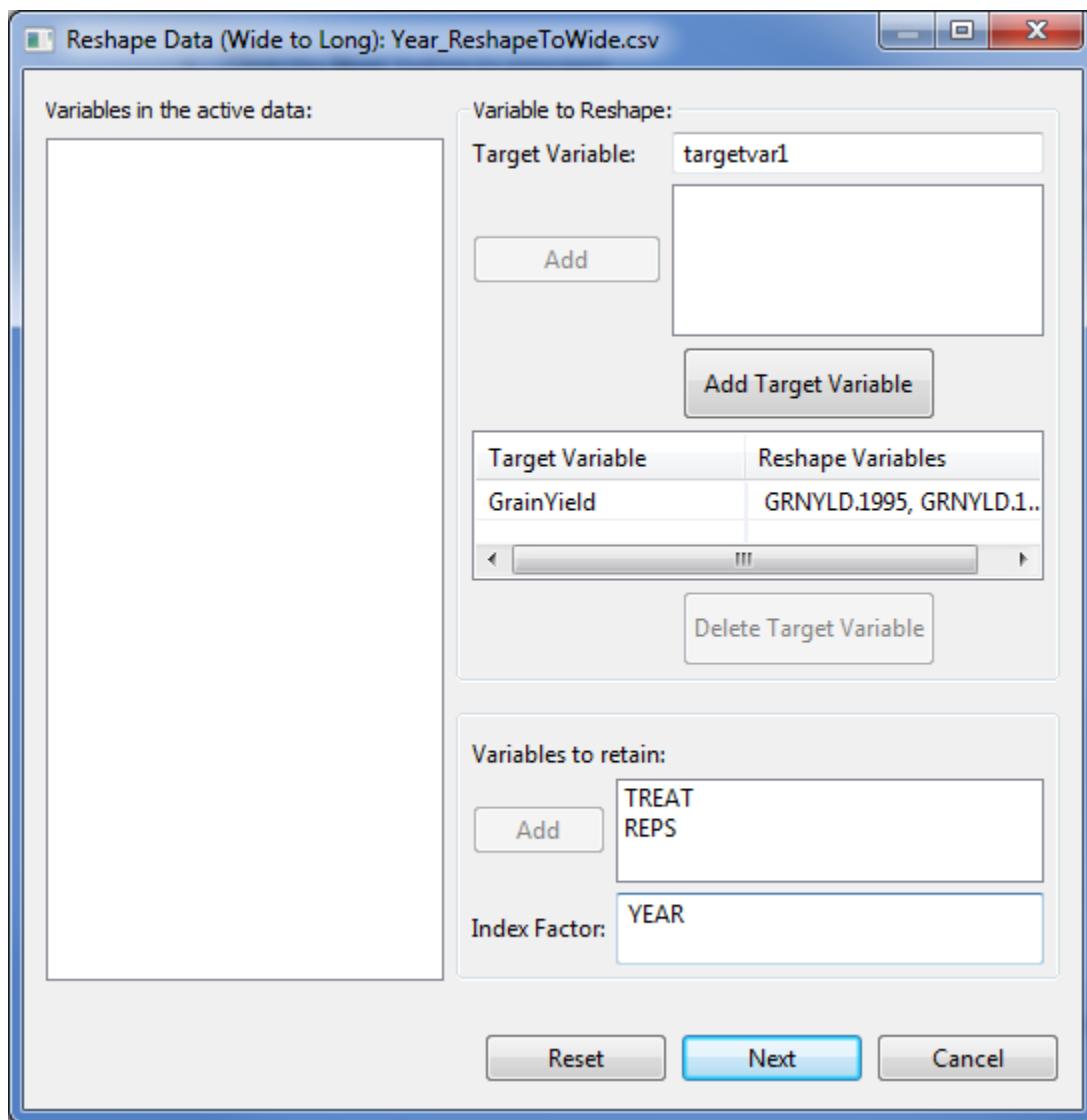
Variables to be Retained

These are the variables that will be included in the restructured data set and pertains to variables that describe the individual cases.

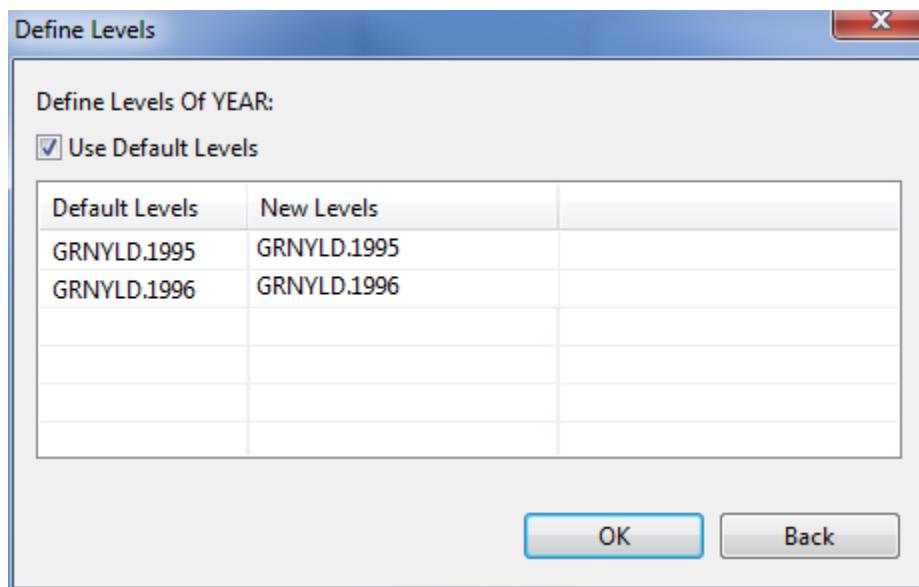
Index Factor

This is the name of the grouping variables that will be created.

For the *example*, the completed dialog box should appear as shown below:

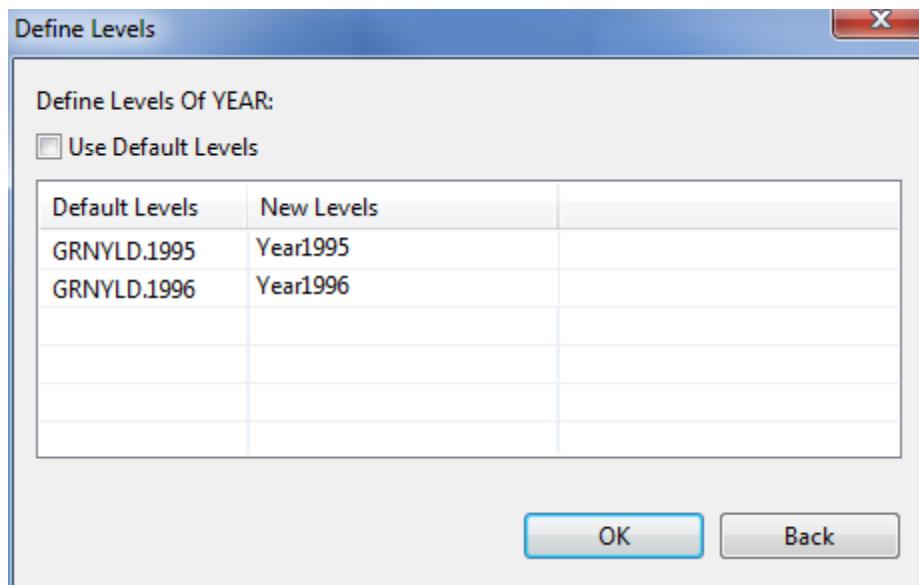


- Click the **Next** button to proceed.
- The **Define Levels** dialog box will appear.

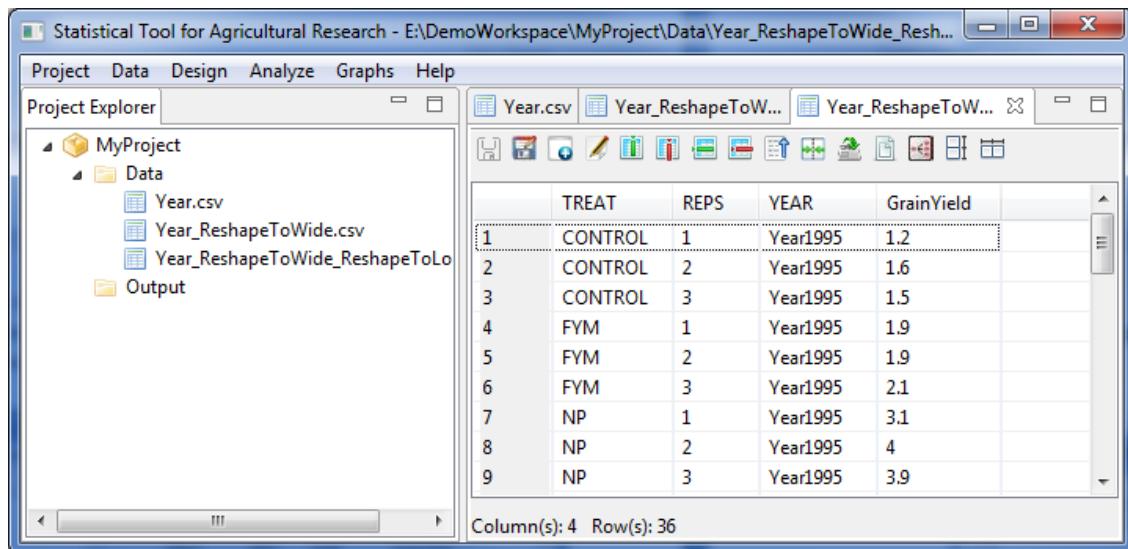


- Specify the required fields. To specify the levels of the Index Factor that will be created, the user can use the default values or type the desired levels in the column labeled 'New Levels'. For now, STAR only accepts non-numeric values for the levels.

For the example, the completed dialog box should appear as shown below:



- Click the **OK** button. The restructured data set is saved in the parent folder of the active data and displayed in the Data Viewer tab. The default filename of the restructured data set follows the format <activeDataFilename>_ReshapeToLong.csv.



Aggregating Data

This feature aggregates group of observations into single observation and creates a new, aggregated data file or creates new columns in the active data file that contain the aggregated data. Observations are aggregated based on the value of one or more grouping variables.

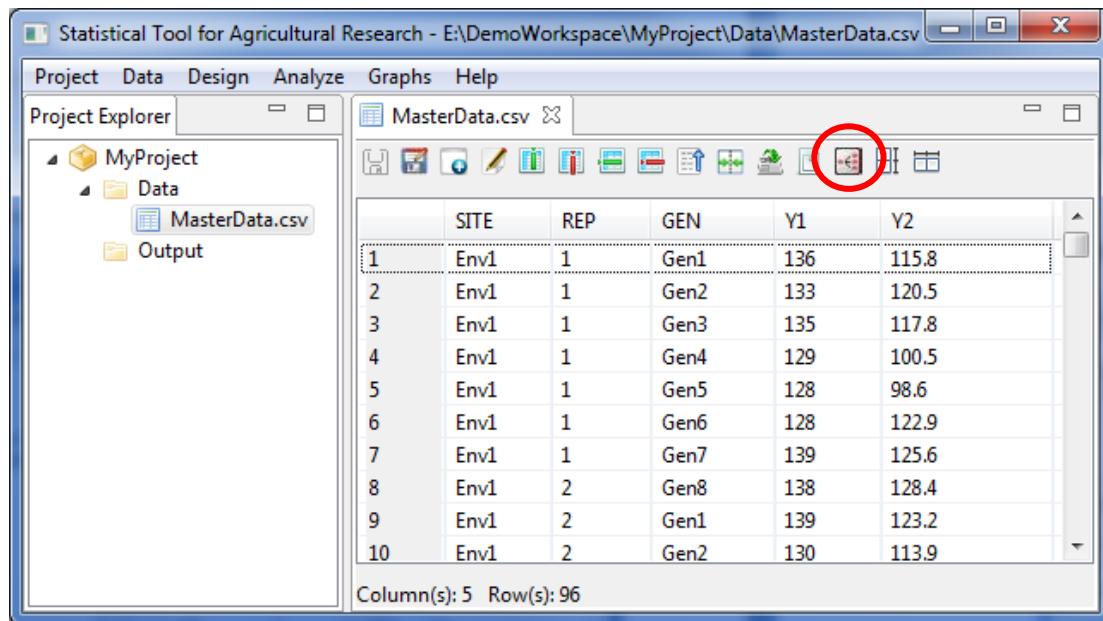
The steps for aggregating the data are listed below:

- On the **Project Explorer**, double-click the file to be used to open and view it in the Data Viewer tab.

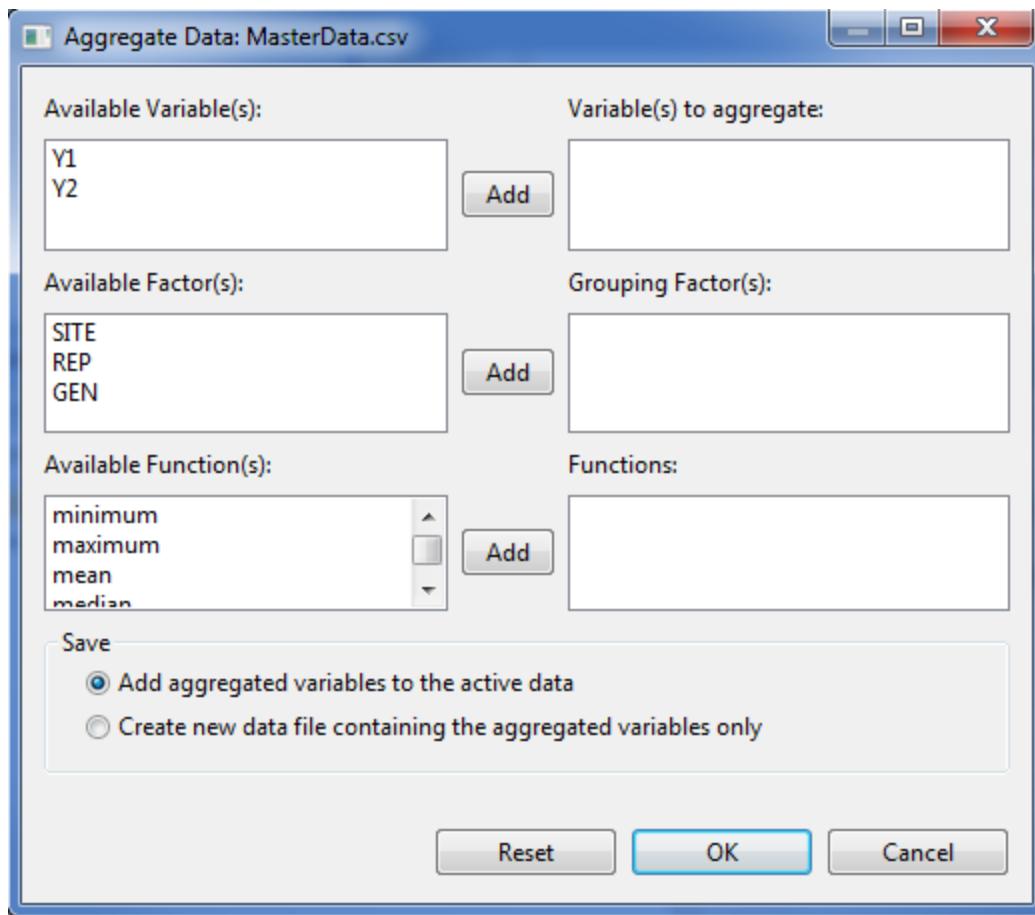
For the example, using the project named *MyProject*, import the data file *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

Opening the data for the first time, *REP* field is regarded by STAR as numerical variable. To change its type, go to **Data | Edit Variable Information** and change the *Variable Type* of *REP* from *Numeric* to *Factor*.

- Choose **Data | Aggregate...** or click on the Aggregate Data icon  in the **Data Viewer** tool bar.



- The **Aggregate Data** dialog box will appear.



- Specify the required fields and appropriate options.

Variable(s) to Aggregate

This is the list of variables whose values will be summarized using the selected functions. At least one entry is needed and should come from the **Available Variable(s)** list box.

Grouping Factor(s)

This field is optional. If specified, it will summarize the values per category of the grouping factor. Entries should come from the **Available Factor(s)** list box.

Functions

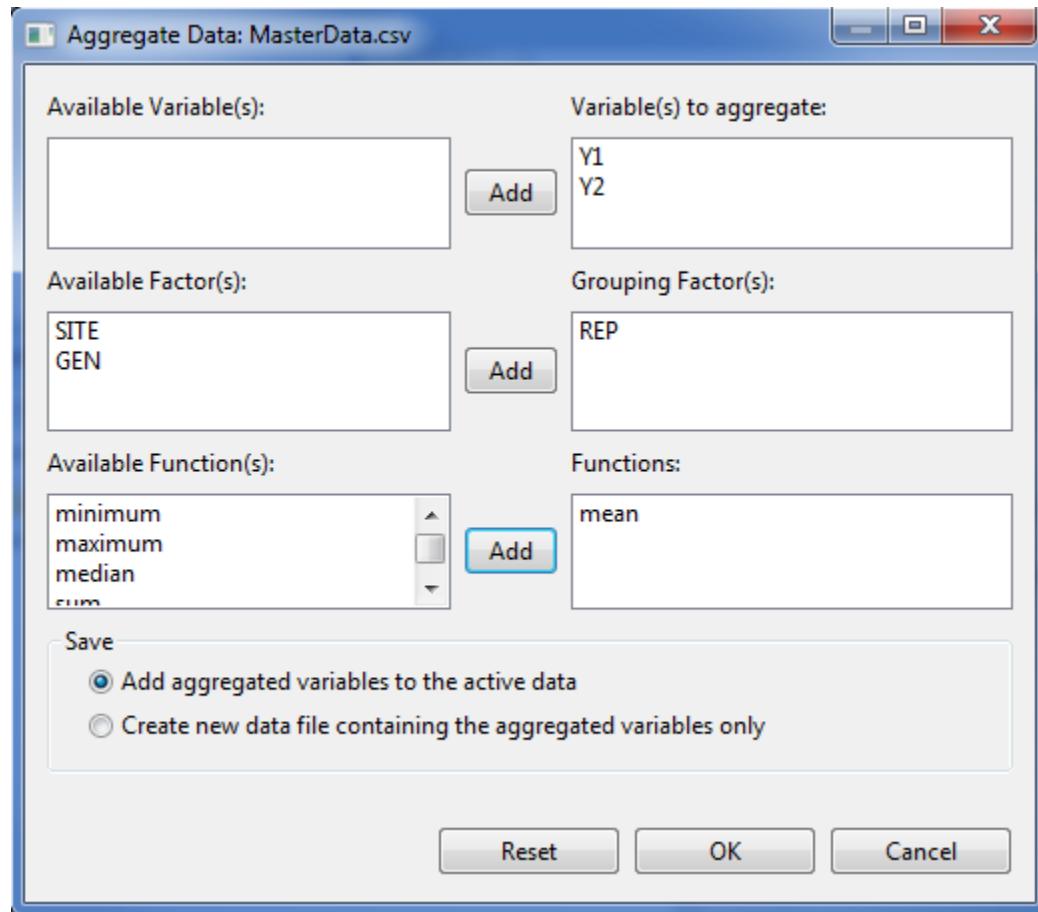
At least one function should be specified. Available functions are minimum, maximum, mean, median, sum, variance and standard deviation.

Save

Specify how the aggregated data will be saved. User can either add aggregated variables to the active data (default options) or create a new data file containing the

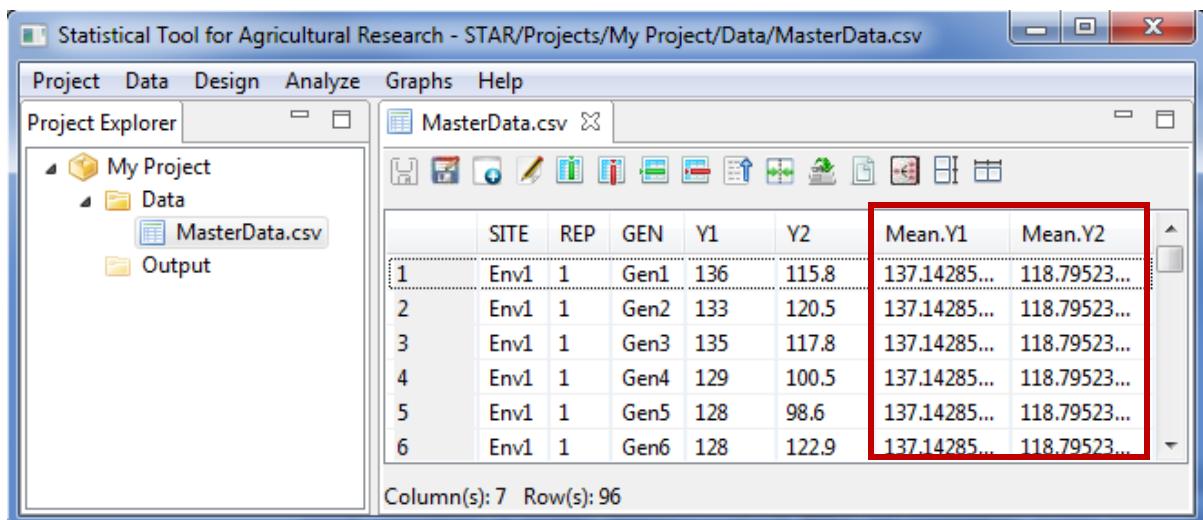
aggregated variables only. If user chooses the default option, the resulting data file is not aggregated. Each observations with the same value(s) of the grouping variable(s) will receive the same values of the new aggregated variables. If user chooses the option create a new data file containing the aggregated variables only, the new data files contains one observation for each category of the grouping variables.

For the *example*, suppose we want to compute the mean of the variables *Y1* and *Y2* for each level of *REP*. The completed dialog box should appear as shown below:

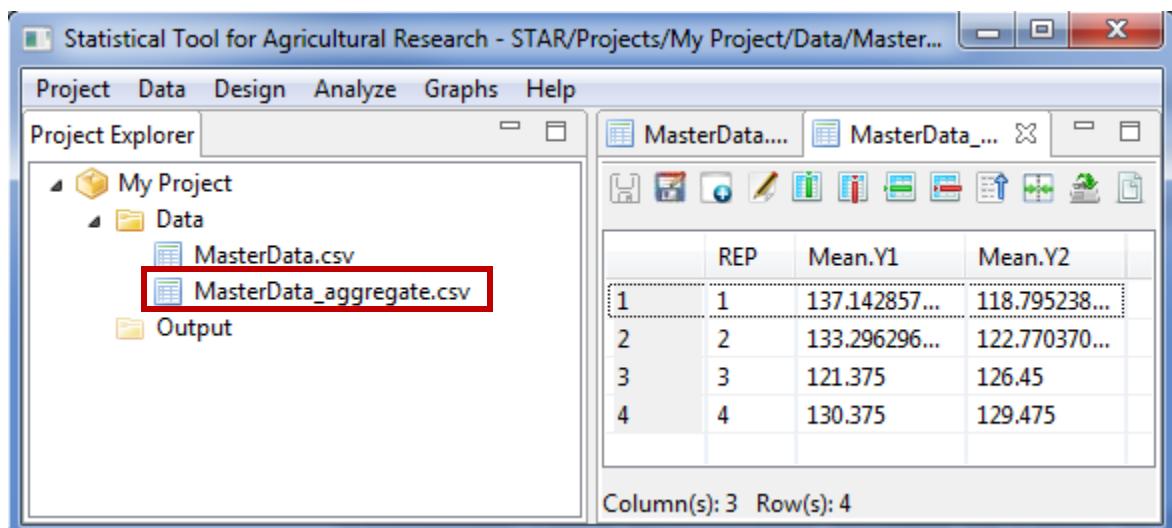


- Click the **OK** button.

If the user chooses the save option: *add aggregated variables to the active data* (default option), the active data is saved with the additional column(s) and displayed in the Data Viewer. The default column name(s) which contain the aggregate data follows the format *<Function>.< Variable to aggregate >*.



If the user chooses the save option: *create a new data file containing the aggregated variables only*, the aggregated data file is saved in the parent folder of the active data and displayed in the Data Viewer. The default column name(s) which contain the aggregate data follows the format <Function>.<Variable name to aggregate> and the default filename of the aggregated data set follows the format <activeDataFilename>_aggregate.csv.



Randomization and Layout

The **Design Menu** allows user to generate randomization and layout for single- and multi-factor designs. If the menu is used for the first-time in the active project, a *Randomization* folder will be created inside the *Output* folder. Sub-folders will be created inside the *Randomization* folder where the generated field book in csv format and a text file will be saved. The default sub-folder name follows the format *<design>_<time stamp>*.

Completely Randomized Design

The randomization that will be generated is for balanced design. Equal number of replicates will be generated for each level of the factor or treatment combination.

The steps to generate randomization are listed below:

- To perform randomization, click **Design | Completely Randomized Design....** The **Randomization and Layout** dialog box will appear.
- Specify the required fields and appropriate options.

Number of Factors

Define the number of factors to be generated with default and minimum value equal to 1 and maximum value equal to 10. The number of factors specified will define the number of rows of the table inside the **Factor Definition** frame. To change the default value, user can either type a value inside the spin box or click the up-arrow key of the spin box to increase the value, down-arrow key, otherwise.

Factor Definition Frame

Name

Contain the default name for each factor which can be change. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.);

- succeeding characters can be a combination of letters, numbers, period(.) and underscore(_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2 while the maximum value of the levels is 500.

Number of Replicates

Define the number of replicates to be generated for each level of the factor or for each treatment combination. The default and minimum value is equal to 2.

Number of Field Rows

Define the number of rows in the field layout for each trial. The default and minimum value is equal to one.

Number of Trial

Define the number of trials to be generated. The default and minimum value is equal to one.

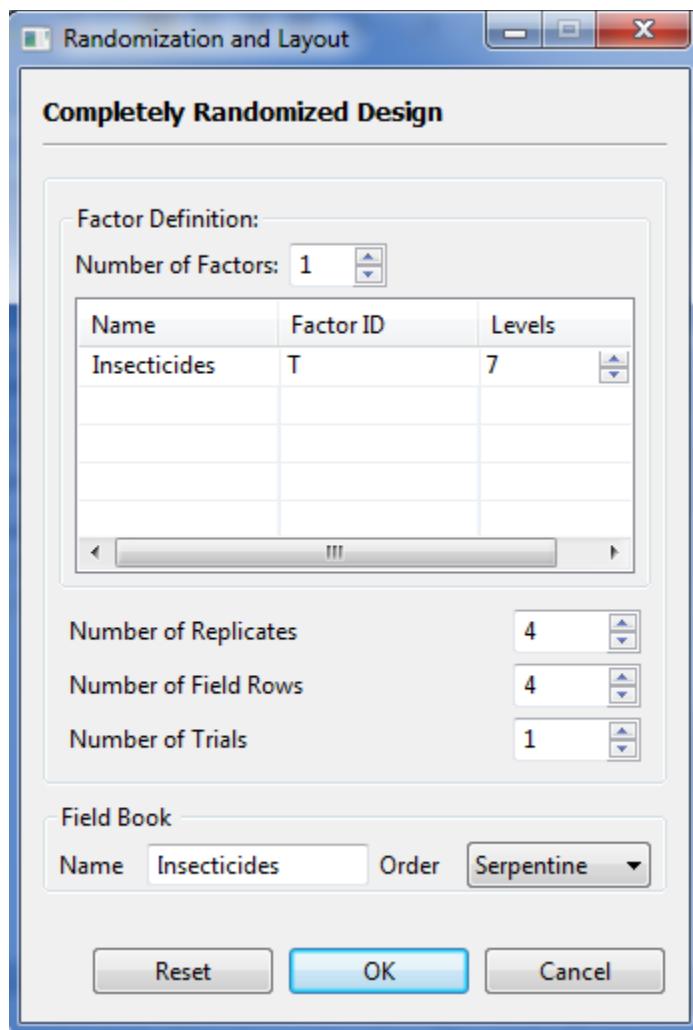
Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved. The default filename is *fieldbookCRD*.

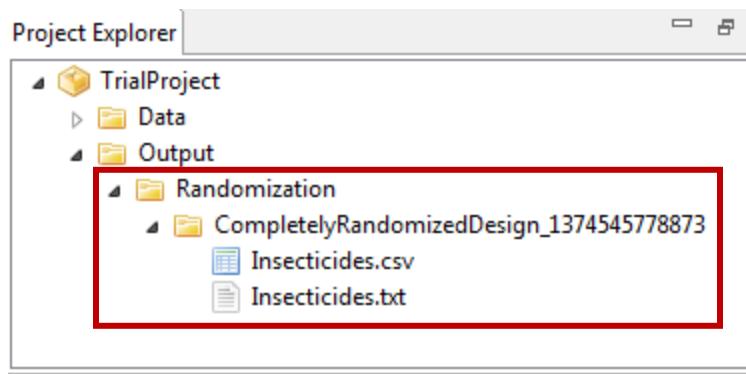
Field Book Order

Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For the example, suppose we want to generate a randomization and layout for an experiment whose aim is to compare the effects of seven different foliar and granular insecticides on the grain yield of a certain rice variety. The experiment will be conducted in completely randomized design with four replicates. The completed dialog box should appear as shown below:



- Click the **OK** button to generate the randomization. The **Randomization** dialog box will be minimized and STAR will automatically display the csv data file and the txt file created in the Data Viewer tab.



Sample *txt* file displayed in the Data Viewer tab is shown below:

DESIGN PROPERTIES:
Single Factor
Completely Randomized Design

DESIGN PARAMETERS:
Number of Trials = 1
Number of Replicates = 4
Treatment Name = Insecticides
Treatment Levels = T1, T2, T3, ..., T7

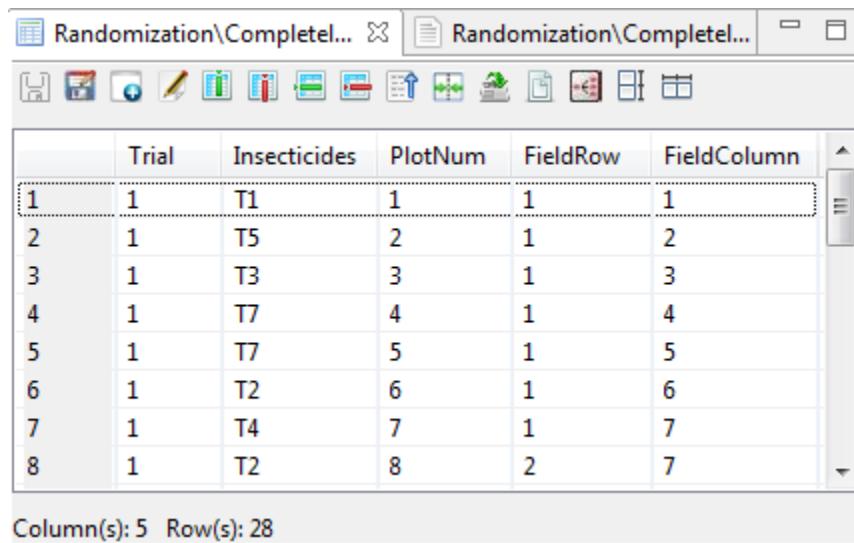
Layout for Completely Randomized Design

Trial = 1

+-----+-----+-----+-----+-----+-----+
1 2 3 4 5 6 7
T2 T7 T3 T6 T4 T5 T1
+-----+-----+-----+-----+-----+-----+
8 9 10 11 12 13 14
T2 T1 T1 T3 T2 T1 T5
+-----+-----+-----+-----+-----+-----+
15 16 17 18 19 20 21
T6 T5 T7 T5 T2 T6 T4
+-----+-----+-----+-----+-----+-----+
22 23 24 25 26 27 28
T3 T6 T4 T4 T7 T7 T3
+-----+-----+-----+-----+-----+-----+

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer tab is shown below:



The screenshot shows a software interface with two tabs labeled "Randomization\Completel...". Below the tabs is a toolbar with various icons. The main area displays a data grid with the following data:

	Trial	Insecticides	PlotNum	FieldRow	FieldColumn
1	1	T1	1	1	1
2	1	T5	2	1	2
3	1	T3	3	1	3
4	1	T7	4	1	4
5	1	T7	5	1	5
6	1	T2	6	1	6
7	1	T4	7	1	7
8	1	T2	8	2	7

Column(s): 5 Row(s): 28

Randomized Complete Block Design

The steps to generate randomization in are listed below:

- To perform randomization, click **Design | Randomized Complete Block Design**. The **Randomization and Layout** dialog box will appear.
- Specify the required fields and appropriate options.

Number of Factors

Define the number of factors to be generated with default and minimum value equal to 1 and maximum value equal to 10. The number of factors specified will define the number of rows of the table inside the **Factor Definition** frame. To change the default value, user can either type a value inside the spin box or click the up-arrow key of the spin box to increase the value, down-arrow key, otherwise.

Factor Definition Frame

Name

Contain the default name for each factor which can be change. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.);

- succeeding characters can be a combination of letters, numbers, period(.) and underscore(_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2 while the maximum value of the levels is 500.

Number of Blocks

Define the number of blocks to be generated. The default and minimum value is equal to 2.

Number of Rows per Block

Define the number of rows per block. The default and minimum value is equal to 1.

Number of Field Rows

Define the number of rows in the field layout for each trial. The default and minimum value is equal to one.

Number of Trial

Define the number of trials to be generated. The default and minimum value is equal to one.

Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved. The default filename is *fieldbookRCBD*.

Field Book Order

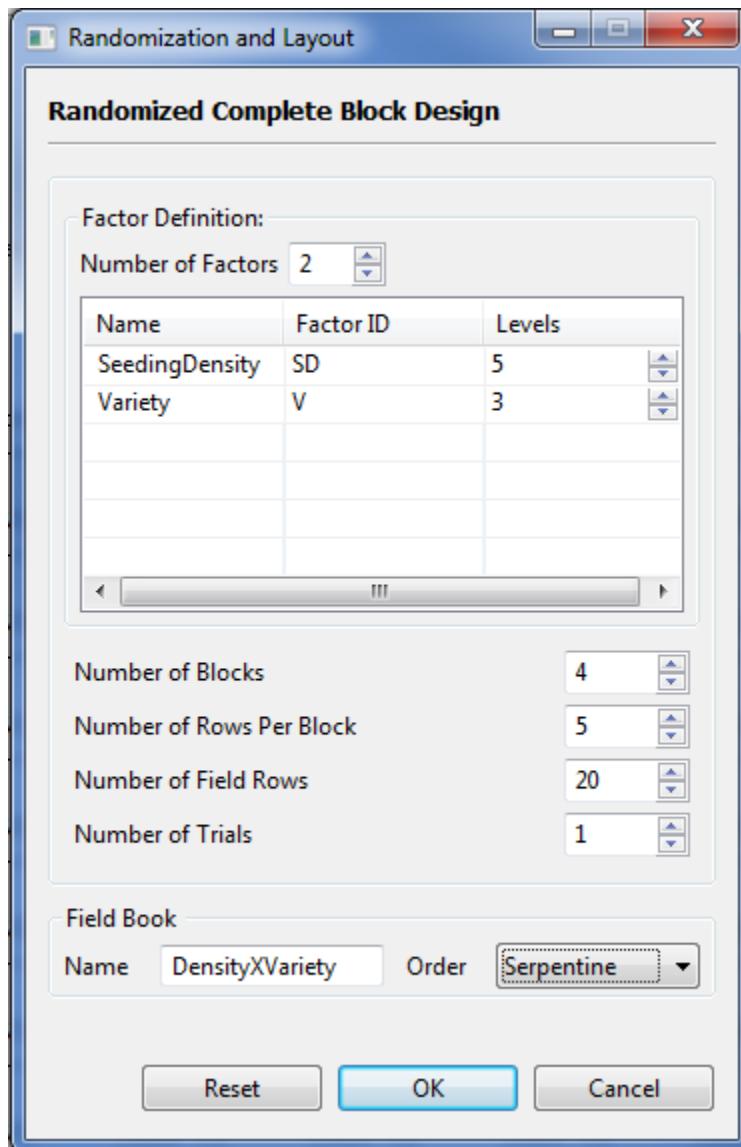
Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For the example, suppose we want to generate a randomization and layout for an experiment whose aim is to compare the effects of five seeding densities on the on grain yield of three rice varieties. The experiment will be conducted in randomized complete block design with four blocks in one trial. The planned field layout is shown below:

	Field Column 1	Field Column 2	Field Column 3
Field Row 1			
Field Row 2			
Field Row 3			
Field Row 4			
Field Row 5			
Field Row 6			
Field Row 7			
Field Row 8			
Field Row 9			
Field Row 10			
Field Row 11			
Field Row 12			
Field Row 13			
Field Row 14			
Field Row 15			
Field Row 16			
Field Row 17			
Field Row 18			
Field Row 19			
Field Row 20			

The diagram shows a field layout with 20 rows and 3 columns. The columns are labeled 'Field Column 1', 'Field Column 2', and 'Field Column 3'. The rows are labeled 'Field Row 1' through 'Field Row 20'. To the right of the grid, there are four curly braces, each enclosing a group of 5 rows. The first brace is labeled 'Block 1', the second 'Block 2', the third 'Block 3', and the fourth 'Block 4'. This indicates that the 20 rows are divided into 4 blocks, with each block containing 5 rows.

The completed dialog box should appear as shown below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the csv data file and the *txt* file created in the Data Viewer tab.

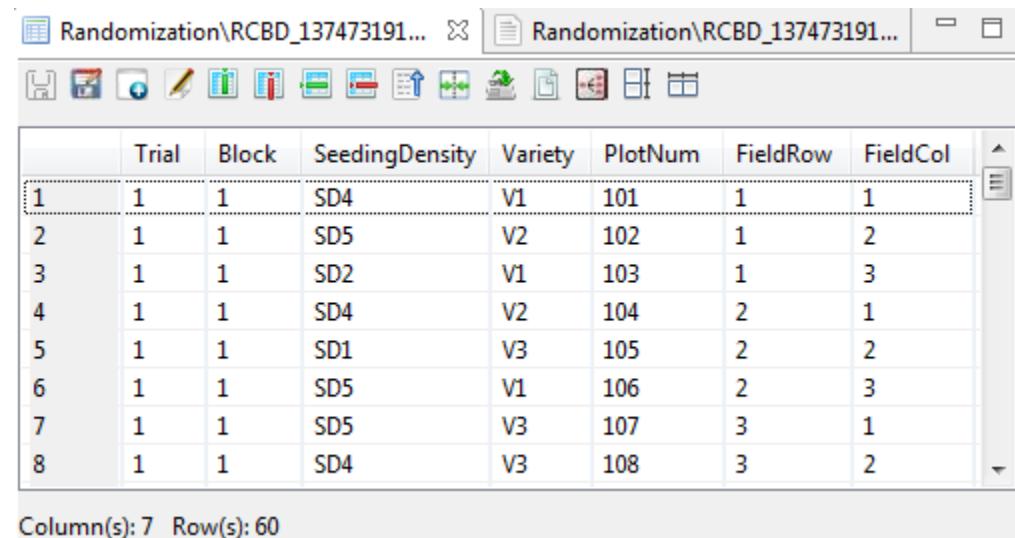
Sample *txt* file displayed in the Data Viewer tab is shown below:

```
DESIGN PROPERTIES:  
    Factorial Design  
    Randomized Complete Block Design  
  
DESIGN PARAMETERS:  
    Number of Trials = 1  
    Number of Blocks = 4  
    Factor 1 = SeedingDensity  
    Levels = SD1, SD2, SD3, SD4, SD5  
    Factor 2 = Variety  
    Levels = V1, V2, V3  
  
    Number of Field Row = 20  
    Number of Field Column = 3  
  
Layout for Randomized Complete Block Design  
  
Trial = 1  
  
          FieldCol1   FieldCol2   FieldCol3  
          +-----+-----+-----+  
FieldRow1 |       101|       102|       103|  
          |SD4 V1     |SD5 V2     |SD2 V1     |  
          +-----+-----+-----+  
FieldRow2 |       104|       105|       106|  
          |SD4 V2     |SD1 V3     |SD5 V1     |  
          +-----+-----+-----+  
FieldRow3 |       107|       108|       109|  
          |SD5 V3     |SD4 V3     |SD1 V1     |  
          +-----+-----+-----+  
FieldRow4 |       110|       111|       112|  
          |SD1 V2     |SD2 V3     |SD3 V2     |  
          +-----+-----+-----+  
FieldRow5 |       113|       114|       115|  
          |SD3 V3     |SD2 V2     |SD3 V1     |  
          +-----+-----+-----+  
FieldRow6 |       201|       202|       203|  
          |SD1 V1     |SD4 V1     |SD5 V1     |  
          +-----+-----+-----+  
FieldRow7 |       204|       205|       206|  
          |SD1 V2     |SD4 V2     |SD3 V1     |  
          +-----+-----+-----+  
FieldRow8 |       207|       208|       209|  
          |SD1 V3     |SD3 V3     |SD2 V1     |  
          +-----+-----+-----+  
FieldRow9 |       210|       211|       212|  
          |SD2 V3     |SD3 V2     |SD2 V2     |  
          +-----+-----+-----+  
FieldRow10|       213|       214|       215|  
          |SD5 V2     |SD4 V3     |SD5 V3     |  
          +-----+-----+-----+  
FieldRow11|       301|       302|       303|  
          |SD5 V3     |SD4 V2     |SD2 V1     |
```

FieldRow12	304	305	306
	SD5 V1	SD5 V2	SD2 V2
FieldRow13	307	308	309
	SD3 V2	SD1 V1	SD1 V3
FieldRow14	310	311	312
	SD3 V3	SD3 V1	SD4 V3
FieldRow15	313	314	315
	SD1 V2	SD4 V1	SD2 V3
FieldRow16	401	402	403
	SD4 V2	SD5 V3	SD2 V3
FieldRow17	404	405	406
	SD3 V1	SD4 V1	SD1 V3
FieldRow18	407	408	409
	SD5 V2	SD4 V3	SD2 V1
FieldRow19	410	411	412
	SD3 V2	SD1 V1	SD5 V1
FieldRow20	413	414	415
	SD3 V3	SD2 V2	SD1 V2

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer tab is shown below:



	Trial	Block	SeedingDensity	Variety	PlotNum	FieldRow	FieldCol
1	1	1	SD4	V1	101	1	1
2	1	1	SD5	V2	102	1	2
3	1	1	SD2	V1	103	1	3
4	1	1	SD4	V2	104	2	1
5	1	1	SD1	V3	105	2	2
6	1	1	SD5	V1	106	2	3
7	1	1	SD5	V3	107	3	1
8	1	1	SD4	V3	108	3	2

Column(s): 7 Row(s): 60

Latin Square Design

The steps to generate randomization are listed below:

- To perform randomization, click **Design | Latin Square Design....** The **Randomization and Layout** dialog box will appear.
- Specify the required fields and appropriate options.

Number of Factors

Define the number of factors to be generated with default and minimum value equal to 1 and maximum value equal to 10. The number of factors specified will define the number of rows of the table inside the **Factor Definition** frame. To change the default value, user can either type a value inside the spin box or click the up-arrow key of the spin box to increase the value, down-arrow key, otherwise.

Factor Definition Frame

Name

Contain the default name for each factor which can be change. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.) ;
- succeeding characters can be a combination of letters, numbers, period (.) and underscore (_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2. The maximum levels of the factor or treatment combination should be equal to 11.

Number of Trial

Define the number of trials to be generated. The default value and minimum value is equal to one.

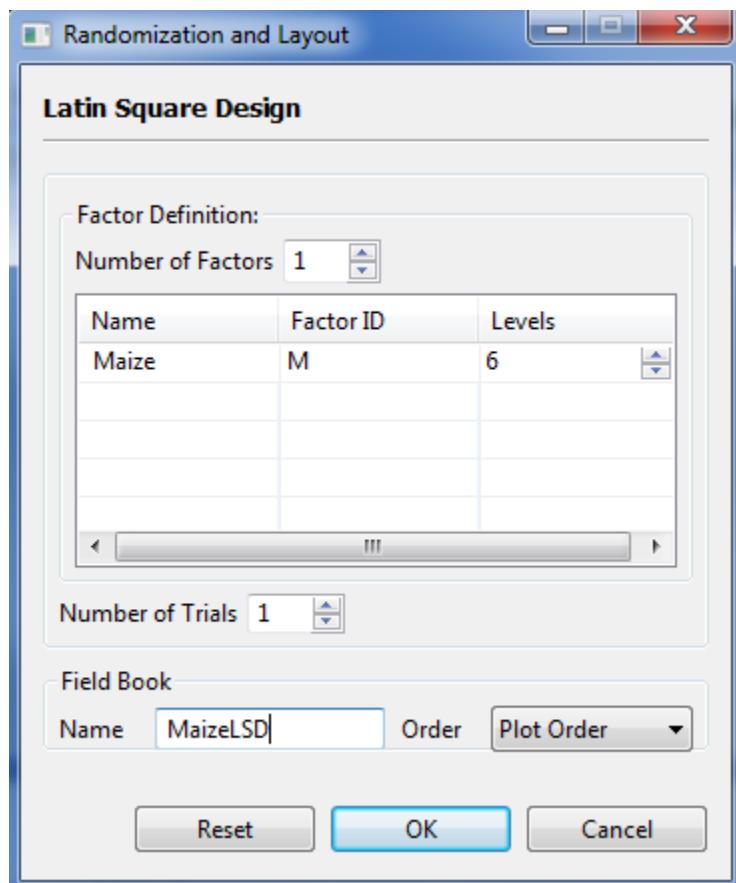
Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved. The default filename is *fieldbookLSD*.

Field Book Order

Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For the example, suppose we want to generate a randomization and layout for an experiment to be laid out in Latin Square design whose aim is to compare the grain yield of five promising maize hybrids and a check variety. The completed dialog box should appear as shown below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the *csv* data file and the *txt* file created in the Data Viewer tab.

Sample *txt* file displayed in the Data Viewer is shown below:

DESIGN PROPERTIES:
Single Factor
Latin Square Design

DESIGN PARAMETERS:
Number of Trials = 1
Treatment Name = Maize
Treatment Levels = M1, M2, M3, ..., M6

Layout for Latin Square Design

Trial = 1

1	2	3	4	5	6
M6	M3	M2	M4	M1	M5
7	8	9	10	11	12
M5	M2	M1	M3	M6	M4
13	14	15	16	17	18
M1	M4	M3	M5	M2	M6
19	20	21	22	23	24
M2	M5	M4	M6	M3	M1
25	26	27	28	29	30
M3	M6	M5	M1	M4	M2
31	32	33	34	35	36
M4	M1	M6	M2	M5	M3

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer is shown below:

	Trial	Row	Column	Maize	PlotNum
1	1	1	1	M1	1
2	1	1	2	M6	2
3	1	1	3	M4	3
4	1	1	4	M5	4
5	1	1	5	M2	5
6	1	1	6	M3	6
7	1	2	1	M3	7
8	1	2	2	M2	8

Column(s): 5 Row(s): 36

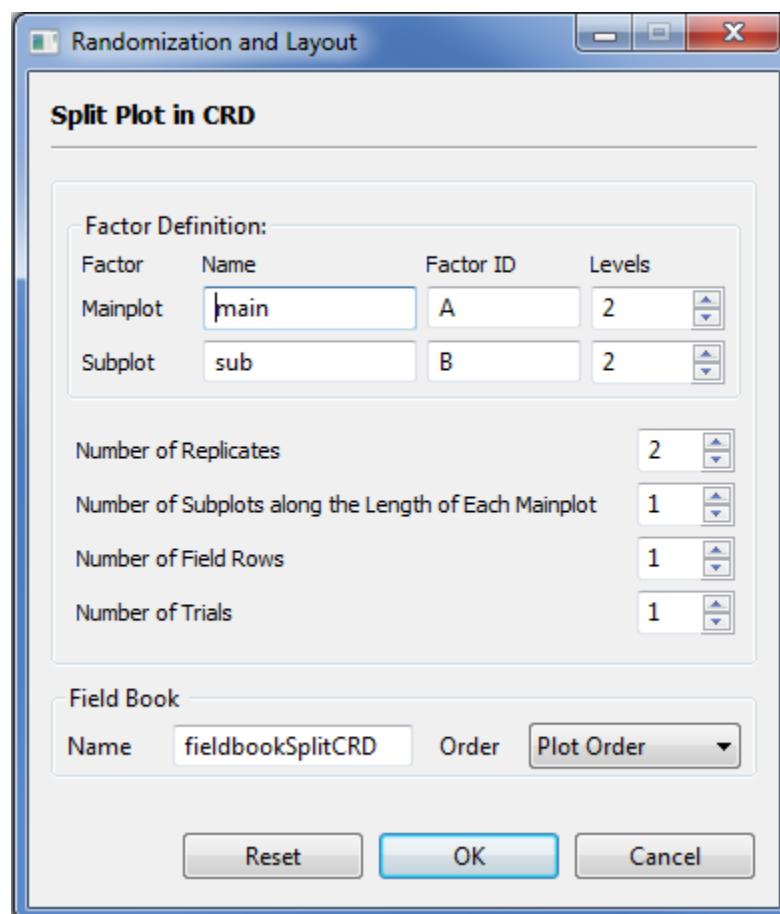
Split Plot Design

This design can be generated for experiments where the main plot is arranged in completely randomized design, randomized complete block design and Latin square design.

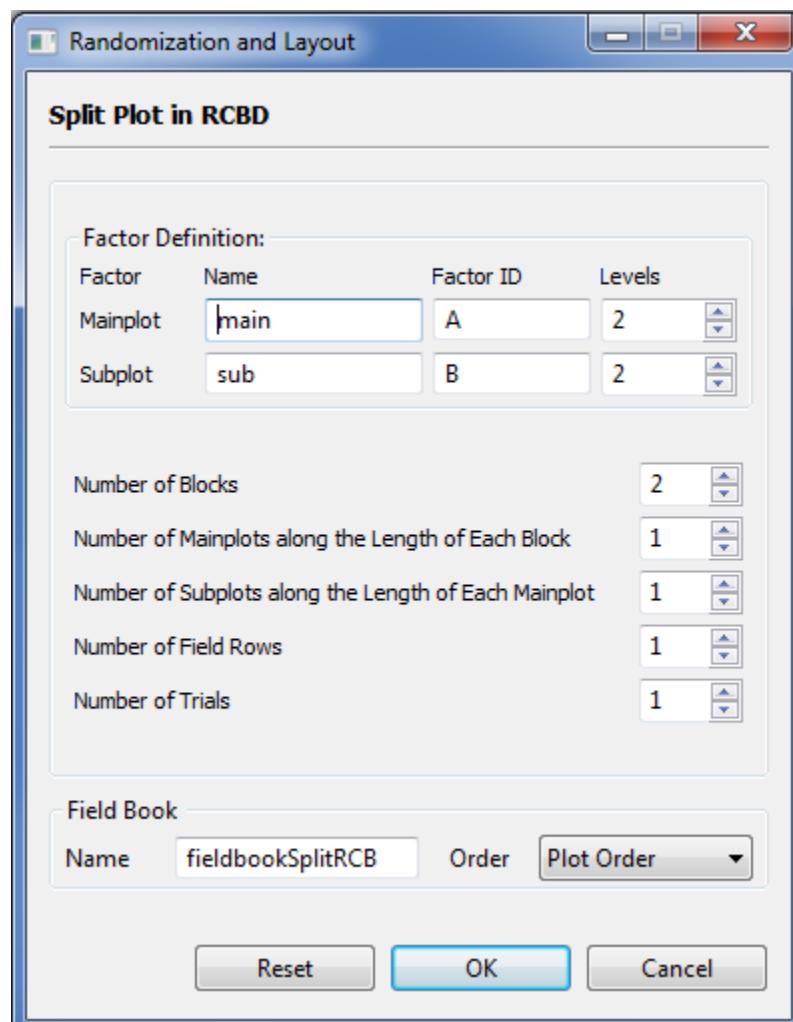
The steps to generate randomization are listed below:

- Click **Design | Split Plot Design** from the main window. Choose the appropriate design for the main plot.
- The **Randomization and Layout** dialog box will appear.

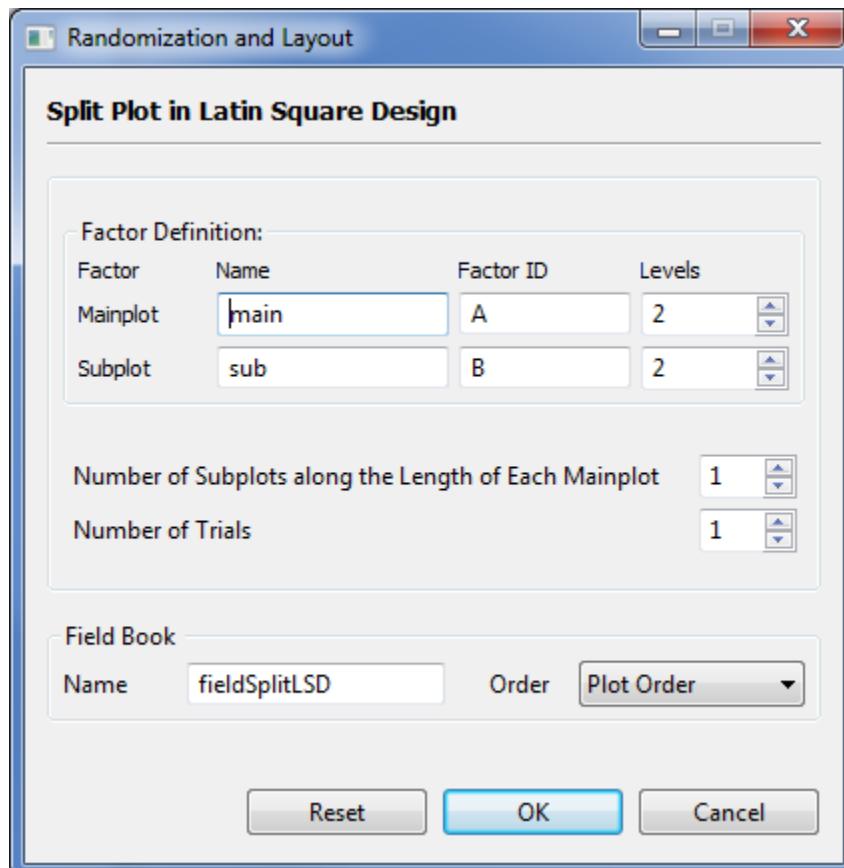
When the main plot is arranged in completely randomized design, the default **Randomization and Layout** dialog box is shown below:



When the main plot is arranged in randomized complete block design, the default **Randomization and Layout** dialog box is shown below:



When the main plot is arranged in Latin Square design, the default **Randomization and Layout** dialog box is shown below:



- Specify the required field and appropriate options.

Factor Definition Frame

Name

Contain the default name for each factor which can be change. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.) ;
- succeeding characters can be a combination of letters, numbers, period (.) and underscore (_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2 while the maximum value of the levels is 500. When the main plot is arranged in Latin Square design, the maximum value of the levels of the main plot factor is 11.

Number of Replicates

Applicable only if completely randomized design is chosen. These define the number of replicates to be generated per trial. The default and minimum value is equal to 2.

Number of Blocks

Applicable only if randomized complete block design is chosen. These define the number of blocks to be generated. The default value and minimum value is equal to 2.

Number of Mainplots along the Length of Each Block

Applicable only if randomized complete block design is chosen. This is defined as the number of mainplots along the length of each block. The default and minimum value is equal to one.

This is defined as the number of rows in each block that would receive the levels of the mainplot factor

Number of Subplots along the Length of Each Mainplot

These define the number of subplots along the length of each mainplot. The default and minimum value is equal to one.

Number of Field Rows

Applicable only if completely randomized or randomized complete block design is chosen. Define the total number of rows in the field layout for each trial. The default and minimum value is equal to one.

Number of Trial

Define the number of trials to be generated. The default value and minimum value is equal to one.

Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved.

Field Book Order

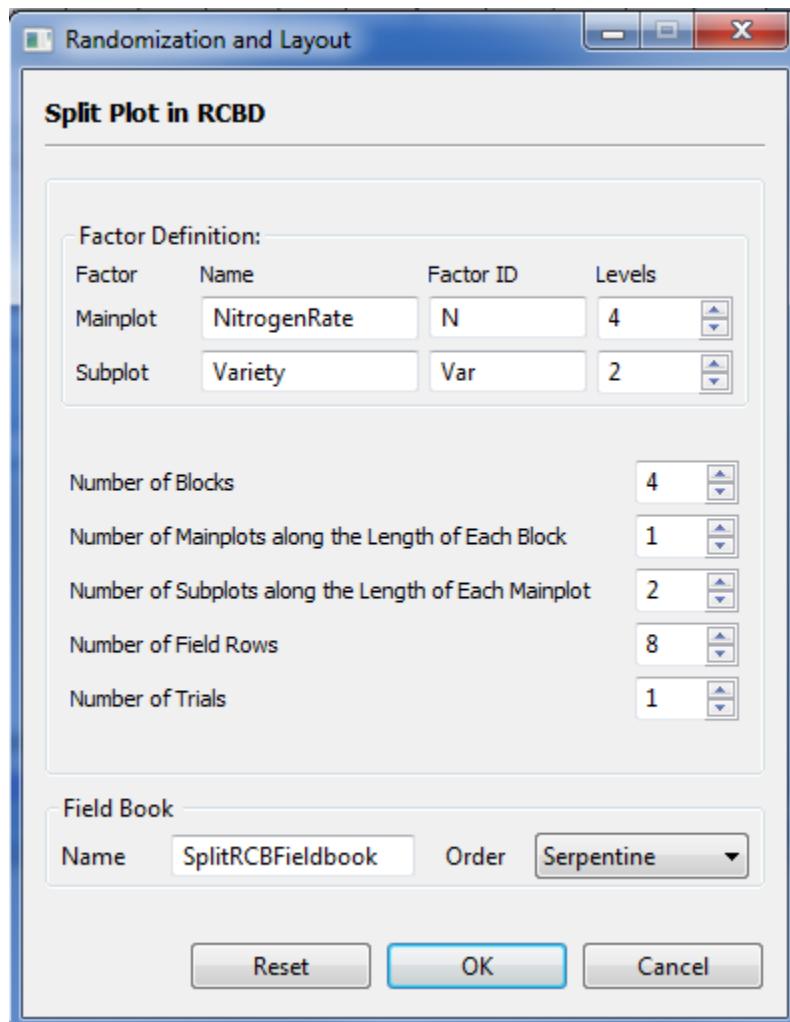
Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For this example, suppose we want to generate a randomization and layout for an experiment whose aim is to compare the effects of four nitrogen rates on grain yield of five rice varieties. The experiment will be conducted using Split Plot with four blocks. Nitrogen rates will be assigned as the mainplot factor and will be laid out in randomized complete block design; and the varieties as the subplot factor. The planned field layout is shown below:

	Field Column 1	Field Column 2	Field Column 3	Field Column 4	
Field Row 1					
Field Row 2					
Field Row 3					
Field Row 4					
Field Row 5					
Field Row 6					
Field Row 7					
Field Row 8					

The diagram illustrates a field layout with 8 rows and 4 columns. The columns are labeled "Field Column 1", "Field Column 2", "Field Column 3", and "Field Column 4". The rows are labeled "Field Row 1" through "Field Row 8". Horizontal dashed lines separate the rows. Vertical lines separate the columns. To the right of the grid, four curly braces group the rows into four blocks: Block 1 (Rows 1-2), Block 2 (Rows 3-4), Block 3 (Rows 5-6), and Block 4 (Rows 7-8).

The completed dialog box should appear as below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the csv data file and the *txt* file created in the Data Viewer tab.

Sample *txt* displayed in the Data Viewer is shown below:

DESIGN PROPERTIES:

Split Plot Design
Randomized Complete Block Design

DESIGN PARAMETERS:

Number of Trials = 1
Number of Blocks = 4
Mainplot Factor = NitrogenRate
Levels = N1, N2, N3, N4

Subplot Factor = Variety
Levels = Var1, Var2

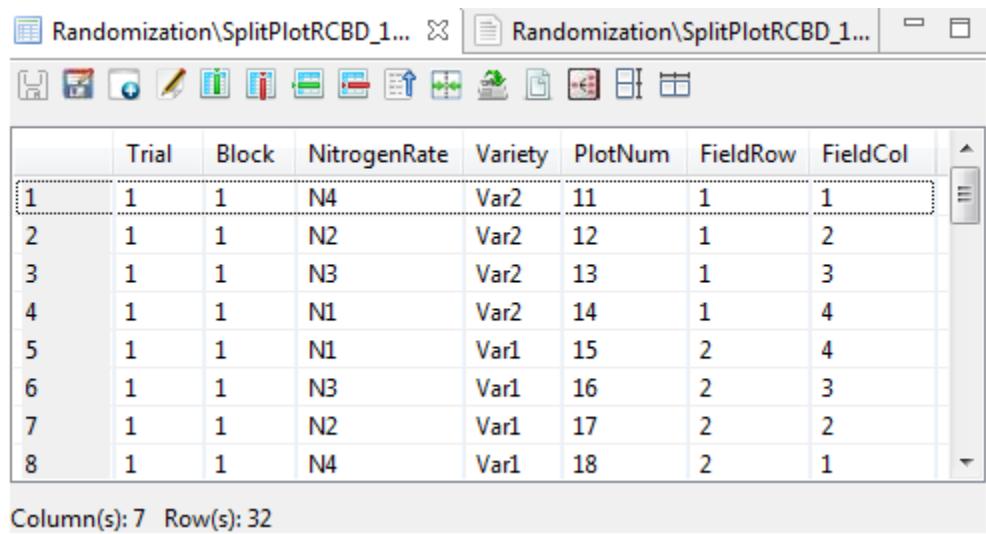
Layout for Split Plot in Randomized Complete Block Design:

Trial = 1

	FieldCol1	FieldCol2	FieldCol3	FieldCol4
FieldRow1	11 N4 Var2	12 N2 Var2	13 N3 Var2	14 N1 Var2
FieldRow2	18 N4 Var1	17 N2 Var1	16 N3 Var1	15 N1 Var1
FieldRow3	21 N3 Var2	22 N4 Var2	23 N2 Var1	24 N1 Var2
FieldRow4	28 N3 Var1	27 N4 Var1	26 N2 Var2	25 N1 Var1
FieldRow5	31 N4 Var1	32 N2 Var2	33 N1 Var2	34 N3 Var2
FieldRow6	38 N4 Var2	37 N2 Var1	36 N1 Var1	35 N3 Var1
FieldRow7	41 N2 Var1	42 N3 Var1	43 N4 Var2	44 N1 Var2
FieldRow8	48 N2 Var2	47 N3 Var2	46 N4 Var1	45 N1 Var1

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer is shown below:



The screenshot shows a software window titled "Randomization\SplitPlotRCBD_1...". The main area is a Data Viewer containing a table with 8 rows and 7 columns. The columns are labeled: Trial, Block, NitrogenRate, Variety, PlotNum, FieldRow, and FieldCol. The data entries are as follows:

	Trial	Block	NitrogenRate	Variety	PlotNum	FieldRow	FieldCol
1	1	1	N4	Var2	11	1	1
2	1	1	N2	Var2	12	1	2
3	1	1	N3	Var2	13	1	3
4	1	1	N1	Var2	14	1	4
5	1	1	N1	Var1	15	2	4
6	1	1	N3	Var1	16	2	3
7	1	1	N2	Var1	17	2	2
8	1	1	N4	Var1	18	2	1

Column(s): 7 Row(s): 32

Strip Plot Design

The steps to generate randomization are listed below:

- Choose **Design | Strip Plot Design...** from the main window. The **Randomization and Layout** dialog box will appear.
- Specify the required fields and appropriate options.

Factor Definition Frame

Name

Contain the default name for each factor which can be change. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.);
- succeeding characters can be a combination of letters, numbers, period (.) and underscore (_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2 while the maximum value of the levels is 500.

Number of Blocks

These define the number of blocks to be generated. The default value and minimum value is equal to 2.

Number of Field Rows

Define the total number of rows in the field layout for each trial. The default and minimum value is equal to 2.

Number of Trial

Define the number of trials to be generated. The default value and minimum value is equal to one.

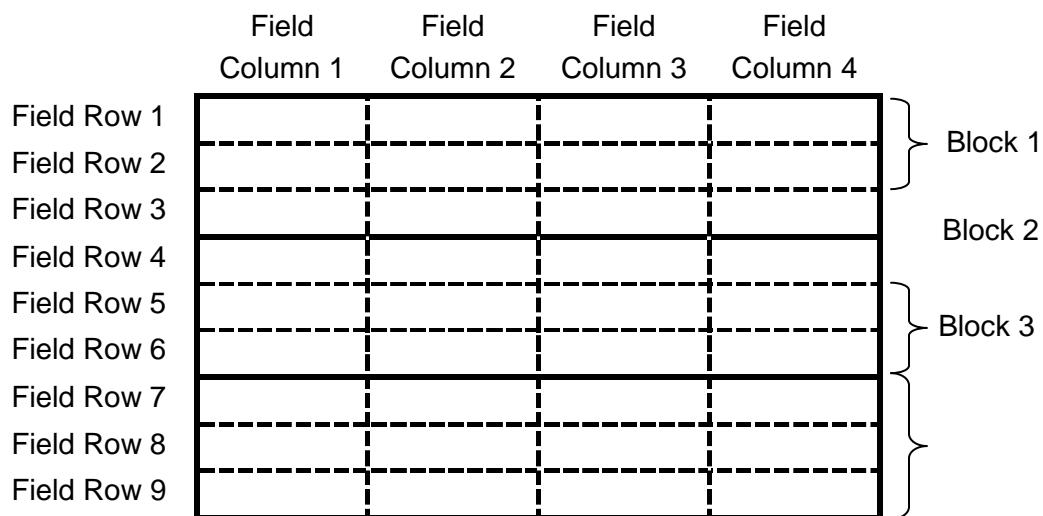
Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved. The default filename is *fieldbookStrip*.

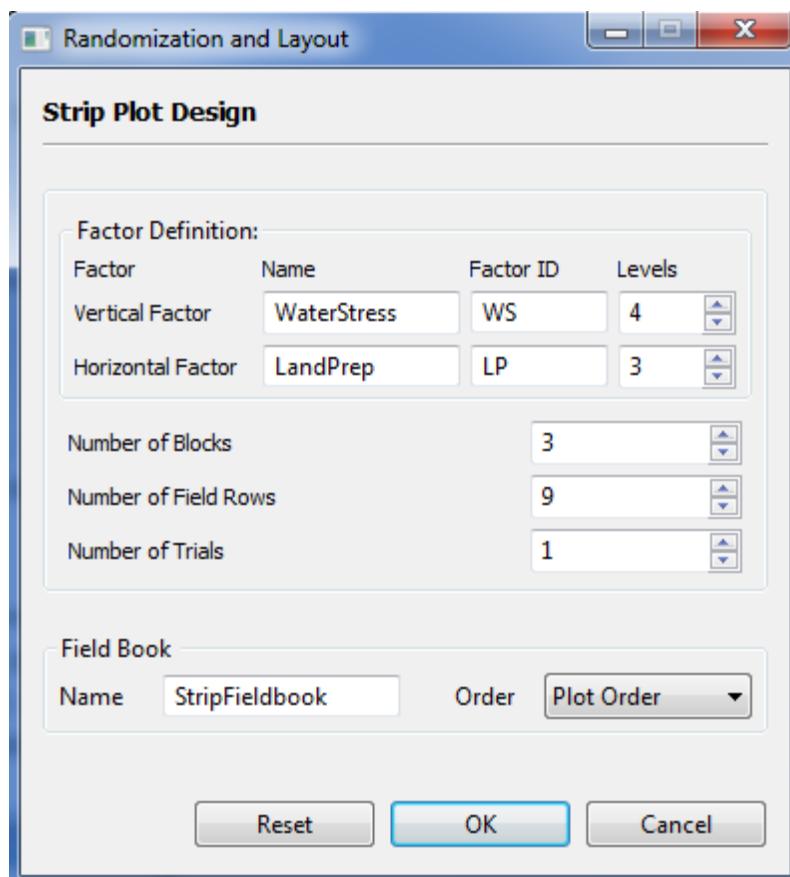
Field Book Order

Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For this example, suppose we want to generate a randomization and layout for an experiment to be conducted using Strip Plot, with land preparation (with 3 levels) as the horizontal factor and water stress (with four levels) as the vertical factor. Three blocks will be used. The planned field layout is shown below:



The completed dialog box should appear as shown below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the csv data file and txt file created in the Data Viewer tab.

Sample *txt* file displayed in the Data Viewer is shown below:

DESIGN PROPERTIES:

Strip Plot Design

DESIGN PARAMETERS:

Number of Trials = 1
Number of Blocks = 3
Vertical Factor = WaterStress
Levels = WS1, WS2, WS3, WS4
Horizontal Factor = LandPrep
Levels = LP1, LP2, LP3

Layout for Strip Plot:

Trial = 1

	FieldCol1	FieldCol2	FieldCol3	FieldCol4
FieldRow1	101	102	103	104
	WS3 LP1	WS2 LP1	WS4 LP1	WS1 LP1
FieldRow2	105	106	107	108
	WS3 LP3	WS2 LP3	WS4 LP3	WS1 LP3
FieldRow3	109	110	111	112
	WS3 LP2	WS2 LP2	WS4 LP2	WS1 LP2
FieldRow4	201	202	203	204
	WS4 LP3	WS1 LP3	WS3 LP3	WS2 LP3
FieldRow5	205	206	207	208
	WS4 LP1	WS1 LP1	WS3 LP1	WS2 LP1
FieldRow6	209	210	211	212
	WS4 LP2	WS1 LP2	WS3 LP2	WS2 LP2
FieldRow7	301	302	303	304
	WS3 LP1	WS1 LP1	WS4 LP1	WS2 LP1
FieldRow8	305	306	307	308
	WS3 LP3	WS1 LP3	WS4 LP3	WS2 LP3
FieldRow9	309	310	311	312
	WS3 LP2	WS1 LP2	WS4 LP2	WS2 LP2

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer is shown below:

	Trial	Block	WaterStress	LandPrep	PlotNum	FieldRow	FieldCol
1	1	1	WS3	LP1	101	1	1
2	1	1	WS2	LP1	102	1	2
3	1	1	WS4	LP1	103	1	3
4	1	1	WS1	LP1	104	1	4
5	1	1	WS3	LP3	105	2	1
6	1	1	WS2	LP3	106	2	2
7	1	1	WS4	LP3	107	2	3
8	1	1	WS1	LP3	108	2	4

Column(s): 7 Row(s): 36

Split-Split Plot Design

This design can be generated for experiments where the main plot is arranged in completely randomized design, randomized complete block design and Latin square design.

The steps to generate randomization are listed below:

- Click **Design | Split-Split Plot Design** from the main window. Choose the appropriate design for the main plot.
- The appropriate **Randomization and Layout** dialog box will appear depending on how the main plot will be arranged.
- Specify the required field and appropriate options.

Factor Definition Frame

Name

Contain the default name for each factor which can be change. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.) ;
- succeeding characters can be a combination of letters, numbers, period (.) and underscore (_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2 while the maximum value of the levels is 500. When the main plot is arranged in Latin Square design, the maximum value of the levels of the main plot factor is 11.

Number of Replicates

Applicable only if completely randomized design is chosen. These define the number of replicates to be generated. The default and minimum value is equal to 2.

Number of Blocks

Applicable only if randomized complete block design is chosen. These define the number of blocks to be generated. The default and minimum value is equal to 2.

Number of Mainplots along the Length of Each Block

Applicable only if randomized complete block design is chosen. These define the number of main plots along the length of each block. The default and minimum value is equal to one.

Number of Subplots along the Length of Each Mainplot

Define the number of subplots along the length of each mainplot. The default and minimum value is equal to one.

Number of Sub-subplots along the Length of Each Subplot

Define the number of sub-subplots along the length of each subplot. The default and minimum value is equal to one.

Number of Field Rows

Applicable only if completely randomized or randomized complete block design is chosen. Define the total number of rows in the field layout for each trial. The default and minimum value is equal to one.

Number of Trial

Define the number of trials to be generated. The default value and minimum value is equal to one.

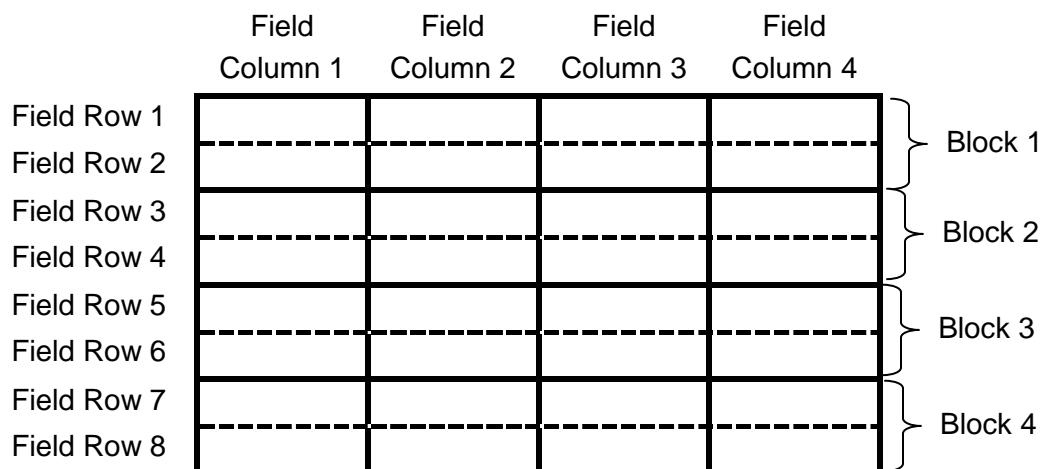
Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved.

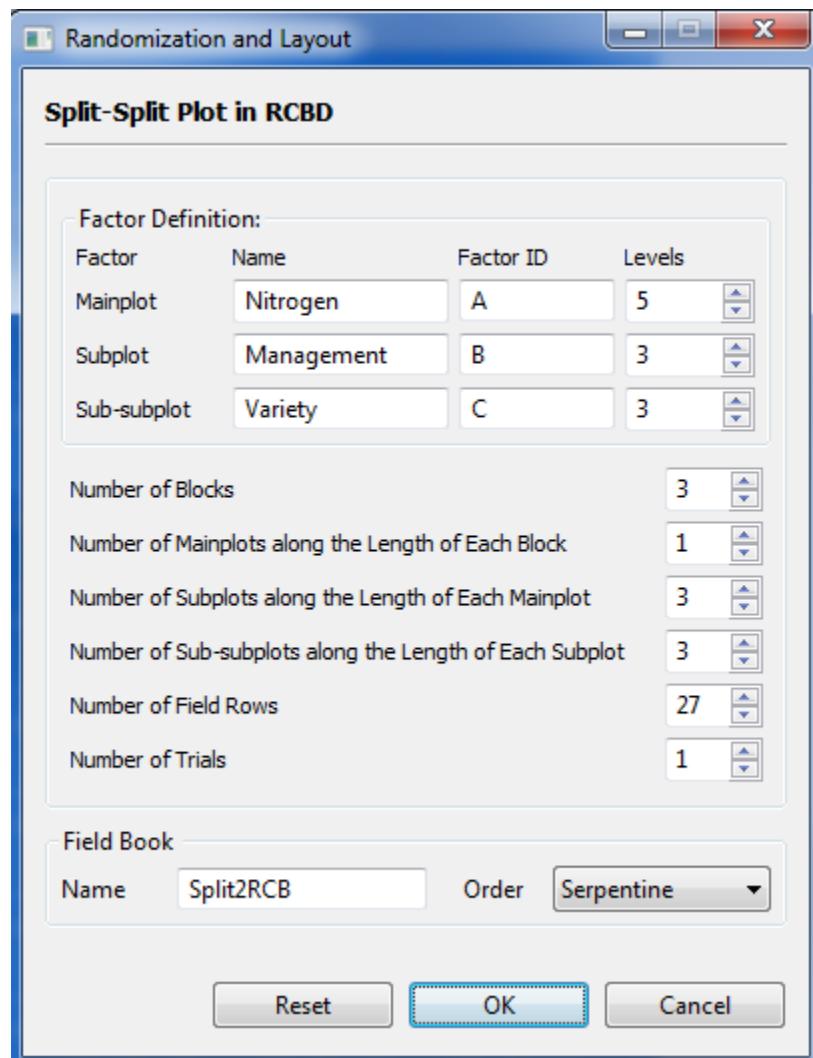
Field Book Order

Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For this example, suppose we want to generate a randomization and layout for an experiment whose aim is to compare the effects of five nitrogen rates and three management practices on the grain yield of three rice varieties which will be conducted using Split-Split Plot with three blocks. Nitrogen rates will be assigned as the mainplot factor and will be arranged in randomized complete block design, management practices as the subplot factor and varieties as the sub-subplot factor. The planned field layout is shown below:



The completed dialog box should appear as shown below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the csv data file created in the Data Viewer.

Sample *txt* data file displayed in the Data Viewer is shown below:

DESIGN PROPERTIES:
 Split-Split Plot Design
 Randomized Complete Block Design

DESIGN PARAMETERS:
 Number of Trials = 1
 Number of Blocks = 3
 Mainplot Factor = Nitrogen
 Levels = A1, A2, A3, A4, A5
 Subplot Factor = Management
 Levels = B1, B2, B3
 Sub-subplot Factor = Variety
 Levels = C1, C2, C3

Layout for Split-Split Plot in Randomized Complete Block Design:

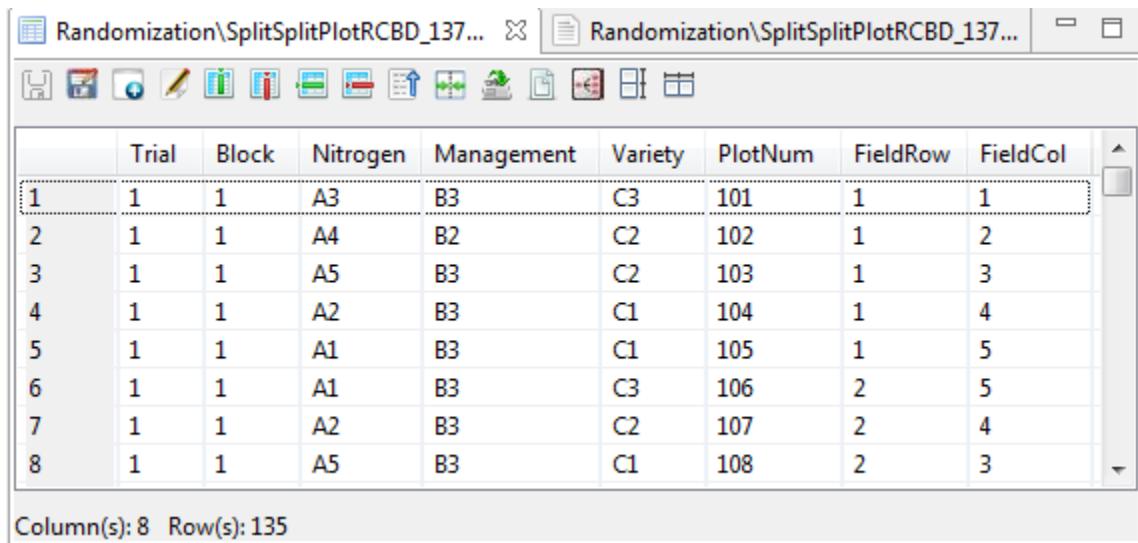
Trial = 1

	FieldCol1	FieldCol2	FieldCol3	FieldCol4	FieldCol5
FieldRow1	101	102	103	104	105
FieldRow2	A3 B3 C3	A4 B2 C2	A5 B3 C2	A2 B3 C1	A1 B3 C1
FieldRow3	111	112	113	114	115
FieldRow4	A3 B3 C1	A4 B2 C3	A5 B3 C3	A2 B3 C3	A1 B3 C2
FieldRow5	120	119	118	117	116
FieldRow6	A3 B2 C2	A4 B3 C2	A5 B2 C2	A2 B1 C3	A1 B1 C3
FieldRow7	121	122	123	124	125
FieldRow8	A3 B2 C1	A4 B3 C1	A5 B2 C3	A2 B1 C1	A1 B1 C1
FieldRow9	130	129	128	127	126
FieldRow10	A3 B2 C3	A4 B3 C3	A5 B2 C1	A2 B1 C2	A1 B1 C2
FieldRow11	131	132	133	134	135
FieldRow12	A3 B1 C2	A4 B1 C3	A5 B1 C2	A2 B2 C3	A1 B2 C3
FieldRow13	140	139	138	137	136
FieldRow14	A3 B1 C1	A4 B1 C2	A5 B1 C3	A2 B2 C2	A1 B2 C2
FieldRow15	141	142	143	144	145
FieldRow16	A3 B1 C3	A4 B1 C1	A5 B1 C1	A2 B2 C1	A1 B2 C1
FieldRow17	201	202	203	204	205
FieldRow18	A1 B1 C1	A3 B2 C3	A5 B2 C2	A2 B2 C3	A4 B1 C1
FieldRow19	210	209	208	207	206
FieldRow20	A1 B1 C3	A3 B2 C2	A5 B2 C1	A2 B2 C1	A4 B1 C2

FieldRow12	211	212	213	214	215
	A1 B1 C2	A3 B2 C1	A5 B2 C3	A2 B2 C2	A4 B1 C3
FieldRow13	220	219	218	217	216
	A1 B2 C3	A3 B3 C1	A5 B1 C3	A2 B3 C1	A4 B2 C2
FieldRow14	221	222	223	224	225
	A1 B2 C1	A3 B3 C2	A5 B1 C1	A2 B3 C3	A4 B2 C3
FieldRow15	230	229	228	227	226
	A1 B2 C2	A3 B3 C3	A5 B1 C2	A2 B3 C2	A4 B2 C1
FieldRow16	231	232	233	234	235
	A1 B3 C3	A3 B1 C3	A5 B3 C3	A2 B1 C2	A4 B3 C3
FieldRow17	240	239	238	237	236
	A1 B3 C2	A3 B1 C2	A5 B3 C2	A2 B1 C3	A4 B3 C2
FieldRow18	241	242	243	244	245
	A1 B3 C1	A3 B1 C1	A5 B3 C1	A2 B1 C1	A4 B3 C1
FieldRow19	301	302	303	304	305
	A2 B3 C3	A4 B1 C2	A1 B3 C1	A3 B2 C2	A5 B2 C3
FieldRow20	310	309	308	307	306
	A2 B3 C1	A4 B1 C3	A1 B3 C2	A3 B2 C3	A5 B2 C2
FieldRow21	311	312	313	314	315
	A2 B3 C2	A4 B1 C1	A1 B3 C3	A3 B2 C1	A5 B2 C1
FieldRow22	320	319	318	317	316
	A2 B2 C1	A4 B2 C1	A1 B2 C1	A3 B1 C3	A5 B3 C3
FieldRow23	321	322	323	324	325
	A2 B2 C2	A4 B2 C3	A1 B2 C2	A3 B1 C1	A5 B3 C1
FieldRow24	330	329	328	327	326
	A2 B2 C3	A4 B2 C2	A1 B2 C3	A3 B1 C2	A5 B3 C2
FieldRow25	331	332	333	334	335
	A2 B1 C1	A4 B3 C2	A1 B1 C3	A3 B3 C1	A5 B1 C2
FieldRow26	340	339	338	337	336
	A2 B1 C3	A4 B3 C1	A1 B1 C2	A3 B3 C3	A5 B1 C3
FieldRow27	341	342	343	344	345
	A2 B1 C2	A4 B3 C3	A1 B1 C1	A3 B3 C2	A5 B1 C1

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer is shown below:



The screenshot shows a software interface titled "Randomization\SplitSplitPlotRCBD_137...". The main area displays a data grid with 8 columns and 135 rows. The columns are labeled: Trial, Block, Nitrogen, Management, Variety, PlotNum, FieldRow, and FieldCol. The first few rows of data are as follows:

Trial	Block	Nitrogen	Management	Variety	PlotNum	FieldRow	FieldCol
1	1	A3	B3	C3	101	1	1
2	1	A4	B2	C2	102	1	2
3	1	A5	B3	C2	103	1	3
4	1	A2	B3	C1	104	1	4
5	1	A1	B3	C1	105	1	5
6	1	A1	B3	C3	106	2	5
7	1	A2	B3	C2	107	2	4
8	1	A5	B3	C1	108	2	3

Column(s): 8 Row(s): 135

Strip-Split Plot Design

The steps to generate randomization are listed below:

- Choose **Design | Strip-Split Plot Design...** from the main window. The **Randomization** dialog box will appear.
- Specify the required fields and appropriate options.

Factor Definition

Name

Contain the default name for each factor. User can change the factor name by typing the desired factor name. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.) ;
- succeeding characters can be a combination of letters, numbers, period (.) and underscore (_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2 while the maximum value of the levels is 500.

Number of Blocks

Define the number of blocks to be generated. The default value and minimum value is equal to 2.

Number of Subplots along the Length of Each VxH plot

These define the number of subplots along the length of VxH plot. The default and minimum value is equal to one.

Number of Field Rows

Define the total number of rows in the field layout for each trial. The default and minimum value is equal to 2.

Number of Trial

Define the number of trials to be generated. The default value and minimum value is equal to one.

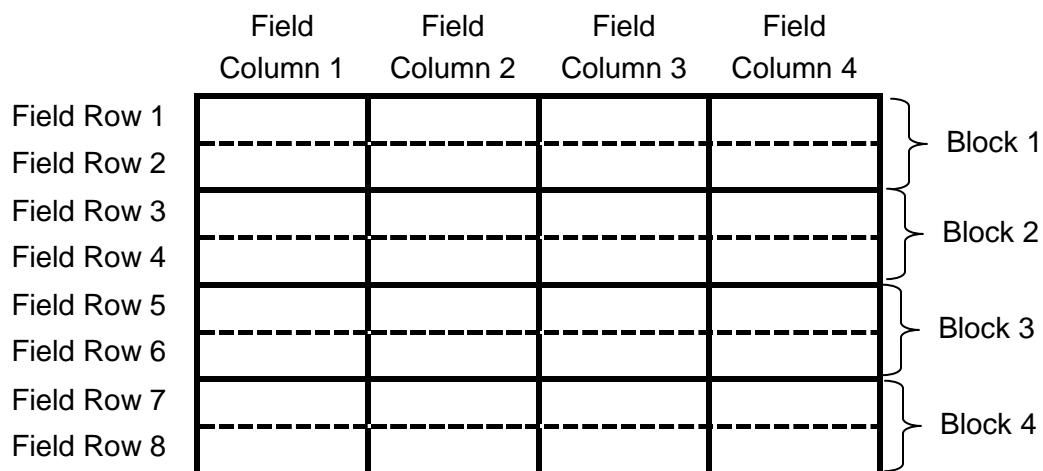
Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved. The default filename is *fieldbookStripSplit*.

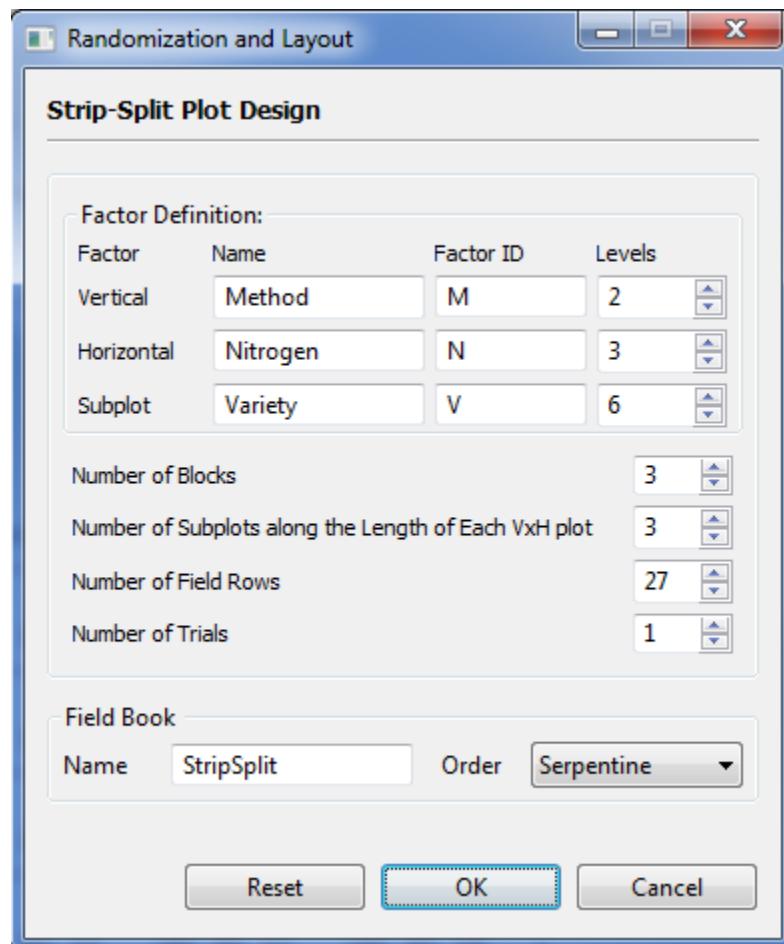
Field Book Order

Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For this example, suppose we want to generate a randomization for an experiment to be conducted using Strip-Split Plot, with planting methods (with 2 levels) as vertical factor, nitrogen rates (with 3 levels) as the horizontal factor and rice variety (with 6 levels) as the subplot factor. Three blocks will be used. The planned field layout is shown below:



The completed dialog box should appear as shown below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the csv data file and the *txt* file created in the Data Viewer.

