

8. Visual Dashboard: Performing Statistical Analyses with Visual Tools

Introduction

Visual Dashboard is one of the Epi Info™ 7 analysis modules. The Visual Dashboard is designed to be intuitive and simple to use. With the use of gadgets, the need for programming code is minimized. Data can be selected, sorted, listed, or manipulated using the various gadgets in Visual Dashboard. The statistical analyses tools available in Visual Dashboard include frequencies, means, and more advanced statistical calculations (e.g., linear regression and logistic regression). Visual Dashboard has graphing functionality to display data as an Epi Curve, Pareto Chart, and several other bar and column charts.

Users can access Visual Dashboard from the Epi Info™ 7 main menu by clicking the Visual Dashboard button, or by selecting **Dashboard** from the main page menu of Enter once an Epi Info project has been loaded.

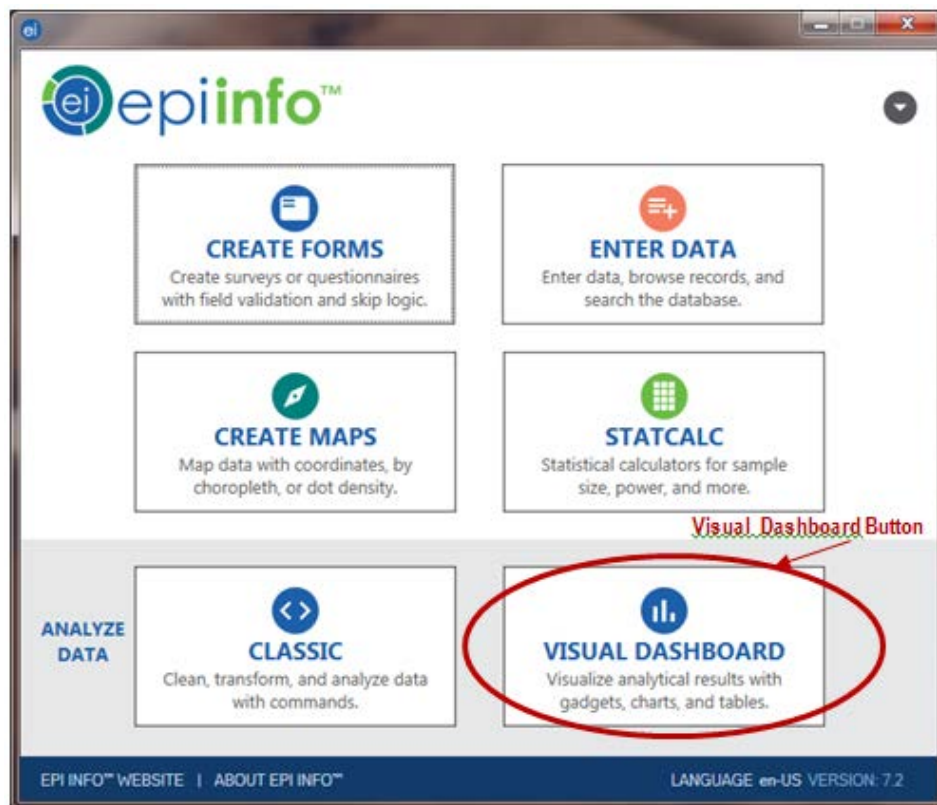


Figure 8.1: Epi Info™ 7 main menu, with Visual Dashboard button highlighted

Navigate the Visual Dashboard Workspace

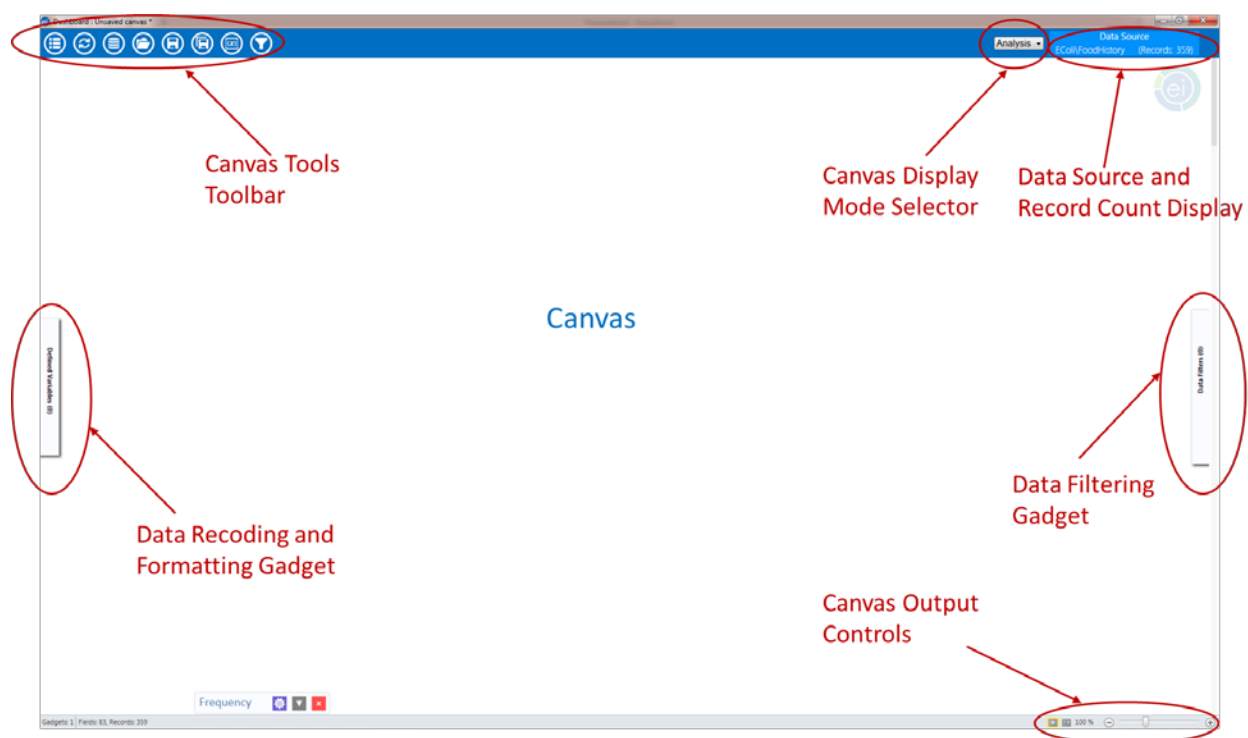


Figure 8.2: Visual Dashboard canvas

The Visual Dashboard contains seven areas: the Canvas, the Canvas Display Mode Selector, the Data Source and Record Count Display, the Canvas Output Controls, the Canvas Tools Toolbar, the Defined Variables gadget, and the Data Filters Gadget.

1. The **Canvas** is the blank space in the center of the Visual Dashboard. It displays output generated by the multiple gadgets available in the tool. The canvas has a function dialog box that appears when a blank space is right-clicked.
2. The **Canvas Display Mode** is located in the upper right corner of the Visual Dashboard just before the data source and record count display. It has three settings:
 - a. **Analysis** (default)—allows the user to create analysis gadgets.
 - b. **Data**—allows the users to browse the data in the data source that is loaded.
 - c. **Variables**—displays a data dictionary for variables in the loaded data set.
3. The **Data Source and Record Count Display**, located at the top right of the screen, displays the source and record count of the data being analyzed. The number of fields (columns) in your dataset will be displayed on the bottom left of the Windows status bar.

4. The **Canvas Output Controls** are located at the bottom right side of the Dashboard workspace. The controls contain three options. The Hide Borders button turns off the external borders for each gadget, which gives the canvas a more report-like appearance. The Auto-arrange button positions the gadgets on the canvas to maximize the display area. The Zoom Percent and Zoom Slider allow you to zoom in and out of the canvas.
5. The **Defined Variables Gadget** in the pull-out drawer on the left side of the canvas allows for the creation, editing, and deletion of variables created during the analysis session.
6. The **Data Filtering Gadget** (labeled as **Data Filters**) in the pull-out drawer on the right side of the canvas allows for canvas-wide filters to be applied. It is used to select a subset of data for analysis.
7. The **Canvas Tools Toolbar** is located at the top left corner of the canvas. It allows a quick way to access Visual Dashboard functions:
 - a. **Options**—opens options menu.
 - b. **Refresh**—refreshes your screen with new data that have been entered.
 - c. **Set Data Source**—sets your data source.
 - d. **Open**—opens a saved canvas.
 - e. **Save**—saves your current canvas back to the original file name and location.
 - f. **Save As**—saves your current canvas to a new name or new location.
 - g. **Variables**—opens the Defined Variables Gadget.
 - h. **Filters**—opens the Data Filtering Gadget.

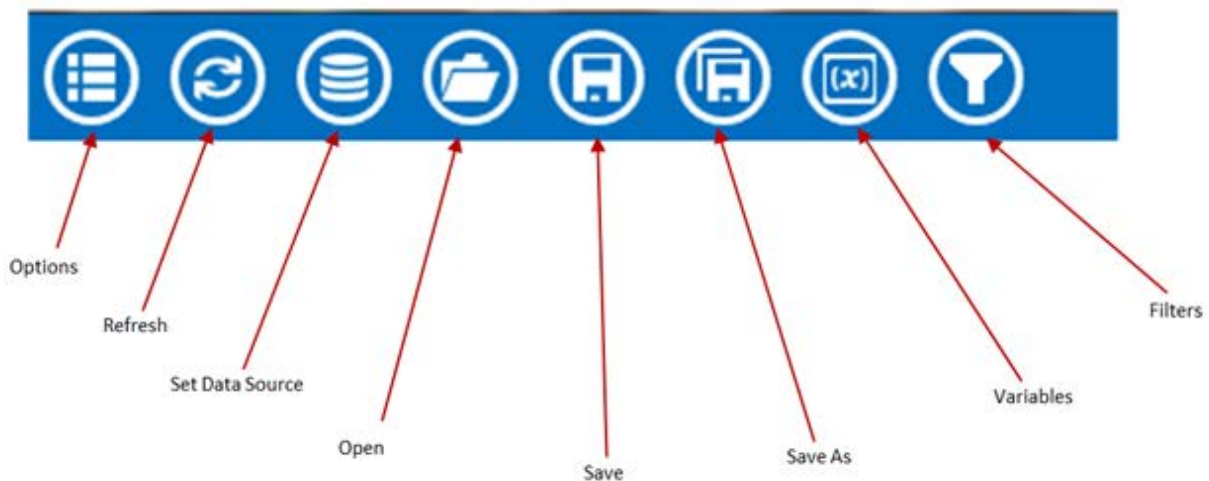


Figure 8.3: Canvas tools toolbar

Canvas Properties

After a data source has been loaded, the **Canvas Properties** are accessed by right clicking on the canvas and selecting **Canvas Properties** from the menu options. These properties are used to:

- Change the data source properties by selecting a new project and table.
- Change HTML output properties and create a custom title, summary, and conclusion for your Visual Dashboard output.
- View technical information about your canvas, such as the number of rows and columns, and cached information.

Note: *Changing the data source in the Canvas Properties is an advanced feature and not recommended unless the schema of both data sources is the same.*

Opening the Options Menu

The **Options** menu icon will open the **Options** menu where user can set a data source.



Figure 8.4: Visual Dashboard toolbar, with **Options** menu icon circled

The **Options** menu has many other options that can be used once a data source is set. There are additional ways to open the Options menu, and these are discussed later in this chapter.

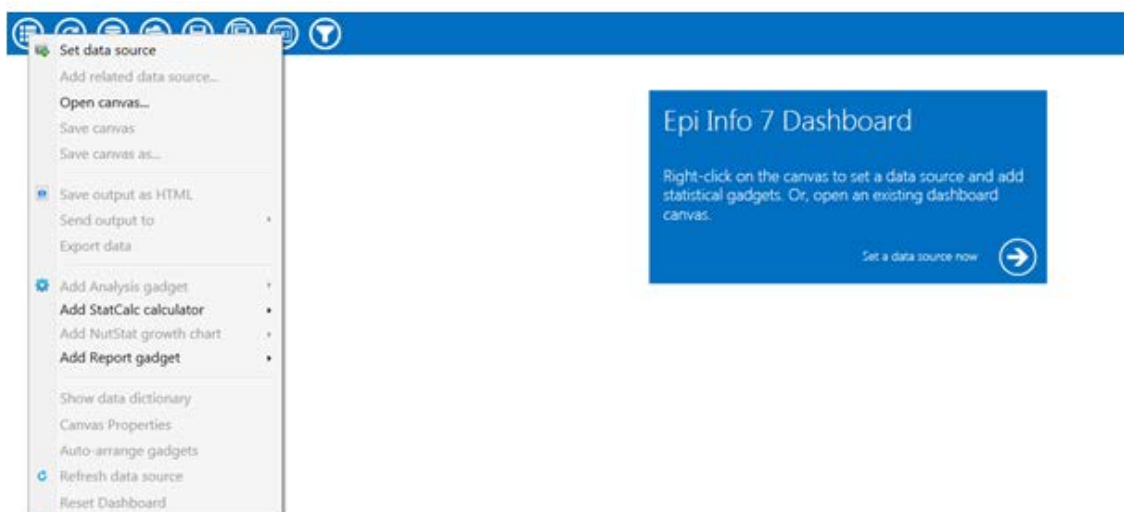


Figure 8.5: Options menu on the Dashboard canvas

Updating the Canvas with Refresh

The Refresh icon updates the gadgets on the canvas with the latest data.



Figure 8.6: Visual Dashboard toolbar, with Refresh icon circled

Selecting a Data Source

Before beginning analysis in Visual Dashboard, select a data source. Data from the following sources can be analyzed in Visual Dashboard: Epi Info™ 7 project files, ASCII Text files, Microsoft Excel, Microsoft Access, Microsoft SQL Server, MySQL, and PostgreSQL. Follow **one** of the first three steps below to select a data source:

1. Click the arrow in the blue box next to **Set a data source now** when you open a new canvas.

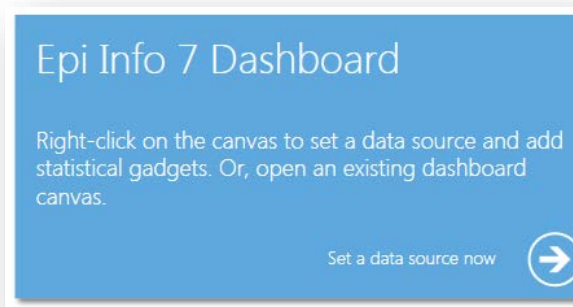


Figure 8.7: Set a data source now option

2. You can also click on the **Set Data Source** icon from the bar at the top of the screen.



Figure 8.8: Visual Dashboard toolbar, with Set Data Source icon circled

3. You can also right click on the canvas and select **Set Data Source** from the Options menu.

4. After selecting the **Set Data Source** option, the **Select Data Source** dialog box appears. Select a **Database Type**, **Data Source**, and form to open data in Visual Dashboard.

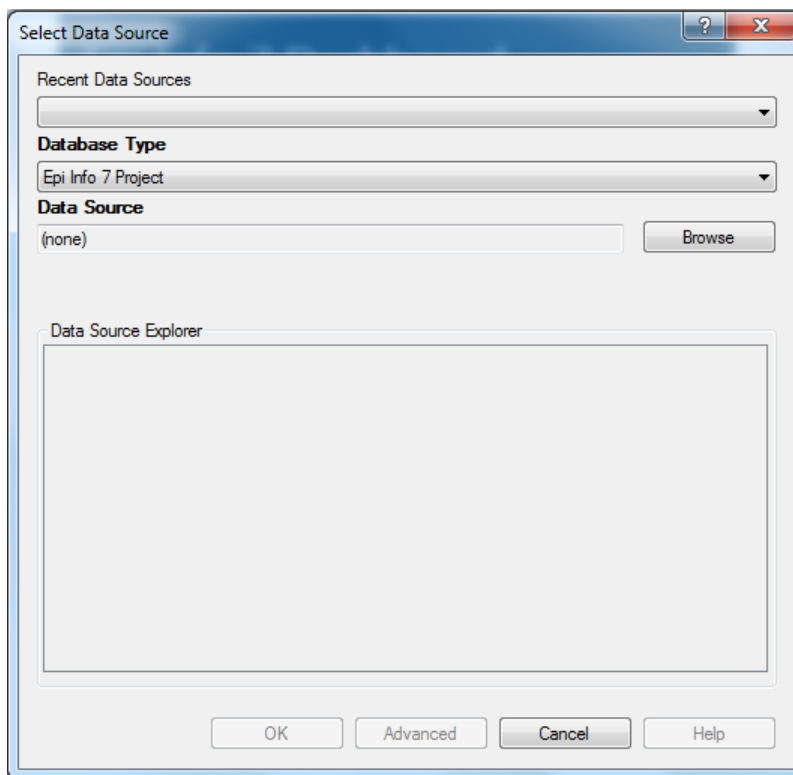


Figure 8.9: Select Data Source dialog box

- The **Recent Data Sources** field provides a list of recently accessed databases in Classic Analysis or Visual Dashboard. By selecting one of the data sources available on the list, you do not need to provide Database Type or Data Source information. If you are reading an Epi Info™ 7 project file from the Data Source section, all available forms will appear in the Data Source Explorer box.
- The **Database Type** field indicates the database file to be loaded (.PRJ, .MDB, .XLS). Specify the data format for the data being accessed. Unless otherwise specified, output files are stored in this destination folder as well.
- The **Data Source** field indicates the file location/path. If you use a SQL Server database, the Data Source field will require server and database names.
- The **Data Source Explorer** field allows you to select a form in your project. This reads the data table associated with your form.
- The **Advanced** mode allows the use of SQL queries to select the data source. Only SELECT statements are allowed.