Sample *txt* file displayed in the Data Viewer is shown below:

```
DESIGN PROPERTIES:
    Strip-Split Plot Design

DESIGN PARAMETERS:
    Number of Trials = 1
    Number of Blocks = 3
    Vertical Factor = Method
    Levels = M1, M2
    Horizontal Factor = Nitrogen
    Levels = N1, N2, N3
    Subplot Factor = Variety
    Levels = V1, V2, V3, ..., V6
```

Layout for Strip-Split Plot

Trial = 1

	FieldCol1	FieldCol2	FieldCol3	FieldCol4
FieldRow1	101	102	103	104
	M2 N1 V1	M2 N1 V3	M1 N1 V6	M1 N1 V4
FieldRow2	108	107	106	105
	M2 N1 V2	M2 N1 V4	M1 N1 V2	M1 N1 V1
FieldRow3	109	110	111	112
	M2 N1 V6	M2 N1 V5	M1 N1 V5	M1 N1 V3
FieldRow4	116	115	114	113
	M2 N3 V2	M2 N3 V3	M1 N3 V6	M1 N3 V1
FieldRow5	117	118	119	120
	M2 N3 V6	M2 N3 V5	M1 N3 V3	M1 N3 V5
FieldRow6	124	123	122	121
	M2 N3 V1	M2 N3 V4	M1 N3 V4	M1 N3 V2
FieldRow7	125	126	127	128
	M2 N2 V2	M2 N2 V5	M1 N2 V1	M1 N2 V3
FieldRow8	132	131	130	129
	M2 N2 V3	M2 N2 V4	M1 N2 V6	M1 N2 V4
FieldRow9	133	134	135	136
	M2 N2 V1	M2 N2 V6	M1 N2 V2	M1 N2 V5
FieldRow10	201	202	203	204

	M1 N3 V1	M1 N3 V2	M2 N3 V6	M2 N3 V3	
FieldRow11	208	207	206	205	
	M1 N3 V4	M1 N3 V5	M2 N3 V2	M2 N3 V4	
FieldRow12	209	210	211	212	
	M1 N3 V6	M1 N3 V3	M2 N3 V1	M2 N3 V5	
FieldRow13	216	215	214	213	
	M1 N1 V2	M1 N1 V1	M2 N1 V6	M2 N1 V2	
FieldRow14	217	218	219	220	
	M1 N1 V3	M1 N1 V5	M2 N1 V5	M2 N1 V3	
FieldRow15	224	223	222	221	
	M1 N1 V4	M1 N1 V6	M2 N1 V4	M2 N1 V1	
FieldRow16	225	226	227	228	
	M1 N2 V4	M1 N2 V2	M2 N2 V3	M2 N2 V4	
FieldRow17	232	231	230	229	
	M1 N2 V3	M1 N2 V6	M2 N2 V1	M2 N2 V5	
FieldRow18	233	234	235	236	
	M1 N2 V5	M1 N2 V1	M2 N2 V6	M2 N2 V2	
FieldRow19	301	302	303	304	
	M1 N1 V5	M1 N1 V3	M2 N1 V6	M2 N1 V2	
FieldRow20	308	307	306	305	
	M1 N1 V4	M1 N1 V1	M2 N1 V4	M2 N1 V3	
FieldRow21	309	310	311	312	
	M1 N1 V2	M1 N1 V6	M2 N1 V1	M2 N1 V5	
FieldRow22	316	315	314	313	
	M1 N2 V1	M1 N2 V2	M2 N2 V2	M2 N2 V5	
FieldRow23	317	318	319	320	
	M1 N2 V5	M1 N2 V6	M2 N2 V1	M2 N2 V6	
FieldRow24	324	323	322	321	
	M1 N2 V3	M1 N2 V4	M2 N2 V3	M2 N2 V4	
FieldRow25	325	326	327	328	
	M1 N3 V6	M1 N3 V2	M2 N3 V2	M2 N3 V4	
FieldRow26	332	331	330	329	
	M1 N3 V1	M1 N3 V4	M2 N3 V3	M2 N3 V1	
FieldRow27	333	334	335	336	
	M1 N3 V3	M1 N3 V5	M2 N3 V5	M2 N3 V6	

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer is shown below:

	Trial	Block	Method	Nitrogen	Variety	PlotNum	FieldRow	FieldCol	
1	1	1	M2	N1	V1	1	1	1	
2	1	1	M2	N1	V3	2	1	2	
3	1	1	M1	N1	V6	3	1	3	
4	1	1	M1	N1	V4	4	1	4	
5	1	1	M2	N1	V2	5	2	1	
6	1	1	M2	N1	V4	6	2	2	
7	1	1	M1	N1	V2	7	2	3	
8	1	1	M1	N1	V1	8	2	4	
9	1	1	M2	N1	V6	9	3	1	

Column(s): 8 Row(s): 108

Split-Split-Split Plot Design

This design can be generated for experiments where the main plot is arranged in completely randomized, randomized complete block and Latin square design.

The steps to generate randomization are listed below:

- Click **Design | Split-Split-Split Plot Design** from the main window. Choose the appropriate design for the main plot.
- The appropriate **Randomization and Layout** dialog box will appear depending on how the main plot will be arranged.
- Specify the required field and appropriate options.

Factor Definition Frame

Name

Contain the default name for each factor which can be change. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.);

- succeeding characters can be a combination of letters, numbers, period(.) and underscore(_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2 while the maximum value of the levels is 500. When the main plot is arranged in Latin Square design, the maximum value of the levels of the main plot factor is 11.

Number of Replicates

Applicable only if completely randomized design is chosen. These define the number of replicates to be generated. The default and minimum value is equal to 2.

Number of Blocks

Applicable only if randomized complete block design is chosen. These define the number of blocks to be generated. The default and minimum value is equal to 2.

Number of Mainplots along the Length of Each Block

Applicable only if randomized complete block design is chosen. These define the number of main plots along the length of each block. The default and minimum value is equal to one.

Number of Subplots along the Length of Each Mainplot

Define the number of subplots along the length of each main plot. The default and minimum value is equal to one.

Number of Sub-subplots along the Length of Each Subplot

Define the number of sub-subplots along the length of each subplot. The default and minimum value is equal to one.

Number of Sub-sub-subplots along the Length of Each Sub-subplot

Define the number of sub-sub-subplots along the length of each sub-subplot. The default and minimum value is equal to one.

Number of Field Rows

Applicable only if completely randomized or randomized complete block design is chosen. Define the total number of rows in the field layout for each trial. The default and minimum value is equal to one.

Number of Trial

Define the number of trials to be generated. The default value and minimum value is equal to one.

Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved.

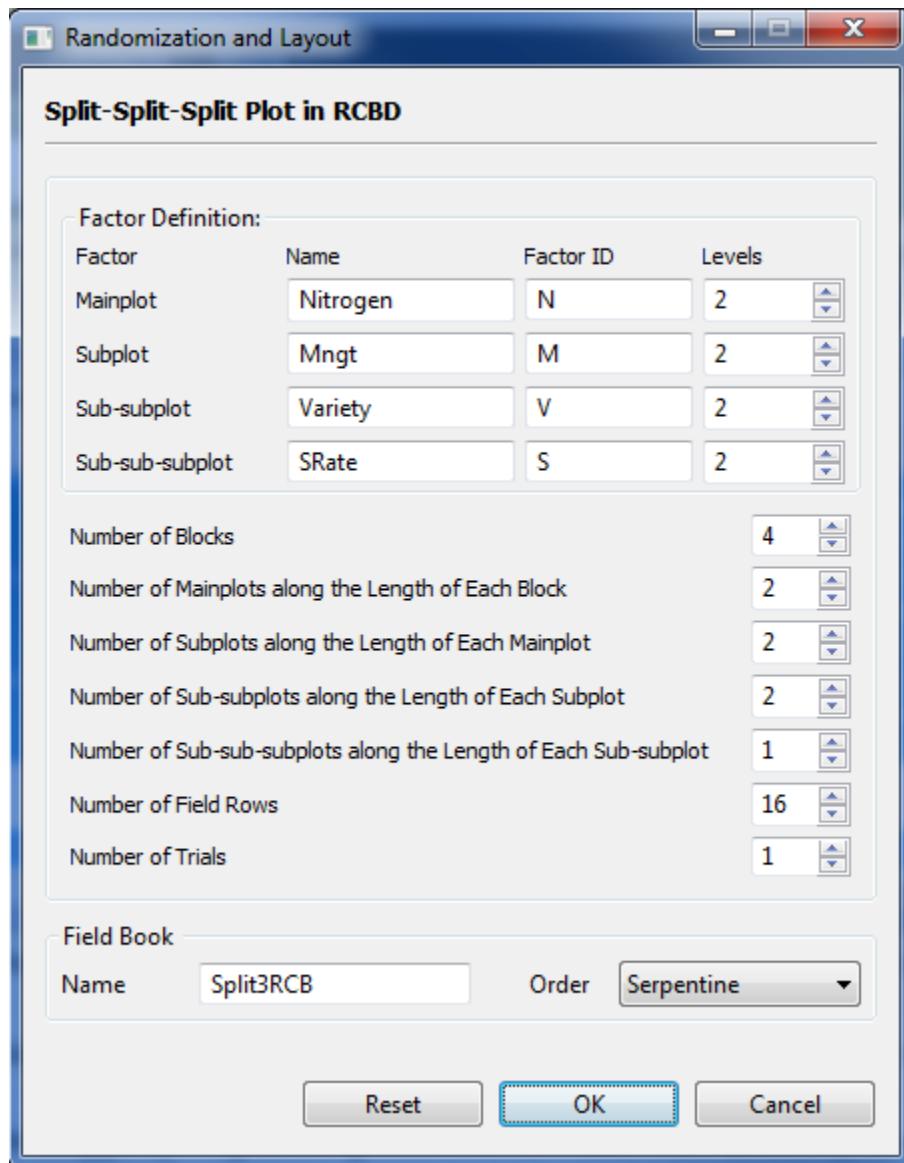
Field Book Order

Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For this example, suppose we want to generate a randomization for an experiment whose aim is to compare the effects of two nitrogen rates, two management practices and two seeding rate on the grain yield of two rice varieties which will be conducted using Split-Split-Split Plot with four blocks. Nitrogen rates will be assigned as the mainplot factor and will be arranged in randomized complete block design, management practices as the subplot factor, varieties as the sub-subplot factor and seeding rate as the sub-sub-subplot factor. The planned field layout is shown below:

	Field Column 1	Field Column 2	Field Column 3	Field Column 4	
Field Row 1					Block 1
Field Row 2					
Field Row 3					
Field Row 4					
Field Row 5					
Field Row 6					
Field Row 7					
Field Row 8					

The completed dialog box should appear as shown below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the csv data file and the *txt* file created in the Data Viewer tab.

Sample *txt* file displayed in the Data Viewer is shown below:

```
DESIGN PROPERTIES:  
Split-Split-Split Plot Design  
Randomized Complete Block Design
```

DESIGN PARAMETERS:

Number of Trials = 1
 Number of Blocks = 4
 Mainplot Factor = Nitrogen
 Levels = N1, N2
 Subplot Factor = Mngt
 Levels = M1, M2
 Sub-subplot Factor = Variety
 Levels = V1, V2
 Sub-sub-subplot Factor = SRate
 Levels = S1, S2

Layout for Split-Split-Split Plot in Randomized Complete Block Design:

Trial = 1

	FieldCol1	FieldCol2	FieldCol3	FieldCol4
FieldRow1	101 N2 M2 V2 S1	102 N2 M2 V2 S2	201 N1 M1 V1 S1	202 N1 M1 V1 S2
FieldRow2	104 N2 M2 V1 S1	103 N2 M2 V1 S2	204 N1 M1 V2 S2	203 N1 M1 V2 S1
FieldRow3	105 N2 M1 V1 S2	106 N2 M1 V1 S1	205 N1 M2 V1 S2	206 N1 M2 V1 S1
FieldRow4	108 N2 M1 V2 S1	107 N2 M1 V2 S2	208 N1 M2 V2 S2	207 N1 M2 V2 S1
FieldRow5	109 N1 M2 V2 S2	110 N1 M2 V2 S1	209 N2 M1 V2 S1	210 N2 M1 V2 S2
FieldRow6	112 N1 M2 V1 S2	111 N1 M2 V1 S1	212 N2 M1 V1 S1	211 N2 M1 V1 S2
FieldRow7	113 N1 M1 V1 S2	114 N1 M1 V1 S1	213 N2 M2 V2 S1	214 N2 M2 V2 S2
FieldRow8	116 N1 M1 V2 S1	115 N1 M1 V2 S2	216 N2 M2 V1 S2	215 N2 M2 V1 S1
FieldRow9	301 N1 M2 V2 S2	302 N1 M2 V2 S1	401 N2 M2 V2 S1	402 N2 M2 V2 S2
FieldRow10	304 N1 M2 V1 S2	303 N1 M2 V1 S1	404 N2 M2 V1 S1	403 N2 M2 V1 S2
FieldRow11	305 N1 M1 V2 S1	306 N1 M1 V2 S2	405 N2 M1 V1 S1	406 N2 M1 V1 S2
FieldRow12	308 N1 M1 V1 S1	307 N1 M1 V1 S2	408 N2 M1 V2 S2	407 N2 M1 V2 S1
FieldRow13	309 N2 M1 V2 S2	310 N2 M1 V2 S1	409 N1 M1 V1 S1	410 N1 M1 V1 S2
FieldRow14	312 N2 M1 V1 S2	311 N2 M1 V1 S1	412 N1 M1 V2 S1	411 N1 M1 V2 S2
FieldRow15	313	314	413	414

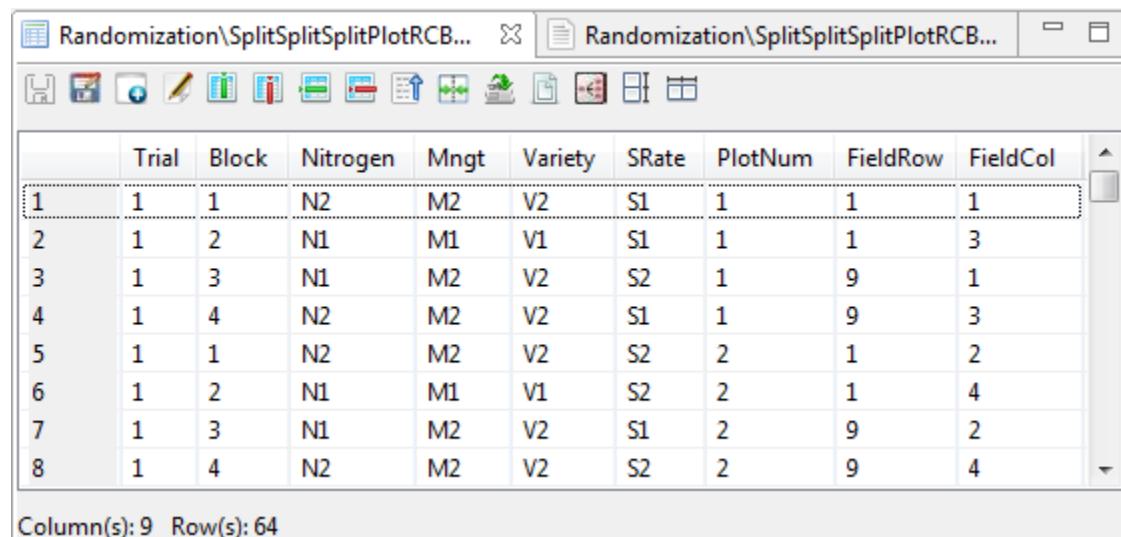
```

|N2 M2 V1 S2 |N2 M2 V1 S1 |N1 M2 V1 S1 |N1 M2 V1 S2 |
+-----+-----+-----+-----+
FieldRow16 |           316|       315|      416|      415|
|N2 M2 V2 S2 |N2 M2 V2 S1 |N1 M2 V2 S1 |N1 M2 V2 S2 |
+-----+-----+-----+-----+

```

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer is shown below:



	Trial	Block	Nitrogen	Mngt	Variety	SRate	PlotNum	FieldRow	FieldCol
1	1	1	N2	M2	V2	S1	1	1	1
2	1	2	N1	M1	V1	S1	1	1	3
3	1	3	N1	M2	V2	S2	1	9	1
4	1	4	N2	M2	V2	S1	1	9	3
5	1	1	N2	M2	V2	S2	2	1	2
6	1	2	N1	M1	V1	S2	2	1	4
7	1	3	N1	M2	V2	S1	2	9	2
8	1	4	N2	M2	V2	S2	2	9	4

Column(s): 9 Row(s): 64

Strip-Split-Split Plot Design

The steps to generate randomization are listed below:

- Choose **Design | Strip-Split-Split Plot Design** from the main window. The **Randomization and Layout** dialog box will appear.
- Specify the required fields and appropriate options.

Factor Definition

Name

Contain the default name for each factor. User can change the factor name by typing the desired factor name. In specifying the factor name, the following rules apply:

- name should not contain any space;
- must begin with a letter or a period (.)

- succeeding characters can be a combination of letters, numbers, period(.) and underscore(_); and
- name must be different from other factor names already specified.

Factor ID

Contain the default code for each factor. User can change the factor ID by typing the desired factor ID. Factor ID should consist of 1 to 4 letters only and should be unique.

Levels

Contain the level for each factor. User can change the level of the factor by either typing a value inside the spin box or clicking the up-arrow key of the spin box to increase the value, down-arrow key, otherwise. The default and minimum value of the levels is equal to 2 while the maximum value of the levels is 500.

Number of Blocks

Define the number of blocks to be generated. The default value and minimum value is equal to 2.

Number of Subplots along the Length of Each VxH plot

These define the number of subplots along the length of VxH plot. The default and minimum value is equal to one.

Number of Sub-subplots along the Length of Each Subplot

These define the number of sub-subplots along the length of subplot. The default and minimum value is equal to one.

Number of Field Rows

Define the total number of rows in the field layout for each trial. The default and minimum value is equal to 2.

Number of Trial

Define the number of trials to be generated. The default value and minimum value is equal to one.

Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved. The default filename is *fieldbookStripSplit2*.

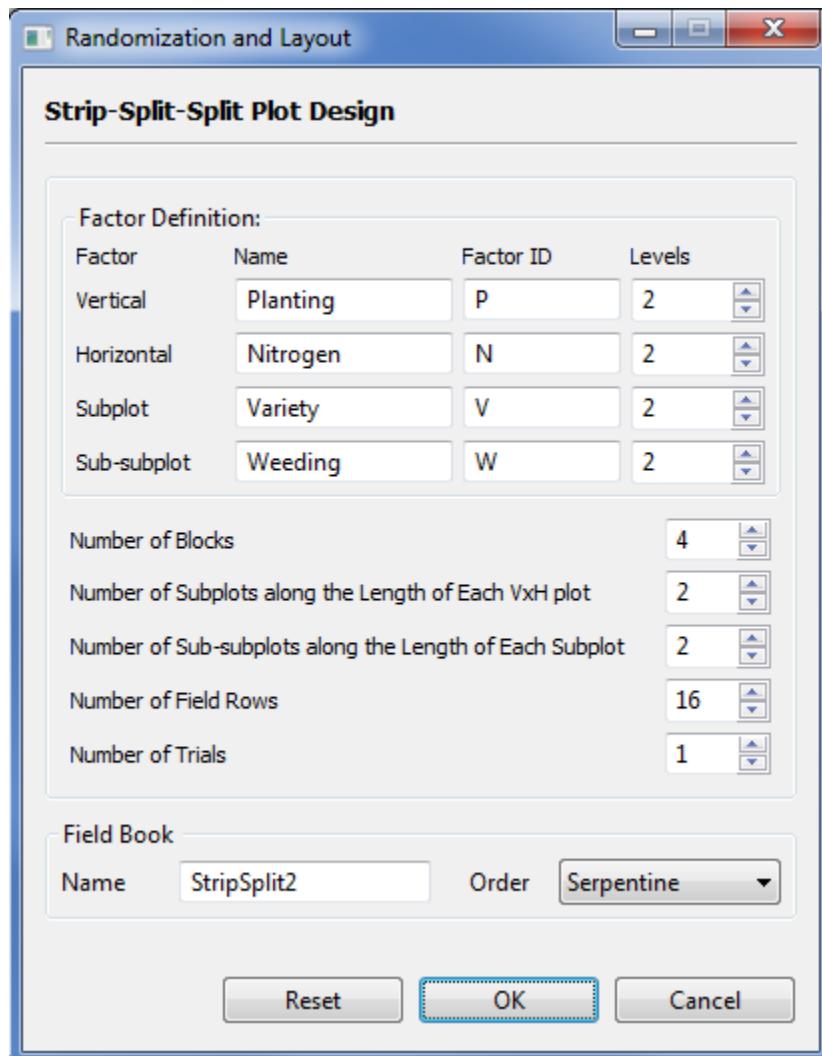
Field Book Order

Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For this example, suppose we want to generate a randomization for an experiment to be conducted using Strip-Split-Split Plot, with planting methods (with 2 levels) as vertical factor, nitrogen rates (with 2 levels) as the horizontal factor, rice variety (with 2 levels) as the subplot factor and weeding method (with and without as levels) as sub-subplot. Four blocks will be used. The planned field layout is shown below:

	Field Column 1	Field Column 2	Field Column 3	Field Column 4	
Field Row 1					Block 1 Block 2 Block 3 Block 4
Field Row 2					
Field Row 3					
Field Row 4					
Field Row 5					
Field Row 6					
Field Row 7					
Field Row 8					

The completed dialog box should appear as shown below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the csv data file and the *txt* file created in the Data Viewer tab.

Sample *txt* file displayed in the Data Viewer is shown below:

```
DESIGN PROPERTIES:  
    Strip-Split-Split Plot Design  
  
DESIGN PARAMETERS:  
    Number of Trials = 1  
    Number of Blocks = 4  
    Vertical Factor = Planting  
    Levels = P1, P2  
    Horizontal Factor = Nitrogen  
    Levels = N1, N2
```

```

Subplot Factor = Variety
Levels = V1, V2
Sub-subplot Factor = Weeding
Levels = W1, W2

```

Layout for Strip-Split-Split Plot:

Trial = 1

	FieldCol1	FieldCol2	FieldCol3	FieldCol4
FieldRow1	101 P2 N2 V2 W2 P1 N2 V1 W2 P2 N1 V2 W1 P1 N1 V2 W1	102 P2 N1 V2 W1 P1 N1 V2 W1 P2 N1 V2 W1 P1 N1 V2 W1	201 P1 N1 V2 W1 P1 N1 V2 W1 P1 N1 V2 W1 P1 N1 V2 W1	202 P1 N1 V2 W1 P1 N1 V2 W1 P1 N1 V2 W1 P1 N1 V2 W1
FieldRow2	104 P2 N2 V2 W1 P1 N2 V1 W1 P2 N1 V2 W2 P1 N1 V2 W2	103 P2 N1 V2 W1 P1 N1 V2 W1 P2 N1 V2 W2 P1 N1 V2 W2	204 P1 N1 V2 W2 P1 N1 V2 W2 P1 N1 V2 W2 P1 N1 V2 W2	203 P1 N1 V2 W2 P1 N1 V2 W2 P1 N1 V2 W2 P1 N1 V2 W2
FieldRow3	105 P2 N2 V1 W1 P1 N2 V2 W1 P2 N1 V1 W1 P1 N1 V1 W2	106 P1 N2 V2 W1 P2 N1 V1 W1 P2 N1 V1 W1 P1 N1 V1 W2	205 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W2	206 P1 N1 V1 W2 P1 N1 V1 W2 P1 N1 V1 W2 P1 N1 V1 W2
FieldRow4	108 P2 N2 V1 W2 P1 N2 V2 W2 P2 N1 V1 W2 P1 N1 V1 W1	107 P1 N2 V2 W2 P2 N1 V1 W2 P2 N1 V1 W2 P1 N1 V1 W1	208 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1	207 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1
FieldRow5	109 P2 N1 V1 W2 P1 N1 V1 W2 P2 N2 V1 W1 P1 N2 V2 W2	110 P1 N1 V1 W2 P2 N2 V1 W1 P1 N2 V2 W2 P1 N2 V2 W2	209 P1 N2 V2 W2 P1 N2 V2 W2 P1 N2 V2 W2 P1 N2 V2 W2	210 P1 N2 V2 W2 P1 N2 V2 W2 P1 N2 V2 W2 P1 N2 V2 W2
FieldRow6	112 P2 N1 V1 W1 P1 N1 V1 W1 P2 N2 V1 W2 P1 N2 V2 W1	111 P1 N1 V1 W1 P2 N2 V1 W2 P1 N2 V2 W1 P1 N2 V2 W1	212 P1 N2 V2 W1 P1 N2 V2 W1 P1 N2 V2 W1 P1 N2 V2 W1	211 P1 N2 V2 W1 P1 N2 V2 W1 P1 N2 V2 W1 P1 N2 V2 W1
FieldRow7	113 P2 N1 V2 W2 P1 N1 V2 W2 P2 N2 V2 W2 P1 N2 V1 W2	114 P1 N1 V2 W2 P2 N2 V2 W2 P1 N2 V1 W2 P1 N2 V1 W2	213 P1 N2 V1 W2 P1 N2 V1 W2 P1 N2 V1 W2 P1 N2 V1 W2	214 P1 N2 V1 W2 P1 N2 V1 W2 P1 N2 V1 W2 P1 N2 V1 W2
FieldRow8	116 P2 N1 V2 W1 P1 N1 V2 W1 P2 N2 V2 W1 P1 N2 V1 W1	115 P1 N1 V2 W1 P2 N2 V2 W1 P1 N2 V1 W1 P1 N2 V1 W1	216 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1	215 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1
FieldRow9	301 P2 N2 V1 W1 P1 N2 V2 W1 P2 N2 V2 W1 P1 N2 V1 W2	302 P1 N2 V2 W1 P2 N2 V2 W1 P1 N2 V1 W1 P1 N2 V1 W2	401 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W2	402 P1 N2 V1 W2 P1 N2 V1 W2 P1 N2 V1 W2 P1 N2 V1 W2
FieldRow10	304 P2 N2 V1 W2 P1 N2 V2 W2 P2 N2 V2 W2 P1 N2 V1 W1	303 P1 N2 V2 W2 P2 N2 V2 W2 P1 N2 V1 W1 P1 N2 V1 W1	404 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1	403 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1 P1 N2 V1 W1
FieldRow11	305 P2 N2 V2 W2 P1 N2 V1 W2 P2 N2 V1 W1 P1 N2 V2 W2	306 P1 N2 V1 W2 P2 N2 V1 W1 P1 N2 V2 W2 P1 N2 V2 W2	405 P1 N2 V2 W2 P1 N2 V2 W2 P1 N2 V2 W2 P1 N2 V2 W2	406 P1 N2 V2 W2 P1 N2 V2 W2 P1 N2 V2 W2 P1 N2 V2 W2
FieldRow12	308 P2 N2 V2 W1 P1 N2 V1 W1 P2 N2 V1 W2 P1 N2 V2 W1	307 P1 N2 V1 W1 P2 N2 V1 W2 P1 N2 V2 W1 P1 N2 V2 W1	408 P1 N2 V2 W1 P1 N2 V2 W1 P1 N2 V2 W1 P1 N2 V2 W1	407 P1 N2 V2 W1 P1 N2 V2 W1 P1 N2 V2 W1 P1 N2 V2 W1
FieldRow13	309 P2 N1 V1 W2 P1 N1 V2 W1 P2 N1 V1 W2 P1 N1 V2 W1	310 P1 N1 V2 W1 P2 N1 V1 W2 P1 N1 V2 W2 P1 N1 V2 W1	409 P1 N1 V2 W2 P1 N1 V2 W2 P1 N1 V2 W2 P1 N1 V2 W1	410 P1 N1 V2 W1 P1 N1 V2 W1 P1 N1 V2 W1 P1 N1 V2 W1
FieldRow14	312 P2 N1 V1 W1 P1 N1 V2 W2 P2 N1 V1 W1 P1 N1 V2 W2	311 P1 N1 V2 W2 P2 N1 V1 W1 P1 N1 V2 W1 P1 N1 V2 W2	412 P1 N1 V2 W1 P1 N1 V2 W1 P1 N1 V2 W1 P1 N1 V2 W2	411 P1 N1 V2 W2 P1 N1 V2 W2 P1 N1 V2 W2 P1 N1 V2 W2
FieldRow15	313 P2 N1 V2 W1 P1 N1 V1 W1 P2 N1 V2 W2 P1 N1 V1 W2	314 P1 N1 V1 W1 P2 N1 V2 W2 P1 N1 V1 W2 P1 N1 V1 W2	413 P1 N1 V1 W2 P1 N1 V1 W2 P1 N1 V1 W2 P1 N1 V1 W2	414 P1 N1 V1 W2 P1 N1 V1 W2 P1 N1 V1 W2 P1 N1 V1 W2
FieldRow16	316 P2 N1 V2 W2 P1 N1 V1 W2 P2 N1 V2 W1 P1 N1 V1 W1	315 P1 N1 V1 W2 P2 N1 V2 W1 P1 N1 V1 W1 P1 N1 V1 W1	416 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1	415 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1 P1 N1 V1 W1

**Note: Cells contain plot numbers on top, treatments/entries below

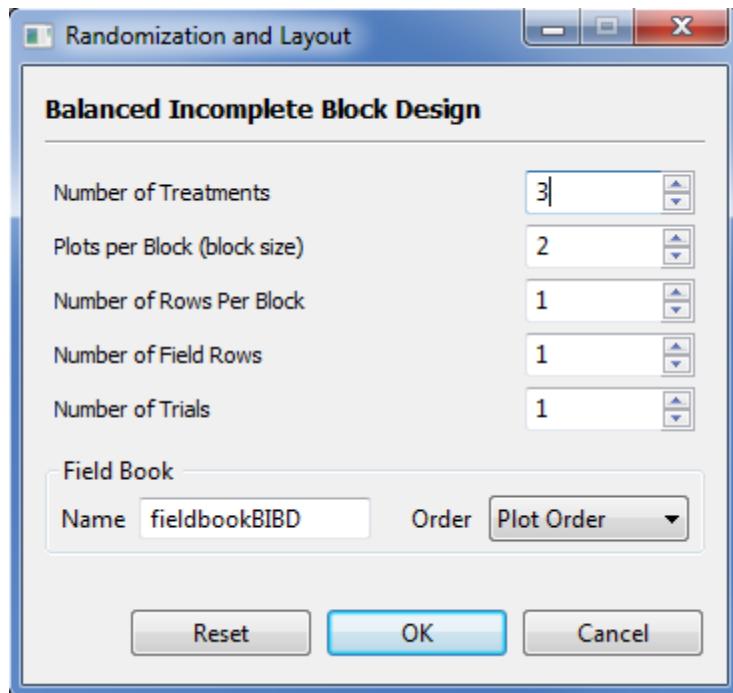
Sample csv data file displayed in the Data Viewer is shown below:

	Trial	Block	PlantingMethod	Nitrogen	Variety	WeedingMethod	PlotNum	FieldRow	FieldCol
1	1	1	P2	N2	V1	W1	1	1	1
2	1	1	P2	N2	V2	W1	2	1	2
3	1	1	P1	N2	V1	W1	3	1	3
4	1	1	P1	N2	V5	W1	4	1	4
5	1	1	P2	N2	V1	W2	5	2	1
6	1	1	P2	N2	V2	W2	6	2	2
7	1	1	P1	N2	V1	W2	7	2	3
8	1	1	P1	N2	V5	W2	8	2	4

Balanced Incomplete Block Design

The steps to generate randomization for Balanced Incomplete Block Design are listed below:

- Click **Design | Balanced Incomplete Block...** from the main window. The **Randomization and Layout** dialog box will appear.



- Specify the required field and appropriate options.

Number of Treatments

Define the number of treatments (trmt) to be included in the randomization. The default and minimum value is 3. The entry should be a greater than the entry for the plots per block.

Plots per Block (block size)

Define the number of replicates to be generated. The default value and minimum value is equal to 2. The entry should be less than the number of treatments.

Number of Rows Per Block

Define the number of rows per block. The default and minimum value is equal to 1.

Number of Field Rows

Define the number of rows in the field layout for each trial. The default and minimum value is equal to one.

Number of Trial

Define the number of trials to be generated. The default value and minimum value is equal to one.

Field Book Filename

This field is required and it specifies the file name where the result of the randomization will be saved. The default filename is *fieldbookBIBD*.

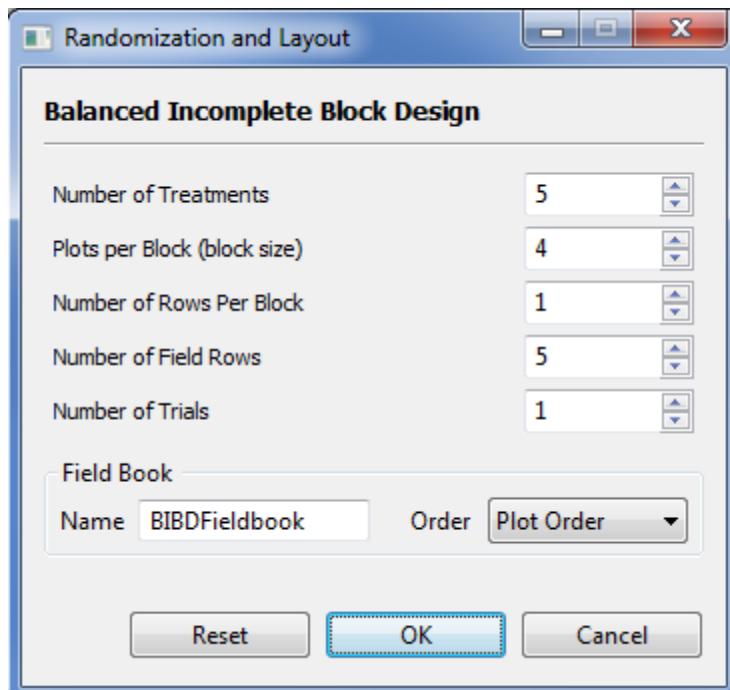
Field Book Order

Define how the plot number will be arranged. There are two options available: *Plot Order* (default value) and *Serpentine*.

For this example, suppose we want to generate a randomization for an experiment with 5 treatment levels and 5 blocks. Each block can accommodate only 4 treatment levels. The planned field layout is shown below:

	Field Column 1	Field Column 2	Field Column 3	Field Column 4	
Field Row 1					Block 1
Field Row 2					Block 2
Field Row 3					Block 3
Field Row 4					Block 4
Field Row 5					Block 5

For the example, the completed dialog box should appear as below:



- Click the **OK** button to generate the randomization. The **Randomization and Layout** dialog box will be minimized and STAR will automatically display the csv data file created in the Data Viewer tab.

Sample *txt* file displayed in the Data Viewer is shown below:

DESIGN PROPERTIES:
Balanced Incomplete Block Design

DESIGN PARAMETERS:
Number of Trials = 1
Number of Treatments = 5
Plots per Block (Block Size) = 4
Number of Blocks = 5
Number of Replicates = 4
Lambda = 3

Number of Field Row = 5
Number of Field Column = 4

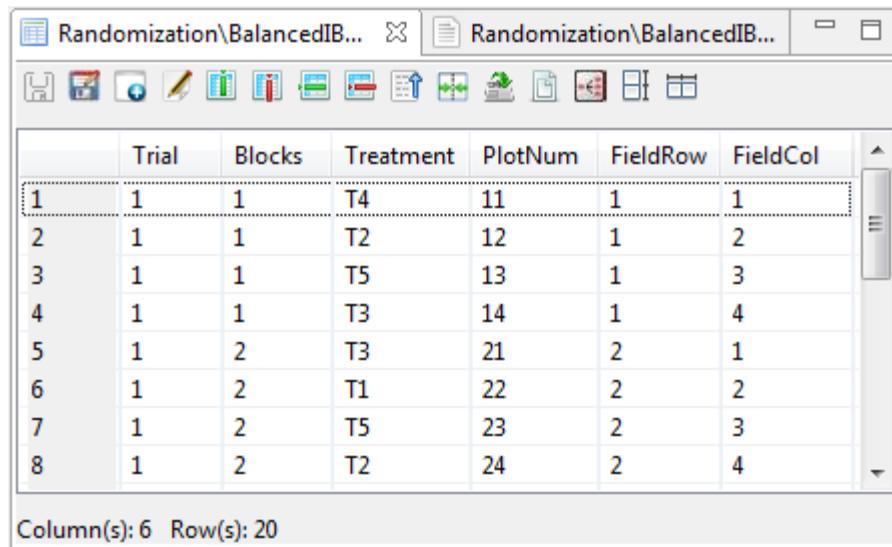
Layout for Balanced Incomplete Block Design:

Trial = 1

	FieldCol1	FieldCol2	FieldCol3	FieldCol4
FieldRow1	11	12	13	14
	T4	T2	T5	T3
FieldRow2	21	22	23	24
	T3	T1	T5	T2
FieldRow3	31	32	33	34
	T4	T2	T3	T1
FieldRow4	41	42	43	44
	T1	T2	T4	T5
FieldRow5	51	52	53	54
	T5	T3	T1	T4

**Note: Cells contain plot numbers on top, treatments/entries below

Sample csv data file displayed in the Data Viewer:



The screenshot shows a software window titled "Randomization\BalancedIB...". The window contains a data grid with 8 rows and 7 columns. The columns are labeled: Trial, Blocks, Treatment, PlotNum, FieldRow, and FieldCol. The data is as follows:

Trial	Blocks	Treatment	PlotNum	FieldRow	FieldCol
1	1	T4	11	1	1
2	1	T2	12	1	2
3	1	T5	13	1	3
4	1	T3	14	1	4
5	1	T3	21	2	1
6	1	T1	22	2	2
7	1	T5	23	2	3
8	1	T2	24	2	4

Column(s): 6 Row(s): 20

Graphs

Bar Graph

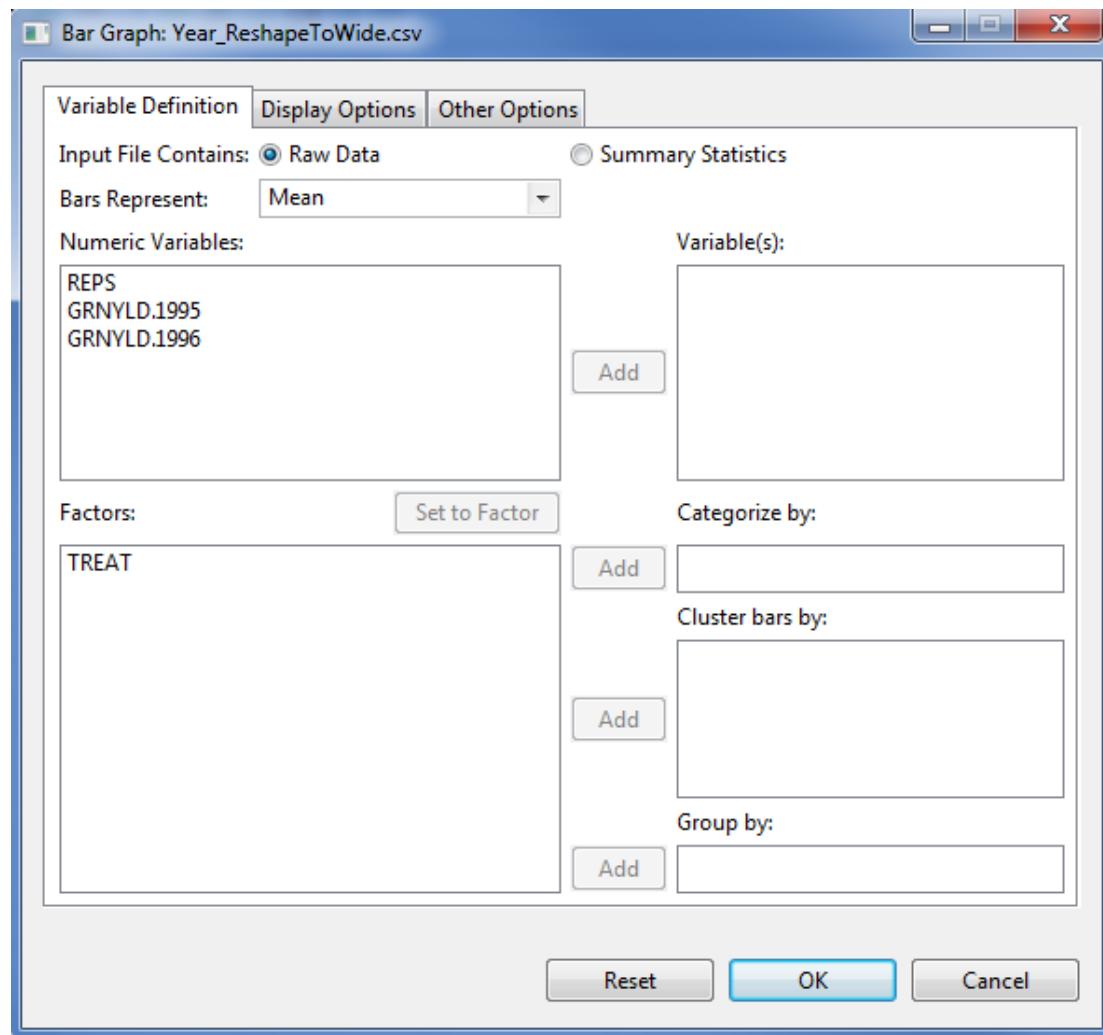
Bar graphs are useful for visualizing frequencies, sums, and statistics such as means and medians. The bar graphs that are created here can come from raw data or from pre-computed summary information.

The steps to create a bar graph are listed below:

- On the **Project Explorer** panel, double-click the file to be used to open it and view it in the Data Viewer tab.

For the example, using the project named *TrialProject*, select the data file *Year_ReshapeToWide.csv*. The file was created in the **Reshaping data from long (serial) to wide (parallel)** section of this user's manual.

- Click **Graphs | Bar Graph....** The **Bar Graph** dialog box will appear.



- Specify the inputs for the required fields and desired options.

Variable Definition Tab

Input File Contains

There are two options available, **Raw Data** (default option) and **Summary Statistics**.

Bars Represent

This item is enabled if the **Input File Contains Raw Data** option is selected. Four options are available for creating the bars: **Frequency**, **Mean**, **Median**, and **Sum**.

Variable(s)

At least one entry should be specified for the graph to be created. Entries should come from the **Numeric Variables** dialog box.

Categorize by

This field is required. The factor specified will define the labels for the bars and, if input file contains raw data, the categories for which values are computed.

Cluster bars by

This field is optional. For this, up to three factors can be specified. A group of bars is going to be generated for each level of the factors combined. Entries should come from the **Factors** list box.

Group by

This field is optional. If a factor is specified, a graph is going to be generated per category of the factor. Entry should come from the **Factors** list box.

Display Options Tab

Main Title

This field will contain the title for the graph.

Category Axis Label

This field will contain the label for the factor. If a factor is specified in the **Cluster bars by** field, the button **Display levels** will be enabled and checking it will add the levels of the combined factors onto the category axis.

Value Axis

The label and limits for the numeric variable(s) can be edited here. Default axis label is the name of the numeric variable(s). Lower and upper limits, if not specified, are determined from the data.

Axis Label Orientation

This option will determine how axis labels are displayed: *parallel to axis*, *horizontal*, *vertical*, and *perpendicular to axis*.

Draw box around plot

If checked, the graph will be enclosed in a box.

Show legend

If checked, a legend will be added to the graph. It may or may not have a title. The legend may be placed in any of nine positions within the graph: *bottom*, *bottom-left*, *bottom-right*, *center*, *left*, *right*, *top*, *top-left*, *top-right*. The items on the legend can be

listed in one or more columns, the maximum of which is the number of levels of the category variable.

Display multiple graphs in a page

If checked, several graphs will be displayed on a single page, based on the specified number of rows and columns and orientation. At most 5 rows and 5 columns of graphs can be fitted into a single page, arranged either from *Top-to-bottom* or *Left-to-right*.

Other Options Tab

Orientation

Bars can either be drawn vertically or horizontally.

Style

This option is useful if several numeric variables are selected. Bars can be clustered (side-by-side) or stacked.

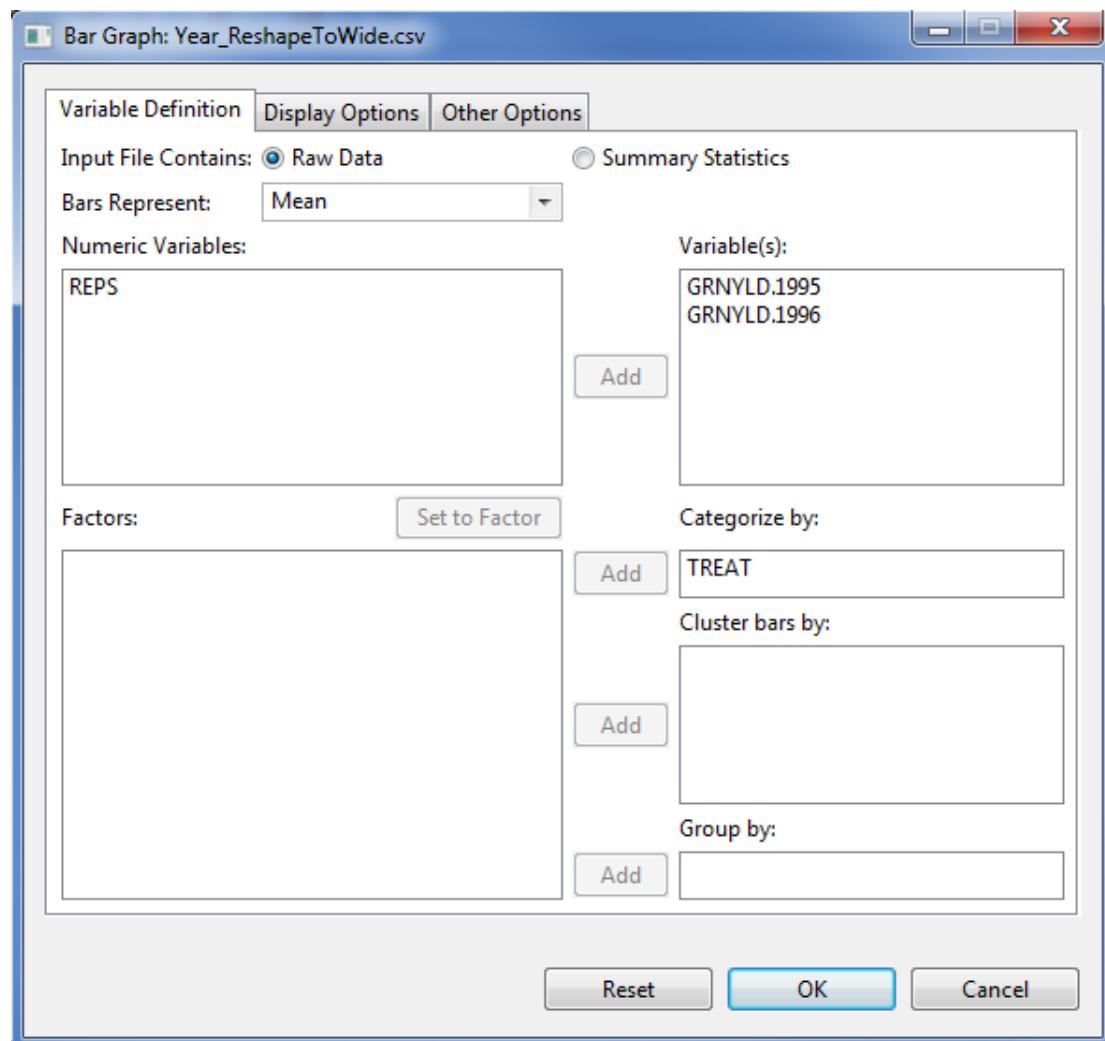
Color and Fill

Colors and fill patterns for the bars may be changed in this section.

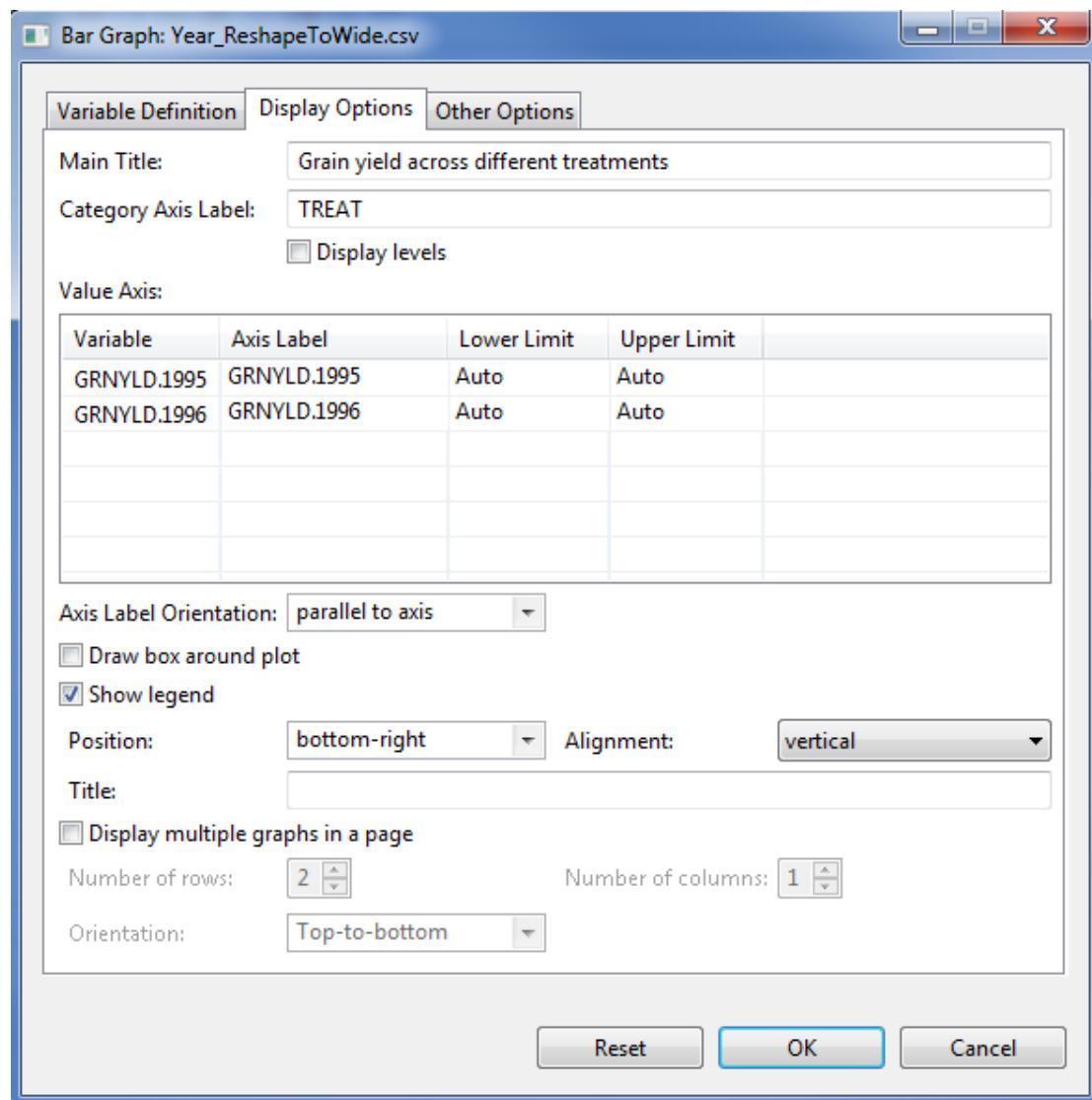
Display Error Bars

If the bars represent means to be computed from raw data, error bars may be added based on one of three options: **confidence level** (default: 95%), **standard error** (with multiplier 1, 2, or 3), and **standard deviation** (with multiplier 1, 2, or 3).

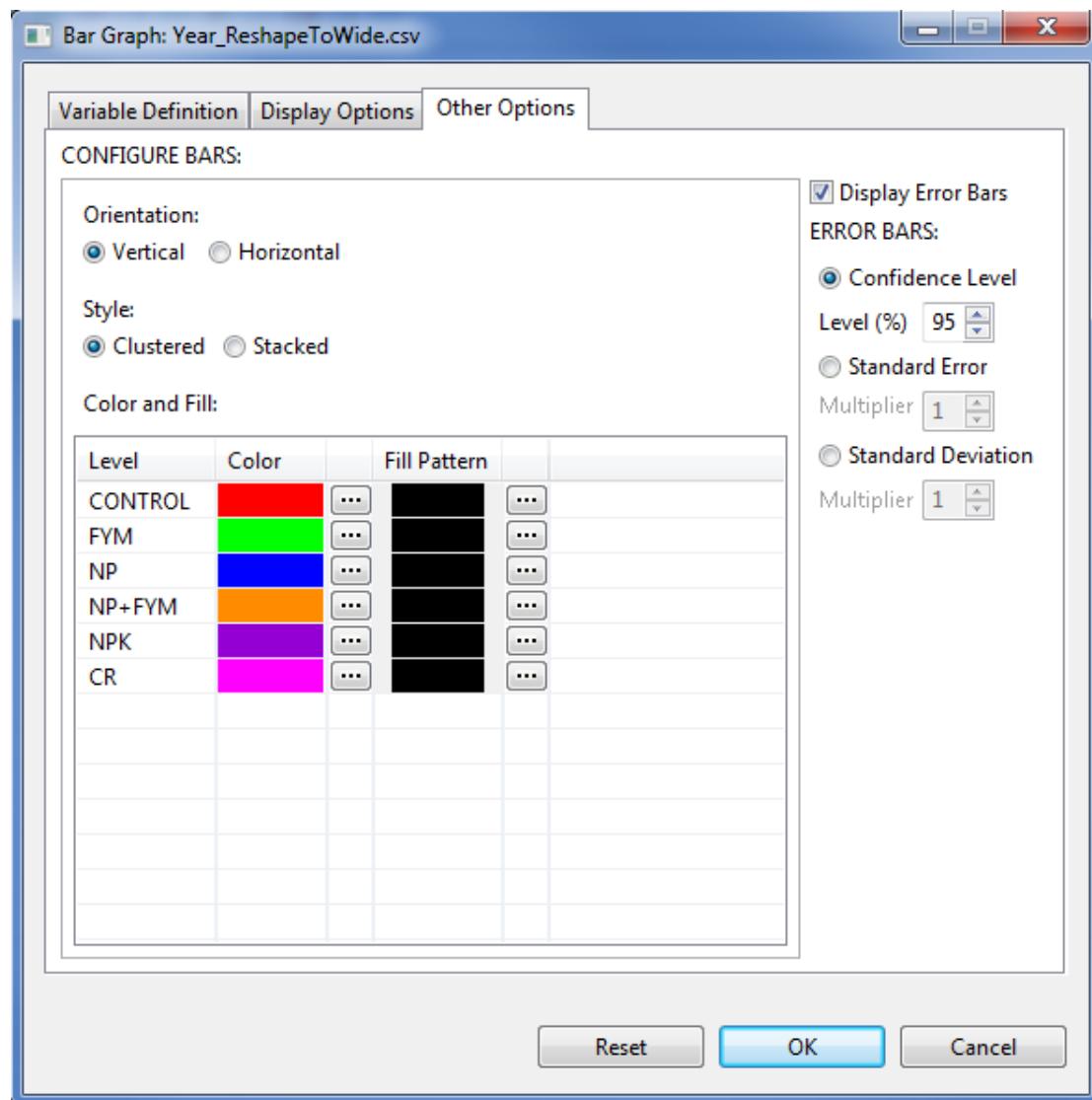
For the example, suppose we want to create a bar graph for the means of *GRNYLD.1995* and *GRNYLD.Y1996* with error bars, the completed **Variable Definition** tab should appear as shown below:



Items in the Display Options tab may be supplied as illustrated:

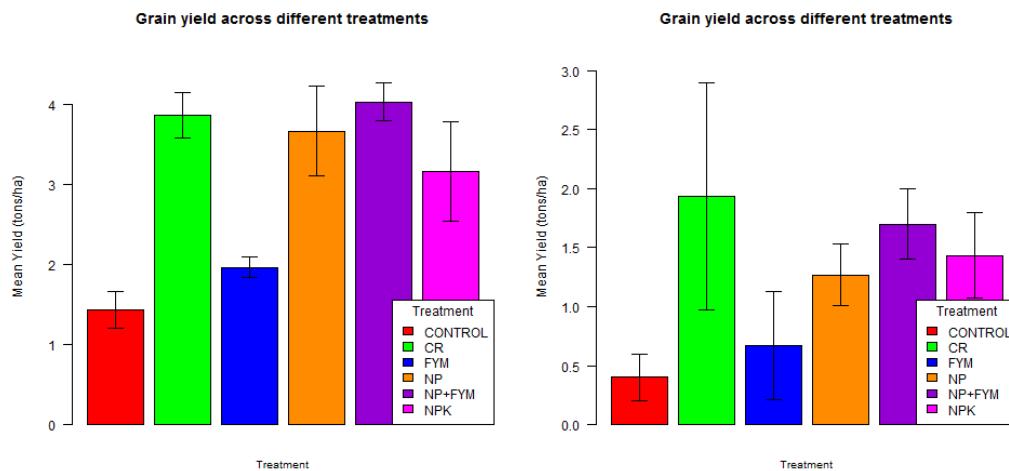


Items in the **Other Options** tab may be supplied as illustrated:



- Click the **OK** button to create the graph. The **Bar Graph** dialog box will be minimized and the **Result Viewer** tab will be activated.

Sample graphs as displayed in two pages in the *Graph* page of the Result Viewer:



Box Plot

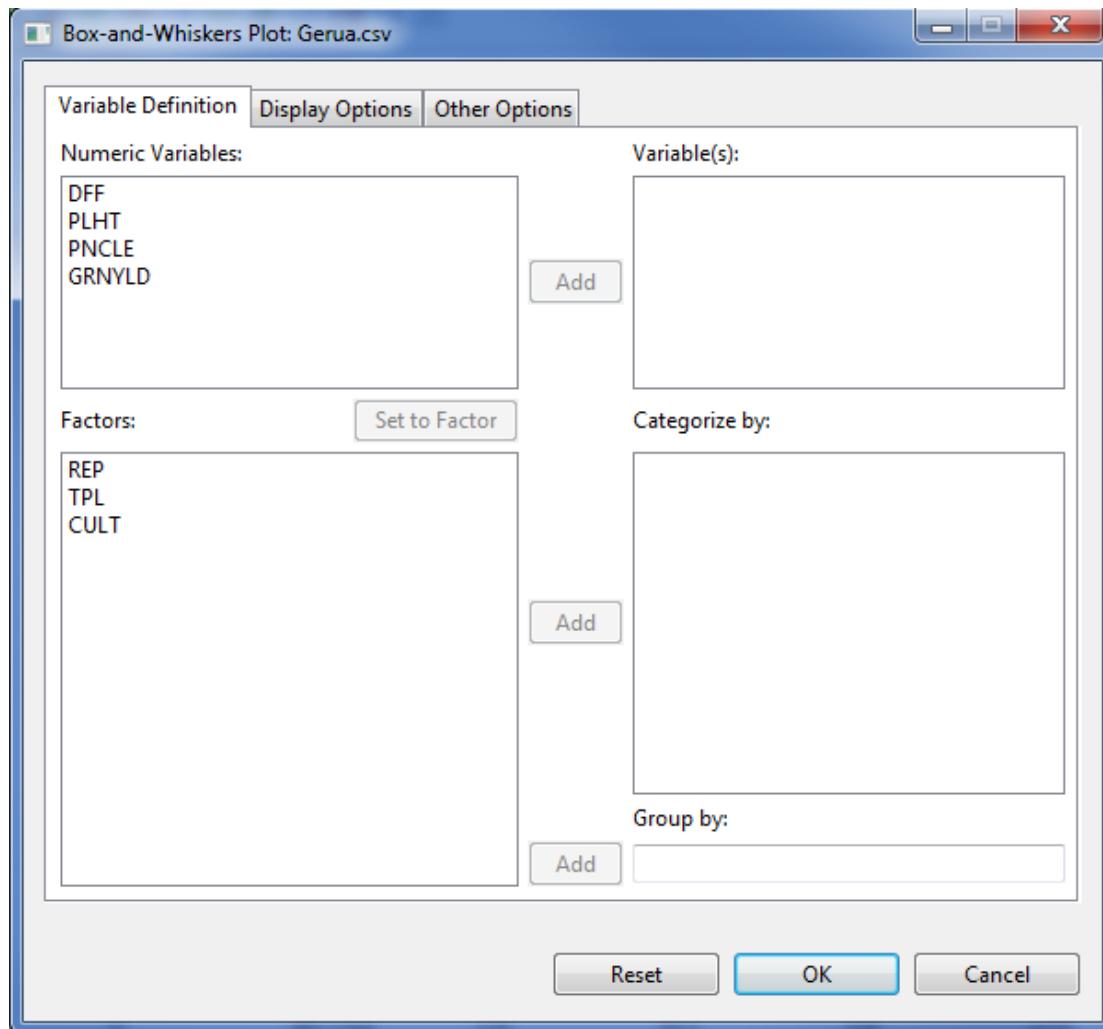
Box plots are useful for displaying distributional properties of data.

The steps to create a box plot are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *Gerua* from STAR Package. See the *Importing Data from Package* section of this user's manual for instructions on how to do this.

- Click **Graphs | Box Plot....** The **Box Plot** (or Box-and-Whiskers Plot) dialog box will appear.



- Specify the inputs for the required fields and desired options.

Variable Definition Tab

Variable(s)

At least one entry should be specified for the graph to be created. Entries should come from the **Numeric Variables** dialog box.

Categorize by

This field, if specified, will define the labels and sub-grouping for the box plots.

Group by

This field is optional. If a factor is specified, a graph is going to be generated per category of the factor. Entry should come from the **Factors** list box.

Display Options Tab

Main Title

This field will contain the title for the graph.

Category Axis Label

This field will contain the label for the factor.

Value Axis

The label and limits for the numeric variable(s) can be edited here. Default axis label is the name of the numeric variable(s). Lower and upper limits, if not specified, are determined from the data.

Axis Label Orientation

This option will determine how axis labels are displayed: *parallel to axis*, *horizontal*, *vertical*, and *perpendicular to axis*.

Orientation

Plots can be displayed vertically (default) or horizontally.

Draw box around plot

If checked, the graph will be enclosed in a box.

Display multiple graphs in a page

If checked, several graphs will be displayed on a single page, based on the specified number of rows and columns and orientation. At most 5 rows and 5 columns of graphs can be fitted into a single page, arranged either from *Top-to-bottom* or *Left-to-right*.

Other Options Tab

BOX

Color for the box outline and fill, as well as the size of the line used for the box may be edited here.

WHISKERS

Color, type and width for the line representing the whiskers may be edited here.

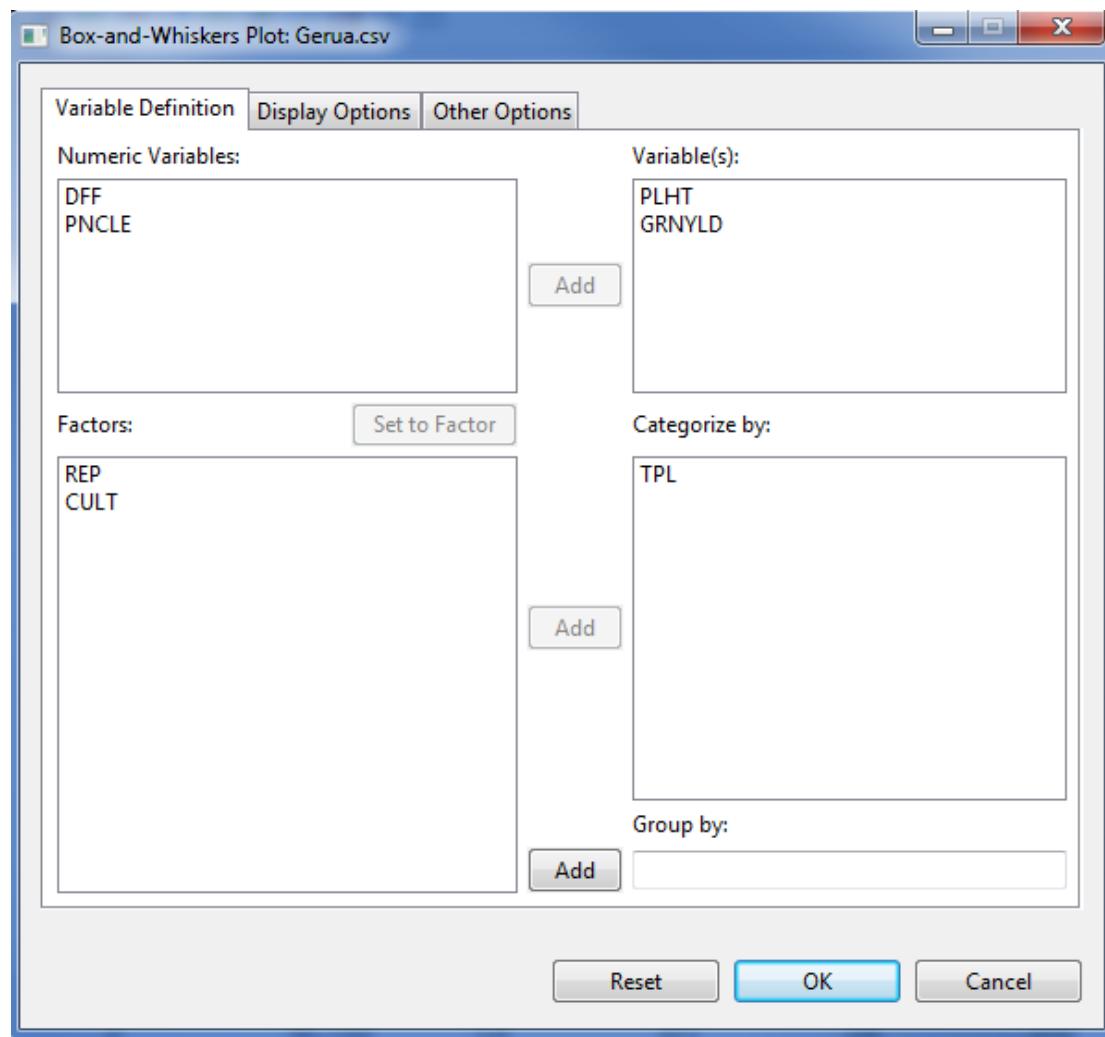
MEDIAN

Color, type and width for the line representing the median may be edited here.

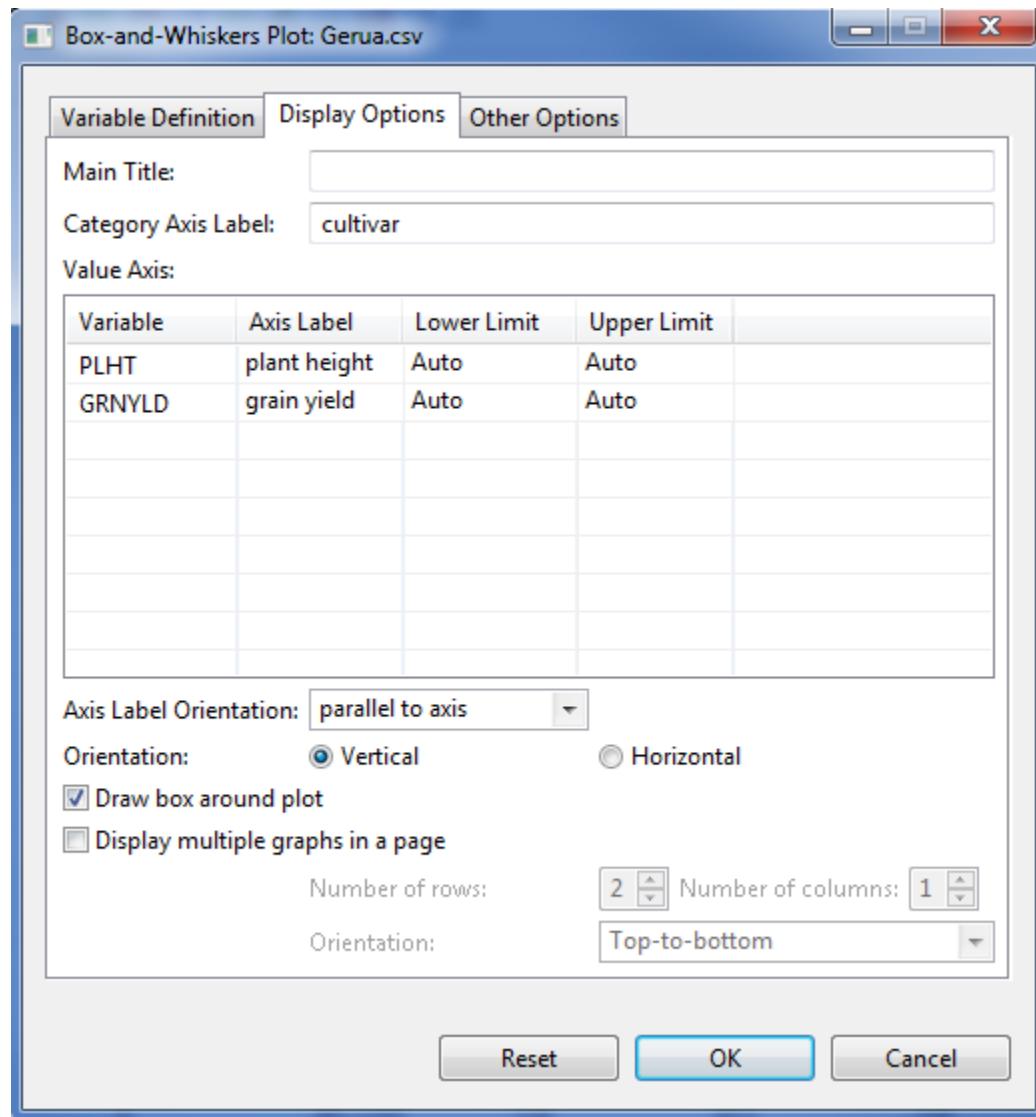
OUTLIERS

Color, symbol and size for the representing the outliers may be edited here.

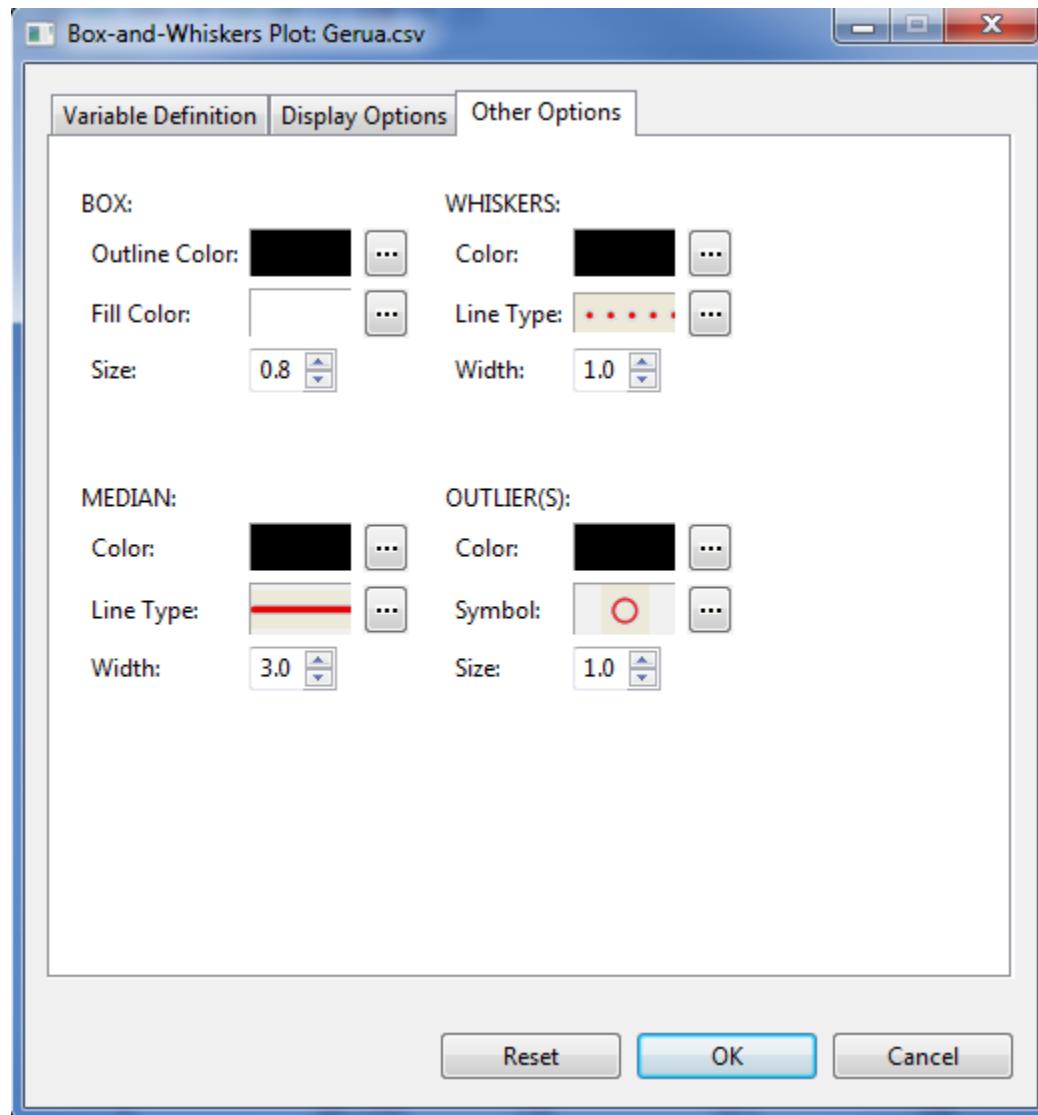
For the example, suppose we want to create box plots for the plant height (*PLHT*) and grain yield (*GRNYLD*) variables, categorized based on transplanting method (*TPL*). The completed **Variable Definition** tab should appear as shown below:



Items in the **Display Options** tab may be supplied as illustrated:

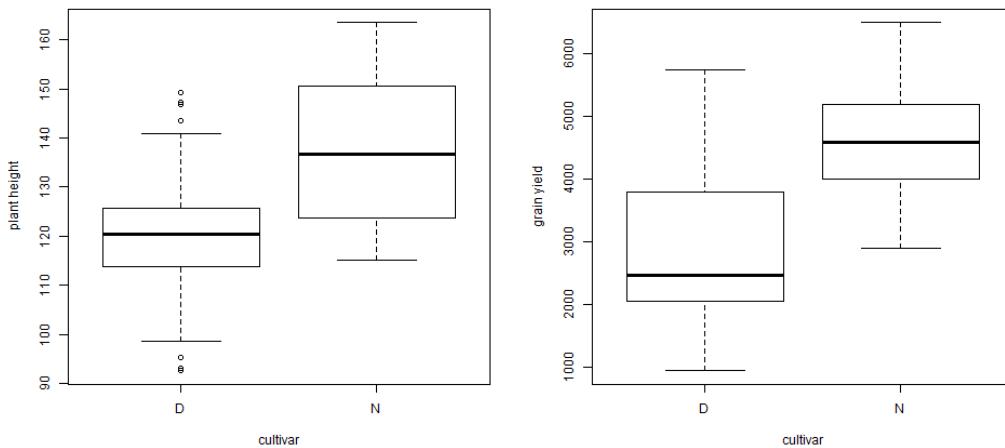


Items in the **Other Options** tab may be supplied as illustrated:



- Click the **OK** button to create the graph. The **Box Plot** dialog box will be minimized and the Result Viewer will be activated.

Sample graph as displayed in the *Graph* page of the Result Viewer:



Histogram

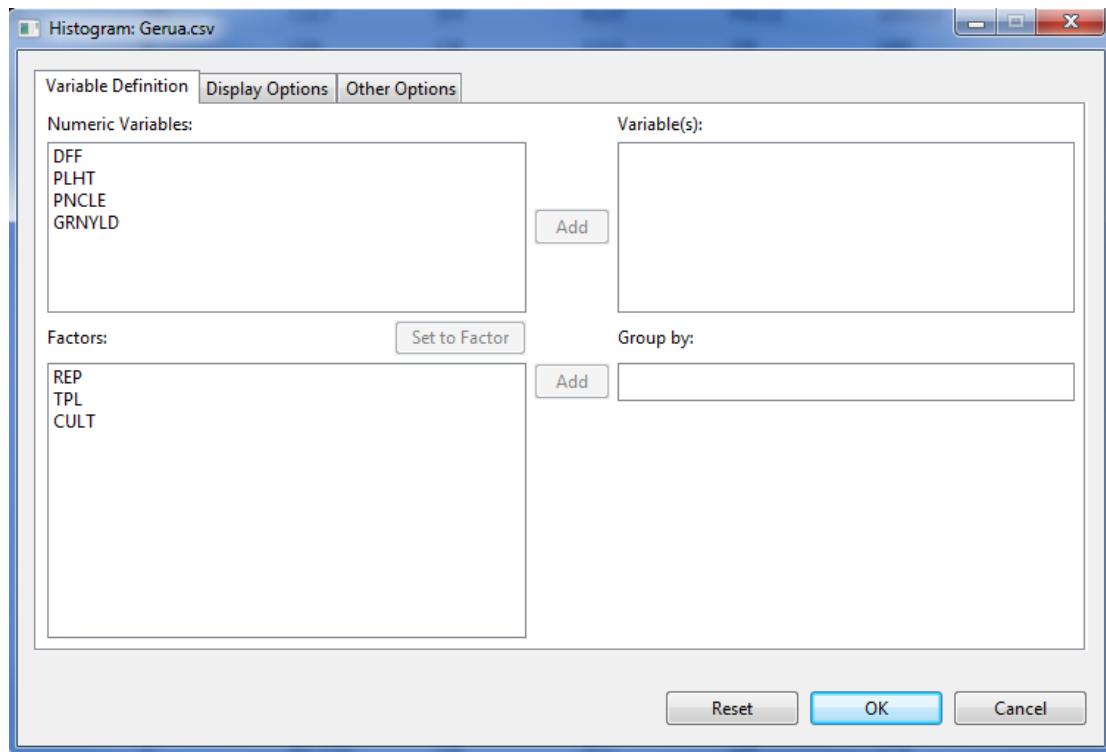
Histograms are useful for displaying distributional properties of data.

The steps to create a histogram are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *Gerua* from STAR Package. See the *Importing Data from Package* section of this user's manual for instructions on how to do this.

- Click **Graphs | Histogram....** The **Histogram** dialog box will appear.



- Specify the inputs for the required fields and desired options.

Variable Definition Tab

Variable(s)

At least one entry should be specified for the graph to be created. Entries should come from the **Numeric Variables** dialog box.

Group by

This field is optional. If a factor is specified, a graph is going to be generated per category of the factor. Entry should come from the **Factors** list box.

Display Options Tab

Main Title

This field will contain the title for the graph.

Use

This is used for determining whether frequencies or probability densities will be used in the y-axis.

X-Axis

The label and limits for the numeric variable(s) can be edited here. Default axis label is the name of the numeric variable(s). Lower and upper limits, if not specified, are determined from the data.

Y-Axis

The label and limits for the y-axis can be edited here. Default axis label is *Frequency* or *Probability*, depending the option chosen in the **Use** subsection above. Lower and upper limits, if not specified, are determined from the data.

Axis Label Orientation

This option will determine how axis labels are displayed: *parallel to axis*, *horizontal*, *vertical*, and *perpendicular to axis*.

Draw box around plot

If checked, the graph will be enclosed in a box.

Display multiple graphs in a page

If checked, several graphs will be displayed on a single page, based on the specified number of rows and columns and orientation. At most 5 rows and 5 columns of graphs can be fitted into a single page, arranged either from *Top-to-bottom* or *Left-to-right*.

Other Options Tab

Number of bars

A suggested number of bars to be used for distributing the data can be specified here. The actual number of bars used by STAR may vary. If not specified, STAR determines this automatically.

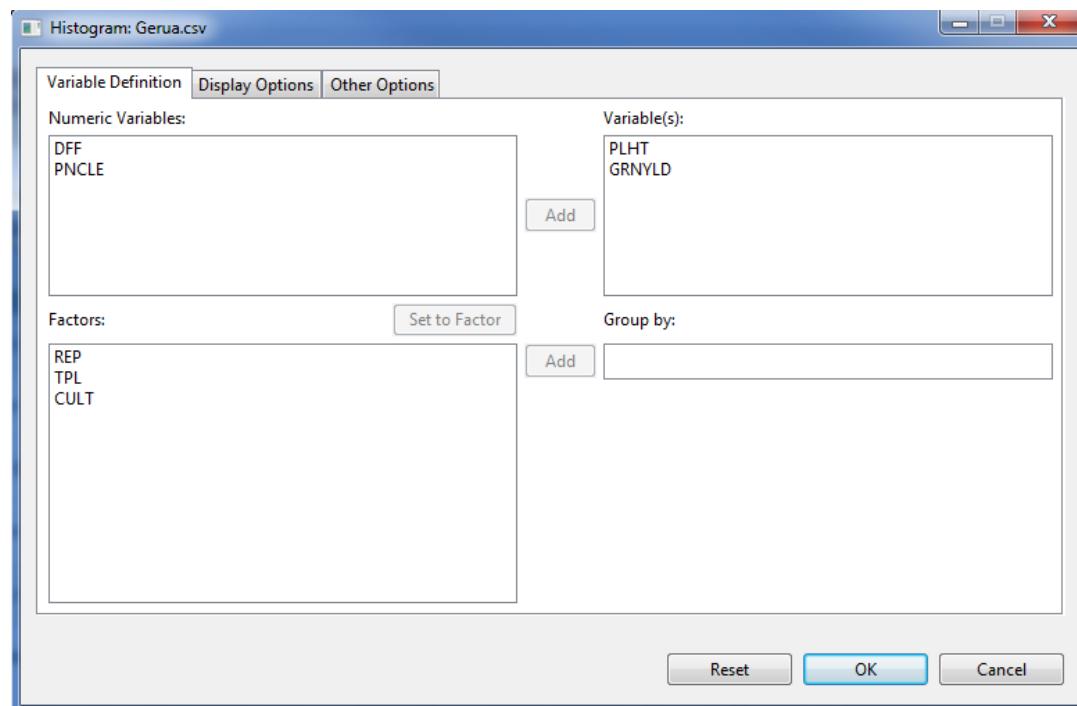
Overlay kernel density estimates

If this option is checked, kernel density estimates are plotted over the bars. Option used in the **Use** subsection above will be automatically set to *probabilities*.

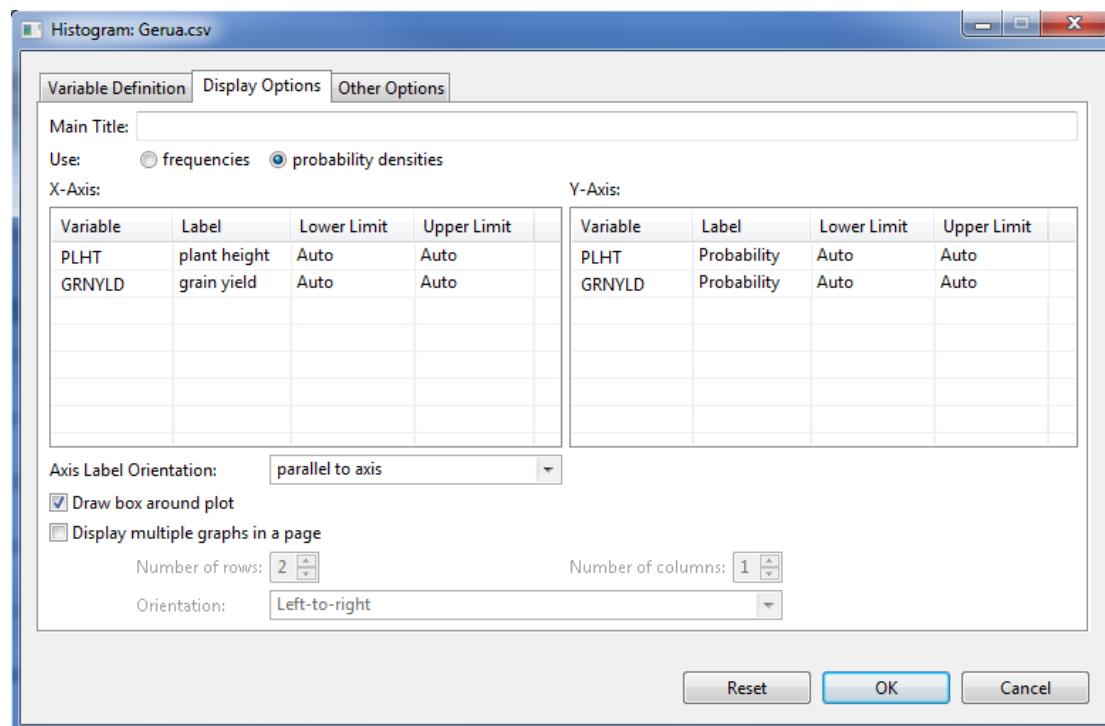
Format Components

Color for the bars can be edited here. If **Overlay kernel density estimates** option is checked, type of line, line color, and width for the curve can also be edited.

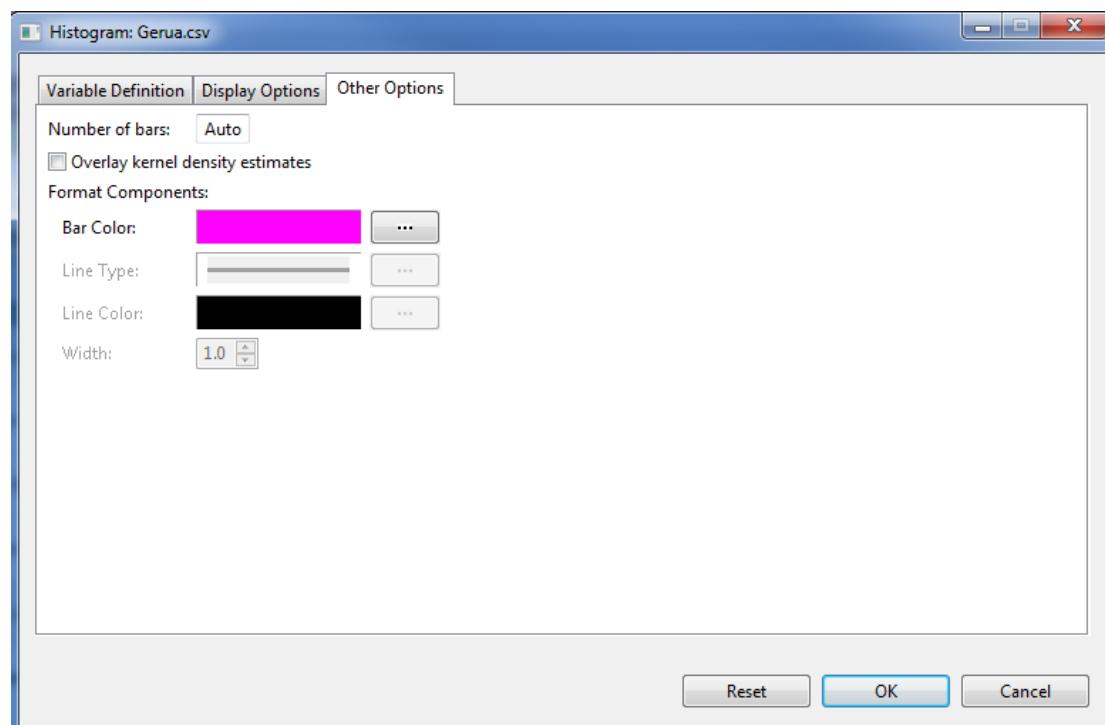
For the example, suppose we want to create histograms for plant height (*PLHT*) and grain yield (*GRNYLD*) variables. The completed **Variable Definition** tab should appear as shown below:



Items in the **Display Options** tab may be supplied as illustrated:

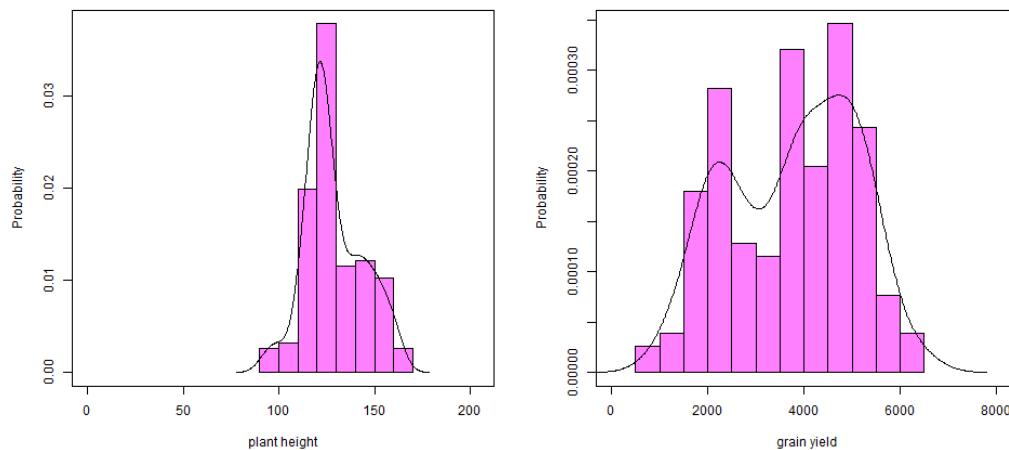


Items in the **Other Options** tab may be supplied as illustrated:



- Click the **OK** button to create the graph. The **Histogram** dialog box will be minimized and the Result Viewer will be activated.

Sample graph as displayed in the *Graph* page of the Result Viewer:



Line Graph

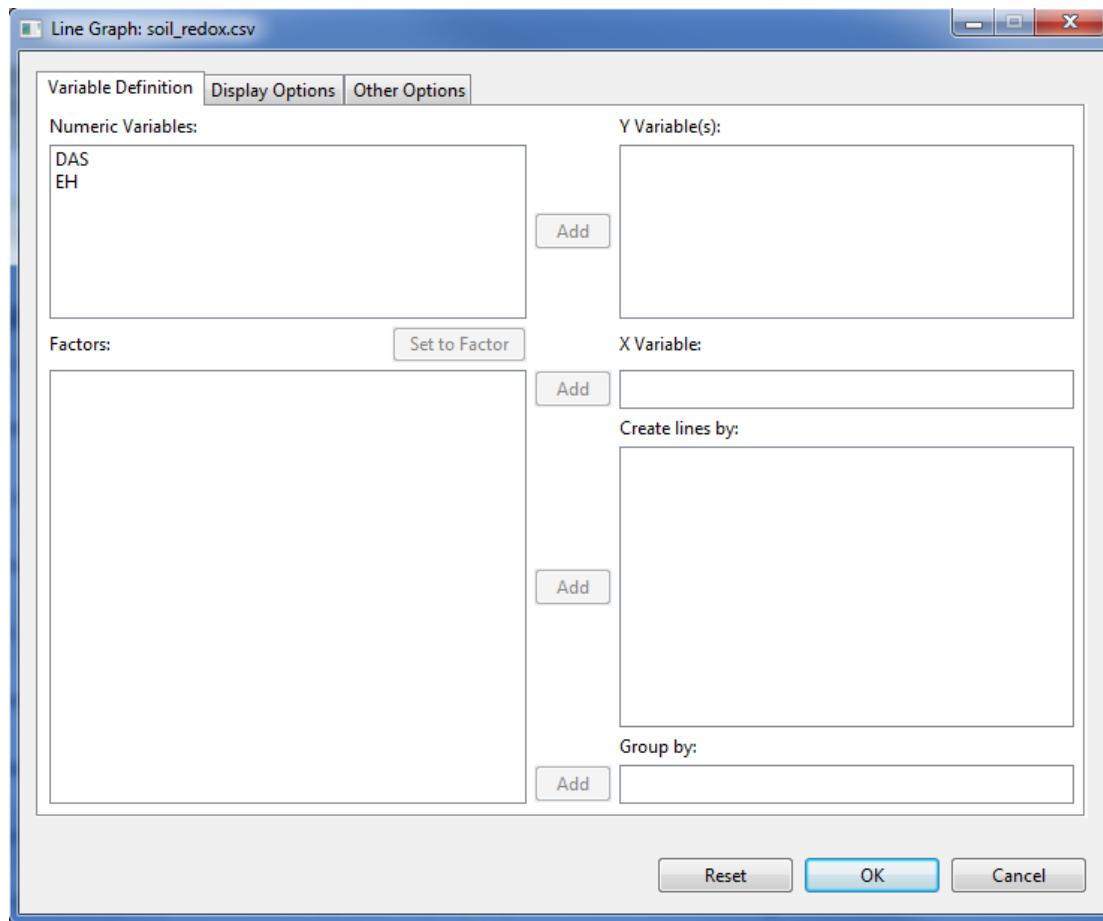
Line graphs are useful for displaying trends using either raw data or statistics.

The steps to create a line graph are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *soil_redox* from STAR Package. See the *Importing Data from Package* section of this user's manual for instructions on how to do this.

- Click **Graphs | Line Graph....** The **Line Graph** dialog box will appear.



- Specify the inputs for the required fields and desired options.

Variable Definition Tab

Y Variable(s)

At least one entry should be specified for the graph to be created. Entries should come from the **Numeric Variables** dialog box.

X Variable(s)

This determines the levels for which means are computed. Entry should come from the **Factors** dialog box.

Create lines by

This field is optional. For this, up to three factors can be specified. Lines will be created for each of the levels of the factors combined. Entries should come from the **Factors** list box.

Group by

This field is optional. If a factor is specified, a line graph is going to be generated per category of the factor. Entry should come from the **Factors** list box.

Display Options Tab

Main Title

This field will contain the title for the graph.

Horizontal Axis Label

This field will contain the label for the factor.

Vertical Axis

The label and limits for the vertical axis can be edited here. Default axis label is the name of the variable. Lower and upper limits, if not specified, are determined from the data.

Axis Label Orientation

This option will determine how axis labels are displayed: *parallel to axis*, *horizontal*, *vertical*, and *perpendicular to axis*.

Draw box around plot

If checked, the graph will be enclosed in a box.

Display line labels

If checked, the name(s) of the numeric variable(s) are displayed on the right side of the corresponding line(s).

This option is useful if there are two or more lines displayed in the graph. If this option is used, **Show legend** is disabled.

Show legend

If checked, a legend will be added to the graph. It may or may not have a title. The legend may be placed in any of nine positions within the graph: bottom, bottom-left, bottom-right, center, left, right, top, top-left, top-right. The items on the legend can be listed in one or more columns, the maximum of which is the number lines displayed in the graph.

Display multiple graphs in a page

If checked, several graphs will be displayed on a single page, based on the specified number of rows and columns and orientation. At most 5 rows and 5 columns of

graphs can be fitted into a single page, arranged either from *Top-to-bottom* or *Left-to-right*.

Other Options Tab

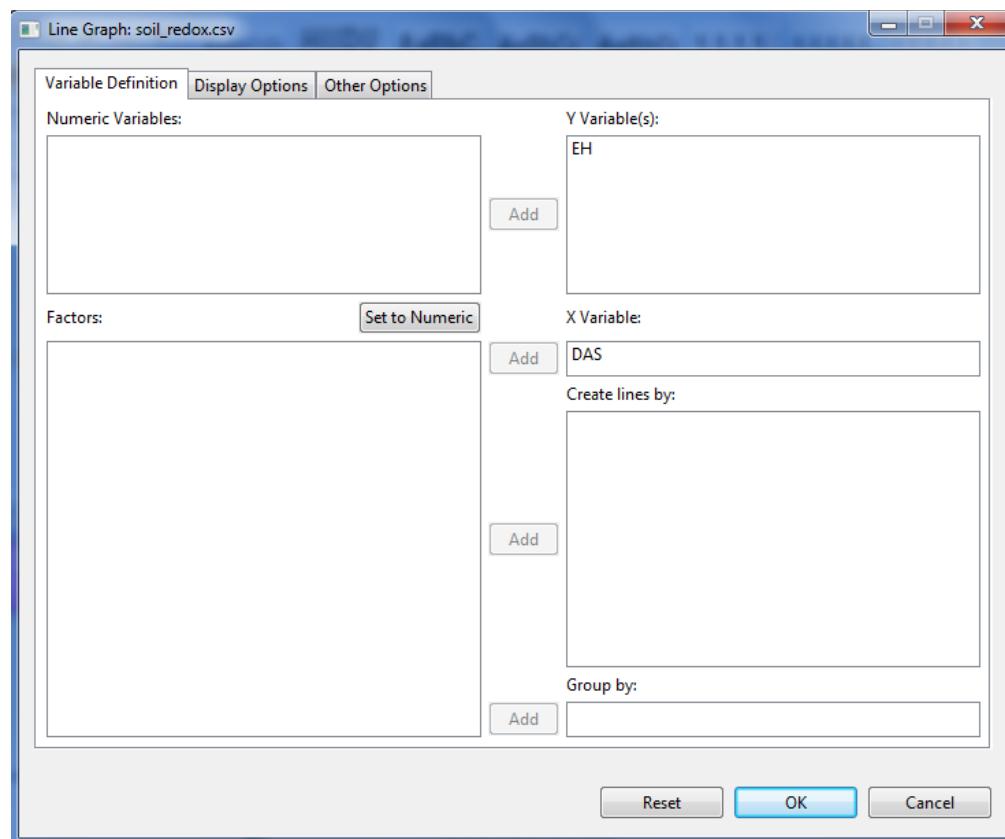
Format Components (lines and points)

Formats of lines and points (if any) can be edited here. Under *Type*, available options are *lines only*, *overplot* (overlapping lines and points), or *both* (non-overlapping lines and points). Default color, width and type for lines, as well as symbol and size for points can be changed.

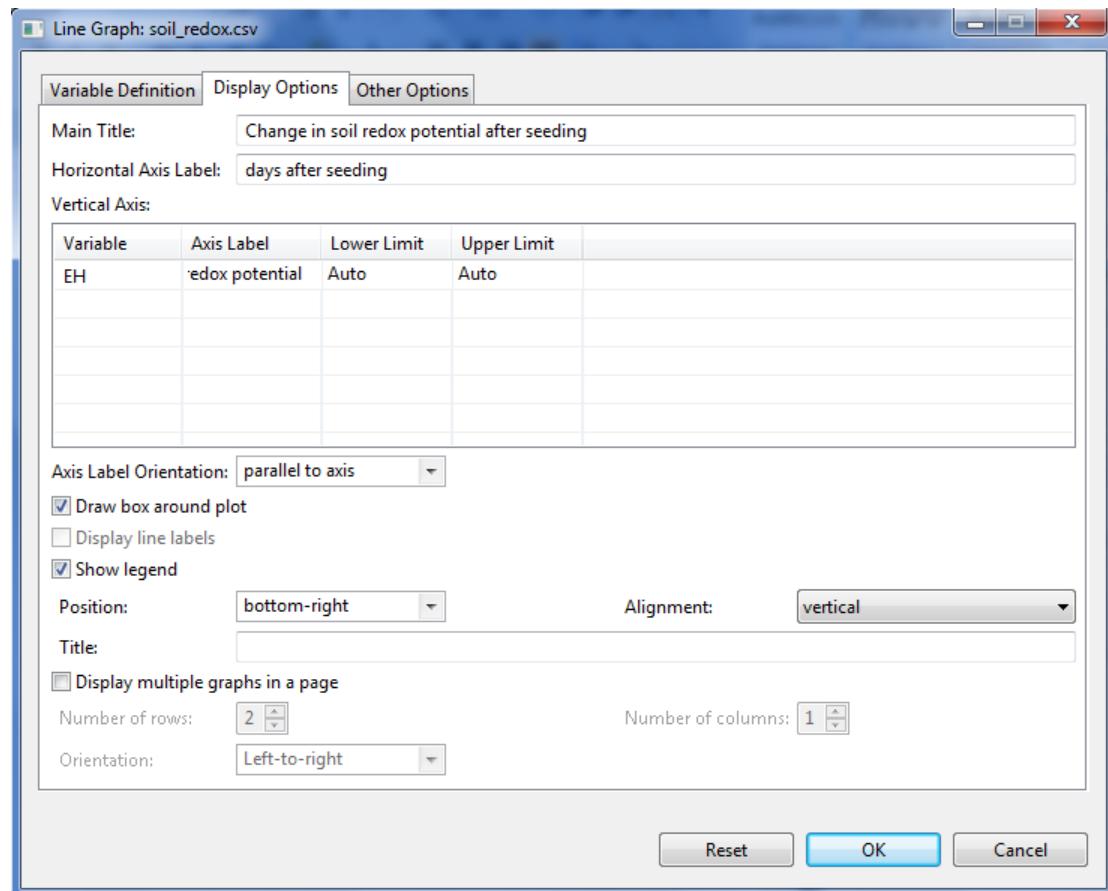
Display Error Bars

Error bars may be added based on one of two options: **confidence level** (default: 95%) and **standard error** (with multiplier 1, 2, or 3).

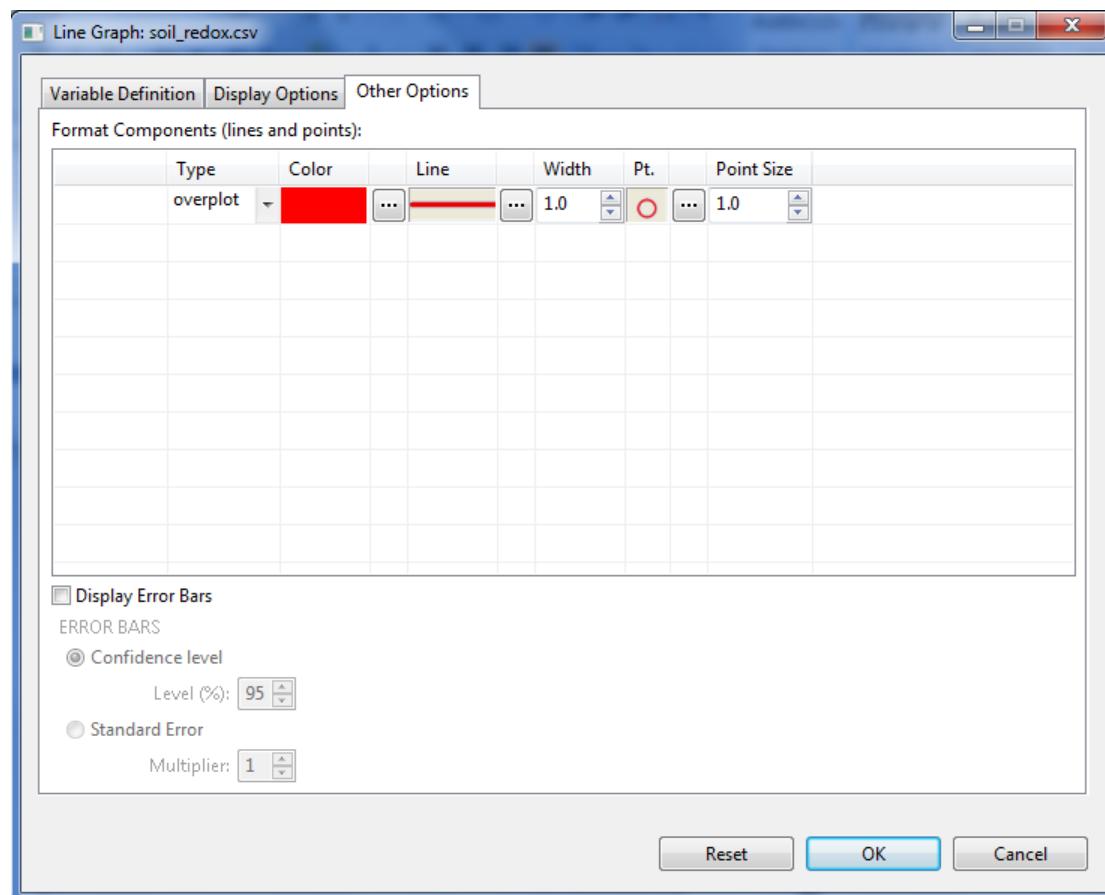
For the example, the **Variable Definition** tab should appear as:



Items in the **Display Options** tab may be supplied as:

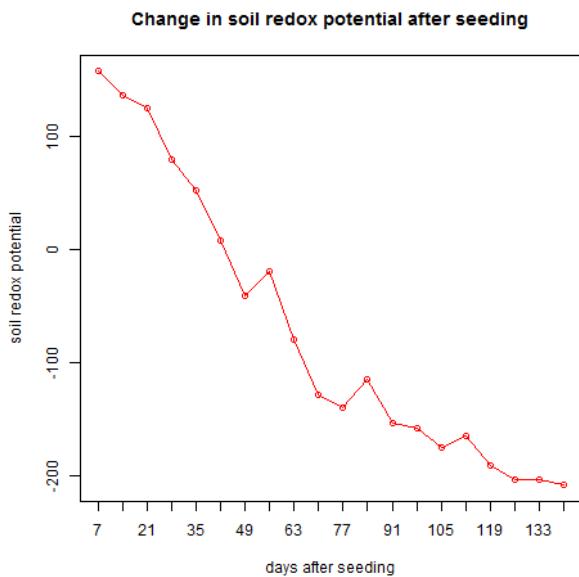


Components in the **Other Options** tab may be edited:



- Click the **OK** button to create the graph. The **Histogram** dialog box will be minimized and the Result Viewer will be activated.

Sample graph as displayed in the *Graph* page of the Result Viewer:



Pie Chart

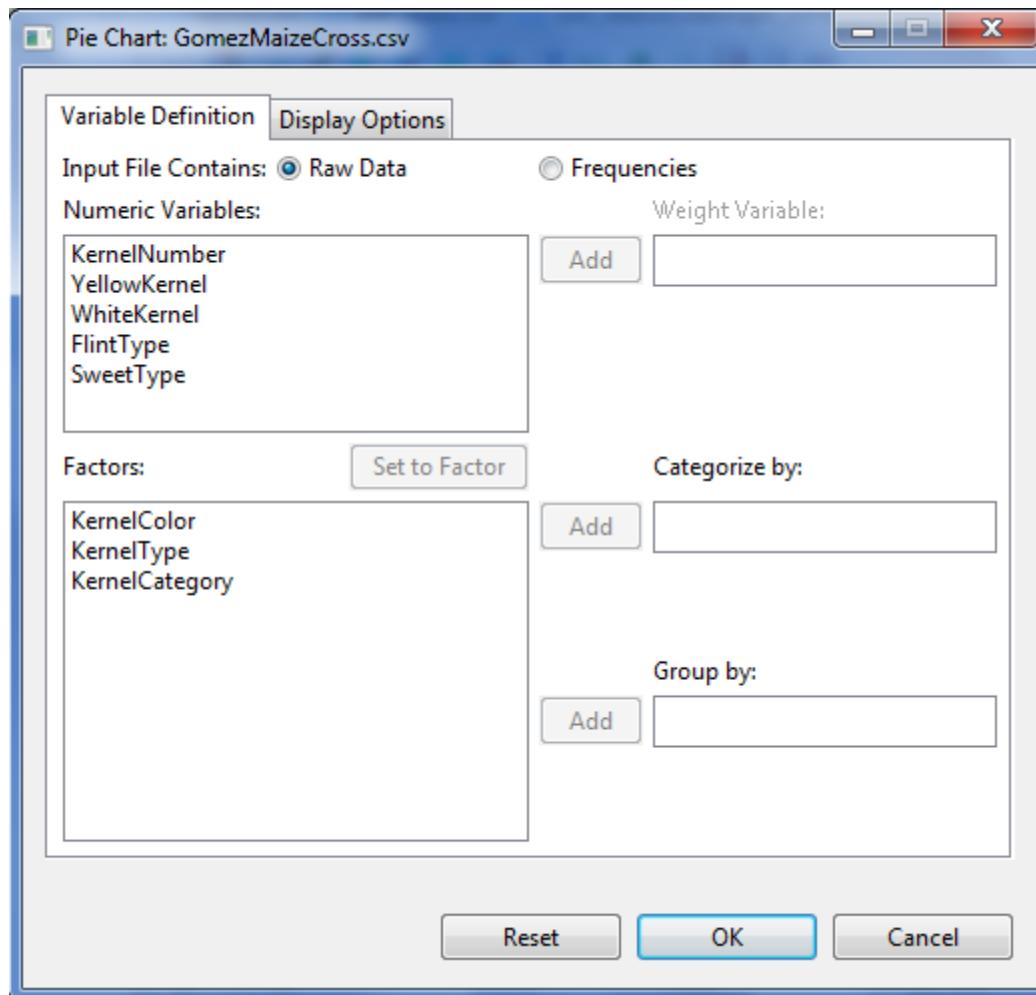
Pie charts are useful for presenting how a data set is divided into different sections or groups. In STAR, pie charts can be created from raw data or summaries (containing frequencies).

The steps to create a pie chart are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *GomezMaizeCross* from STAR Package. See the *Importing Data from Package* section of this user's manual for instructions on how to do this.

- Click **Graphs | Pie Chart....** The **Pie Chart** dialog box will appear.



- Specify the inputs for the required fields and desired options.

Variable Definition Tab

Input file contains

There are two options available, **Raw Data** (default option) and **Frequencies**. If the default option is used, the percentages (or frequencies) are determined for the category variable, otherwise, values from the frequencies are plotted directly.

Weight Variable

This item is enabled if **Input file contains Frequencies** is selected. This will contain the frequencies variable.

Categorize by

This will determine the sections for partitioning the pie. This field is required.

Group by

This field is optional. If a factor is specified, a pie chart is going to be generated per category of the factor. Entries should come from the **Factors** list box.

Display Options Tab

Main Title

This field will contain the title for the graph.

Format slices

In this section, changes to the formats (color and fill pattern) of the slices may be made.

Under **Labels**, available options are: none, counts, percentages, levels, both counts and levels and both levels and percentages.

Draw box around plot

If checked, the graph will be enclosed in a box.

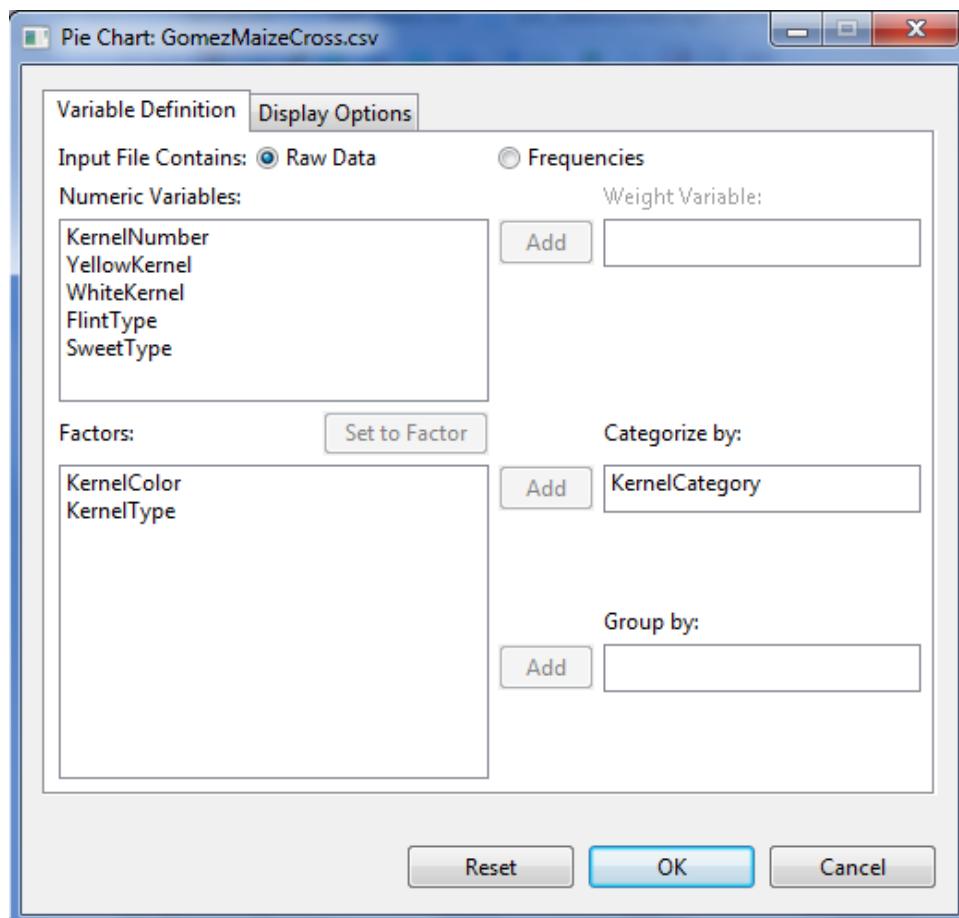
Show legend

If checked, a legend will be added to the graph. It may or may not have a title. The legend may be placed in any of nine positions within the graph: bottom, bottom-left, bottom-right, center, left, right, top, top-left, top-right. The items on the legend can be listed in one or more columns, the maximum of which is the number of levels of the category variable.

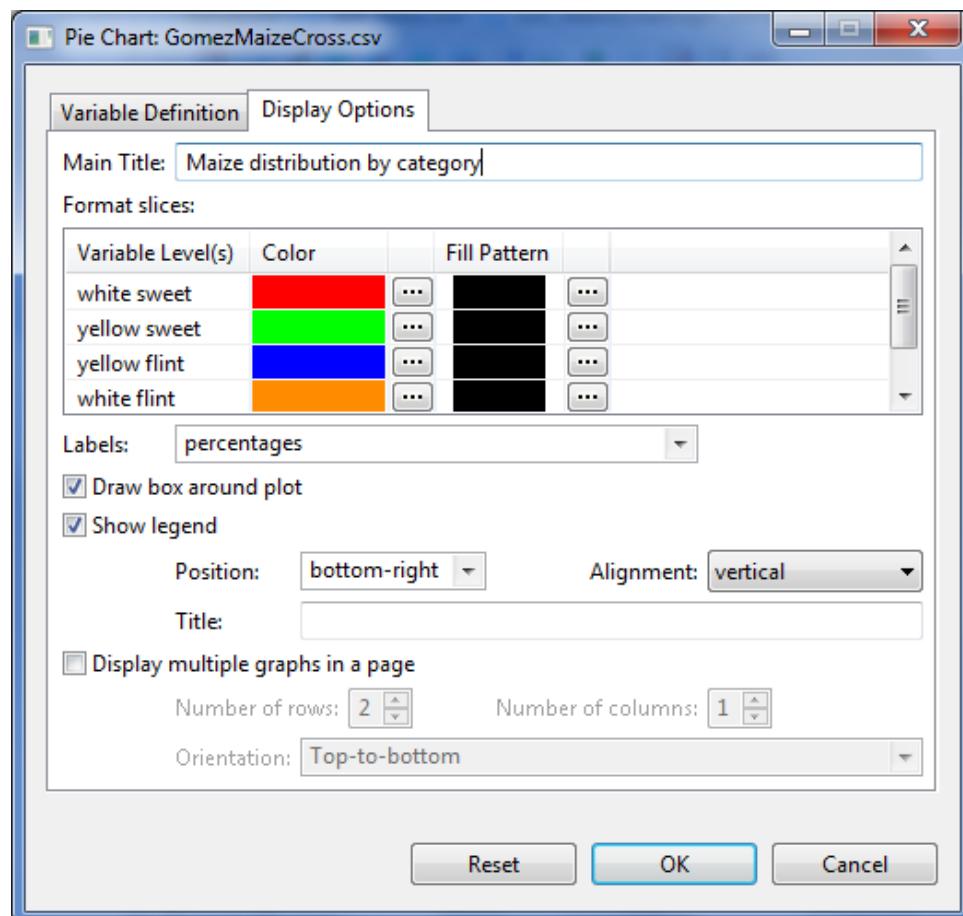
Display multiple graphs in a page

If checked, several graphs will be displayed on a single page, based on the specified number of rows and columns and orientation. At most 5 rows and 5 columns of graphs can be fitted into a single page, arranged either from *Top-to-bottom* or *Left-to-right*.

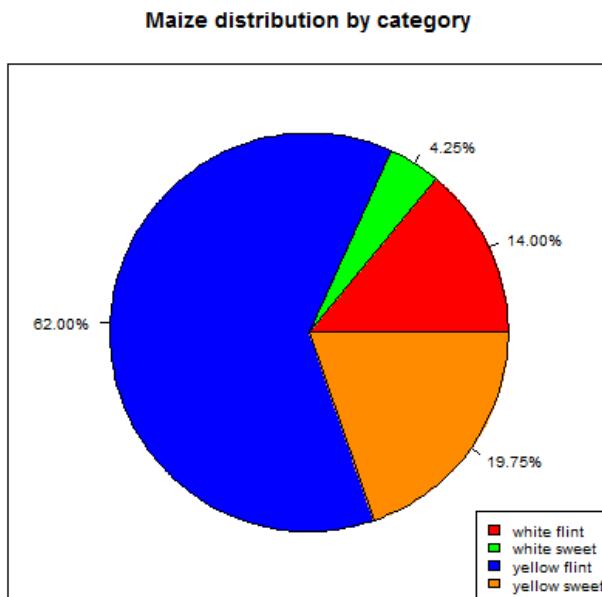
For the example, the **Variable Definition** tab should appear as illustrated:



The **Display Options** Tab may be edited:



Sample graph as displayed in the *Graph* page of the Result Viewer:



Scatter Plot

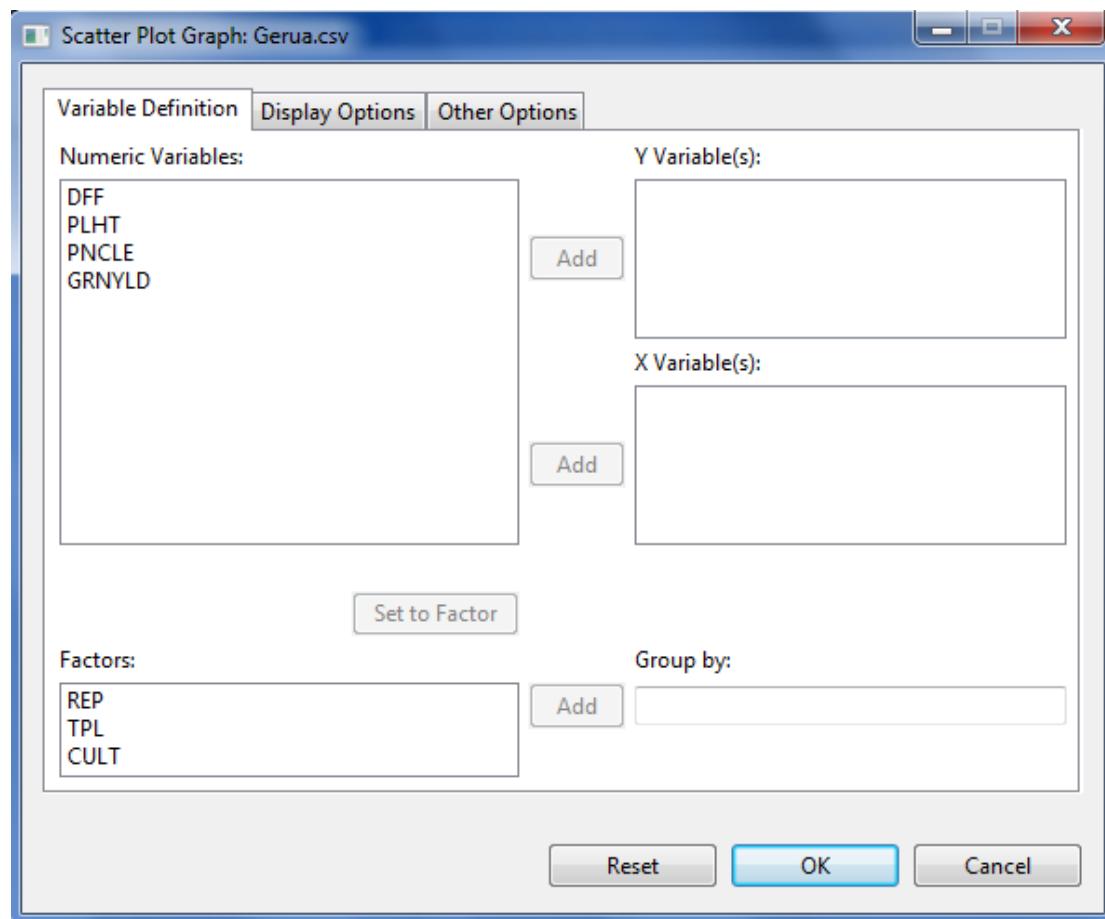
Scatter plots are useful for describing relationship between a pair of variables.

The steps to create a scatter plot are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *Gerua* from STAR Package. See the *Importing Data from Package* section of this user's manual for instructions on how to do this.

- Click **Graphs | Scatter Plot....** The **Scatter Plot** dialog box will appear.



- Specify the inputs for the required fields and desired options.

Variable Definition Tab

Y Variable(s)

At least one entry should be specified for the graph to be created. Entries should come from the **Numeric Variables** dialog box.

X Variable(s)

At least one entry should be specified for the graph to be created. Entries should come from the **Numeric Variables** dialog box.

Group by

This field is optional. If a factor is specified, a graph is going to be generated per category of the factor. Entry should come from the **Factors** list box.

Display Options Tab

Main Title

This field will contain the title for the graph.

Format Axes

The label and limits for the numeric variable(s) can be edited here. Default axis label is the name of the numeric variable. Lower and upper limits, if not specified, are determined from the data.

Axis Label Orientation

This option will determine how axis labels are displayed: *parallel to axis*, *horizontal*, *vertical*, and *perpendicular to axis*.

Format Points

In this section, changes to the formats (symbol, size, and color) of the points may be made.

Draw box around plot

If checked, the graph will be enclosed in a box.

Display multiple graphs in a page

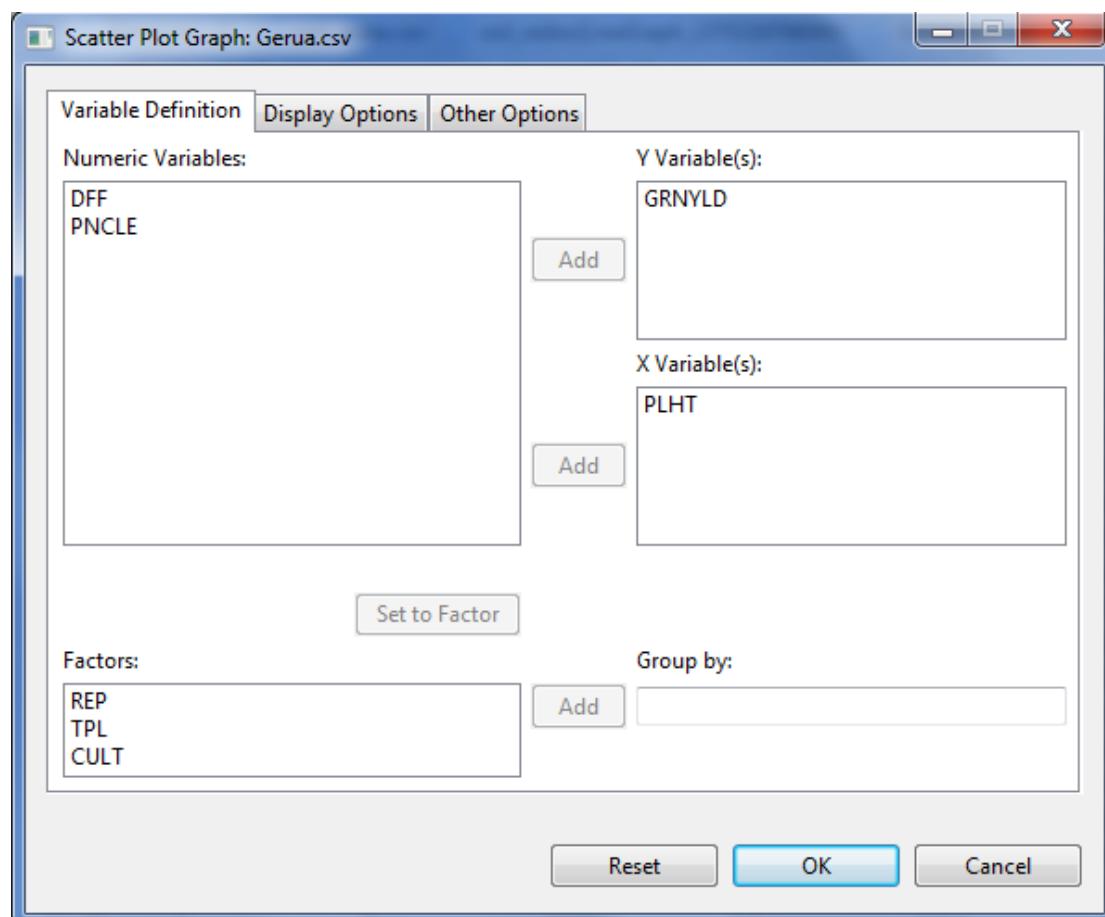
If checked, several graphs will be displayed on a single page, based on the specified number of rows and columns and orientation. At most 5 rows and 5 columns of graphs can be fitted into a single page, arranged either from *Top-to-bottom* or *Left-to-right*.

Other Options Tab

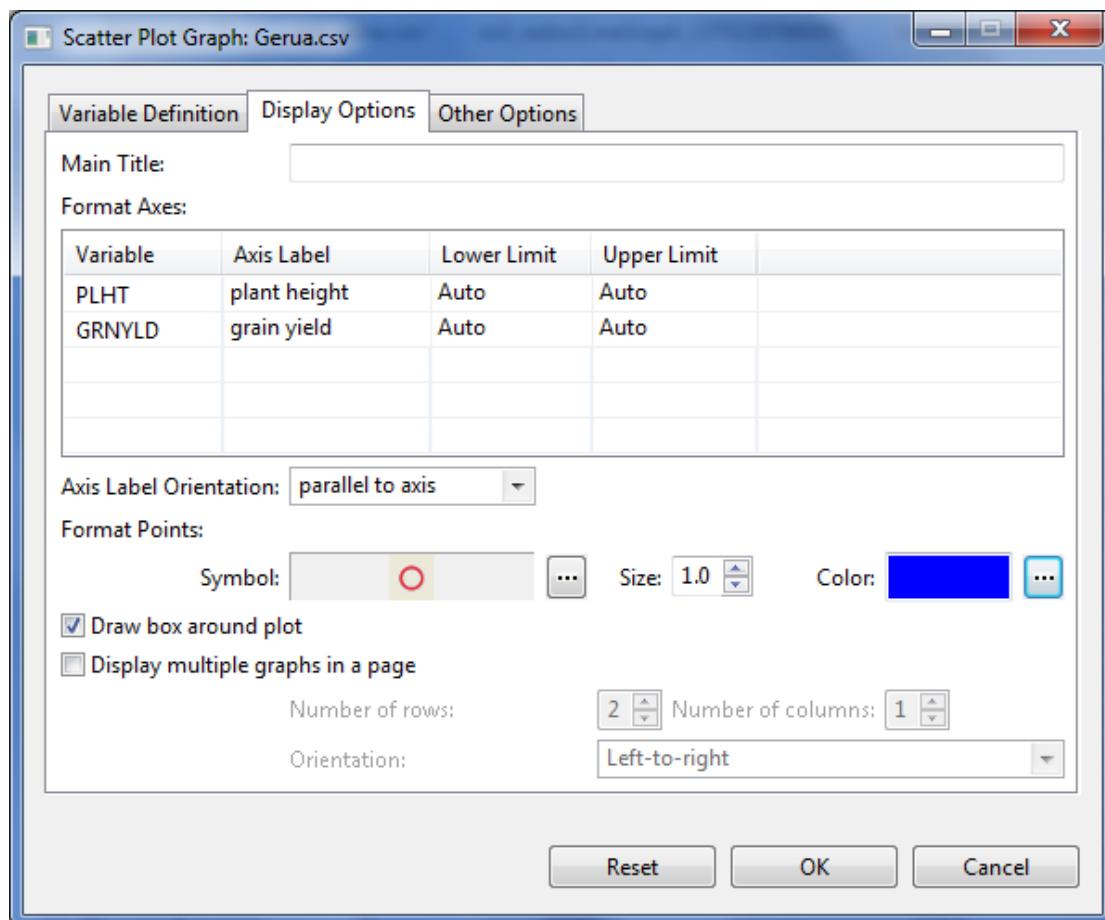
Plot regression line

If this option is checked, an estimated regression line is plotted over the points. If checked, type, width, and color of line may be edited. In addition, the *R-squared* and *p-value* may also be displayed (*bottom*, *bottom-left*, *bottom-right*, *center*, *left*, *right*, *top*, *top-left*, or *top-right* position).

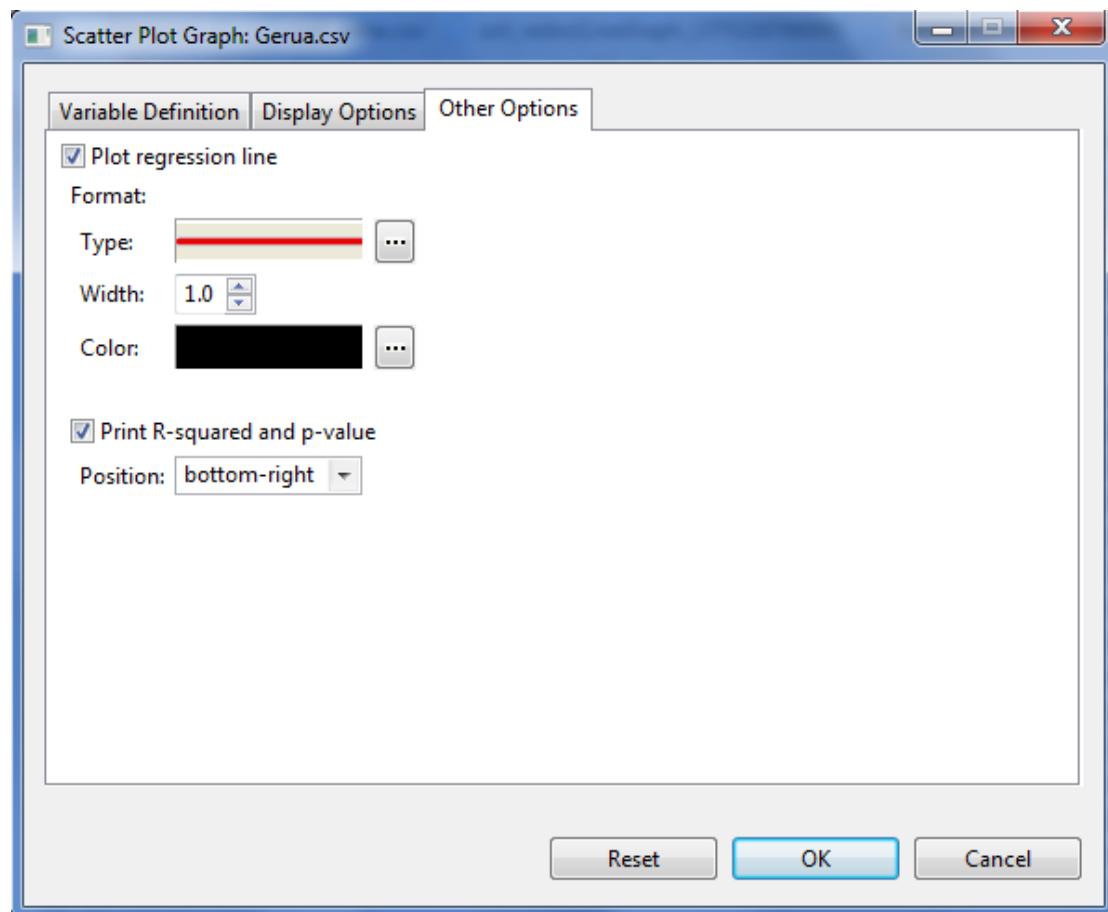
For the example, suppose we want to create a scatter plot for plant height (*PLHT*) and grain yield (*GRNYLD*). The completed **Variable Definition** tab should appear as shown below:



Items in the **Display Options** tab may be supplied as illustrated:

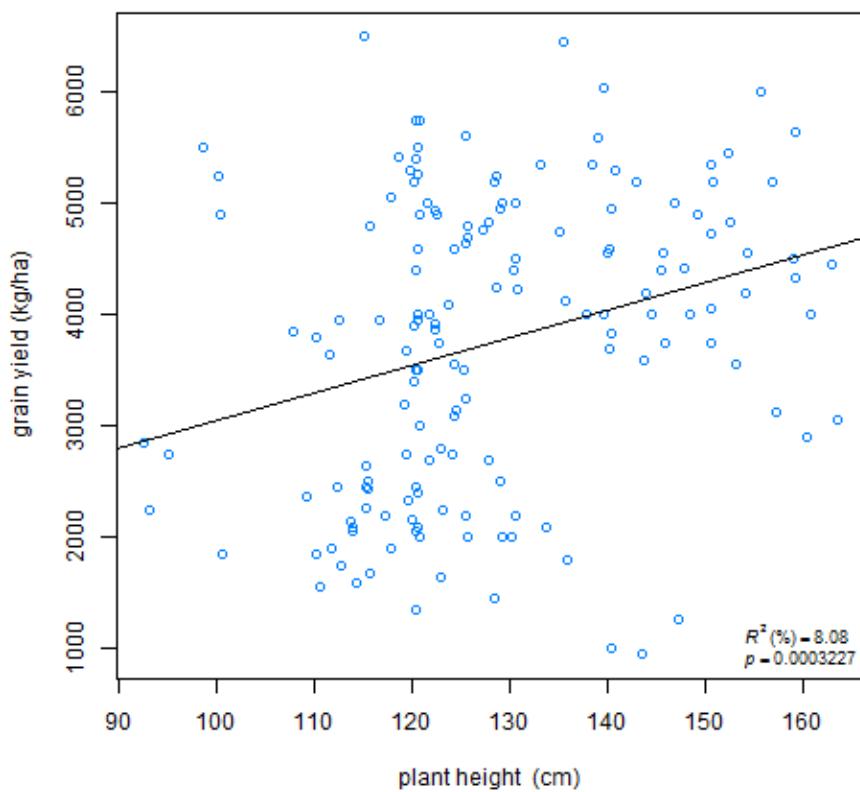


Items in the **Other Options** tab may be supplied as illustrated:



- Click the **OK** button to create the graph. The **Scatter Plot** dialog box will be minimized and the Result Viewer will be activated.

Sample graph as displayed in the **Graph** tab of the **Result Viewer**:



Descriptive Statistics

The **Descriptive** procedure performs univariate summary statistics for several variables and/or per levels of grouping variable(s).

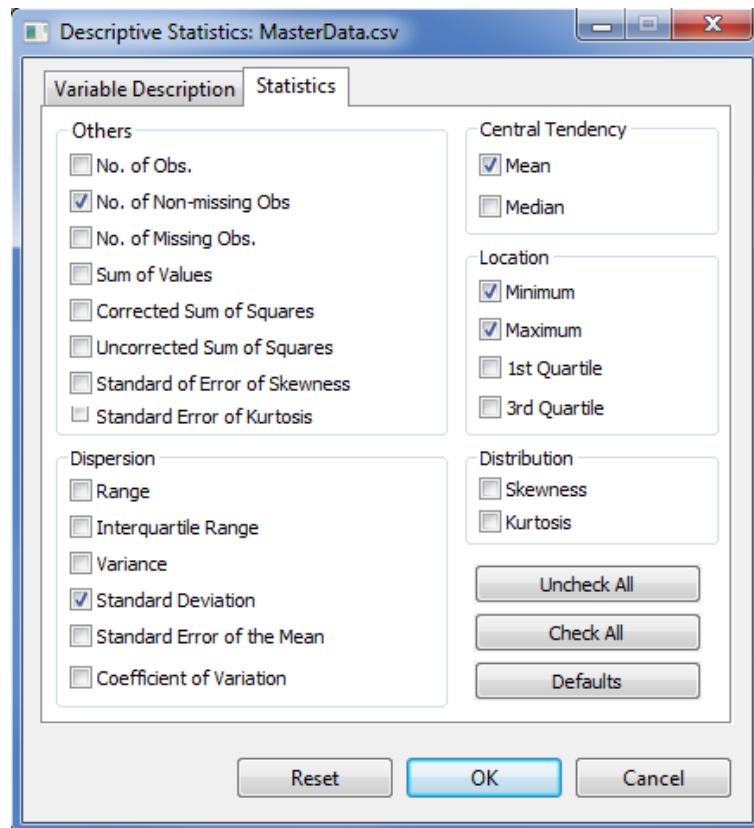
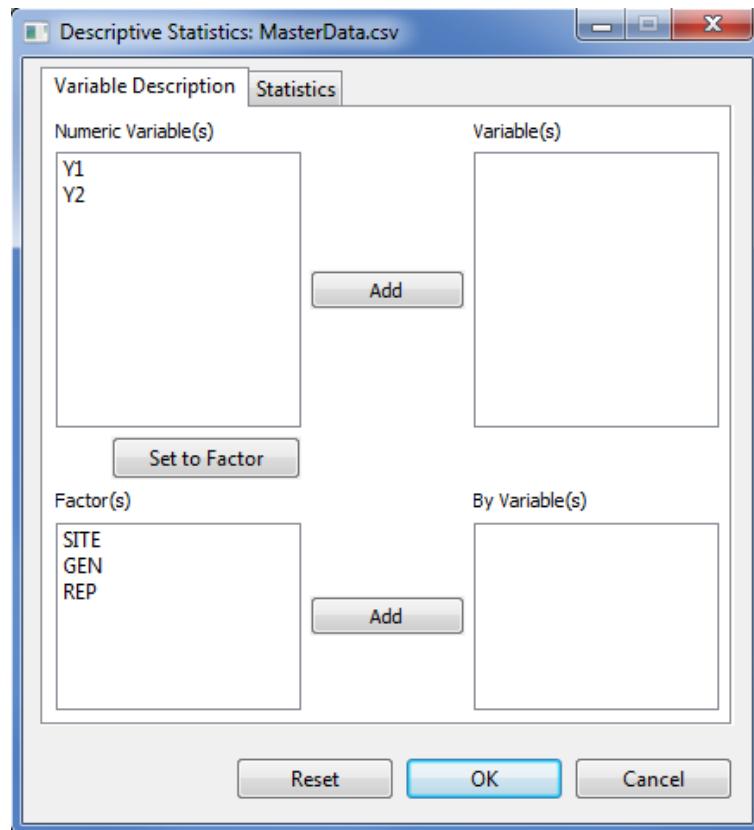
The steps to obtain numerical descriptive measures are listed below:

- On the **Project Explorer** panel, double-click the file to be used for the analysis to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

Opening the data for the first time, *REP* field is regarded by STAR as numerical variable. To change its type, choose the variable *REP* and click on the **Set to Factor** button.

- Click **Analyze | Descriptive Statistics | Descriptive....** The **Descriptive Statistics** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Variable Description Tab

This tab lets the user identify the variables and grouping variable(s) where descriptive statistics will be computed.

Variable(s)

At least one entry is needed for the analysis to proceed. Entries should come from the **Numeric Variable(s)** list box.

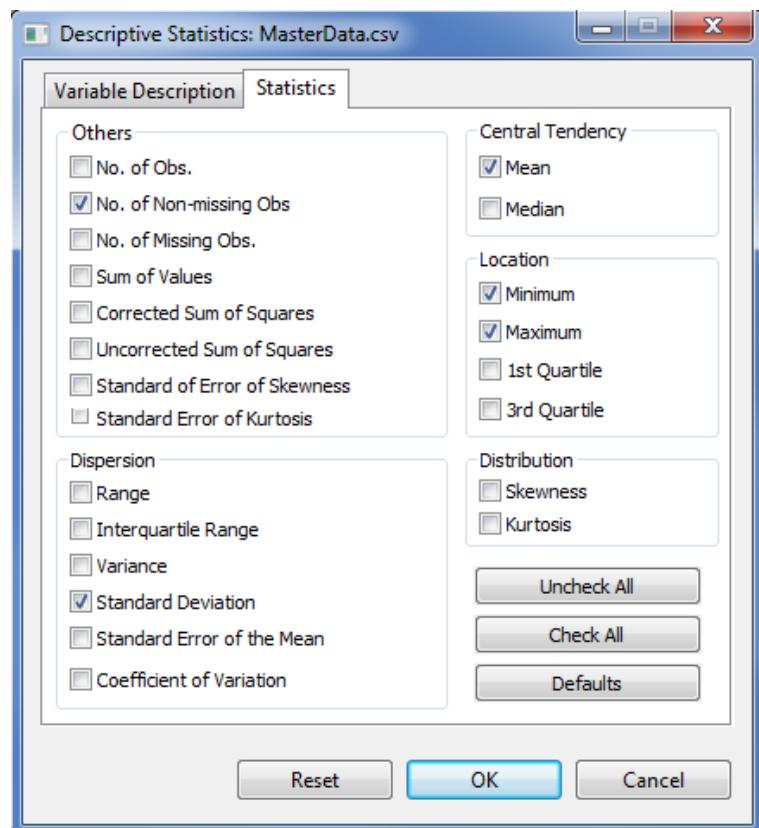
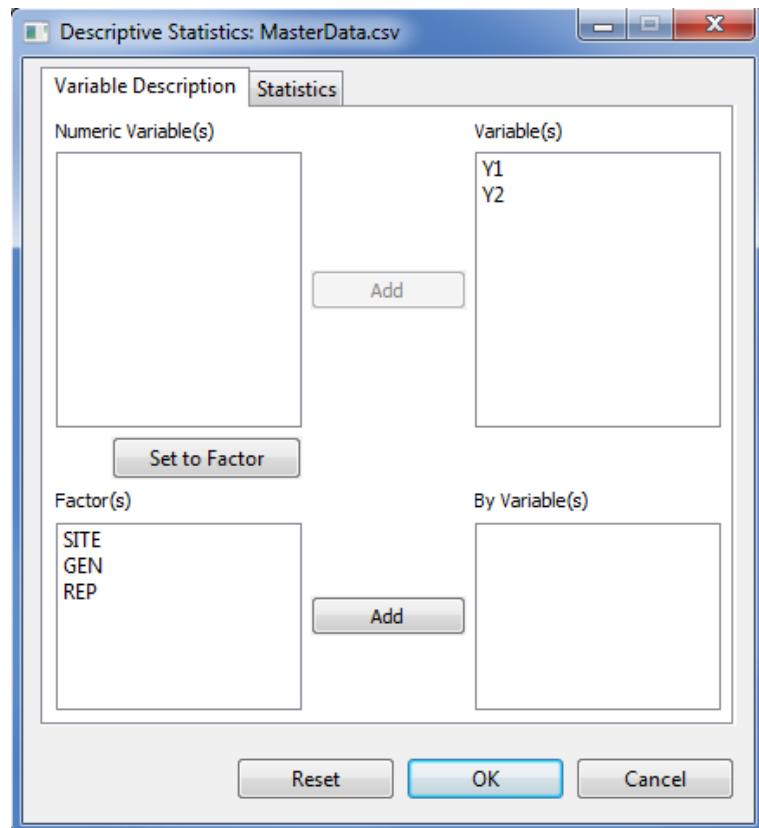
By Variable(s)

This field is optional. If a **By Variable(s)** is specified, it will generate descriptive statistic per category of the **By Variable**. Entries should come from the **Factor(s)** list box.

Statistics Tab

This tab lets the user identify the numerical descriptive measures to be displayed. By default, the procedure will display five statistics, namely: number of non-missing observations (NNMissObs), minimum (min), maximum (max), mean, and standard deviation (stdDev). To request for all statistics, click the **Check All** button of the **Statistics** tab. User may click on the **Uncheck All** button to remove previously selected items and start with a new selection. At least one statistics should be checked to the analysis to proceed.

For the example, suppose we want to compute the descriptive statistics of variables Y1 and Y2. The completed **Variable Description** tab and **Statistics** tab should appear as shown:

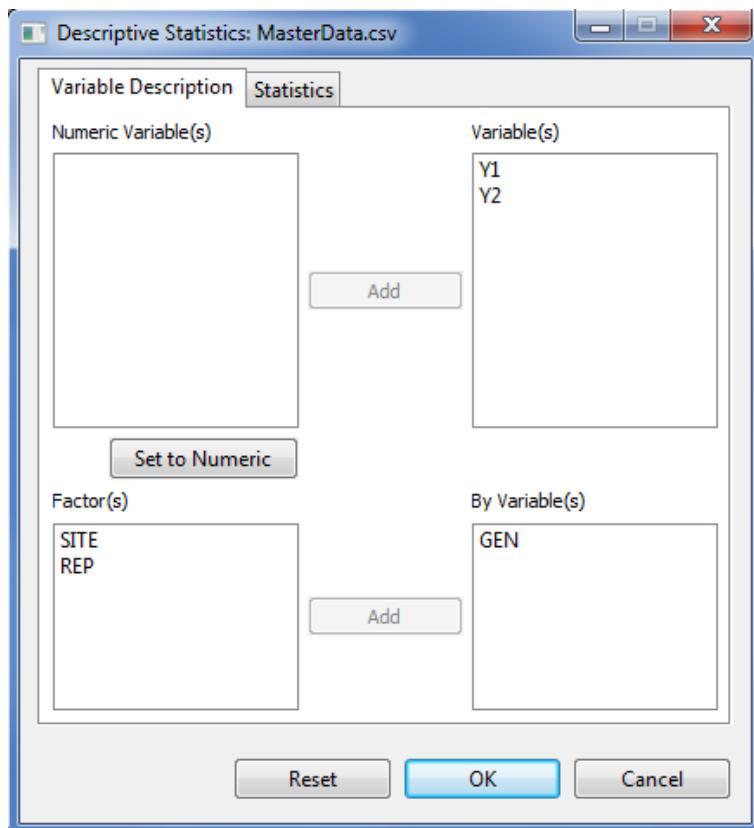


- Click the **OK** button to generate the summary statistics. The **Descriptive Statistics** dialog box will be minimize and STAR activates the **Output** page of the **Result Viewer** tab. Sample output displayed in the **Output** page is shown below:

Descriptive Statistics

Variable	N_NonMissObs	Min	Max	Mean	StdDev
Y1	96	14.00	151.00	130.43	16.05
Y2	96	92.60	163.50	124.50	15.05

Suppose we want to compute the descriptive statistics for variables Y1 and Y2 per level of GEN. The completed **Variable Description** tab should appear as shown below:



- Click the **OK** button to generate the summary statistics. Sample output is shown below:

Descriptive Statistics

Variable	GEN	N_NonMissObs	Min	Max	Mean	StdDev
Y1	Gen1	12	113.00	144.00	132.33	10.50
Y1	Gen2	12	106.00	150.00	134.17	11.07
Y1	Gen3	12	116.00	148.00	132.42	9.46
Y1	Gen4	12	14.00	151.00	122.92	36.36
Y1	Gen5	12	106.00	145.00	129.42	11.84
Y1	Gen6	12	121.00	143.00	133.25	6.84
Y1	Gen7	12	109.00	149.00	130.00	12.05
Y1	Gen8	12	106.00	144.00	128.92	12.26
Y2	Gen1	12	92.60	163.50	128.87	20.06
Y2	Gen2	12	95.20	149.30	119.24	14.21
Y2	Gen3	12	93.10	148.40	120.09	15.03
Y2	Gen4	12	100.50	150.70	124.54	16.45
Y2	Gen5	12	98.60	160.70	125.22	16.83
Y2	Gen6	12	110.20	130.60	119.24	5.54
Y2	Gen7	12	117.90	159.30	130.05	13.78
Y2	Gen8	12	111.70	150.90	128.72	13.98

Cross Tabulation

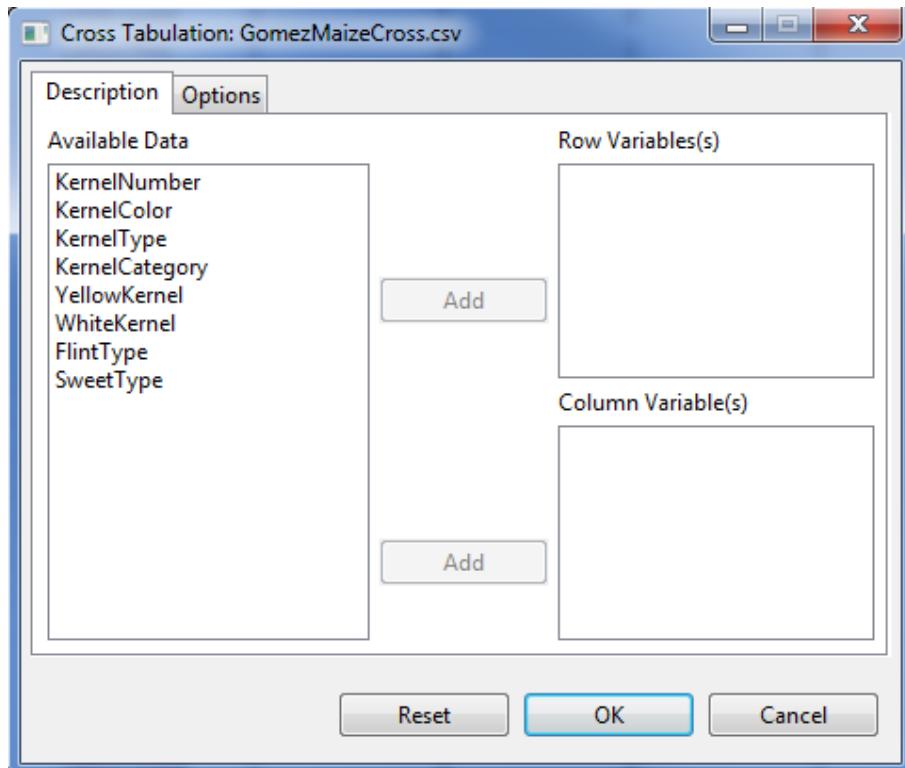
The cross tabulation forms two-way tables. By default, the observed frequency is displayed.

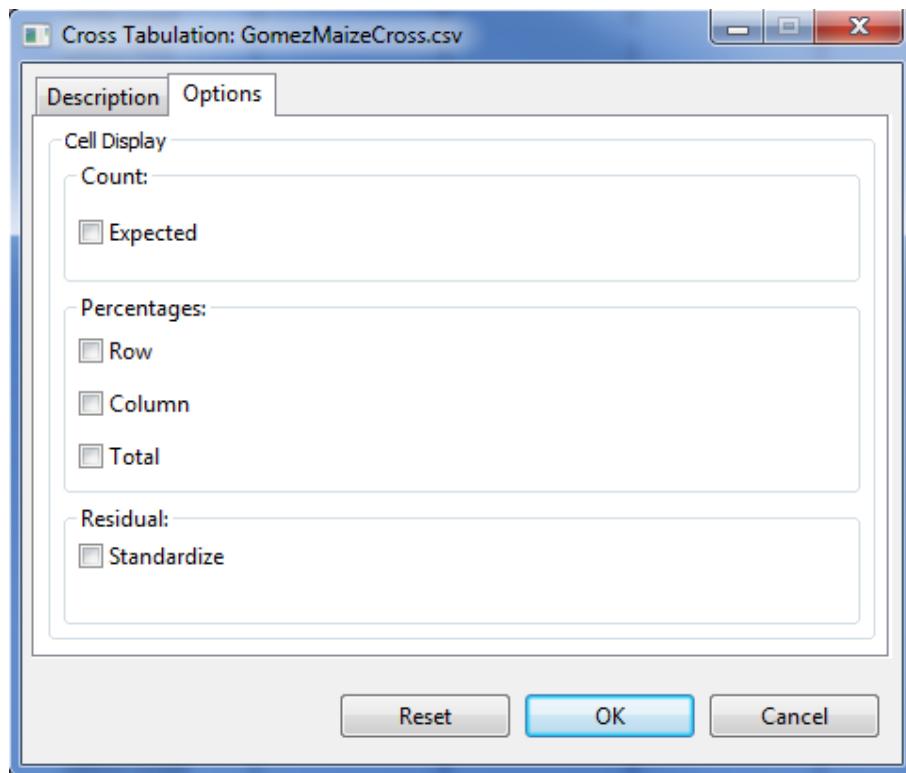
The steps to perform Cross Tabulation are listed below:

- On the **Project Explorer** panel, double-click the file to be used for analysis to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *GomezMaizeCross.csv* from the package. To import data from the package, see *Importing Data from Package* section of this user's manual.

- To generate cross table, from the main window click **Analyze | Descriptive Statistics | Cross Tabulation....** The **Cross Tabulation** dialog box will appear.





- Specify the required field and the appropriate options for the analysis.

Description Tab

This tab lets the user identify the variables where cross tabulation will be performed.

Row Variable(s)

At least one entry is need for the analysis to proceed.

Column Variable(s)

At least one entry is need for the analysis to proceed.

Options Tab

This tab lets the user identify the additional output that will be displayed.

Expected Count

If this option is selected, the expected frequency will be displayed.

Row Percentages

If this option is selected, the row percentage will be displayed.

Column Percentages

If this option is selected, the column percentage will be displayed.

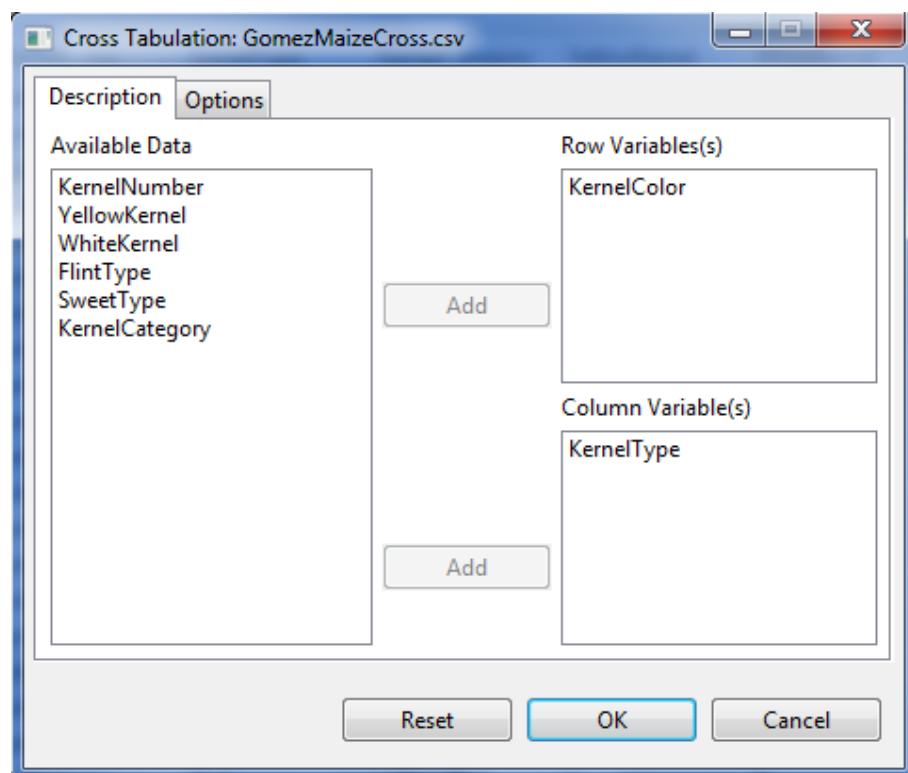
Total Percentages

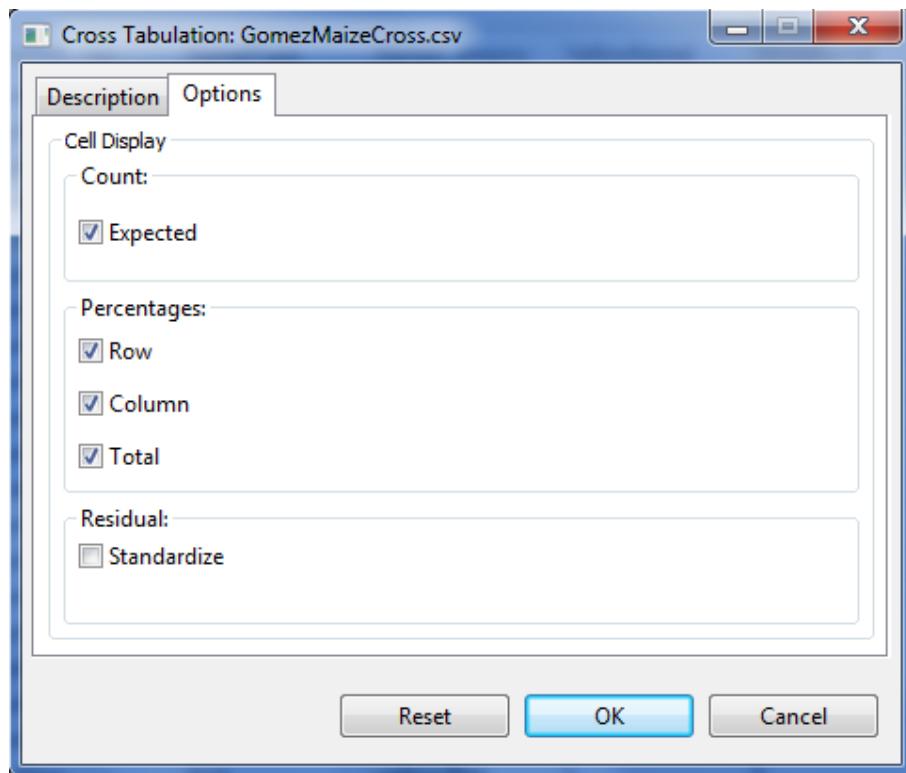
If this option is selected, the total percentage will be displayed.

Standardize Residual

If this option is selected, the standardize residual will be displayed.

For the example, the completed **Variable Description** tab and **Options** tab should appear as shown below:





- Click the **OK** button. The **Cross Tabulation** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab. Sample output of the analysis is shown below:

CROSS TABULATION

Table of KernelColor by KernelType

		KernelType		
		flint	sweet	Total
white	Obs Freq	1121	34	146
	Expected Freq	110.9600	35.0400	146.0000
	% within KernelColor	76.7123	23.2877	100.0000
	% within KernelType	18.4211	17.7083	
	% of Total	14.0000	4.2500	18.2500
yellow	Obs Freq	496	158	654
	Expected Freq	497.0400	156.9600	654.0000
	% within KernelColor	75.8410	24.1590	100.0000
	% within KernelType	81.5789	82.2917	
	% of Total	62.0000	19.7500	81.7500
Total	Obs Freq	608	192	800
	Expected Freq	608.0000	192.0000	800.0000
	% within KernelType	100.0000	100.0000	
	% of Total	76.0000	24.0000	100.0000

Normality Test

The Test for Normality menu display normality test for at least one numeric variable. There are five procedures available in this menu, namely: Shapiro-Wilk, Shapiro-Francia, Lilliefors (Kolmogorov-Smirnov), Cramer-Von Mises and the Anderson-Darling test for normality.

The Shapiro-Wilk test for normality can be used if the number of non-missing observations is between 3 and 5000. While the Shapiro-Francia test for normality can be used if the number of non-missing observations is between 5 and 5000. The Lilliefors test for normality can be used if the number of non-missing observations is greater than 4. The Cramer-Von Mises and the Anderson-Darling test for normality can both be used if the number of non-missing observations is greater than 7.

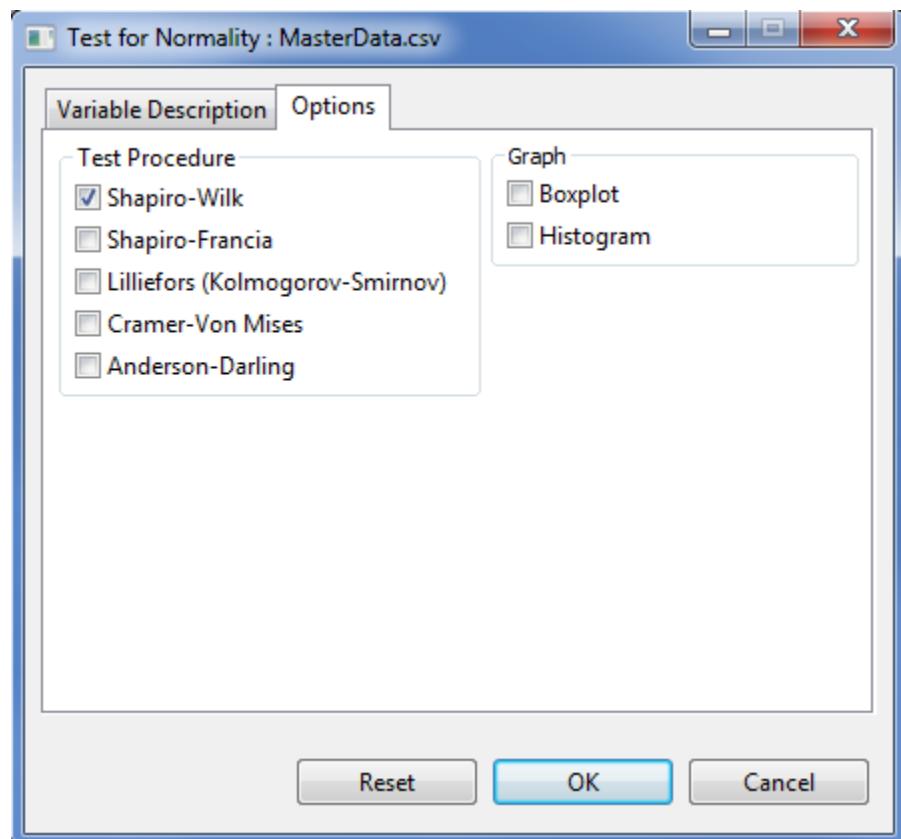
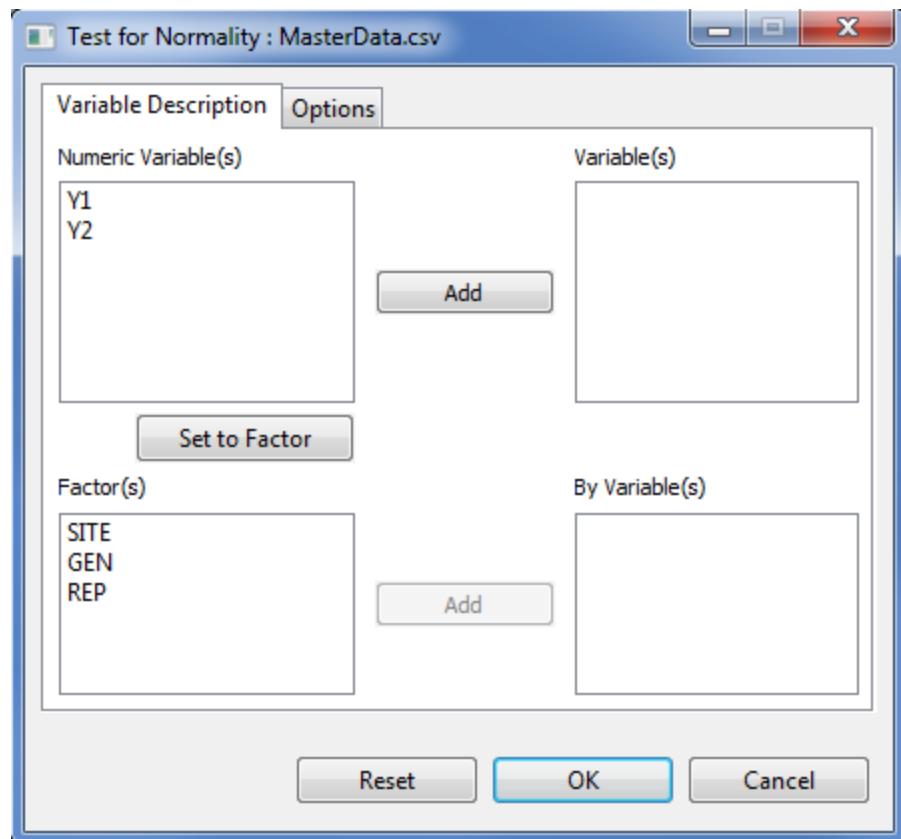
The steps to perform test for normality are listed below:

- On the **Project Explorer** panel, double-click the file to be used for the analysis to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

Opening the data for the first time, *REP* field is regarded by STAR as numerical variable. To change its type, choose the variable *REP* and click on the **Set to Factor** button.

- Click **Analyze | Descriptive Statistics | Normality Test....** The **Normality Test** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Variable Description Tab

This tab lets the user identify the variables and grouping variable(s) where test for normality will be performed.

Variables

At least one entry should be specified for the analysis to proceed. Entries for this list box should come from the **Numeric Variable(s)** list box.

By Variables

This field is optional. If a by variable(s) is specified, it will perform test for normality per level of the **By Variable(s)**. Entries for this list box should come from the **Factor(s)** list box.

Options Tab

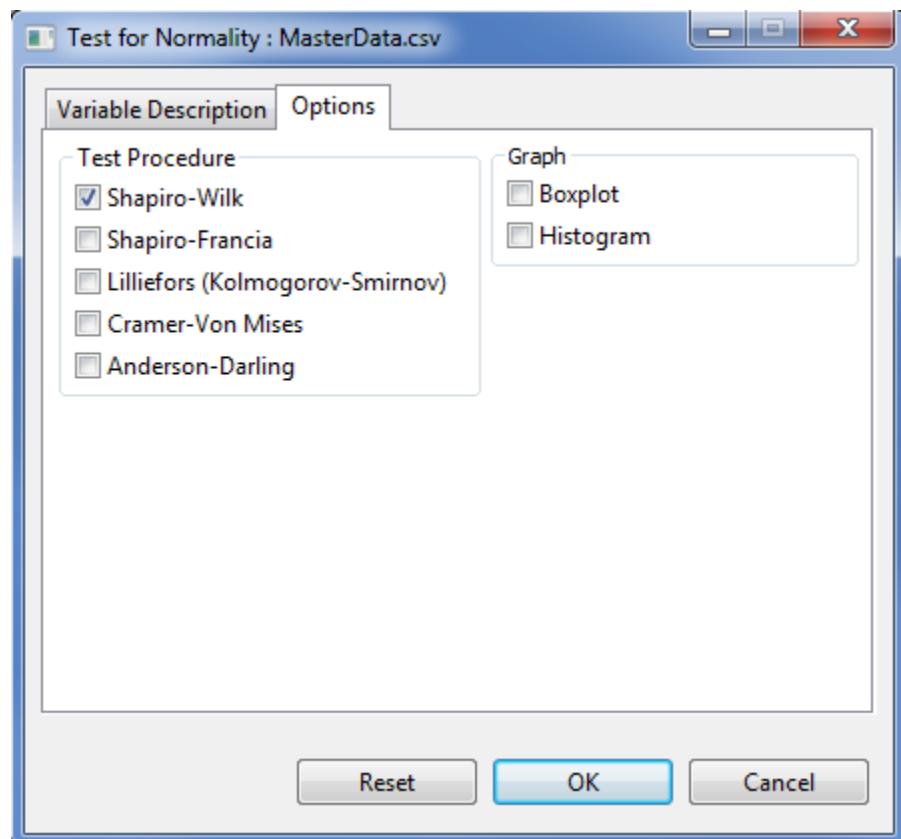
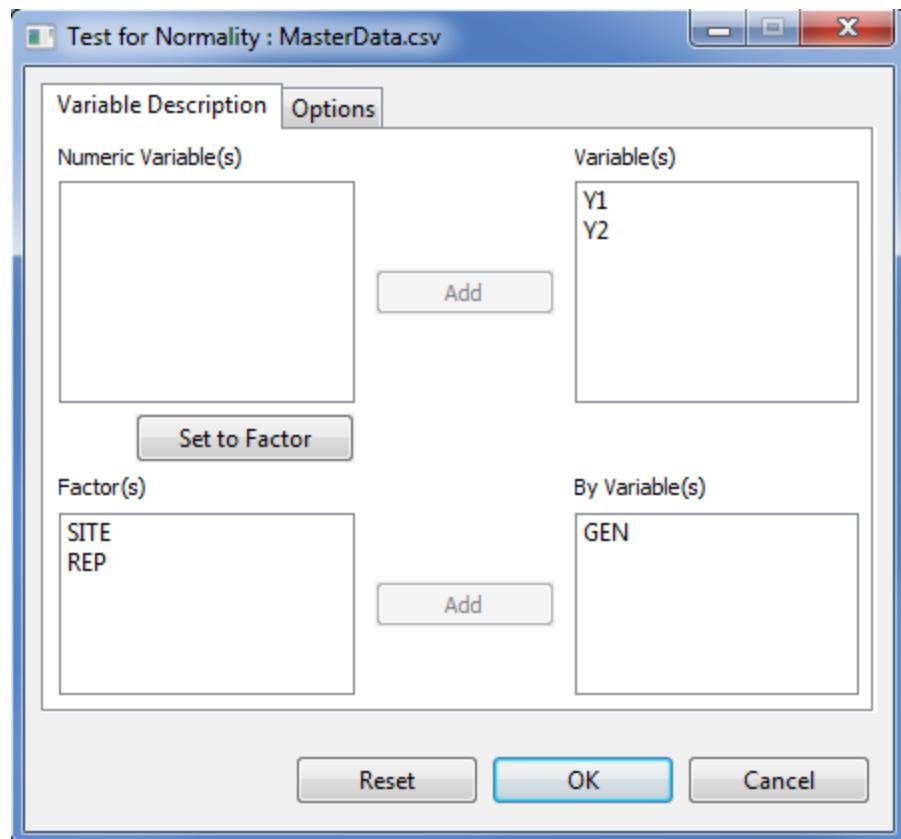
Test Procedure

This option is required. At least one test procedure should be selected for the analysis to proceed. By default, the Shapiro-Wilk test is selected.

Graph

The user has the option to display the box plot and the histogram.

For the example, suppose we want to perform normality test using Shapiro-Wilk test for variables *Y1* and *Y2* per level of *GEN*. The completed **Variable Description** tab and **Option** tab should appear as shown:

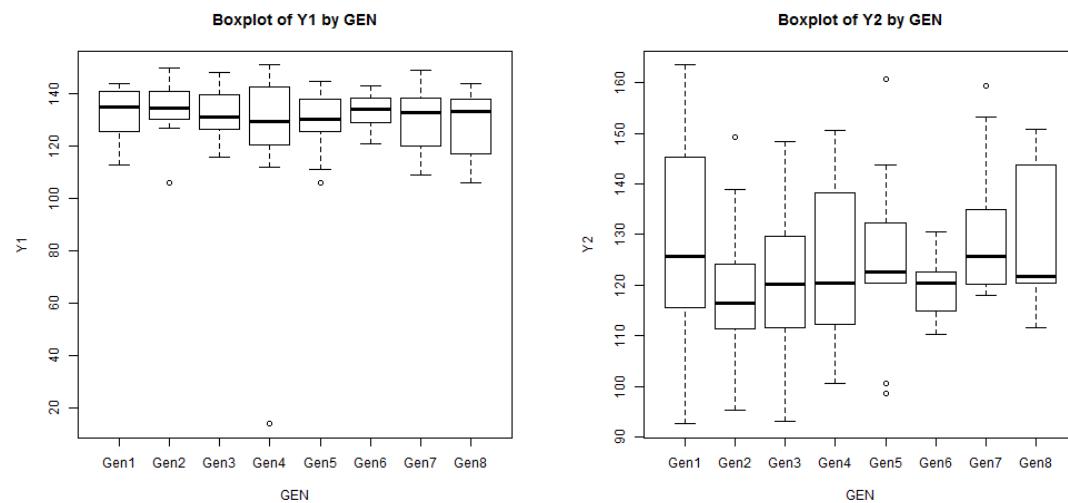


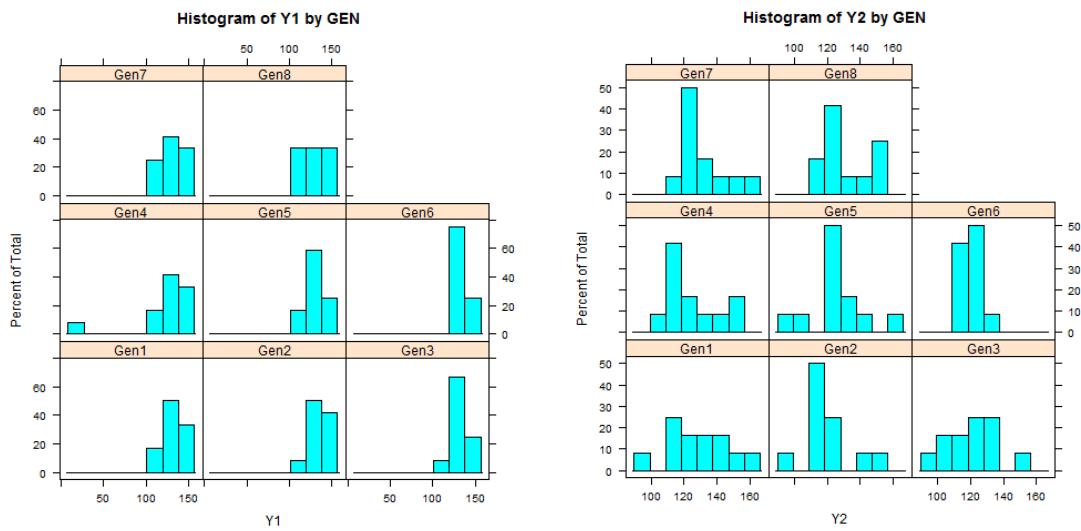
- Click the **OK** button to perform the test. The **Normality Test** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample text output displayed in the **Output** page is shown below:

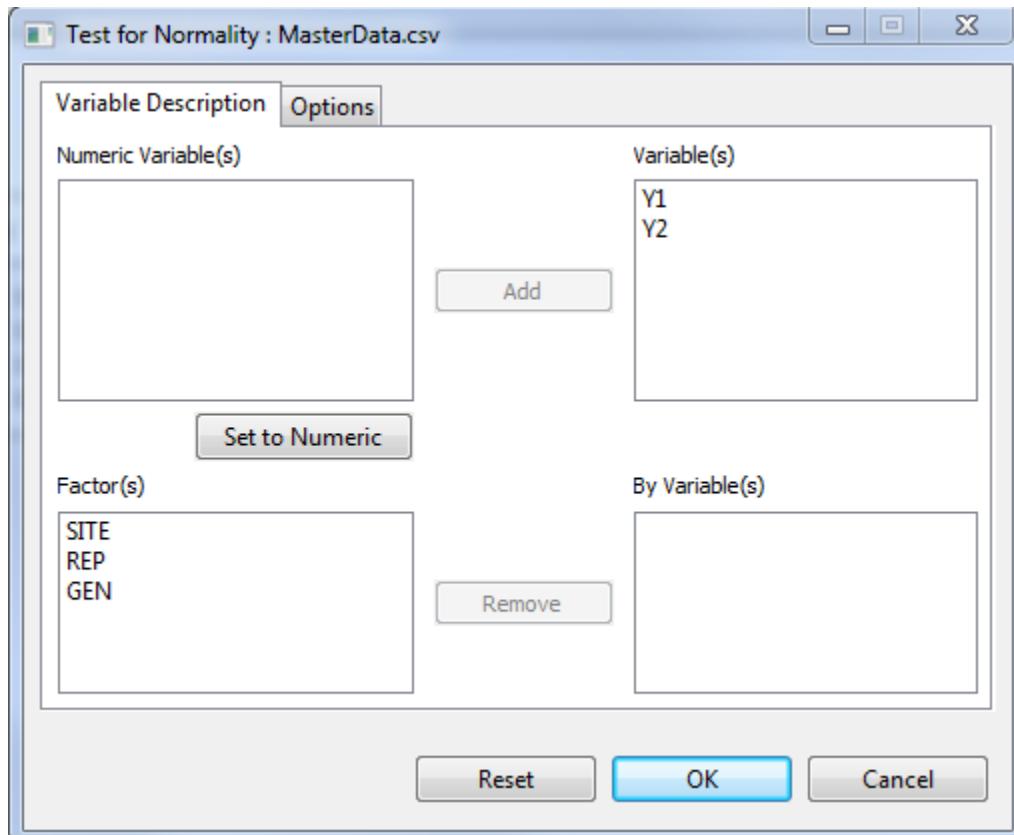
Test for Normality					
Grp	Level	Variable	Method	W Value	Pr (< W)
GEN	Gen1	Y1	Shapiro-Wilk	0.8914	0.1228
GEN	Gen2	Y1	Shapiro-Wilk	0.8947	0.1353
GEN	Gen3	Y1	Shapiro-Wilk	0.9560	0.7259
GEN	Gen4	Y1	Shapiro-Wilk	0.6460	0.0003
GEN	Gen5	Y1	Shapiro-Wilk	0.9268	0.3474
GEN	Gen6	Y1	Shapiro-Wilk	0.9645	0.8454
GEN	Gen7	Y1	Shapiro-Wilk	0.9618	0.8097
GEN	Gen8	Y1	Shapiro-Wilk	0.9020	0.1681
GEN	Gen1	Y2	Shapiro-Wilk	0.9700	0.9107
GEN	Gen2	Y2	Shapiro-Wilk	0.9341	0.4251
GEN	Gen3	Y2	Shapiro-Wilk	0.9811	0.9876
GEN	Gen4	Y2	Shapiro-Wilk	0.9200	0.2860
GEN	Gen5	Y2	Shapiro-Wilk	0.9314	0.3956
GEN	Gen6	Y2	Shapiro-Wilk	0.9450	0.5658
GEN	Gen7	Y2	Shapiro-Wilk	0.8047	0.0106
GEN	Gen8	Y2	Shapiro-Wilk	0.8613	0.0508

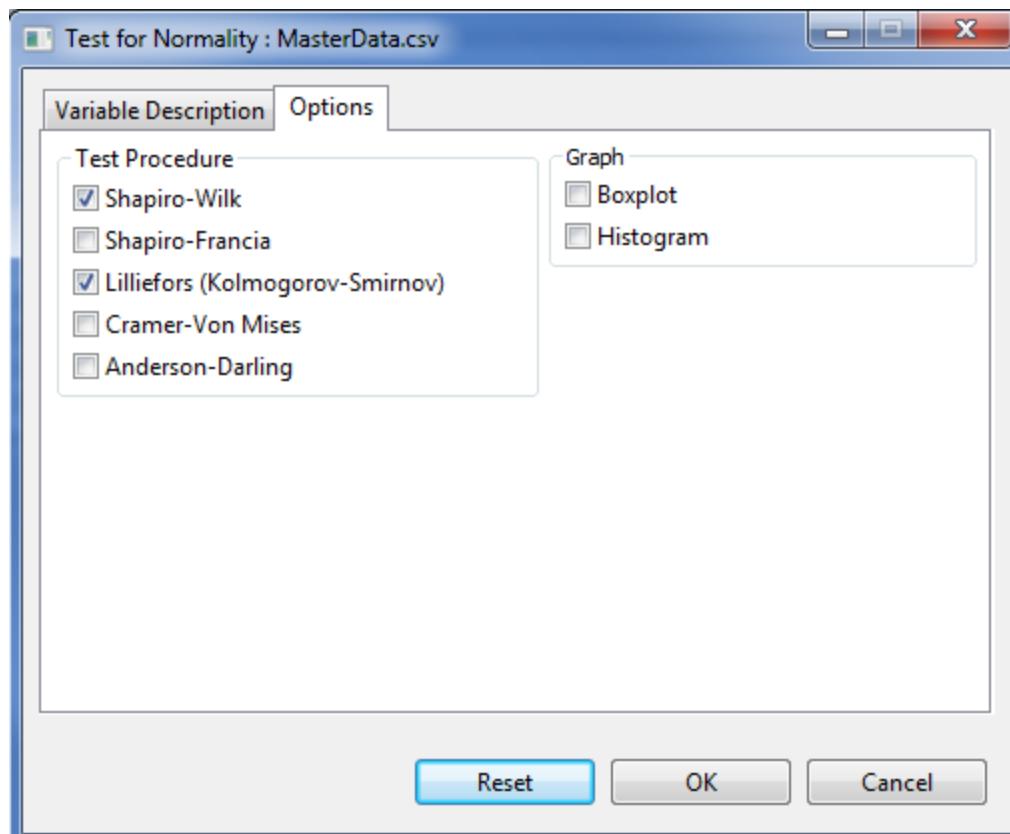
Sample graphical output displayed in the **Graph** page is shown below:





Suppose we want to perform normality test using Shapiro-Wilk and Kolmogorov-Smirnov for variables Y1 and Y2. The completed **Variable Description** and **Option** tab should appear as shown below:





Sample text output that will be displayed in the **Output** page is shown below:

```
Test for Normality
-----
Variable Method Statistic Value Prob p Value
-----
Y1      Shapiro-Wilk   W    0.6908 Pr(< W) 0.0000
Y2      Shapiro-Wilk   W    0.9522 Pr(< W) 0.0015
Y1      Kolmogorov-Smirnov D    0.1795 Pr(> D) 0.0000
Y2      Kolmogorov-Smirnov D    0.1281 Pr(> D) 0.0005
```


Homogeneity of Variances

The Homogeneity of Variances menu perform test for equally of variances across groups. There are two procedures available in this menu, namely: Bartlett's test and the Levene's test.