Read an Epi Info™7 Project File

1. Execute the **Set a Data Source** option using one of the first three steps previously noted (click **Set a Data Source now** arrow, click the **Set a Data Source** icon, or right click and select **Set a Data Source**). The **Select Data Source** dialog box appears.
2. From the **Database Type** drop-down list, select **Epi Info 7 Project**.
3. Click **Browse** next to the **Data Source** field. The **Epi Info 7 Project** folder will appear.
4. Browse for the project file you would like to read into Visual Dashboard. If your projects are in folders, you will have to first open the correct folder and then open the desired project (.PRJ) file. Some sample projects, such as **Ecoli.PRJ**, are included with the installation of Epi Info™ 7.
5. Click **Open**.
6. The **Data Source Explorer** box populates with a list of all the forms that are associated with the project. Select a form.
7. Click **OK**. The record count and data source appear in the toolbar at the top right of the Visual Dashboard canvas.
8. Click **OK**. The “Processing Finished” message appears indicating that the system has cached the data locally. Data from the two tables can now be used for analysis.

Visual Dashboard Analysis Gadgets

The Visual Dashboard contains a number of gadgets for analysis, charts, **and advanced statistics**.

General Analysis Gadgets

- **Line list** - provides a view of data in tabular format
- **Frequency** - shows how many records have each value of a field
- **Word cloud** - used for text fields; shows the most common words used in a font size proportional to the frequency the word is used
- **Combined frequency** - illustrates the frequency and percentage of grouped variables
- **M x N / 2 x 2 table** – allows you to examine the relationship between two or more categorical values
- **Matched pair case-control** – compares an Exposure Variable, the Outcome Variable (i.e., case vs. control), and a Match variable that links each case to its control
- **Means** - calculates the average for a continuous numeric variable
- **Duplicates list** – helps you identify potentially duplicate records
- **Charts** – produces histograms, scatter plots, pie charts, and bar and line graphs directly from data files.
 - o Column chart
 - o Line chart
 - o Area chart

- Pie chart
- Aberration detection chart
- Pareto chart
- Scatter chart
- Epi curve chart
- **Advanced statistics** – provides options to determine the relationship between one dependent variable with one or more independent variables; one or more explanatory variables; or analyze data collected using simple random sampling or unbiased systematic sampling, and/or through more complex sampling strategies such as stratification, cluster sampling, and the use of unequal sampling fractions.
 - Linear regression
 - Logistic regression
 - Complex sample frequencies
 - Complex sample means
 - Complex sample tables

Analysis Gadget > Line List

The **Line List** option creates a list of the current dataset. This allows you to view the details of the data in your current data source. You can select variables to display and then sort or group those variables accordingly. **Line List** supports adding groups or pages. All fields in a group or on a page will be displayed in such cases. This option is available at the bottom of the **Variables to display** list within the **Line List** Variables property panel (below the **Fields** list).

1. Select the **EColi.PRJ** Data Source. Open the **FoodHistory** form from the **Data Source Explorer** menu.
2. Right click on the canvas and select **Add Analysis Gadget > Line list**.
3. The Line List gadget's configuration window opens to the Variables property panel.

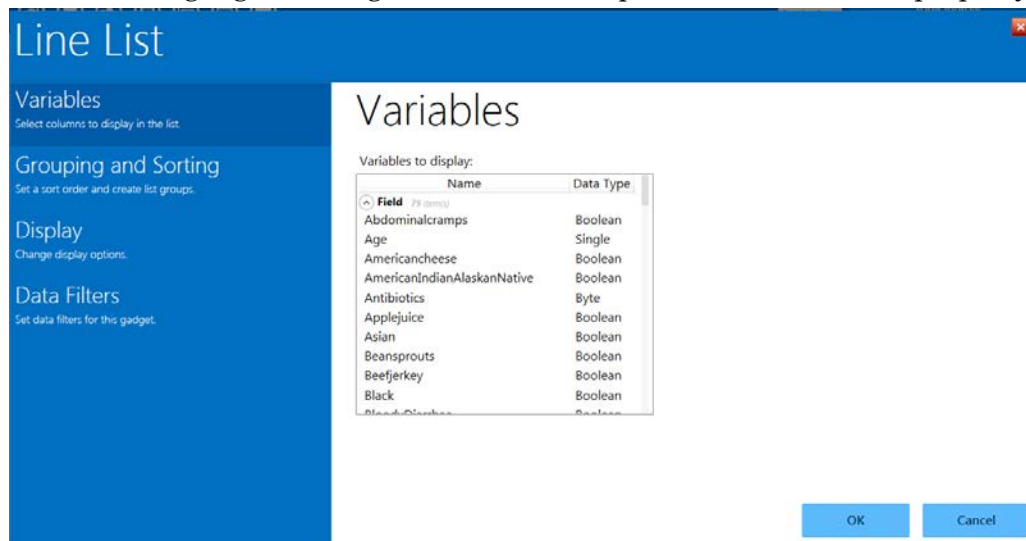
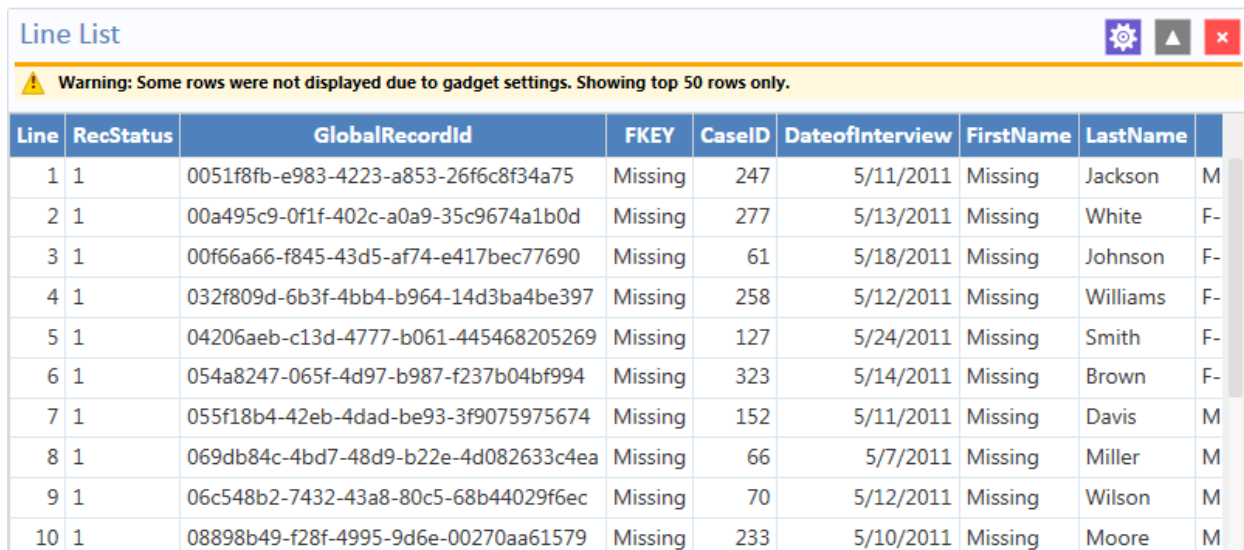


Figure 8.10: Line List Gadget

- To select multiple variables from **Variables to display**, hold down the **Ctrl** key and click on the variables you want to display. You can also hit **Ctrl + A** to select all variables at once.
- Click the **OK** button. The line list appears on the canvas. The graphic below illustrates the list with all variables selected.



The screenshot shows a window titled "Line List" with a yellow warning banner at the top that reads: "Warning: Some rows were not displayed due to gadget settings. Showing top 50 rows only." Below the banner is a table with the following columns: Line, RecStatus, GlobalRecordId, FKEY, CaseID, DateofInterview, FirstName, LastName, and an unlabeled column. The table contains 10 rows of data.

Line	RecStatus	GlobalRecordId	FKEY	CaseID	DateofInterview	FirstName	LastName	
1	1	0051f8fb-e983-4223-a853-26f6c8f34a75	Missing	247	5/11/2011	Missing	Jackson	M
2	1	00a495c9-0f1f-402c-a0a9-35c9674a1b0d	Missing	277	5/13/2011	Missing	White	F-
3	1	00f66a66-f845-43d5-af74-e417bec77690	Missing	61	5/18/2011	Missing	Johnson	F-
4	1	032f809d-6b3f-4bb4-b964-14d3ba4be397	Missing	258	5/12/2011	Missing	Williams	F-
5	1	04206aeb-c13d-4777-b061-445468205269	Missing	127	5/24/2011	Missing	Smith	F-
6	1	054a8247-065f-4d97-b987-f237b04bf994	Missing	323	5/14/2011	Missing	Brown	F-
7	1	055f18b4-42eb-4dad-be93-3f9075975674	Missing	152	5/11/2011	Missing	Davis	M
8	1	069db84c-4bd7-48d9-b22e-4d082633c4ea	Missing	66	5/7/2011	Missing	Miller	M
9	1	06c548b2-7432-43a8-80c5-68b44029f6ec	Missing	70	5/12/2011	Missing	Wilson	M
10	1	08898b49-f28f-4995-9d6e-00270aa61579	Missing	233	5/10/2011	Missing	Moore	M

Figure 8.11: Generated Line List of all variables

Note that there is a yellow warning at the top of the output. By default, the line list will only show the top 50 rows. The row limit exists to improve performance when working with large datasets. The Max Rows to Display setting can be modified in the Display property panel, which is explained later.

Sorting Records

Sometimes it is important to have the records in a file arranged in a particular order, such as alphabetically by name, or numerically by age. There are two ways to sort the records: by clicking on the column name in the line list or by using the sort options in the Line List configuration window's Sorting and Grouping property panel.

Sorting a List by Column Name

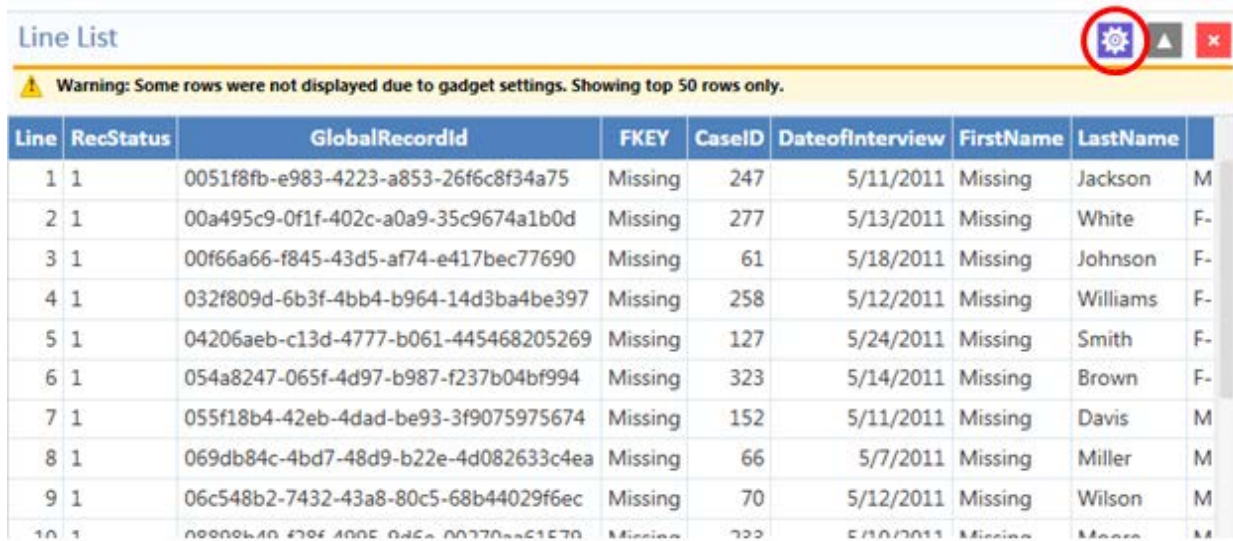
Follow the steps below to sort the list by Age.

- Select the **EColi.PRJ** Data Source. Open the **FoodHistory** form from the **Data Source Explorer** menu.
- Right click on the canvas and select **Add Analysis Gadget > Line list**.
- The Line List gadget's configuration window opens to the Variables property panel.
- Select all variables from the **List Variables to Display** box by pressing **Ctrl + A** on your keyboard.
- Click **OK**.
- The **Line List** results appear on the Visual Dashboard canvas.

- Click on the **Age** column heading (you may have to scroll over to see the **Age** column). The records are displayed by **Age** in ascending order. The **Age** column header also turns red to indicate the data are currently sorted by that variable.
- To sort the column in descending order, click the **Age** column header again. The **Age** column header turns green to show it is sorted in descending order.

Sorting a List Using Line List Grouping and Sorting Properties

- Select the **EColi.PRJ** Data Source. Open the **FoodHistory** form from the **Data Source Explorer** menu.
- Right click on the canvas and select **Add Analysis Gadget > Line list**.
- The Line List gadget's configuration window opens to the Variables property panel.
- Select all variables from **List Variables to Display** box by pressing **Ctrl + A** on your keyboard.
- Click **OK**.
- The **Line List** results appear on the Visual Dashboard canvas.
- The list data are not sorted in any particular order. To sort variables, first click on the gadget gear icon at the top of the Line List.



Line	RecStatus	GlobalRecordId	FKEY	CaseID	DateofInterview	FirstName	LastName	
1	1	0051f8fb-e983-4223-a853-26f6c8f34a75	Missing	247	5/11/2011	Missing	Jackson	M
2	1	00a495c9-0f1f-402c-a0a9-35c9674a1b0d	Missing	277	5/13/2011	Missing	White	F-
3	1	00f66a66-f845-43d5-af74-e417bec77690	Missing	61	5/18/2011	Missing	Johnson	F-
4	1	032f809d-6b3f-4bb4-b964-14d3ba4be397	Missing	258	5/12/2011	Missing	Williams	F-
5	1	04206aeb-c13d-4777-b061-445468205269	Missing	127	5/24/2011	Missing	Smith	F-
6	1	054a8247-065f-4d97-b987-f237b04bf994	Missing	323	5/14/2011	Missing	Brown	F-
7	1	055f18b4-42eb-4dad-be93-3f9075975674	Missing	152	5/11/2011	Missing	Davis	M
8	1	069db84c-4bd7-48d9-b22e-4d082633c4ea	Missing	66	5/7/2011	Missing	Miller	M
9	1	06c548b2-7432-43a8-80c5-68b44029f6ec	Missing	70	5/12/2011	Missing	Wilson	M
10	1	069db84c-4bd7-48d9-b22e-4d082633c4ea	Missing	66	5/7/2011	Missing	Miller	M

Figure 8.12: Line List with Gadget Configuration Icon Noted

- Select **Grouping and Sorting** in the configuration window and then select Sorting Method of Ascending or Descending.
- Double click the desired sort variable from the **Available Variables** list. You can select multiple variables on which to sort by double-clicking each one. With each double click operation, the selected variable will move to the **Sort Order** box.

10. Double click on a variable in the **Sort Order** box to remove that sort variable from the list.

Grouping and Sorting

Grouping

Group By

Sub-Group By

Sorting

Method:

Available Variables	Sort Order:
Grapes	Age (ascending)
Headache	Antibiotics (ascending)
HomePhone	ILL (ascending)
HospitalAdmissionDate	
Hospitalized	
Icecream	
Irritable	

OK Cancel

Figure 8.13: Line List Grouping and Sorting

11. If you want to group the list by a particular field, select the field from the **Group By** drop-down list.

Grouping and Sorting

Grouping

Group By

Sub-Group By

Sorting



Method:

Available Variables	Sort Order:
Abdominalcramps	
Age	
Americancheese	
AmericanIndianAlaskanNative	
Antibiotics	
Applejuice	
Asian	

OK Cancel

Figure 8.14: Line List/Grouping options

12. Click **OK**. Visual Dashboard generates and displays a separate line list, grouped by the selected variable.



Line List   

Warning: Some rows were not displayed due to gadget settings. Showing top 50 rows only.

Sex	DOB	Age	EthnicityGroup	PatientAddress	State	Latitude	Longitude	Occupation	EmailAddress
-Male	1/1/1986	25	Missing	Missing	GA	33.594	-83.475	Software developer	Missing
-Male	10/10/1985	25	Missing	Missing	GA	33.525	-83.228	Soldier	Missing
-Male	1/1/1986	25	Missing	Missing	GA	33.74	-83.303	Service worker	Missing
Female	10/4/1980	30	Missing	Missing	GA	33.627	-83.61	Manufacturing	Missing
Female	1/1/1981	30	Missing	Missing	GA	33.621	-83.681	Lawyer	Missing
-Male	4/16/1981	30	Missing	Missing	GA	33.599	-83.457	Researcher	Missing
Female	3/11/1963	48	Missing	Missing	GA	33.585	-83.175	Unemployed	Missing
Female	8/28/1962	48	Missing	Missing	GA	33.401	-83.587	Manufacturing	Missing
Female	4/19/1991	20	Missing	Missing	GA	33.603	-83.63	Police officer	Missing
Female	3/20/1987	24	Missing	Missing	GA	33.575	-83.185	Sales	Missing
Female	2/4/1959	52	Missina	Missina	GA	33.603	-83.472	Lawver	Missina

Figure 8.15: Line List with Grouping results

13. You can hide a list from view by clicking on the arrow to the left of the list heading.

Line List   

Warning: Some rows were not displayed due to gadget settings. Showing top 50 rows only.

Line	RecStatus	GlobalRecordId	FKEY	CaseID	Dateofinterview	FirstName	LastName	F-
49	1	240ca8c2-2504-4f8a-b0cf-9ade00287f0e	Missing	92	5/1/2011	Missing	Evans	F-
4/17/2011 12:00:00 AM 2								
2	1	00a495c9-0f1f-402c-a0a9-35c9674a1b0d	Missing	277	5/13/2011	Missing	White	F-
11	1	090f4bf5-a8f9-4048-85be-b167a4635e92	Missing	275	5/13/2011	Missing	Taylor	F-
5/12/2011 12:00:00 AM 1								
5/17/2011 12:00:00 AM 1								
4/25/2011 12:00:00 AM 1								
4/28/2011 12:00:00 AM 1								
18	1	0c9df877-6df1-4e58-818d-a59fef41b84a	Missing	30	5/17/2011	Missing	Thompson	F-
5/3/2011 12:00:00 AM 1								
21	1	0f5627fd-3a7d-4ac3-8d38-08f3cf4d4615	Missing	45	5/16/2011	Missing	Robinson	F-
5/19/2011 12:00:00 AM 1								
23	1	0fdc8da1-dd50-4ce8-9176-2d5bf9cf2039	Missing	151	5/9/2011	Missing	Rodriguez	F-

Figure 8.16: Line List hidden view, with circle indicating “hide” arrow

Remove Sort Criteria

Remove sort criteria by following the steps below.

1. Right click on the **Line List**.
2. Select **Remove All Sort Criteria**.
3. The Line List will appear with the sort criteria removed.

Other Line List Properties

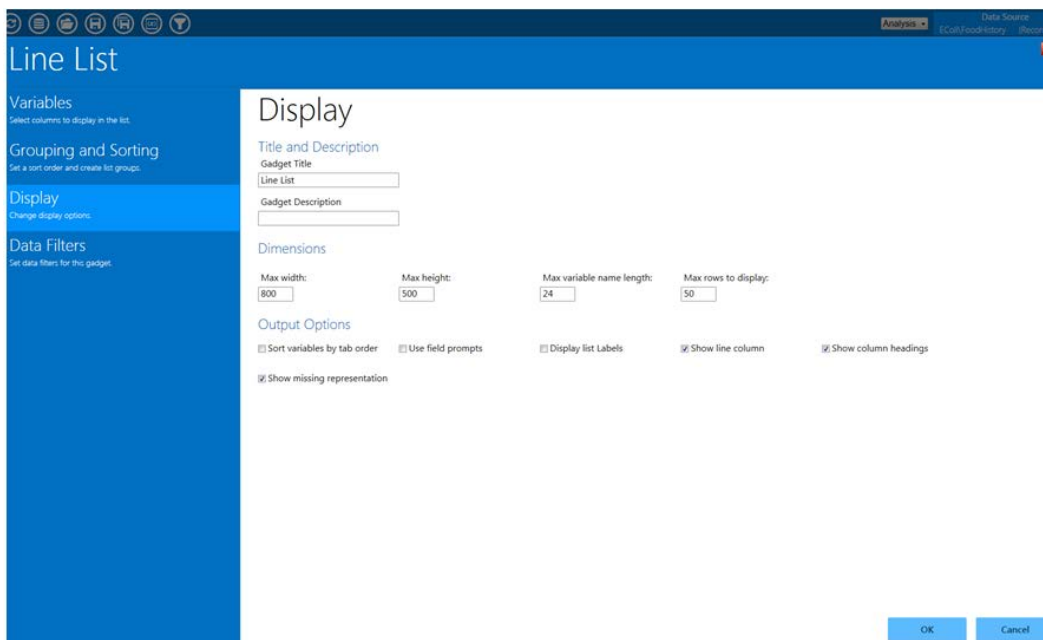


Figure 8.17: Line List Display options

Other options for the **Line List** gadget are available by selecting **Display** in the configuration window.

- The **Max width** indicates the maximum Line List display width in pixels.
- The **Max height** indicates the maximum Line List display height in pixels.
- The **Max variable name length** text box allows for the truncation of long column names in the list output. It is set to 24 by default.
- The **Max rows to display** text box allows the user to change the maximum number of rows that the line list gadget will display. It is set to 50 by default. A maximum of 2,000 rows may be displayed.

Output Options (located at the bottom of the Display screen)

- The **Sort variables by tab order** checkbox forces the columns in the output to be sorted by their tab order.* It is not checked by default.
- The **Use field prompts** checkbox will use the field's prompt as the column heading, rather than the field's name.* It is not checked by default.
- The **Display list labels** checkbox will use the label value for option fields and comment legal fields rather than the underlying value stored in the database.*

* Note: These first three output options are available only when the Data Source is an Epi Info 7 project. If the Data Source is not an Epi Info 7 project, such as an Excel spreadsheet, these options have no effect.

- The **Show line column** option adds the Line column to the list and displays the line number when the checkbox is selected.
- The **Show column headings** option displays the line list column headings when the checkbox is selected.
- The **Show missing representation** option displays the word or symbol that represents a missing value for any field that does not have a value. By default, the word “Missing” is used, but other options include “Unknown”, “(.)” or another custom setting as specified in the Tools > Options > Analysis tab. If the checkbox is not selected, a field without a value will be left blank.

Exporting a Line List

Once the Line List is created, the following export options are available: **Copy List Data to Clipboard**, **Send List Data to Web Browser**, and **Send List Data to Excel**.

The next three steps explain how to **Copy List Data to Clipboard**.

1. Right click on the **Line List**.
2. Select **Copy List Data to Clipboard**.

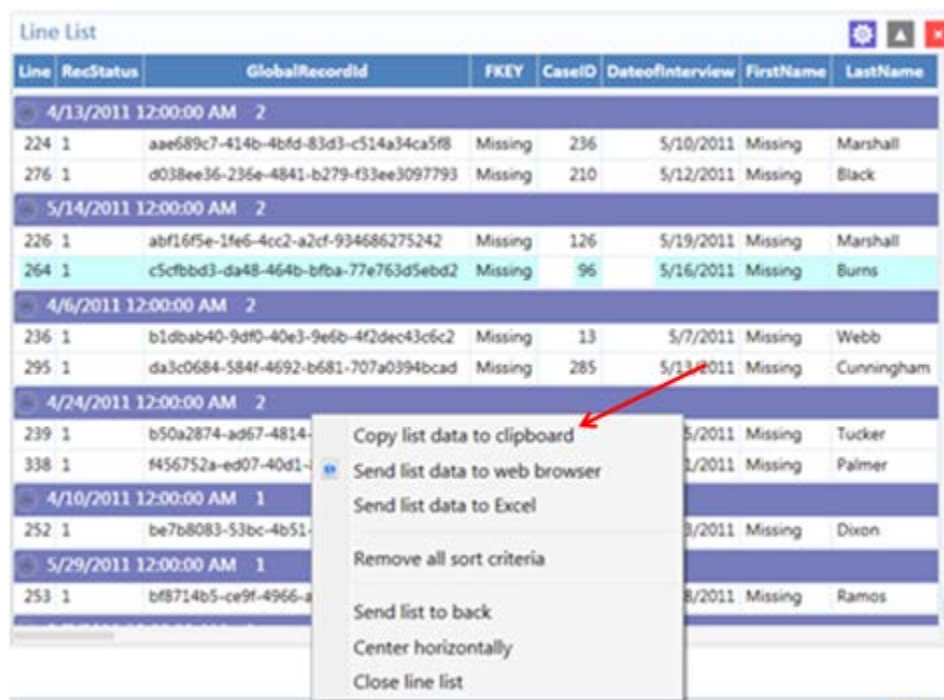


Figure 8.18: Copy list data to clipboard

3. The data are now on the clipboard and may be exported.

The next three steps explain how to **Send List Data to Web Browser**.

1. Right click on the **Line List**.
2. Select **Send List Data to Web Browser**.

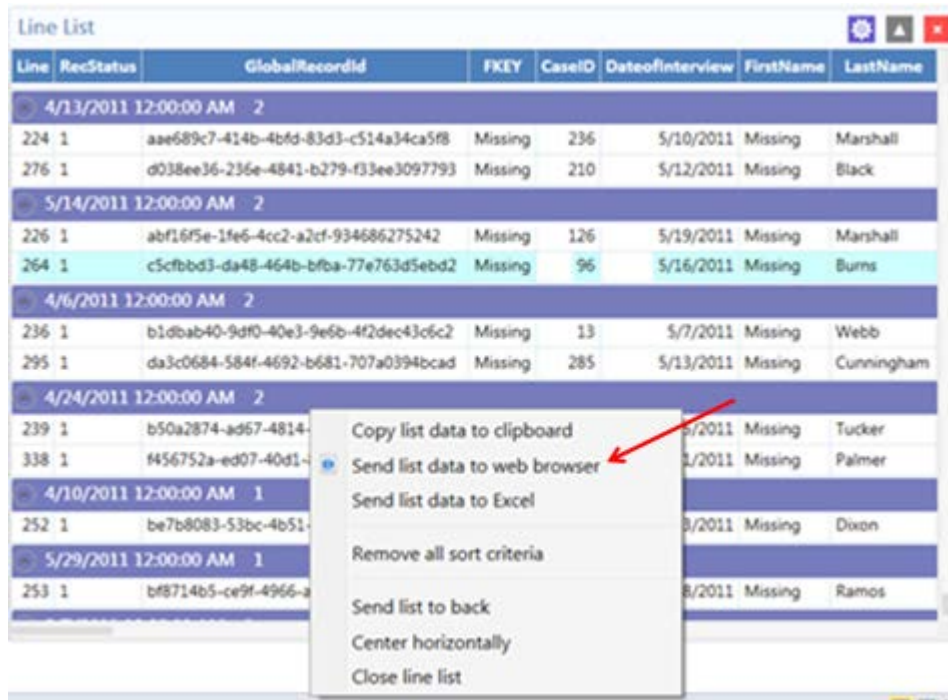


Figure 8.19: Send list data to web browser

3. The data now appear in your default web browser.

Line List

DoctorVisitDate =

Line	RecStatus	GlobalRecordId	FKEY	CaseID	DateofInterview	FirstName	LastName	Sex	DOB	Age	EthnicityGroup	PatientAddress	State	Latitude	Longitude	Occupation
1	1	0051f8fb-e983-4223-a853-26f6c8f34a75		247	5/11/2011 12:00:00 AM	Jackson		M-Male	1/1/1986 12:00:00 AM	25			GA	33.594	-83.475	Software developer
3	1	00f66a66-f845-43d5-a774-e417bec77690		61	5/18/2011 12:00:00 AM	Johnson		F-Female	3/11/1963 12:00:00 AM	48			GA	33.585	-83.175	Unemployed
4	1	032f809d-6b3f-4bb4-b964-14d3ba4be397		258	5/12/2011 12:00:00 AM	Williams		F-Female	4/19/1991 12:00:00 AM	20			GA	33.603	-83.63	Police officer
6	1	054a8247-065f-4a97-b987-f237b04bf994		323	5/14/2011 12:00:00 AM	Brown		F-Female	2/4/1959 12:00:00 AM	52			GA	33.603	-83.472	Lawyer
8	1	069db84c-4bd7-48a9-b22e-4d082633c4ea		66	5/7/2011 12:00:00 AM	Miller		M-Male	1/1/1990 12:00:00 AM	21			GA	33.627	-83.61	Student
9	1	06c548b2-7432-43ab-80c5-68b44029f6ec		70	5/12/2011 12:00:00 AM	Wilson		M-Male	1/1/1980 12:00:00 AM	31			GA	33.587	-83.469	Sales

Figure 8.20: List Data in web browser

The next three steps explain how to **Send List Data to Excel**.

1. Right click on the **Line List**.
2. Select **Send List Data to Excel**.

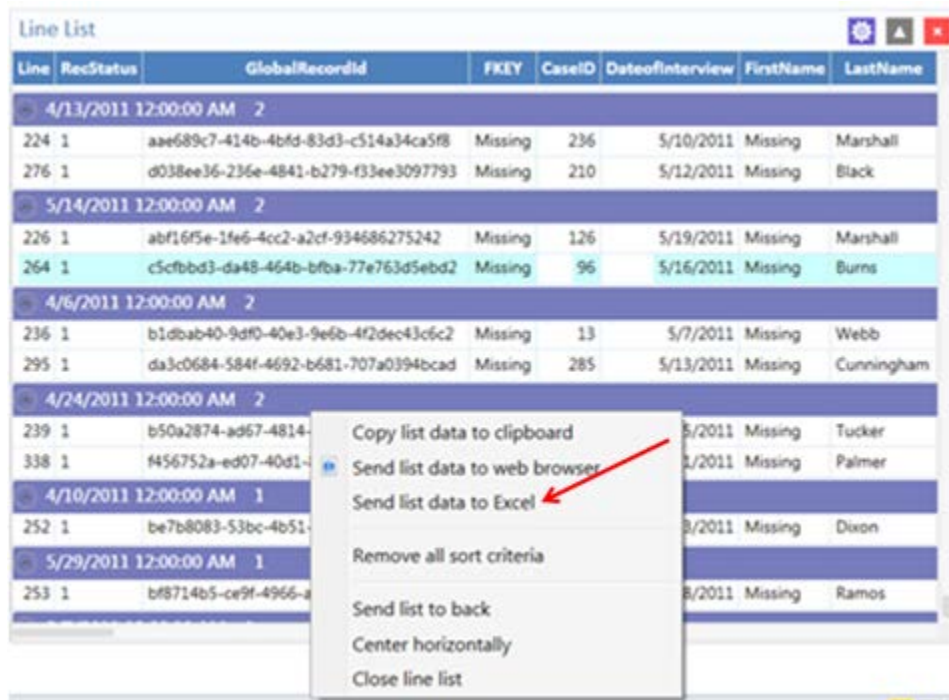


Figure 8.21: Send list data to Excel

3. The list data appear in a Microsoft Excel spreadsheet.

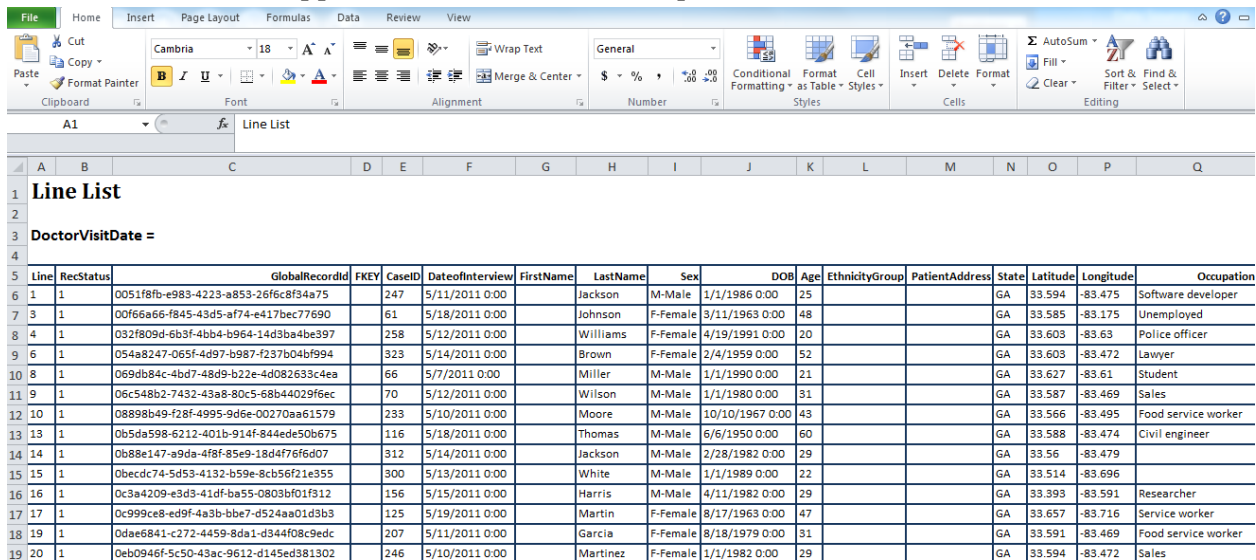


Figure 8.22: List Data in Excel

Note: You must have Microsoft Excel installed to use this feature.