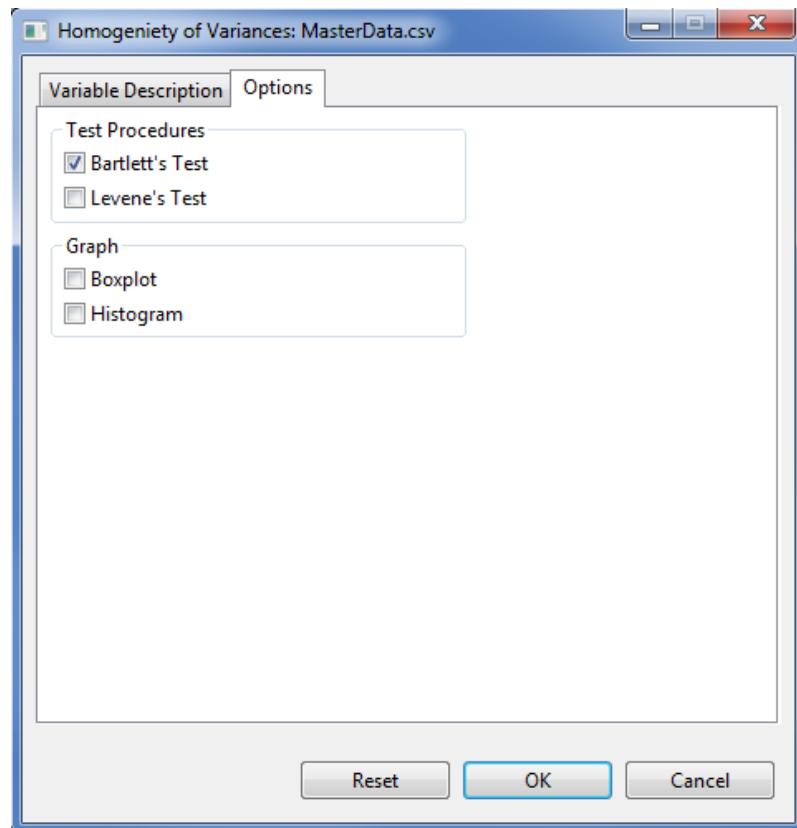
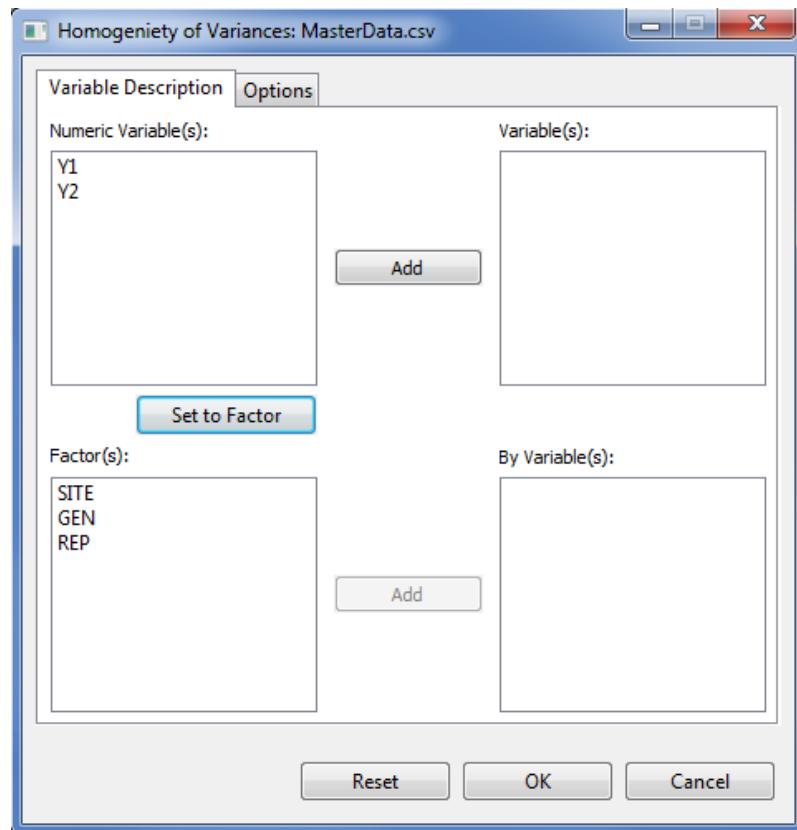
The steps to perform the test are listed below:

- On the **Project Explorer** panel, double-click the file to be used for the analysis to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

Opening the data for the first time, *REP* field is regarded by STAR as numerical variable. To change its type, choose the variable *REP* and click on the **Set to Factor** button.

- Click **Analyze | Descriptive Statistics | Homogeneity of Variance....** The **Homogeneity of Variance** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Variable Description Tab

This tab lets the user identify the variables and grouping variable(s) where test for homogeneity of variances will be performed.

Variables

At least one entry should be specified for the analysis to proceed. Entries must come from the **Numeric Variable(s)** list box.

By Variables

At least one entry should be specified for the analysis to proceed. Entries specified here should contain two or more levels and must come from the **Factor(s)** list box.

Options Tab

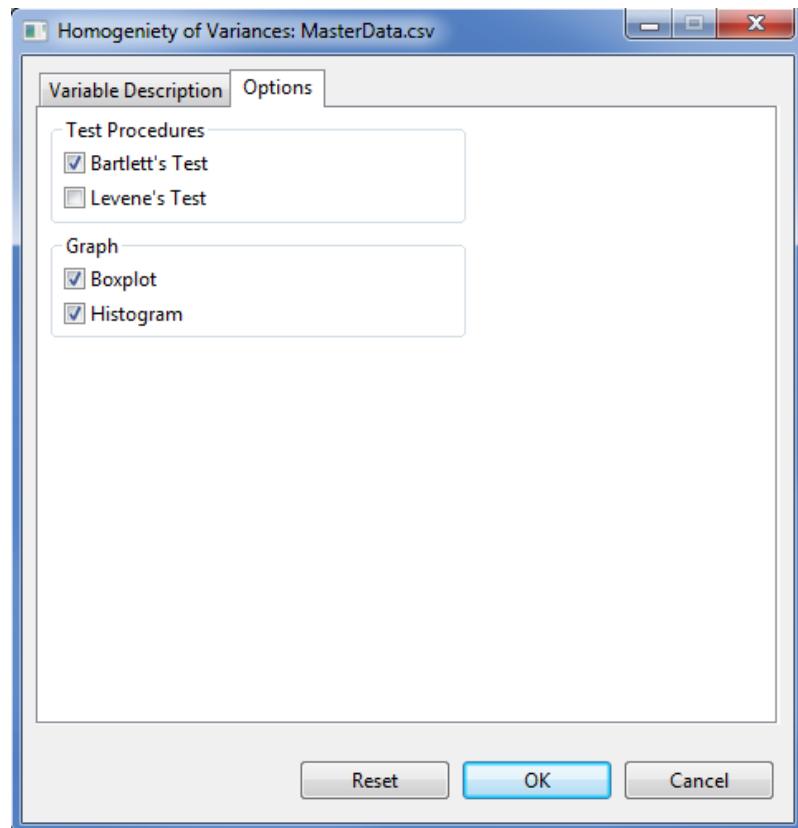
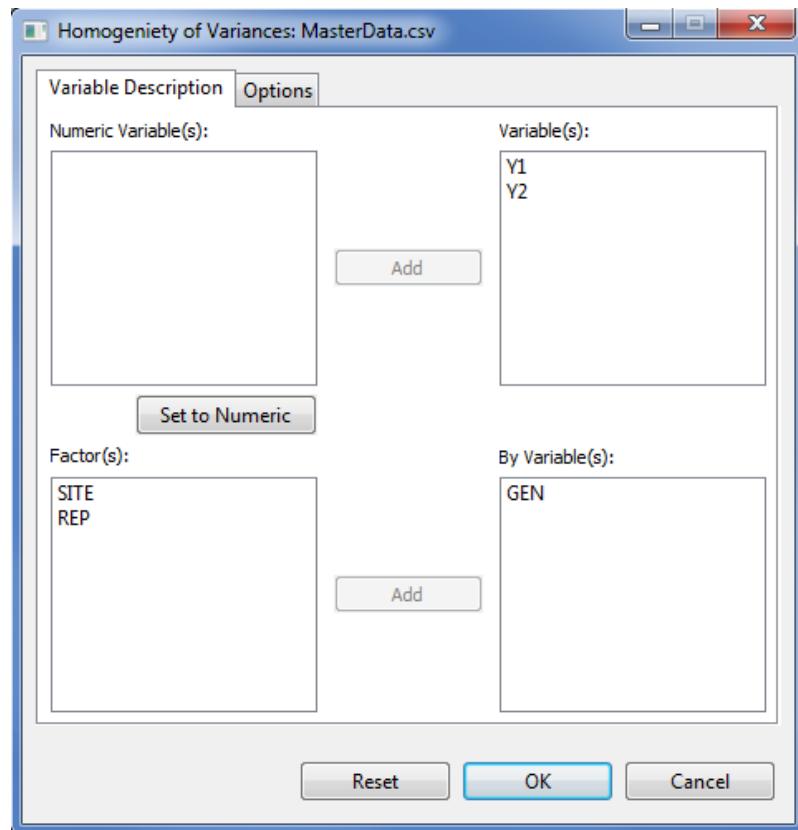
Test Procedure

At least one test procedure should be selected for the analysis to proceed. By default, the Bartlett's test for homogeneity of variances is selected.

Graph

The user has the option to display the box plot and the histogram.

For the example, suppose we want to determine whether the variances are equal among the levels of *GEN* for variables *Y1* and *Y2*. The completed **Variable Description** tab and completed **Option** tab should appear as below:



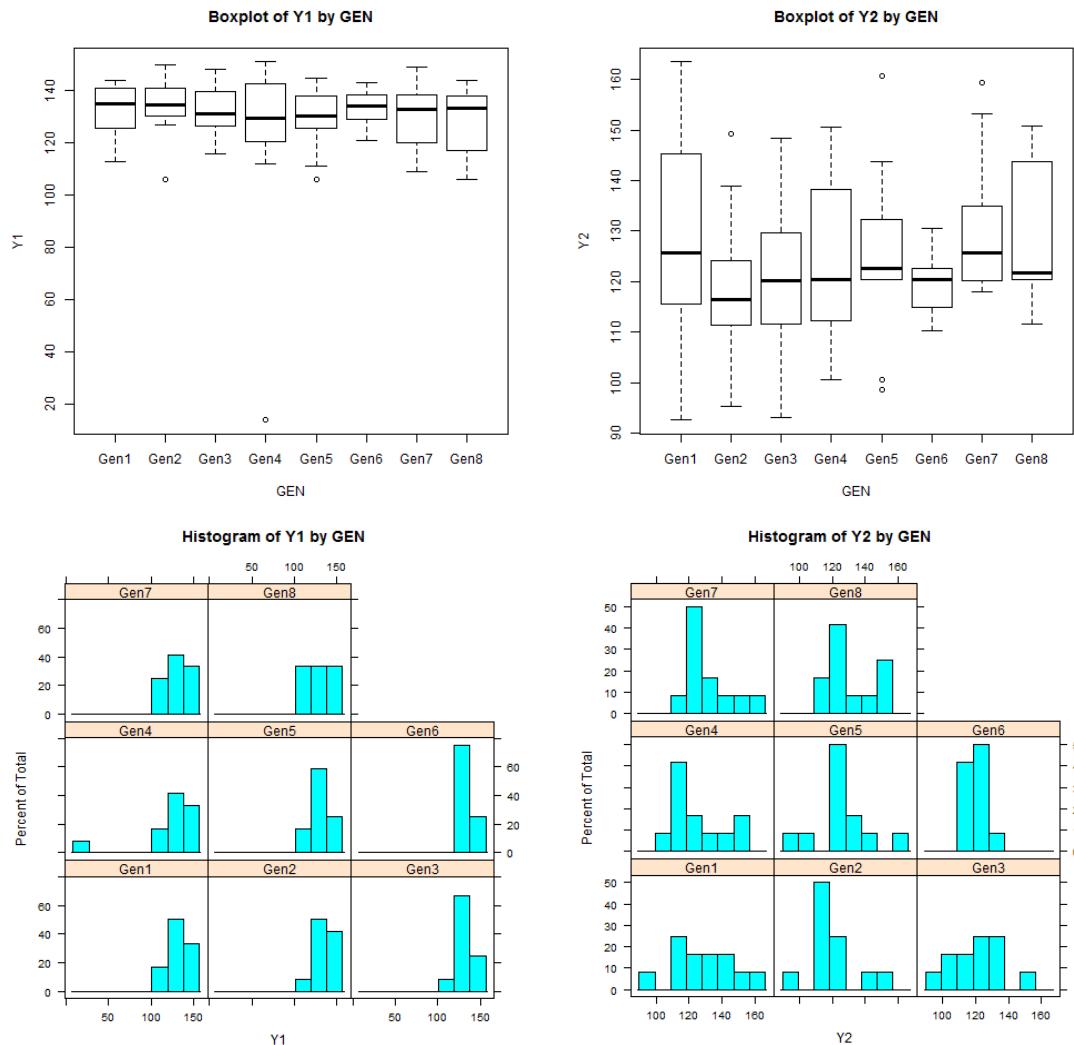
- Click the **OK** button to perform the test. The **Homogeneity of Variances** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample text output as displayed in the **Output** page:

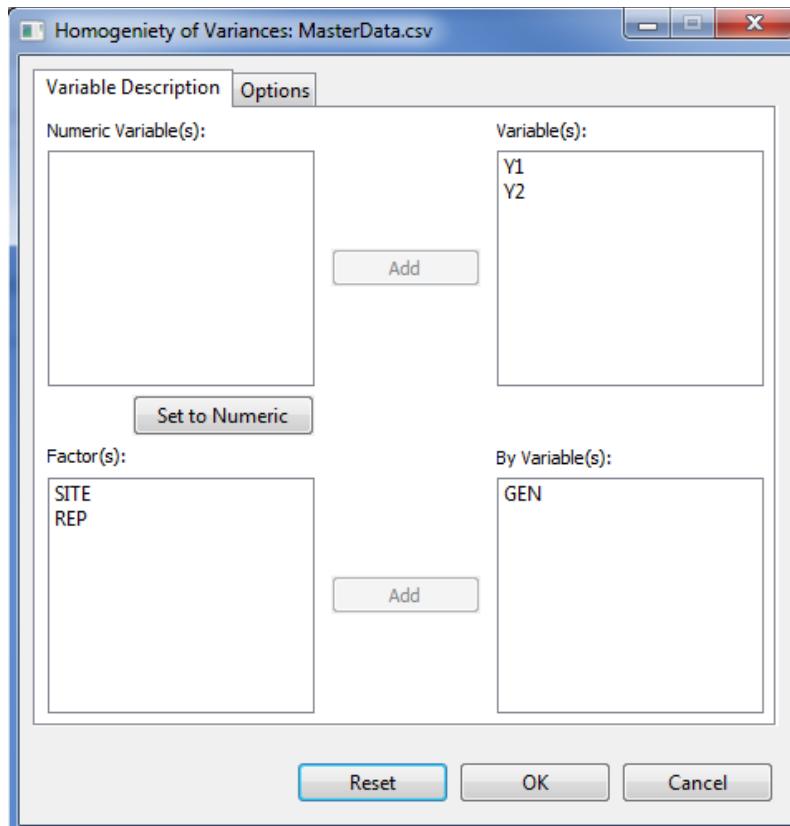
Test for Homogeneity of Variances

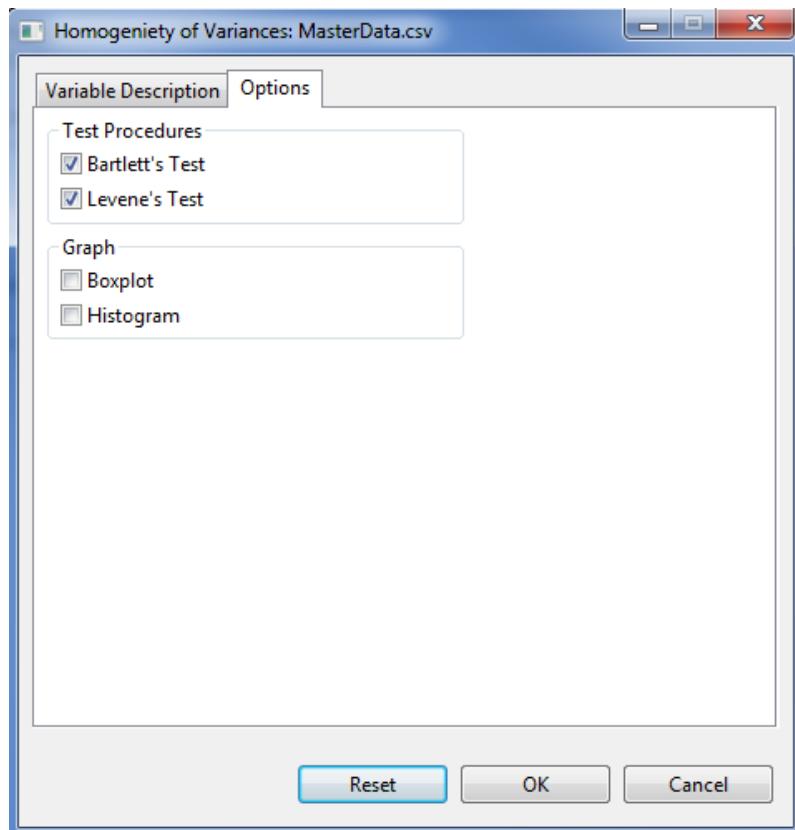
Grp	Variable	Method	DF	Chisq	Value	Pr (>Chisq)
GEN	Y1	Bartlett	7	49.91	0.0000	
GEN	Y2	Bartlett	7	15.00	0.0359	

Sample graphical output displayed in the **Graph** page:



Suppose we want to determine whether the variances are equal among the levels of **GEN** for variables **Y1** and **Y2** using Bartlett's and Levene's test. The completed **Variable Description** tab and completed **Option** tab should appear as below:





Sample output as displayed in the **Output page** tab is displayed below:

Test for Homogeneity of Variances						
Grp	Variable	Method	DF	Statistic	Value	Prob
GEN	Y1	Bartlett	7	Chisq	49.91	Pr(>Chisq)
GEN	Y1	Levene	7	F	1.17	Pr(>F)
GEN	Y2	Bartlett	7	Chisq	15.00	Pr(>Chisq)
GEN	Y2	Levene	7	F	1.25	Pr(>F)

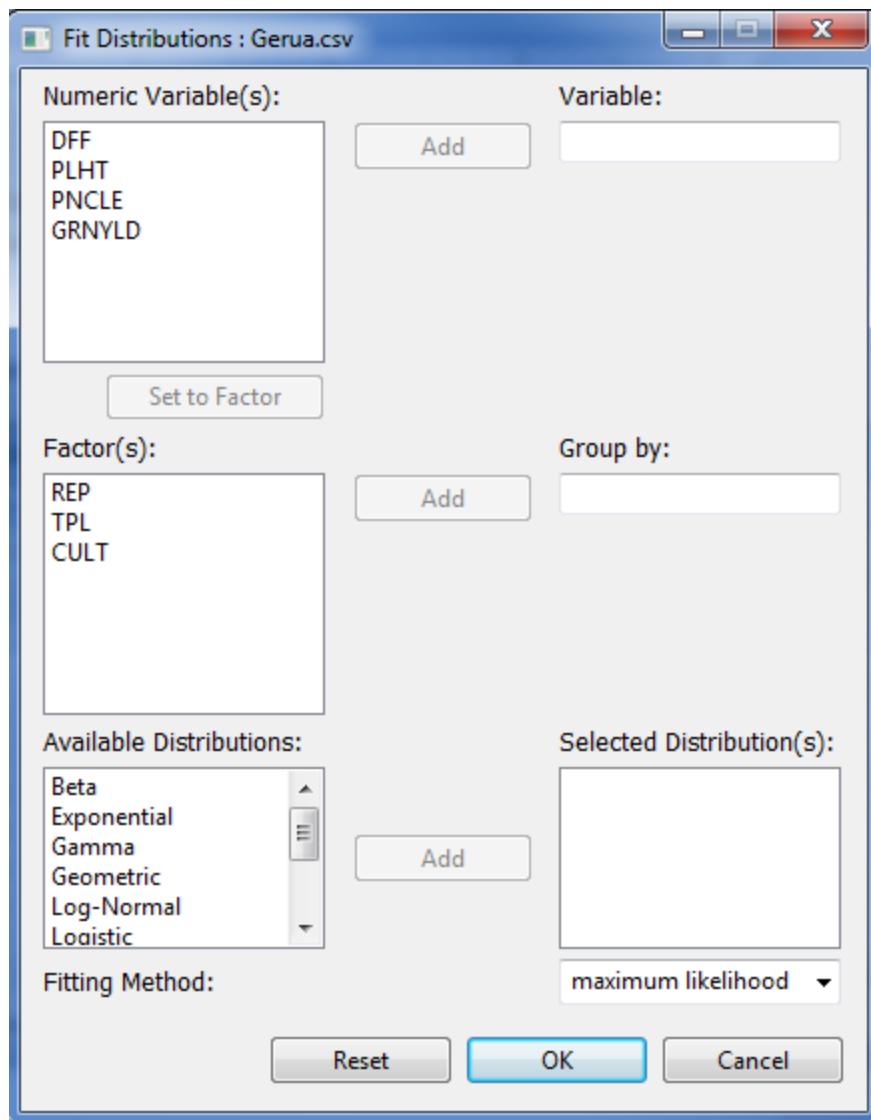
Fitting Distributions

The steps to fit a distribution are listed below:

- On the **Project Explorer** panel, double-click the file to be used for the analysis to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *Gerua* from the package. To import data from the package, see *Importing Data from package* section of this user's manual.

- Click **Analyze | Fit Distributions....** The **Fit Distributions** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Variable

This field is required. Entry should come from the **Numeric Variable(s)** list box.

Group by

This field is optional. If a factor is specified, analysis is going to be done per category of the factor. Entry should come from the **Factor(s)** list box.

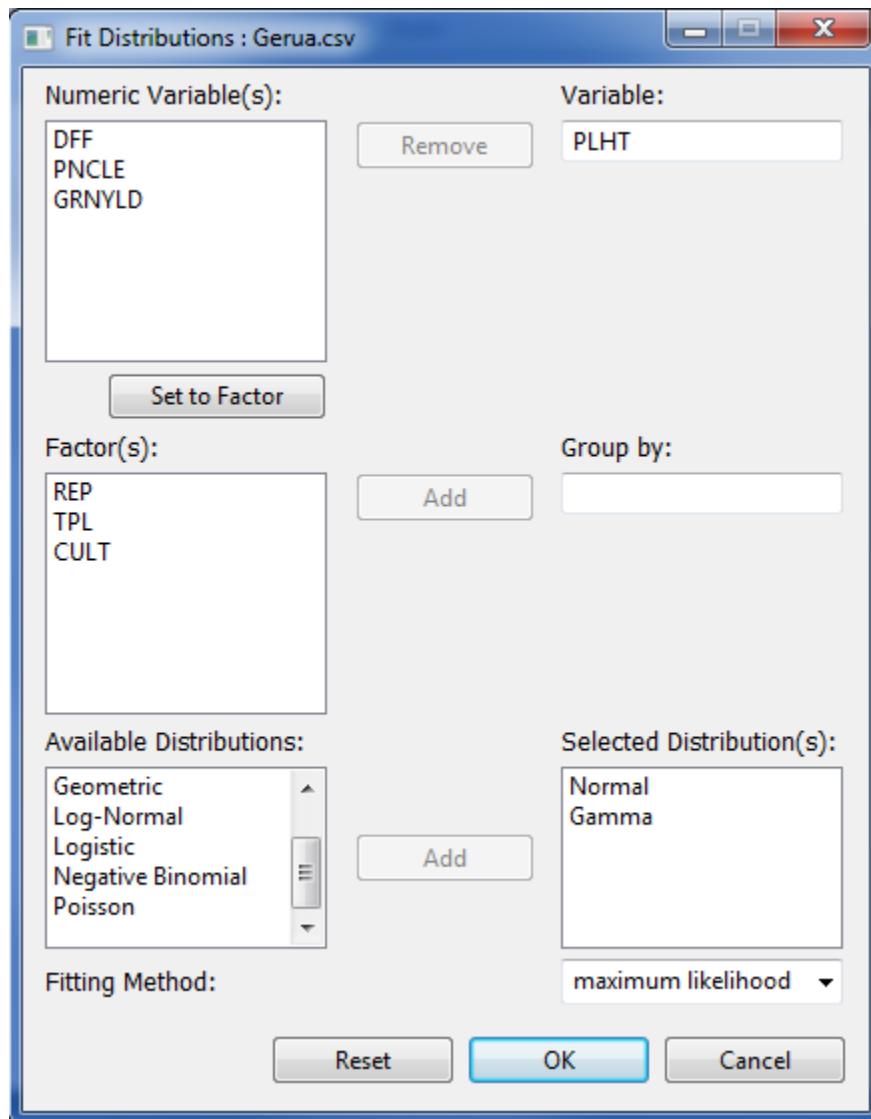
Selected Distribution(s)

This field is required. One or more entries should be selected from the **Available Distributions** list box.

Fitting Method

Available methods for fitting are: maximum likelihood, moment matching, and maximum goodness-of-fit.

For the example, the completed dialog box should appear as below:



- Click the **OK** button to fit the distributions. The **Fit Distributions** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample text output as displayed in the **Output** page.

DISTRIBUTION FITTING
METHOD: maximum likelihood estimation

VARIABLE: PLHT

DISTRIBUTION: Normal

Fitting of the distribution 'norm' by maximum likelihood
Parameters:

	estimate	Std. Error
mean	128.47756	1.2203140
sd	15.24172	0.8628923

Goodness-of-fit

	Statistic	Decision
Cramer-von Mises	0.5887	NULL
Anderson-Darling	2.9911	NULL
Kolmogorov-Smirnov	0.1364	rejected

DISTRIBUTION: Gamma

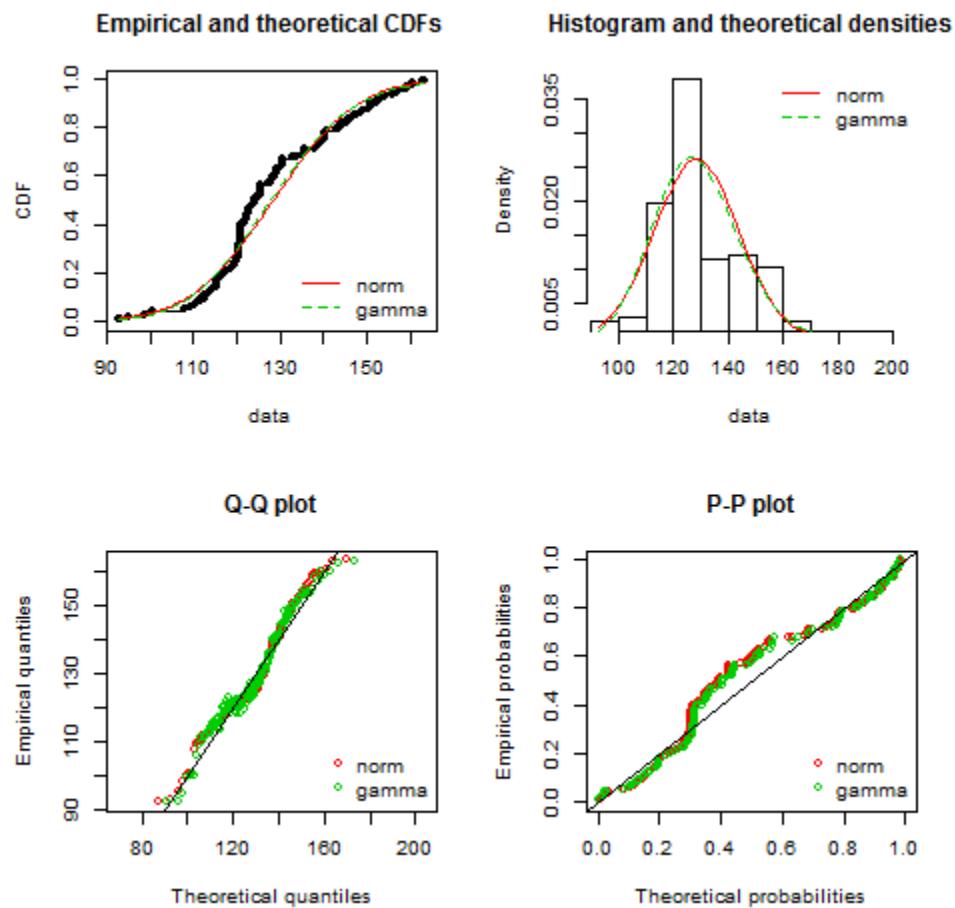
Fitting of the distribution 'gamma' by maximum likelihood
Parameters:

	estimate	Std. Error
shape	71.8758920	8.11707165
rate	0.5594382	0.06339853

Goodness-of-fit

	Statistic	Decision
Cramer-von Mises	0.4672	rejected
Anderson-Darling	2.4234	rejected
Kolmogorov-Smirnov	0.1219	rejected

Sample graphical output as displayed in the **Graph** page.



t-Test

One Sample t-test

The One Sample *t*-test procedure determine whether the mean of one variable differ from the hypothesized value. It assumes that the samples are randomly taken from the population of interest and that the observations are obtained from a normal distribution.

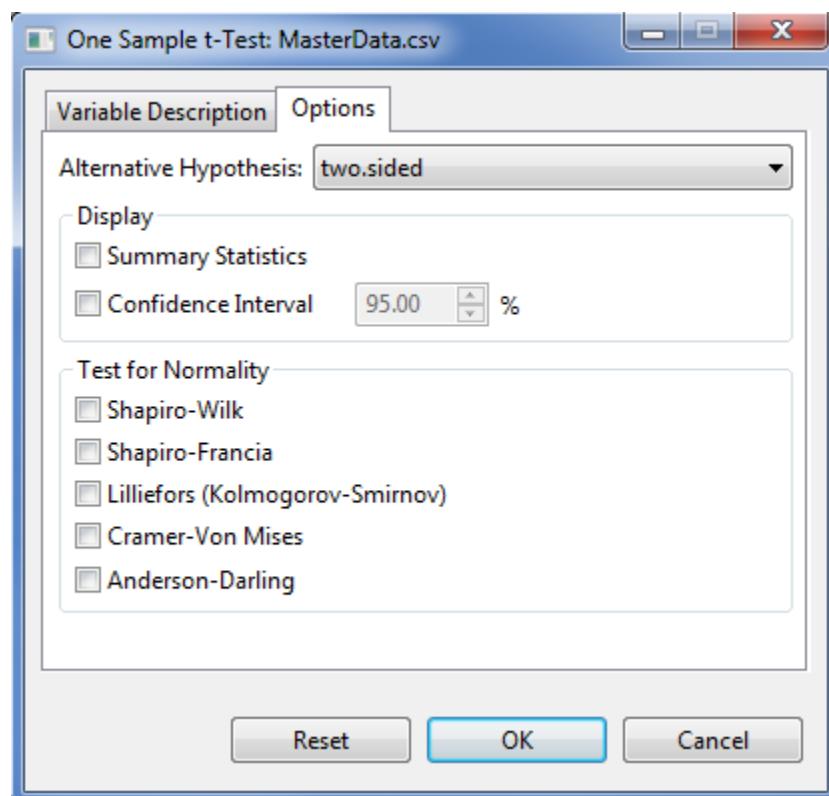
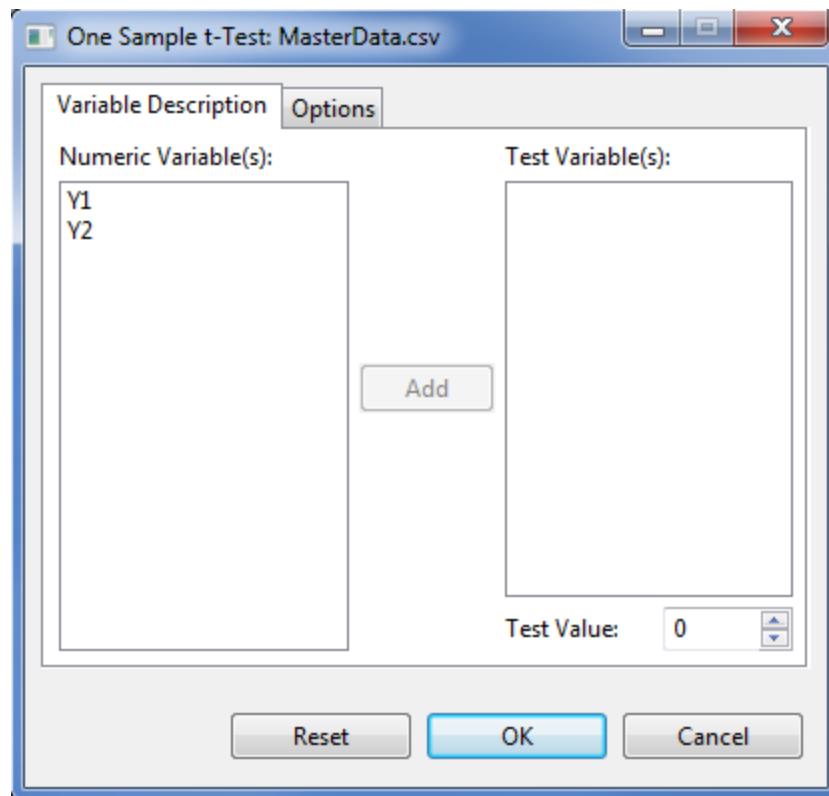
The steps to perform the one sample t-test are listed below:

- On the **Project Explorer** panel, double-click the file to be used for the analysis to open it and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *MasterData.csv* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

Opening the data for the first time, *REP* field is regarded by STAR as numerical variable. To change its type, choose the variable *REP* and click on the **Set to Factor** button.

- Click **Analyze | t-Test | One Sample....** The **One Sample** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Variable Description Tab

Test Variable(s)

At least one entry should be specified for the analysis to proceed. Entries should come from the **Numeric Variable(s)** list box.

Test Value

This field is required and will only accept one entry at a time. The default value is 0. User can change the value by typing a value inside the spin box or by pressing the up-arrow key to increment the value, down-arrow key, otherwise.

Options Tab

Alternative Hypothesis

User can select a one-tailed or two-tailed test. The default alternative hypothesis used is that for the two-sided test.

Summary Statistics

If this option is selected, a summary table with number of observations, minimum, maximum, mean and standard deviation will be displayed.

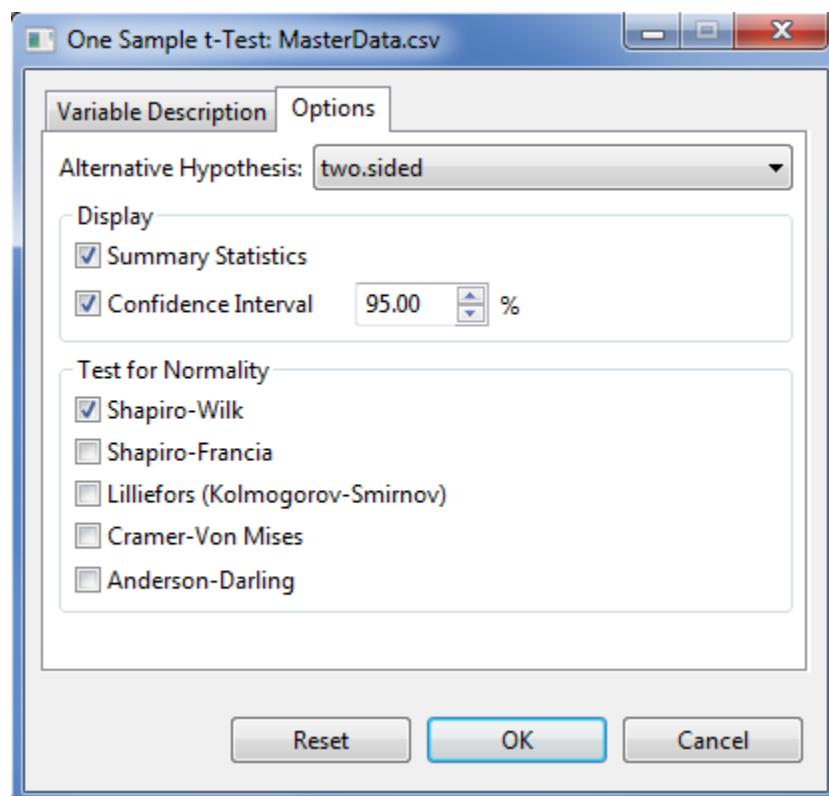
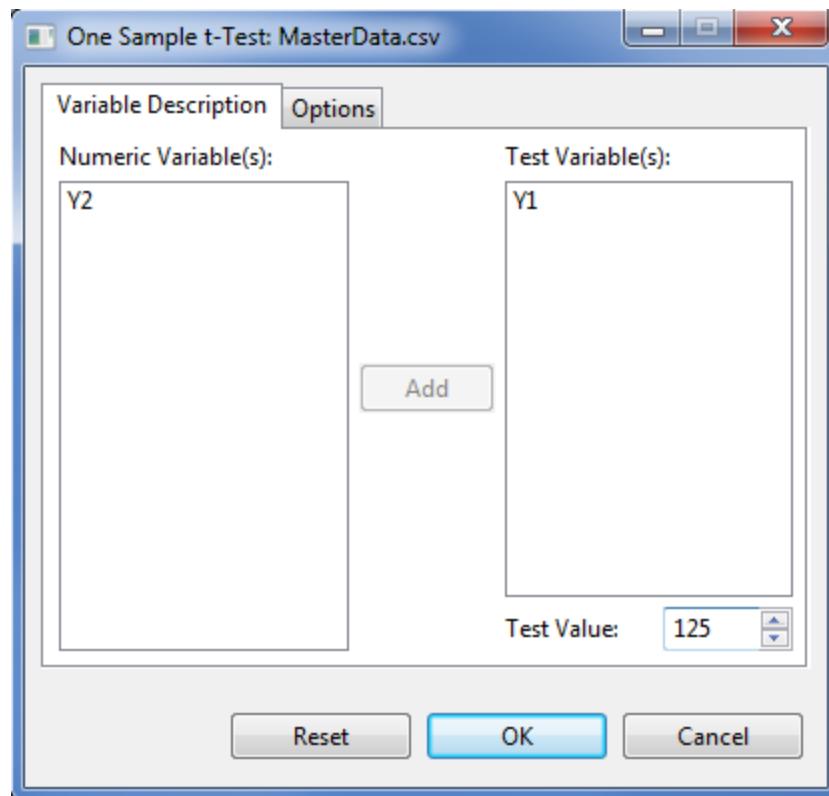
Confidence Interval

If this option is selected, a confidence interval of the mean will be displayed. If this is selected, by default, a 95% confidence interval is displayed.

Test for Normality

If at least one option is selected, if display the test for normality test for each level of the **Grouping Variable**. There are five available tests for normality procedure.

For the example, suppose we want to determine whether the mean of Y1 is significantly different from 125. The completed **Variable Description** tab and **Option** tab should appear as shown below:



- Click the **OK** button to perform the test. The **One Sample t-Test** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample output as displayed in the **Output** page.

```
Test for Normality
-----
Variable Method      W Value   Pr(< W)
-----
Y1      Shapiro-Wilk 0.6908   0.0000
-----

Descriptive Statistics
-----
Variable N Lower CI*    Mean   Upper CI*   StdDev  SE_Mean
-----
Y1      96  127.18    130.43    133.68    16.05   1.64
-----
* At 95% Confidence Level.

One Sample t-Test, h0: mean = 125
-----
Variable DF t Value   Pr(>|t|) 
-----
Y1      95  3.31     0.0013
-----
```

Paired Sample t-tests

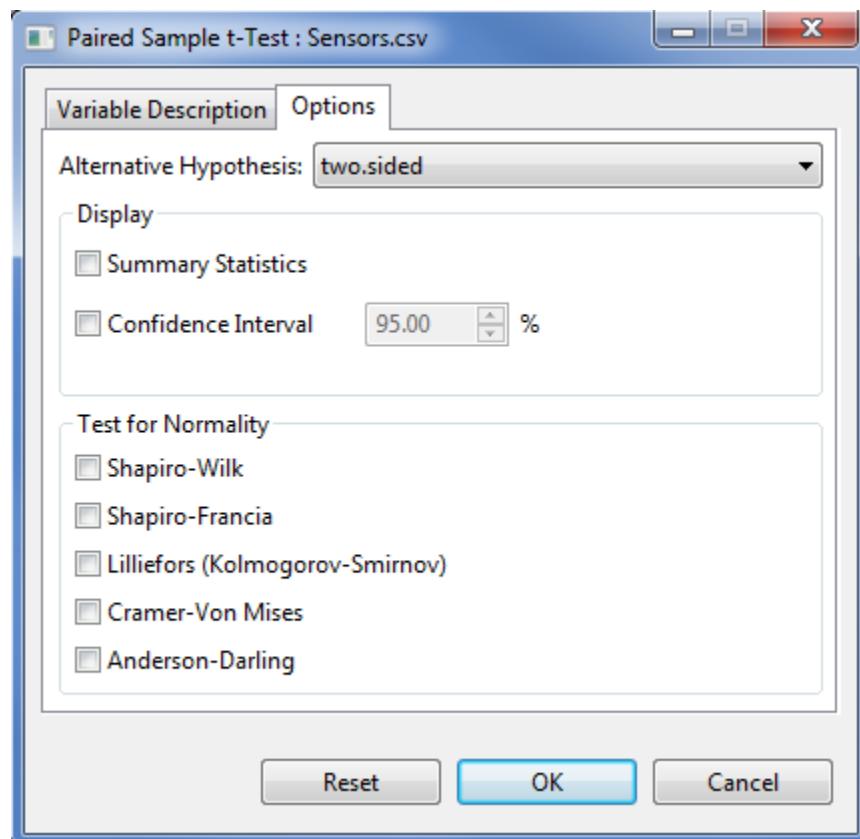
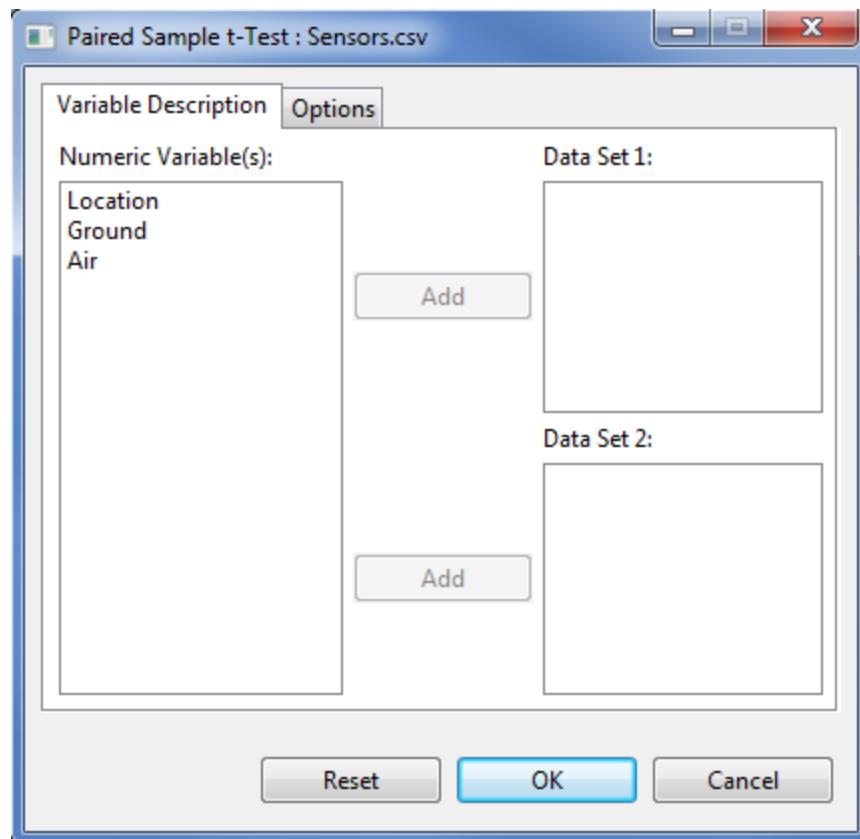
The paired sample *t*-test is used to compares the means of paired or related samples. It assumes that the samples are randomly taken from the population of interest and the population of differences is normally distributed.

The steps to perform paired *t*-test are listed below:

- On the **Project Explorer** panel, double-click the file to be used for the analysis to open it and view it in the Data Viewer tab.

For the example, using the project named *My Project*, import the data *Sensors* from the package. To import data from the package, see *Importing Data from the Package* section of this user's manual.

- Click **Analyze | t-Tests | Paired Samples...** from the main window. The **Paired Samples t-Test** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Variable Description Tab

Data Set 1

This field needs at least one entry for the analysis to proceed and must have the same number of entries as **Data Set 2**.

Data Set 2

This field needs at least one entry for the analysis to proceed and must have the same number of entries as **Data Set 1**.

Options Tab

Alternative Hypothesis

User can select a one-tailed or two-tailed test. The default alternative hypothesis used is that for the two-sided test.

Summary Statistics

If this option is selected, a summary table with number of observations, minimum, maximum, mean and standard deviation will be displayed.

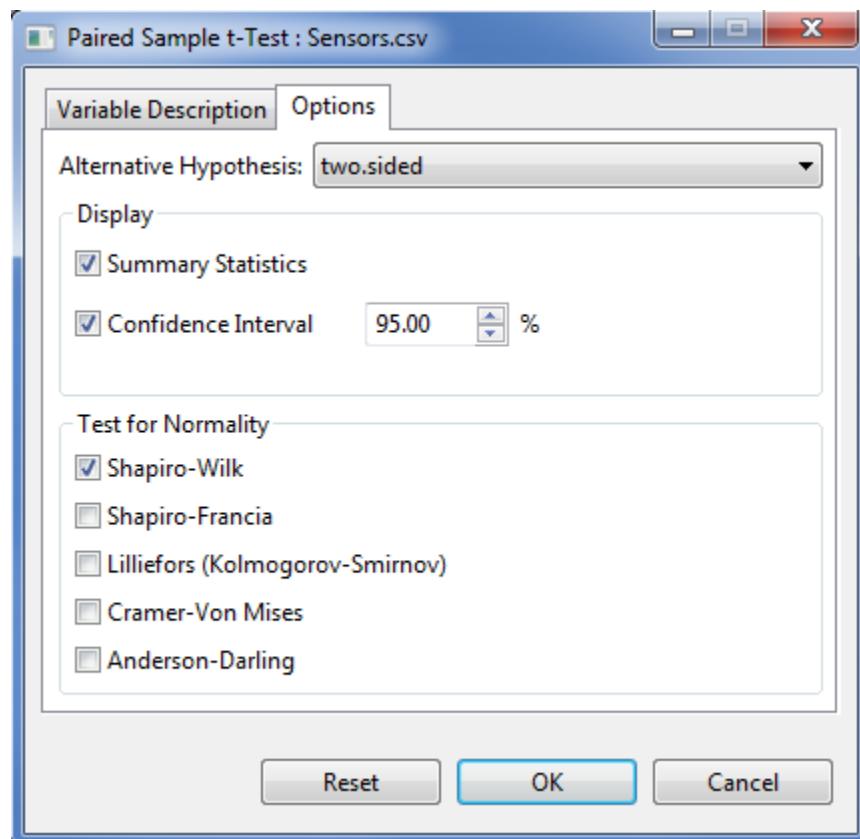
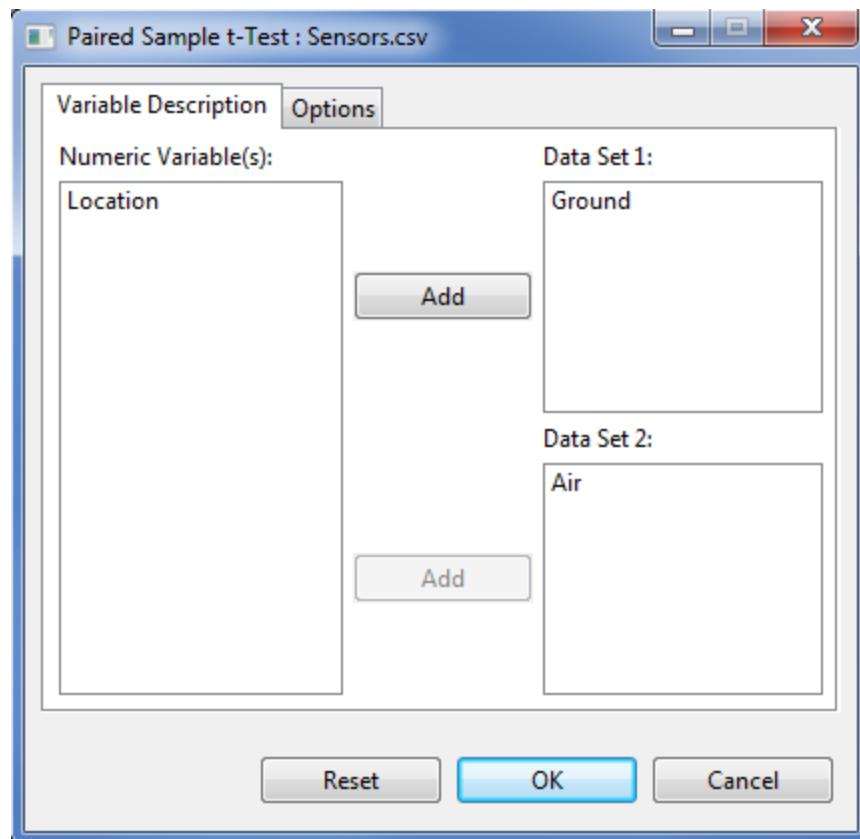
Confidence Interval

If this option is selected, a confidence interval of the mean will be displayed. If this is selected, by default, a 95% confidence interval is displayed.

Test for Normality

If at least one option is selected, if display the test for normality test for the difference between each pair of variables. There are five available tests for normality procedure.

For the example, the completed **Variable Description** tab and **Options** tab should appear as shown below:



- Click the **OK** button to perform the analysis. The **Paired Sample t-Test** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample output of the analysis as shown in the Output page is displayed below:

Test for Normality

Difference	Method	W	Value	Pr(< W)
Ground - Air	Shapiro-Wilk	0.9403	0.5562	

Descriptive Statistics

Difference	N	Lower CI*	Mean	Upper CI*	StdDev	SE_Mean
Ground - Air	10	-2.10	-1.55	-0.9987	0.7706	0.2437

* At 95% Confidence Level.

Paired Sample t-Test, h_0 : mean diff = 0

Difference	DF	t	Value	Pr(> t)
Ground - Air	9	-6.36	0.0001	

Independent Sample t-tests

The independent sample *t*-test is used to determine whether the means between two groups or populations are equal. It assumes that (i) the samples are drawn independently from each population; (ii) the samples from each population are randomly taken; and (iii) the observations from each population are obtained from a normal distribution.

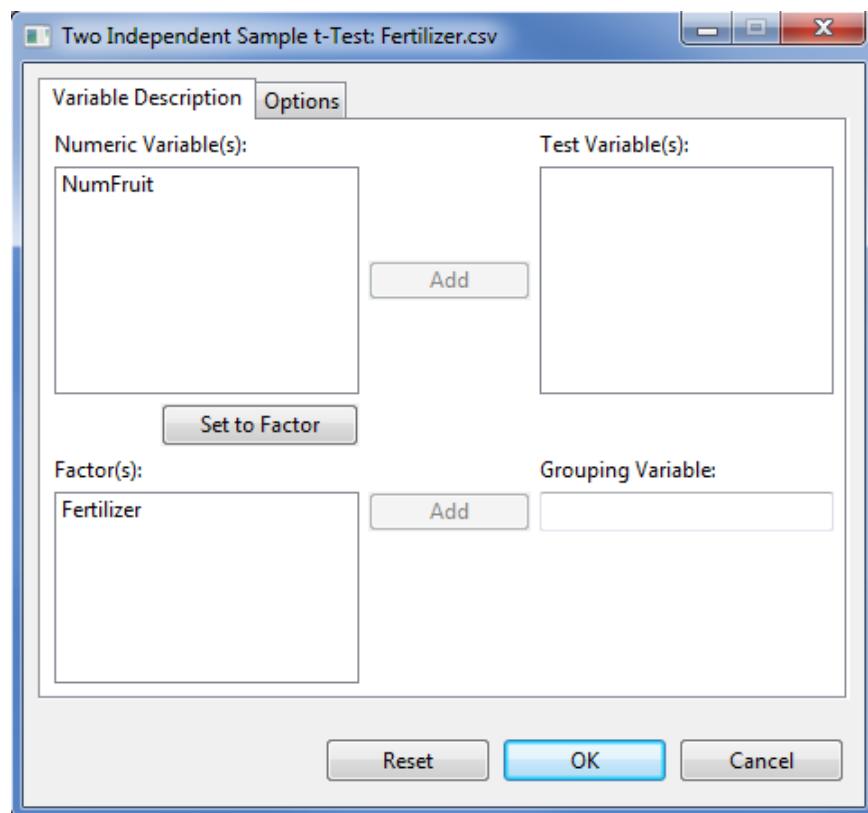
Default output of the *t*-test for two independent samples in STAR includes the test for equality of variances. Depending on the result of the equality of variances, pooled *t*-test or a *t*-test with satterwaite adjustment is computed.

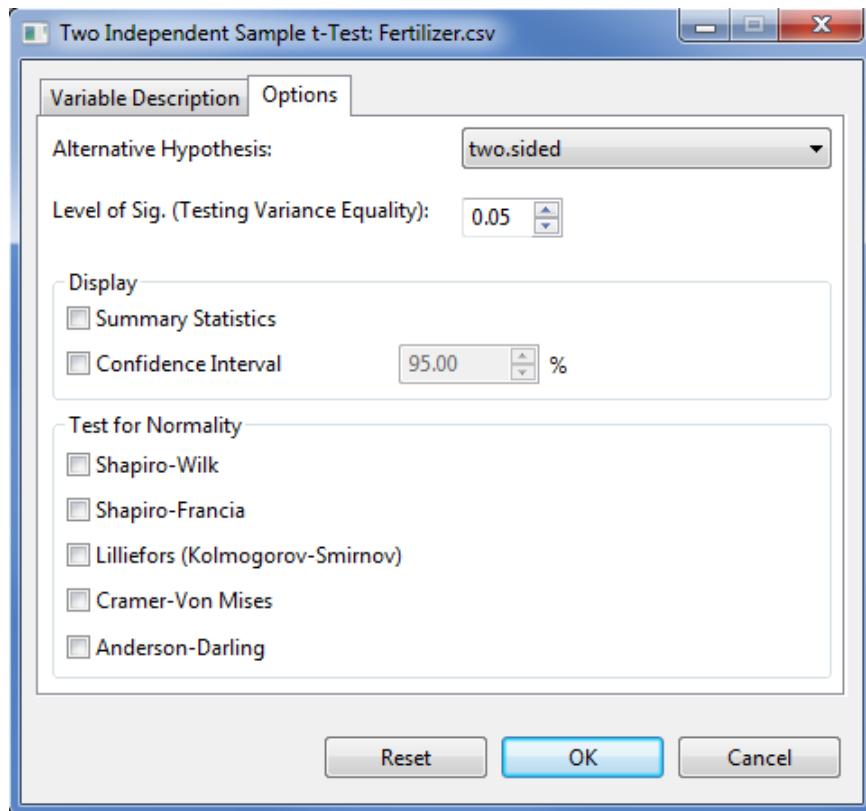
The steps to perform *t*-Test for two independent samples are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *My Project*, import the data *Fertilizer* from package. To import that data from package, see *Importing from Package* section of this user's manual.

- Click **Analyze | t-test | Independent Sample...** from the main window. The **Two Independent Sample t-Test** dialog box will appear.





- Specify the required field and appropriate options for the analysis.

Variable Description Tab

Test Variable(s)

This field need at least one entry for the analysis to proceed and should come from the **Numeric Variable(s)** list box.

Grouping Variable

This field accepts one entry at a time and needs an entry for the analysis to proceed. The entry must come from the **Factor(s)** list box. The entry should have only two levels.

Options Tab

Alternative Hypothesis

User can select a one-tailed or two-tailed test. The default alternative hypothesis used is that for the two-sided test.

Level of Significance

This will be used to determine if the two groups have equal variance. Its default value is 0.05. The user can change this value by specifying a numeric value from zero to 1.

Summary Statistics

If this option is selected, a summary table with number of observations, minimum, maximum, mean and standard deviation will be displayed.

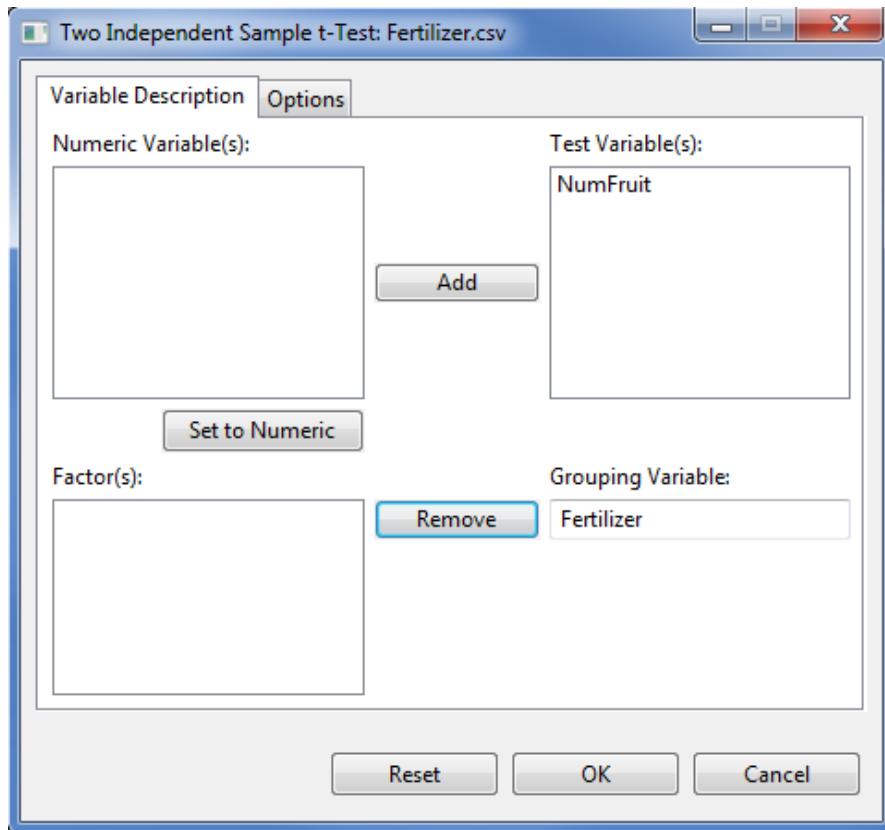
Confidence Interval

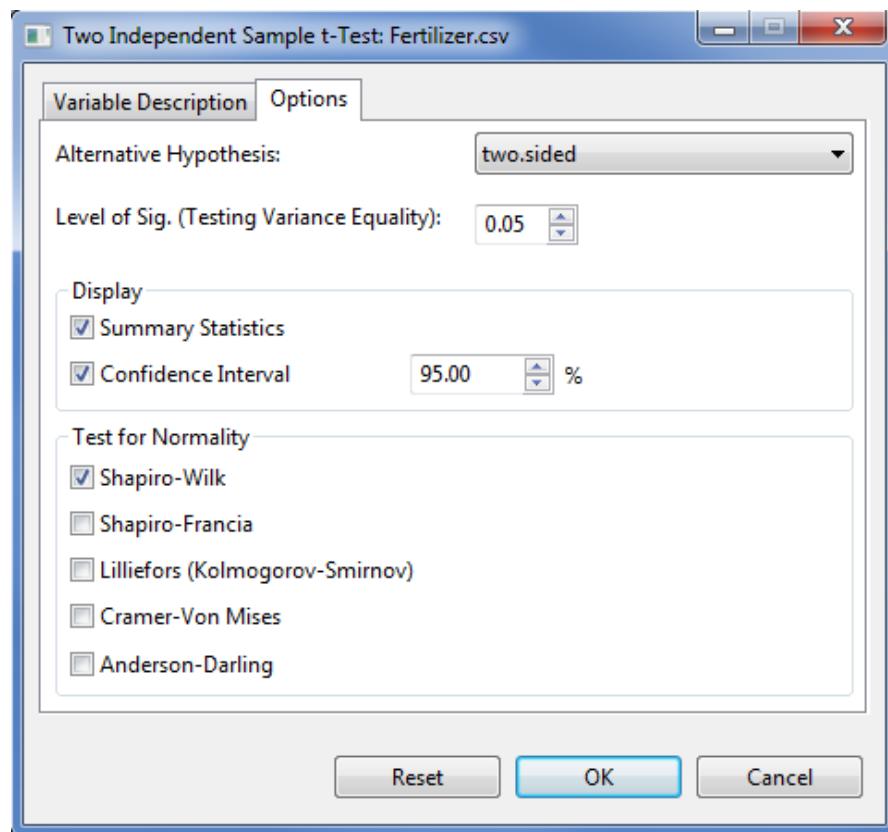
If this option is selected, a confidence interval of the mean will be displayed. If this is selected, by default, a 95% confidence interval is displayed.

Test for Normality

If at least one option is selected, it displays the test for normality test for each level of the **Grouping Variable**. There are five available tests for normality procedure.

For the example, suppose we want to test the hypothesis that the mean number of fruits from trees sprayed by the two fertilizers are the same. The completed **Variable Description** tab and **Option** tab should appear as shown below:





- Click the **OK** button to perform the analysis. The **Two Independent Sample t-test** dialog box will be minimized and STAR actives the **Output** page of the **Result Viewer** tab.

Sample output of the analysis as shown in the Output page is displayed below.

Test for Normality						
Grp	Level	Variable	Method	W	Value	Pr(< W)
Fertilizer	A	NumFruit	Shapiro-Wilk	0.9726	0.8950	
Fertilizer	B	NumFruit	Shapiro-Wilk	0.9198	0.1916	

Descriptive Statistics						
Variable	Fertilizer	N	Lower CI*	Mean	Upper CI*	StdDev
NumFruit	A	15	21.51	23.13	24.75	2.92
NumFruit	B	15	19.25	20.87	22.49	3.54
NumFruit	Diff(A-B)		-0.16	2.27	4.70	3.25
SE_Mean						
						1.19

* At 95% Confidence Level.

Homogeneity of Variances

Variable	Method	Num DF	Den DF	F Value	Pr(> F)
NumFruit	Folded F	14	14	1.47	0.4820

Two Independent Sample t-Test, $h_0:$ mean diff = 0

Variable	Method*	Variances	DF	t Value	Pr(> t)
NumFruit	Pooled	Equal	28	1.91	0.0663

* At 0.05 level of significance.

Chi-Square Test

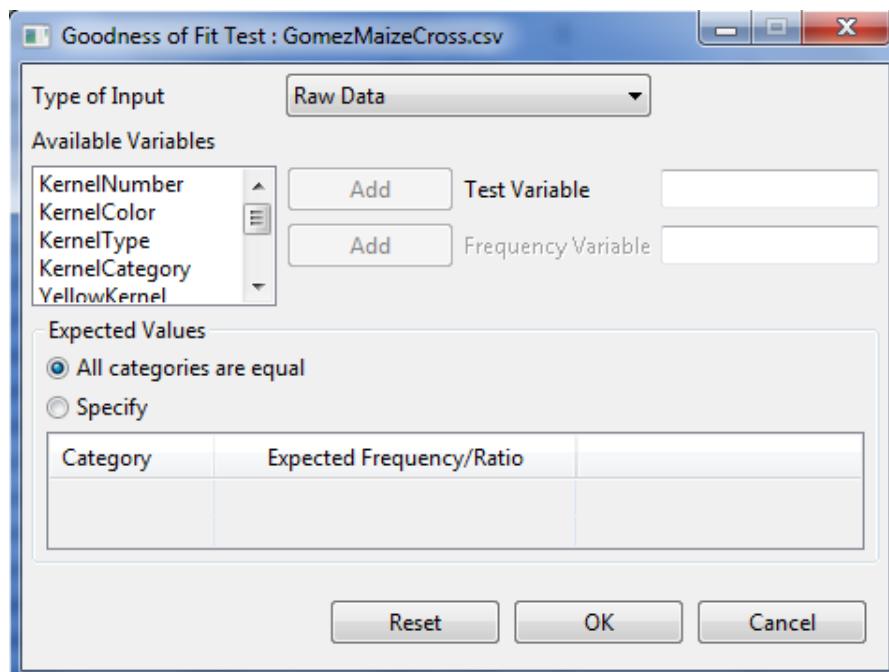
Goodness of fit Test

The steps to perform chi-square goodness of fit test are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *GomezMaizeCross* from the package. To import the data from the package, see *Importing from Package* section of this user's manual.

- Click **Analyze | Chi-Square | Goodness of Fit...** from the main window. The **Goodness of Fit Test** dialog box will appear as shown below:



- Specify the required field and appropriate options for the analysis.

Type of Input

There are two options available, *Raw Data* (default option) and *Summary Data*.

Test Variable

This variable is required.

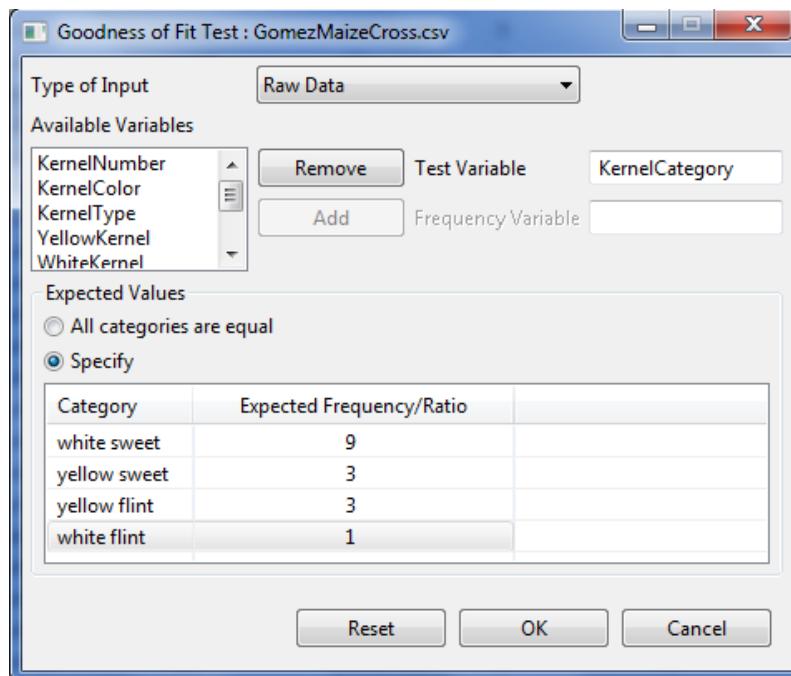
Frequency Variable

If the *Summary Data* is chosen as type of input, this variable is required.

Expected Values

There are two options available, *All categories are equal* (default option) and *specify*. If *specify* option is selected, table will be enable wherein user need to indicate the expected frequency/ratio for each of the unique categories of the test variable. The default and minimum expected frequency/ratio for each category is 1.

For the example, the completed dialog box should appear as shown below:



- Click **OK** button to perform the analysis. The **Goodness of Fit Test** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample text output is shown below.

Frequency Table:

KernelCategory	Obs	Freq	Exp Freq
white flint	112	50	
white sweet	34	150	
yellow flint	496	450	
yellow sweet	158	150	

Chi-Square Goodness of Fit Test

Chi-Square	171.7156
DF	3
Pr > Chi-Square	0.0000

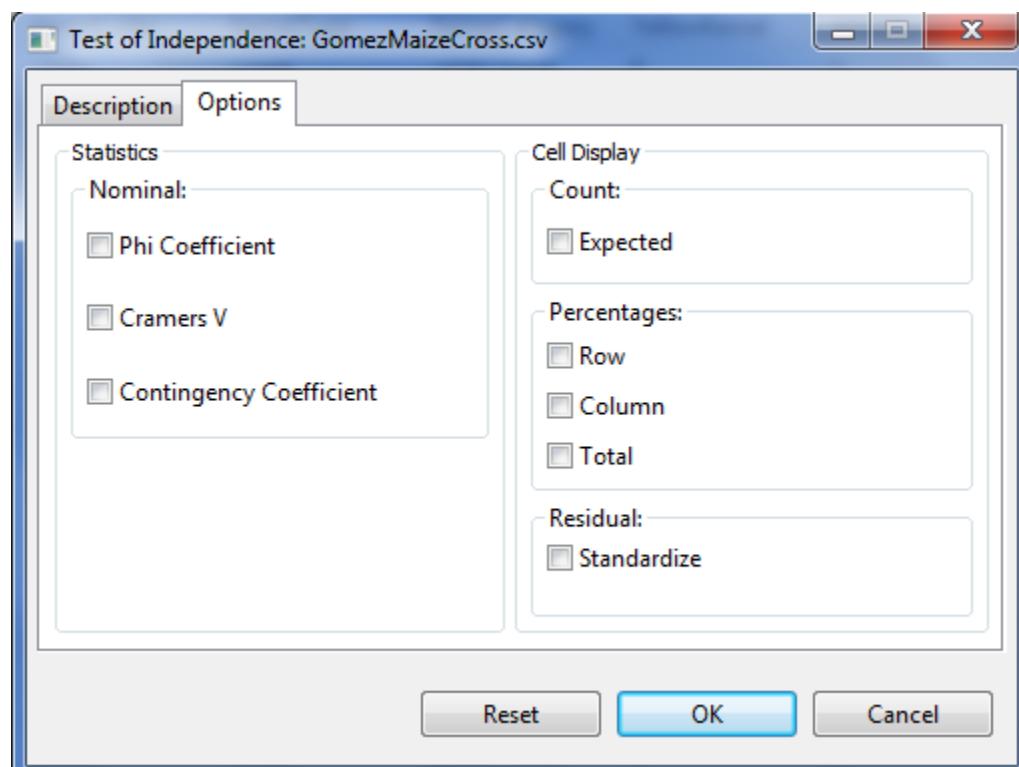
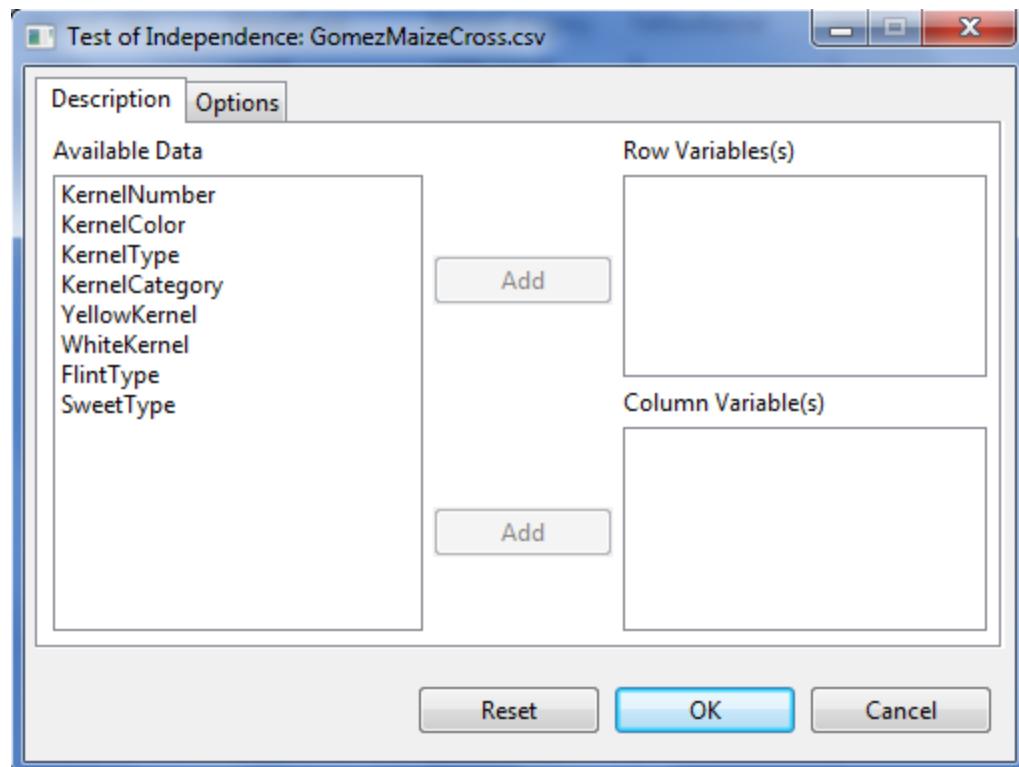
Test of Independence

The steps to perform Chi-Square test of independence are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *GomezMaizeCross* from the package. To import the data the package, see *Importing from Package* section of this user's manual.

- Click **Analyze | Chi-Square | Test of Independence...** from the main window. The **Test of Independence** dialog box will appear as shown below:



- Specify the required field and appropriate options for the analysis.

Description Tab

Type of Input

There are two options available, *Raw Data* (default option) and *Summary Data*.

Row Variable(s)

If *Raw Data* is selected as **Type of Input**, at least one entry is need for the analysis to proceed. If *Summary Data* is selected, only one entry should be entered in the list box.

Column Variable(s)

If *Raw Data* is selected as **Type of Input**, at least one entry is need for the analysis to proceed. If *Summary Data* is selected, only one entry should be entered in the list box.

ID Variable(s)

This field is required and visible if *Summary Data* is chosen as **Type of Input**. Only one entry should be entered in this entry box Entries for this entry box should be a numeric variable.

Options Tab

Phi Coefficient

If this option is selected, the phi coefficient will be displayed.

Cramer's V

If this option is selected, the cramer's v coefficient will be displayed.

Contingency Coefficient

If this option is selected, the contingency coefficient will be displayed.

Expected Count

If this option is selected, the expected frequency will be displayed.

Row Percentages

If this option is selected, the row percentage will be displayed.

Column Percentages

If this option is selected, the column percentage will be displayed.

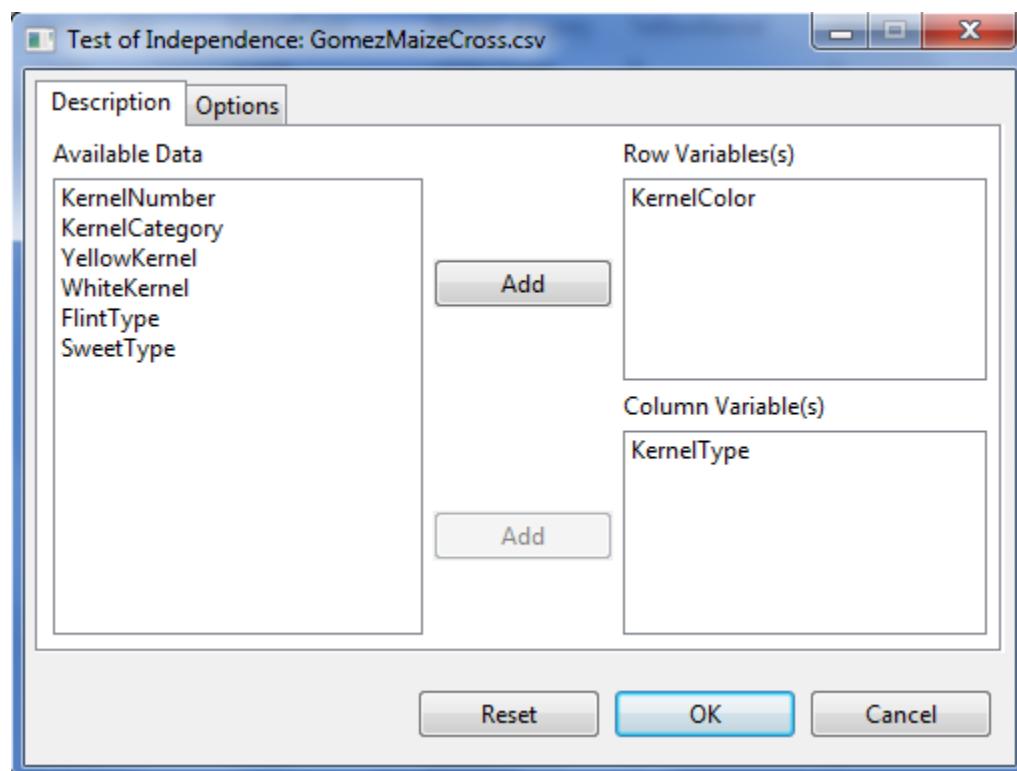
Total Percentages

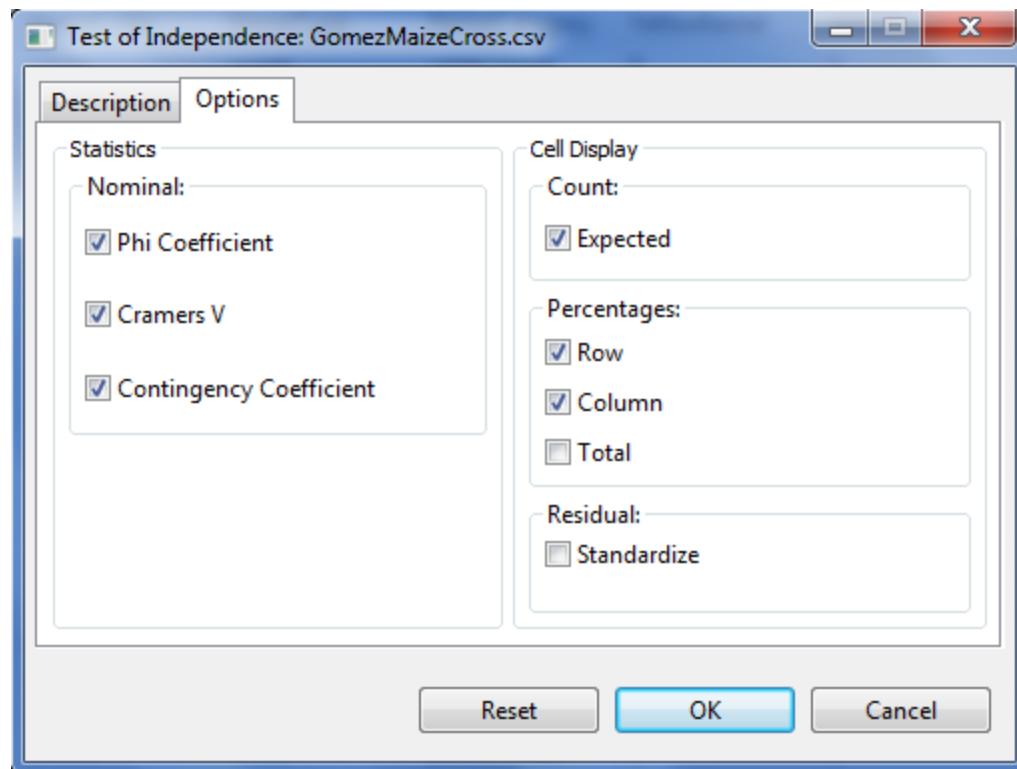
If this option is selected, the total percentage will be displayed.

Standardize Residual

If this option is selected, the standardize residual will be displayed.

For the example, the completed dialog box is shown below:





- Click the **OK** button to perform the analysis. The **Test of Independence** dialog box will be minimized and STAR actives the **Output** page of the **Result Viewer** tab.

Sample text output as displayed in the **Output** page is shown below:

		KernelType		
		flint	sweet	Total
white	Obs Freq	112	34	146
	Expected Freq	110.9600	35.0400	146.0000
	% within KernelColor	76.7123	23.2877	100.0000
	% within KernelType	18.4211	17.7083	
	% of Total	14.0000	4.2500	18.2500
yellow	Obs Freq	496	158	654
	Expected Freq	497.0400	156.9600	654.0000
	% within KernelColor	75.8410	24.1590	100.0000
	% within KernelType	81.5789	82.2917	
	% of Total	62.0000	19.7500	81.7500
Total	Obs Freq	608	192	800
	Expected Freq	608.0000	192.0000	800.0000
	% within KernelType	100.0000	100.0000	
	% of Total	76.0000	24.0000	100.0000

Statistics for Table KernelColor by KernelType

Statistics	DF	Value	Prob
<hr/>			
Pearson Chi-Square	1	0.0134	0.9079
Likelihood Ratio Chi-Square	1	0.0499	0.8232
Phi Coefficient		0.0482	
Contingency Coefficient		0.0482	
Cramer's V		0.0482	
Fisher's Exact Test			0.9147

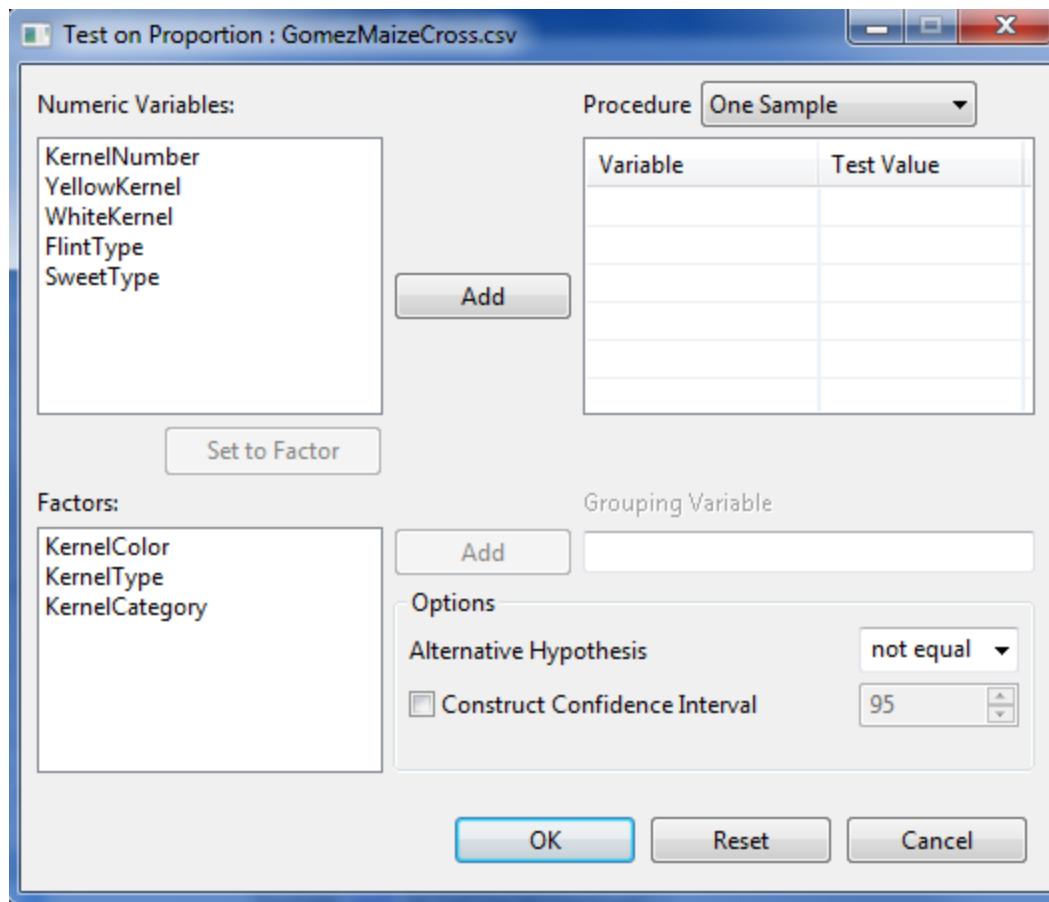
Test on Proportion

The steps to perform test on proportion are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *GomezMaizeCross* from the package. To import the data, see *Importing from Package* section of this user's manual.

- Click **Analyze | Test on Proportion ...** from the main window. The **Test on Proportion** dialog box will appear as shown below:



- Specify the required field and appropriate options for the analysis.