Analysis Gadget > Frequency

The Frequency gadget counts each occurrence of a value for a specified variable and gives the absolute and relative frequencies for each value. This option then produces a frequency table that shows how many records have a value for each variable, the percentage of the total, a cumulative percentage and upper and lower confidence intervals.

1. Select the **EColi.PRJ** Data Source. Open the **FoodHistory** form from the **Data Source Explorer** menu.
2. Right click on the canvas and select **Add Analysis Gadget > Frequency**. The Frequency gadget configuration window opens to the Variables property panel.

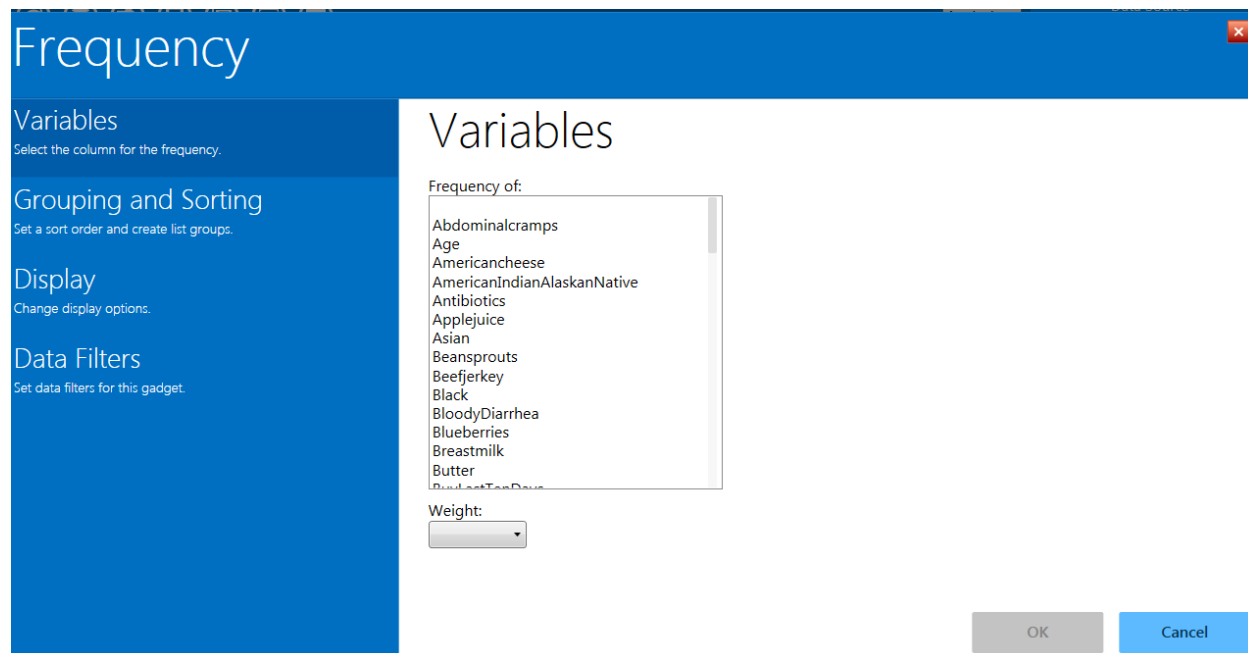


Figure 8.23: Frequency gadget

3. From the **Frequency Of** list, select a **variable** from the data table. For this example, select Sex as the desired variable.

Variables

Frequency of:

Abdominalcramps
 Age
 Americancheese
 AmericanIndianAlaskanNative
 Antibiotics
 Applejuice
 Asian
 Beansprouts
 Beefjerkey
 Black
 BloodyDiarrhea
 Blueberries
 Breastmilk
 Butter
 BudLeatTeaDeus

Weight:

OK
Cancel

Figure 8.24: Selecting frequency of variable

- Click **OK**. The results appear on the Visual Dashboard canvas.

Frequency						
Sex	Frequency	Percent	Cum. Percent	Wilson 95% LCL	Wilson 95% UCL	
F-Female	186	51.81 %	51.81 %	46.65 %	56.93 %	<div style="width: 50%; height: 10px; background-color: #4a90e2;"></div>
M-Male	173	48.19 %	100.00 %	43.07 %	53.35 %	<div style="width: 50%; height: 10px; background-color: #4a90e2;"></div>
TOTAL	359	100.00 %	100.00 %			<div style="width: 50%; height: 10px; background-color: #4a90e2;"></div>

Figure 8.25: Frequency results

The **Frequency** column provides the count of individuals that are either female or male. The **Percent** column indicates the percentage of female or male. The **Wilson 95% LCL (Lower Confidence Limits)** are a range of values that indicates the likely location of the true value of a measure, meaning (in this instance) that the percentage of females is likely to be between 46.64% and 56.93%.

Frequency Additional Options

Additional options are available in the **Frequency** gadget:

- Select the **EColi.PRJ** Data Source. Open the **FoodHistory** form from the **Data Source Explorer** menu.
- Right click on the canvas and select **Add Analysis Gadget > Frequency**. The Frequency gadget configuration window opens to the Variables property panel.

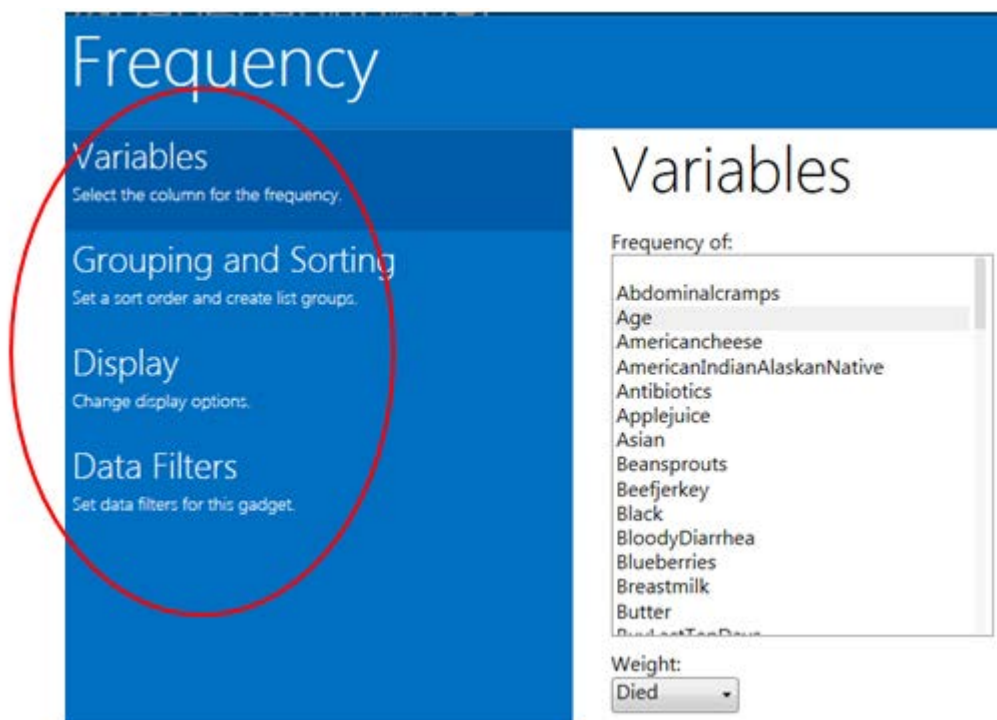





Figure 8.26: Frequency gadget functions menu

The following options are available in the Frequency gadget, under the various functions located in the left-hand column:

- **Weight** (available under **Variables**)—The field to use as the weight variable, for example, if aggregate data are being analyzed.
 - **Stratify By** (available under **Grouping and Sorting**)—The field to use for stratifying (grouping) the output.
 - **Sort high to low** (available under **Grouping and Sorting**)— displays the results sorted from high to low based on the selected Frequency of variable.
 - **Display all list values** (available under **Display**)—Only applicable when the frequency variable is a drop-down list field in an Epi Info™ 7 project. When selected, the output will show all of the drop-down list values even if they have a count of zero.
 - **Display list labels** (available under **Display**)—Only applicable when the frequency variable is a Comment Legal field or Option Field in an Epi Info™ 7 project. When selected, the output will show the value and the label for Comment Legal fields and the option button label for Option fields.
 - **Include Missing Values** (available under **Display**)—This option, which is unchecked by default, displays the count of rows having missing values.
3. From the **Frequency of** list in **Variables**, select **Age** from the data table.
 4. From the **Weight** drop-down list below, select **Died** from the data table.
 5. Select **Hospitalized** from the **Stratify by** list.
 6. Click **OK**. Results appear on the canvas.

Frequency   

Was the patient hospitalized? = Yes



Age	Frequency	Percent	Cum. Percent	Exact 95% LCI	Exact 95% UCL	
7	0	0.00 %	0.00 %	0.00 %	36.94 %	
11	0	0.00 %	0.00 %	0.00 %	36.94 %	
12	0	0.00 %	0.00 %	0.00 %	36.94 %	
13	0	0.00 %	0.00 %	0.00 %	36.94 %	
14	2	25.00 %	25.00 %	3.19 %	65.09 %	
15	0	0.00 %	25.00 %	0.00 %	36.94 %	
16	0	0.00 %	25.00 %	0.00 %	36.94 %	
17	0	0.00 %	25.00 %	0.00 %	36.94 %	
18	0	0.00 %	25.00 %	0.00 %	36.94 %	
19	0	0.00 %	25.00 %	0.00 %	36.94 %	
20	0	0.00 %	25.00 %	0.00 %	36.94 %	
21	0	0.00 %	25.00 %	0.00 %	36.94 %	
22	1	12.50 %	37.50 %	0.32 %	52.65 %	
23	0	0.00 %	37.50 %	0.00 %	36.94 %	
24	0	0.00 %	37.50 %	0.00 %	36.94 %	
25	0	0.00 %	37.50 %	0.00 %	36.94 %	

Figure 8.27: Frequency Line List stratified view #1

Was the patient hospitalized? = No

Age	Frequency	Percent	Cum. Percent	Exact 95% LCI	Exact 95% UCL	
7	0	0.00 %	0.00 %	0.00 %	97.50 %	
11	0	0.00 %	0.00 %	0.00 %	97.50 %	
12	0	0.00 %	0.00 %	0.00 %	97.50 %	
13	0	0.00 %	0.00 %	0.00 %	97.50 %	
14	0	0.00 %	0.00 %	0.00 %	97.50 %	
15	0	0.00 %	0.00 %	0.00 %	97.50 %	
16	0	0.00 %	0.00 %	0.00 %	97.50 %	
17	0	0.00 %	0.00 %	0.00 %	97.50 %	
18	0	0.00 %	0.00 %	0.00 %	97.50 %	
19	0	0.00 %	0.00 %	0.00 %	97.50 %	
20	0	0.00 %	0.00 %	0.00 %	97.50 %	
21	0	0.00 %	0.00 %	0.00 %	97.50 %	
22	0	0.00 %	0.00 %	0.00 %	97.50 %	
23	0	0.00 %	0.00 %	0.00 %	97.50 %	
24	0	0.00 %	0.00 %	0.00 %	97.50 %	
25	0	0.00 %	0.00 %	0.00 %	97.50 %	

Figure 8.28: Frequency Line List stratified view #2

The results display the frequency of deceased patients by age and are stratified by whether the patient was hospitalized or not. The age, frequency, percent, cumulative percent and upper and lower 95% confidence intervals are displayed. The last column is the percent bar.

You can also open the Display property panel to show additional properties for the Frequency gadget. The following options are available in Display:

- **Gadget Title**—Allows you to enter a title for the current gadget. This title will appear at the top of the gadget results on the canvas.
- **Gadget Description**— Allows you to enter a description for the current gadget. This description will appear at the top of the gadget results on the canvas, just below the title.
- **Use field prompts**—Specifies whether to use the field question or prompt instead of the field name. This property is checked by default.
- **Draw borders**—Setting for drawing the frequency’s border.
- **Draw header row**—Setting for displaying the frequency’s headers.
- **Draw total row**—Setting for displaying the frequency’s totals.
- **Decimals to display**—Specifies the number of decimals that will be displayed in the results. The possible values range from zero to four. This is set to two by default.
- **Maximum rows to display**—Specifies number of rows to display.
- **Max width of percent bar**—Width for percent bar column using a % value.

Analysis Gadget > Word cloud

The **Word cloud** option provides a visual snapshot of terms used in the current dataset. You can select a variable to display (parse). **Word cloud** returns a graphic of words and terms used in that field, with font sizes illustrating the relative use of that term compared to others in the search. This gadget is most valuable for analyzing text fields and multi-line text fields. For example, it can condense a field with patients’ chief complaints from verbose responses to just the most frequently used words. **Word cloud** can also analyze other text-based fields such as Legal Value, Comment Legal, and Codes fields.

The gadget allows you to exclude frequent mundane words from the search with an editable list of default words for Epi Info™ to ignore (i.e., exclude from the analysis): *a am an and are as at but by for her his I if in is it me my not of on or our pm that the their this to us was we were will with would you your.*

1. Select the **Sample.PRJ** Data Source. Open the **Surveillance** form from the **Data Source Explorer** menu.
2. Right click on the canvas and select **Add Analysis Gadget >Word cloud**. The **Word Cloud** gadget configuration window opens to the Variables property panel.

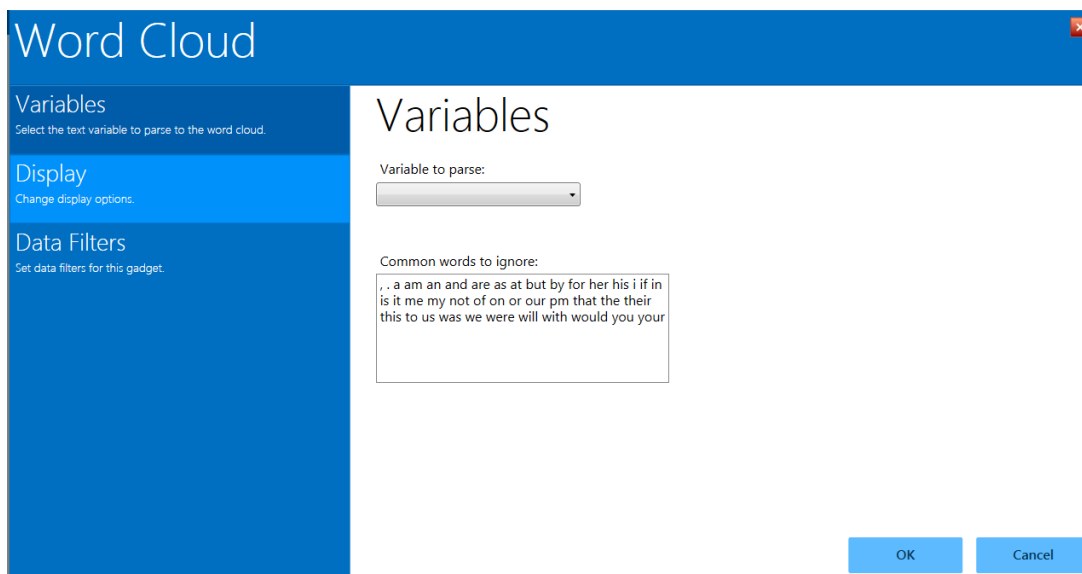


Figure 8.29: Word cloud gadget

3. Choose the desired variable on which to parse words from the drop down list. For this example, select **Race**. The **Common words to ignore** can remain as listed in the default.
4. Click **OK**. The results appear on the Visual Dashboard canvas. The font size for each word in the selected variable group denotes the use of that word relative to the others in the group. In this example, then, the most common race can be deduced to be “White,” followed by “Black.”



Figure 8.30: Word cloud results

Analysis Gadget > Combined Frequency

The Combined Frequency option produces a table that illustrates the frequency and percentage of affirmative responses to check box and Yes-No fields contained within a group field (i.e., checked boxes are affirmative, as are “Yes” answers to Yes-No fields). All other field types are ignored. In the following example, the Demographic Information group contains a number of check box fields for the individual’s race and it allows you to “Check all that apply”.

1. Select the **EColi.PRJ** Data Source. Open the **FoodHistory** form from the **Data Source Explorer** menu.
2. Right click on the canvas and select **Add Analysis Gadget > Combined frequency**. The **Combined Frequency** gadget configuration window opens to the Variables property panel. The **Variables to display** list only shows Group variables.

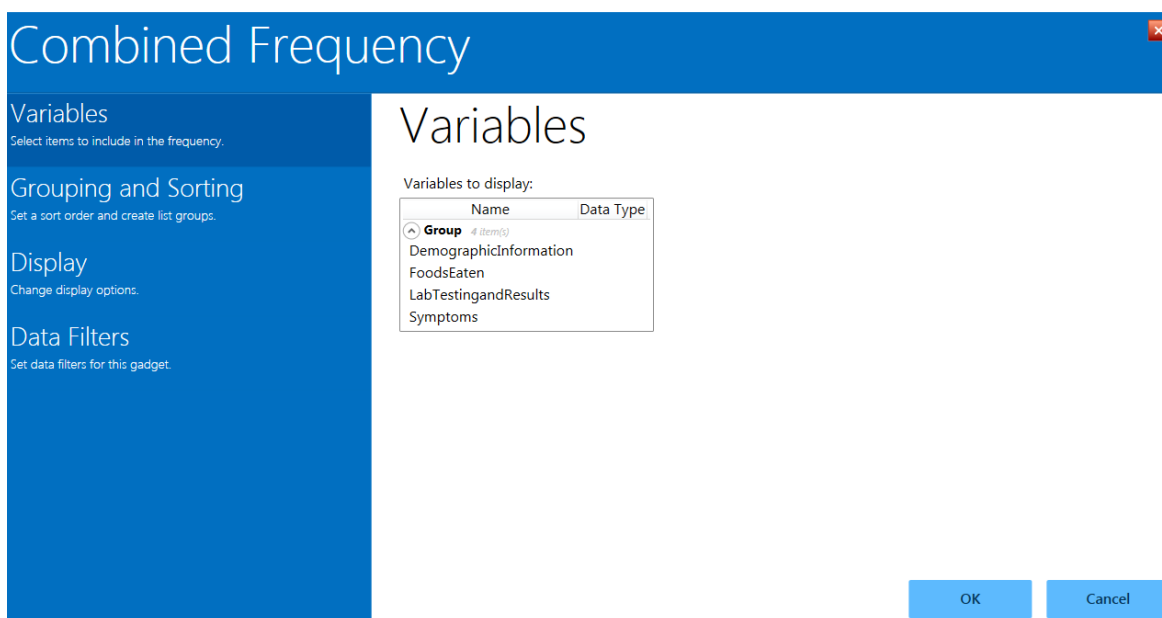


Figure 8.31: Combined Frequency gadget

3. From the **Variables to Display** list under Group, select **Symptoms**.
4. Click **OK**. The results appear on the Visual Dashboard canvas.




Combined Frequency   		
Value	Frequency	Percent
Fever	311	86.63 %
ILL	276	76.88 %
Headache	269	74.93 %
Vomiting	241	67.13 %
Non-bloody diarrhea	218	60.72 %
Bloody diarrhea	218	60.72 %
Nausea	208	57.94 %
Chills	190	52.92 %
DoctorVisit	180	50.14 %
Irritable	174	48.47 %
Abdominal cramps	172	47.91 %
Hospitalized	137	38.16 %
Poor feeding	125	34.82 %
Antibiotics	114	31.75 %
Died	11	3.06 %
Fields are boolean.		
Denominator = 359		

Figure 8.32: Combined Frequency results

Combined Frequency Advanced Options

To use the Combined Frequency advanced options:

1. Select the **EColi.PRJ** Data Source. Open the **FoodHistory** form from the **Data Source Explorer** menu.
2. Right click on the canvas and select **Add Analysis Gadget > Combined Frequency**. The **Combined Frequency** gadget configuration window opens to the **Variables** property panel.
3. In the **Variables** option, select **DemographicInformation** from the **Group** field.

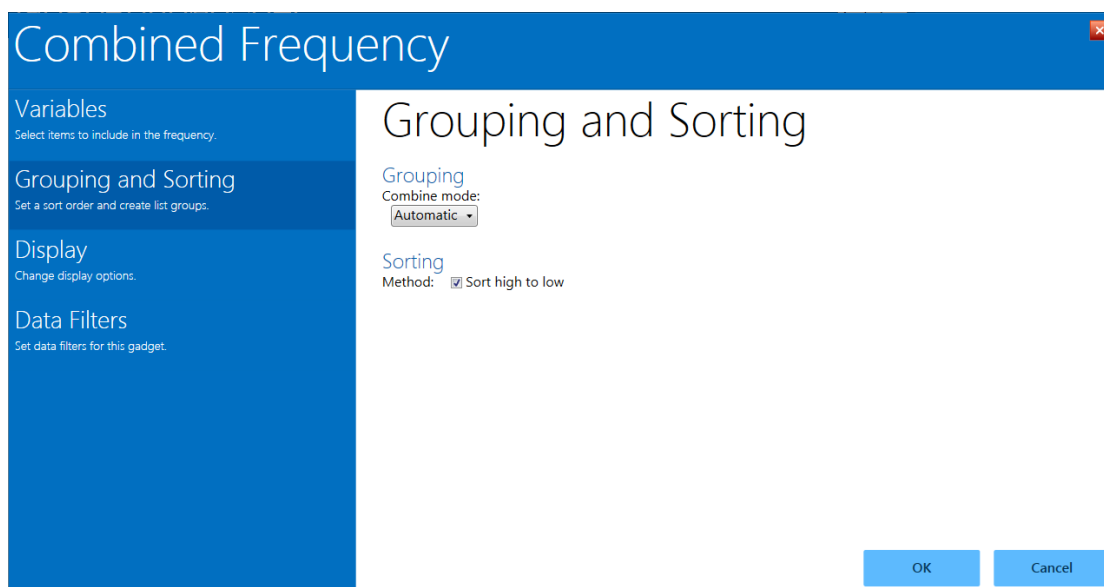


Figure 8.33: Combined Frequency Grouping and Sorting options

4. Click **Grouping and Sorting** and select **Automatic** from the **Combine Mode** drop-down list. Here are the functions of each option in that drop-down:
 - **Boolean**—use this mode for Yes/No and Checkbox fields, or if you wish for the fields in the group to be treated as such.
 - Value to treat as true – This field appears when the Combine Mode is Boolean. The specified value is treated as the true or positive value and all other values are treated as false or negative values.
 - **Categorical**—use this mode if your data are grouped according to similar characteristics in a way that shows the relative frequencies of each group or category.
 - **Automatic**—will automatically determine the appropriate combination mode of either Boolean or Categorical.
5. The **Sort High to Low** checkbox is checked by default. This option will display the results sorted from highest group field frequency to the lowest.
6. Click **Display** and mark your selection for **Show Denominator**. This option will display the total number of records in the results. The box is checked by default.
7. Click **OK**. The **Combined frequency** results appear on the Visual Dashboard canvas.

Value	Frequency	Percent
Fever	311	86.63 %
ILL	276	76.88 %
Headache	269	74.93 %
Vomiting	241	67.13 %
Non-bloody diarrhea	218	60.72 %
Bloody diarrhea	218	60.72 %
Nausea	208	57.94 %
Chills	190	52.92 %
DoctorVisit	180	50.14 %
Irritable	174	48.47 %
Abdominal cramps	172	47.91 %
Hospitalized	137	38.16 %
Poor feeding	125	34.82 %
Antibiotics	114	31.75 %
Died	11	3.06 %

Fields are boolean.
Denominator = 359

Figure 8.34: Combined Frequency results with advanced options

The result displays the frequency and percentage of each symptom check box that is checked and each Yes-No field with a value of “Yes”. Note the denominator included at the bottom of the results window.

Analysis Gadget > M x N/ 2 x 2 Table

Epidemiologists frequently use 2 x 2 tables to examine the relationship between two or more categorical values. In these tables, an exposure variable is usually considered as the risk factor. The outcome variable is considered the disease of consequence (e.g., the person had the disease or outcome of interest, or they did not). Values of the first variable will appear on the left margin of the table, and those of the second will be across the top of the table. Normally, cells contain counts of records matching the values in corresponding marginal labels. The M x N / 2 x 2 Table analysis gadget in Visual Dashboard is demonstrated below:

1. Select the **Sample.PRJ** Data Source. Open the **Oswego** form from the **Data Source Explorer** menu. Click **OK**.
2. Right click on the canvas and select **Add Analysis Gadget > M x N / 2 x 2 table**. The **Crosstabulation Properties** gadget configuration window opens to the **Variables** property panel.
3. From the **Exposure** drop-down list, select **VANILLA**.
4. From the **Outcome** drop-down list, select **ILL**.

Other options in the gadget allow you to add a weight variable or **Stratify** the data by another variable. This example does not use these options.

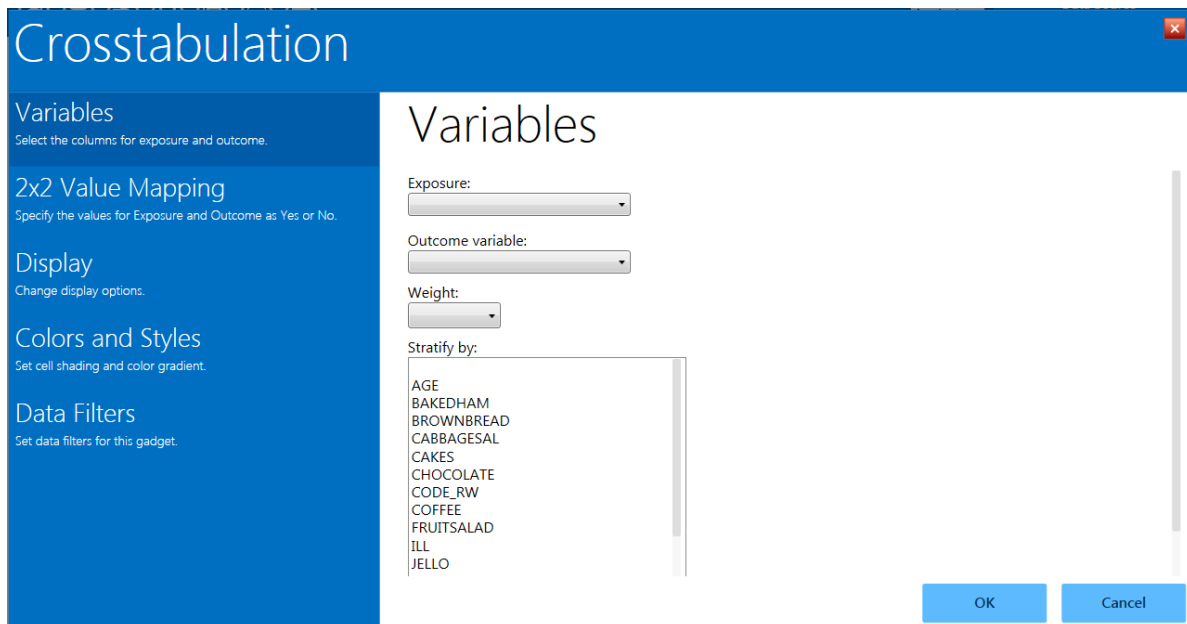


Figure 8.35: Crosstabulation gadget

5. Click **OK**. Results appear on the Visual Dashboard canvas.

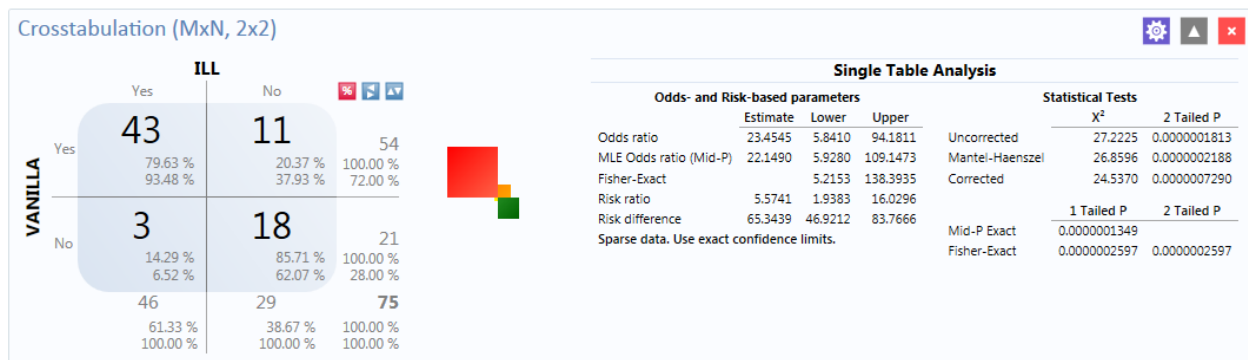


Figure 8.36: Crosstabulation results

The results display a 2 x 2 table and a single table analysis of all records crosstabulated by exposure to vanilla where the outcome was ill. Out of the 75 records in this dataset, 43 patients were exposed to vanilla and became ill. Eleven patients were exposed to vanilla and did not become ill. Of the 21 patients who were not exposed to vanilla, only three became ill.

The Single Table Analysis chart on the right displays Odds Ratios, Risk Ratios, and Statistical Tests results for this example.

Analysis Gadget > Matched pair case-control

The **Matched pair case-control** option in Visual Dashboard is for use with pair-matched case-control studies. Each case must have exactly one matched control. For this reason, it is generally incorrect to perform the analysis using 2x2 tables since it does not take matching into consideration. To use this option, you need to specify an Exposure Variable, the Outcome Variable (i.e., case vs. control), and a Match variable that links each case to its control. The following example demonstrates a matched pair-case control analysis.

The dataset below is from a matched case-control study of the consumption of chicken, where the outcome resulted in the patient becoming ill. The primary exposure was the consumption of chicken (AnyChkn), the case control is **CaCO**, and the **Pair Group ID** is **Matched pairs**.

1. Select the **Case Control DatabaseExample.xlsx** Data Source found in the Sample project folder. Note that this will require updating the **Database Type** to Microsoft Excel 2007 Workbook (.xlsx) before browsing.
2. Select **Section_1_3 & 8\$** form in the **Data Source Explorer** menu.
3. Click **OK**.
4. Select **Add Analysis Gadget > Matched pair case-control**.
5. The **Matched Pair Case-Control Properties** gadget appears.