Procedure

There are three available procedures available: *One Sample* (default value), *Two Related Samples* and *Independent Sample*.

Variable

At least one entry is need for the analysis to proceed. Entries for this list box should come from the **Numeric Variable** list box.

Test Value

For every variable entered in the table, a default test value is specified. User can change the value by typing a value inside the spin box or by pressing the up-arrow key to increment the value, down-arrow key, otherwise. Valid value is between 0 to 1.

Grouping Variable

This field is required and visible if the procedure is *Independent Samples* is chosen. Entries for this list box should come from the **Factors** list box.

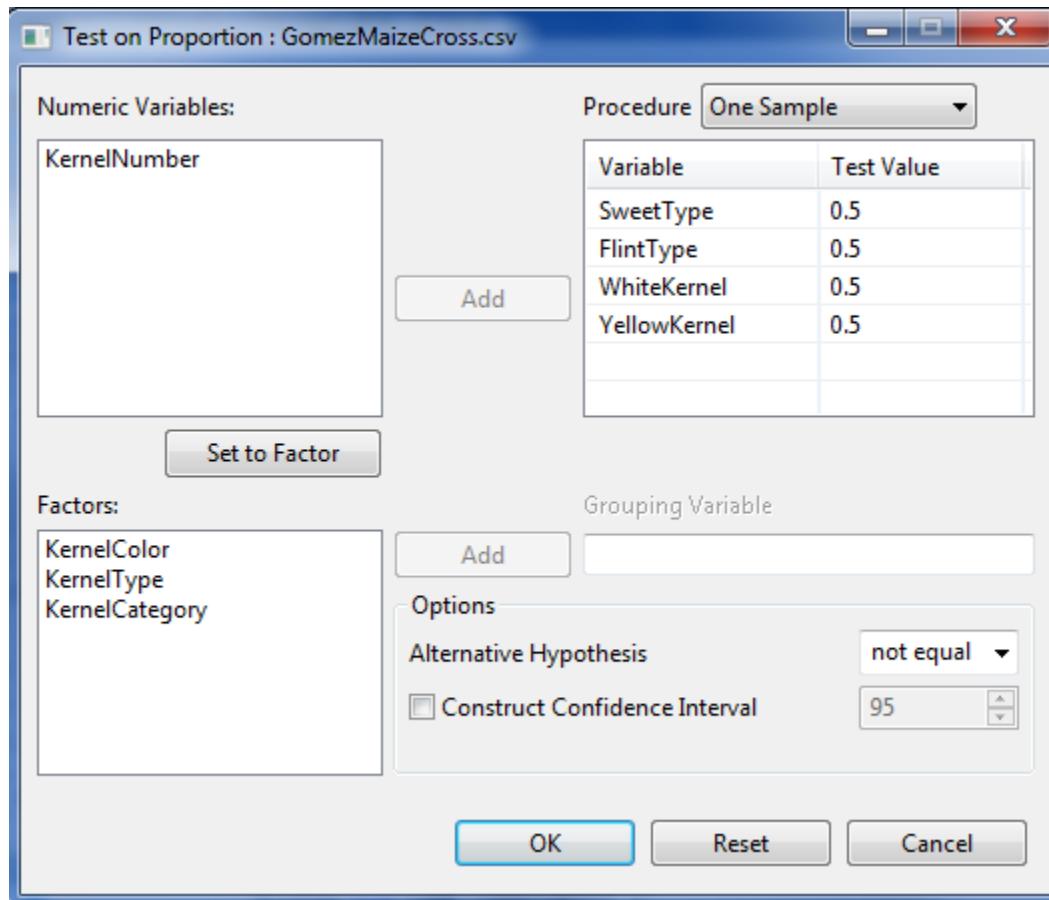
Alternative hypothesis

There are three available options: *less*, *greater* and *not equal* (default value).

Construct Confidence Interval

This option is available only if the procedure chosen is either *One Sample* or *Independent Sample*. If this option is selected, a confidence interval will be constructed.

For the example, the completed dialog box is shown below:



- Click the **OK** button to perform the analysis. The **Test on Proportion** dialog box will be minimized and STAR actives the **Output** page of the **Result Viewer** tab.

Sample output of the analysis as displayed in the Output page is shown below.

Summary

Variable	Num Success	Num Trial	Proportion
SweetType	192	800	0.2400
FlintType	608	800	0.7600
WhiteKernel	146	800	0.1825
YellowKernel	654	800	0.8175

Test on One Proportion

Variable	Procedure	TestValue	Statistics	Prob
SweetType	Chi-Square	0.5000	216.32	0.0000
SweetType	Exact Binomial	0.5000	192.00	0.0000
SweetType	Approx Z	0.5000	-14.71	0.0000
FlintType	Chi-Square	0.5000	216.32	0.0000
FlintType	Exact Binomial	0.5000	608.00	0.0000
FlintType	Approx Z	0.5000	14.71	0.0000
WhiteKernel	Chi-Square	0.5000	322.58	0.0000
WhiteKernel	Exact Binomial	0.5000	146.00	0.0000
WhiteKernel	Approx Z	0.5000	-17.96	0.0000
YellowKernel	Chi-Square	0.5000	322.58	0.0000
YellowKernel	Exact Binomial	0.5000	654.00	0.0000
YellowKernel	Approx Z	0.5000	17.96	0.0000

* alternative hypothesis = 'two.sided'

Analysis of Variance

The Analysis of Variance is a statistical technique which partitions the total variation into explained variation and unexplained variation.

The **Analysis of Variance submenu** can be used to perform analysis of variance for data from experimental designs with single error (e.g., RCB) and with 2 or more errors (e.g., Split Plot).

The Analysis of Variance in STAR uses fixed model and can be used for balanced data set. For data set with less than 10% missing observations, STAR can still perform the analysis of variance by first estimating the missing observations.

Treatment means or means of the treatment combination will be displayed depending on the result of the analysis of variance. If the analysis is performed on single factor and the treatment effect is not significant, the table of means is displayed. If the analysis is performed on multi-factor experiment and the highest interaction is not significant, the mean of the treatment combination are displayed.

Depending on the results of the analysis of variance, pairwise comparison means will be automatically generated for all effects that are significant using one applicable pairwise comparison procedure. If the factor being compared has at most 5 levels, the default pairwise comparison procedure used in STAR is the Least Significant Difference (LSD) test; otherwise the Honestly Significant Difference (HSD) or Tukey's test is used. Aside from LSD and HSD, Duncan's Multiple Range test, Student Newmann-Keul's test and Scheffe's test maybe also be requested.

Contrast analysis maybe requested in this menu. There are three available contrast analyses that maybe request for each factor, namely: compare with control, orthogonal polynomial contrast and user specified contrast.

Boxplot of the response variable and two diagnostic plots are automatically displayed in the *Graph* page of the **Result Viewer** tab. The *Residuals vs. Fitted* is for assessing the distribution of the residuals. An ideal plot should reveal that the points are distributed

around 0. The residuals should not be affected by the size of the fitted values. Problematic data may result to a plot wherein the points increase as the fitted values increase, thus forming a “funnel-shaped” distribution. This will indicate a violation of the homogeneity of variances requirement for analysis of variance. The *Normal Q-Q* plot is used for assessing normality. The points should lie about the line.

A csv data named *residualData* is automatically created in the output folder that contains the original dataset, predicted and residual value of all response variable used in the analysis.

Completely Randomized Design

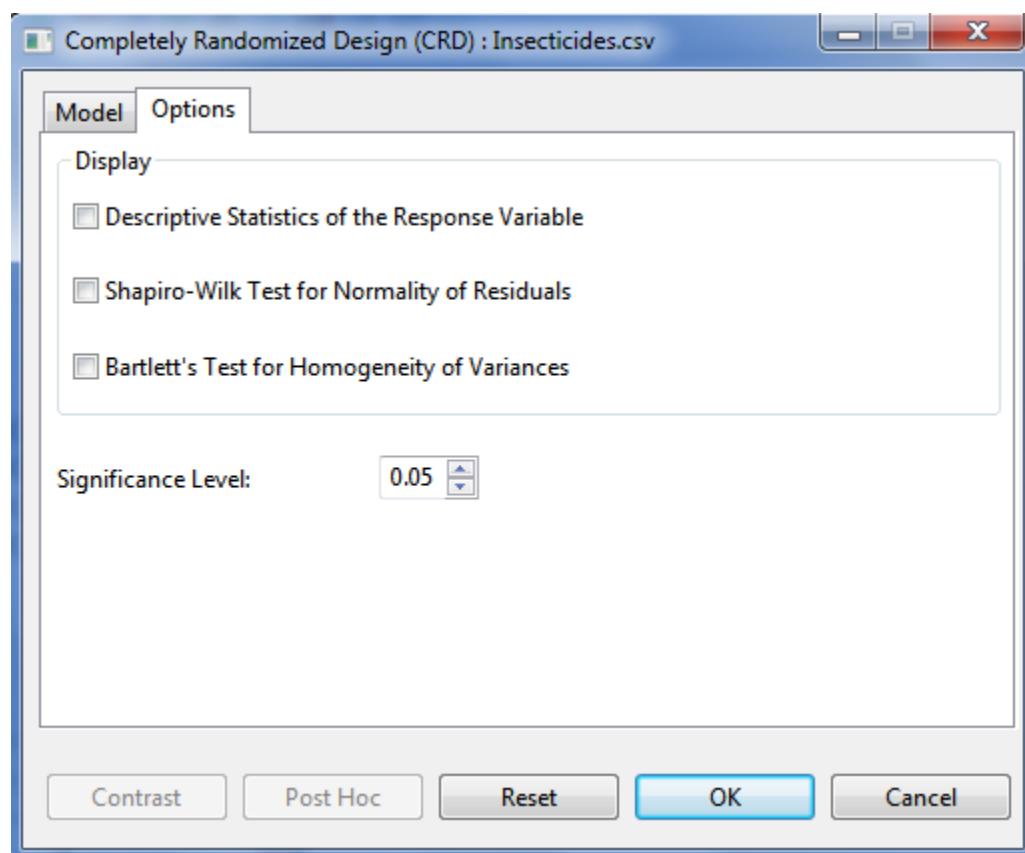
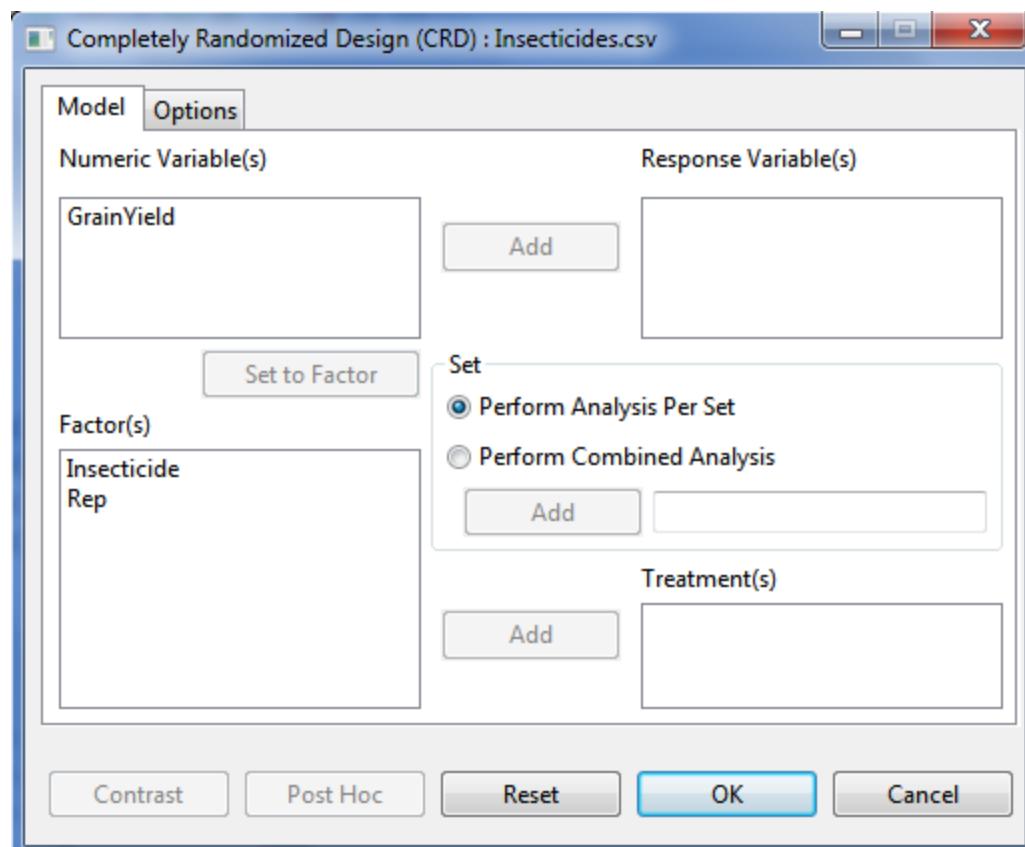
This submenu can be used to perform analysis of variance for single or multi factor experiment.

The steps to perform the analysis are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *Insecticides* from package. To import the data from the package, see *Importing from Package* section of this user’s manual.

- Click **Analyze | Analysis of Variance | Completely Randomized Design...** from the main window. The **Completely Randomized Design** dialog box will appear as below:



- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and no **Set** variable is specified, STAR will perform the analysis as if the entire data came from only one set variable. If *perform combined analysis* is selected, a **Set** variable with at least two levels must be specified. The item should come from the **Factor(s)** list box.

Treatment(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Options Tab

Descriptive Statistics of the Response Variable

If this option is selected, a summary table with number of non-missing observations, minimum, maximum, mean and standard deviation of the response variable will be displayed.

Shapiro-Wilk Test for Normality of Residuals

If this option is selected, test for normality of residuals using Shapiro-Wilk will be displayed, if the number of observation is between 3 and 5000.

Bartlett's Test for Homogeneity of Variances

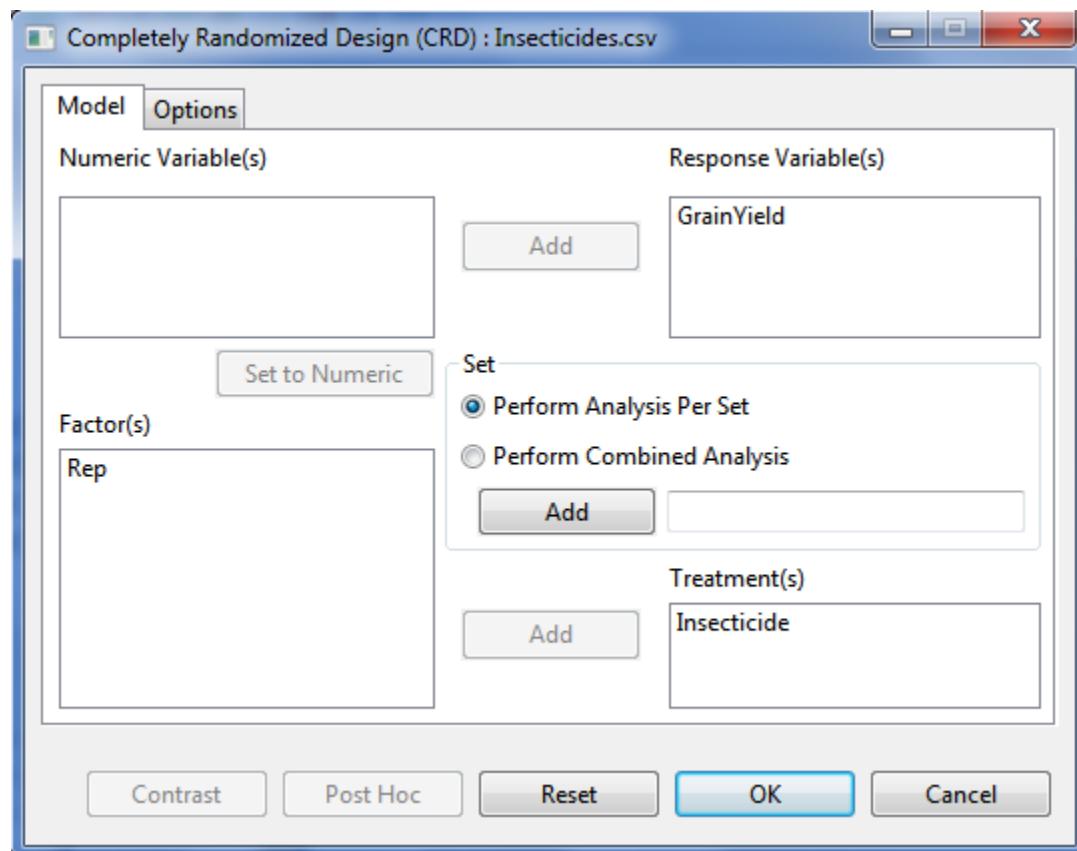
This option is visible if *perform analysis per set* is selected. If this option is selected, test for homogeneity of variances using Bartlett's Test will be displayed.

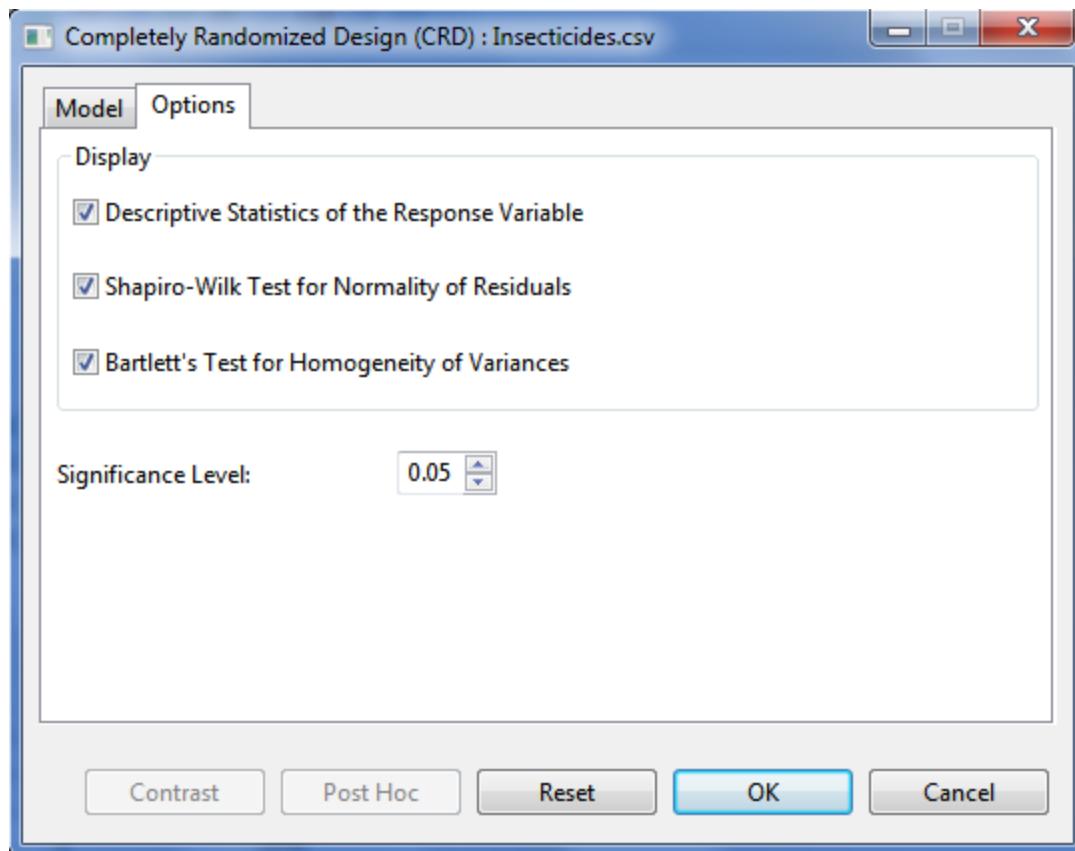
Significance Level

This will be used to determine if the treatment effects and interaction effects are significant in the Analysis of Variance. This is also the level of significance to be used

when pairwise mean comparison is performed. Its default value is 0.05. The user can change this value by specifying a numeric value from 0 to 1.

For the example, the completed **Model** tab and **Option** tab should appear as shown below:





- Click **OK** button to perform the analysis. The **Completely Randomized Design** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample text output is shown below.

```
Analysis of Variance
Completely Randomized Design
=====
ANALYSIS FOR RESPONSE VARIABLE: GrainYield
=====

Summary Information
-----
FACTOR      NO. OF LEVELS   LEVELS
-----
Insecticide  7           Azodin, Control, ..., Dol-Mix (2kg)
-----
Number of Observations Read and Used: 28
```

Descriptive Statistics

Variable	N_Obs	Min	Max	Mean	StdDev
GrainYield	28	1077	3366	2039.64	529.76

Test for Homogeneity of Variances

Method	DF	Chisq	Value	Pr(>Chisq)
Bartlett	6	5.56		0.4744

Test for Normality

Variable	Method	W	Value	Pr(< W)
GrainYield_resid	Shapiro-Wilk	0.9857		0.9587

ANOVA TABLE

Response Variable: GrainYield

Source	DF	Sum of Square	Mean Square	F Value	Pr(> F)
Insecticide	6	5587174.9286	931195.8214	9.83	0.0000
Error	21	1990237.5000	94773.2143		
Total	27	7577412.4286			

Summary Statistics

CV(%)	GrainYield	Mean
15.09		2039.64

Standard Errors

Effects	StdErr
Insecticide	217.68

Pairwise Mean Comparison of Insecticide

Tukeys's Honest Significant Difference (HSD) Test

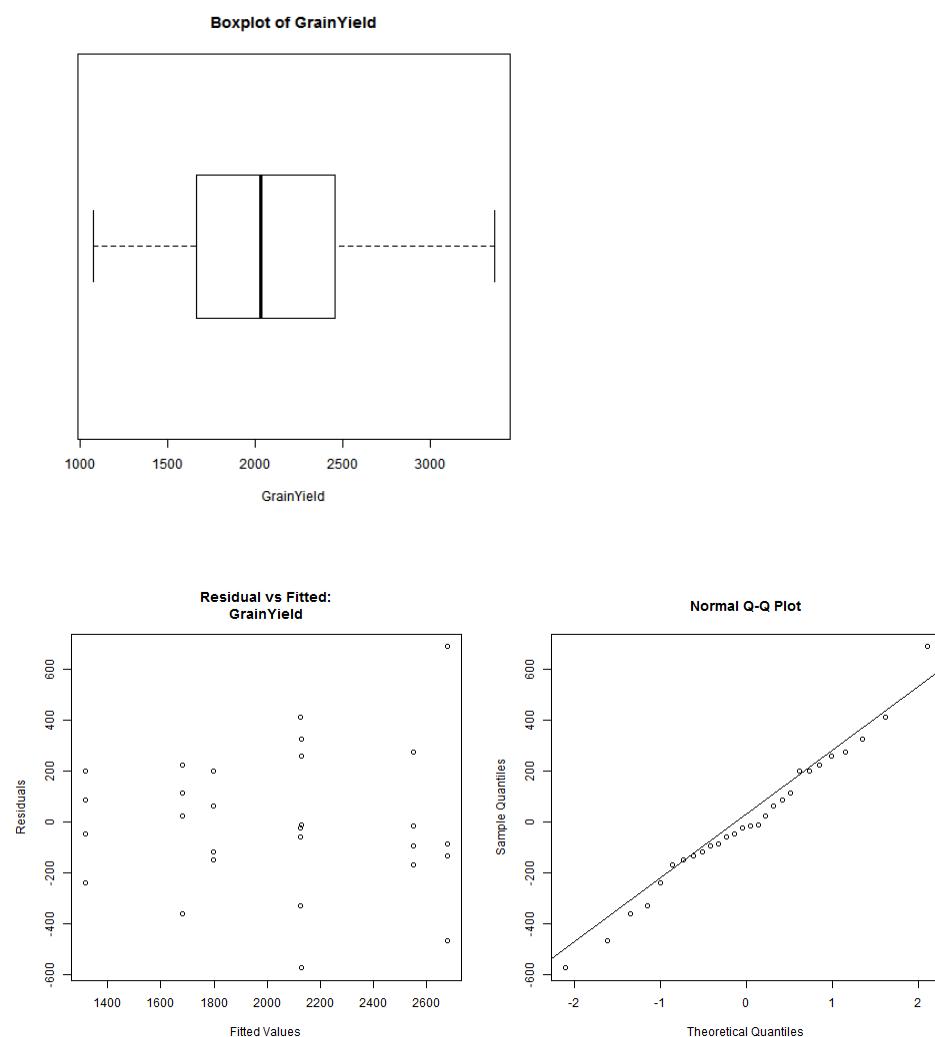
Alpha	0.05
Error Degrees of Freedom	21
Error Mean Square	94773.2143
Critical Value	4.5973
Test Statistics	707.6456

Summary of the Result:

Insecticide	means	N group
Azodin	2128.00	4 ab
Control	1316.00	4 c
DDT + phi-BHC	2551.75	4 a
Dimercon-Boom	1796.00	4 bc
Dimercon-Knap	1681.00	4 bc
Dol-Mix (1kg)	2126.75	4 ab
Dol-Mix (2kg)	2678.00	4 a

Means with the same letter are not significantly different.

Sample graphical output displayed in the **Graph** page.



Additional Options

If there are still modifications to be made, the following option buttons may be used:

- **Post Hoc**

Use this option if other pairwise comparison procedures, other than what was presented in the **Output** page, are desired. There are five pairwise mean comparison procedure available, namely: Least Significant Difference Test, Duncan Multiple Range Test, Tukey's Honest Significant Test, Student Newmann-Keul's Test and the Scheffe's Test. This button will be disabled if *F*-tests conducted using ANOVA yielded results that are not significant. Use of this button will be discussed in Pairwise Comparison section on this module.

- **Contrast**

This button will be enabled if there is at least one factor that has more than two levels. Use of this button will be discussed in Partitioning Sum of Squares section of this module.

Randomized Complete Block Design

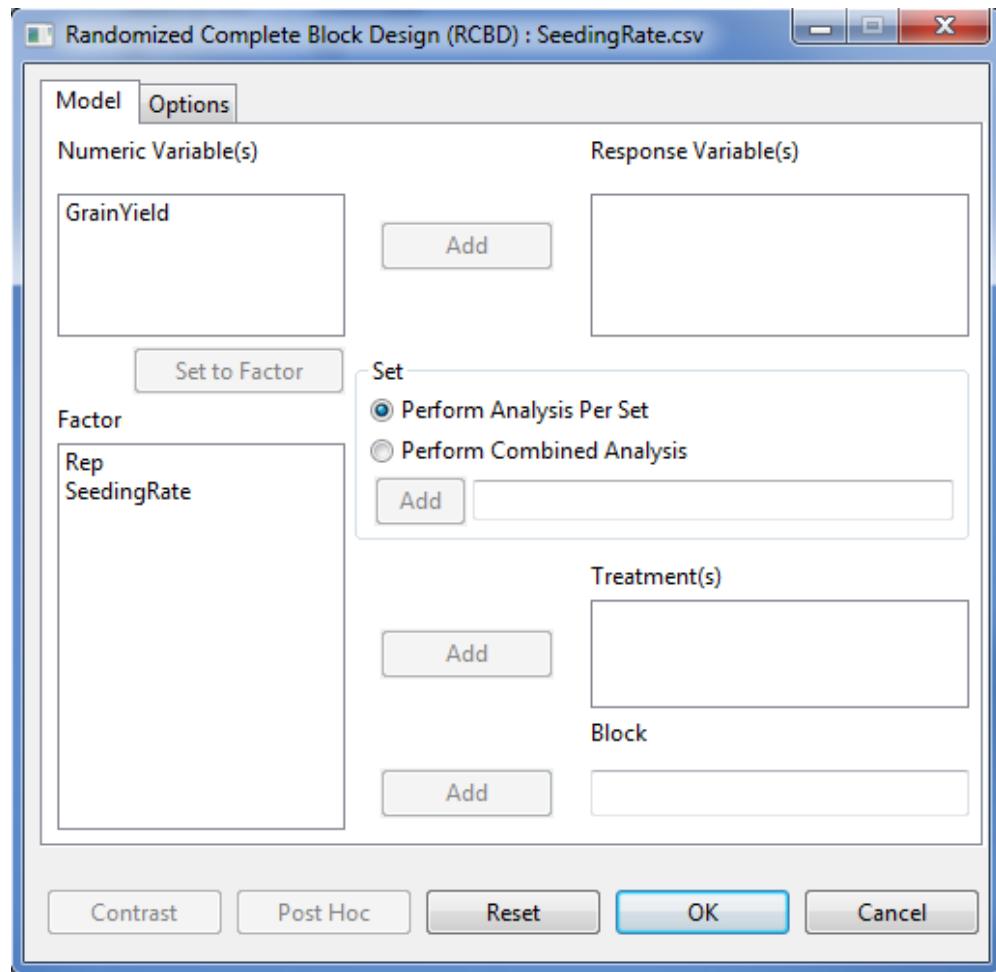
This submenu can be used to perform analysis of variance for single or multi factor experiment.

The steps to perform the analysis are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *SeedingRate* from package. To import that data, see *Importing from Project* section of this user's manual.

- From the main window of STAR click **Analyze | Analysis of Variance | Randomized Complete Block Design....** The **Randomized Complete Block Design** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only Items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default option) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and no **Set** variable is specified, STAR will perform the analysis as if the entire data came from only one set variable. If *perform combined analysis* is selected, a **Set** variable with at least two levels must be specified. The item should come from the **Factor(s)** list box.

Treatment(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

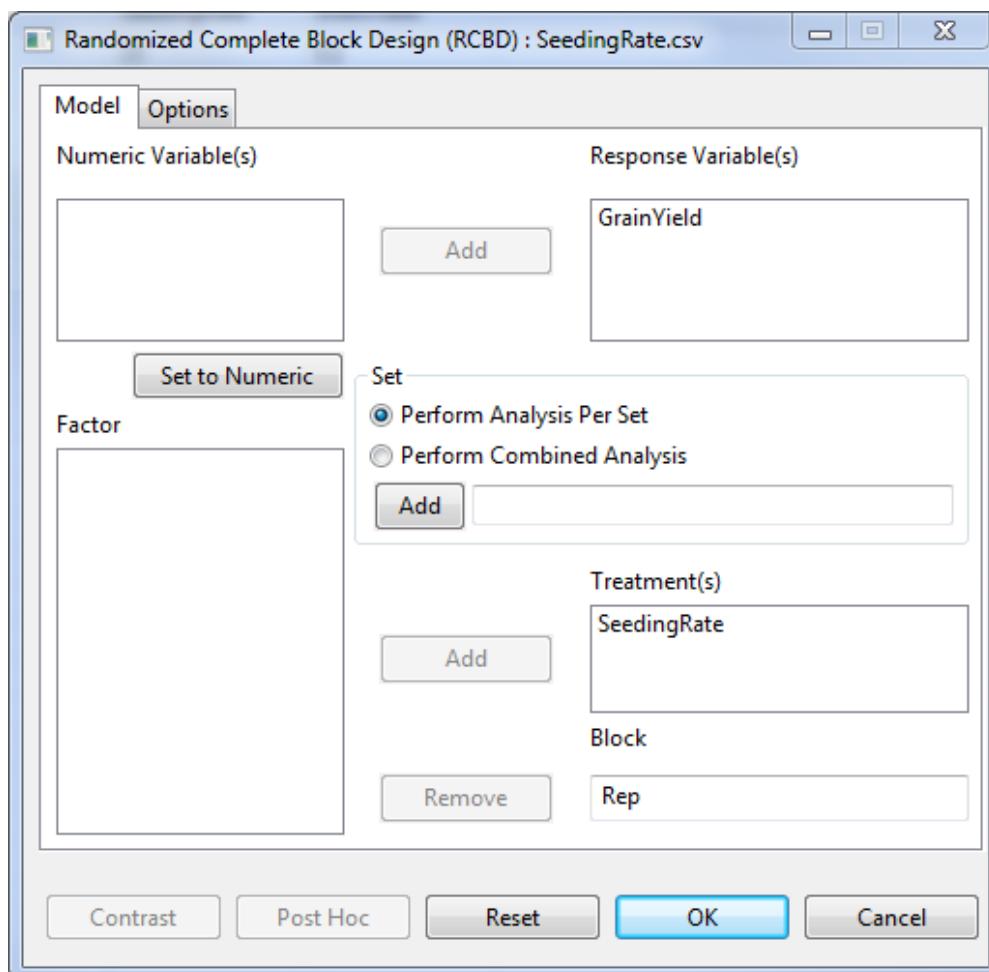
Block

Only one item can be added in this field and should come from the **Factor(s)** list box.

Options Tab

This tab lets the user identify additional output to be displayed and set the level of significance to be used. To know more about the options in detail, see the discussion under *Analysis of Variance > Completely Randomized Design > Options tab* of this user's manual.

For the example, the completed **Model** tab should appear as shown below:



- Click **OK** button to perform the analysis. The **Randomized Complete Block Design (RCBD)** dialog box will be minimized and STAR activates the **Output Viewer** page of the **Result Viewer** tab.

Sample text output is shown below:

```
Analysis of Variance
Randomized Complete Block Design
=====
ANALYSIS FOR RESPONSE VARIABLE: GrainYield
=====

Summary Information
-----
FACTOR      NO. OF LEVELS   LEVELS
-----
SeedingRate  6            25, 50, ..., 150
Rep          4            1, 2, 3, 4
-----
Number of Observations Read and Used: 24

Descriptive Statistics
-----
Variable     N_Obs    Min    Max    Mean    StdDev
-----
GrainYield   24       2.80   6      4.65   0.7735
-----
Test for Homogeneity of Variances
-----
Method      DF   Chisq Value  Pr(>Chisq)
-----
Bartlett    5      4.76    0.4462
-----
Test for Normality
-----
Variable      Method      W Value  Pr(< W)
-----
GrainYield_resid Shapiro-Wilk  0.9346  0.1235
-----
ANOVA TABLE
Response Variable: GrainYield
-----
Source        DF   Sum of Square  Mean Square  F Value  Pr(> F)
-----
Rep           3      1.7667      0.5889      4.87    0.0147
SeedingRate   5      10.1800     2.0360      16.84   0.0000
Error          15     1.8133      0.1209
Total          23     13.7600
-----
```

```
Summary Statistics
-----
CV(%)  GrainYield Mean
-----
7.48      4.65
-----

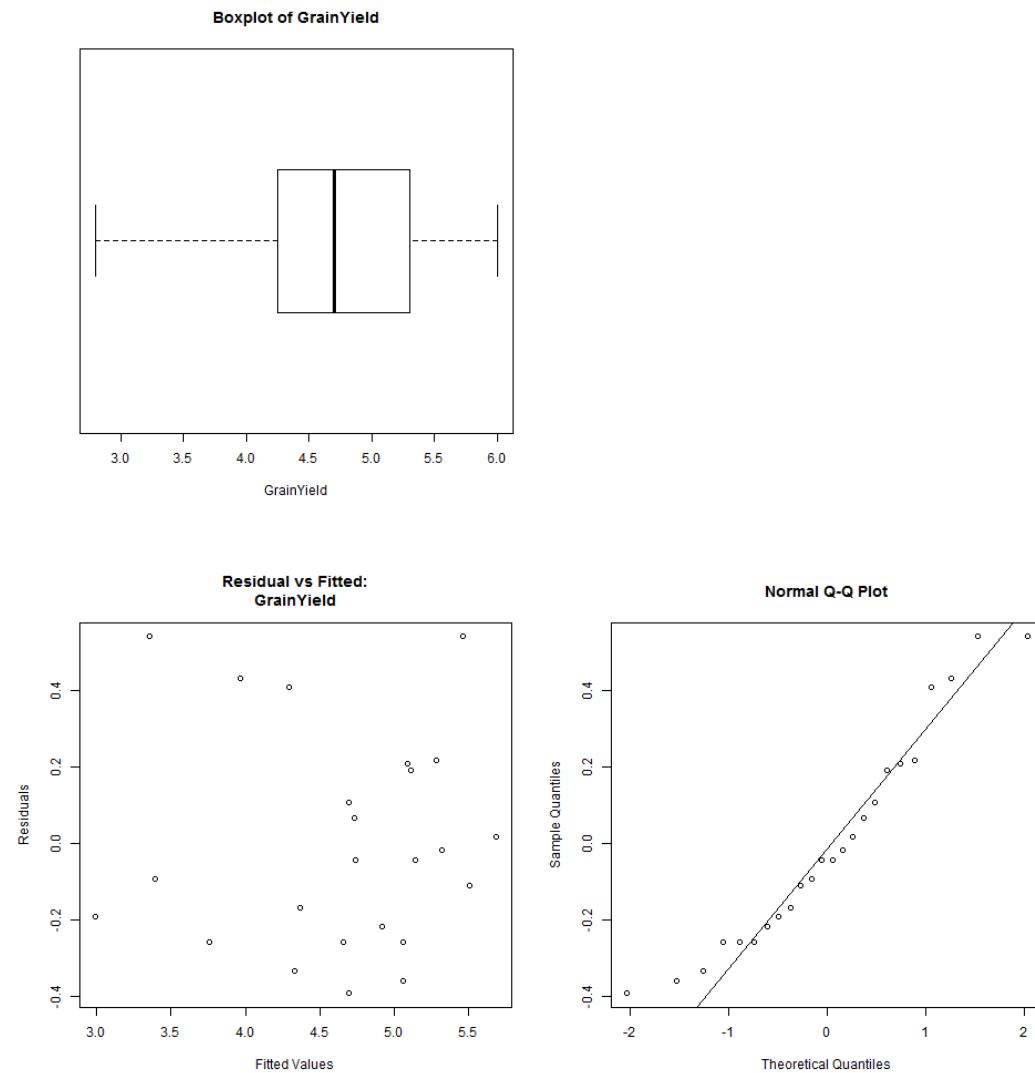
Standard Errors
-----
Effects      StdErr
-----
Rep          0.2007
SeedingRate  0.2459
-----

Pairwise Mean Comparison of SeedingRate
Tukeys's Honest Significant Difference (HSD) Test

Alpha           0.05
Error Degrees of Freedom    15
Error Mean Square       0.1209
Critical Value         4.5947
Test Statistics        0.7988

Summary of the Result:
-----
SeedingRate   means      N group
-----
25            5.12      4  ab
50            5.08      4  ab
75            5.30      4  a
100           4.35      4  b
125           4.67      4  ab
150           3.38      4  c
-----
Means with the same letter are not significantly different.
```

Sample graphical output of the analysis displayed in the Graph Viewer tab is shown below:



Modifications can still be made on the analysis submitted. To know more about the modification that can be made, see the discussion under *Analysis of Variance > Completely Randomized Design > Additional Options* of this user's manual.

Latin Square Design

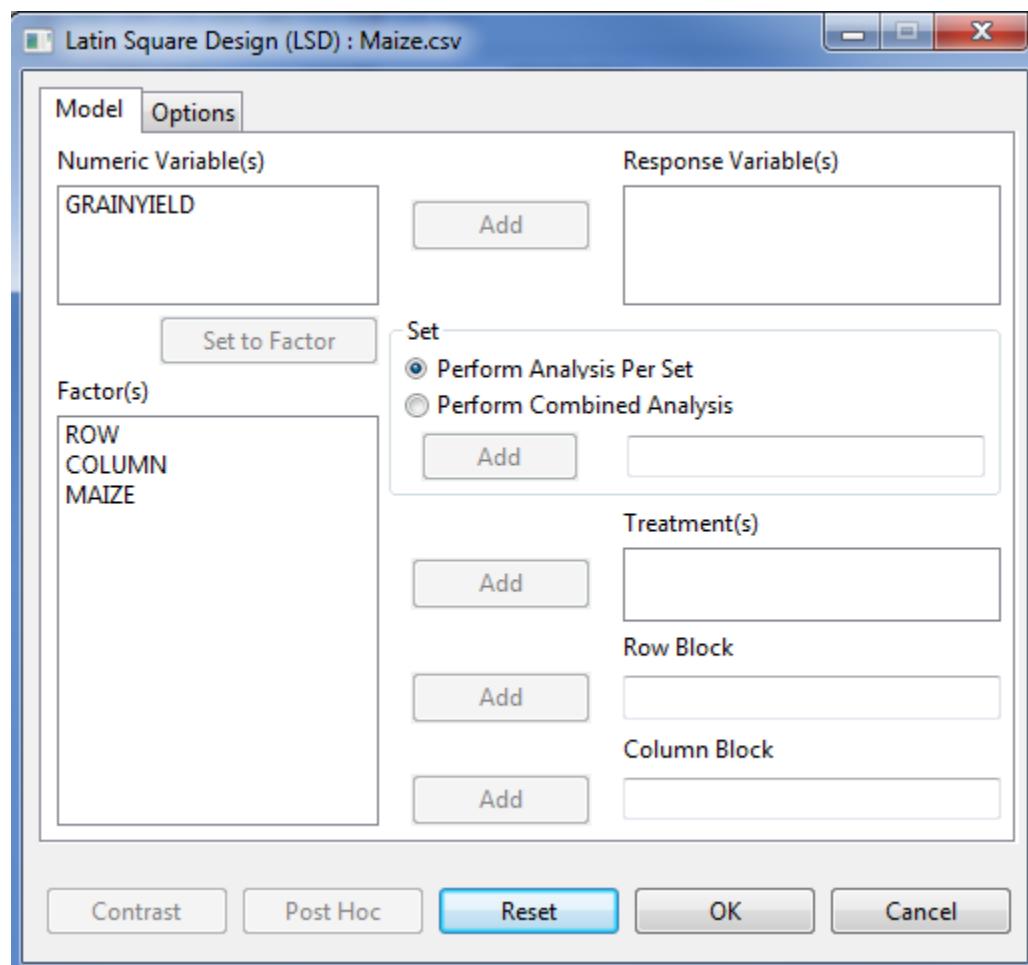
This submenu can be used to perform analysis of variance for single or multi factor experiment.

The steps to perform the analysis are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *Maize* from package. To import that data, see *Importing from Package* section of this user's manual.

- From the main window of STAR click **Analyze | Analysis of Variance | Latin Square Design....** The **Latin Square Design** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only Items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default option) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and no **Set** variable is specified, STAR will perform the analysis as if the entire data came from only one set variable. If *perform combined analysis* is selected, a **Set** variable with at least two levels must be specified. The item should come from the **Factor(s)** list box.

Treatment(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Row Block

Only one item can be added in this field and should come from the **Factor(s)** list box.

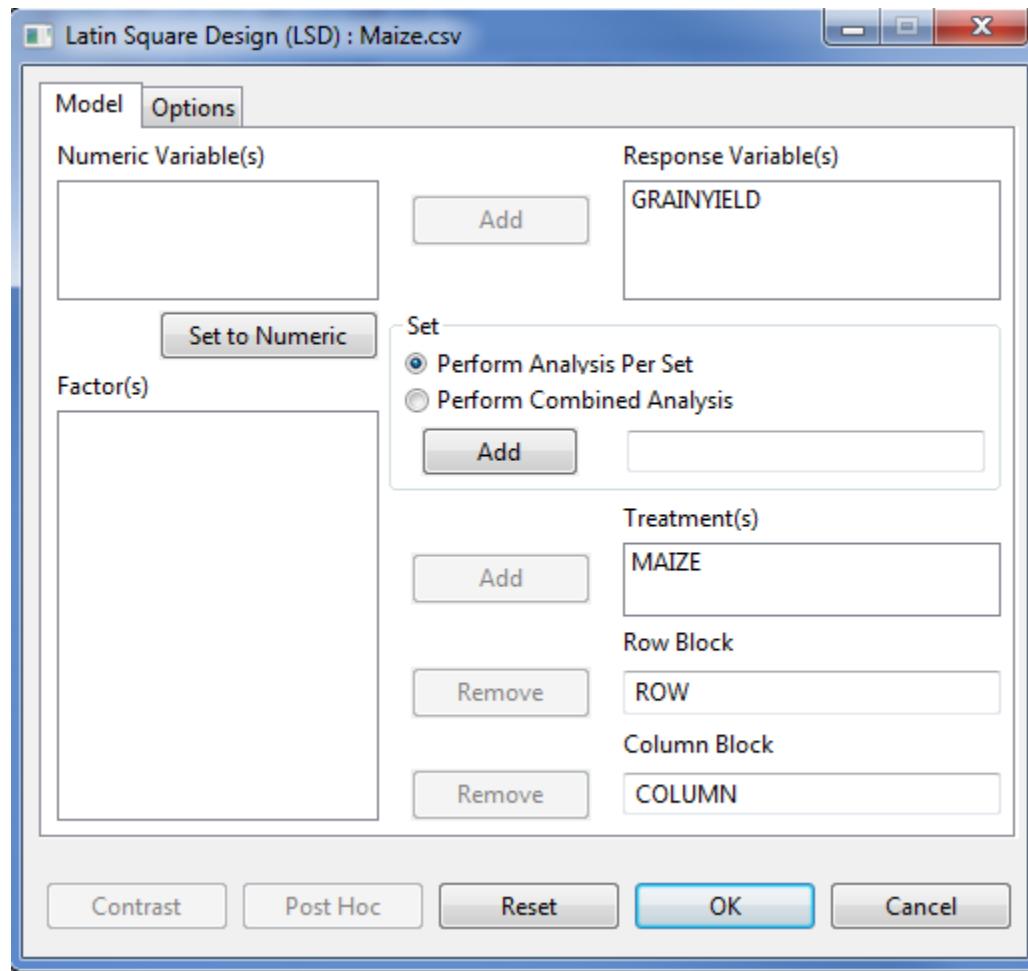
Column Block

Only one item can be added in this field and should come from the **Factor(s)** list box.

Options Tab

This tab lets the user identify additional output to be displayed and set the level of significance to be used. To know more about the options in detail, see the discussion under *Analysis of Variance > Completely Randomized Design > Options tab* of this user's manual.

For the example, the completed **Model** tab should appear as illustrated below:



- Click **OK** button to perform the analysis. The **Latin Square Design** dialog box will be minimized and STAR activates the **Output Viewer** tab and the **Graph Viewer** tab.

Sample text output is shown below:

```
Analysis of Variance
Latin Square Design
=====
ANALYSIS FOR RESPONSE VARIABLE: GRAINYIELD
=====

Summary Information
-----
FACTOR    NO. OF LEVELS   LEVELS
-----
MAIZE      4             A, B, C, D
ROW         4             1, 2, 3, 4
COLUMN     4             1, 2, 3, 4
-----
Number of Observations Read and Used: 16

Descriptive Statistics
-----
Variable      N_Obs      Min      Max      Mean      StdDev
-----
GRAINYIELD    16       0.6600   1.67    1.34     0.3070
-----

Test for Homogeneity of Variances
-----
Method        DF   Chisq Value  Pr(>Chisq)
-----
Bartlett     3      3.81      0.2825
-----

Test for Normality
-----
Variable      Method      W Value  Pr(< W)
-----
GRAINYIELD_resid  Shapiro-Wilk  0.9250  0.2026
-----

ANOVA TABLE
Response Variable: GRAINYIELD
-----
Source      DF   Sum of Square   Mean Square   F Value  Pr(> F)
-----
ROW         3      0.0302      0.0101      0.47    0.7170
COLUMN     3      0.8273      0.2758     12.77    0.0051
MAIZE       3      0.4268      0.1423      6.59    0.0251
Error        6      0.1296      0.0216
Total       15      1.4139
-----
```

Summary Statistics

CV(%) GRAINYIELD Mean

11.01 1.34

Standard Errors

Effects StdErr

ROW 0.1039
COLUMN 0.1039
MAIZE 0.1039

Pairwise Mean Comparison of MAIZE

Least Significant Difference (LSD) Test

Alpha	0.05
Error Degrees of Freedom	6
Error Mean Square	0.0216
Critical Value	2.4469
Test Statistics	0.2543

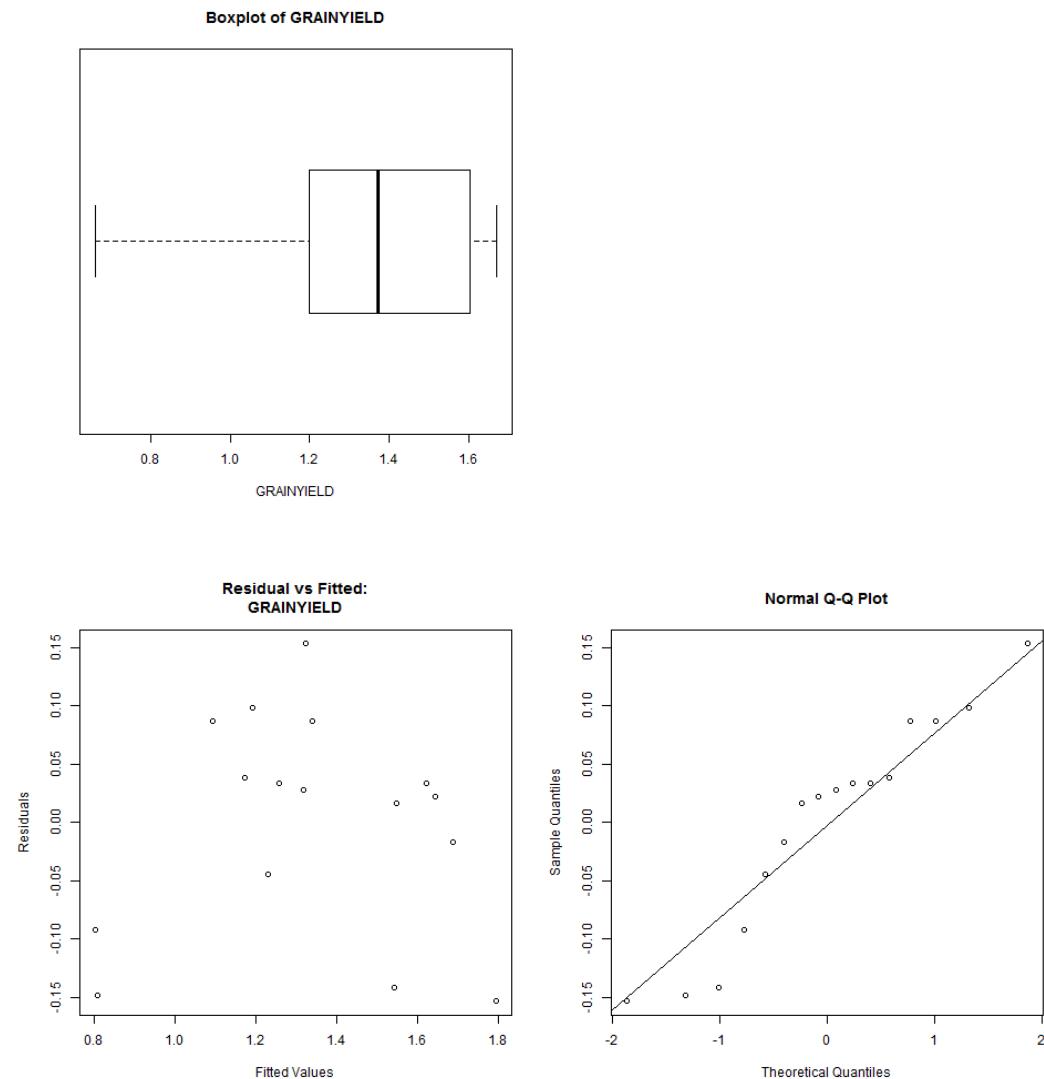
Summary of the Result:

MAIZE means N group

A 1.46 4 a
B 1.47 4 a
C 1.07 4 b
D 1.34 4 a

Means with the same letter are not significantly different.

Sample graphical output of the analysis displayed in the Graph Viewer tab is shown below:



Modifications can still be made on the analysis submitted. To know more about the modification that can be made, see the discussion under *Analysis of Variance > Completely Randomized Design > Additional Options* of this user's manual.

Split Plot Design

This analysis can be performed for experiments where the main plot is arranged in completely randomized, randomized complete block or Latin square design.

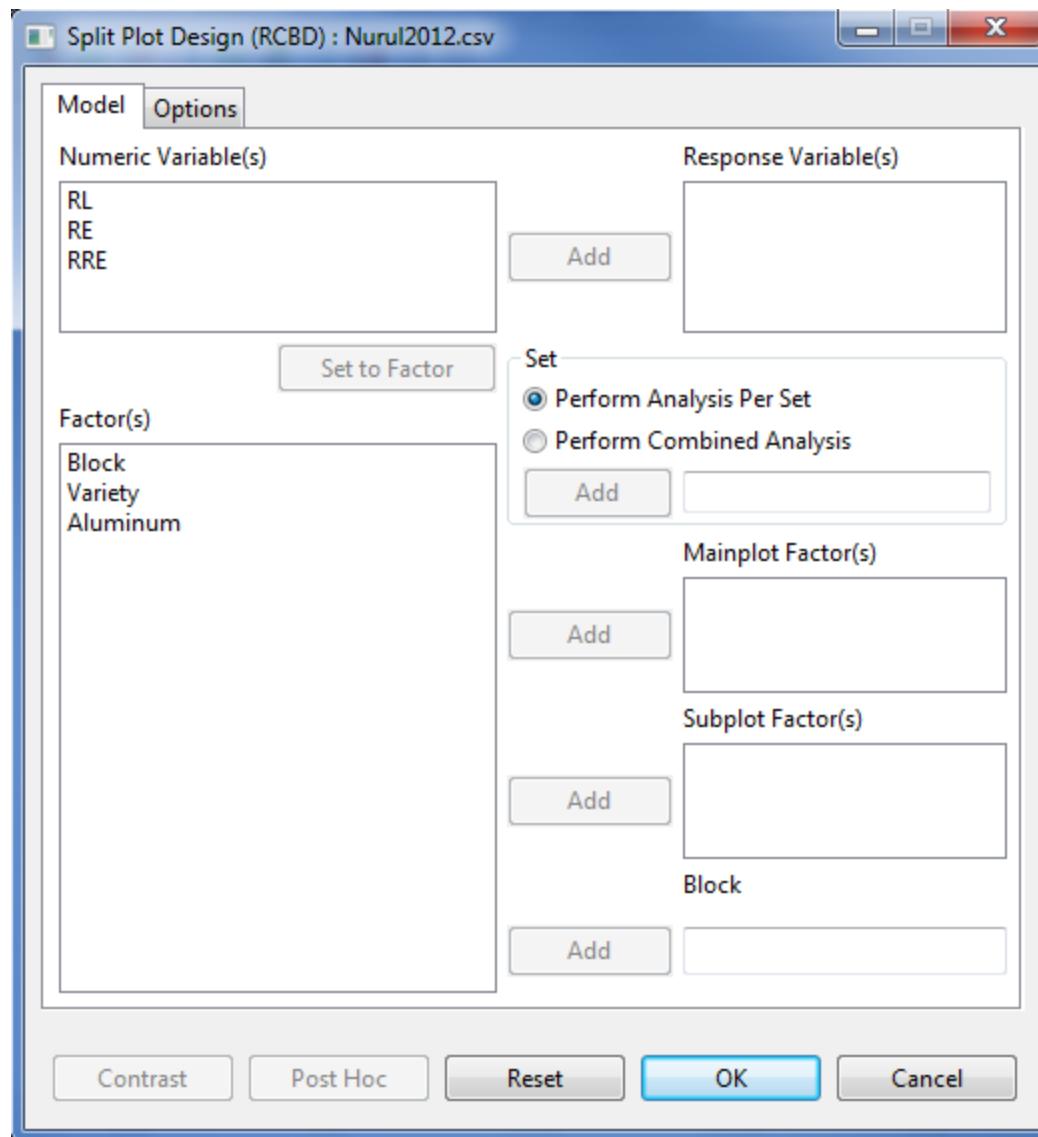
The steps to perform the Analysis of Variance using Split Plot Design are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *My Project*, import the data *Nurul2012* from the package. To import the data from the package, see *Importing from Package* section of this user's manual.

- From the main window of STAR click **Analyze | Analysis of Variance | Split Plot Design**. Choose the appropriate design for the main plot. The **Split Plot Design** dialog box will appear.

For the example, choose **Analyze | Analysis of Variance | Split Plot Design | Randomized Complete Block Design ...** from the main window.



- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and no **Set** variable is specified, STAR will perform the analysis as if the

entire data came from only one set variable. If *perform combined analysis* is selected, a **Set** variable with at least two levels must be specified. The item should come from the **Factor(s)** list box.

Mainplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Subplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Replicate

Applicable only if completely randomized design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Block

Applicable only if randomized complete block design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Row Block

Applicable only if Latin square design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

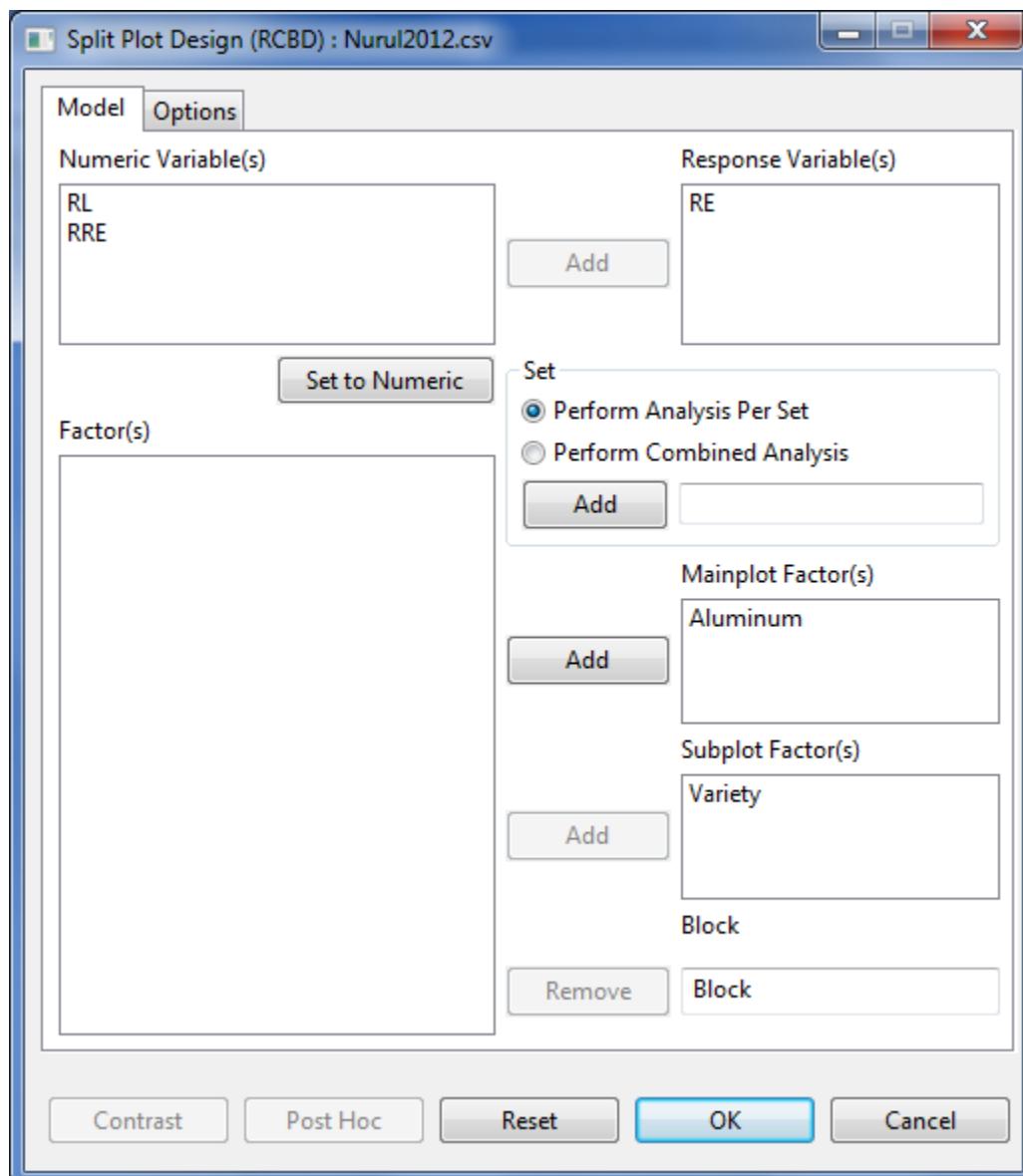
Column Block

Applicable only if Latin square design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Options Tab

This tab lets the user identify additional output to be displayed and set the level of significance to be used. To know more about the options in detail, see the discussion under *Analysis of Variance > Completely Randomized Design > Options tab* of this user's manual.

For the example, the completed **Model** tab should appear as illustrated:



- Click **OK** button to perform the analysis. The **Split Plot Design** dialog box will be minimized and STAR activates the **Output Viewer** tab and the **Graph Viewer** tab.

Sample text output is shown below:

```
Analysis of Variance
Split Plot in Randomized Complete Block Design
=====
ANALYSIS FOR RESPONSE VARIABLE: RE
=====

Summary Information
-----
FACTOR      NO. OF LEVELS   LEVELS
-----
Aluminum    4              0, 540, 750, 1300
Variety     8              Azucena112854, Azucena47125, ..., Pokalli
Block       2              I, II
-----
Number of Observations Read and Used: 64

Descriptive Statistics
-----
Variable   N_Obs      Min      Max      Mean     StdDev
-----
RE         64        13.62    107.62   55.93    23.38
-----

Test for Homogeneity of Variances
-----
Method      DF   Chisq Value  Pr(>Chisq)
-----
Bartlett    31      28.15      0.6136
-----

Test for Normality
-----
Variable  Method          W Value  Pr(< W)
-----
RE_resid  Shapiro-Wilk    0.9926    0.9675
-----

ANOVA TABLE
Response Variable: RE
-----
Source           DF   Sum of Square  Mean Square   F Value  Pr(> F)
-----
Block            1      647.2572    647.2572    177.11    0.0009
Aluminum        3     11826.5606   3942.1869    1078.68    0.0000
Error(a)        3      10.9639     3.6546
Variety         7     17768.2497   2538.3214    64.04    0.0000
Aluminum:Variety 21     3085.8037   146.9430     3.71    0.0007
Error(b)        28     1109.9079    39.6396
Total           63     34448.7430
-----
```

Summary Statistics

CV(a) %	CV(b) %	RE Mean
3.42	11.26	55.93

Comparison of Aluminum at each level of Variety

Least Significant Difference (LSD) Test

Alpha	0.05
Error Degrees of Freedom	28
Error Mean Square	35.1415
Critical Value	2.0462
Test Statistic	12.1297

Summary:

Aluminum	N	Variety = Azucena112854 group	...
0	2	65.1400 a	...
540	2	54.9500 ab	...
750	2	44.6000 b	...
1300	2	23.0750 c	...

Means with the same letter are not significantly different

Comparison of Variety at each level of Aluminum

Tukeys's Honest Significant Difference (HSD) Test

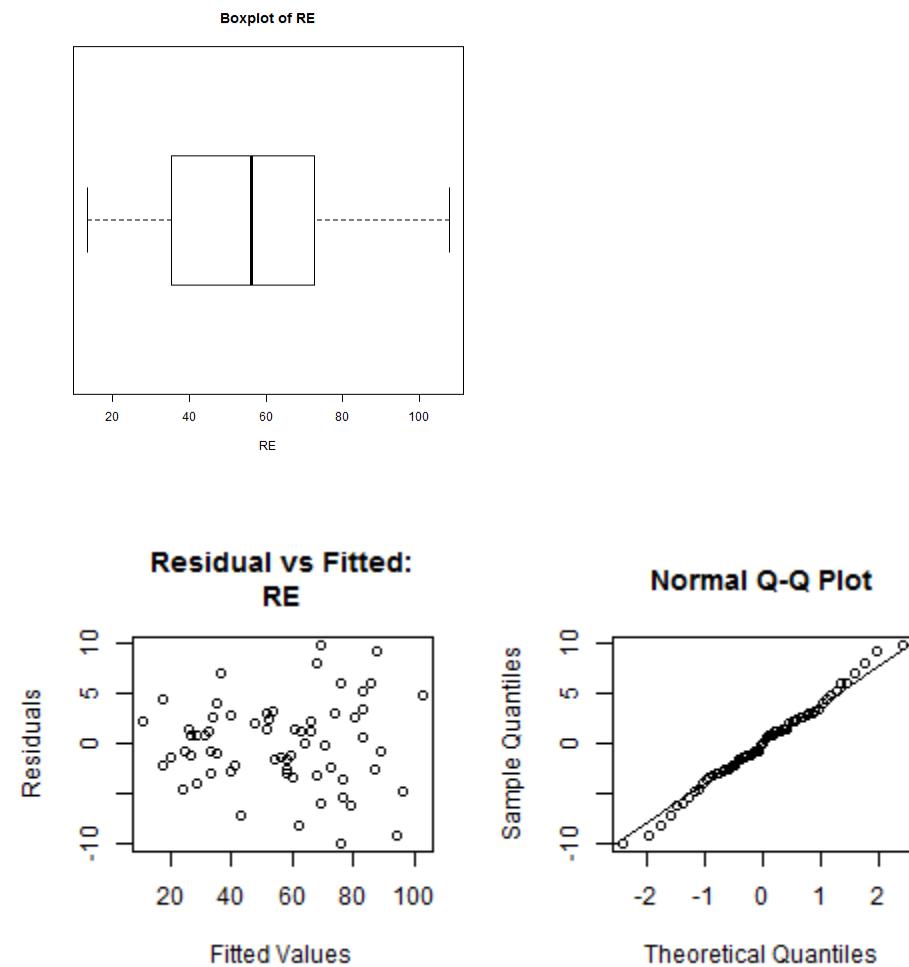
Alpha	0.05
Error Degrees of Freedom	28
Error Mean Square	39.6396
Critical Value	4.6248
Test Statistic	20.5893

Summary:

Variety	N	Aluminum = 0 group	Aluminum = 540 group	...
Azucena112854	2	65.1400 bc	54.9500 bc	...
Azucena47125	2	79.6300 ab	82.3350 a	...
Azucena52992	2	99.3100 a	86.0500 a	...
Chadungda	2	54.8650 c	36.9400 c	...
IR20	2	32.0500 d	40.0300 c	...
IR64	2	56.9600 c	70.9350 ab	...
IR74	2	67.3950 bc	83.8500 a	...
Pokalli	2	91.1000 a	79.6000 a	...

Means with the same letter are not significantly different

Sample graphical output of the analysis displayed in the Graph Viewer tab is shown below:



Modifications can still be made on the analysis submitted. To know more about the modification that can be made, see the discussion under *Analysis of Variance > Completely Randomized Design > Additional Options* of this user's manual.

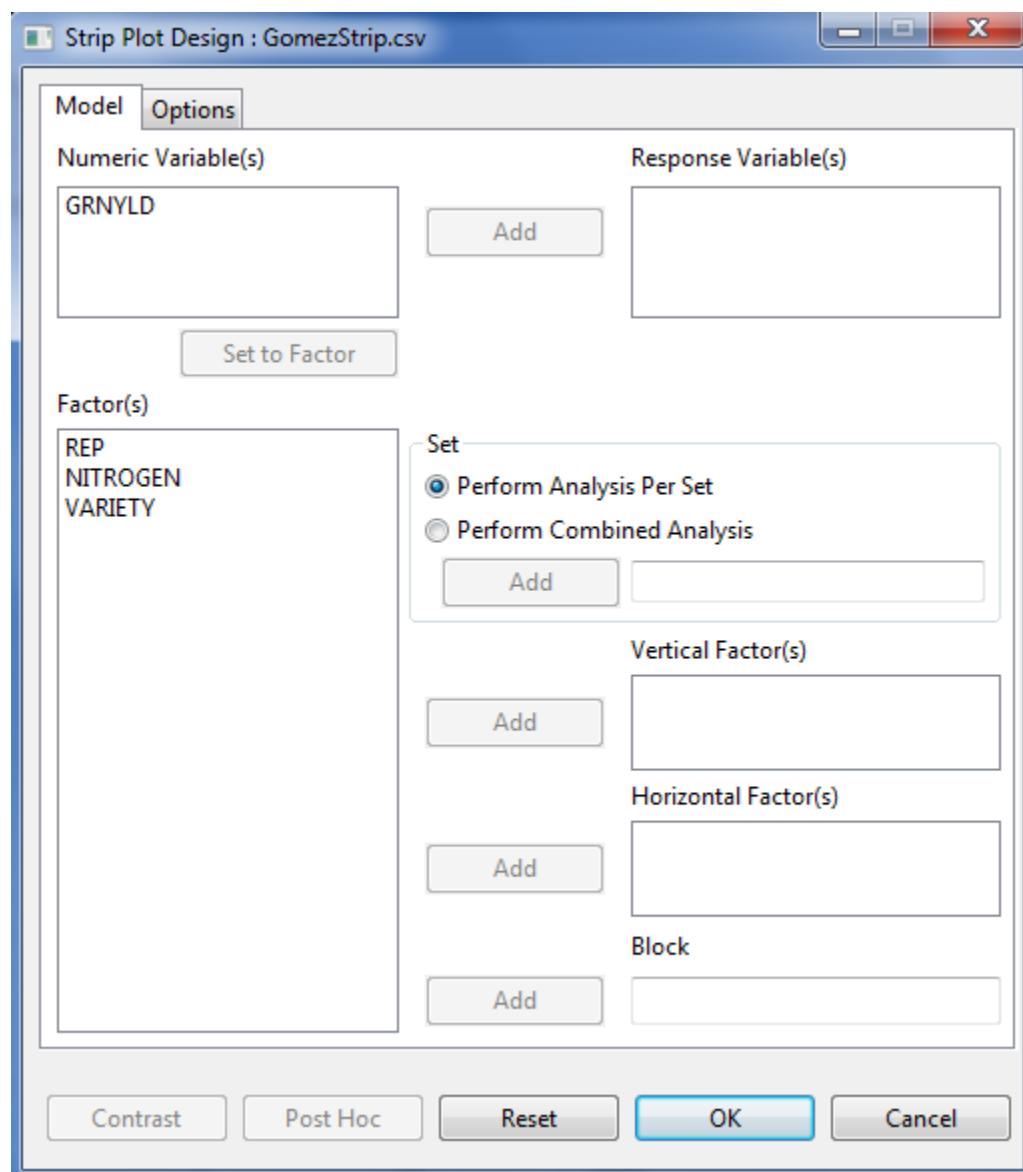
Strip Plot Design

The steps to perform the Analysis of Variance in Strip Plot Design are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *GomezStrip* from package. To import the data from the package, see *Importing from Package* section of this user's manual.

- From the main window of STAR click **Analyze | Analysis of Variance | Strip Plot Design**. The **Strip Plot** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only Items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default option) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and the **Set** variable is not specified, STAR will perform the analysis as if the entire data came from only one site. If **perform combined analysis is selected**, a Set variable must be specified.

Vertical Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Horizontal Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

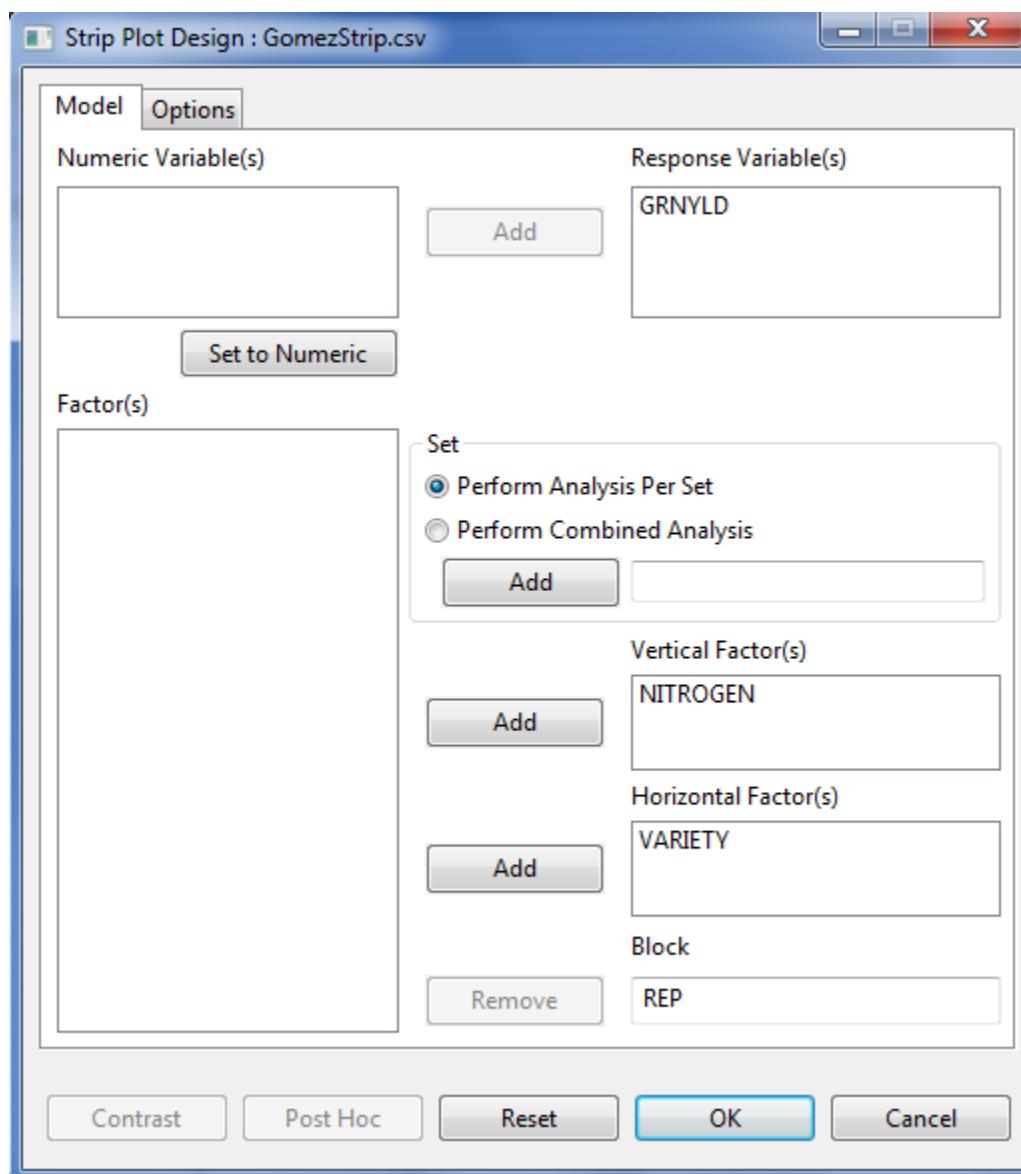
Block

Only one item can be added in this field and should come from the **Factor(s)** list box.

Options Tab

This tab lets the user identify additional output to be displayed and set the level of significance to be used. To know more about the options in detail, see the discussion under *Analysis of Variance > Completely Randomized Design > Options tab* of this user's manual.

For the example, the completed **Model** Specification tab should appear as shown below:



- Click **OK** button to perform the analysis. The **Strip Plot Design** dialog box will be minimized and STAR activates the **Output Viewer** tab and the **Graph Viewer** tab.

Sample text output is shown below:

```
Analysis of Variance
Strip Plot Design
=====
ANALYSIS FOR RESPONSE VARIABLE: GRNYLD
=====

Summary Information
-----
FACTOR      NO. OF LEVELS   LEVELS
-----
NITROGEN    3              0, 60, 120
VARIETY     6              V1, V2, ..., V6
REP          3              1, 2, 3
-----
Number of Observations Read and Used: 54

Descriptive Statistics
-----
Variable   N_Obs   Min     Max      Mean    StdDev
-----
GRNYLD     54     1556   8611   5289.94  1775.12
-----

Test for Homogeneity of Variances
-----
Method      DF   Chisq Value  Pr(>Chisq)
-----
Bartlett    17     21.30    0.2130
-----

Test for Normality
-----
Variable      Method      W Value  Pr(< W)
-----
GRNYLD_resid  Shapiro-Wilk  0.9732  0.2665
-----
```

ANOVA TABLE

Response Variable: GRNYLD

Source	DF	Sum of Square	Mean Square	F Value	Pr(> F)
REP	2	9220962.3333	4610481.1667	6.20	0.0595
NITROGEN	2	50676061.4444	25338030.7222	34.07	0.0031
Error(a)	4	2974907.8889	743726.9722		
VARIETY	5	57100201.2778	11420040.2556	7.65	0.0034
Error(b)	10	14922619.2222	1492261.9222		
NITROGEN:VARIETY	10	23877979.4444	2387797.9444	5.80	0.0004
Error(c)	20	8232917.2222	411645.8611		
Total	53	167005648.8333			

Summary Statistics

CV(a) %	CV(b) %	CV(c) %	GRNYLD Mean
16.30	23.09	12.13	5289.94

Comparison of NITROGEN at each level of VARIETY

Least Significant Difference (LSD) Test

Alpha	0.05
Error Degrees of Freedom	22
Error Mean Square	466992.7130
Critical Value	2.0716
Test Statistic	1155.8848

Summary:

NITROGEN	N	VARIETY = V1 group	VARIETY = V2 group	VARIETY = V3 group
0	3	3571.6667 c	4934.3333 b	4249.6667 c ...
60	3	5132.0000 b	6713.6667 a	6122.3333 b ...
120	3	7548.0000 a	7211.3333 a	7868.3333 a ...

Means with the same letter are not significantly different

Comparison of VARIETY at each level of NITROGEN

Tukey's Honest Significant Difference (HSD) Test

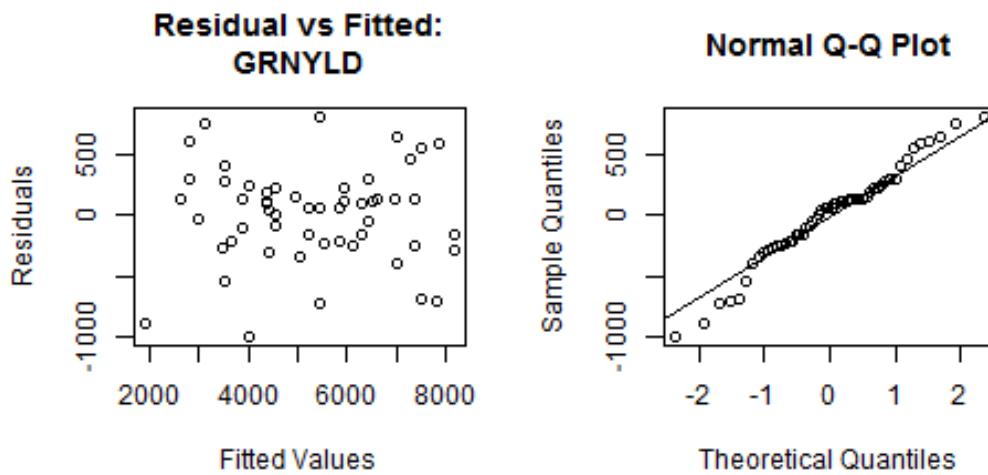
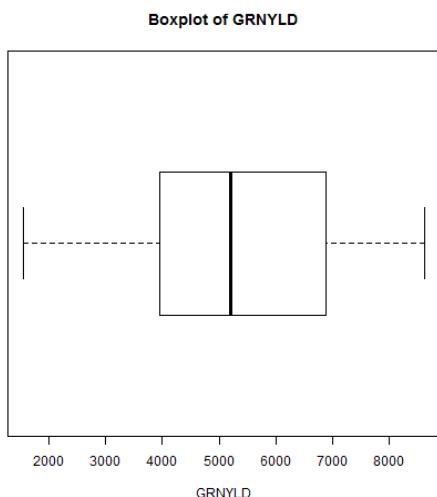
Alpha	0.05
Error Degrees of Freedom	10
Error Mean Square	1492261.9222
Critical Value	4.9120
Test Statistic	3464.3491

Summary:

VARIETY	N	NITROGEN = 0 group	NITROGEN = 60 group	NITROGEN = 120 group
V1	3	3571.6667 a	5132.0000 a	7548.0000 a
V2	3	4934.3333 a	6713.6667 a	7211.3333 a
V3	3	4249.6667 a	6122.3333 a	7868.3333 a
V4	3	4059.0000 a	5553.6667 a	7094.3333 a
V5	3	4101.6667 a	5633.3333 a	6012.0000 a
V6	3	3207.3333 a	3714.3333 a	2492.0000 b

Means with the same letter are not significantly different

Sample graphical output of the analysis displayed in the Graph Viewer tab is shown below:



Modifications can still be made on the analysis submitted. To know more about the modification that can be made, see the discussion under *Analysis of Variance > Completely Randomized Design > Additional Options* of this user's manual.

Split-Split Plot Design

This analysis can be performed for experiments where the main plot is arranged in completely randomized, randomized complete block or Latin square design.

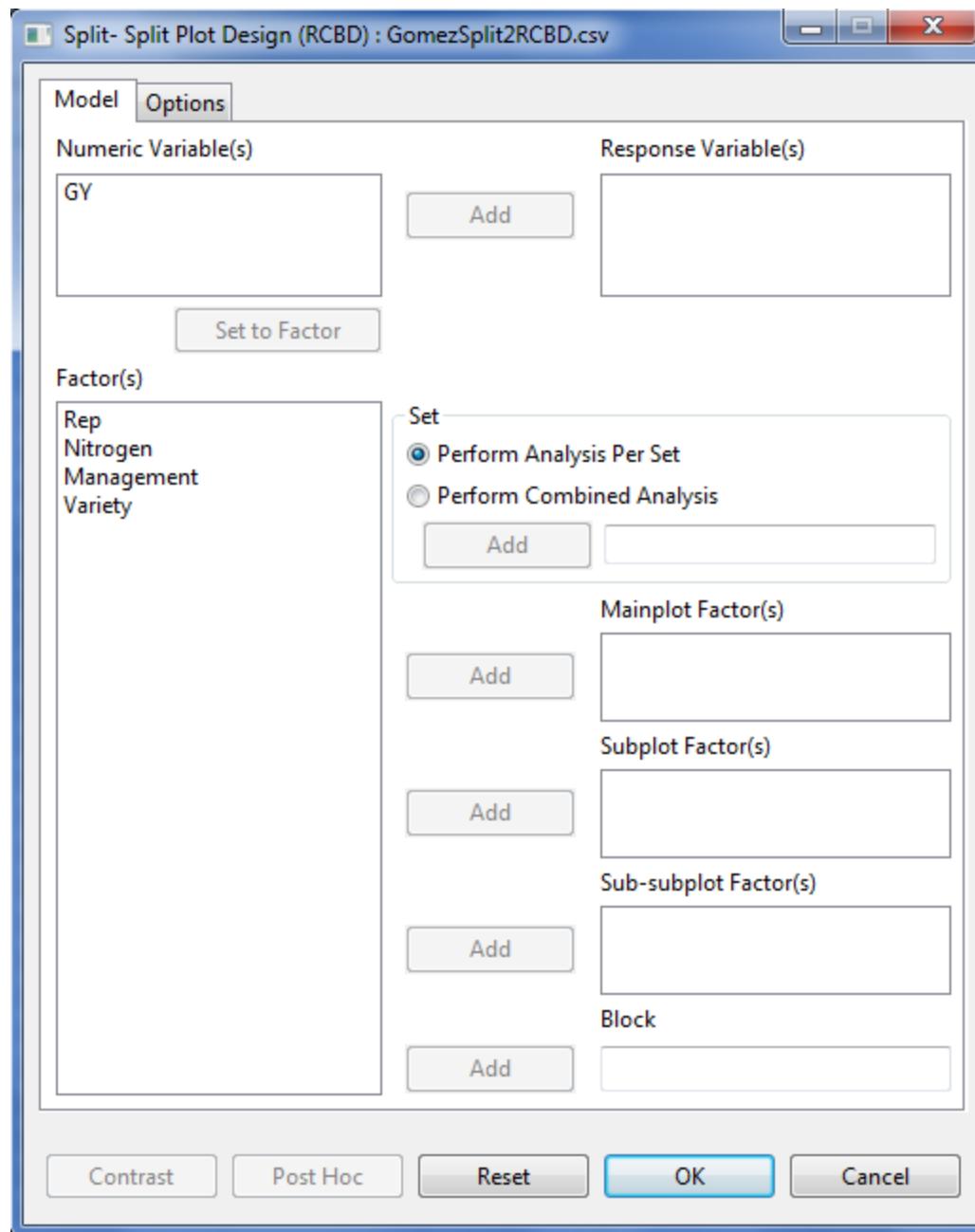
The steps to perform the Analysis of Variance using Split-Split Plot Design are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *GomezSplit2RCBD* from the package. To import the data from the package, see *Importing from Package* section of this user's manual.

- From the main window of STAR click **Analyze | Analysis of Variance | Split-Split Plot Design**. Choose the appropriate design for the mainplot. The **Split-Split Plot Design** dialog box will appear.

For the example, choose **Analyze | Analysis of Variance | Split-Split Plot Design | Randomized Complete Block Design ...** from the main window.



- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only Items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and no **Set** variable is specified, STAR will perform the analysis as if the entire data came from only one set variable. If *perform combined analysis* is selected, a **Set** variable with at least two levels must be specified. The item should come from the **Factor(s)** list box.

Mainplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Subplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Sub-subplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Replicate

Applicable only if completely randomized design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Block

Applicable only if randomized complete block design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Row Block

Applicable only if Latin square design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Column Block

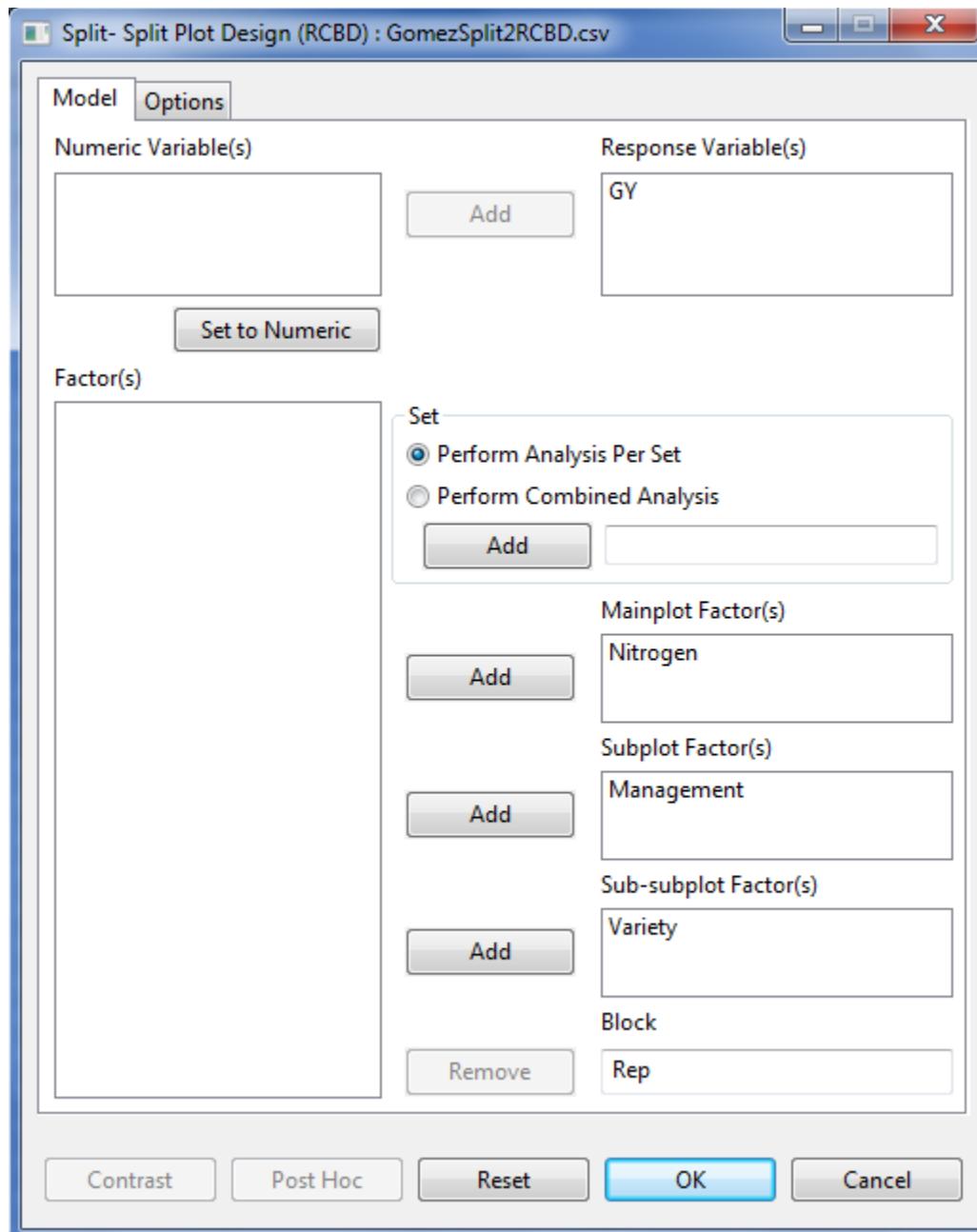
Applicable only if Latin square design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Options Tab

This tab lets the user identify additional output to be displayed and set the level of significance to be used. To know more about the options in detail, see the discussion

under *Analysis of Variance > Completely Randomized Design > Options tab* of this user's manual.