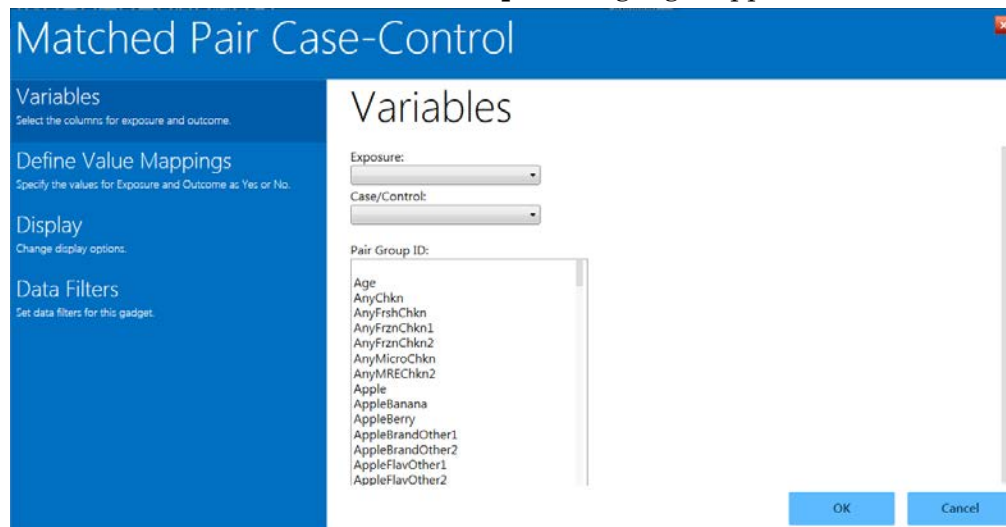


Figure 8.37: Matched Pair Case-Control gadget

6. From the **Exposure** drop-down list, select **AnyChkn**.
7. From the **Case/Control** drop-down list, select **CaCo**.
8. From the **Pair Group ID** list, select **Matched pairs**.
9. Click on **Define Value Mappings**.
10. Click on **1**.
11. Click on the > button on the **Yes values:** section.
12. Click on **0**.
13. Click on the > button on the **No values:** section.
14. Click **OK**.

The **Matched Pair Case-Control** results appear on the canvas.

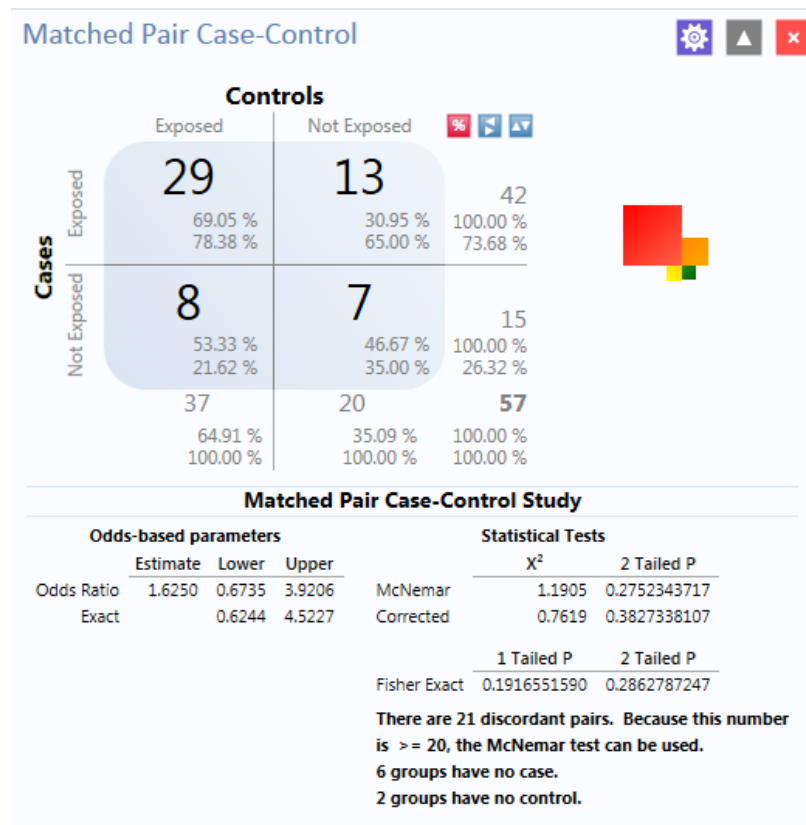


Figure 8.38: Matched Pair Case-Control results

The results display that out of 57 matched pairs, 29 had both the case and the control exposed to AnyChkn, 13 had the case but not the control exposed, eight pairs had the control exposed but the case unexposed, and seven had neither the case nor the control exposed.

The Odds-Ratio is greater than 1, which suggests that eating chicken may be more likely to result in being ill. However, the confidence intervals contain 1 and the P values are less than 0.05, indicating that there is no significant association between eating chicken and becoming ill.

Analysis Gadget > Means

The Means gadget calculates the average for a continuous numeric variable. A **Yes/No** field returns numeric values (Yes=1, No=0), which allows the **Means** gadget to calculate the proportion of respondents answering yes. For this situation, **Analysis Gadget > Means**, has two formats:

- If only one variable is selected, the program produces a table similar to one produced by the Frequency option with descriptive statistics.
 - If two variables are selected, the first variable contains the data to be analyzed and the second indicates how groups will be distinguished. The output of this format is a table similar to one produced by the M x N/ 2 x 2 table with descriptive statistics.
1. Select the **Sample.PRJ** Data Source. Open the **Oswego** form from the **Data Source Explorer** menu.
 2. Click **OK**.
 3. Right click on the canvas and select **Add Analysis Gadget > Means**. The **Means Properties** gadget configuration window opens to the Variables property panel.

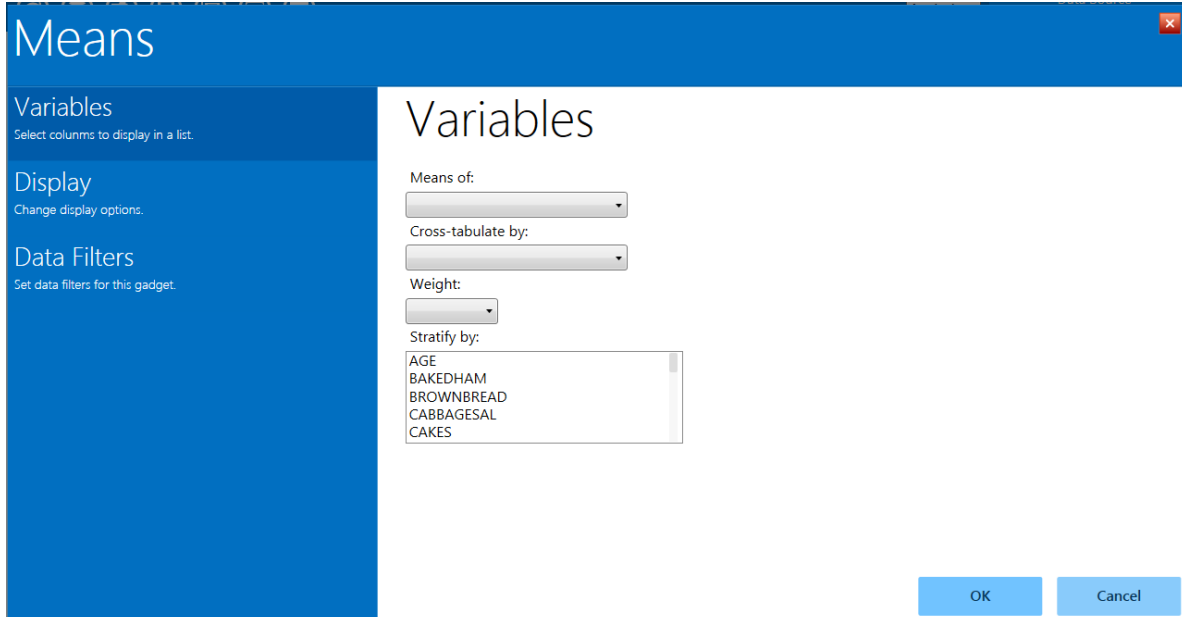


Figure 8.39: Means gadget

4. From the **Means Of** drop-down list, select **Age**.
5. Other choices that can be used in the **Variables** option include:
 - The **Cross-tabulate by** drop-down list contains a variable to help determine if the means of a group are equal.
 - The **Weight** drop-down list contains a variable for weighted analysis.
 - The **Stratify by** list contains a variable to act as a grouping variable.

These options are not used in this example.

6. Click **OK**. The results are displayed on the canvas.

Additional choices for the means gadget are available in the **Display** option and include:

- **Output columns to display**—specifies the columns to include in the current analysis.
- **Decimals to display**—specifies the number of decimals that will be displayed in the results (possible values range from 0–4 and are set to 2 by default). This option does not apply to the descriptive statistics produced when a cross-tabulate variable is used. The descriptive statistics are always displayed to 4 decimal places.
- The **Display ANOVA statistics** checkbox specifies whether to display the analysis of variance statistics. This option is selected by default. If you are working with large data sets and the ANOVA statistics are not required, performance may be improved by turning off (un-checking) this feature.



	Obs	Total	Mean	Var	Std Dev	Min	25%	Median	75%	Max	Mode
AGE	75	2761	36.8133	460.1809	21.4518	3.0000	16.0000	36.0000	58.0000	77.0000	11.0000

Figure 8.40: Means table

The mean age of individuals in the dataset is 36.8.

Analysis Gadget > Duplicates List

The **Duplicates List** gadget in Visual Dashboard is used to list potentially duplicate records. In some situations, as with a patient registry for example, it is important each individual has only one record. The **Duplicates List** gadget allows you to select a few fields that should be unique, and a few other non-unique fields to display. When the gadget runs, it counts the number of occurrences of the key fields for duplicate checking. The gadget displays the message, “No records match the selection criteria” if all rows in the database are unique for the key fields. If any database rows have key fields that match another row, those matching rows are listed in the results.

In the example below, assume a patient registry with a small population, where two individuals sharing the same last name and birthdate is very unlikely. The goal is to check the database to see if there are duplicate records based on the Date of Birth field [DOB] and the Last Name field [LastName]. If we find a match, we might also want to see the few other fields, such as **CaseID**, **DateofInterview**, and **OnsetDate** in order to conclude that individual is actually duplicated in the registry.

1. Open the **EColi.PRJ** project and select the **Food History** form.
2. Right click on the canvas and select **Add Analysis Gadget > Duplicates List**. The **Duplicates List** gadget configuration window opens to the Variables property panel.

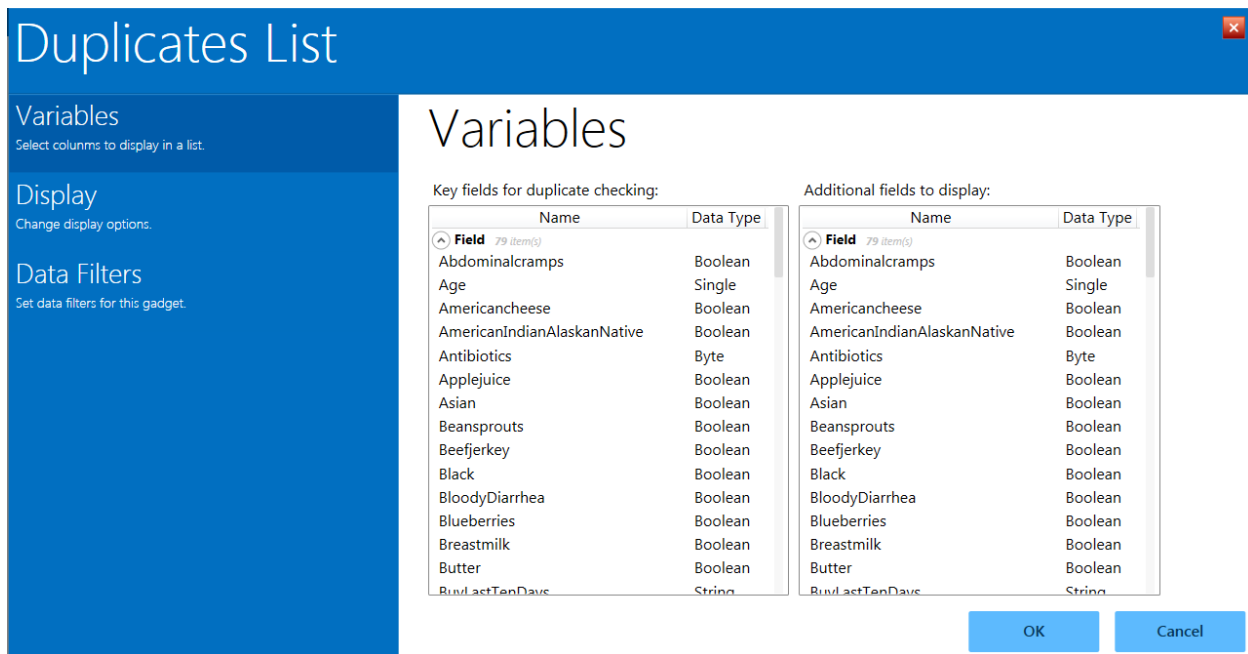


Figure 8.41: Duplicates list gadget

3. In the **Key fields for duplicate checking** list, select **DOB** and **LastName** by holding down the CTRL key and clicking on the respective field names.
4. In the **Additional fields to display** list, select the other fields to help determine if the record is a true duplicate such as **CaseID**, **DateofInterview**, **DoctorVisit**, **DoctorVisitDate**, and **OnsetDate**. Hold down the CTRL key and click of the field names.
5. Click **OK**.
6. The system returns any duplicate records in the database. In this example, patients Johnson and Garcia have duplicate records in the database. They were each interviewed on different days and assigned different case IDs.

Duplicates List							
	DOB	LastName	CaseID	DateofInterview	DoctorVisit	DoctorVisitDate	OnsetDate
1	1/1/1979	Johnson	288	5/13/2011	1		5/6/2011
2	1/1/1979	Johnson	304	5/14/2011	1	4/22/2011	5/7/2011
3	1/1/1984	Garcia	244	5/10/2011	1	4/19/2011	5/3/2011
4	1/1/1984	Garcia	309	5/14/2011	1	4/29/2011	5/7/2011

Figure 8.42: Duplicates list results

Analysis Gadget > Chart

Visual Dashboard produces histograms, scatter plots, pie charts, and bar and line graphs directly from data files. This option produces charts based on the types of data available in the project. The Chart option is capable of generating the following chart types:

- **Column** charts use vertical bars to represent the count or weight for each value of the main variable(s). Each series results in an additional vertical bar at each point.
- **Line** charts connect X and Y data points with straight lines. Each series is represented by a different style of line and both variables must be numeric.
- **Area** charts lay series on top of one another for comparison purposes and can be used to illustrate volume as well as counts.
- **Pie** charts use a circle to represent each series and each value of the main variable has a slice of the circle proportional to the value associated with it.
- **Aberration detection** charts use line charts to monitor changes in the distribution or frequency of health events when compared to historical data.
- **Pareto** charts use vertical bars to represent the count or weight for each value of the main variable(s) and a dashed line to represent the accumulated percentage. Each series results in an additional vertical bar at each point.
- **Scatter** charts display variables along an X-Y axis as a scatter plot. The X variable is the independent variable and Y variable is the dependent variable. Each series is represented by a different point style.
- **Epi Curve** charts use vertical bars to represent the count or weight for each value of the main variable. Each series creates an additional vertical bar at each point. The main variable must be either numeric or date/time. This graph differs from the bar graph because adjacent bars represent equal ranges of the main variable.

The examples below demonstrate how to create each of the charts.

Chart > Column Chart

Follow the steps below to create a bar graph using data from the Sample.PRJ project.

1. Open the **Sample.PRJ** project and select the **Oswego** form.
2. Right click on the canvas and select **Add Analysis Gadget > Charts > Column chart**. The **Column Chart** gadget configuration window opens to the Variables property panel.
3. From the **Main variable** drop-down list, select the variable **Age** to use for the independent variable.
4. Leave the **Weight** and **One graph for each value of** option at the default setting (none).
5. Click **OK**. The graph appears on the canvas, displaying in columns the number of patients by age.

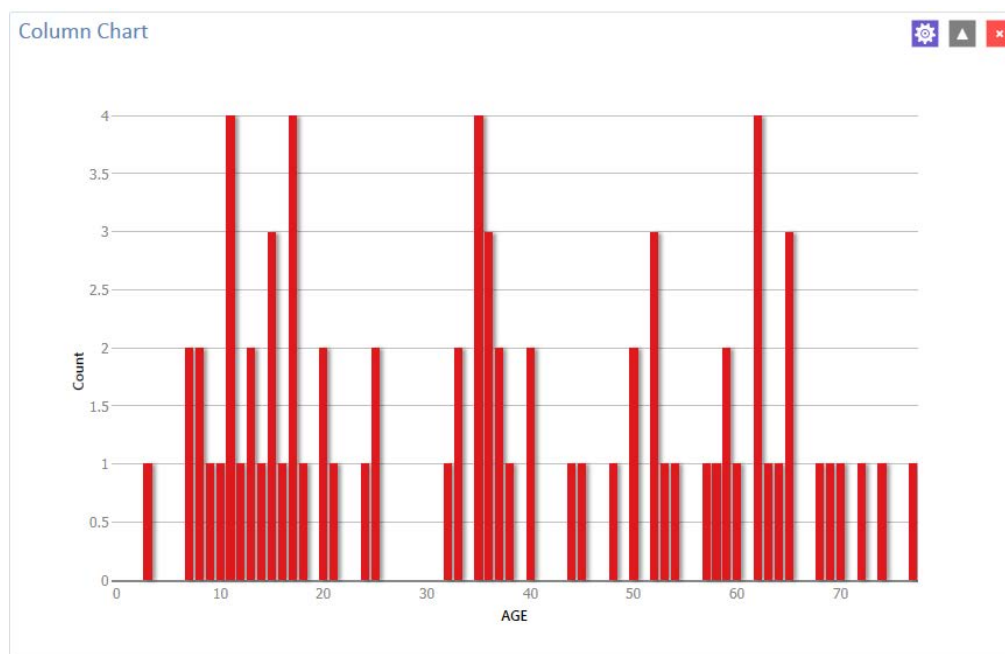


Figure 8.43: Column chart

Chart > Line Chart

Follow the steps below to create a line chart showing counts by date of onset of illness.

1. Open the **EColi.PRJ** project and select the **Food History** form.
2. Right click on the canvas and select **Add Analysis Gadget > Charts > Line chart**. The **Line Chart** gadget configuration window opens to the Variables property panel.
3. Select **Onset Date** from the **Main variables** list.
4. The system returns a line chart of onset dates between April and May from the dataset.

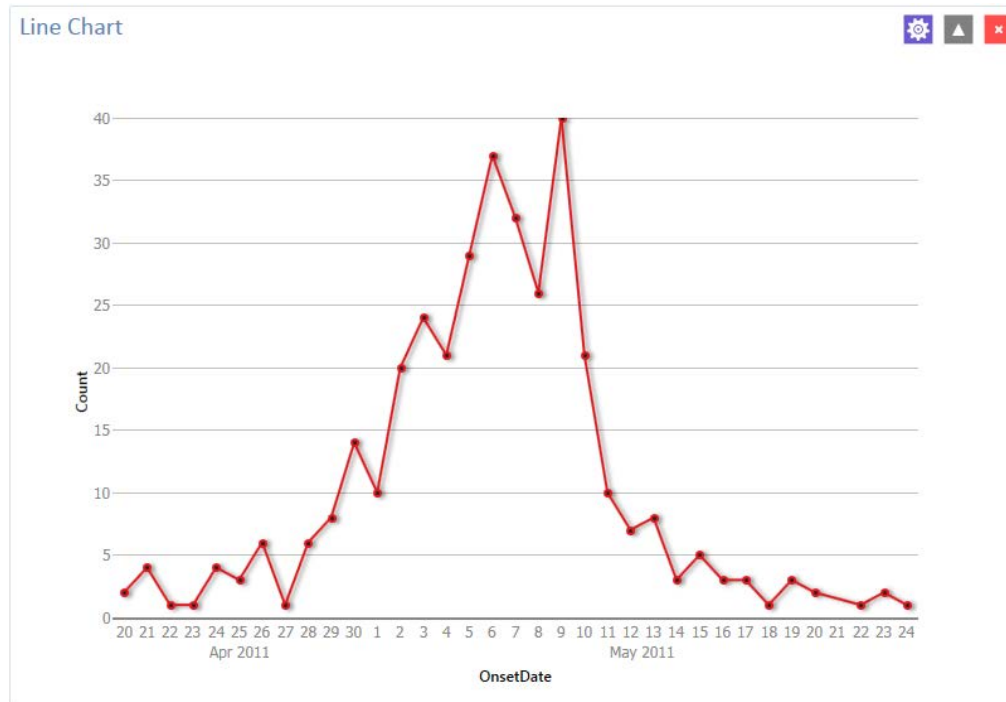


Figure 8.44: Line chart

Chart > Area Chart

Follow the steps below to create an area chart showing fever temperatures of church supper attendees.

1. Open the **EColi.PRJ** project and select the **Food History** form.
2. Right click on the canvas and select **Add Analysis Gadget > Charts > Area chart**. The **Area Chart** gadget configuration window opens to the Variables property panel.
3. Select **OnsetDate** from the **Main variables** list.
4. The system returns a line chart of onset dates between April and May from the dataset.

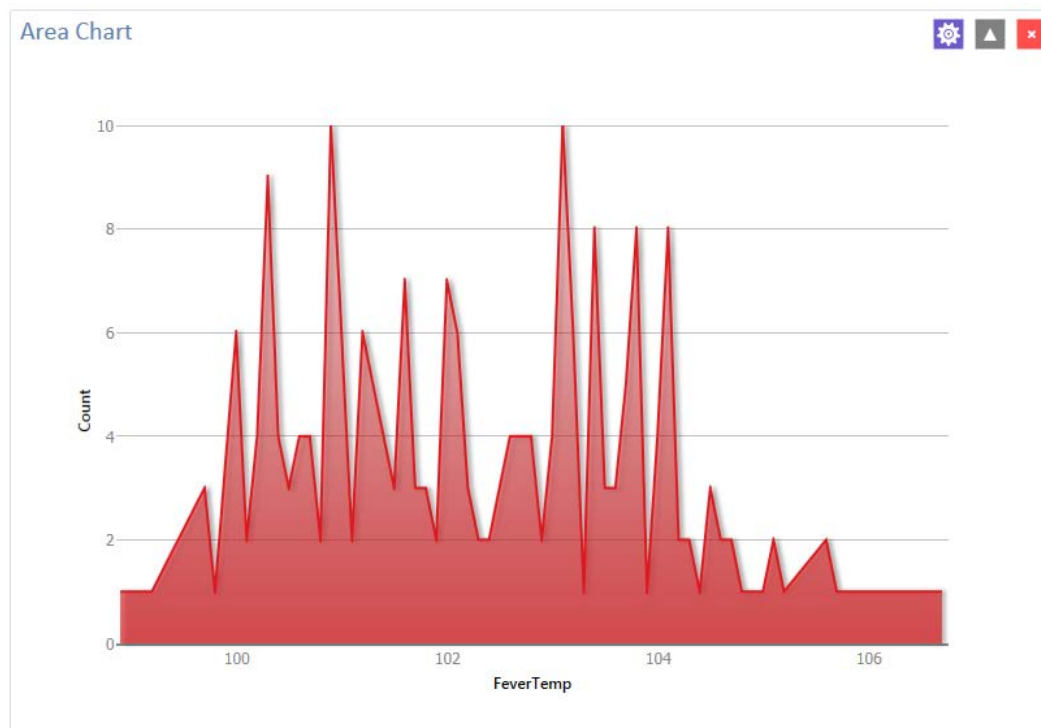


Figure 8.45: Area chart

Chart > Pie Chart

Follow the steps below to create a pie chart showing age groups of church supper attendees. The dataset has individual ages. To create age groups, you can define a new variable with recoded values making age groups from age 0 to age 70 grouped by 10 years per group. Detailed instructions about how to use the Defined Variables gadget are in the Create a New Variable section later in this guide.

1. Open the **Sample.PRJ** project and select the **Oswego** form.

Recode the **Age** variable using the **Defined Variables** gadget.

 - On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget. The gadget expands and becomes fully visible.
 - Click the **New Variable** button.
 - Select **With Recoded Value**. The **Add Recoded Variable** dialog box appears.
 - From the **Source** field drop-down list, select **Age**. Notice the **Destination** field is populated with **Age_RECoded**
 - Click the **Fill Ranges** button. The **Fill Ranges** dialog box appears.
 - Enter **0** for **Start value**, **70** for **End value**, and **10** for **By**.
 - Click **OK**. The **From**, **To**, and **Representation** columns populate within the **Add Recoded Variable** window.
 - Click **OK**. The variables are recoded and are listed in the **Defined Variables** gadget window.
2. Right click on the canvas and select **Add Analysis Gadget > Charts > Pie chart**. The **Pie Chart** gadget configuration window opens to the Variables property panel.
3. From the **Main variable** drop-down list, select the variable **AGE_RECoded**.

4. Leave the **Weight** and **One graph for each value of** option at the default setting (none).
5. Click **OK**. The graph appears on the canvas.

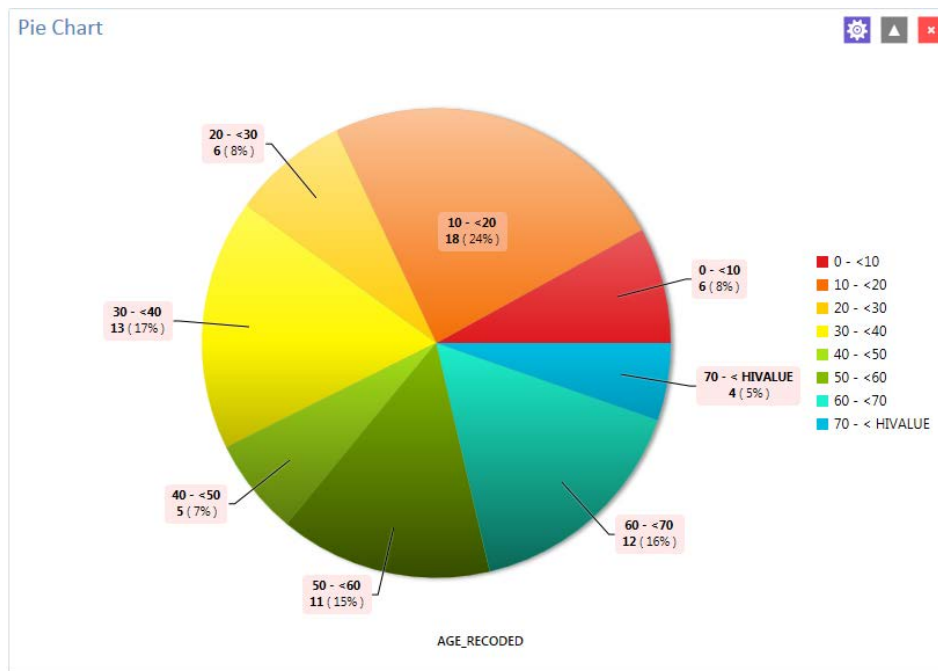


Figure 8.46: Pie chart

The chart shows the age groups of church supper attendees.

Chart > Aberration Detection

The Aberration Detection Analysis gadget can be used with syndromic or national surveillance data. It is designed to monitor changes in the distribution or frequency of health events when compared to historical data. This gadget uses an algorithm implemented by the Early Aberration Reporting System (EARS), which is a tool developed by CDC's Division of Bioterrorism Preparedness and Response to assist local-level health offices with early detection of bioterrorism events or possible outbreaks.

The example described here uses an Excel spreadsheet named EARS.xls that comes with Epi Info™ 7 and is located in the Projects\Sample project folder.

1. Right click on the canvas and choose “Set data source,” or click the data source icon in the toolbar to **Select Data Source**.
2. Click the **Database Type** drop-down to select Microsoft Excel 97-2003.
3. Click **Browse** and then **Browse** again to browse for the EARS.xls file.
4. Browse to the Epi Info 7\Projects\Sample folder and select the **EARS.xls** file.
5. Click on **Sheet 1\$**, then click OK.

- Right click on the canvas and select **Add Analysis Gadget**. Scroll to highlight **Charts**, then select **Aberration Detection**. The **Aberration Detection** gadget configuration window opens to the Variables property panel.

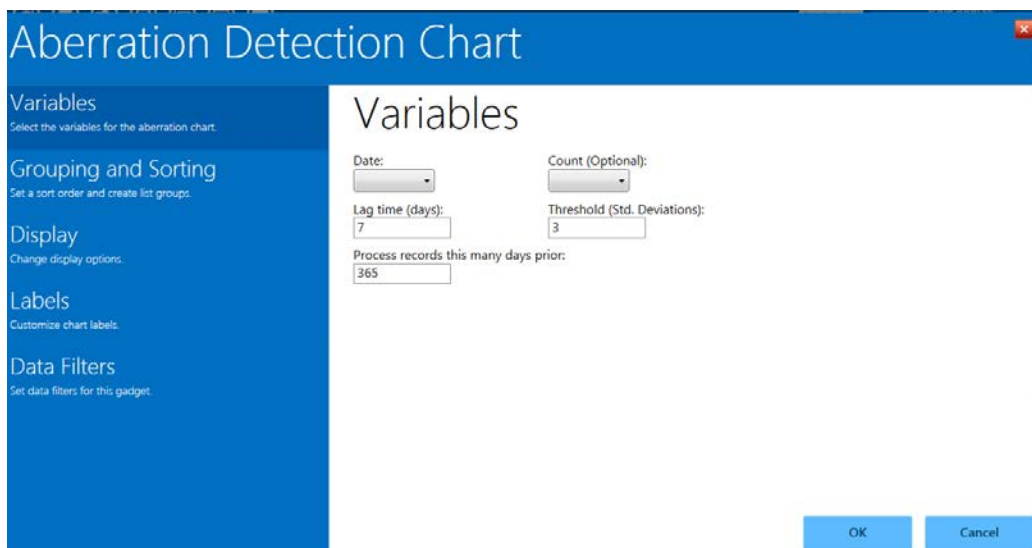


Figure 8.47: Aberration Detection gadget

- From the **Date** drop-down list, select the variable, **DATE**.
- From the **Count (Optional)** drop-down list, select the variable, **COUNT**. Leave the default settings for **Lag Time**, **Threshold**, and **Process Records This Many Days Prior**.
- Click **Grouping and Sorting** to that property panel.
- From the **Indicators** list, click the variable **SYNDROME**.
- Click **OK**. The Aberration Detection line graph displays on the canvas.

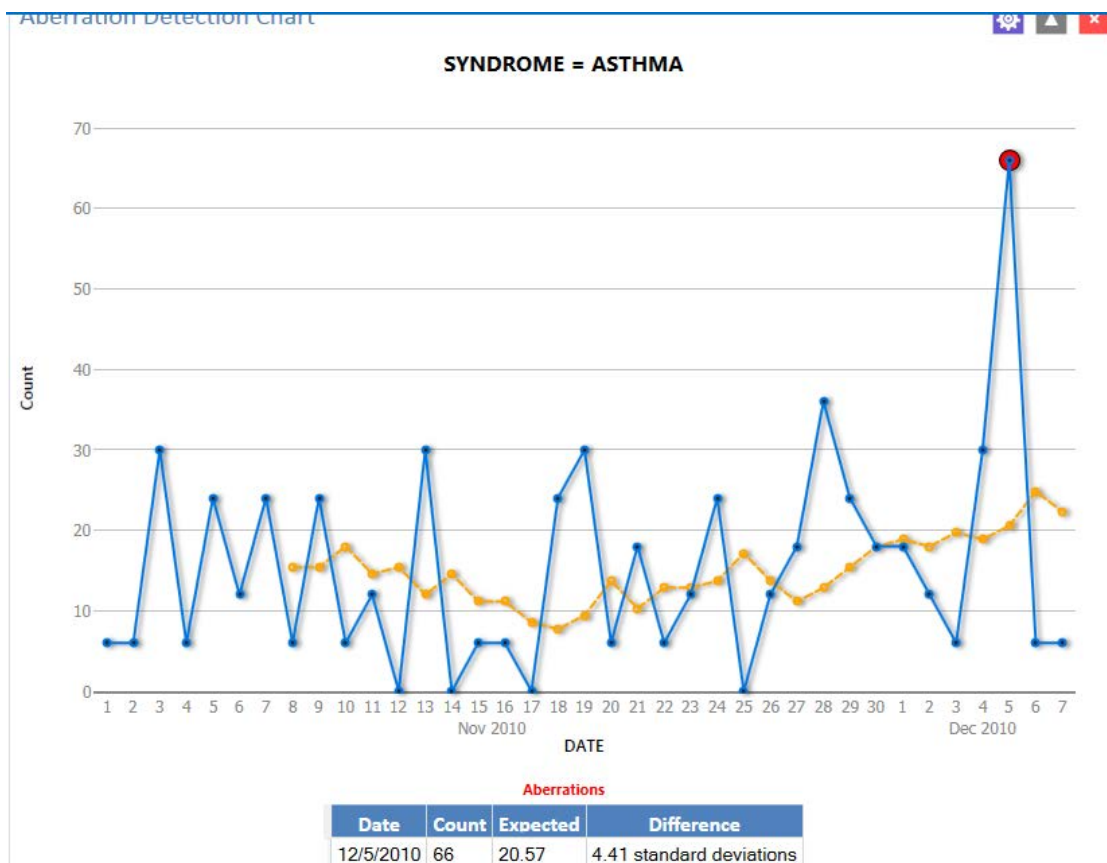


Figure 8.48: Aberration Detection line graph

The chart and table display aberrations found, date, count, expected count, and the standard deviations. The current number of patient visits with Asthma is higher (Count: 66) than the defined threshold (Expected; 20.57), hence the aberration detected on 12/05/2010. Aberrations may need to be evaluated by an epidemiologist (or other public health professional) to determine whether they signify an event of public health importance warranting further investigation.