For the example, the completed **Model** tab should appear as illustrated:



- Click **OK** button to perform the analysis. The **Split-Split Plot Design** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample text output is shown below:

Analysis of Variance
Split-Split Plot in Randomized Complete Block Design

=====

ANALYSIS FOR RESPONSE VARIABLE: GY

=====

Summary Information

FACTOR	NO. OF LEVELS	LEVELS
--------	---------------	--------

Nitrogen	5	N1, N2, N3, N4, N5
Management	3	M1, M2, M3
Variety	3	V1, V2, V3
Rep	3	1, 2, 3

Number of Observations Read and Used: 135

Descriptive Statistics

Variable	N_Obs	Min	Max	Mean	StdDev
GY	135	3.13	10.36	6.55	1.67

Test for Homogeneity of Variances

Method	DF	Chisq	Value	Pr(>Chisq)
Bartlett	44	44.66		0.4440

Test for Normality

Variable	Method	W	Value	Pr(< W)
GY_resid	Shapiro-Wilk	0.9951		0.9279

ANOVA TABLE

Response Variable: GY

Source	DF	Sum of Square	Mean Square	F Value	Pr(> F)
Rep	2	0.7320	0.3660	0.66	0.5439
Nitrogen	4	61.6408	15.4102	27.70	0.0001
Error(a)	8	4.4514	0.5564		
Management	2	42.9361	21.4681	82.00	0.0000
Nitrogen:Management	8	1.1030	0.1379	0.53	0.8226
Error(b)	20	5.2363	0.2618		
Variety	2	206.0132	103.0066	207.87	0.0000
Nitrogen:Variety	8	14.1445	1.7681	3.57	0.0019
Management:Variety	4	3.8518	0.9629	1.94	0.1149
Nitrogen:Management:Variety	16	3.6992	0.2312	0.47	0.9538
Error(c)	60	29.7325	0.4955		
Total	134	373.5407			

Summary Statistics

CV(a) %	CV(b) %	CV(c) %	GY Mean
11.38	7.81	10.74	6.55

Table of Means

Management	Variety	Nitrogen	N1	N2	N3	N4	N5
	V1	3.897000	4.232000	5.226000	4.650333	4.061667	
M1	V2	5.346000	5.921667	5.979667	6.252000	6.775667	
	V3	5.378333	6.781333	7.933333	7.452667	8.618000	
	V1	4.317333	4.576333	5.732333	5.447333	5.200000	
M2	V2	4.711667	5.815000	6.500000	7.004000	7.084333	
	V3	6.496000	7.816667	8.567333	8.624667	9.399333	
	V1	5.325000	5.482667	6.546333	6.236333	5.971667	
M3	V2	5.431000	6.312000	7.286000	7.517667	8.005333	
	V3	7.560000	9.045333	9.190667	9.251333	9.989333	

Comparison of Nitrogen at each level of Variety

Least Significant Difference (LSD) Test

Alpha	0.05
Error Degrees of Freedom	43
Error Mean Square	0.5158
Critical Value	2.0160
Test Statistic	0.6826

Summary:

Nitrogen	N	Variety = V1 group	Variety = V2 group	Variety = V3 group
N1	9	4.5131 c	5.1629 d	6.4781 d
N2	9	4.7637 bc	6.0162 c	7.8811 c
N3	9	5.8349 a	6.5886 bc	8.5638 b
N4	9	5.4447 ab	6.9246 ab	8.4429 bc
N5	9	5.0778 bc	7.2884 a	9.3356 a

Means with the same letter are not significantly different

Comparison of Variety at each level of Nitrogen

Least Significant Difference (LSD) Test

Alpha	0.05
Error Degrees of Freedom	60
Error Mean Square	0.4955
Critical Value	2.0003
Test Statistic	0.6638

Summary:

Variety	N	Nitrogen = N1 group	Nitrogen = N2 group	Nitrogen = N3 group ...
V1	9	4.5131 b	4.7637 c	5.8349 c ...
V2	9	5.1629 b	6.0162 b	6.5886 b ...
V3	9	6.4781 a	7.8811 a	8.5638 a ...

Means with the same letter are not significantly different

Pairwise Mean Comparison of Management

Least Significant Difference (LSD) Test

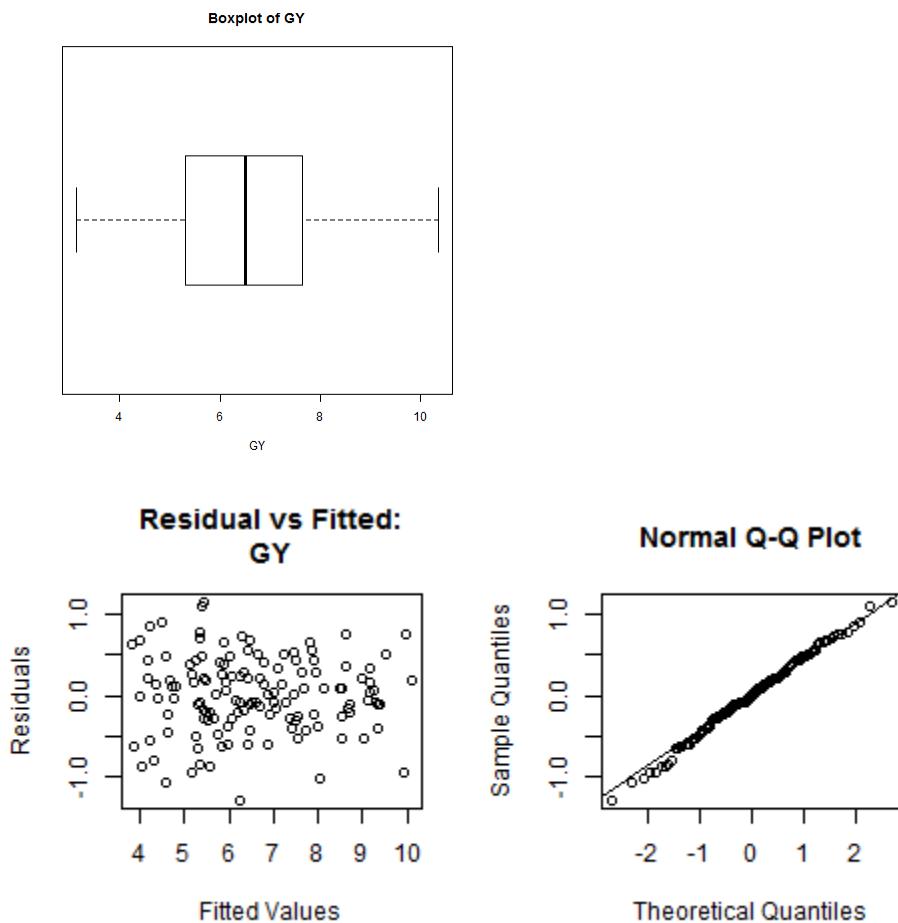
Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	0.2618
Critical Value	2.0860
Test Statistics	0.2250

Summary of the Result:

Management	means	N group
M1	5.90	45 c
M2	6.49	45 b
M3	7.28	45 a

Means with the same letter are not significantly different.

Sample graphical output of the analysis displayed in the Graph Viewer tab is shown below:



Modifications can still be made on the analysis submitted. To know more about the modification that can be made, see the discussion under *Analysis of Variance > Completely Randomized Design > Additional Options* of this user's manual.

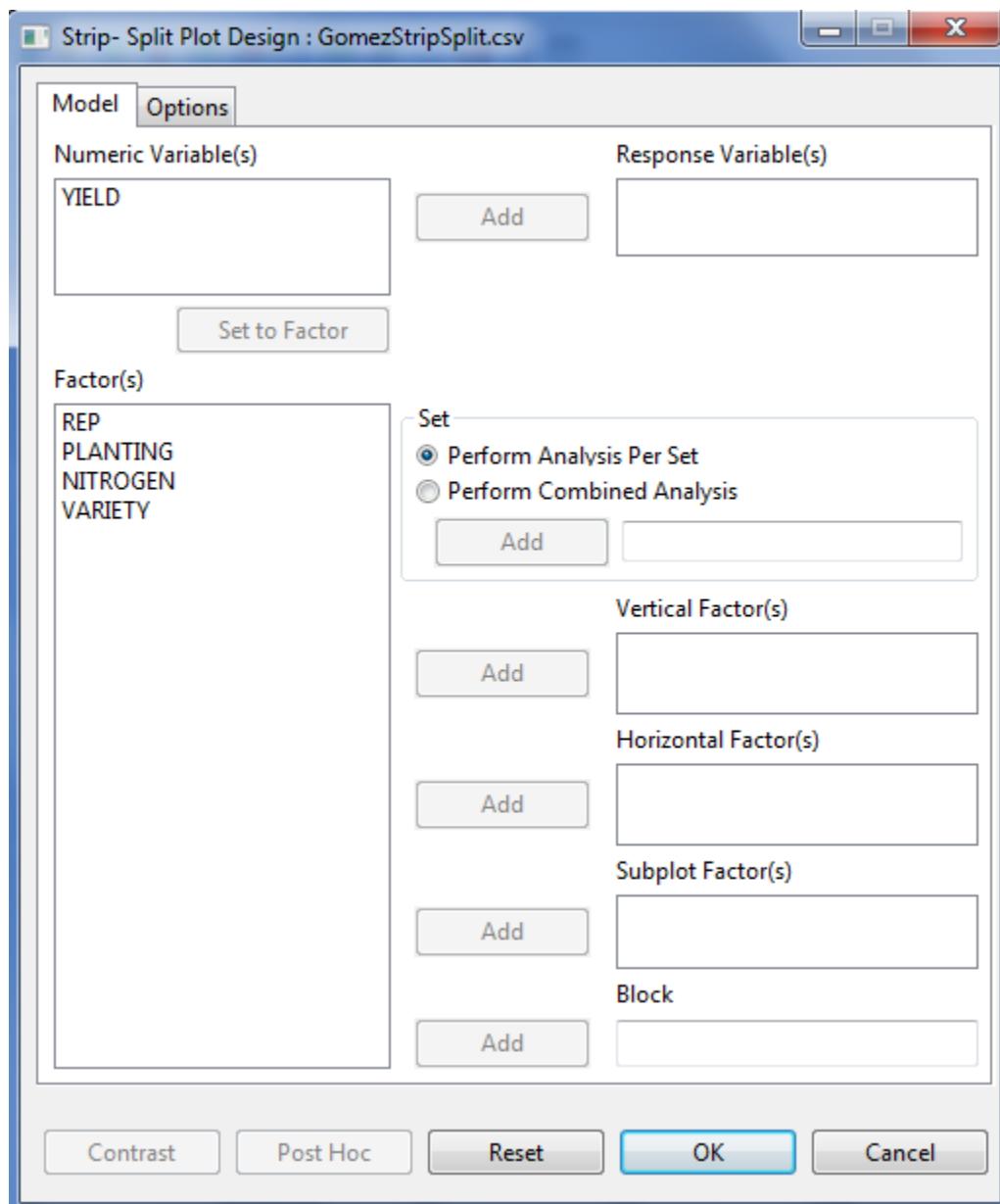
Strip-Split Plot Design

The steps to perform the Analysis of Variance in Strip-Split Plot Design are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *GomezStripSplit* from the package. To import the data from the package, see *Importing from Package* section of this user's manual.

- From the main window of STAR click **Analyze | Analysis of Variance | Strip-Split Plot Design**. The **Strip-Split Plot Design** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only Items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default option) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and the **Set** variable is not specified, STAR will perform the analysis as if the entire data came from only one site. If **perform combined analysis is selected**, a Set variable must be specified.

Vertical Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Horizontal Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Subplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

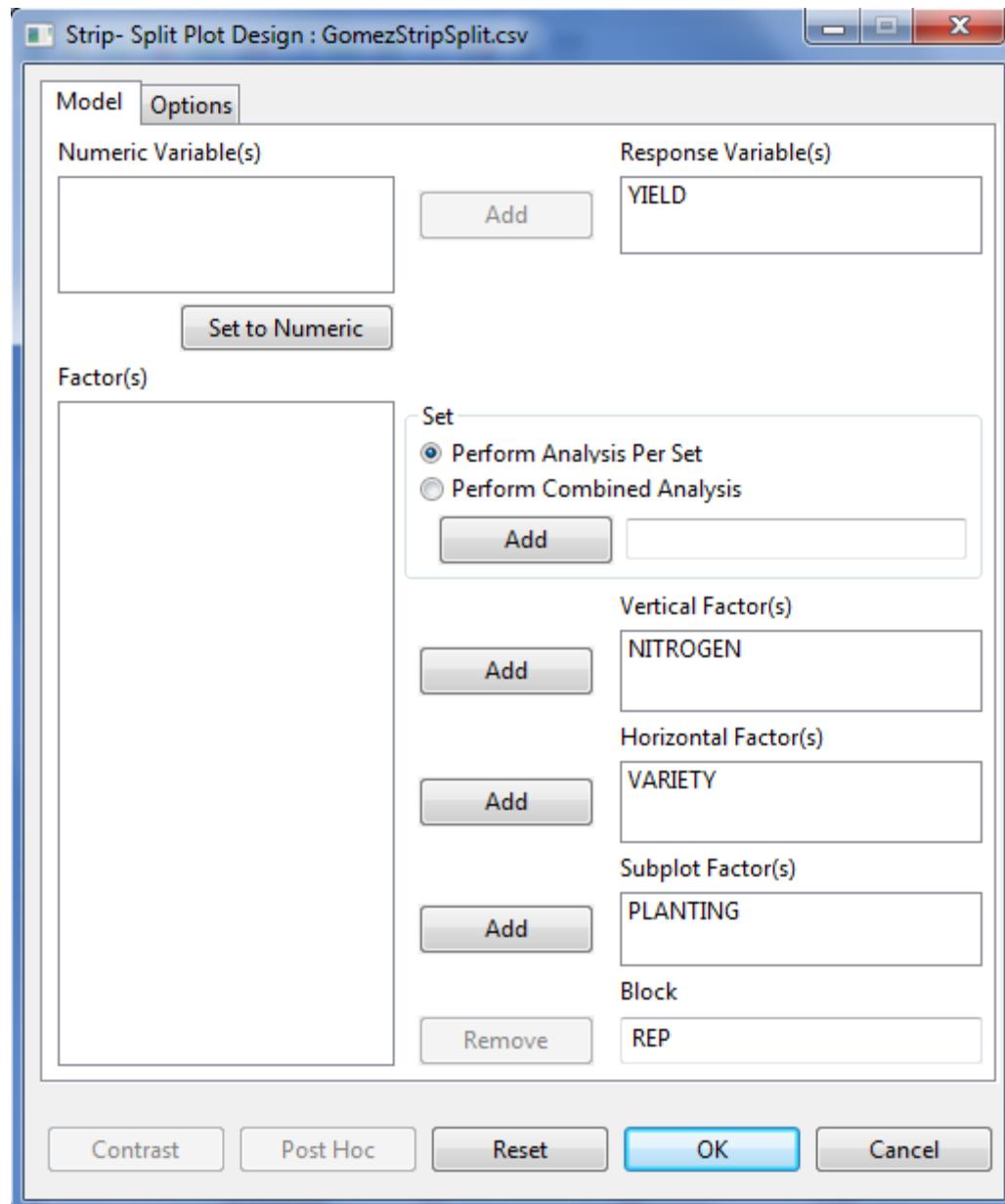
Block

Only one item can be added in this field and should come from the Factor(s) list box.

Options Tab

This tab lets the user identify additional output to be displayed and set the level of significance to be used. To know more about the options in detail, see the discussion under *Analysis of Variance > Completely Randomized Design > Options tab* of this user's manual.

For the example, the completed Model tab is shown below:



- Click **OK** button to perform the analysis. The **Strip-Split Plot Design** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample text output is shown below:

```
Analysis of Variance
Strip-Split Plot Design
=====
ANALYSIS FOR RESPONSE VARIABLE: YIELD
=====

Summary Information
-----
FACTOR      NO. OF LEVELS   LEVELS
-----
NITROGEN    3              0, 60, 120
VARIETY     6              V1, V2, ..., V6
PLANTING    2              Broadcast, Transplanted
REP         3              I, II, III
-----
Number of Observations Read and Used: 108

Descriptive Statistics
-----
Variable   N_Obs   Min     Max      Mean     StdDev
-----
YIELD      108    1556   9838   5371.77   1694.76
-----

Test for Homogeneity of Variances
-----
Method      DF   Chisq Value  Pr(>Chisq)
-----
Bartlett    35    57.79    0.0090
-----

Test for Normality
-----
Variable      Method      W Value  Pr(< W)
-----
YIELD_resid  Shapiro-Wilk  0.9938  0.9089
-----

ANOVA TABLE
Response Variable: YIELD
-----
Source          DF   Sum of Square   Mean Square   F Value  Pr(> F)
-----
REP             2   15289498.1296   7644749.0648   4.81    0.0863
NITROGEN        2   116489166.1296   58244583.0648   36.62   0.0027
Error(a)        4   6361491.0370   1590372.7593
VARIETY         5   49119269.6019   9823853.9204   3.68    0.0379
Error(b)        10  26721827.9815   2672182.7981
NITROGEN:VARIETY 10  24595730.6481   2459573.0648   2.57    0.0344
Error(c)        20  19106733.1852   955336.6593
PLANTING        1   723079.3426   723079.3426   1.71    0.1986
NITROGEN:PLANTING 2   2468131.9074   1234065.9537   2.93    0.0664
VARIETY:PLANTING 5   23761441.3796   4752288.2759   11.27   0.0000
NITROGEN:VARIETY:PLANTING 10  7512072.2037   751207.2204   1.78    0.1000
Error(d)        36  15179353.6667   421648.7130
```

Total	107	307327795.2130
-------	-----	----------------

Summary Statistics

CV(a) %	CV(b) %	CV(c) %	CV(d) %	YIELD Mean
23.48	30.43	18.20	12.09	5371.77

Table of Means

PLANTING	NITROGEN	VARIETY	V1	V2	V3	V4	V5	V6
			3571.667	4934.333	4249.667	4059.000	4101.667	3207.333
Broadcast	0		5132.000	6713.667	6122.333	5553.667	5633.333	3714.333
	60		7548.000	7211.333	7868.333	7094.333	6012.000	2492.000
	120		2786.333	4501.000	4340.333	5170.667	3730.000	4509.667
Transplanted	0		4983.000	5509.667	6233.000	5545.000	4772.667	4620.667
	60		6924.333	6608.000	7716.333	7878.667	6015.000	6320.333

Comparison of NITROGEN at each level of VARIETY

Least Significant Difference (LSD) Test

Alpha	0.05
Error Degrees of Freedom	22
Error Mean Square	1061176.0093
Critical Value	2.0693
Test Statistic	1230.7388

Summary:

NITROGEN	N	VARIETY = V1 group	VARIETY = V2 group
0	6	3179.0000 c	4717.6667 b
60	6	5057.5000 b	6111.6667 a
120	6	7236.1667 a	6909.6667 a

Means with the same letter are not significantly different

Comparison of VARIETY at each level of NITROGEN

Tukey's Honest Significant Difference (HSD) Test

Alpha	0.05
Error Degrees of Freedom	10
Error Mean Square	2672182.7981
Critical Value	4.9120
Test Statistic	3278.0623

Summary:

VARIETY	N	NITROGEN = 0 group	NITROGEN = 60 group	NITROGEN = 120 group
V1	6	3179.0000 a	5057.5000 a	7236.1667 ab
V2	6	4717.6667 a	6111.6667 a	6909.6667 ab
V3	6	4295.0000 a	6177.6667 a	7792.3333 a
V4	6	4614.8333 a	5549.3333 a	7486.5000 ab

V5	6	3915.8333 a	5203.0000 a	6013.5000 ab
V6	6	3858.5000 a	4167.5000 a	4406.1667 b

Means with the same letter are not significantly different

Comparison of VARIETY at each level of PLANTING

Tukeys's Honest Significant Difference (HSD) Test

Alpha	0.05
Error Degrees of Freedom	13
Error Mean Square	1546915.7556
Critical Value	4.6728
Test Statistic	1937.2651

Summary:

VARIETY	N	PLANTING = Broadcast group	PLANTING = Transplanted group
V1	9	5417.2222 a	4897.8889 a
V2	9	6286.4444 a	5539.5556 a
V3	9	6080.1111 a	6096.5556 a
V4	9	5569.0000 a	6198.1111 a
V5	9	5249.0000 a	4839.2222 a
V6	9	3137.8889 b	5150.2222 a

Means with the same letter are not significantly different

Comparison of PLANTING at each level of VARIETY

Least Significant Difference (LSD) Test

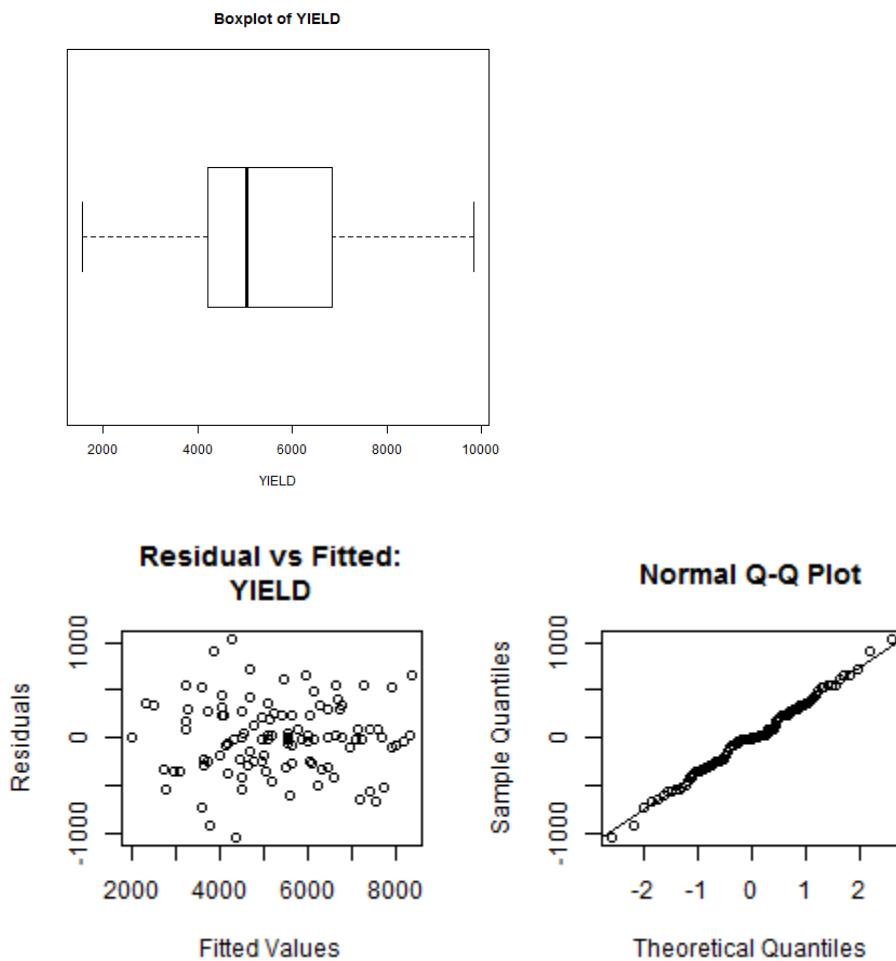
Alpha	0.05
Error Degrees of Freedom	36
Error Mean Square	421648.7130
Critical Value	2.0281
Test Statistic	620.8079

Summary:

PLANTING	N	VARIETY = V1 group	VARIETY = V2 group
Broadcast	9	5417.2222 a	6286.4444 a
Transplanted	9	4897.8889 a	5539.5556 b

Means with the same letter are not significantly different

Sample graphical output displayed in the Graph Viewer page is shown below:



Modifications can still be made on the analysis submitted. To know more about the modification that can be made, see the discussion under *Analysis of Variance > Completely Randomized Design > Additional Options* of this user's manual.

Split-Split-Split Plot Design

This analysis can be performed for experiments where the main plot is arranged in completely randomized, randomized complete block or Latin square design.

The steps to perform the Analysis of Variance using Split-Split-Split Plot Design are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.
- From the main window of STAR click **Analyze | Analysis of Variance | Split-Split-Split Plot Design**. Choose the appropriate design for the mainplot. The **Split-Split-Split Plot Design** dialog box will appear.
- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default option) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and the **Set** variable is not specified, STAR will perform the analysis as if the entire data came from only one site. If **perform combined analysis is selected**, a Set variable must be specified.

Mainplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Subplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Sub-subplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Sub-sub-subPlot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Replicate

Applicable only if completely randomized design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Block

Applicable only if randomized complete block design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Row Block

Applicable only if Latin square design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Column Block

Applicable only if Latin square design is chosen. Only one item can be added in this field and should come from the **Factor(s)** list box.

Options Tab

This tab lets the user identify additional output to be displayed and set the level of significance to be used. To know more about the options in detail, see the discussion under *Analysis of Variance > Completely Randomized Design > Options tab* of this user's manual.

- Click **OK** button to perform the analysis. The **Split-Split-Split Plot Design (RCBD)** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Modifications can still be made on the analysis submitted. To know more about the modification that can be made, see the discussion under *Analysis of Variance > Completely Randomized Design > Additional Options* of this user's manual.

Strip-Split-Split Plot Design

The steps to perform the Analysis of Variance in Strip-Split-Split Plot Design are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

- From the main window of STAR click **Analyze | Analysis of Variance | Strip-Split-Split Plot Design**. The **Strip-Split-Split Plot** dialog box will appear.
- Specify the required field and appropriate options for the analysis.

Model Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only Items from the **Numeric Variable(s)** can be added in this list box.

Set

There are two options available, *perform analysis per set* (default option) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and the **Set** variable is not specified, STAR will perform the analysis as if the entire data came from only one site. If **perform combined analysis is selected**, a Set variable must be specified.

Vertical Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Horizontal Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Subplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Sub-subplot Factor(s)

At least one item should be added in this list box and should come from the **Factor(s)** list box.

Block

Only one item can be added in this field and should come from the Factor(s) list box.

Options Tab

This tab lets the user identify additional output to be displayed and set the level of significance to be used. To know more about the options in detail, see the discussion under *Analysis of Variance > Completely Randomized Design > Options tab* of this user's manual.

- Click **OK** button to perform the analysis. The **Strip-Split-Split Plot Design** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Modifications can still be made on the analysis submitted. To know more about the modification that can be made, see the discussion under *Analysis of Variance > Completely Randomized Design > Additional Options* of this user's manual.

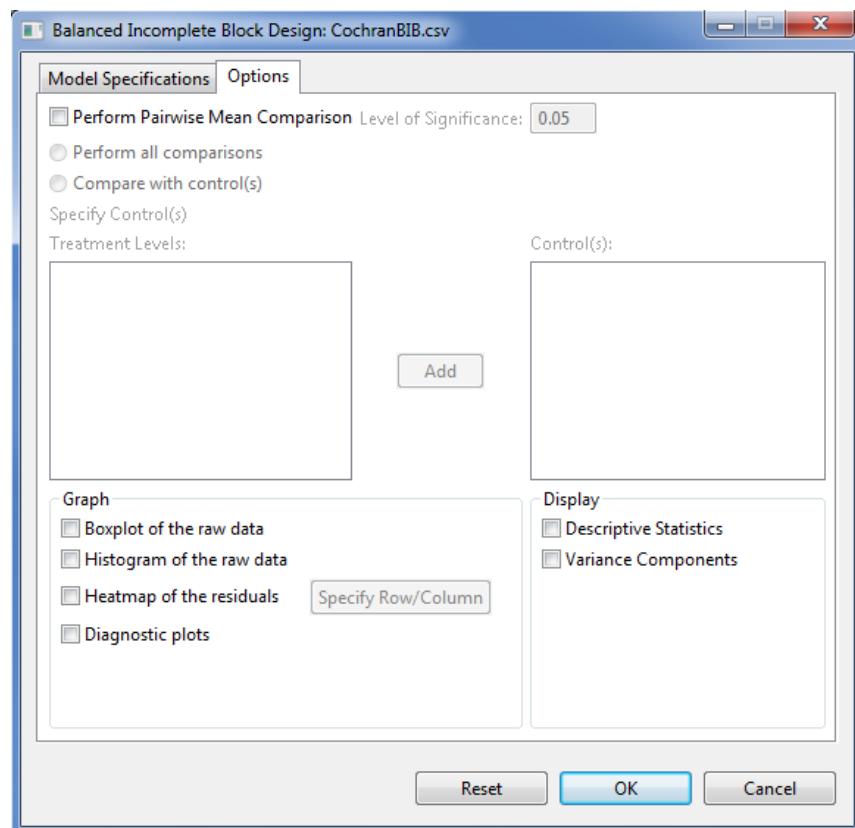
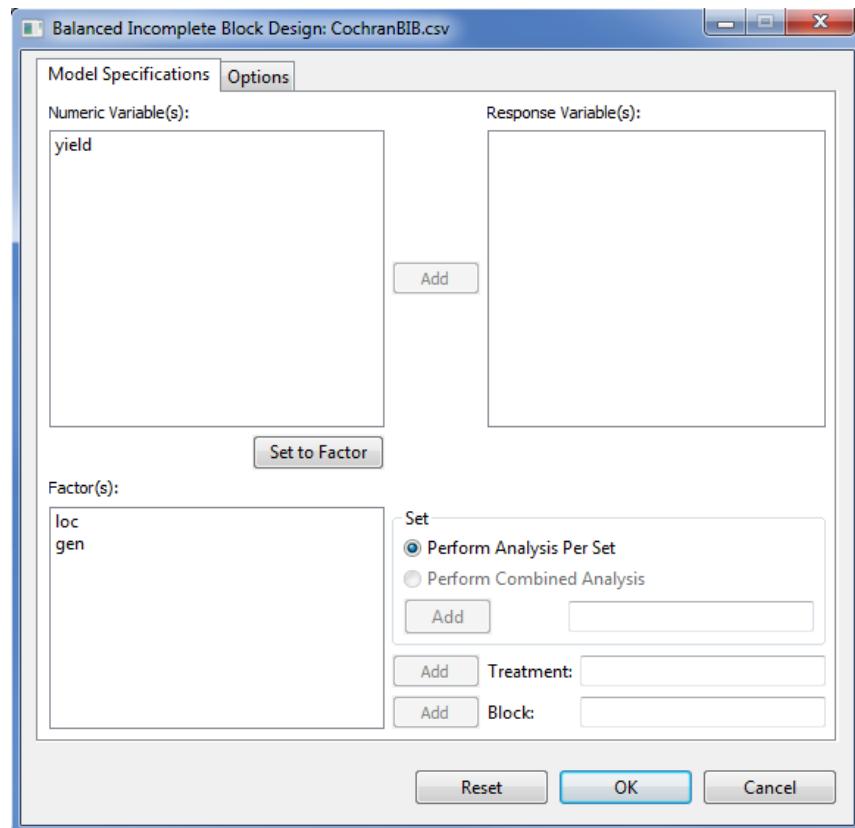
Balanced Incomplete Block Design

The steps to perform analysis for balanced incomplete block design are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *CochranBIB* from the package. To import the data from the package, see *Importing from Package* section of this user's manual.

- Choose **Analyze | Incomplete Block | Balanced Incomplete Block Design...** from the main window. The **Balanced Incomplete Block Design** dialog box will appear.



- Specify the required fields and appropriate options for the analysis.

Model Specifications Tab

Response Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. The items to be entered should come from the **Numeric Variable(s)** list box.

Set

There are two options available, *perform analysis per set* (default) and *perform combined analysis*. If default option is selected and a **Set** variable is specified, STAR will perform analysis for each level of the **Set** variable. If the default option is selected and no **Set** variable is specified, STAR will perform the analysis as if the entire data came from only one set variable. If *perform combined analysis* is selected, a **Set** variable with at least two levels must be specified. Only one item can be added in this entry box and should come from the **Factor(s)** list box.

Treatment

This field is required. Only one item can be added in this entry box and should come from the **Factor(s)** list box.

Block

This field is required. Only one item can be added in this entry box and should come from the **Factor(s)** list box.

Options Tab

Perform pairwise mean comparisons

If this option is selected, user can either choose *Perform all comparisons* or *Compare with control(s)*. The option *Perform all comparisons* is not recommended when the number of treatment levels is very large. For now, STAR executes this option if the treatment levels is at most fifteen. If the treatment has more than 15 levels, user can only choose the option *Compare with control(s)*. If the option *Compare with control(s)* was selected, the user should specify the control levels to be compared with the rest of the treatment levels. This can be done by selecting the level(s) of the treatment that are considered as the control(s) then click the **Add** button.

Level of Significance

This will be enabled if *Perform pairwise mean comparison* is selected. This will be the level of significance to be used when pairwise mean comparison is performed. Its default value is 0.05.

Display Descriptive Statistics

If selected, a summary table with number of missing observations, mean and standard deviation will be displayed.

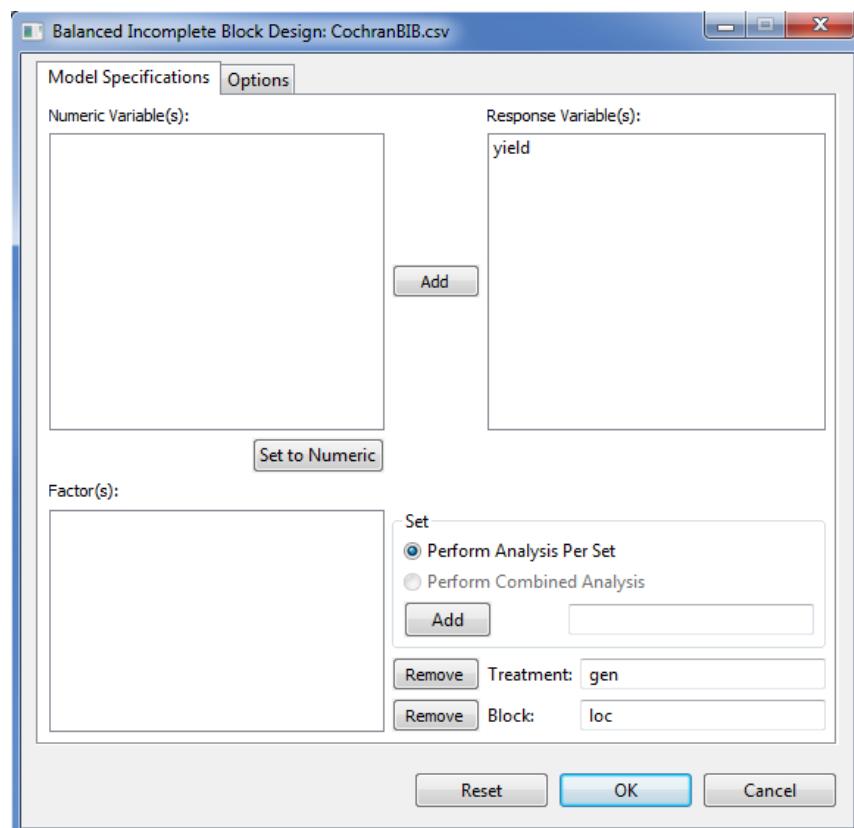
Display Variance Components

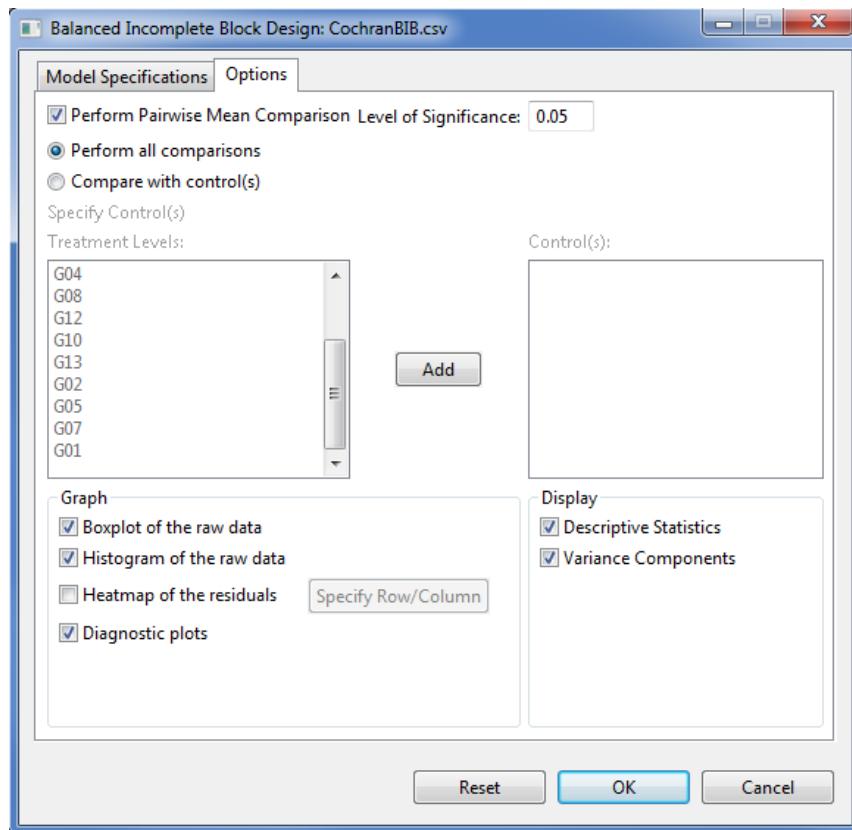
If selected, a table with the variance estimates of the some model components will be displayed.

Graph

Graphs like boxplot, histogram, heatmap and diagnostic plots can be generated as part of the output of the analysis. If heatmap is selected, user needs to specify a row and a column variable.

For the example, the completed **Model Specification** tab and **Options** tab should appear as shown:





- Click **OK** button to perform the analysis. The **Augmented RCB** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

Sample text output of the analysis is shown below:

```

DESIGN: Balanced Incomplete Block Design (BIBD)
=====
ANALYSIS FOR RESPONSE VARIABLE: yield
=====

Descriptive Statistics
-----
Variable   N_NonMissObs   Mean   StdDev
-----
yield          52    29.78     5.52
-----

Summary Information
-----
FACTOR   NO. OF LEVELS   LEVELS
-----
loc      13             B01, B02, ..., B13
gen      13             G01, G02, ..., G13
-----
Number of Observations Read and Used: 52

```

```
Estimates of Variance Components
-----
Groups      Variance   StdDev
-----
loc          6.0527    2.4602
Residual    19.9340    4.4647
-----
```

Testing for the Significance of Genotypic Effect

Models:

```
model2: yield ~ 1 + (1|loc)
model1: yield ~ 1 + gen + (1|loc)
```

```
-----
Source     Df       AIC       BIC      logLik  Chisq Chi Df Pr(>Chisq)
-----
model2      3  326.1376  331.9913  -160.0688
model1     15  328.4084  357.6770  -149.2042  21.73      12      0.0407
-----
```

```
-----
Source     Df  Sum Sq  Mean Sq  F value    Prob
-----
gen        12  400.63   33.39     1.67  0.0407
-----
```

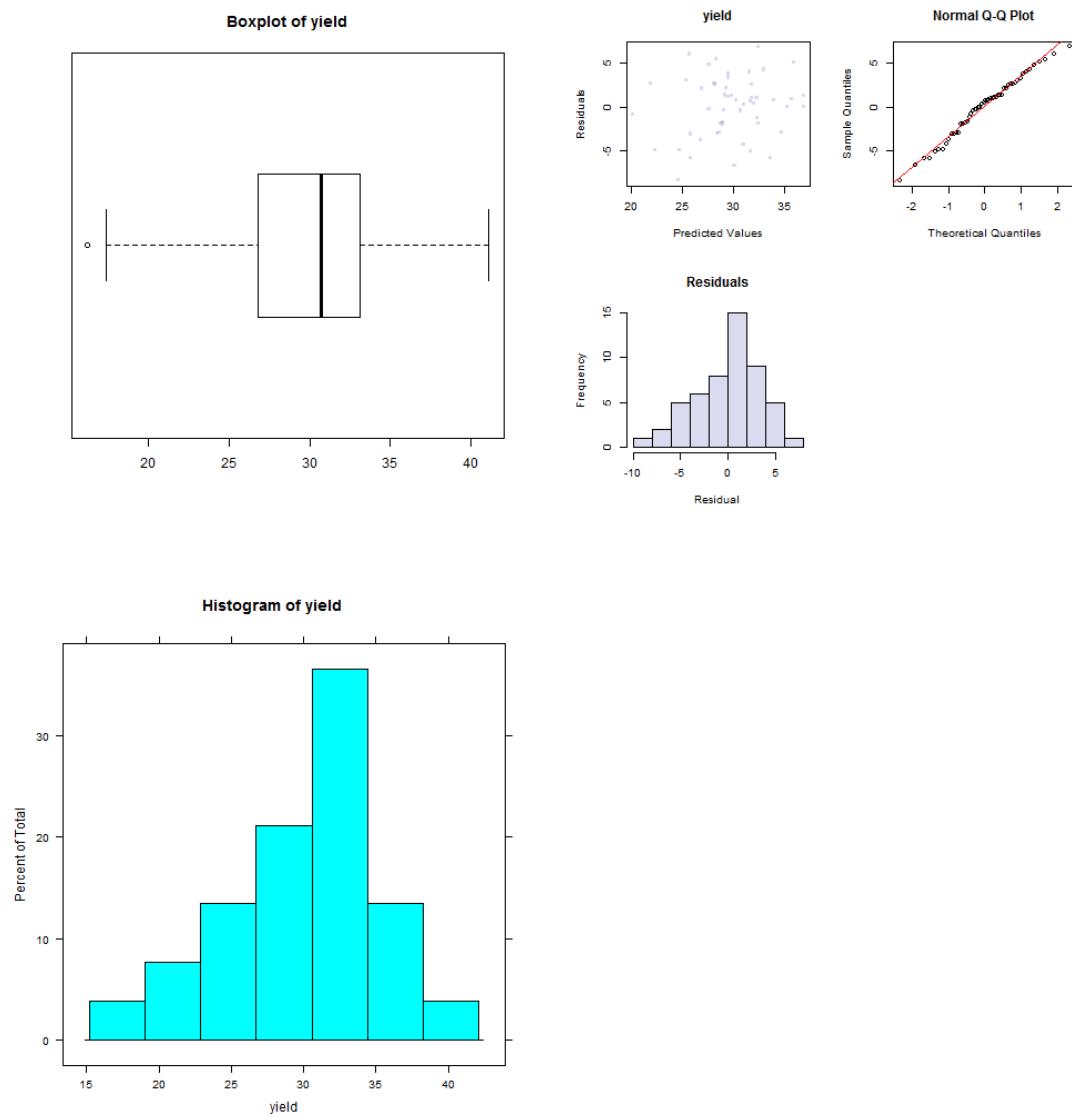
gen LSMeans and Standard Error

```
-----
gen      LSMean  StdErrMean
-----
G01      34.17    2.44
G02      29.04    2.44
G03      30.11    2.44
G04      28.08    2.44
G05      30.34    2.44
G06      27.59    2.44
G07      30.76    2.44
G08      32.75    2.44
G09      28.56    2.44
G10      28.10    2.44
G11      23.47    2.44
G12      28.99    2.44
G13      35.18    2.44
-----
```

List of gen pairs that are significantly different using Tukey at alpha = 0.05

```
-----
gen[i]  gen[j]  Estimate    lwr    upr
-----
G13     G11      11.71  0.6738  22.74
-----
```

Sample graphical output of the analysis is shown below:



Pairwise Mean Comparison

The **Post Hoc** option is used to request for other pairwise comparison procedures. This button will be enabled if there is at least one effect that is significant.

The steps to request for other pairwise mean comparison are listed below:

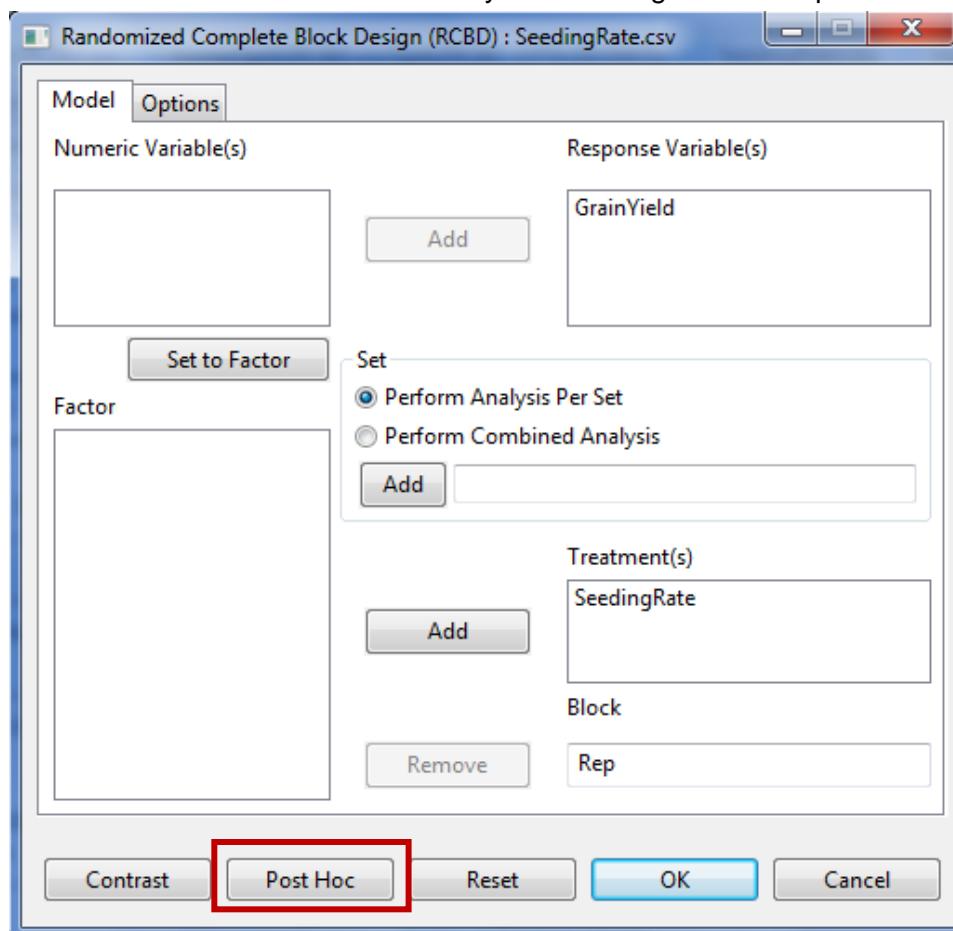
- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *SeedingRate* from STAR package. To import the data from the package, see *Importing from Package* section of this user's manual.

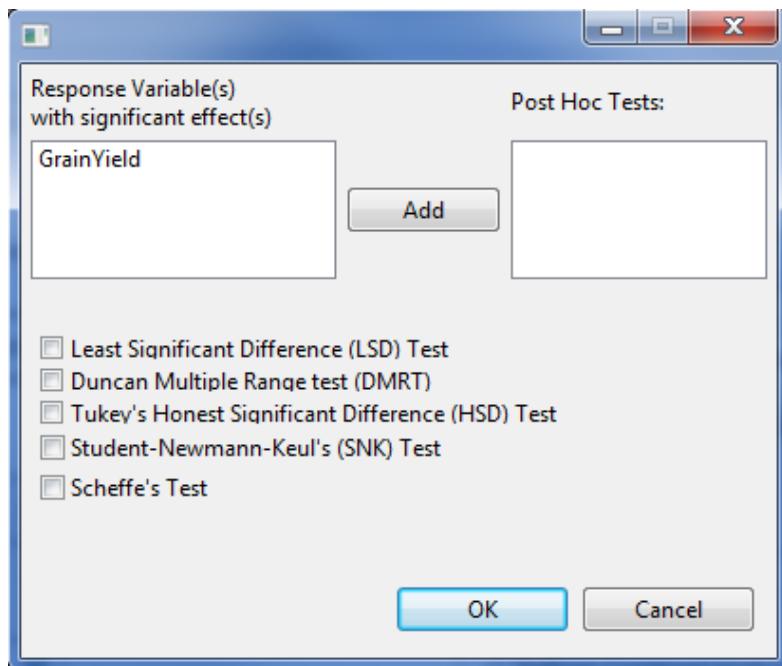
- Perform the Analysis of Variance.

For the example, follow the discussion in **Randomized Complete Block Design** of this module with *SeedingRate* as Factor and *Rep* as Block.

- After performing the Analysis of Variance, the **Post hoc** button will be enabled if there is at least one effect in the analysis that is significant at specified level.



- Click the **Post Hoc** button. A dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Response Variable(s)

This list box displays all response variables with at least one significant effect. User need to select one item from this list and transfer to the **Post Hoc Test** list box to proceed with the analysis

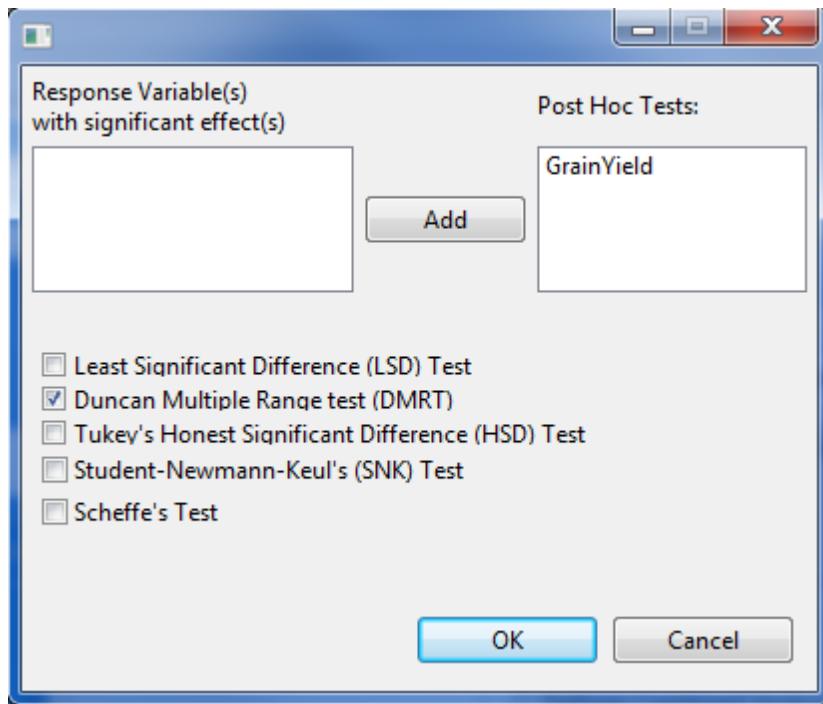
Post Hoc Tests

This list box should contain at least one item that will come from the **Response Variable(s)** list box. STAR will perform the additional pairwise comparison procedure for all effects of the response variable(s) selected.

Pairwise Comparison Procedure

At least one of the checkbox should be ticked for the analysis to proceed.

For the example, the completed dialog should appear as illustrated below:



- Click the **OK** button to perform the pairwise mean comparison. The **Pairwise Comparison** dialog box will be closed. The *PairwiseOutput.txt* is created in the parent folder within the *Output* folder and STAR activates the **Result Viewer** tab.

Sample content of the *PairwiseOutput.txt* displayed in the **Result Viewer** tab is shown below:

Response Variable: GrainYield

Pairwise Mean Comparison of SeedingRate

Duncan's Multiple Range Test (DMRT)

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.1209

Number of Means	2	3	4	5	6
-----------------	---	---	---	---	---

Tabular Value	3.01	3.16	3.25	3.31	3.36
Test Statistics	0.52	0.55	0.57	0.58	0.58

Summary of the Result:

SeedingRate	means	N	group
100	4.35	4	c
125	4.67	4	bc
150	3.38	4	d
25	5.12	4	ab
50	5.08	4	ab
75	5.30	4	a

Means with the same letter are not significantly different.

- After performing the pairwise comparison test, user can click the **OK** button to embed the requested pairwise mean comparison procedure with the Analysis of Variance output.

Sample content of the text file when user clicks the **OK** button after performing the pairwise comparison.

Analysis of Variance
Randomized Complete Block Design

=====
ANALYSIS FOR RESPONSE VARIABLE: GrainYield
=====

Summary Information

FACTOR NO. OF LEVELS LEVELS

SeedingRate 6 25, 50, ..., 150
Rep 4 1, 2, 3, 4

Number of Observations Read and Used: 24

ANOVA TABLE

Response Variable: GrainYield

Source DF Sum of Square Mean Square F Value Pr > F

Rep 3 1.7667 0.5889 4.87 0.0147
SeedingRate 5 10.1800 2.0360 16.84 0.0000
Error 15 1.8133 0.1209
Total 23 13.7600

Summary Statistics

CV (%) GrainYield Mean

7.48 4.65

Standard Errors

Effects StdErr

Rep 0.2007
SeedingRate 0.2459

Pairwise Mean Comparison of SeedingRate

Duncan's Multiple Range Test (DMRT)

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.1209

Number of Means	2	3	4	5	6
Tabular Value	3.0143	3.1598	3.2502	3.3118	3.3560
Test Statistics	0.5240	0.5493	0.5650	0.5757	0.5834

Summary of the Result:

SeedingRate	means	N	group
25	5.12	4	ab
50	5.08	4	ab
75	5.30	4	a
100	4.35	4	c
125	4.67	4	bc
150	3.38	4	d

Means with the same letter are not significantly different.

Partitioning Sum of Squares

The **Contrast** option is used in partitioning sum of squares that decompose the variability of the response(s) into various components to have a better view of the nature of variability. There are different selections in partitioning sum of squares such as, “Compare with Control”, “User Specified Contrast (Group Comparison)” and “Orthogonal Polynomial Contrast”.

Group Comparison

The steps to perform the Group Comparison are listed below:

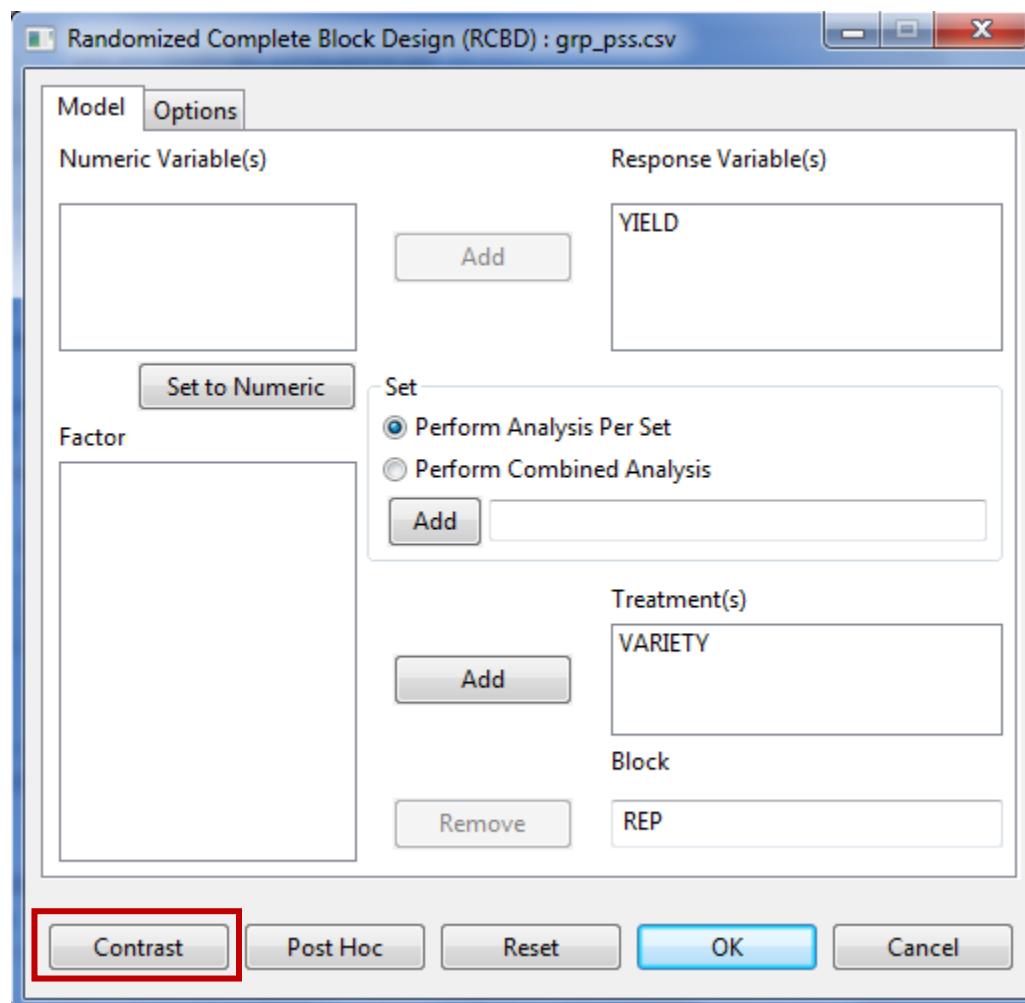
- On the **Project Explorer**, double-click the file to use for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *grp_pss* from STAR package. To import the data from the package, see *Importing from Package* section of this user's manual.

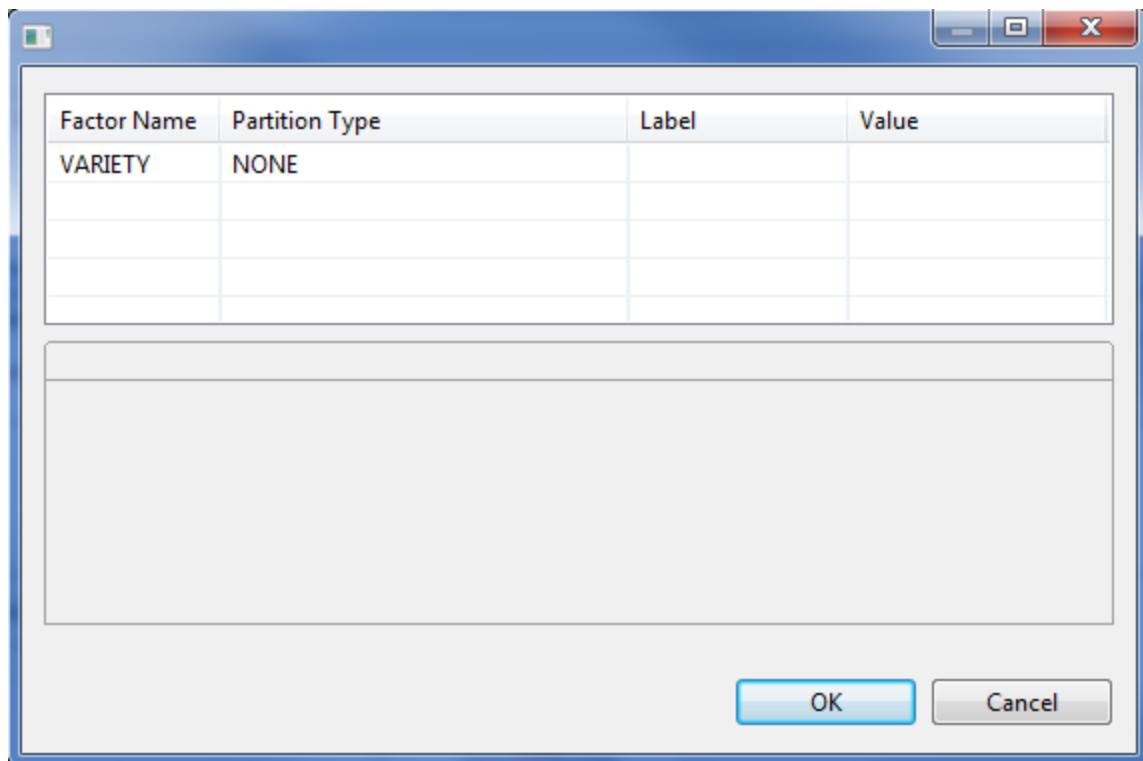
- Perform the Analysis of Variance.

For the example, follow the discussion in **Randomized Complete Block Design** of this module with *VARIETY* as Factor and *REP* as Block.

- After performing the Analysis of Variance, the **Contrast** button will be enabled if there is at least one factor in the analysis with more than two levels.



- Click the **Contrast** button. A dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Factor Name column

All factors included in the analysis with more than two levels will be displayed in this column.

Partition Type column

Specify the type of partitioning to be performed by clicking the drop-down menu. If the levels of the factor are quantitative, the 3 choices are available, otherwise; only the choices *Compare with Control* and *User-Specified Contrast* will be available for that factor.

Value column

If *Compare with Control* is chosen, this column specifies which level of the factor is considered as control level. By default, the first level of the factor is considered the control. To change the control level, click the drop-down box to display the different levels of the factor.

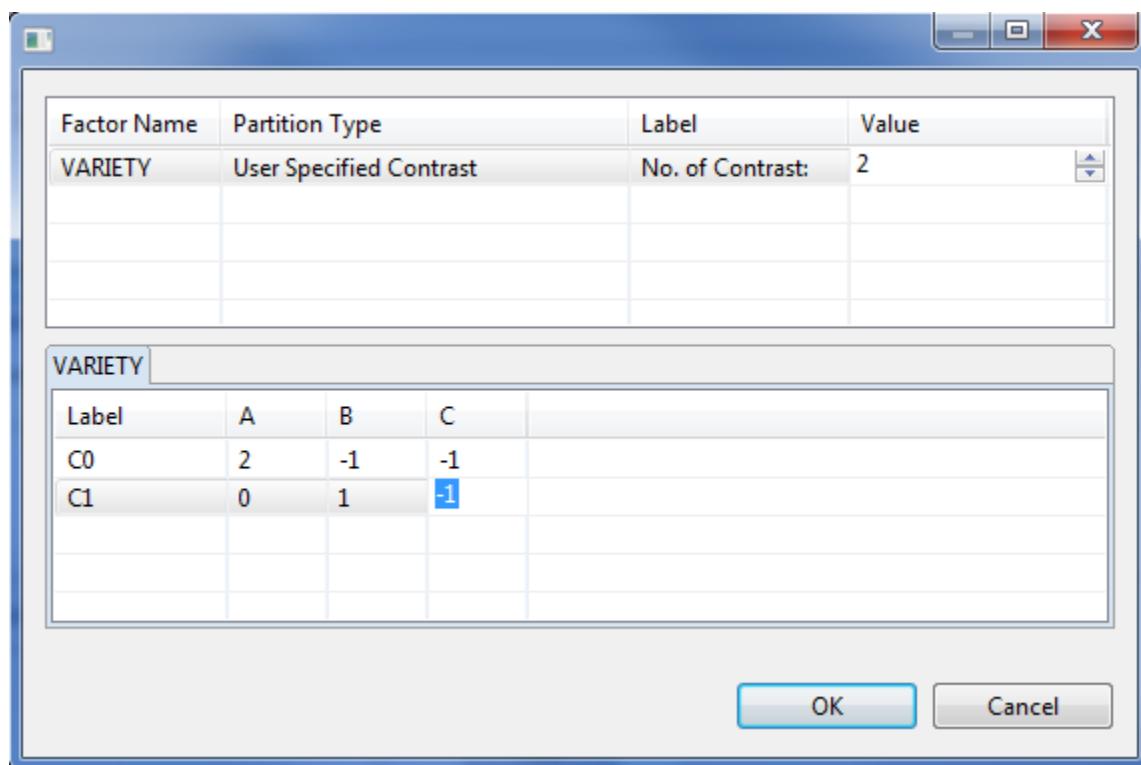
If *User Specified Contrast* or *Orthogonal Polynomial Contrast* is chosen, this column specifies the number of contrast that will be used for the analysis. The default value is 1 while the maximum number is equal to the treatment degrees of freedom.

If **User Specified Contrast** is selected, a tab with the factor name as label will be displayed at the lower part of the **Contrast Analysis** dialog box to specify the coefficients for each contrast. The sum of each contrast should be equal to zero. If there are 2 or more contrasts, the sum of the cross product of each pair wise contrast should be equal to zero.

For the example, suppose we want to determine the following:

- i. If the mean of the traditional variety (A) is significantly different from the new varieties (B and C).
- ii. If the mean of the new varieties (B and C) are significantly different.

The completed dialog box should appear as shown below:



- Click the **OK** button to perform Group Comparison. The **Contrast Analysis** dialog box will be closed. The *ContrastAnalysisOutput.txt* is created in the parent folder within the *Output* folder and STAR actives the **Result Viewer** tab.

Sample content of the *ContrastAnalysisOutput.txt* displayed in the **Result Viewer** tab is shown below:

Result of Contrast Analysis

ANOVA TABLE

Response Variable: YIELD

Source	DF	Sum of Square	Mean Square	F Value	Pr > F
<hr/>					
REP	3	0.5867	0.1956	2.20	0.1889
VARIETY	2	1.0400	0.5200	5.85	0.0390
VARIETY: C0	1	0.9600	0.9600	10.80	0.0167
VARIETY: C1	1	0.0800	0.0800	0.90	0.3794
Error	6	0.5333	0.0889		
Total	11	2.1600			

Trend Comparison

This procedure is applicable and available if the levels of the factor are quantitative. The steps to perform the Group Comparison are listed below:

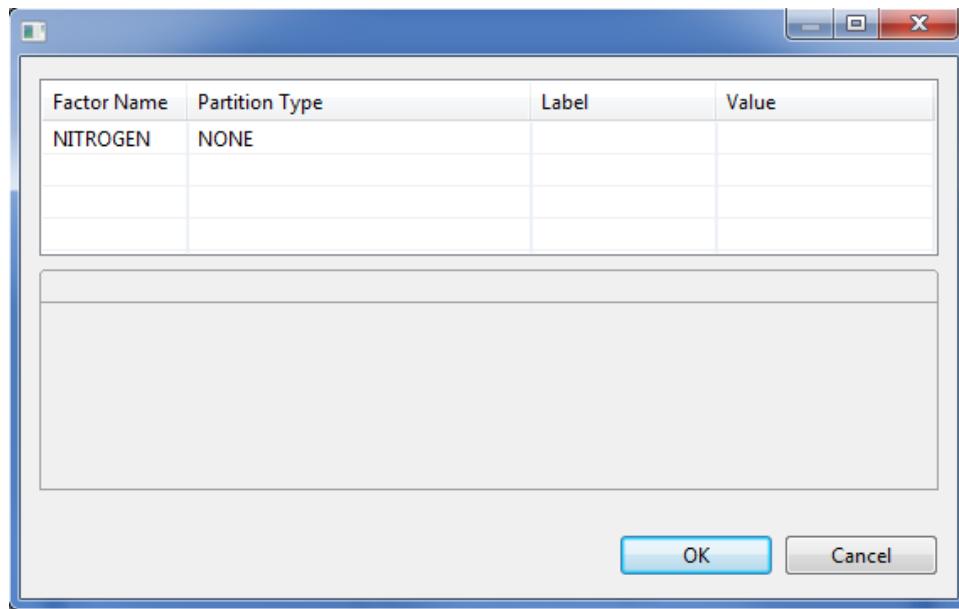
- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *OP_pss* from STAR package. To import the data from the package, see *Importing from Package* section of this user's manual.

- Perform the Analysis of Variance.

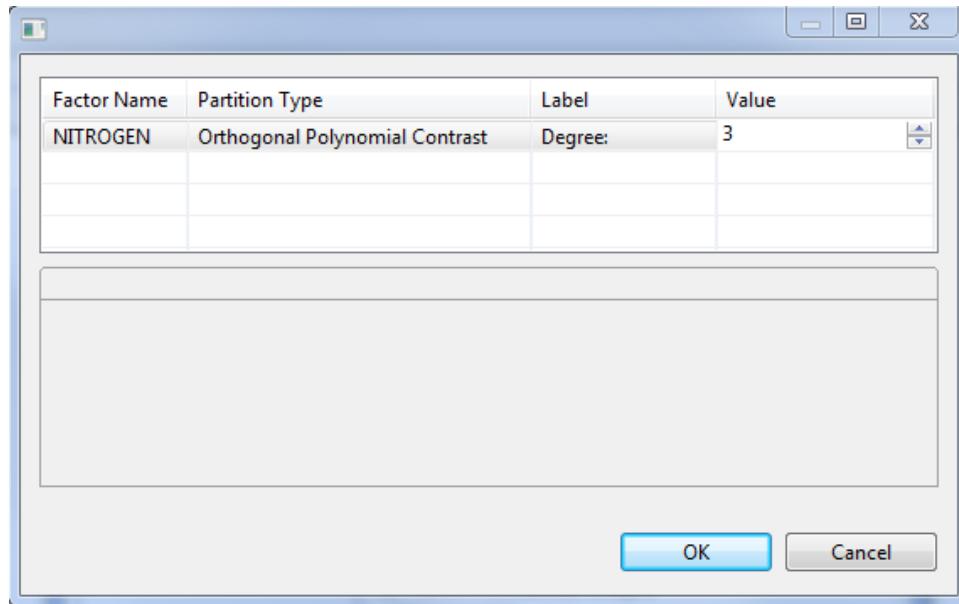
For the example, follow the discussion in **Randomized Complete Block Design** of this module with *NITROGEN* as Factor and *REP* as Block.

- After performing the Analysis of Variance, the **Contrast** button will be enabled if there is at least one factor in the analysis with more than two levels.
- Click the **Contrast** button. A dialog box will appear.



- Specify the required field and appropriate options for the analysis. To know more about the options in detail, see the discussion under *Analysis of Variance > Partitioning Sum of Squares > Group Comparison* of this user's manual.

For the example, suppose we want to know the highest degree of relationship between the yield response and Nitrogen Rate. The completed dialog box should appear as illustrated below:



- Click the **OK** button to perform Group Comparison. The **Contrast Analysis** dialog box will be closed. The *ContrastAnalysisOutput.txt* is created in the parent folder within the *Output* folder and STAR actives the **Result Viewer** tab.

Sample text output of the analysis displayed in the Output Viewer tab is shown below:

ANOVA TABLE						
Response Variable: YIELD						
Source	DF	Sum of Square	Mean Square	F Value	Pr > F	
REP	2	0.4067	0.2033	0.65	0.5557	
NITROGEN	3	11.0225	3.6742	11.73	0.0064	
NITROGEN: 1	1	10.6562	10.6562	34.01	0.0011	
NITROGEN: 2	1	0.1913	0.1913	0.61	0.4643	
NITROGEN: 3	1	0.1750	0.1750	0.56	0.4831	
Error	6	1.8800	0.3133			
Total	11	13.3092				

Correlation Analysis

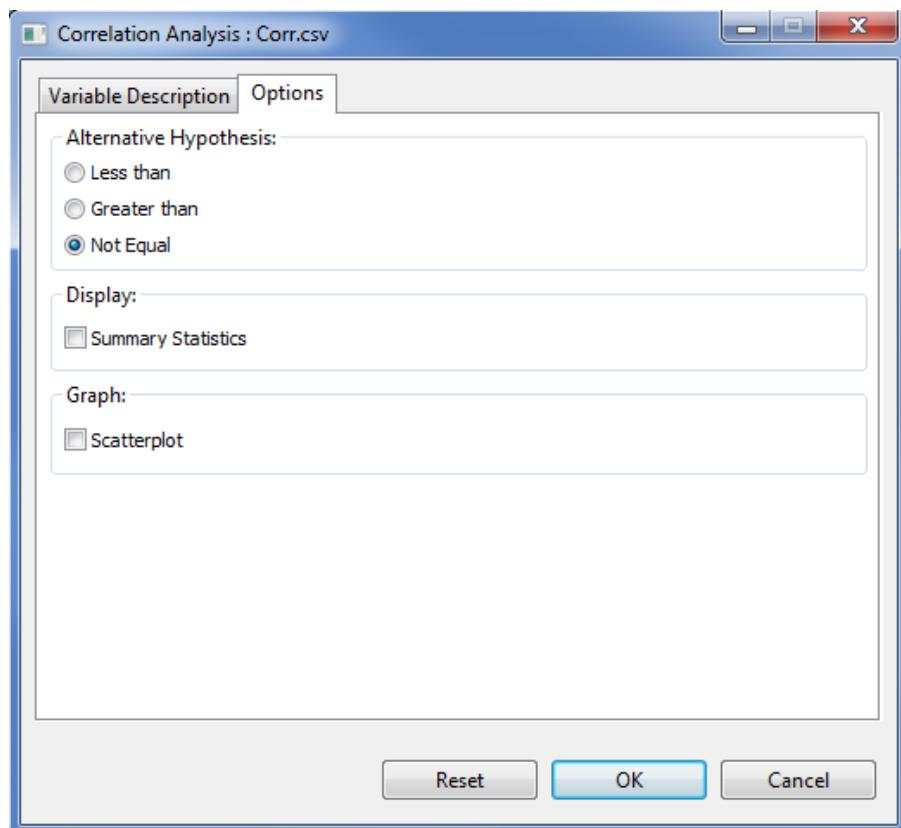
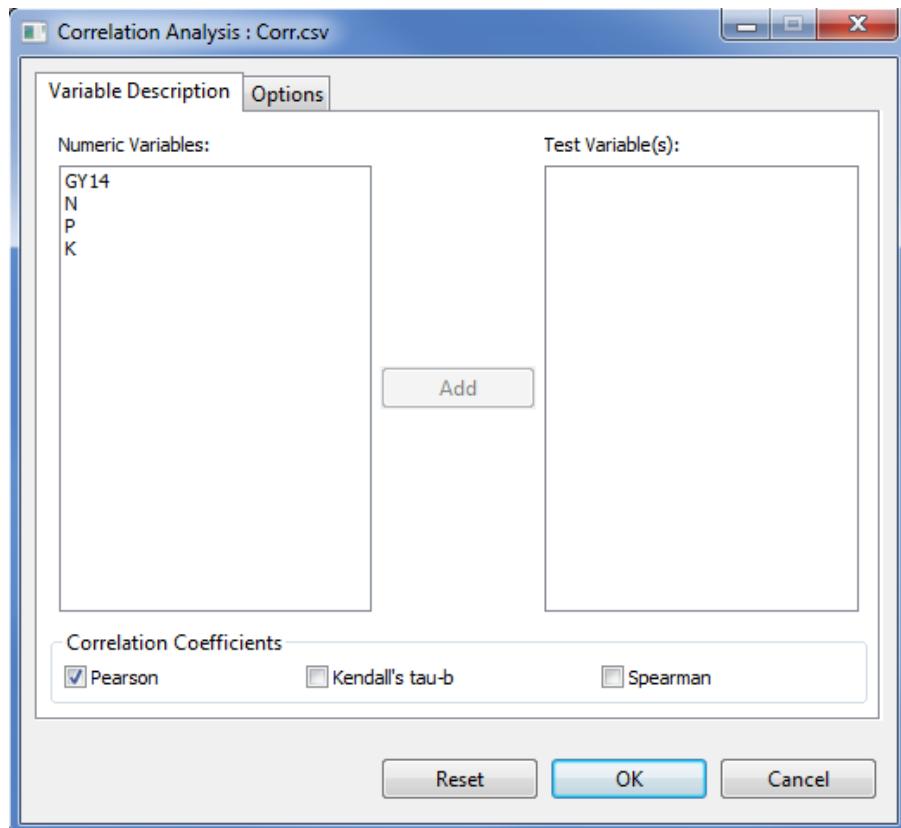
The **Correlation Analysis** is used to determine the strength of linear relationship between variables. The **Correlation Analysis** submenu performs bivariate correlation using the pearson's correlation coefficient, spearman's rho and kendall tau. By default, it uses the pearson's correlation coefficient.

The steps to perform Correlation Analysis are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis open and view it in the Data Viewer tab.

For the example, using the project named *MyProject*, import the data *Corr* from the package. To import the data from the package, see *Importing from Package* section of this user's manual.

- Click **Analyze | Correlation...** from the main window. The **Correlation Analysis** dialog box will appear.



- Specify the required field and appropriate options for the analysis:

Variable Description Tab

Test Variable(s)

For the analysis to proceed, this list box should have at least two entries.

Correlation Coefficient

For the analysis to proceed, at least one method should be selected. By default, Pearson's correlation coefficient is chosen. For quantitative and normally distributed variables, choose the Pearson correlation coefficient. If the data are not normally distributed or ordinal level, choose Kendall's tau or Spearman, which measures the association between ranks. The correlation coefficient ranges from -1 to +1.

Option Tab

Alternative Hypothesis

User can select a one-tailed (less than or greater than) or two-tailed test (not equal). The default alternative hypothesis used is that for the two-sided test.

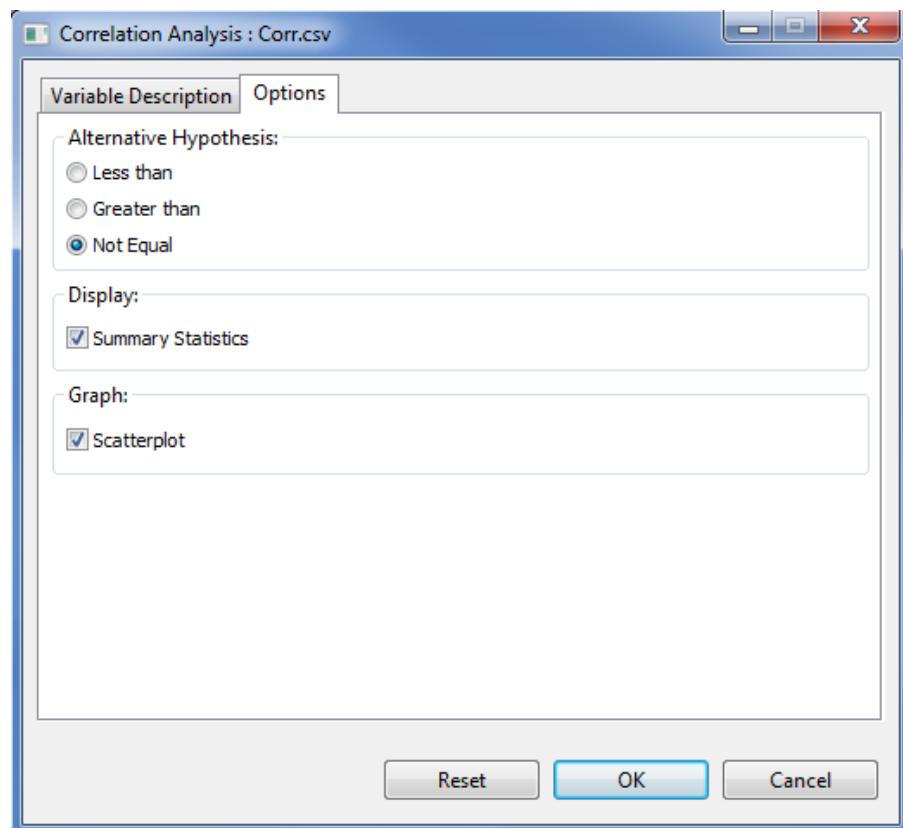
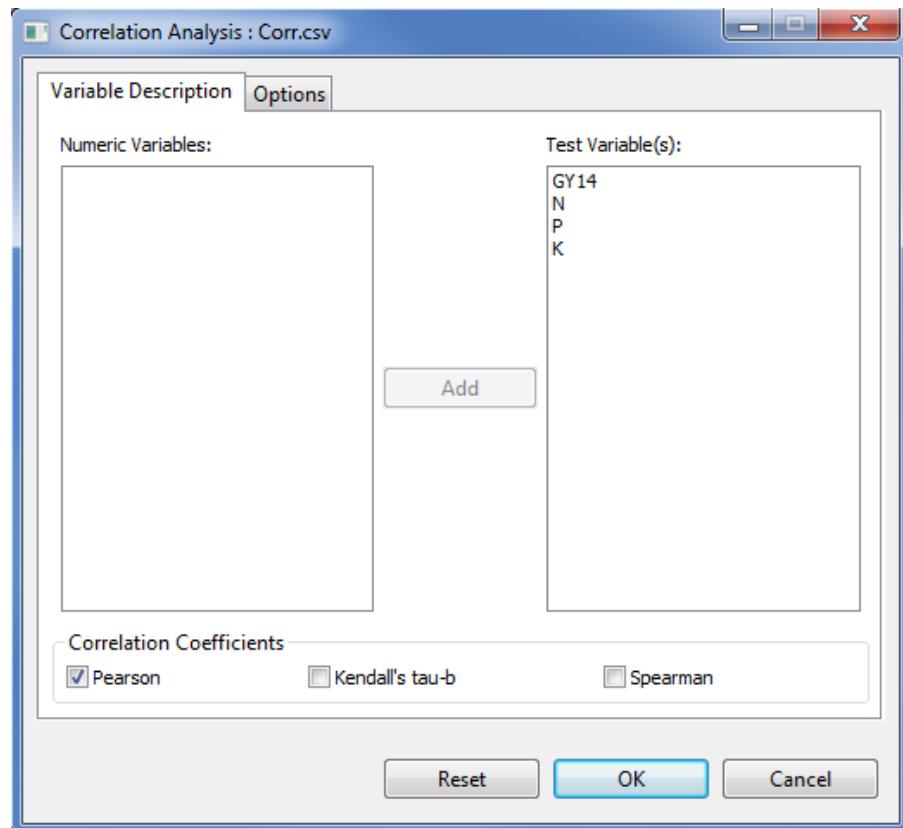
Display Summary Statistics

If this option is selected, a summary table with number of observations, minimum, maximum, mean and standard deviation will be displayed.

Scatterplot

If this option is selected, scatter plot for all pair of variables listed in the **Test Variable(s)** list box will be created.

For the example, the completed **Variable Description** tab and **Option** tab should appear as shown:



- Once all options have been specified, click the **OK** button. The **Correlation Analysis** dialog box will be minimized and STAR activates the **Result Viewer** tab which shows the results of the analysis.

Sample output of the Correlation Analysis using Pearson's product-moment correlation is shown below.

Result of Correlation Analysis

DESCRIPTIVE STATISTICS

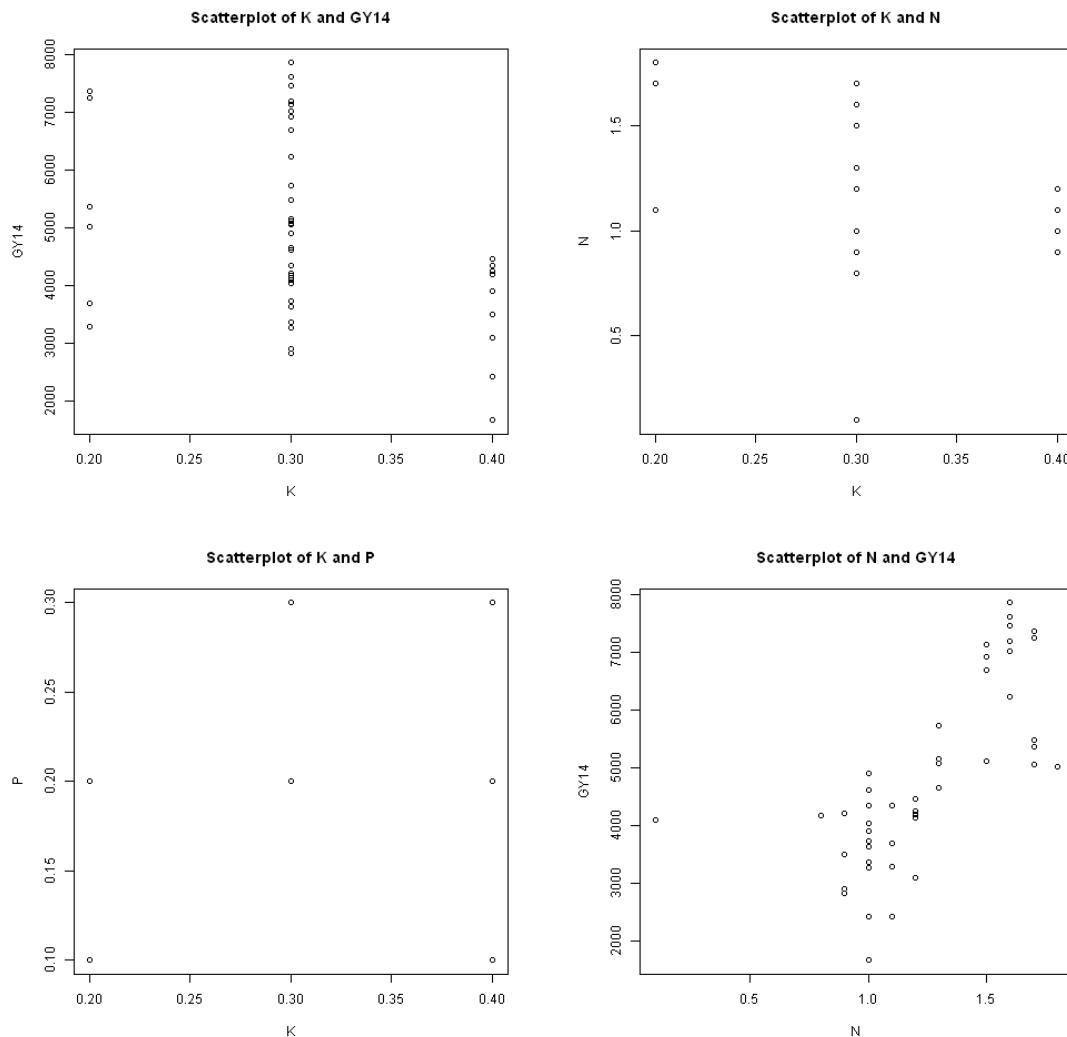
Variable	N	NonMissObs	Min	Max	Mean	StdDev
GY14	48	1678.00	7860.00	4788.67	1552.27	
N	48	0.10	1.80	1.23	0.33	
P	48	0.10	0.30	0.23	0.08	
K	48	0.20	0.40	0.31	0.06	

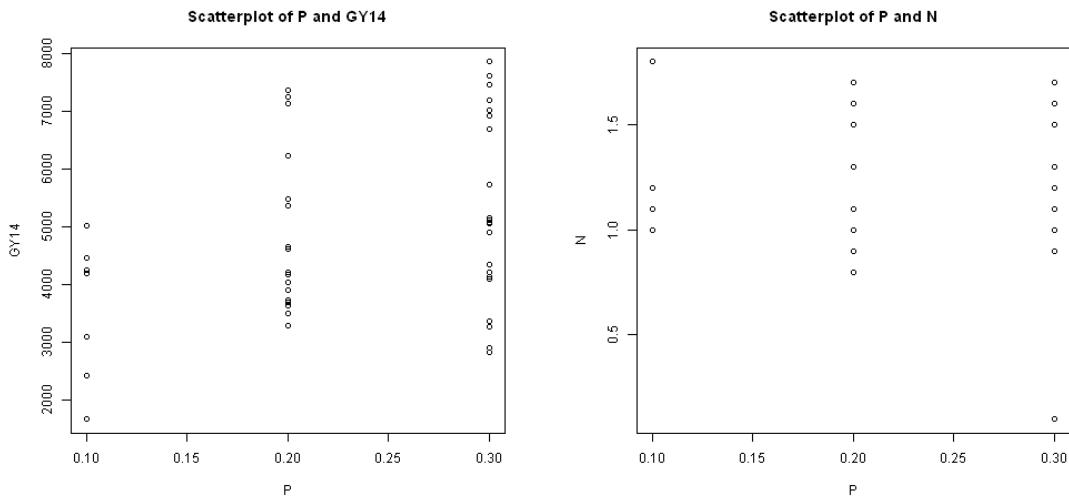
CORRELATION ANALYSIS

Pearson's product-moment correlation, Prob > |r|

		GY14	N	P	K
GY14	coef	1.00	0.72	0.38	-0.40
	p-value		0.0000	0.0080	0.0053
	n	48	48	48	48
N	coef	0.72	1.00	0.02	-0.34
	p-value	0.0000		0.8686	0.0180
	n	48	48	48	48
P	coef	0.38	0.02	1.00	-0.35
	p-value	0.0080	0.8686		0.0161
	n	48	48	48	48
K	coef	-0.40	-0.34	-0.35	1.00
	p-value	0.0053	0.0180	0.0161	
	n	48	48	48	48

If a scatter plot is requested, the graph can be viewed in the **Graph Viewer**. Sample generated graphs are shown below:





Regression Analysis

Regression Analysis is the statistical technique used to determine the strength of linear relationship between variables.

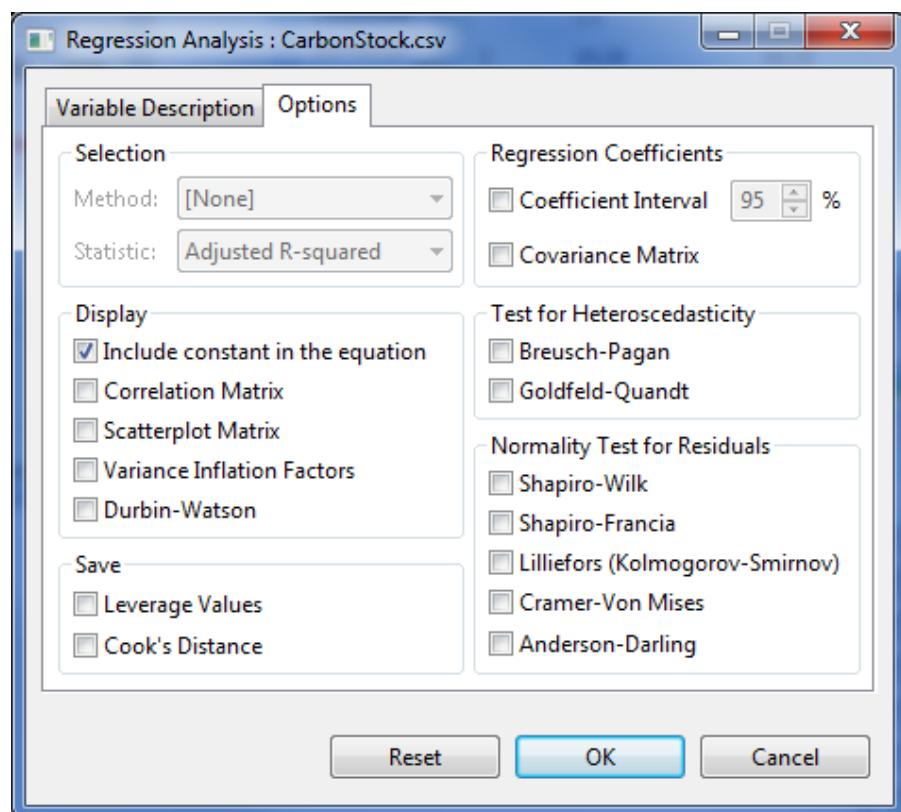
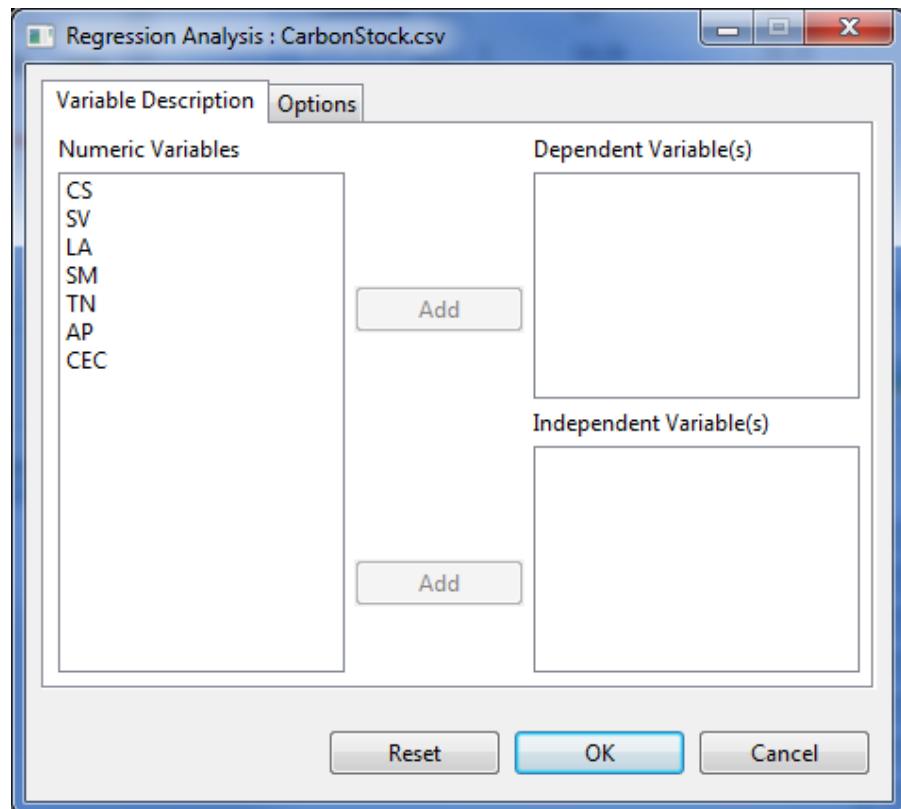
Linear Regression

The steps to perform Linear Regression Analysis are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *My Project*, import the data *CarbonStock* from the package. To import the data from the package, see *Importing from Package* section of this user's manual.

- Click **Analyze | Regression | Linear...** from the main window. The **Regression Analysis** dialog box will appear.



- Specify the required field and appropriate options for the analysis:

Variable Description Tab

Dependent Variable(s)

For the analysis to proceed, this list box should have at least one entry.

Independent Variable(s)

For the analysis to proceed, this list box should have at least one entry.

Options Tab

Selection

By default, no selection method is used for linear regression analysis in STAR. If at least two independent variables were specified, a selection method from one of the following can be selected: *All possible regression*, *Forward selection*, *Backward elimination*, *Stepwise regression*. If *all possible regression* is selected, user can choose one among *Adjusted R-squared*, *R-squared*, and *Mallow's C(p)* as Statistic.

Include constant in the equation

By default, the regression model includes a constant term in the equation. Deselecting this option forces regression through the origin, this is rarely done.

Correlation Matrix

If this is selected, a matrix of correlation coefficients for all numeric variables will be displayed.

Scatterplot Matrix

If this is selected, a matrix of scatterplots for all numeric variables will be displayed. This option is disabled if all possible regression selection method is selected.

Variance Inflation Factors

If this is selected, it displays the variance inflation factors (VIF) which measures the Collinearity (or multicollinearity). This option is disabled if all possible regression selection method is selected.

Durbin-Watson

If this is selected, it displays the Durbin-Watson test for serial correlation of the residuals. This option is disabled if all possible regression selection method is selected.

Leverage Values

If this is selected, the leverage values will be computed and included in the output data file containing the raw data, residuals, and predicted values. This option is disabled if all possible regression selection method is selected.

Cook's Distance

If this is selected, Cook's distance will be computed and included in the output data file containing the raw data, residuals, and predicted values. This option is disabled if all possible regression selection method is selected.

Regression Coefficients

User can choose to display the confidence interval and the covariance matrix. If at least one of this is selected, the result is embedded in the table of parameter estimates. For the confidence interval, the default value is 95%. This option is disabled if all possible regression selection method is selected.

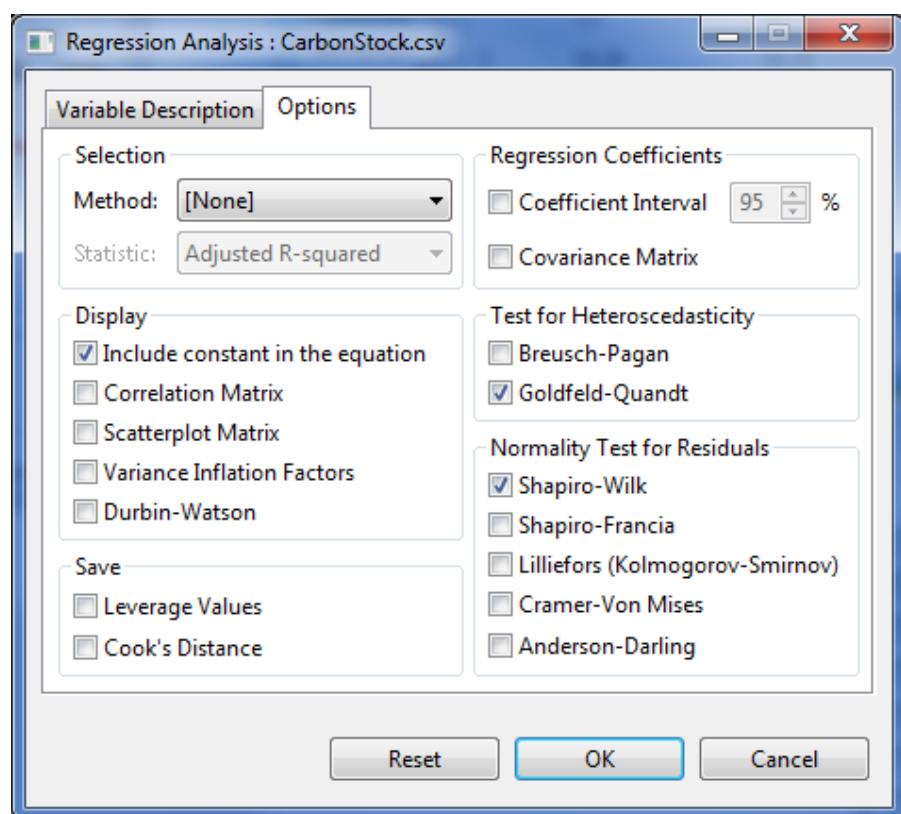
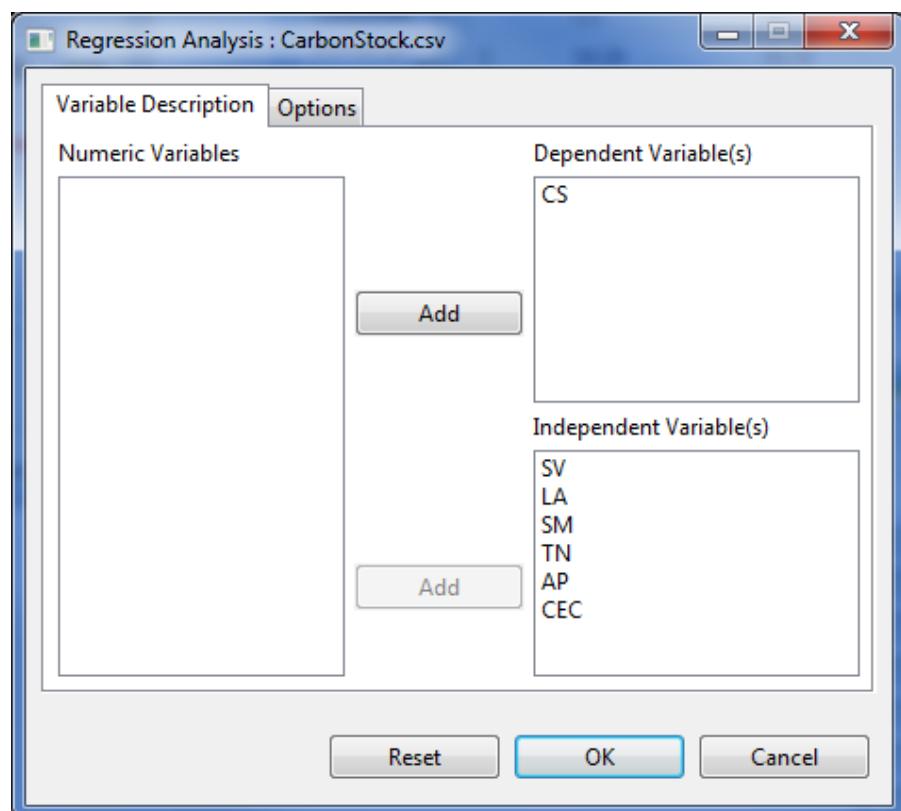
Test for Heteroskedasticity

There are two available procedures namely: Breusch-Pagan and Goldfeld-Quandt. This option is disabled if all possible regression selection method is selected.

Normality Test

This option will perform normality test for the residuals. There are five procedures available, namely: Shapiro-Wilk, Shapiro-Francia, Lilliefors, Cramer-Von Mises and Anderson-Darling. This option is disabled if all possible regression selection method is selected.

For the example, the completed **Variable Description** tab and **Option** tab should appear as shown:



- Once all options have been specified, click the **OK** button. The **Regression Analysis** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

The Linear Regression Analysis procedure in STAR saves all text output in the filename *LinearRegOutput.txt*. By default, the text output includes the descriptive statistics, Analysis of Variance table, a table of model summary and the parameter estimates:

Descriptive Statistics

Variable	N_NonMissObs	Mean	StdDev	SE_Mean
CS	50	20.28	3.99	0.56
SV	50	31.63	10.58	1.50
LA	50	7.75	0.35	0.05
SM	50	7.40	0.21	0.03
TN	50	0.16	0.04	0.01
AP	50	5.31	1.48	0.21
CEC	50	16.29	2.01	0.28

LINEAR REGRESSION ANALYSIS

Model Fitted: CS ~ SV + LA + SM + TN + AP + CEC

Analysis of Variance Table

Source	DF	Sum of Square	Mean Square	F Value	Pr(> F)
Model	6	750.3633	125.0606	186.74	0.0000
Error	43	28.7975	0.6697		
Total	49	779.1608			

Model Summary:

Root MSE	CS Mean	CV (%)	R-Square	Adj R-Sq
0.8184	20.28	19.67	0.9630	0.9579

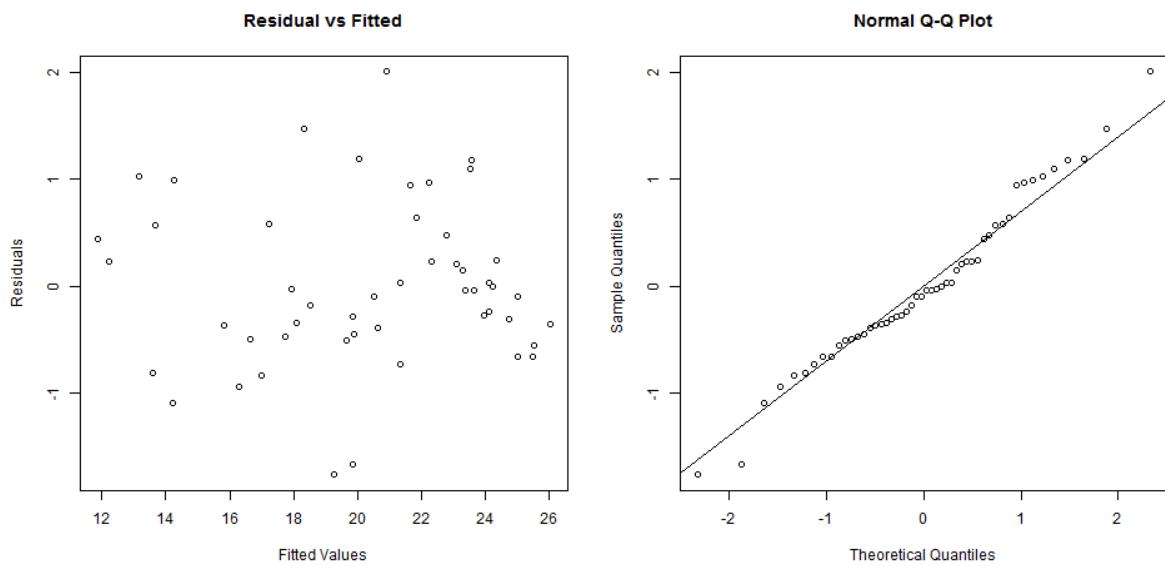
Parameter Estimates:

Variable	Estimate	Std. Error	t value	Pr(> t)
Intercept	-48.91	6.93	-7.06	0.0000
SV	0.04	0.02	2.73	0.0091
LA	0.79	0.53	1.48	0.1471
SM	6.41	1.11	5.75	0.0000
TN	33.23	7.75	4.29	0.0001
AP	0.16	0.08	1.90	0.0642
CEC	0.49	0.13	3.82	0.0004

```
Test for Normality
-----
Variable      Method          W Value   Pr(< W)
-----
residual     Shapiro-Wilk    0.9792    0.5207
-----
```

```
Test for Heteroskedasticity
-----
Method           DF1    DF2    GQ Value   Pr (>GQ)
-----
Goldfeld-Quandt test    18     18     1.22    0.6797
-----
```

By default, a *png* file is created for each dependent variable used in the analysis. The file follows the filename format as: <dependent variable>_DiagPlot.png. Each *png* file contains two diagnostics plots. The *Residuals vs. Fitted* plot is for assessing the distribution of the residuals. An ideal plot should reveal that the points are distributed around 0. The residuals should not be affected by the size of the fitted values. Problematic data may result to a plot wherein the points increase as the fitted values increase, thus forming a “funnel-shaped” distribution. This will indicate a violation of the homogeneity of variances requirement for analysis of variance. The *Normal Q-Q Plot* is used for assessing normality. The points should lie about the line. Sample generated graphs are shown below:



This procedure will also create the *LinearRegNewData.csv*. This file contains the original data with the residual and predicted values for each of the dependent variables used in the analysis. This file will also contain the *Leverage Values* and *Cooks's Distance*, if user chose to save these values.

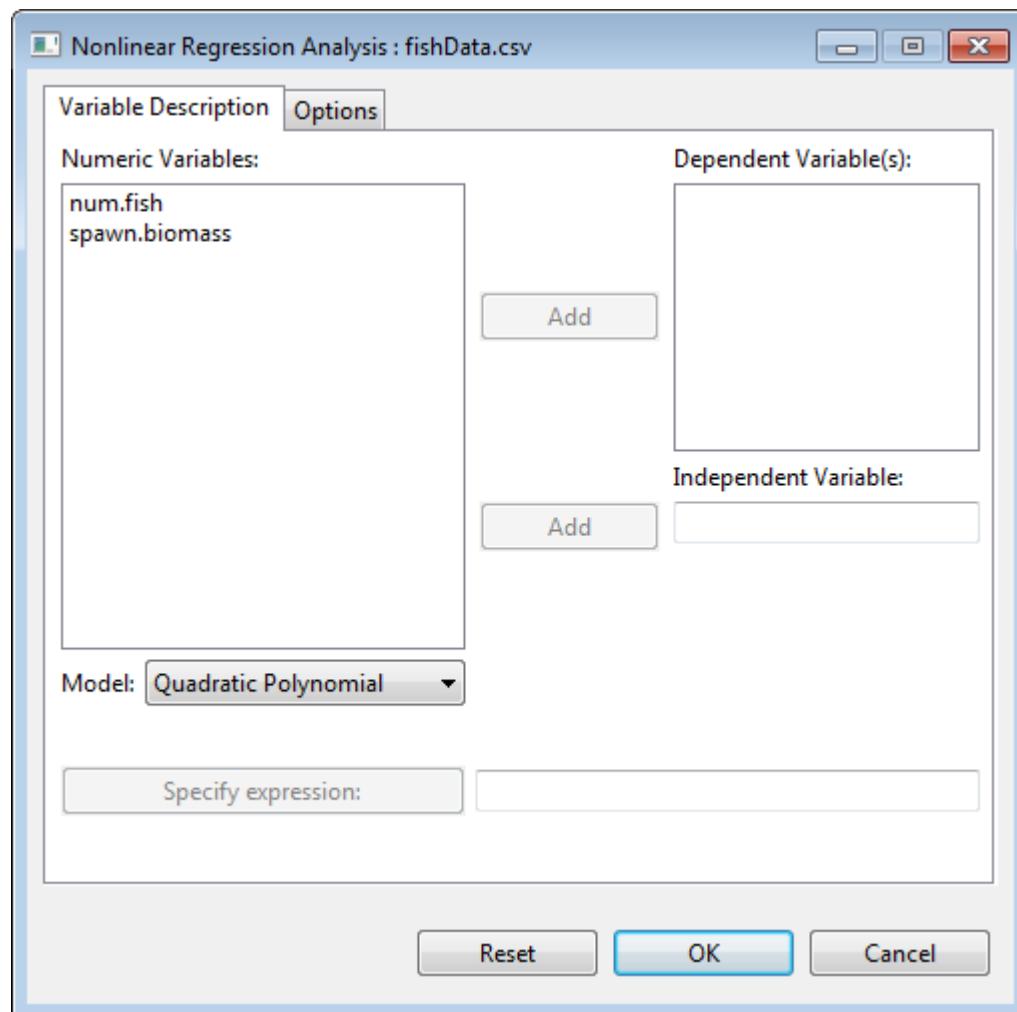
Non-Linear Regression

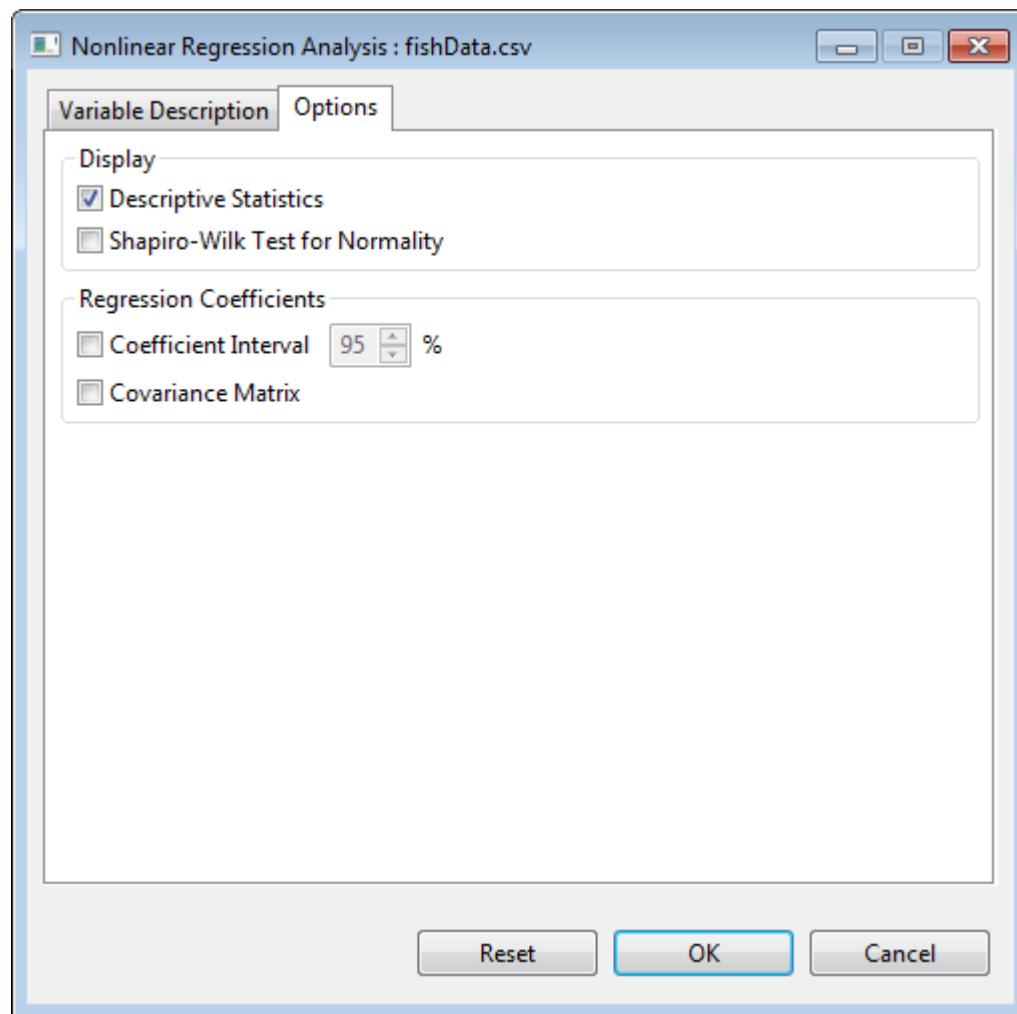
The steps to perform Non-Linear Regression Analysis are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer tab.

For the example, using the project named *My Project*, import the data *fishData* from the package. To import the data from the package, see *Importing from Package* section of this user's manual.

- Click **Analyze | Regression | Non Linear...** from the main window. The **Nonlinear Regression** dialog box will appear.





- Specify the required field and appropriate options for the analysis:

Variable Description Tab

Dependent Variable(s)

For the analysis to proceed, this list box should have at least one entry.

Independent Variable(s)

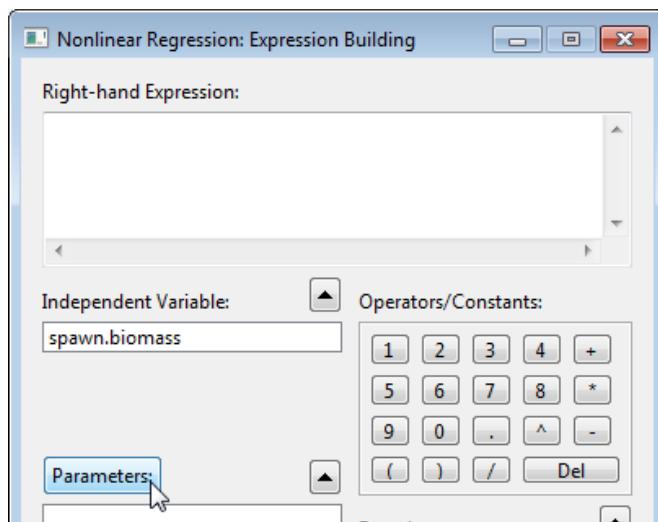
For the analysis to proceed, this list box should have one entry.

Model

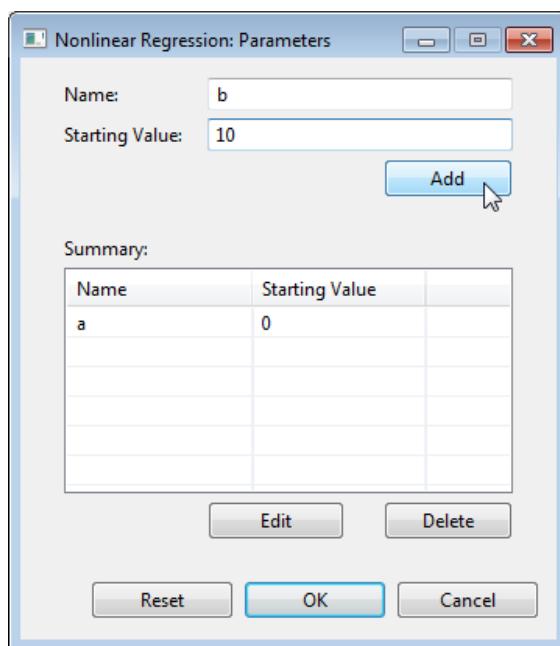
The following models are available for performing the analysis: *Quadratic Polynomial*, *Cubic Polynomial*, *Logistic*, *Four-parameter logistic*, *Gompertz Growth*, or *Weibull Growth Curve*. If the user needs to use a model that is not listed above, the [Custom] option is selected. The user defines the model by clicking on the **Specify expression** button.

Specify expression

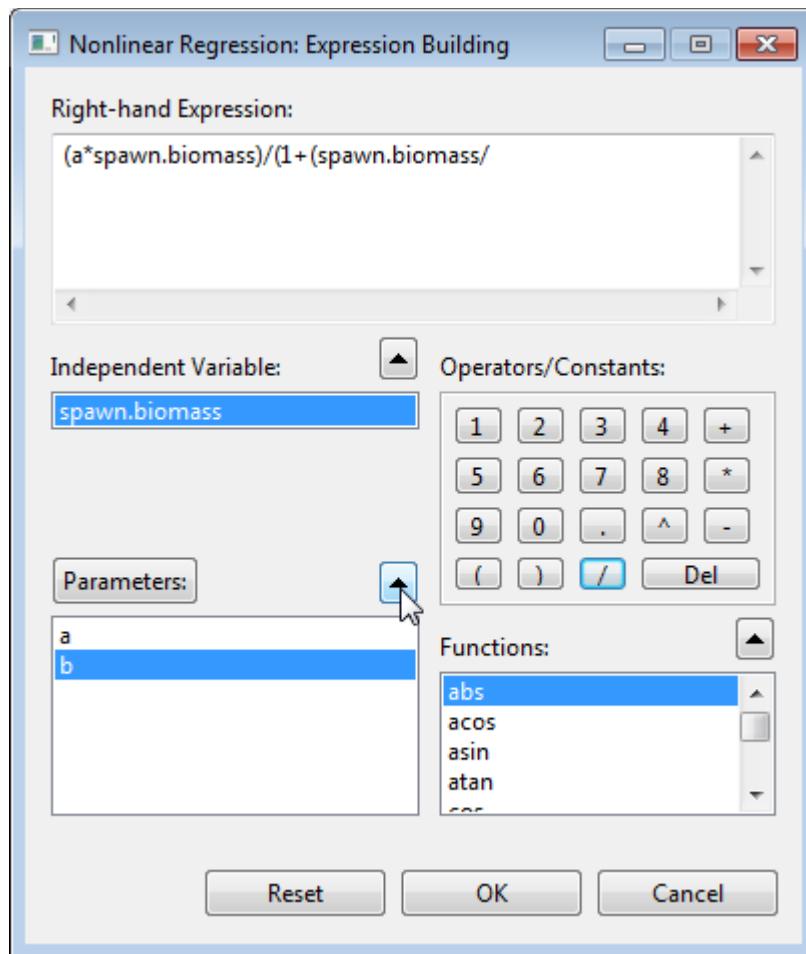
When the **Specify expression** button is pressed, the **Nonlinear Regression: Expression Building** dialog box appears. Prior to defining the expression, parameters need to be defined by clicking on the **Parameters** button.



In the **Nonlinear Regression: Parameters** dialog box, user needs to indicate a name and starting value for each parameter. This is done by typing on the **Name** and **Starting value** text fields and clicking on the **Add** button for each parameter. When all parameters are specified, the **Ok** button is pressed to return to the **Nonlinear Regression: Expression Building** dialog box.



In the Nonlinear Regression: Expression Building dialog box, the **Right-hand Expression** is defined by pressing on the appropriate buttons/items for the independent variable, operator, constant, parameter, or function. The available functions are: abs, acos, asin, atan, cos, exp, log, log10, sin, sqrt and tan. When the expression is specified, the **Ok** button is pressed to return to the main dialog box.



Options Tab

Descriptive Statistics

If this option is selected, a summary table with number of observations, minimum, maximum, mean and standard deviation of the variables is displayed.

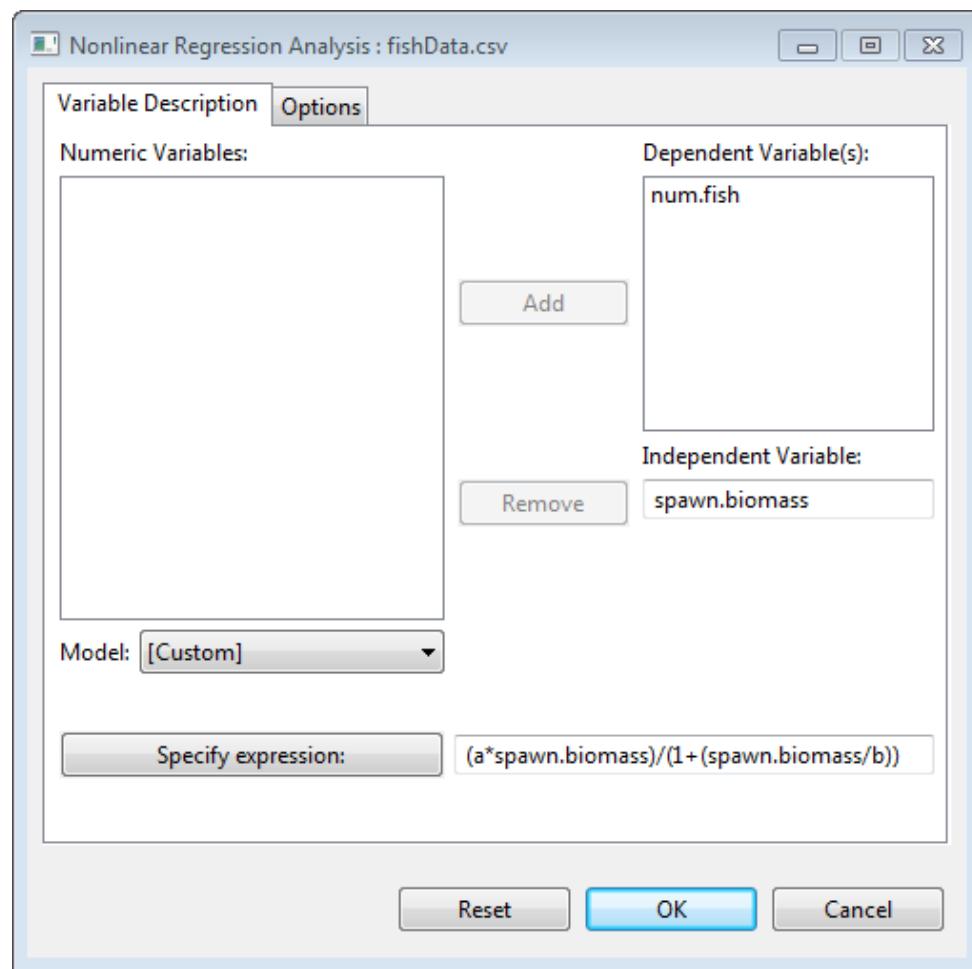
Shapiro-Wilk Test for Normality

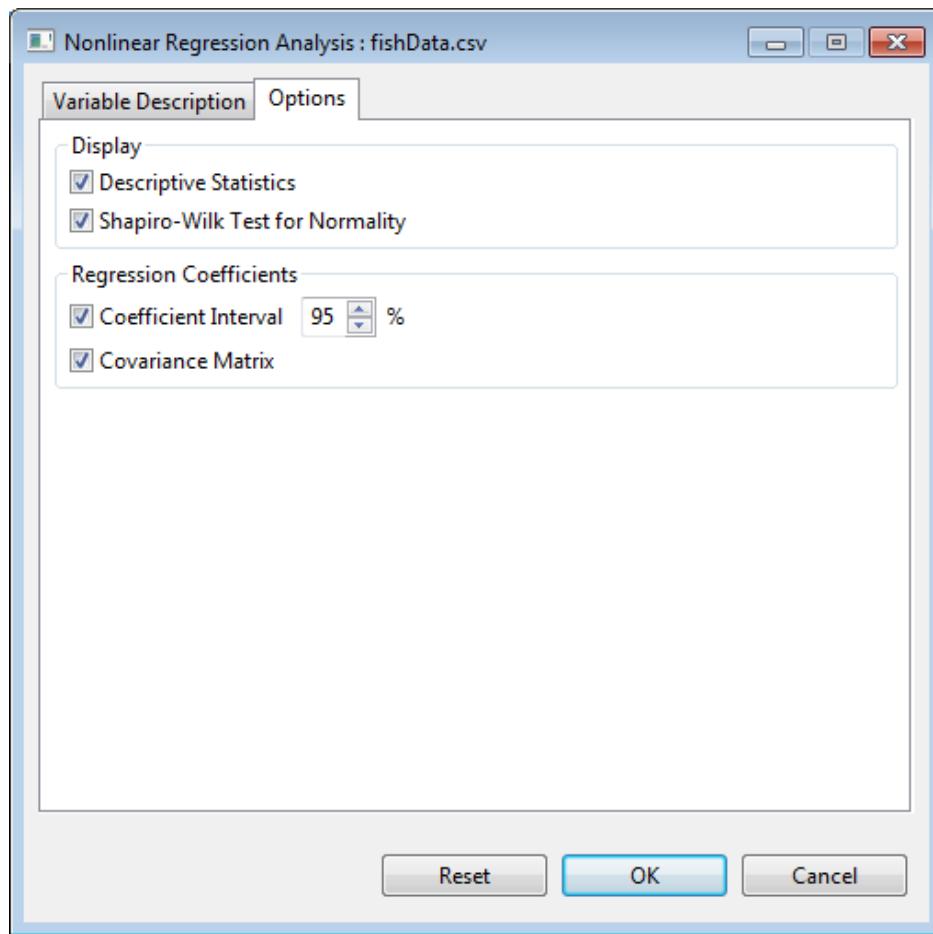
If this option is selected, Shapiro-Wilk Test for Normality will be performed.

Regression Coefficients

User can choose to display the confidence interval and the covariance matrix. If at least one of this is selected, the result is embedded in the table of parameter estimates. For the confidence interval, the default value is 95%.

For the example, the completed **Variable Description** tab and **Option** tab should appear as shown:





- Once all options have been specified, click the **OK** button. The **Nonlinear Regression Analysis** dialog box will be minimized and STAR activates the **Output** page of the **Result Viewer** tab.

The NonLinear Regression Analysis procedure in STAR saves all text output in the filename *NonLinearRegOutput.txt*.

```
Descriptive Statistics
-----
Variable      N_NonMissObs      Mean      StdDev      SE_Mean
-----
num.fish          15     84.53      27.55      7.11
spawn.biomass      15     29.90      16.96      4.38
-----
```

NONLINEAR REGRESSION ANALYSIS

Dependent Variable: num.fish

Test for Normality

Variable	Method	W	Value	Pr(< W)
----------	--------	---	-------	---------

```

residual      Shapiro-Wilk      0.9550      0.6063
-----
Formula: num.fish ~ (a*spawn.biomass) / (1+(spawn.biomass/b))

Parameter Estimates:
-----
Variable      Estimate   Std. Error   t value   LL CI*   UL CI*   Pr(>|t|)

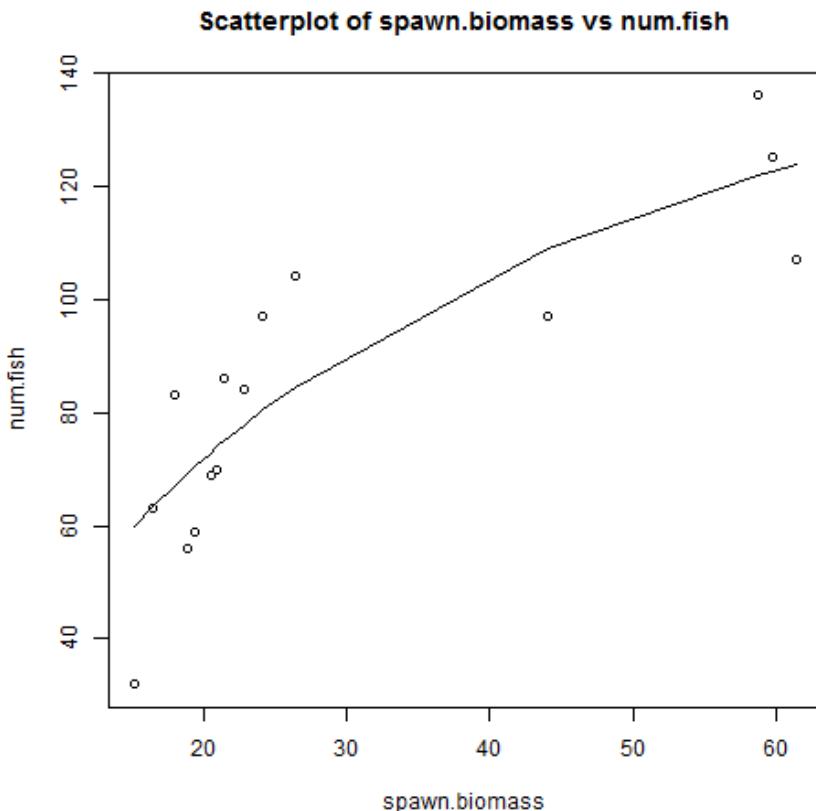
a             5.75       1.03        5.57     3.52     7.98    0.0001
b            33.16      11.41       2.91     8.52    57.80    0.0122
-----
* At 95% Confidence Interval

Residual Sum of Squares: 2809.0013

Coefficient Variance-Covariance Matrix
      a         b
a  1.0659 -11.428
b -11.4277 130.103

```

Sample generated graph is shown below:



Nonparametric Tests

Nonparametric Tests procedure provides several tests that do not require assumptions about the shape of the underlying distribution.

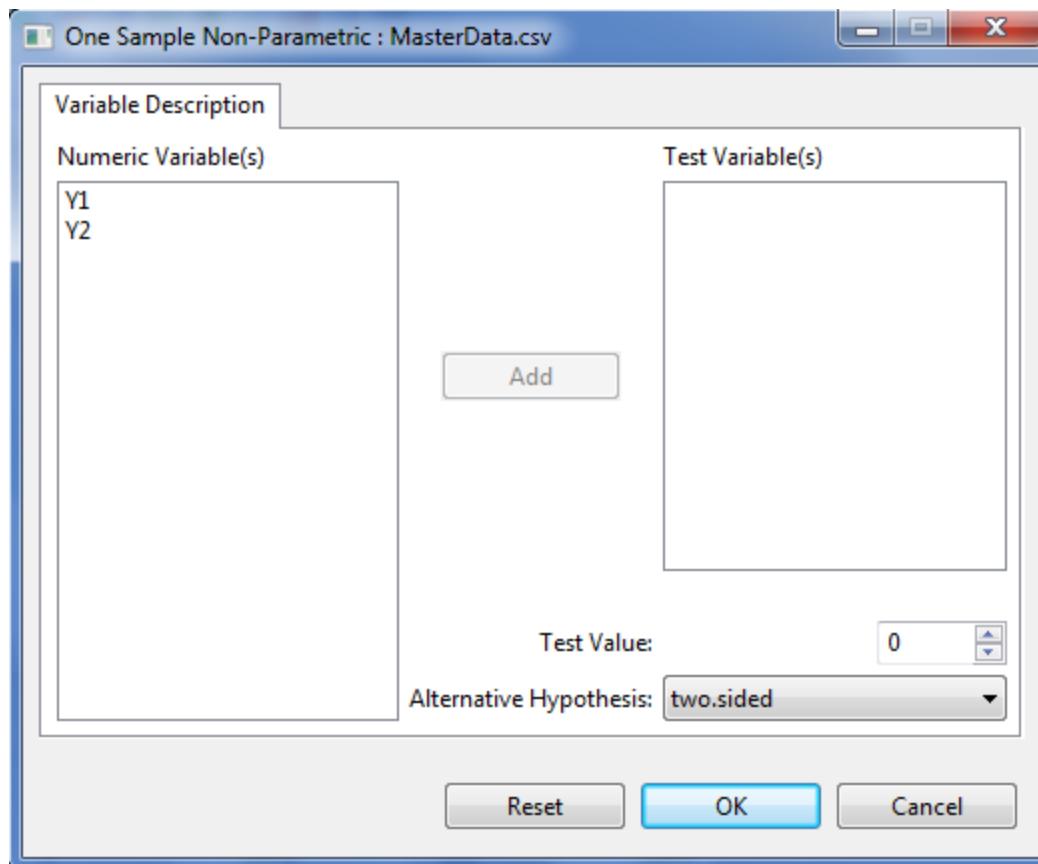
One Sample

The steps to perform one sample non-parametric test are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *MasterData* from the *Data* folder of the project named *SampleProject*. To import data from the *SampleProject*, see *Importing Data from Other Project* section of this user's manual.

- From the main window of STAR, click **Analyze | Non-Parametric Test | One Sample....** The **One Sample Non-Parametric** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Test Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this dialog box.

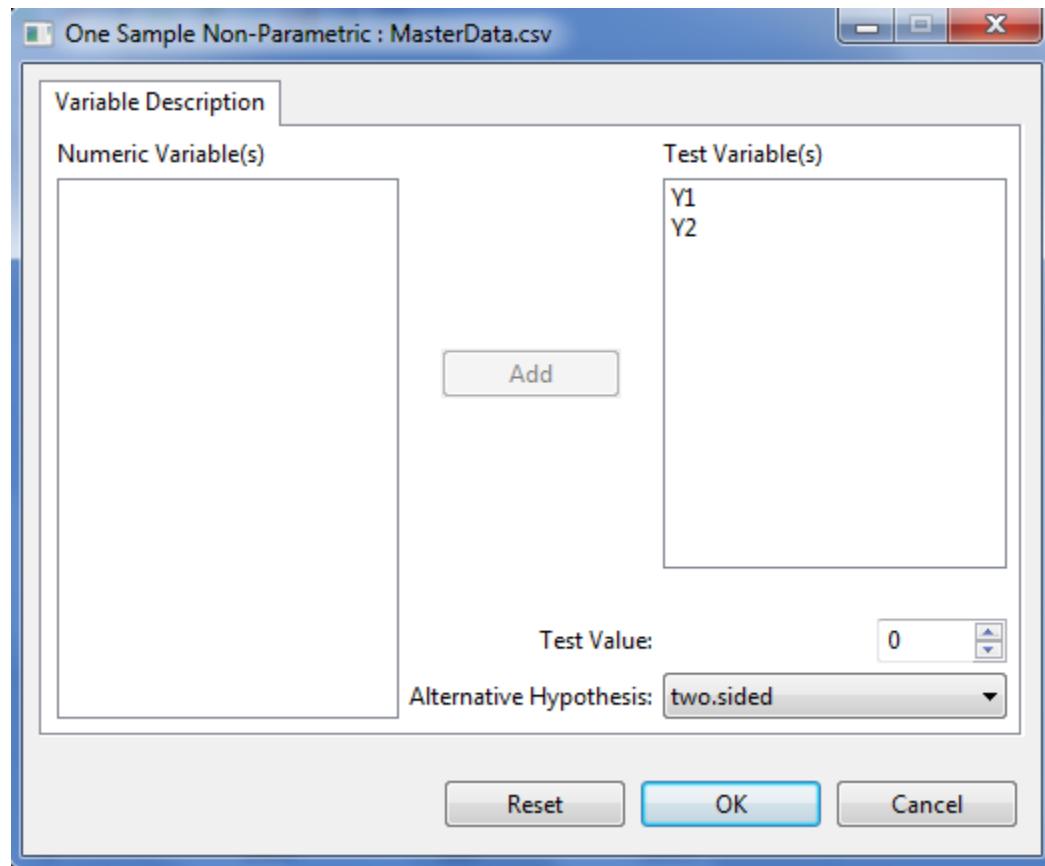
Test Value

This field is required and will only accept one entry at a time. The default value is 0. User can change the value by typing a value inside the spin box or by pressing the up-arrow key to increment the value, down-arrow key, otherwise.

Alternative Hypothesis

User can select a one-tailed or two-tailed test. The default alternative hypothesis used is that for the two-sided test.

For the example, the completed dialog box should appear as shown below:



- Click **OK** button to perform the analysis. The **One Sample Non-Parametric** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample output is shown below:

```
Sigh Test: Md = 0
-----
Variable Statistics Prob > |V| LL_CI * Median UL_CI *
-----
Y1      4656     0.0000    132.00   132.00   132.00
Y2      4656     0.0000    122.95   122.99   123.00
-----
* 95% Confidence Interval
```

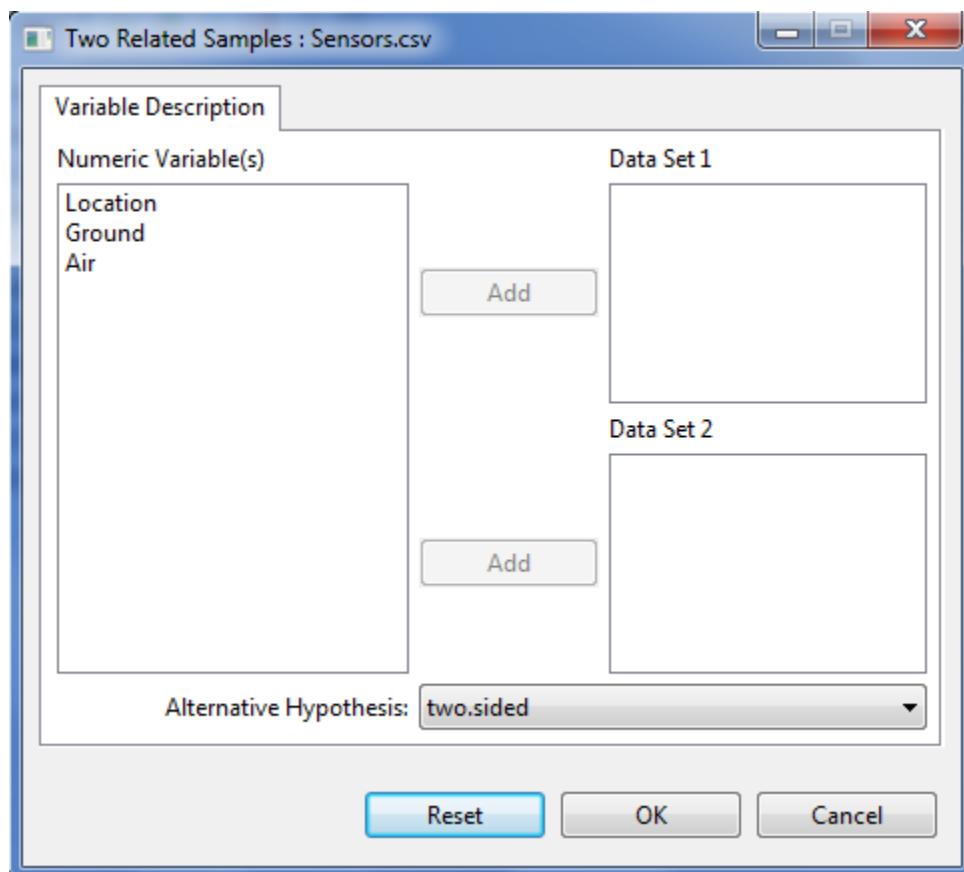
Two Related Samples

The steps to perform Kruskal Wallis test are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *Sensors* from package. To import the data from the package, see *Importing from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | Non-Parametric Test | Two Related Samples....** The **Two Related Samples** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Data Set 1

This field needs at least one entry for the analysis to proceed and must have the same number of entry as **Data Set 2**.

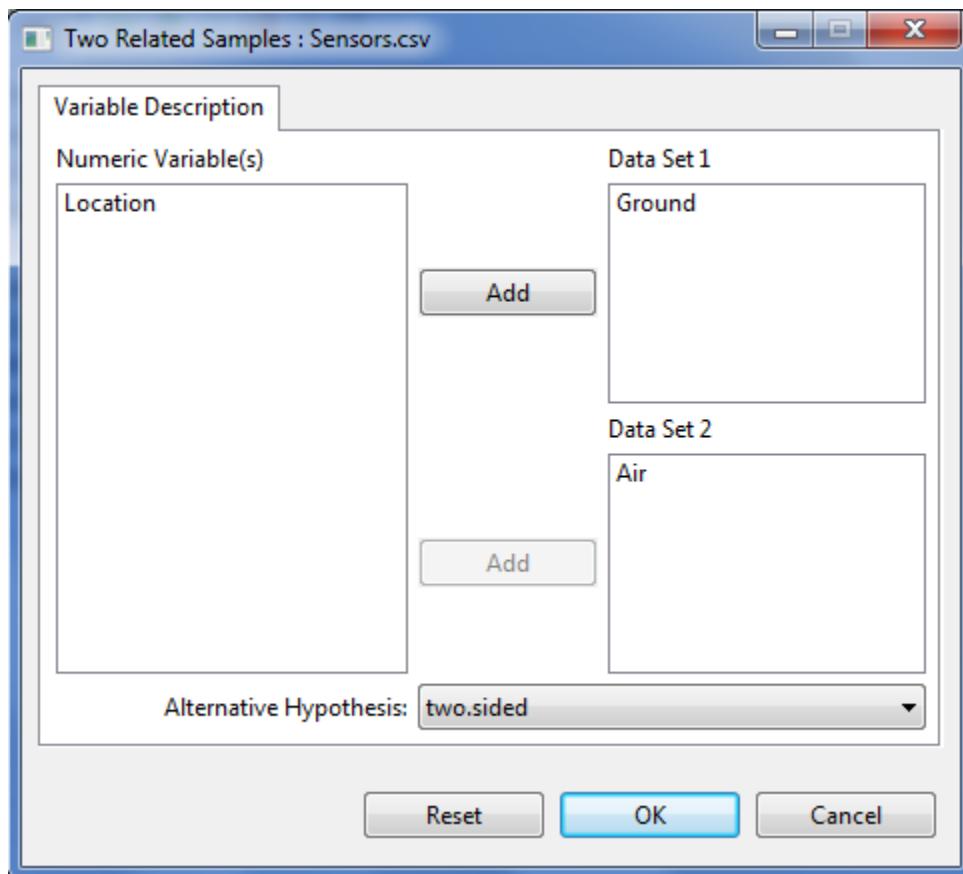
Data Set 2

This field needs at least one entry for the analysis to proceed and must have the same number of entry as **Data Set 1**.

Alternative Hypothesis

User can select a one-tailed or two-tailed test. The default alternative hypothesis used is that for the two-sided test.

For the example, the completed dialog box should appear as shown below:



- Click **OK** button to perform the analysis. The **Two Related Samples** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample output is shown below:

Wilcoxon Signed Rank Test

```
-----  
Variable      Statistics   Prob > |V|   Median  
-----  
Ground - Air    0.0000     0.0020   -1.60  
-----
```

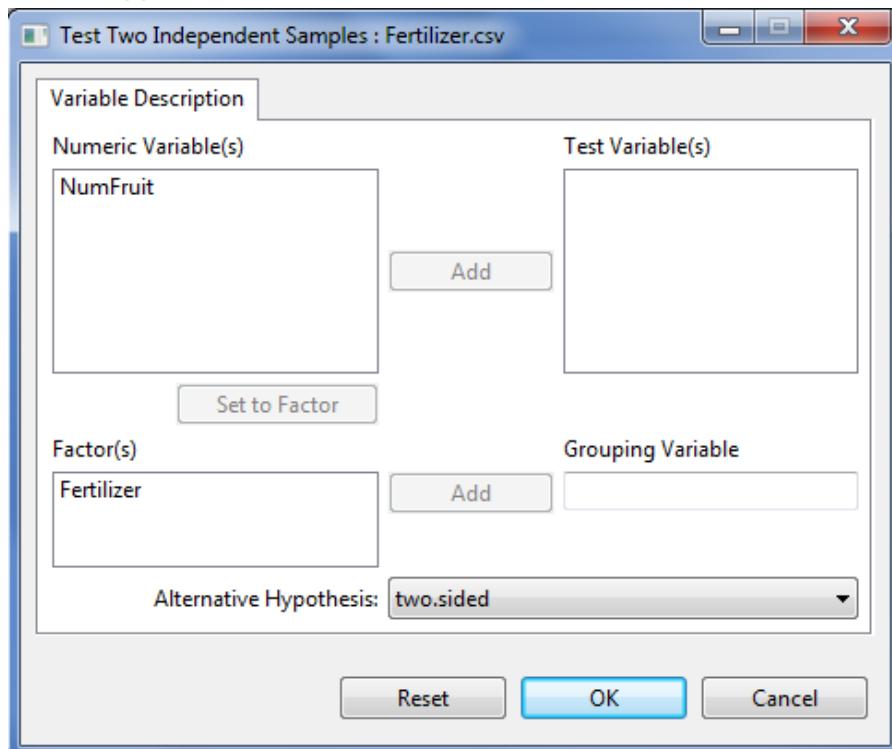
Two Independent Samples

The steps to perform Kruskal Wallis test are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *Fertilizer* from package. To import the data from the package, see *Importing from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | Non-Parametric Test | Test for Two Independent Samples....** The **Test for Two Independent Samples** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Test Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this dialog box.

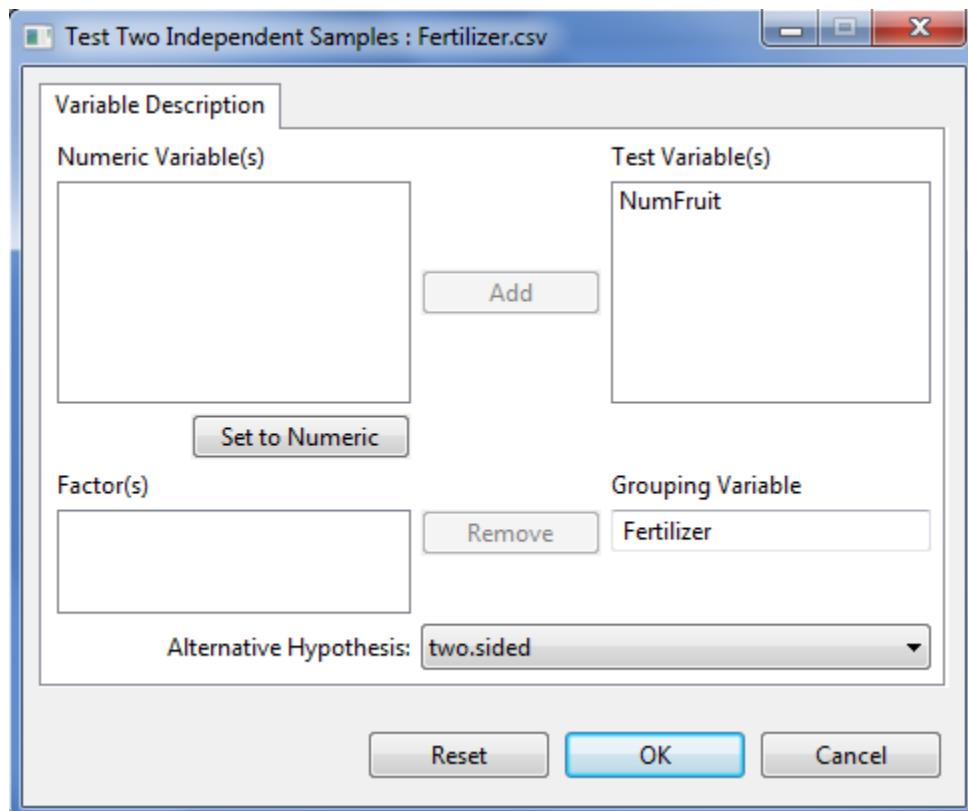
Grouping Variable

Only one item can be added in this field and should come from the **Factor(s)** list box.

Alternative Hypothesis

User can select a one-tailed or two-tailed test. The default alternative hypothesis used is that for the two-sided test.

For the example, the completed dialog box should appear as shown below:



- Click **OK** button to perform the analysis. The **Test for Two Independent Samples** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample output is shown below:

```
Mann-Whitney (Wilcoxon Rank Sum) Test

-----
Variable Statistics Prob > |W| LL_CI * Median UL_CI *
-----
NumFruit      163     0.0371   0.0000    3.00    5.00
-----
* 95% Confidence Interval
```

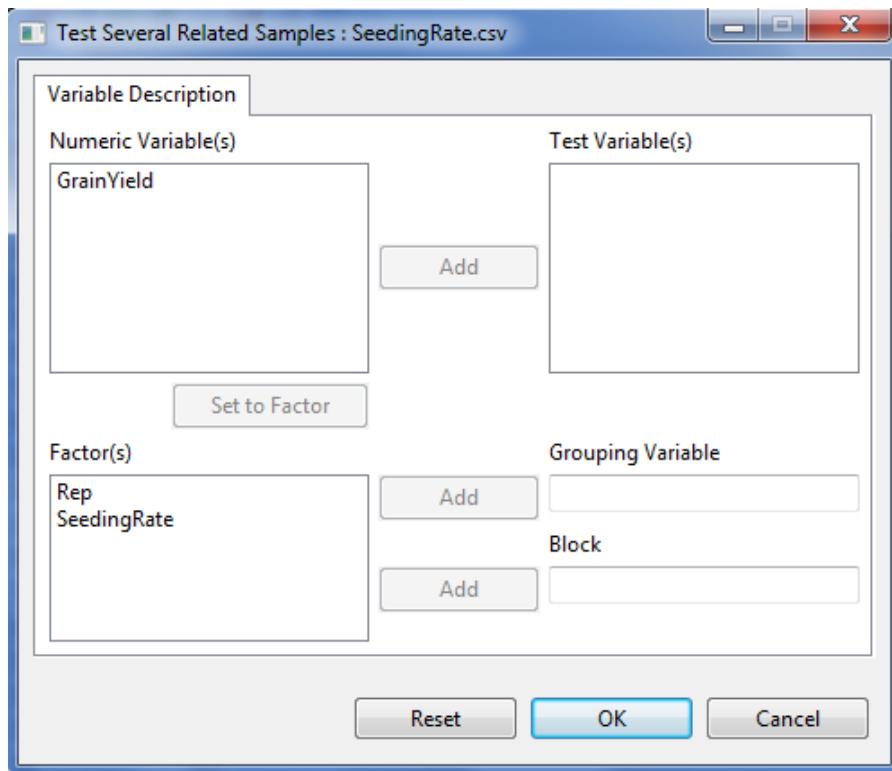
Test for Several Related Samples

The steps to perform Kruskal Wallis test are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *Insecticides* from package. To import the data from the package, see *Importing from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | Non-Parametric Test | Test for Several Independent Samples....** The **Test for Several Independent Samples** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Test Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this dialog box.

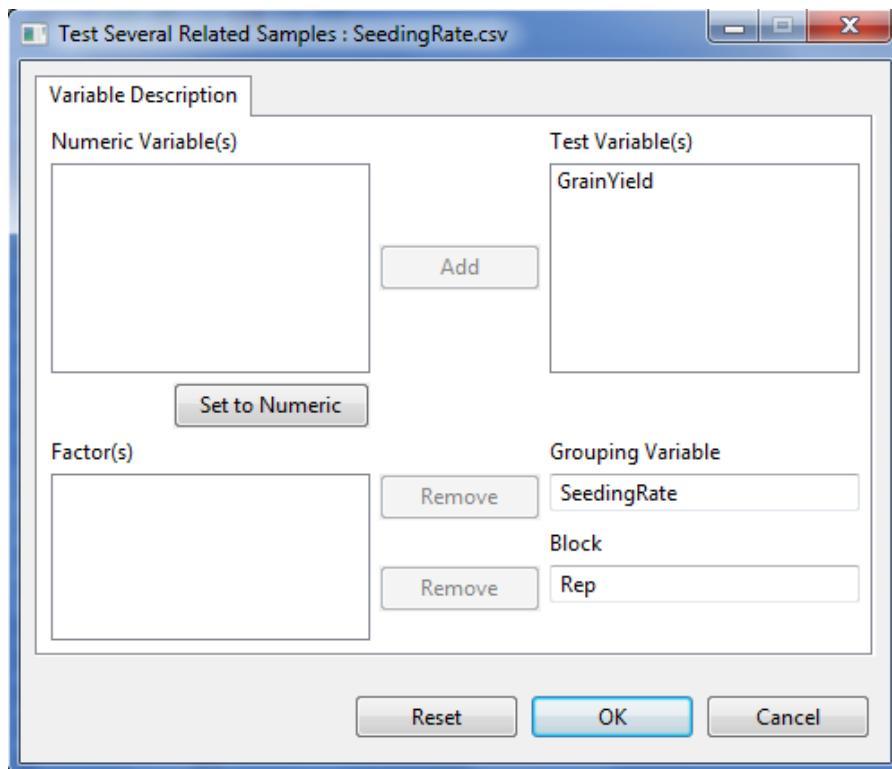
Grouping Variable

Only one item can be added in this field and should come from the **Factor(s)** list box.

Block

Only one item can be added in this field and should come from the **Factor(s)** list box.

For the example, the completed dialog box should appear as:



- Click **OK** button to perform the analysis. The **Test for Several Independent Samples** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample output is shown below:

```
Rank Sums for Variable GrainYield
Classified by SeedingRate
-----
SeedingRate      N   Sum of Ranks   Mean of Ranks
-----
25              4    20.50          5.12
50              4    20.30          5.08
75              4    21.20          5.30
100             4    17.40          4.35
125             4    18.70          4.67
150             4    13.50          3.38
-----
* Average scores are used for ties.

Friedman rank sum test
-----
Chi-Square        15.2985
DF                  5
Pr > Chi-Square     0.0092
-----
```

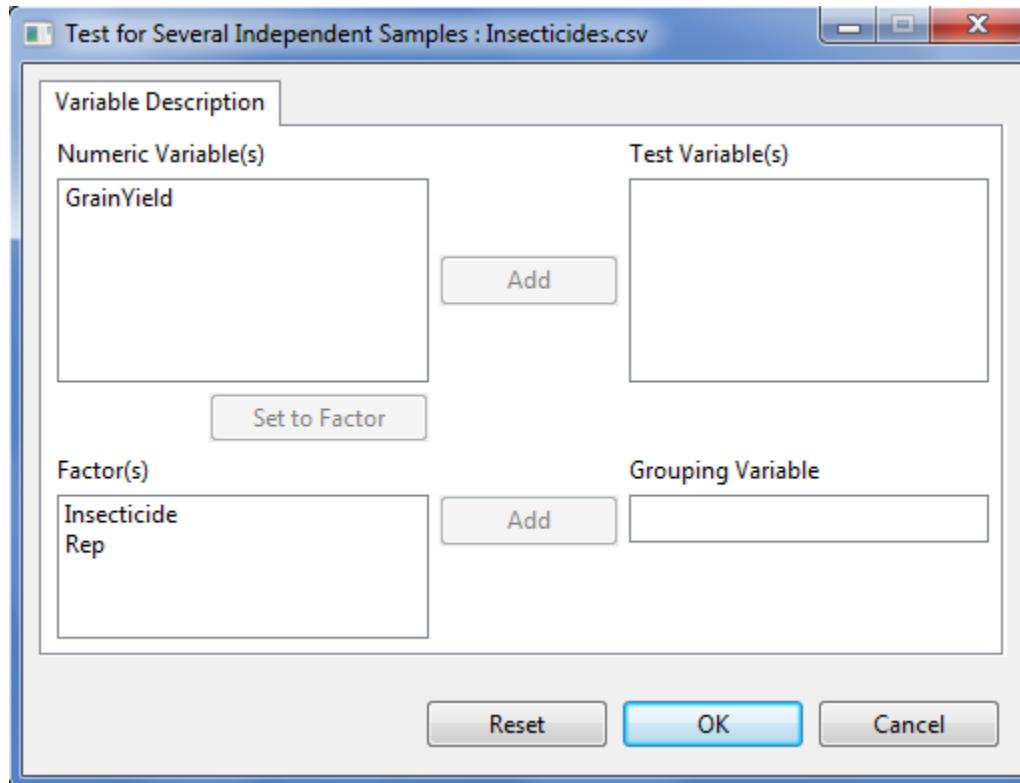
Test for Several Independent Samples

The steps to perform Kruskal Wallis test are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *Insecticides* from package. To import the data from the package, see *Importing from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | Non-Parametric Test | Test for Several Independent Samples....** The **Test for Several Independent Samples** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

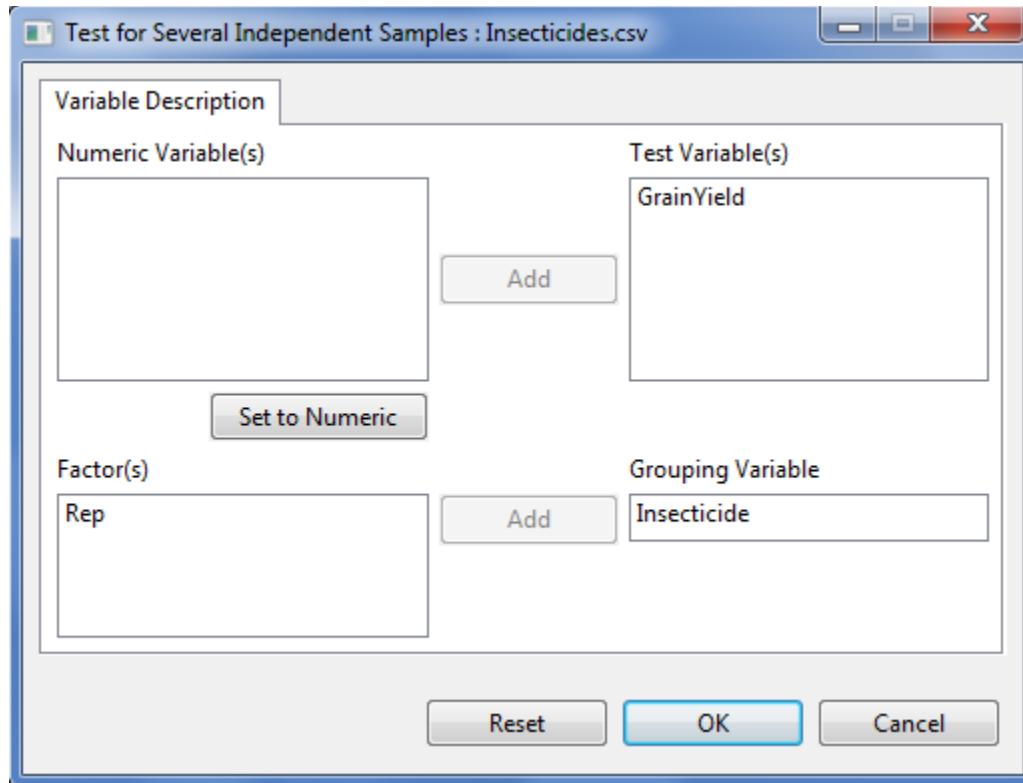
Test Variable(s)

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this dialog box.

Grouping Variable

Only one item can be added in this field and should come from the **Factor(s)** list box.

For the example, the completed dialog box should appear as shown below:



- Click **OK** button to perform the analysis. The **Test for Several Independent Samples** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample output is shown below:

Rank Sums for Variable YIELD
Classified by Variable TREATMENT

TREATMENT	N	Sum of Ranks	Mean of Ranks
Azodin	4	64	16.00
Control	4	12	3.00
DDT + phi-BHC	4	91	22.75
Dimercon-Boom	4	41	10.25
Dimercon-Knap	4	35	8.75
Dol-Mix (1kg)	4	66	16.50
Dol-Mix (2kg)	4	97	24.25

* Average scores are used for ties.

Kruskal-Wallis rank sum test

Chi-Square	20.8522
DF	6
Pr > Chi-Square	0.0020

Multivariate Analysis

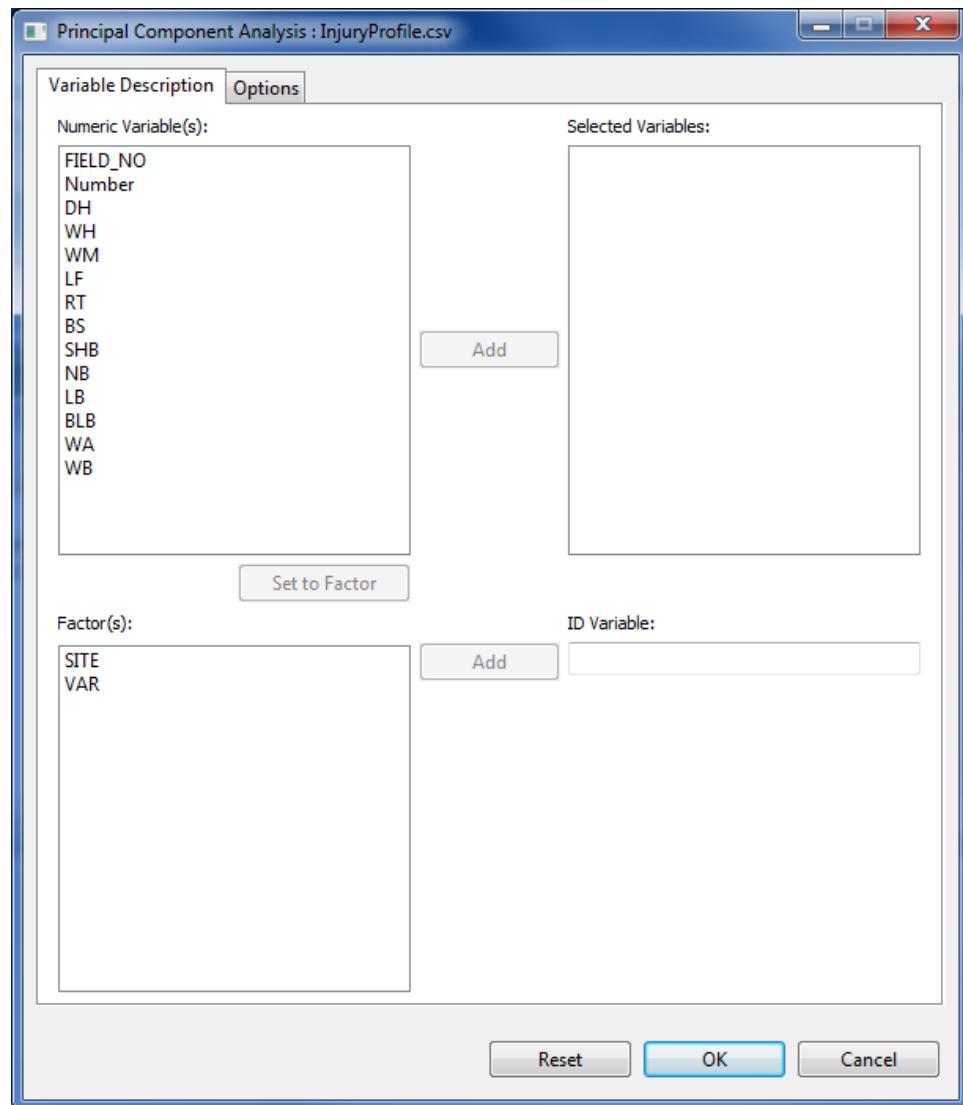
Principal Component Analysis

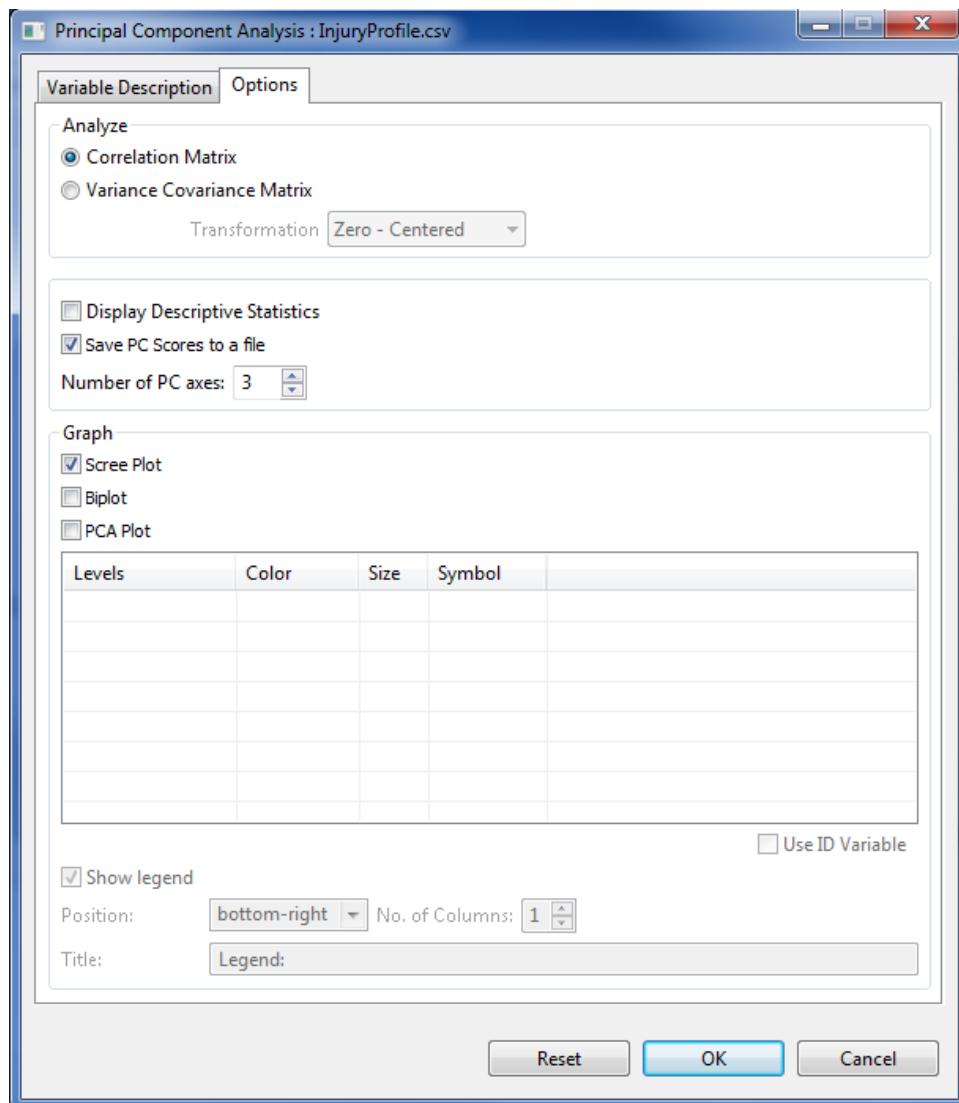
The steps to perform Multivariate Analysis of Variance are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *InjuryProfile* from package. To import data from the package, see *Importing Data from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | Multivariate Analysis | Principal Component Analysis....** The **Principal Component Analysis** dialog box will appear.





- Specify the required field and appropriate options for the analysis.

Variable Description Tab

Selected Variables:

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this dialog box.

ID Variable

Only one item can be entered in this entry box and should come from the **Factor(s)** list box.

Options Tab

Analyze

The user has the option to use Correlation matrix and Variance Covariance matrix for the analysis. By default, Correlation matrix is selected. If the user choose Variance Covariance matrix, the transformation field will be enabled. The user has the option to choose whether the transformation to be used is Zero-Centered, Unit Variance or no transformation. By default, Zero-centered is selected.

Display Descriptive Statistics

If checked, the Descriptive Statistics of the selected variables will be displayed.

Save PC Scores to a file

The user has the option to save PC scores to a file. By default, this option is selected.

Number of PC Axes

The user has the option to specify the number of PCs to be displayed in the PC scores and graphs. By default, the number of PC axes is 3.

Scree Plot

If checked, Scree plot will be displayed. By default, this option is selected.

Biplot

If checked, Biplot will be displayed.

PCA Plot

The user has the option to display PCA plot. If the user specified an ID Variable and also selected PCA Plot, the levels of the ID variable will be automatically displayed on the format table. In this table, user can modify the PCA plot by changing the color, size and symbol.

Use ID Variable

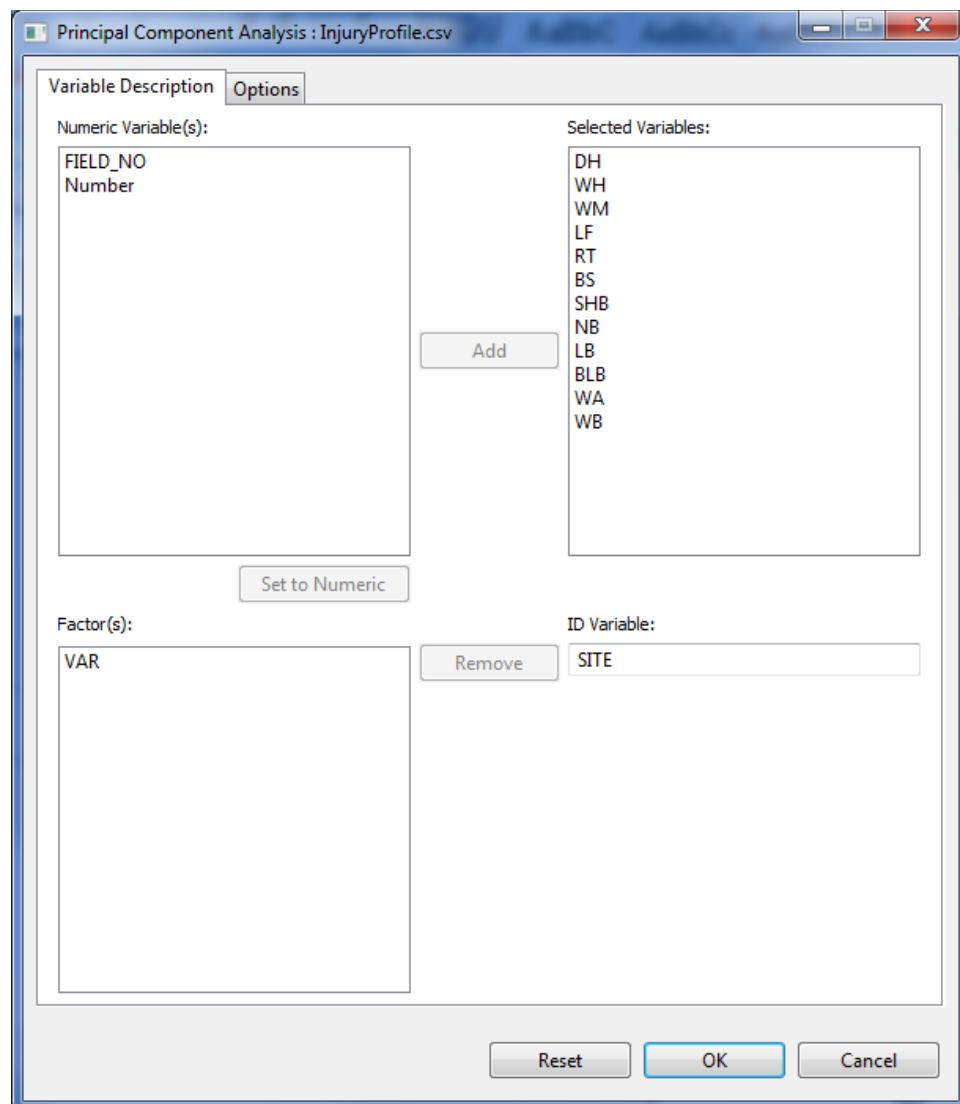
This option will use the levels of the ID variable as points in the PCA plot. This option will only be enabled if the user specifies an ID variable.

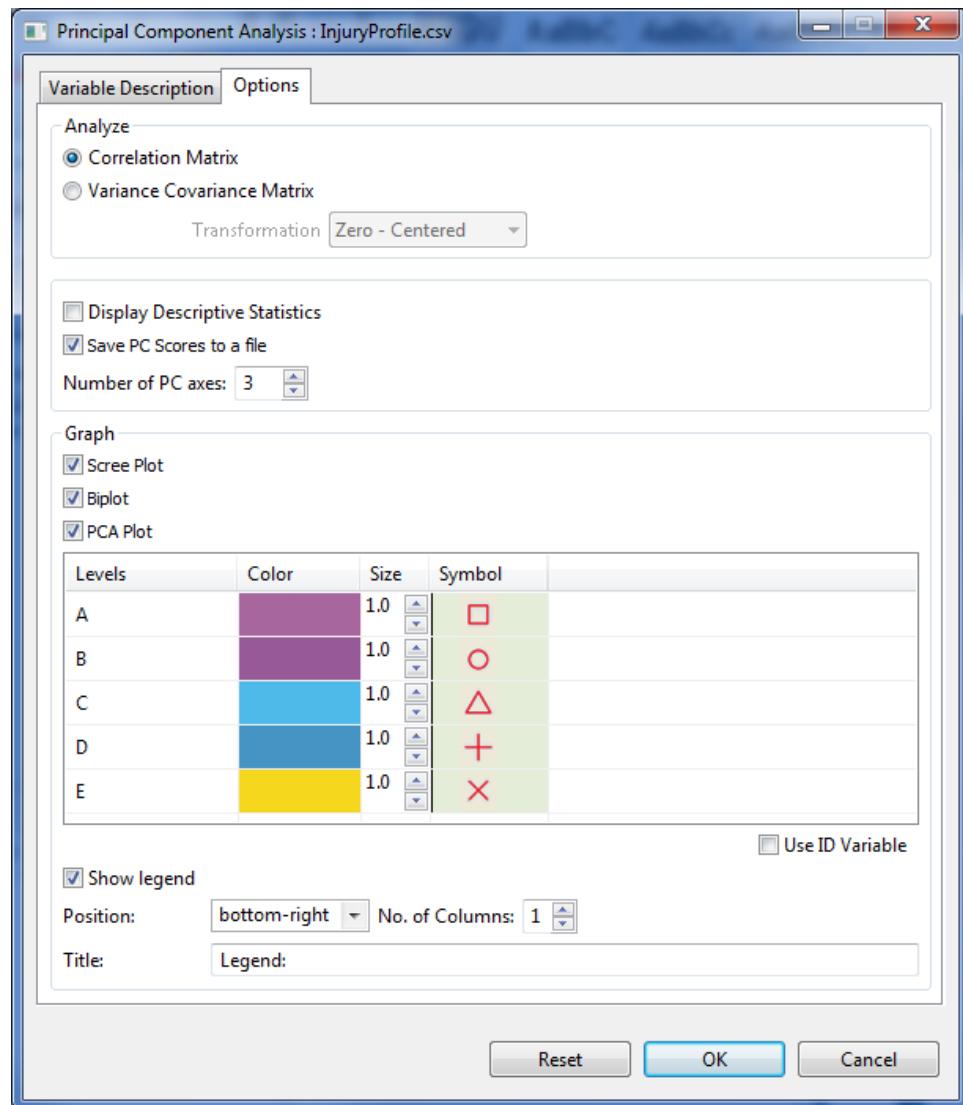
Show legend

If checked, a legend will be added to the graph. By default the title is *Legend*. The legend may be placed in any of nine positions within the graph: *bottom*, *bottom-left*, *bottom-right*, *center*, *left*, *right*, *top*, *top-left*, *top-right*. The items on the legend can be

listed in one or more columns, the maximum of which is the number of levels of the category variable.

For the example, the completed dialog box should appear as shown below:





- Click **OK** button to perform the analysis. The **Principal Component Analysis** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample output is shown below:

Number of Observations: 200

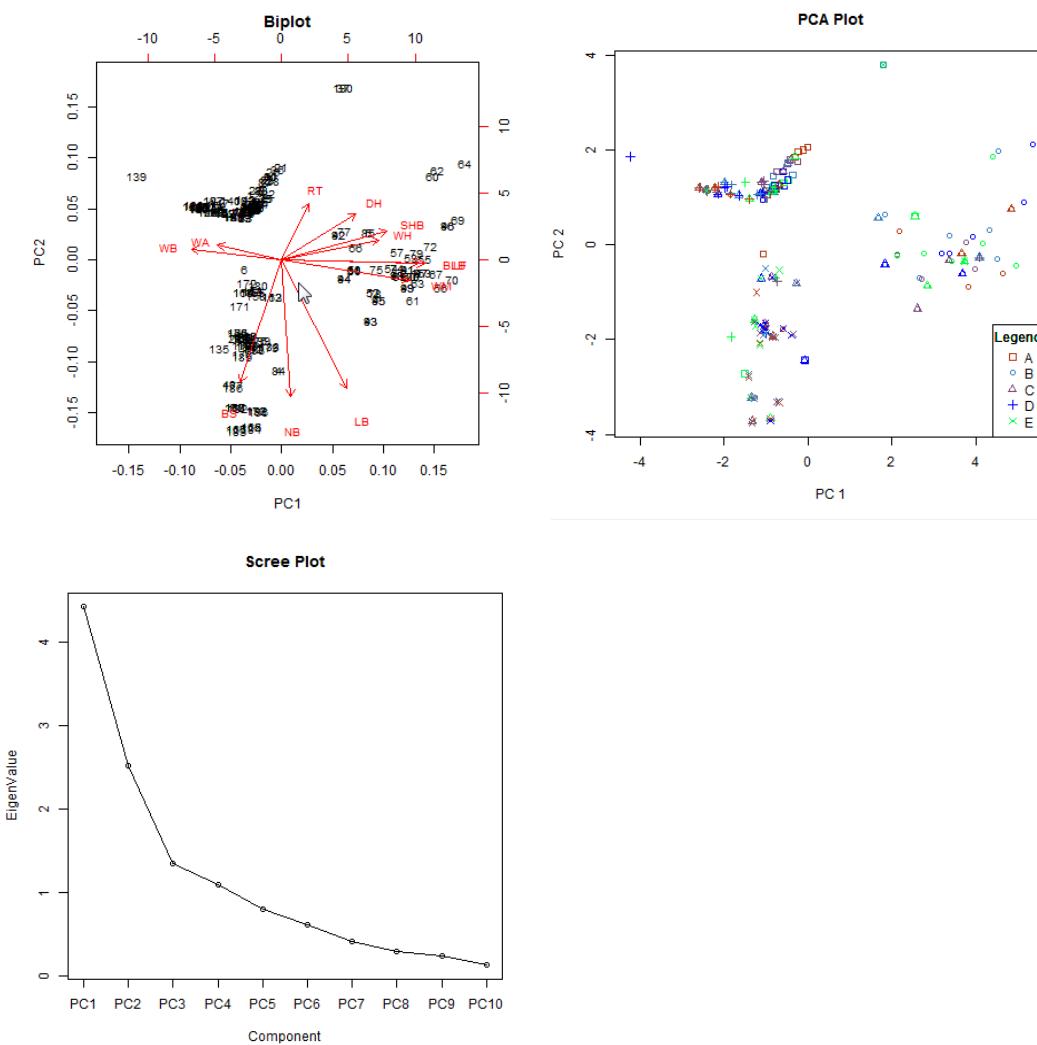
PRINCIPAL COMPONENT ANALYSIS

Statistics	PC1	PC2	PC3	...	PC12
Standard deviation	2.1013	1.5853	1.1604	...	0.1885
Proportion of Variance	0.3680	0.2094	0.1122	...	0.0030
Cumulative Proportion	0.3680	0.5774	0.6896	...	1.0000
EigenValues	4.4155	2.5131	1.3466	...	0.0355

EIGENVECTORS

Variables	PC1	PC2	PC3	...	PC12
DH	0.2339	0.1939	-0.2837	...	0.0123
WH	0.3061	0.0842	0.1799	...	-0.1294
WM	0.4030	-0.0838	0.1610	...	-0.2288
LF	0.4498	-0.0142	0.2328	...	0.8264
RT	0.0846	0.2345	-0.5976	...	0.0673
BS	-0.1302	-0.5115	-0.1688	...	0.0078
SHB	0.3306	0.1190	-0.1340	...	-0.1195
NB	0.0294	-0.5701	-0.1607	...	-0.0285
LB	0.2036	-0.5350	0.0086	...	0.0383
BLB	0.4348	-0.0133	0.2524	...	-0.4755
WA	-0.2021	0.0634	0.4050	...	0.0051
WB	-0.2850	0.0439	0.3871	...	0.0251

Sample graphical output of the analysis is shown below:



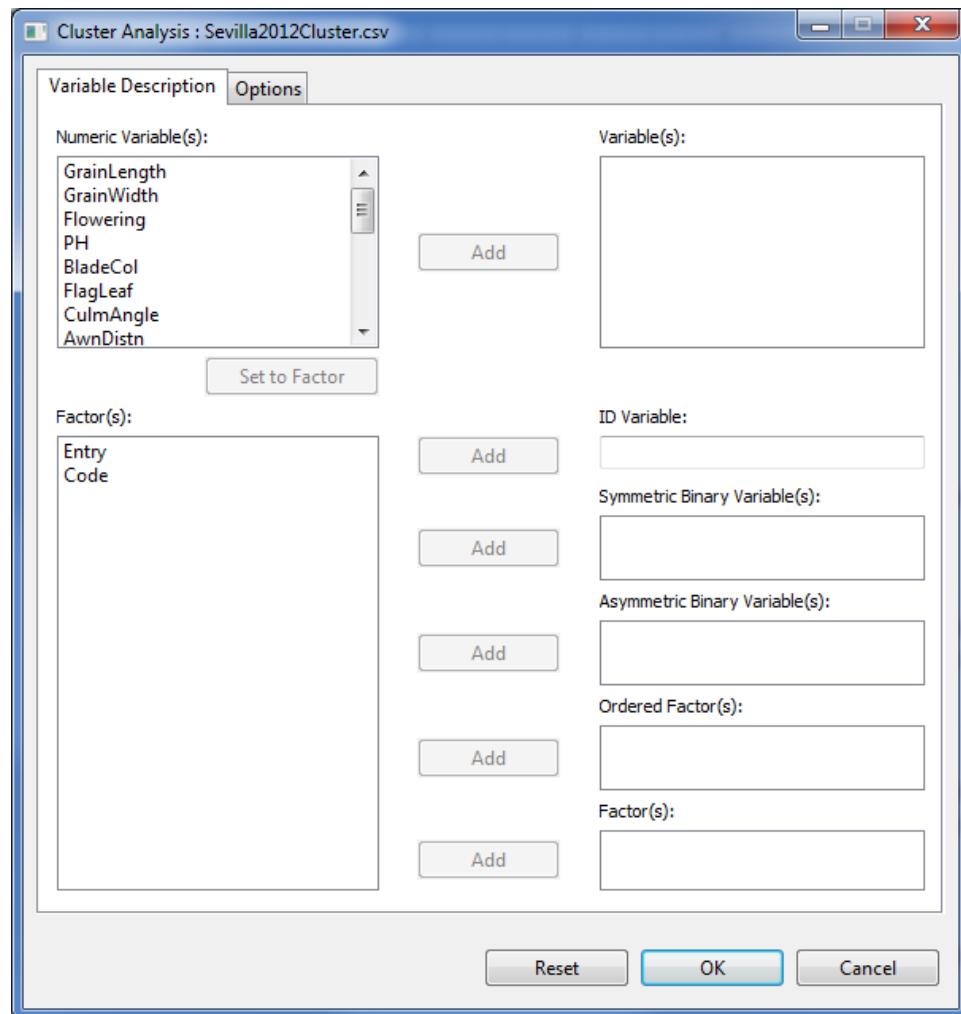
Cluster Analysis

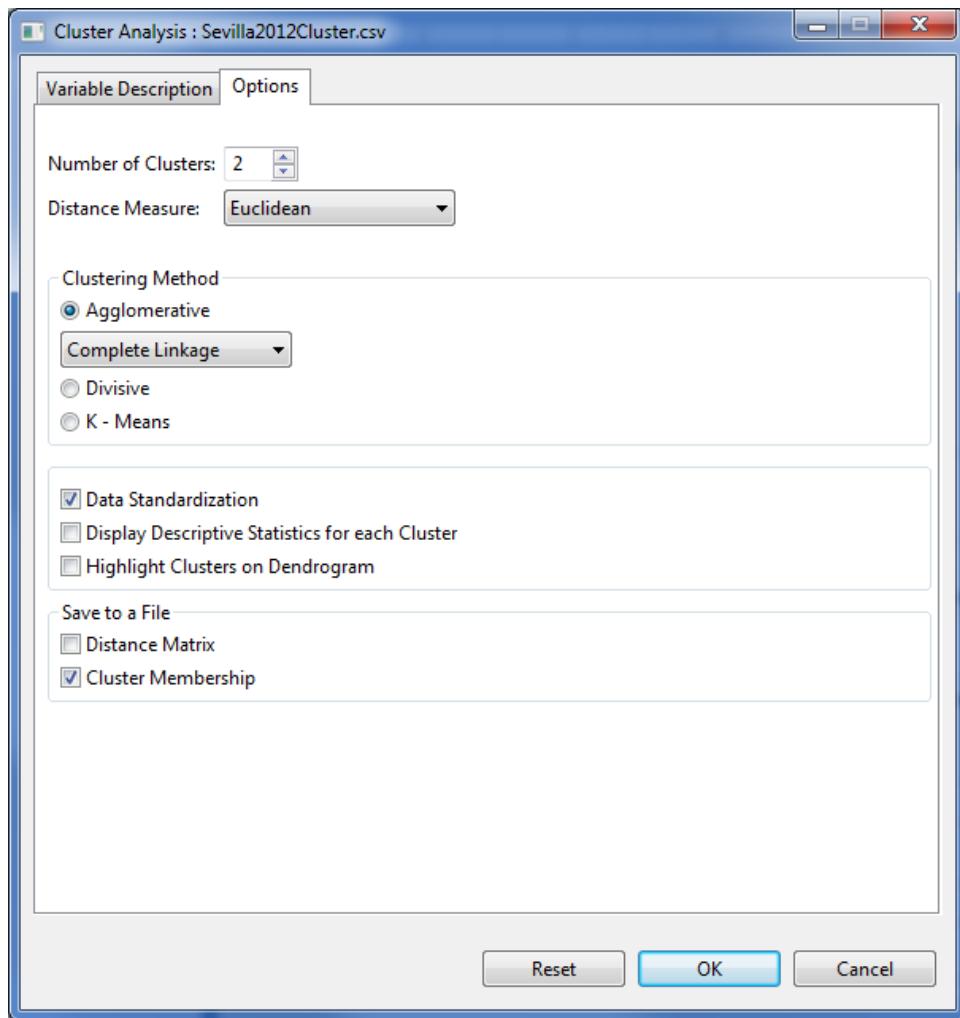
The steps to perform Multivariate Analysis of Variance are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *Sevilla2012Cluster* from the package. To import data from the package, see *Importing Data from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | Multivariate Analysis | Cluster Analysis....** The **Cluster Analysis** dialog box will appear.





- Specify the required field and appropriate options for the analysis.

Variable Description Tab

Variable(s):

For the analysis to proceed, at least one item should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this list box.

ID Variable

Only one item can be added in this field and should come from the **Factor(s)** entry box. ID Variable must have unique observations.

Symmetric Binary Variable(s)

At least one item can be entered in the list box. Only items from the **Factors(s)** list box can be added in this list box. The items should have only two levels denoted by 0 and 1.

Asymmetric Binary Variable(s)

At least one item can be entered in the list box. Only items from the **Factors(s)** list box can be added in this list box. The items should have only two levels denoted by 0 and 1.

Ordered Factor(s)

At least one item can be entered in the list box. Only items from the **Factors(s)** list box can be added in this list box.

Factors(s)

At least one item can be entered in the list box. Only items from the **Factors(s)** list box can be added in this list box.

Options Tab

Number of Clusters

This option will determine the number of clusters to be used in the analysis. By default the number of clusters is 2.

Distance Measure

The distance measure procedures will depend on the type of variables that will be used in the analysis. If the variables are numeric only, the available distance measures are *Euclidean, Maximum, Manhattan, Minkowski and Canberra*. If the variables are symmetric binary, the available distance measures are *Simple Matching, Sokal & Sneath and Hamann Coefficient*. For asymmetric binary variables, *Jaccard and Dice* are the available distance measures. If the variables selected are of different levels or data type, the *Gower* distance measure procedure will be used.

Clustering Method

There are three available types of method. The user can choose Agglomerative, Divisive or K – means. If Agglomerative is selected, the user has the option to choose the clustering method. The available methods are *Single Linkage, Complete Linkage, Average Linkage, Ward's method and Centroid method*. For K – means procedure, this will only be used if the variables are numeric only.

Data Standardization

If checked, the data to be used in the analysis will be standardized. This option is available for numeric variables only.

Display Descriptive Statistics for each Cluster

If checked, the Descriptive Statistics of the selected variables will be displayed. This option is available for numeric variables only.

Highlight Clusters on Dendrogram

If checked, the clusters will be highlighted in the Dendrogram.

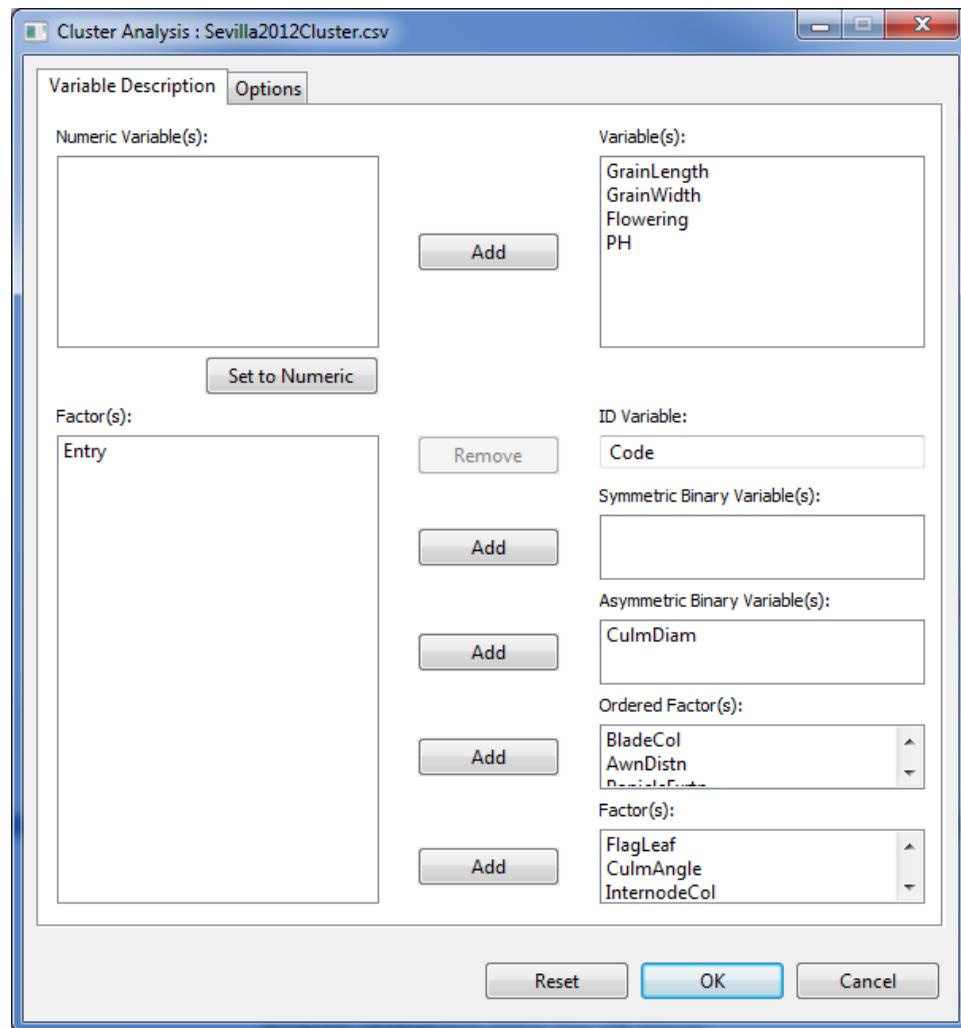
Distance Matrix

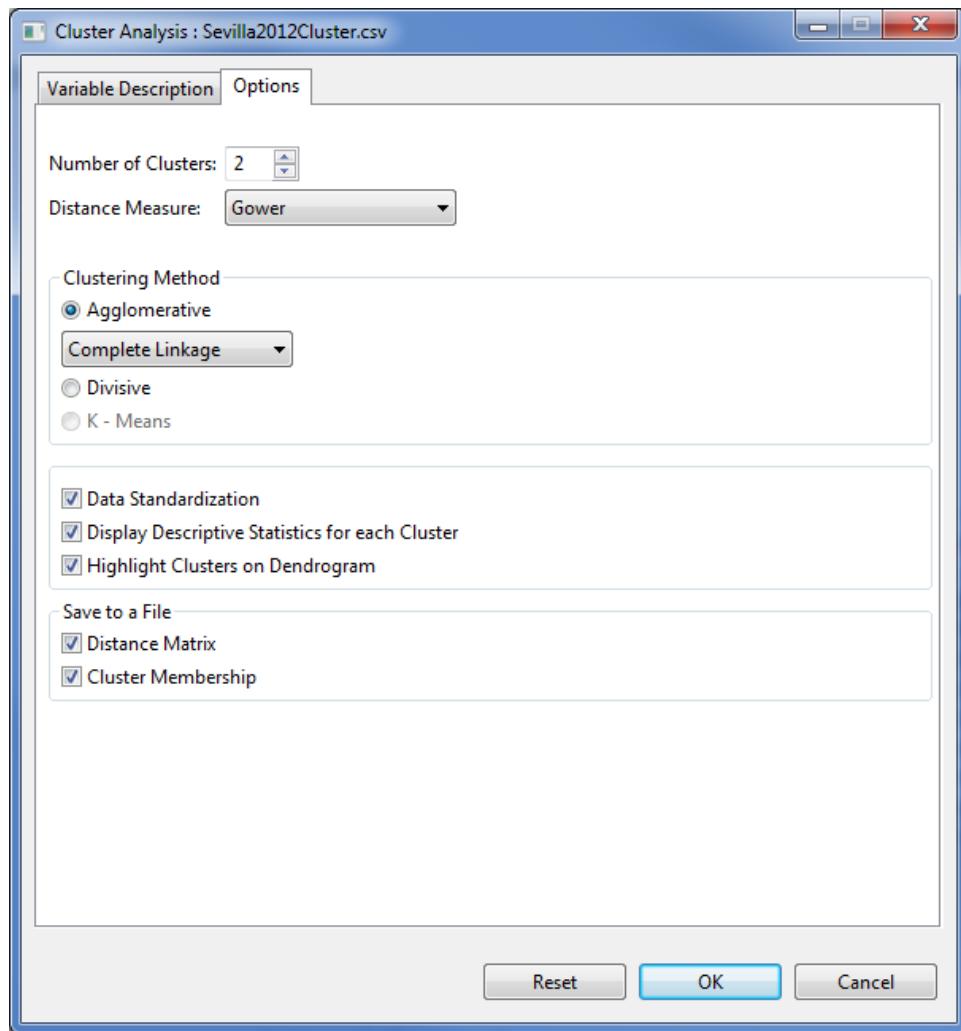
If checked, the Distance Matrix will be saved to a file.

Cluster Membership

If checked, the Cluster Membership will be saved to a file. By default, this option is selected.

For the example, the completed dialog box should appear as shown below:





- Click **OK** button to perform the analysis. The **Multivariate Analysis of Variance** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample text output displayed in the Output Viewer tab is shown below:

Number of Observations: 74

AGGLOMERATIVE CLUSTER ANALYSIS

SPECIFICATIONS:

Distance Method:	Gower
Clustering Method:	Single
Number of Clusters:	2

Descriptive Statistics

Variable	Cluster	Min	Max	Mean	StdDev
GrainLength	1	7.90	10.40	9.59	0.65
GrainLength	2	2.20	10.40	9.43	1.07
GrainWidth	1	2.40	3.00	2.77	0.15
GrainWidth	2	1.90	2.70	2.11	0.15
Flowering	1	82.00	92.00	85.21	3.12
Flowering	2	82.00	96.00	89.20	3.56
PH	1	79.20	101.20	92.27	6.49
PH	2	76.00	116.20	98.52	7.88

CLUSTER MEMBERSHIP SUMMARY

Member of Cluster 1

1 2 3 64 65 66 67 68 69 70 71 72 73 74

Member of Cluster 2

4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
34 35 36 37 38 39 40 41 42 43 44 45 46 47 48
49 50 51 52 53 54 55 56 57 58 59 60 61 62 63

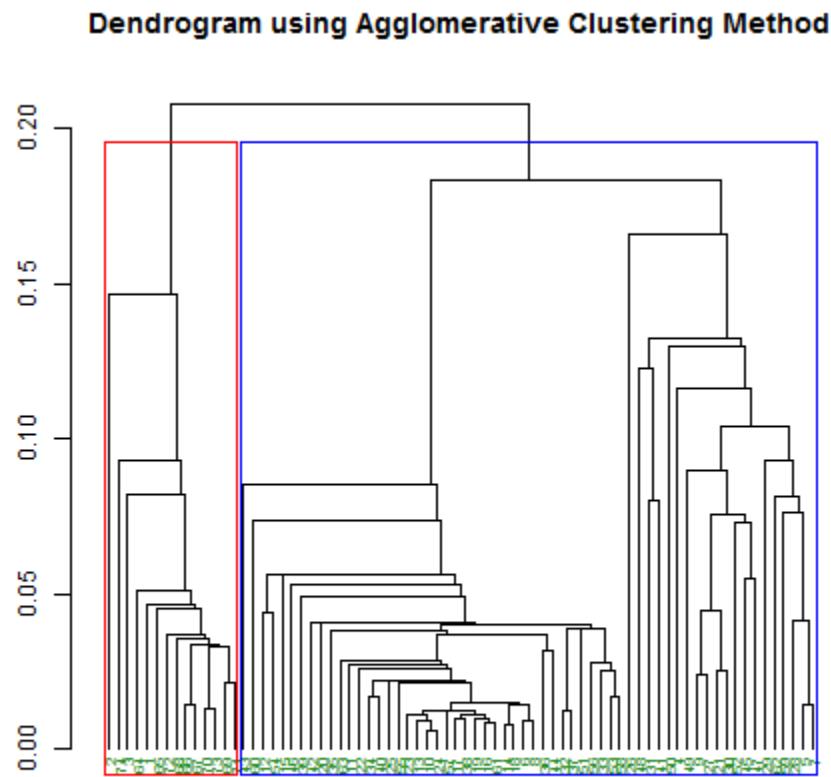
Number of members in each cluster

Cluster Size

1 14
2 60

COPHENETIC CORRELATION COEFFICIENT = 0.882

Sample graphical output displayed in the Graphical Viewer tab is shown below:



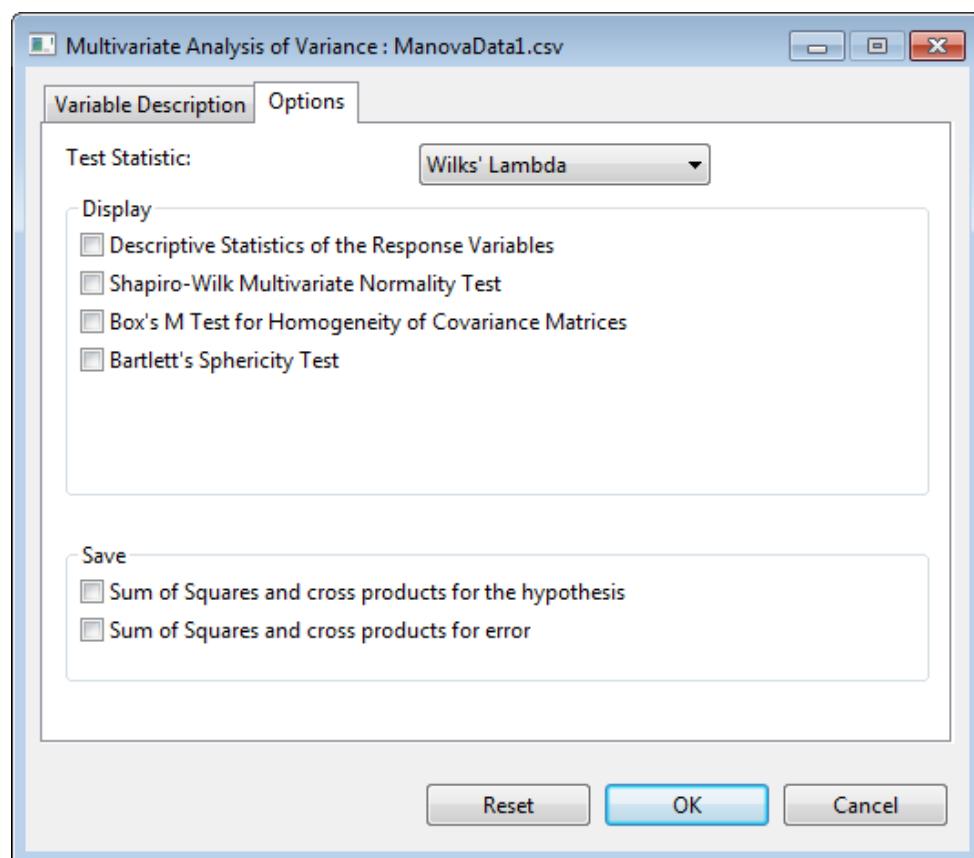
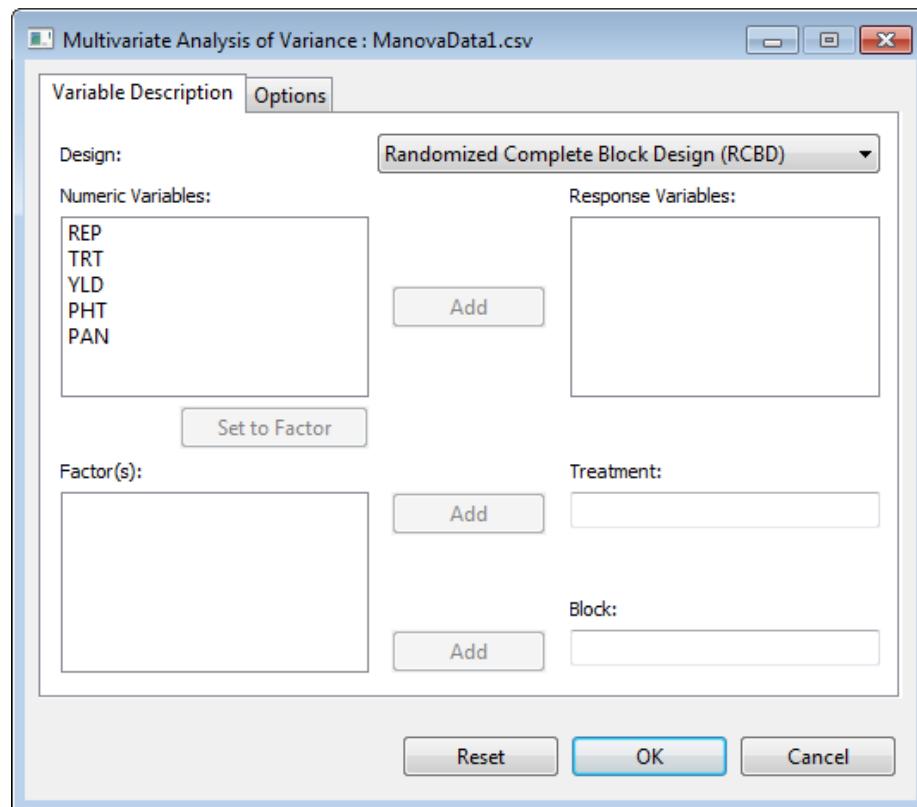
Multivariate Analysis of Variance

The steps to perform Multivariate Analysis of Variance are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *ManovaData1* from package. To import data from the package, see *Importing Data from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | MANOVA....** The **Multivariate Analysis of Variance** dialog box will appear.



- Specify the required field and appropriate options for the analysis.

Variable Description Tab

Response Variable:

For the analysis to proceed, at least two items should be entered in the list box. Only items from the **Numeric Variable(s)** can be added in this dialog box.

Treatment

For the analysis to proceed, one item from the **Factor(s)** list should be entered into this field.

Block

For the analysis to proceed, one item from the **Factor(s)** list should be entered into this field.

Options Tab

Test Statistic

By default, the test statistic used in the analysis is Wilks' Lambda. This can be changed into Pillai's Trace, Hotelling-Lawley Trace, or Roy's maximum-root.

Descriptive Statistics of the Response Variables

If this option is selected, a summary table with number of observations, minimum, maximum, mean and standard deviation of all selected numeric variables is displayed.

Shapiro-Wilk Multivariate Normality Test

If this option is selected, Shapiro-Wilk Test for Multivariate Normality will be performed.

Box's M Test for Homogeneity of Covariance Matrices

If this option is selected, Box's M Test for Homogeneity of Covariance Matrices will be performed.

Bartlett's Sphericity Test

If this option is selected, Bartlett's Sphericity Test will be performed.

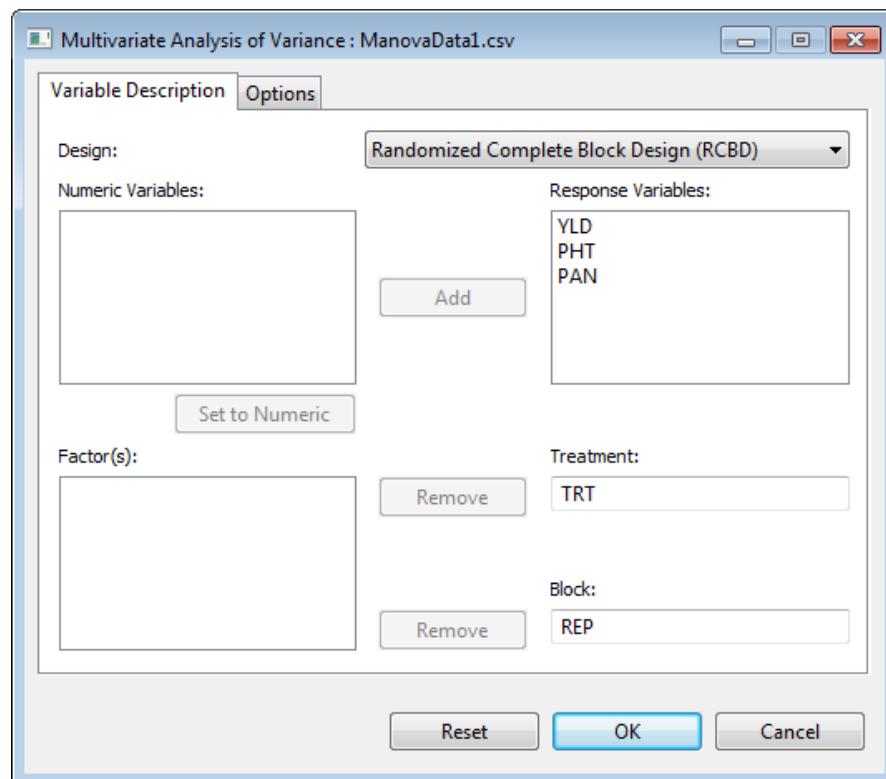
Sum of Squares and cross products for the hypothesis

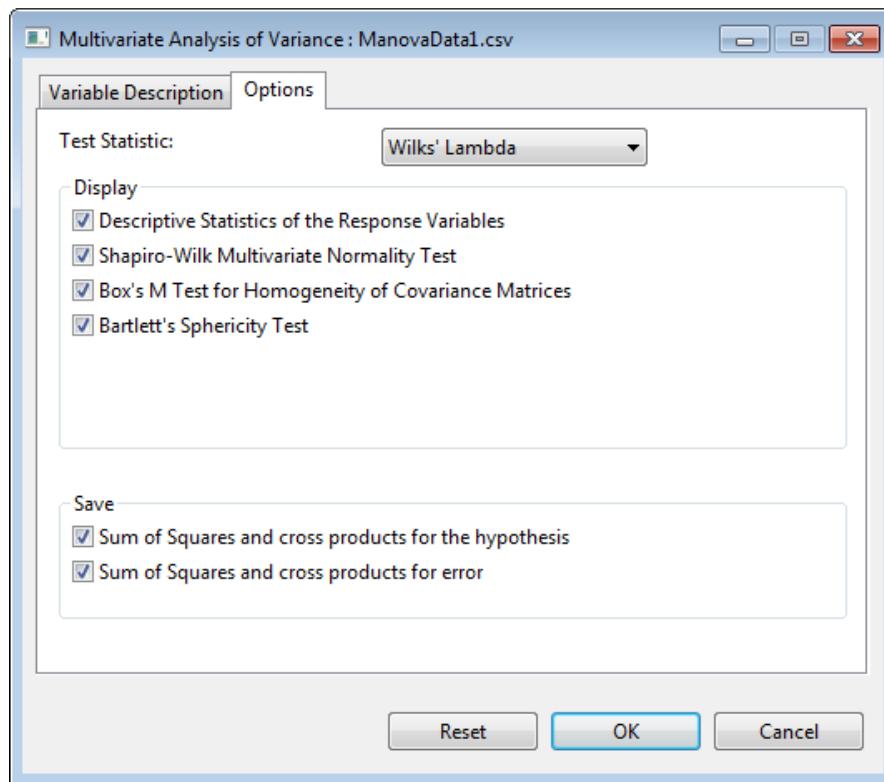
If this option is selected, sum of squares and cross products for the hypothesis will be saved to a csv file.

Sum of Squares and cross products for error

If this option is selected, error sum of squares and cross products will be saved to a csv file.

For the example, the completed dialog box should appear as shown below:





- Click **OK** button to perform the analysis. The **Multivariate Analysis of Variance** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample output is shown below:

```

Summary Information
-----
FACTOR      NO. OF LEVELS      LEVELS
-----
TRT          4                  1, 2, 3, 4
REP          4                  1, 2, 3, 4
-----
Number of Observations: 16

Descriptive Statistics
-----
Variable      N_Obs      Min      Max      Mean      StdDev
-----
YLD           16       3.52     7.85     5.37     1.19
PHT           16     107.30    137.50   119.09    9.58
PAN           16      21.80    26.40     24.32    1.52
-----
```

MULTIVARIATE NORMALITY TEST

Statistic	Value	Prob
Wilk-Shapiro	0.9729	0.8825

BOX'S M TEST FOR HOMOGENEITY OF COVARIANCE MATRICES

Box M	F	df1	df2	Prob
24.9723	0.7120	18	508	0.8001

BARTLETT'S SPHERICITY TEST

Statistic	Value	Df	Prob
Chi-Square	5.9049	3	0.1163

SUM OF SQUARES AND CROSS PRODUCTS FOR THE HYPOTHESIS

TERM: TRT

	YLD	PHT	PAN
YLD	9.32182	42.2462	4.3300
PHT	42.24619	318.6969	35.9675
PAN	4.33000	35.9675	4.9850

TERM: REP

	YLD	PHT	PAN
YLD	9.71707	-52.5421	10.5915
PHT	-52.54206	733.0869	-137.3950
PAN	10.59150	-137.3950	26.6350

SUM OF SQUARES AND CROSS PRODUCTS FOR ERROR

	YLD	PHT	PAN
YLD	2.18436	-6.84594	-1.90325
PHT	-6.84594	324.72562	14.89000
PAN	-1.90325	14.89000	2.83000

MULTIVARIATE ANALYSIS OF VARIANCE IN RCBD

	Df	Wilks	approx F	num Df	den Df	Pr (>F)
REP	3	0.004461	15.742	9	17.19	1.25e-06 ***
TRT	3	0.025898	6.659	9	17.19	0.000409 ***
Residuals	9					

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

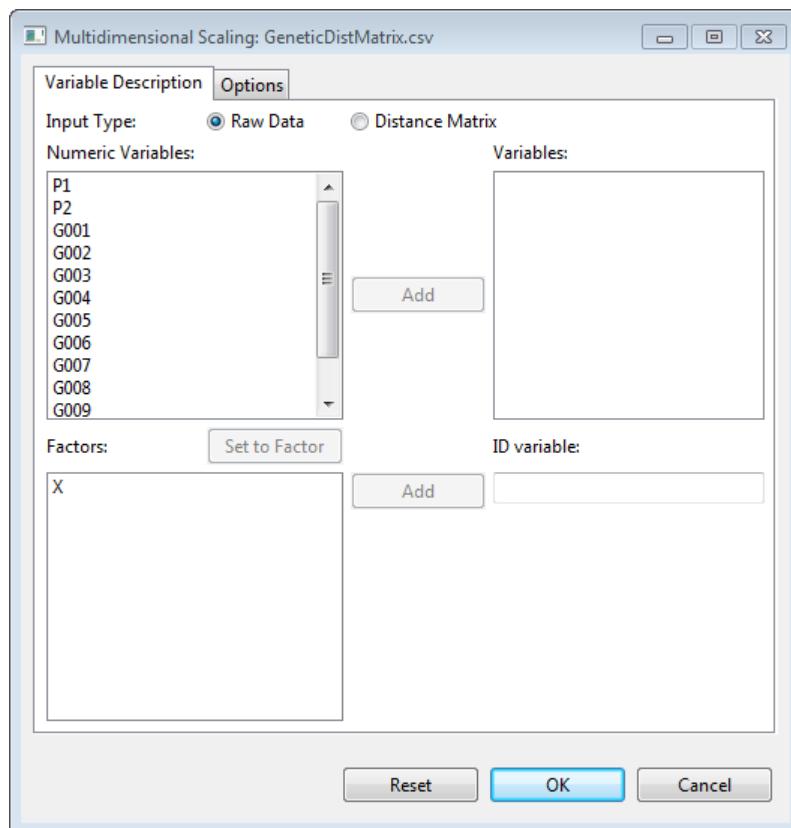
Multidimensional Scaling

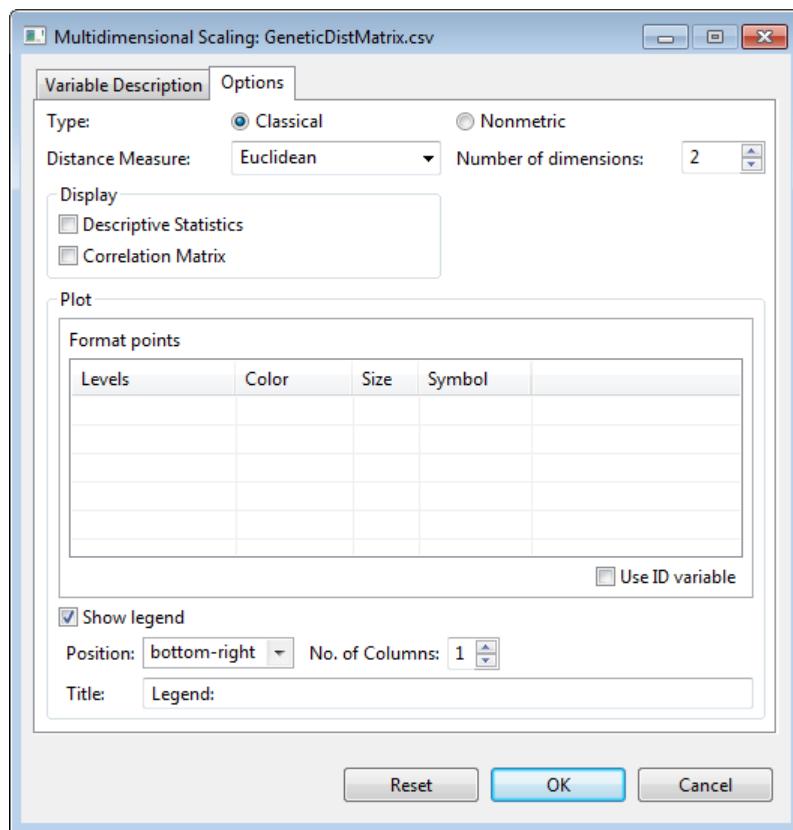
The steps to perform Multidimensional Scaling (MDS) are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data *GeneticDistMatrix* from the package. To import data from the package, see *Importing Data from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | MDS....** The **Multidimensional Scaling** dialog box will appear.





- Specify the required field and appropriate options for the analysis.

Variable Description Tab

Input Type

By default, the type of input data for MDS is raw data. If a symmetric matrix containing distances is used, the second option is selected.

Variable(s)

If input type is raw data, at least two items from the **Numeric Variable(s)** should be entered in this list box, for the analysis to proceed. Otherwise, this part is left blank.

ID Variable

If input type is raw data, an item from the **Factors** list should be entered in this list box, for the analysis to proceed. Otherwise, the levels of the ID variable are obtained automatically from the first column.

Options Tab

Type

The user can perform classical (metric) or nonmetric MDS. The default type is classical MDS.

Distance Measure

The available distance measures are: *Euclidean*, *Maximum*, *Manhattan*, *Canberra*, and *Minkowski* for both classical and nonmetric MDS. In addition, *Bray* is available for nonmetric MDS. This option is disabled if input type is distance matrix.

Number of dimensions

The allowable number of dimensions for the analysis is from two to five. By default, the number of dimensions used is 2.

Descriptive Statistics

If this option is selected, a summary table with number of observations, minimum, maximum, mean and standard deviation of all selected numeric variables is displayed. This option is disabled if input type is distance matrix.

Correlation Matrix

If this option is selected, a matrix of correlation coefficients, together with *p*-value and number of observations, of all selected numeric variables is displayed. This option is disabled if input type is distance matrix.

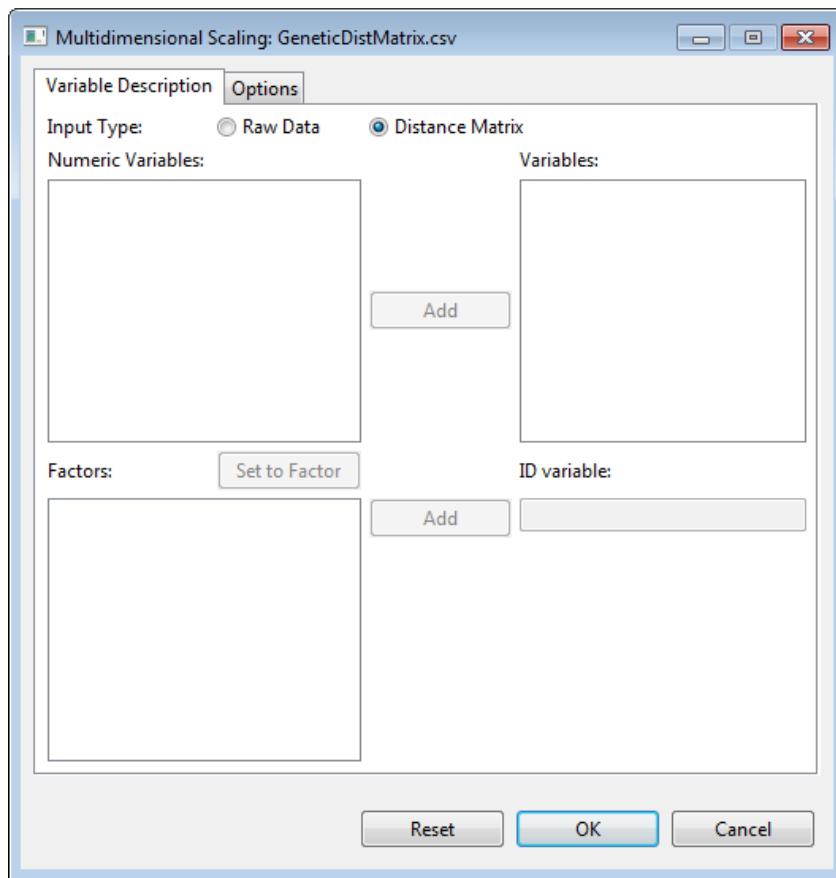
Plot

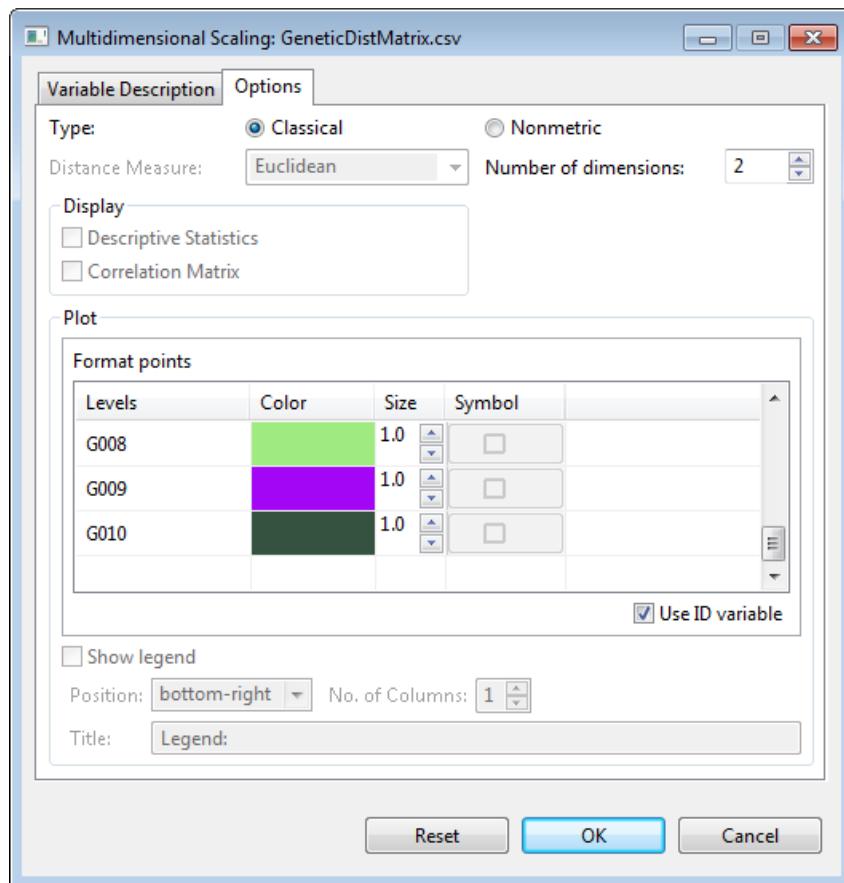
Formats of the points corresponding to the levels of the ID variable can be edited for color, size, and symbol. If the options **Use ID variable** is checked, the symbol column is disabled.

Show legend

If checked, a legend will be added to the graph. Its default title is "Legend." It may be placed in any of nine positions within the graph: bottom, bottom-left, bottom-right, center, left, right, top, top-left, top-right. The items on the legend can be listed in one to five columns.

For the example, the completed dialog box should appear as shown below:





- Click **OK** button to perform the analysis. The **Multidimensional Scaling** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample text output displayed in the Output Viewer is shown below:

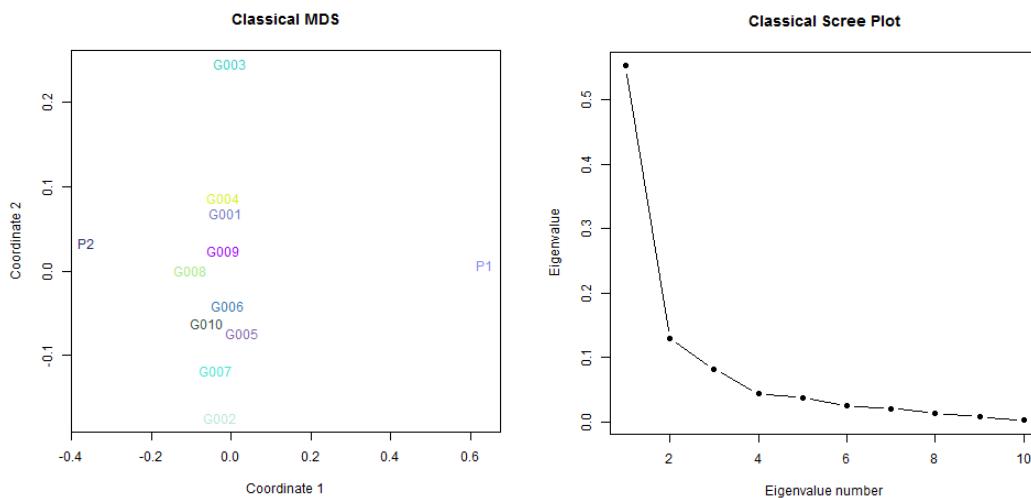
CLASSICAL MULTIDIMENSIONAL SCALING

POINTS

	MDS1	MDS2
1	0.63530755	0.00731054
2	-0.36151720	0.03349883
3	-0.01483443	0.06890020
4	-0.02852569	-0.17409966
5	-0.00402664	0.24548248
6	-0.02019590	0.08630418
7	0.02762699	-0.07304992
8	-0.00976805	-0.04107086
9	-0.04048116	-0.11783893
10	-0.10213215	0.00081293
11	-0.02093113	0.02498110
12	-0.06052219	-0.06123087

P_2 criterion: 0.7472

Sample graphical output displayed in the **Graph** Viewer is displayed below:



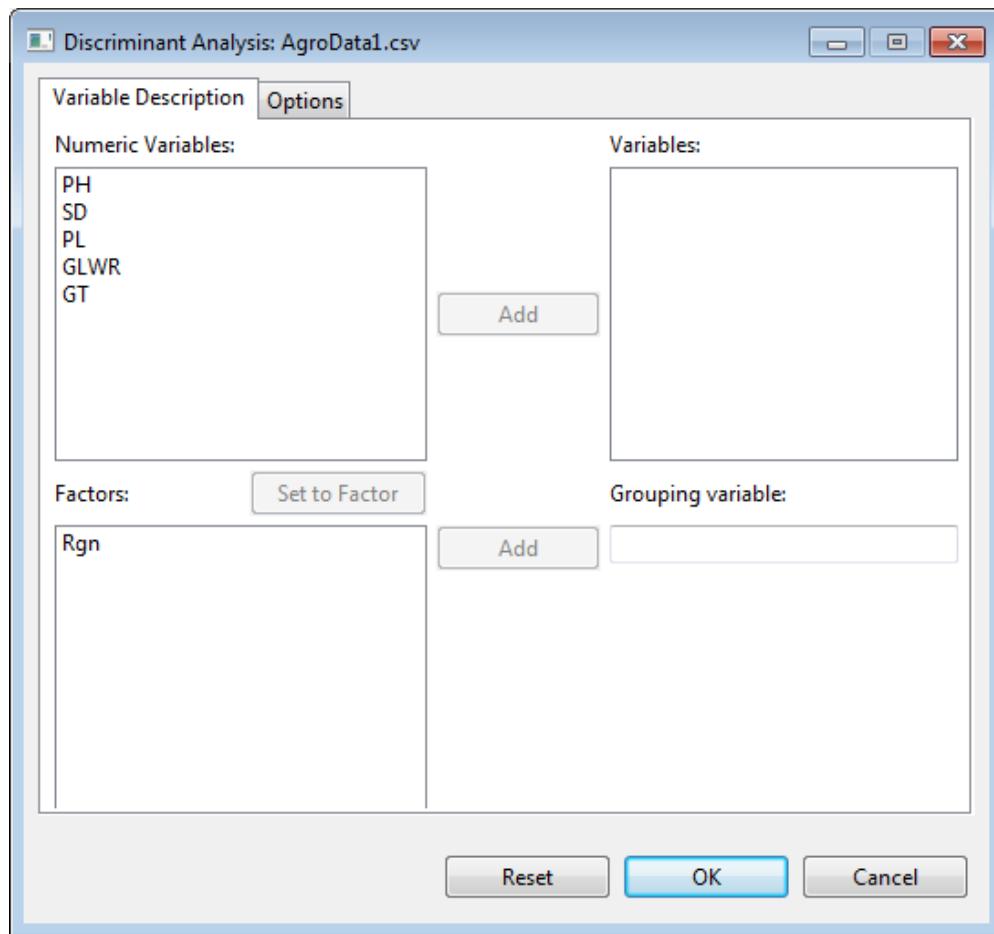
Discriminant Analysis

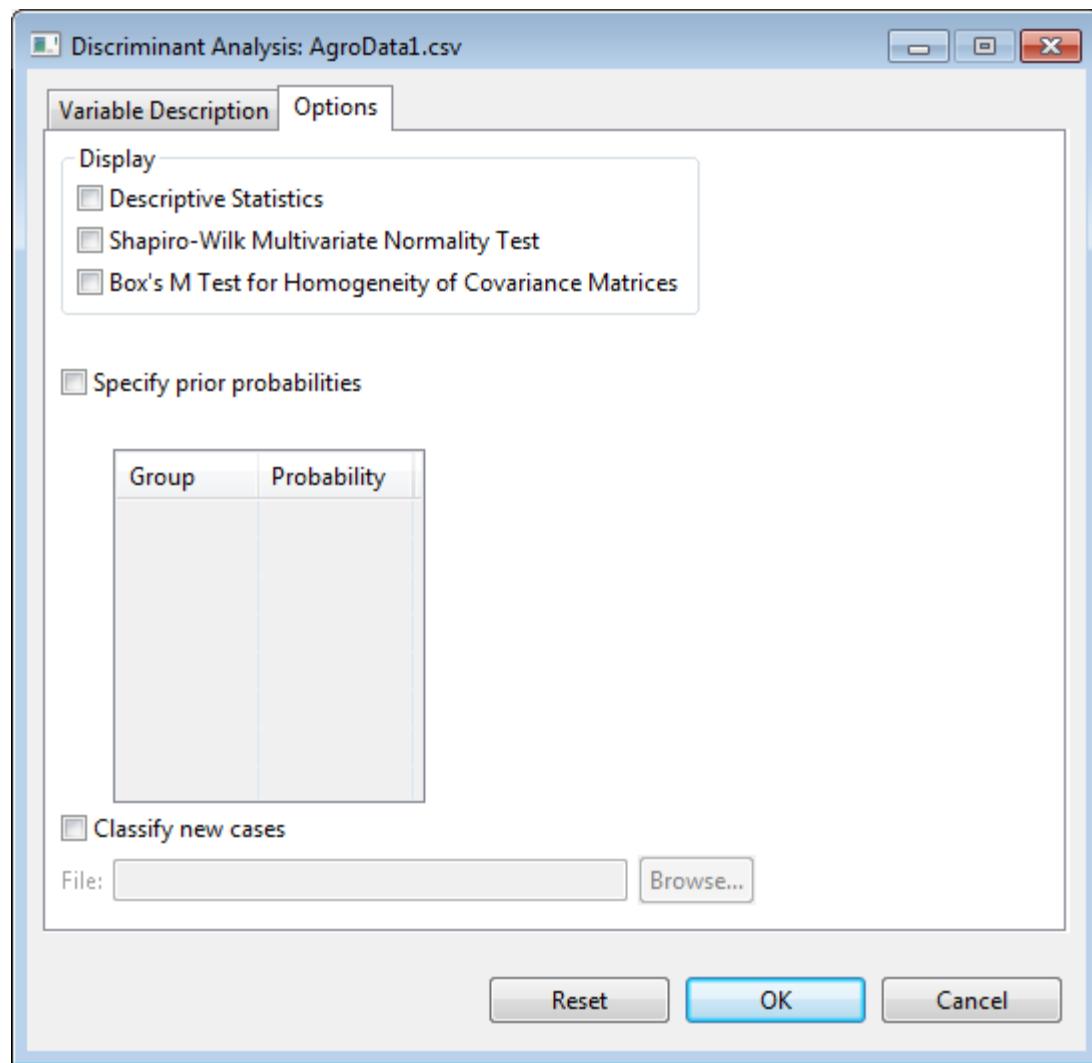
The steps to perform Discriminant Analysis are listed below:

- On the **Project Explorer**, double-click the file to be used for the analysis to open and view it in the Data Viewer.

For the example, using the project named *MyProject*, import the data files *AgroData1* and *AgroData2* from the package. To import data from the package, see *Importing Data from Package* section of this user's manual.

- From the main window of STAR, click **Analyze | Discriminant Analysis....** The **Discriminant Analysis** dialog box will appear.





- Specify the required field and appropriate options for the analysis.

Variable Description Tab

Variable(s)

For the analysis to proceed, at least two items should be entered in the list box. Only items from the **Numeric Variable(s)** can be added into this dialog box.

Grouping Variable

For the analysis to proceed, one item from the **Factors** list should be entered into this list box.

Options Tab

Descriptive Statistics

If this option is selected, a summary table with number of observations, minimum, maximum, mean and standard deviation of all selected numeric variables is displayed.

Shapiro-Wilk Multivariate Normality Test

If this option is selected, Shapiro-Wilk Test for Multivariate Normality will be performed.

Box's M Test for Homogeneity of Covariance Matrices

If this option is selected, Box's M Test for Homogeneity of Covariance Matrices will be performed.

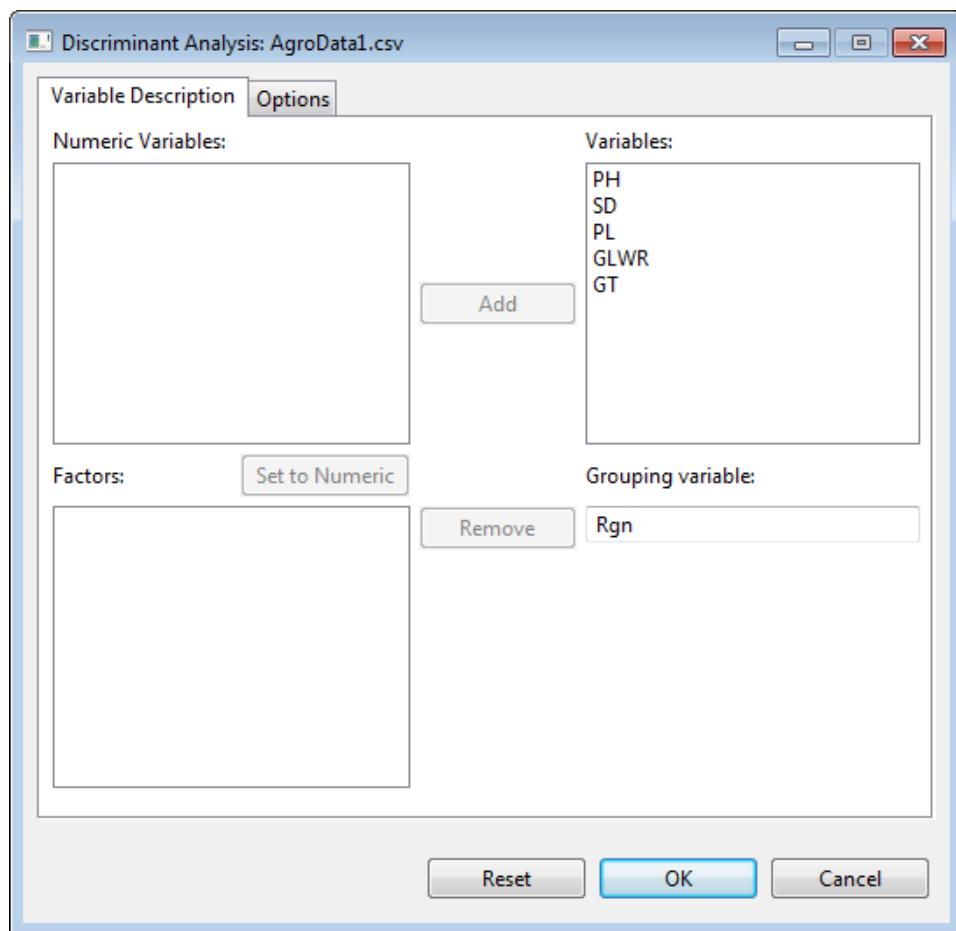
Specify prior probabilities

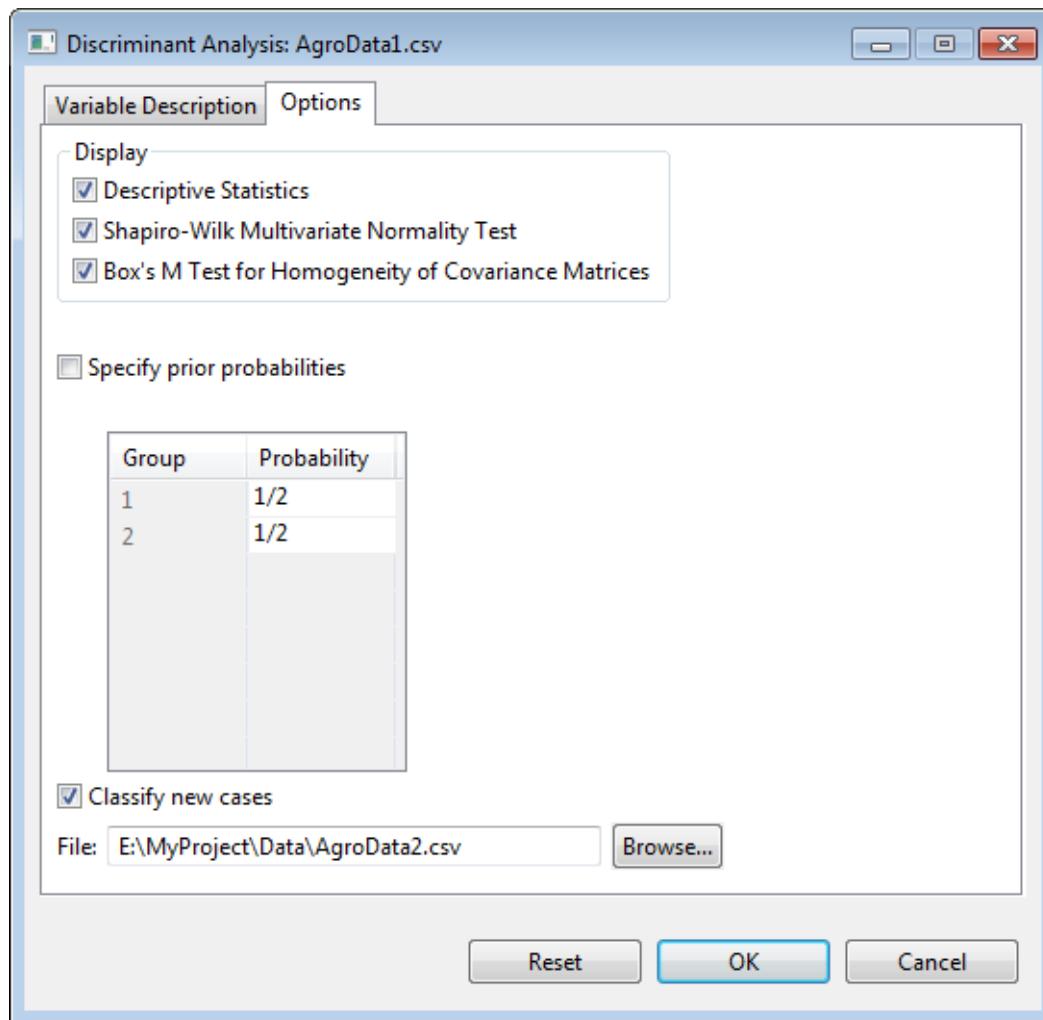
By default prior probabilities are based on the observed distribution of the grouping variable. The user can change this by selecting this option and specifying the corresponding probabilities (fraction or decimal form) in the table. By default, the prior probabilities are all equal.

Classify new cases

If the user wants to classify other cases, this option is selected. The file specified in the field should contain cases with those variables specified in the **Variable(s)** list.

For the example, the completed dialog box should appear as shown below:





- Click **OK** button to perform the analysis. The **Discriminant Analysis** dialog box will be minimized and STAR activates the **Output Viewer** tab of the **Result Viewer** tab.

Sample text output displayed in the Output Viewer is shown below:

Descriptive Statistics					
Variable	N_Obs	Min	Max	Mean	StdDev
PH	38	82.80	145.30	106.98	14.92
SD	38	3.27	8.10	5.23	1.09
PL	38	15.84	28.99	23.26	3.12
GLWR	38	2.24	4.50	3.32	0.48
GT	38	1.77	2.46	2.08	0.16

MULTIVARIATE NORMALITY TEST

```
-----  
Statistic      Value   Prob  
Wilk-Shapiro  0.9662  0.2988
```

BOX'S M TEST FOR HOMOGENEITY OF COVARIANCE MATRICES

```
-----  
Box's M Chi-Square df  Prob  
27.3195  23.2555  15  0.9211
```

DISCRIMINANT ANALYSIS

SUBGROUP MEANS

	PH	SD	PL	GLWR	GT
1	98.0333	5.0311	23.7872	3.3583	2.0433
2	115.0250	5.4160	22.7850	3.2800	2.1180

TEST OF SIGNIFICANCE OF THE DISCRIMINANT FUNCTION(S)

```
Df   Wilks approx F num Df den Df Pr(>F)  
Rgn       1 0.5883      4.48      5     32 0.00331 **  
Residuals 36  
---  
Signif. codes:  0 '****' 0.001 '***' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

DISCRIMINANT FUNCTION COEFFICIENTS

	LD1
PH	0.0806
SD	0.4547
PL	-0.0403
GLWR	-0.4968
GT	0.1055

CLASSIFICATION RESULTS

Original\Predicted (Counts)		
	1	2 Total
1	14	4 18
2	6	14 20

Original\Predicted (%)		
	1	2 Total
1	77.78	22.22 100
2	30.00	70.00 100

73.68% of original cases correctly classified

Appendix

Data from STAR Package

AgroData1

Description

This file contains hypothetical data on some agronomic traits of 38 rice lines coming from two regions.

Formats

This data contains the following columns:

Rgn	a factor with two levels (1, 2)
PH	a numeric vector denoting plant height in cm
SD	a numeric vector denoting stem diameter in mm
PL	a numeric vector denoting panicle length in cm
GLWR	a numeric vector denoting grain length/width ratio
GT	a numeric vector denoting grain thickness in mm

AgroData2

Description

This file contains hypothetical data on some agronomic traits of nine rice lines.

Formats

This data contains the following columns:

PH	a numeric vector denoting plant height in cm
SD	a numeric vector denoting stem diameter in mm
PL	a numeric vector denoting panicle length in cm
GLWR	a numeric vector denoting grain length/width ratio
GT	a numeric vector denoting grain thickness in mm

CarbonStock

Description

This file contains ecological data collected from 125 sampling locations in Eastern Region of Ghana. The objective of the study was to identify explanatory variables of carbon stocks in fallow vegetation aged 3 to 8 years, using ecological parameters that are known to influence primary productivity.

Formats

This data contains the following columns:

CS	a numeric vector of carbon stock ($MgCha^{-1}$)
SV	a numeric vector of stem volume (m^3ha^{-1})
LA	a numeric vector of leaf area (m^2kg^{-1})
SM	a numeric vector of soil moisture (%)
TN	a numeric vector of total nitrogen (%)
AP	a numeric vector of available phosphorus ($mgkg^{-1}$)
CEC	a numeric vector of cation exchange capacity ($Cmol_c kg^{-1}$)

Source

Attua, E. M.

AugLS

Description

Formats

This data contains the following columns:

Row	a factor with 15 levels (1 to 15) denoting the row
Column	a factor with 5 levels (1, 2, 3, 4, 5) denoting the column
Entry	a factor with 55 levels (1 to 55) denoting the treatment with levels 1 to 50 as the unreplicated entries and levels 51 to 55 as the replicated entries
Yield	a numeric vector of grain yield

Source

AugRCB

Description

Formats

This data contains the following columns:

BLOCK	a factor with 4 levels (1, 2, 3, 4) denoting the block
ENTRY	a factor with 55 levels (1, 2, ..., and 66) denoting the treatment with levels 1 to 60 as the unreplicated entries and levels 61 to 66 as the replicated entries
YIELD	a numeric vector of grain yield

Source

CochranBIB

Description

This file contains data of corn from an experiment conducted in 1943 in North Carolina Agricultural Experiment Station, United States Department of Agriculture which uses balanced incomplete block design. Each location or block has 4 genotypes. The blocks are planted at different location.

Formats

This data contains the following columns:

loc	a factor with 13 levels denoting the location or the block
gen	a factor with 13 levels (G01, G02, G03, G04, G05, G06, G07, G08, G09, G10, G11, G12 and G13) denoting the genotype
yield	a numeric vector of grain yield (pounds per plot)

Source

Cochran, W.G. and Cox, G.M. (1957), Experimental Designs, 2nd ed., Wiley and Sons, New York, p. 448.

Corr

Description

The data is from an experiment that was conducted to investigate the relationship between grain yield (GY14, in kg/ha), the grain's nitrogen (N), phosphorus (P) and potassium (K) content in percent. The data were collected from 48 experimental plots with grain yield adjusted to 14% moisture content.

Formats

This data contains the following columns:

GY14	a numeric vector of grain yields in kg/ha adjusted to 14% moisture content
N	a numeric vector of grain's nitrogen
P	a numeric vector of grain's phosphorus
K	a numeric vector of grain's potassium content in percent

Source

BEDDA Training

CropStatRCBDOverSite

Description

Formats

This data contains the following columns:

NTRT	a factor with 6 levels (1, 2, 3, 4, 5, 6) denoting the nitrogen
SITE	a factor with 3 levels (1, 2, 3) denoting the site
REP	a factor with 3 levels (1, 2, 3) denoting the block
YIELD	a numeric vector denoting the grain yield

Source

CropStat Manual

Dule_AlphaGxE

Description

Formats

This data contains the following columns:

Env	a factor with 6 levels (A01WSCM, A03WS-0, A03WS-1, AWS01-D, UWS02-2, UWS02-8) denoting the environment
Rep	a factor with 3 levels (1, 2, 3) denoting the replication
Block	a factor with 5 levels (1, 2, 3, 4, 5) denoting the block
Plot	a numeric vector denoting the plot number
Geno	a factor with 40 levels (AUS196, AZUCENA, B6144F-MR-6-0-0, C22, CT13370-12-2-M, CT13377-4-2-M, CT13382-8-3-M, CT6510-24-1-2, CT6516-24-3-2, DINORADO, IR47686-30-3-2, IR55419-04, IR55423-01, IR60080-46A, IR65261-09-1-B, IR65907-116-1-B, IR66417-18-1-1, IR66421-062-1-1-2, IR66424-1-2-1-5, IR68702-072-1-4-B, IR70358-84-1-1, IR70360-38-1-B-1, IR71524-44-1-1, IR71525-19-1-1, IR72768-15-1-1, IRAT170, IRAT177, IRAT212, IRAT216, MARAVILHA, PALAWAN, PRIMAVERA, UPLRI-5, UPLRI-7,

	VANDANA, WAB181-18, WAB56-125, WAB638-1, WAB96-1-1, WAYRAREM)
FLOWER	a numeric vector
Duration	a numeric vector
PlantHt	a numeric vector
CropBF	a numeric vector
HI	a numeric vector
YLD	a numeric vector

Source

Dune

Description

This file contains 20 rows and 31 columns.

Formats

This data contains the following columns:

Species	a numeric vector of
Belper	a numeric vector of
Empnig	a numeric vector of
Junbuf	a numeric vector of
Junart	a numeric vector of
Airpra	a numeric vector of
Elepa	a numeric vector of
Rumace	a numeric vector of
Viclat	a numeric vector of
Brarut	a numeric vector of
Ranfla	a numeric vector of
Cirarv	a numeric vector of
Hyprad	a numeric vector of

Leayout	a numeric vector of
Potpal	a numeric vector of
Poapra	a numeric vector of
Calcus	a numeric vector of
Tripra	a numeric vector of
Trirep	a numeric vector of
Antodo	a numeric vector of
Salrep	a numeric vector of
Achmil	a numeric vector of
Poatri	a numeric vector of
Chealb	a numeric vector of
Elyrep	a numeric vector of
Sagpro	a numeric vector of
Plalan	a numeric vector of
Agrsto	a numeric vector of
Lolper	a numeric vector of
Alogen	a numeric vector of
Brohor	a numeric vector of

Source

Fertilizer

Description

Formats

This data contains the following columns:

Fertilizer	a factor with two levels (A and B) denoting the fertilizer used
NumFruit	a numeric vector of number of fruits

Source

fishData

Description

The file contains stock-recruitment data of hake (*Merluccius merluccius*) for 15 years.

Formats

This data contains the following columns:

num.fish a numeric vector of the number of fish in millions

spawn.biomass a numeric vector of spawning biomass in thousand tonnes

Source

M.merluccius data of the *nlrwr* package in R

GeneticDistMatrix

Description

This file contains a genetic distance matrix for genotypes $P_1, P_2, G_{001}, G_{002}, \dots, G_{010}$

Formats

This is a symmetric matrix containing genetic distances for genotypes $P_1, P_2, G_{001}, G_{002}, \dots, G_{010}$. First column contains the labels for the rows ($P_1, P_2, G_{001}, G_{002}, \dots, G_{010}$). Columns 2 to 13 have these labels as headings in the same order.

Source

Introduction to Multivariate Analysis Training

Gerua

Description

An experiment was conducted to determine the effect of two transplanting method using 26 cultivars on the grain yield, plant height and panicle number. The experiment was conducted using split plot design with transplanting method as the mainplot factor and cultivar as the subplot factor, replicated three times.

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and 3) denoting the blocks
TPL	a factor with two levels (D and N) denoting the transplanting method; considered as the main plot factor
CULT	a factor with 26 levels denoting the cultivar; considered as the subplot factor
DFF	a numeric vector denoting the number of days after flowering
PLHT	a numeric vector denoting the plant height in cm
PNCLE	a numeric vector denoting the panicle number
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training

GomezMaizeCross

Description

"A plant breeder is studying a cross between a sweet maize inbred line with yellow kernels and a flint maize inbred line with white kernels. He would like to know whether the ratio of kernel type and color in the F2 population follows the normal di-hybrid ratio of 9:3:3:1. From the F1 plants produced by crossing the two inbred lines, he obtain F2 kernels and classifies them into four categories according to kernel color (yellow or white) and kernel type (flint or sweet) as follows: yellow flint, yellow sweet, white flint and white sweet." -- Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 460.

Formats

This data contains the following columns:

KernelNumber	a numeric vector denoting the kernel id number
KernelColor	a factor with 2 levels (white and yellow) denoting the color of the kernel
KernelType	a factor with 2 levels (flint and sweet) denoting the kernel type
KernelCategory	a factor with 4 levels (white flint, white sweet, yellow flint, and yellow sweet) denoting the kernel type and color
YellowKernel	a numeric vector of values 0 and 1, with 1 denoting that the kernel is of color yellow
WhiteKernel	a numeric vector of values 0 and 1, with 1 denoting that the kernel is of color white
FlintType	a numeric vector of values 0 and 1, with 1 denoting that the kernel is of type flint
SweetType	a numeric vector of values 0 and 1, with 1 denoting that the kernel is of type sweet

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 460.

GomezMaizeCrossSummary

Description

This data is a summary of *GomezMaizeCross*.

Formats

This data contains the following columns:

KernelCategory	a factor with 4 levels (white flint, white sweet, yellow flint, and yellow sweet) denoting the kernel type and color
Frequency	a numeric vector denoting the number of observation in that follows in that kernel category

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 460.

GomezSplitRCBD

Description

Grain Yield of Data of Four Rice Varieties Grown in Six Levels of Nitrogen in a Split-Plot Design with three replications.

Formats

This data contains the following columns:

Rep	a factor with 3 levels (1 to 3) denoting the blocks
Nitrogen	main plot factor with 6 levels (0, 60, 90, 120, 150 and 180, in kg N/ha) denoting the Nitrogen Levels
Variety	subplot factor with 4 levels (C4-63, IR5, IR8 and Peta) denoting the Rice Variety
Yield	a numeric vector of grain yield in kg/ha

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 102.

GomezSplit2RCBD

Description

Grain Yield of Three Rice Varieties Grown under Three Management Practices and Five Nitrogen Levels; in a split-split plot design in RCBD with three replications.

Formats

This data contains the following columns:

Rep	a factor with 3 levels (1 to 3) denoting the blocks
Nitrogen	main plot factor with 5 levels (N1, N2, N3, N4, and N5) denoting the Nitrogen Levels
N1	0 kg N/ha
N2	50 kg N/ha

	N3	80 kg N/ha
	N4	110 kg N/ha
	N5	140 kg N/ha
Management	subplot factor with 3 levels (M1, M2 and M3) denoting the Management Practices	
	M1	Minimum
	M2	Optimum
	M3	Intensive
Variety	sub-subplot factor with 3 levels (V1, V2 and V3) denoting the Rice Variety	
GY	a numeric vector of grain yield in t/ha	

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 143.

GomezStrip

Description

Grain Yield of Six Varieties of Rice, Broadcast Seeded and Grown with Three Nitrogen Rates in a Strip-plot Design with Three Replications

Formats

This data contains the following columns:

REP	a factor with 3 levels (1 to 3) denoting the blocks	
NITROGEN	vertical factor with 3 levels (0, 60 and 120, kg N/ha) denoting the Nitrogen Levels	
VARIETY	horizontal factor with 6 levels (V1, V2, V3, V4, V5 and V6) denoting the Rice Variety	
	V1	IR8
	V2	IR127-80
	V3	IR305-4-12

V4	IR400-2-5
V5	IR665-58
V6	Peta
GRNYLD	a numeric vector of grain yield in kg/ha

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 110.

GomezStripSplit

Description

Grain yield of six rice varieties tested under two planting methods and three nitrogen rates in a Strip-Split plot Design with three replicates.

Formats

This data contains the following columns:

REP	a factor with 3 levels (1 to 3) denoting the blocks
NITROGEN	vertical factor with 3 levels (0, 60 and 120, kg N/ha) denoting the Nitrogen Levels
VARIETY	horizontal factor with 6 levels (V1, V2, V3, V4, V5 and V6) denoting the Rice Variety
V1	IR8
V2	IR127-80
V3	IR305-4-12
V4	IR400-2-5
V5	IR665-58
V6	Peta
PLANTING	subplot factor with two levels (Broadcast and Transplanted) denoting the planting methods
YIELD	a numeric vector of grain yield in kg/ha

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 155.

GomezTenureStatus

Description

"An agricultural economist studying factors affecting the adoption of the newly introduced high-yielding rice varieties wishes to know if adoption is affected by the tenure status of farmers. With three distinct classes of tenure status (the first classification criterion) – owner operator, share-rent farmer, and fixed-rent farmer – and two classes of adoption status (the second classification criterion) – adopter and nonadopter – the resulting data form a 3 x 2 contingency table." -- Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 462.

Formats

This data contains the following columns:

TenureStatus a factor with 3 levels (fixed-rent, owner, and share-rent)
FarmerClassif a factor with 2 levels (adopter and nonadopter)

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 463.

GomezTwoCropSeasonRCB

Description

Grain yield of rice tested with five rates of nitrogen in two crop seasons.

Formats

This data contains the following columns:

Season	a factor with 2 levels (Dry and Wet) denoting the season
Rep	a factor with 3 levels (I, II and III) denoting the block
NitrogenRate	a factor with 5 levels (0, 60, 90, 120 and 150 kg/ha) denoting the Nitrogen Rate
GrainYield	a numeric vector of grain yield, t/ha

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 318.

grp_pss

Description

Three (3) varieties, a traditional variety (A) and two new varieties B and C, were tested to determine if their yield means are equal. Analysis of variance showed that difference among the varieties exists. The researcher would like to know the nature of the differences among the groups.

Formats

This data contains the following columns:

VARIETY	a factor with three levels denoting the variety
REP	a factor with four levels (1 to 4) denoting block
YIELD	a numeric vector of grain yield

Source

IASRICorrReg

Description

The data was collected through a pilot survey on Hybrid Jowar crop on yield and biometrical characters.

Formats

This data contains the following columns:

sampleNum	a numeric vector denoting the identification number of the sample
PP	a numeric vector denoting the average plant population
PH	a numeric vector denoting the average plant height
NGL	a numeric vector denoting the average number of green leaves
Yield	a numeric vector denoting the yield (kg/plot)

Source

Design Resources Server. *Indian Agricultural Statistics Research Institute (ICAR), New Delhi 110 012, India.* www.iasri.res.in/design (accessed lastly on December 20, 2012).

IASRILSD

Description

An experiment was conducted at Agricultural Research Station, Kopurgaon, Maharashtra on cotton during the year 1969 – 1970 using a Latin Square design to study the effects of foliear application of urea in combination with insecticidal sprays on the cotton yield.

Formats

This data contains the following columns:

row	a factor with six levels (1 to 6) denoting the row
col	a factor with six levels (1 to 6) denoting the column
trt	a factor with six levels denoting the treatment
T1	Control (No nitrogen and no insecticides)
T2	100 kg N/ha as urea (half at final thinning and half at flowering as top dressing)
T3	100 kg N/ha as urea (80 kg N/ha in 4 equal split does as spray and 20 kg N/ha at final thinning)
T4	100 kg N/ha as CAN (half at final thinning and half at flowering as top dressing)

T5	T2 + six insecticidal sprays
T6	T4 + six insecticidal sprays
yield	a numeric vector denoting the yield in kg

Source

Design Resources Server. *Indian Agricultural Statistics Research Institute (ICAR)*, New Delhi 110 012, India. www.iasri.res.in/design (accessed lastly on December 20, 2012).

IASRITomar

Description

An experiment was conducted to study compare the effect of natural pollination and hand pollination under open field conditions.

Formats

This data contains the following columns:

group	a factor with two levels (1 and 2) denoting the type of pollination used
	1 natural pollination
	2 hand pollination
fs45	a numeric vector of number of fruit set for the period of 45 days
fw	a numeric vector of fruit weight in kilograms (kg)
syp	a numeric vector of seed yield per plant in grams
sl	a numeric vector of seedling length in cm

Source

Tomar, B.S. Descriptive Statistics: Design Resources Server. *Indian Agricultural Statistics Research Institute (ICAR)*, New Delhi 110 012, India. www.iasri.res.in/design (accessed lastly on December 20, 2012).

IASRITomar2

Description

An experiment was conducted to study compare the effect of hand pollination on hybrid of seed production in bottle gourd under open field and poly house conditions.

Formats

This data contains the following columns:

Group	a factor with two levels (2 and 3) denoting the type of conditions
	2 hand pollination in open field condition
	3 hand pollination in poly house condition
nmfp	a numeric vector of total number of male flowers per plant
fs45	a numeric vector of number of fruit set for the period of 45 days
fweight	a numeric vector of fruit weight in kilograms (kg)
flength	a numeric vector of fruit length in centimeter (cm)
slyp	a numeric vector of seed yield per plant in grams
sl	a numeric vector of seedling length in centimeter

Source

Tomar, B.S. Descriptive Statistics: Design Resources Server. *Indian Agricultural Statistics Research Institute (ICAR)*, New Delhi 110 012, India. www.iasri.res.in/design (accessed lastly on December 20, 2012).

InjuryProfile

Description

A study was conducted to assess harmful effects of pests termed as injury variables. The primary concern of the study is to determine the possibility of yield-reducing effects.

Formats

This data contains the following columns:

SITE	a factor with 5 levels (A, B, C, D and E) denoting the site
------	---

FIELD_NO	a numeric vector of field number
Number	a numeric vector of field number
VAR	a factor with 4 levels (HYB1, HYB2, MV1 and MV2) denoting the variety
DH	a numeric vector of deadheart
WH	a numeric vector of whitehead
WM	a numeric vector of whorl maggot
LF	a numeric vector of leaffolder
RT	a numeric vector of rice tungro disease
BS	a numeric vector of brown spot
SHB	a numeric vector of sheath blight
NB	a numeric vector of neck blast
LB	a numeric vector of leaf blast
BLB	a numeric vector of bacterial leaf blast
WA	a numeric vector of weeds above canopy
WB	a numeric vector of weeds below canopy

Source

Inoculation

Description

An experiment was conducted on the effect of 6 times of inoculation of *S. linicola* on the oil content of Redwing Flaxseed. The experiment was conducted using an RCB design with 4 replications.

Formats

This data contains the following columns:

Rep	a factor with three levels (1, 2 and 3) denoting the blocks
InoculationTime	a factor with six levels (Seedling, Early Bloom, Full Bloom, Full Bloom (1/100), Ripening and Uninoculated) denoting the time of inoculation

OilContent a numeric vector denoting the oil content

Source

Steel, R.G.D., J.H. Torrie, and D.A. Dickey. Principles and Procedures of Statistics: A Biometrical Approach.

Insecticides

Description

Grain yield of rice resulting from use of different foliar and granular insecticides for the control of brown planthoppers and stem borers from a Completely Randomized Design Experiment with Four Replicates and 7 treatments. The data contains 28 rows and 3 columns.

Formats

This data contains the following columns:

Treatment	a factor with 7 levels (Azodin, DDT + γ -BHC, Dimercon-Boom, Dimercon-Knap, Dol-Mix (1 kg), Dol-Mix (2 kg) and Control) denoting the Different Foliar and Granular Insecticides
Rep	a factor with 4 levels (1 to 4) denoting the replicates
GrainYield	a numeric vector of grain yield in kg/ha

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 14.

JohnAlpha

Description

A spring oats trial grown in Craibstone, near Aberdeen was conducted with 24 varieties in 3 replicates, each consisting of 6 blocks of 4 plots. The objective is to determine the performance of the 24 new varieties.

Formats

This data contains the following columns:

plot	plot number
rep	a factor with 3 levels (R1, R2, R3) denoting the replicate
block	a factor with 6 levels (B1, B2, B3, B4, B5 and B6) denoting the block
gen	a factor with 24 levels (G01, G02, G03, G04, G05, G06, G07, G08, G09, G10, G11, G12, G13, G14, G15, G16, G17, G18, G19, G20, G21, G22, G23, G24) denoting the genotype (variety)
yield	a numeric vector denoting the dry matter yield in tones/ha

Source

John JA, Williams ER. (1995) Cyclic and computer generated designs, Chapman and Hall, London. Page 146.

R agricidat package

KemptonRowCol

Description

Wheat yields of 35 genotypes, 2 reps, planted in a row-column design. Two plots are missing.

Formats

This data contains the following columns:

rep	a factor with 2 levels (R1 and R2) denoting the replicate
row	a factor with 5 levels (1 to 5) denoting the row
col	a factor with 7 levels (1 to 7) denoting the column
gen	a factor with 35 levels (G01, G02, G03, G04, G05, G06, G07, G08, G09, G10, G11, G12, G13, G14, G15, G16, G17, G18, G19, G20, G21, G22, G23, G24, G25, G26, G27, G28, G29, G30, G31, G32, G33, G34, G35) denoting the genotype
yield	a numeric vector denoting the yield

Source

R A Kempton and P N Fox, Statistical Methods for Plant Variety Evaluation, Chapman and Hall, 1997.

R agridat package.

KijojiDS2010Cluster

Description

Trait means of the 201 genotypes from an experiment conducted in IRRI upland during dry season of 2010.

Formats

This data contains the following columns:

GENO	a factor with 202 levels denoting the genotype
DTF	a numeric vector of days after flowering
YIELD	a numeric vector of mean grain yield in g/m ²
PK	a numeric vector of in %
SBM	a numeric vector of in g/m ²
HI	a numeric vector of harvest index
PH	a numeric vector of plant height
TILLER	a numeric vector of grain yield in kg/ha

Source

Kijoji, Abubakari

Maize

Description

Grain yield of three promising maize hybrids and a check variety from an experiment using Latin Square design. The data contains 24 rows and 3 columns.

Formats

This data contains the following columns:

Maize	a factor with 4 levels denoting the Maize Hybrids (A, B, and D) and a Check Variety (C)
Row	a factor with 4 levels (1 to 4) denoting the row blocking factor
Column	a factor with 4 levels (1 to 4) denoting the column blocking factor
GrainYield	a numeric vector of grain yield in kg/ha

Source

Gomez, Kwanchai A. and Arturo A. Gomez. (1984) *Statistical Procedures for Agricultural Research 2nd Edition*. John Wiley & Sons, Inc. Page 33.

ManovaData1

Description

Yield and other characteristics obtained from an experiment using Randomized Complete Block design. The data contains 16 rows and 5 columns.

Formats

This data contains the following columns:

REP	a factor with four levels (1, 2, 3 and 4) denoting the blocks
TRT	a factor with four levels (1, 2, 3 and 4) denoting the treatment
YLD	a numeric vector denoting yield in tons/ha
PHT	a numeric vector denoting plant height in cm
PAN	a numeric vector denoting panicle length in cm

Source

Introduction to Multivariate Analysis Training

Nurul2012

Description

Formats

This data contains the following columns:

Block	a factor with levels I II
Variety	a factor with levels Azucena112854 Azucena47125 Azucena52992 Chadungda IR20 IR64 IR74 Pokalli
Aluminum	a factor with levels 0 540 750 1300
RL	a numeric vector
RE	a numeric vector
RRE	a numeric vector

Source

OP_pss

Description

Grain yield of three promising maize hybrids and a check variety from an experiment using Latin Square design. The data contains 24 rows and 3 columns.

Formats

This data contains the following columns:

NITROGEN	a factor with levels 0 60 90 120
REP	a factor with levels 1 2 3
YIELD	a numeric vector

Source

PGTYPIC

Description

Formats

This data contains the following columns:

Block	a factor with levels 1 2 3 4
Geno	a factor with levels 1 2 3 4 5 6 7 8
X1	a numeric vector
X2	a numeric vector
X3	a numeric vector
X4	a numeric vector

Source

SeedingRate

Description

A hypothetical data of grain yield in Randomized Complete Block Design.

Formats

This data contains the following columns:

Rep	a factor with levels 1 2 3 4
SeedingRate	a factor with levels 25 50 75 100 125 150
GrainYield	a numeric vector

Source

Sensors

Description

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and B) denoting the blocks
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training

Sevilla2012Cluster

Description

A multivariate data gathered from the entries of the Green Super Rice (GSR) nursery trials that underwent seed increase before distribution to target countries. Each entry were characterized based on agromorphological traits using the new rice descriptors published by the Bioversity International. The data contains 74 rows and 21 columns.

Formats

This data contains the following columns:

Entry	a factor with 202 levels denoting the genotype
Code	a numeric vector of __
Grain_Length	a numeric vector of grain length
Grain_Width	a numeric vector of grain width
Flowering	a numeric vector of __
PH	a numeric vector of __
Blade_Col	a numeric vector of __

Flag_Leaf	a numeric vector of __
Culm_Angle	a numeric vector of __
Awn_Distn	a numeric vector of __
Panicle_Exrtn	a numeric vector of __
Panicle_Axis	a numeric vector of __
Internode_Col	a numeric vector of __
Lodging_Resist	a numeric vector of __
Shattering	a numeric vector of __
Leaf_Senscene	a numeric vector of __
Leaf_Length	a numeric vector of leaf length
Leaf_Width	a numeric vector of leaf width in __
Culm_Length	a numeric vector of __
Culm_Diam	a numeric vector of __
Panicle_Length	a numeric vector of panicle length in __

Source

Sevilla, Maria Anna Lyn <verify from lei>

soil_redox

Description

The file contains the data of number of days after seeding (DAS) as independent variable and soil redox potential (EH) as dependent variable. Based on this data, simple linear regression analysis is to be conducted to assess how the days after seeding linearly affects soil redox potential, on the average.

Formats

This data contains the following columns:

DAS	a numeric vector denoting the number of days after seeding
EH	a numeric vector denoting the soil redox potential

Source

BEDDA Training

soil_var

Description

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and B) denoting the blocks
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training

Survival

Description

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and B) denoting the blocks
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training

tf_rcb

Description

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and B) denoting the blocks
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training

upland

Description

Formats

This data contains the following columns:

Env	a factor with three levels (UDSO2-14, UDSO2-6 and UDSO2-6) denoting the environments
Rep	a factor with three levels (1, 2 and 3) denoting the replicate
Block	a factor with four levels (1, 2, 3 and 4) denoting the blocks
Plot	a numeric vector denoting the plots
Geno	a factor with 44 levels (1, 2 and B) denoting the genotypes
Yield	a numeric vector denoting the grain yield

Source

BEDDA Training

White_heads

Description

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and B) denoting the blocks
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training

Year

Description

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and B) denoting the blocks
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training

Yield94

Description

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and B) denoting the blocks
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training

Description

Formats

This data contains the following columns:

REP	a factor with three levels (1, 2 and B) denoting the blocks
GRNYLD	a numeric vector denoting the grain yield in kg/ha

Source

BEDDA Training