Chart > Pareto Chart

Follow the steps below to create a Pareto chart showing age groups of church supper attendees. The dataset has individual ages. To create age groups, you can define a new variable with recoded values making age groups from age 0 to age 70 grouped by 2 years per group. Detailed instructions about how to use the Defined Variables gadget are in the Create a New Variable section later in this guide.

1. Open the **Ecoli.PRJ** project and select the **FoodHistory** form.
2. Recode the Age variable using the **Defined Variables** gadget.
 - On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget. The gadget expands and becomes fully visible.
 - Click the **New Variable** button.

- Select **With Recoded Value**. The **Add Recoded Variable** dialog box appears.
 - From the **Source** field drop-down list, select **Age**. Notice the **Destination** field is populated with **Age_RECDED**.
 - Click the **Fill Ranges** button. The **Fill Ranges** dialog box appears.
 - Enter **0** for **Start value**, **70** for **End value**, and **2** for **By**.
 - Click **OK**. The **Add Recoded Variable** window appears with the **From**, **To**, and **Representation** columns populated.
 - Click **OK**.
3. Right click on the canvas and select **Add Analysis Gadget > Charts > Pareto chart**. The **Pareto Chart** gadget configuration window opens to the Variables property panel.
 4. From the **Main variable** drop-down list, select the new variable, **Age_RECDED**.
 5. Click the **Labels** option to customize chart labels and axes.
 6. Update the **X-axis angle** setting to **-90**. This step will permit the date values to display fully on the X-axis.
 7. Click the **Display** option.
 8. Change the width setting to **1000**. If your computer screen is too small, you might need to change the width setting of the chart to a higher number until the system displays the X-axis labels.
 9. Click **OK**. The Pareto chart appears on the canvas.

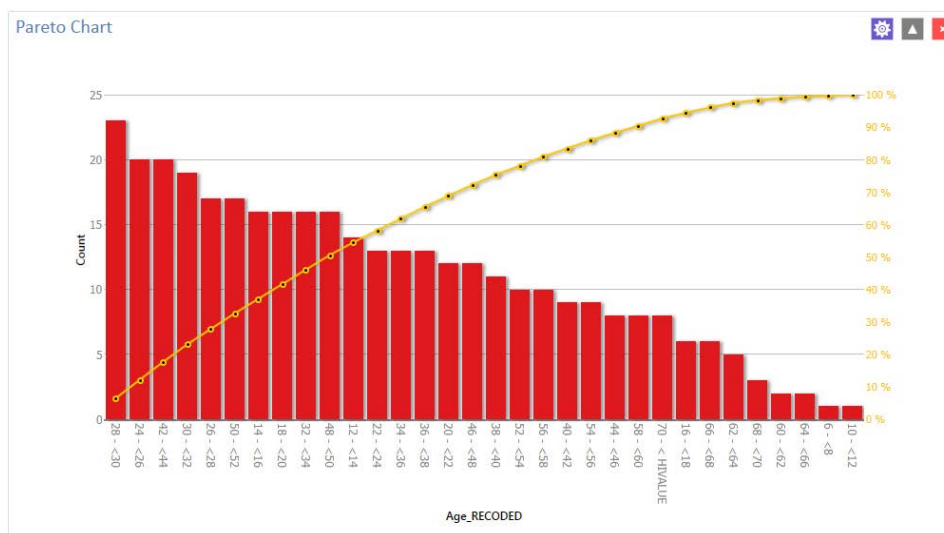


Figure 8.49: Pareto chart

Chart > Scatter Chart

Follow the steps below to create a scatter XY graph using data from the Sample.prj project.

1. Open the **Sample.PRJ** project and select the **EstriolAndBirthweight** form.
2. Right click on the canvas and select **Add Analysis Gadget > Charts > Scatter chart**. The **Scatter Chart** gadget configuration window opens to the Variables property panel.

- From the **Main variable** drop-down list, select the variable **Estriol** to use for the independent variable.
- From the **Outcome variable** drop-down list, select the variable **Birthweight** to use for the independent variable.
- Click **OK**. The graph appears on the canvas, displaying the scatter plot of data points showing the level of Estriol along the X-axis by Birthweight along the Y-axis. A line representing the Linear Regression of the two variables is also included on the chart.

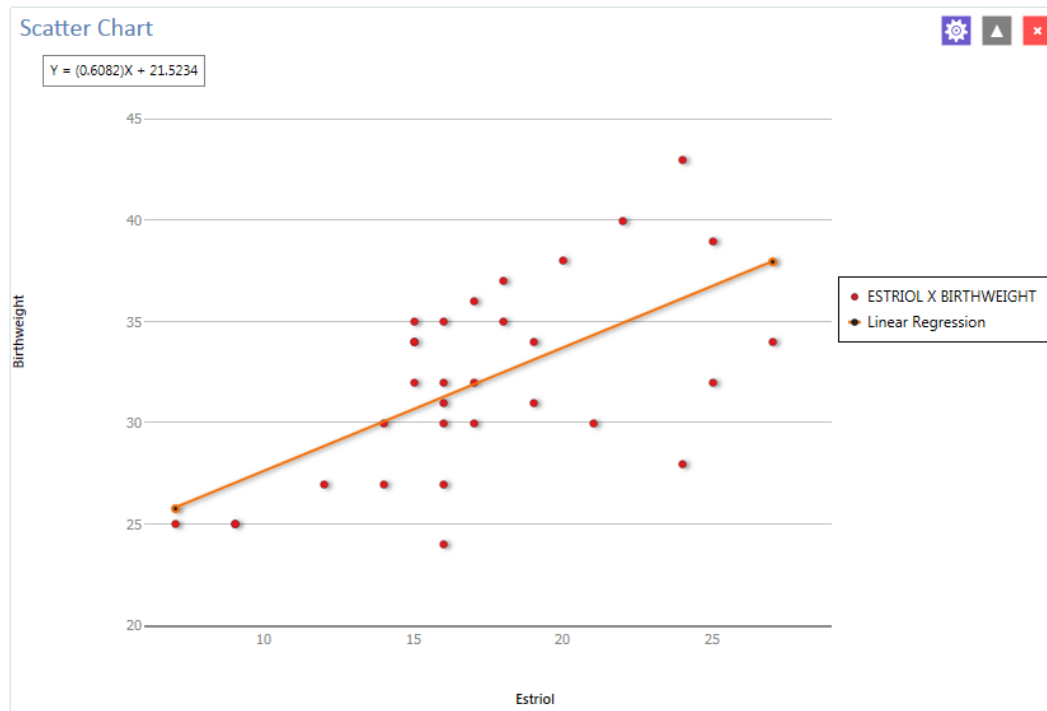


Figure 8.50: Scatter chart

Chart > Epi Curve

Follow the steps below to create an Epi Curve showing the incubation time for an unknown pathogen.

- Open the **Ecoli.PRJ** project and select the **FoodHistory** form.
- Right click on the canvas and select **Add Analysis Gadget > Chart > Epi curve chart**. The **Epi Curve Chart** gadget configuration window opens to the Variables property panel.
- From the **Main variable** drop-down list, select **OnsetDate**.
- Click the **Labels** option to customize chart labels and axes.
- Update the **X-axis angle** setting to -90. This step will permit the date values to display fully on the X-axis.
- Click the **Display** button. The Display property panel appears.

7. Change the width setting to **1000**. If your computer screen is too small, you might need to change the width setting of the chart to a higher number until the system displays the values.
8. Click the **Legend** button. The Legend property panel appears.
9. Click the checkmark to select **Show Legend**.
10. Click the **Grouping and Sorting** button. The Grouping and Sorting property panels appears.
11. In the **Stratify by:** list, scroll down and select the variable Sex.
12. Click **OK**. The Epi Curve chart appears on the canvas. Note that hovering over any bar on the chart will display the date and count for the associated data series.

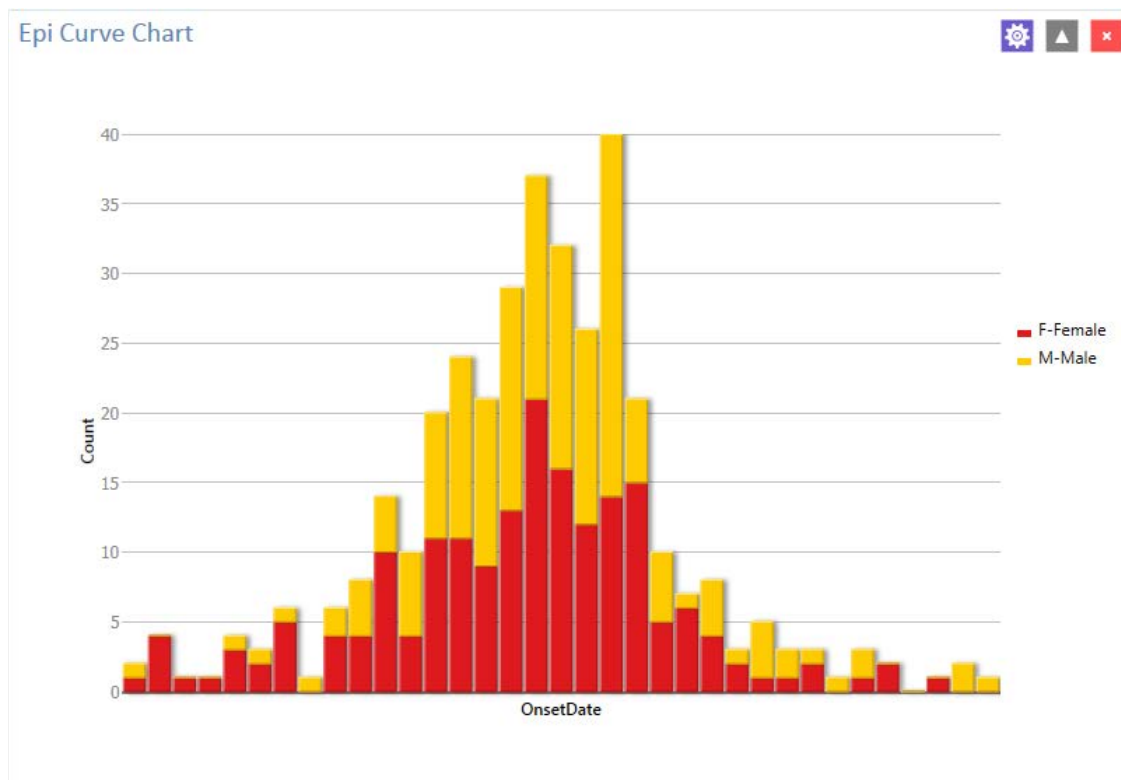


Figure 8.51: Epi Curve chart

Analysis Gadget > Advanced Statistics

Advanced Statistics > Linear Regression

Linear regression is an analysis tool that identifies a relationship between a continuous variable and one or more independent variables. It can be used for simple linear regression (only one independent variable), for multiple linear regression (more than one independent variable), and for quantifying the relationship between two continuous variables (correlation). You can use linear regression when you want to determine the relationship between one dependent variable with one or more independent variables.

In the example below, linear regression is used to determine if systolic blood pressure can be predicted based on an infant's birth weight and age. The dependent variable is systolic blood pressure, and the independent variables are birth weight in ounces and age in days.

1. Set the Data Source to the **Sample.PRJ** project. Open the **BabyBloodPressure** form from the **Data Source Explorer** menu.
2. Click **OK**.
3. Right click on the canvas and select **Add Analysis Gadget > Advanced Statistics > Linear Regression**. The **Regression Properties** gadget configuration window opens to the Variables property panel. (See the subsequent **Linear Regression Properties** section for a screen shot of the gadget.)
4. From the **Outcome Variable** drop-down list, select **SystolicBlood**.
5. From the **Independent variables** drop-down list, select **AgeInDays** and **Birthweight**.
6. Click **OK**. The linear regression results appear in the Visual Dashboard canvas.

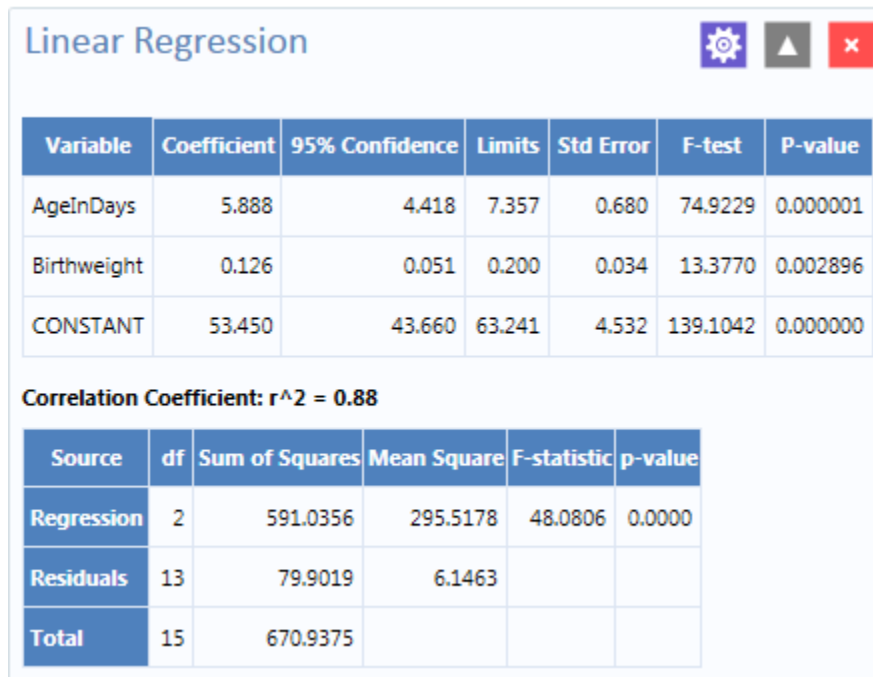


Figure 8.52: Linear Regression results

Coefficients for both independent variables are positive, and the F-statistic associated with each independent variable is highly significant ($p\text{-value} < 0.01$). These results suggest that each variable (AgeInDays and Birthweight) is a predictor of higher systolic blood pressure.

Linear Regression Properties

Figure 8.53: Linear Regression gadget

- The **Outcome variable** is the dependent variable for the regression. The outcome variable must be numeric or a Yes/No field.
- A **Weight** variable may be selected to use in weighted analyses.
- The **Confidence Limits** option specifies the probability level at which Epi Info computes confidence limits.
- If the user checks the **No Intercept** box, the regression is performed without a constant term, which forces the regression line through the origin.
- The **Fields** drop-down list contains the predictor (independent) variables.
- When selected, the predictor variables will appear in the list box below the **Fields** drop-down. Double click on a variable to remove it from the list box.
- Click on a variable from the **Fields** list box to activate the **Make Dummy** button. If you click **Make Dummy**, the selected variable will appear in the **Dummy Variable** list box. Double click on a variable to remove it from the list box.
- **Interaction Terms** are defined with the **Make Interaction** button. The **Make Dummy** button turns to **Make Interaction** when two or more variables are selected at the same time from the **Fields** list box (hold the CTRL key to select more than one variable in this list). If you click **Make Interaction**, the relationship populates the **Interaction terms** list box. **Make Interaction** adds all possible combinations of the selected variables to the regression as interaction terms.

Advanced Statistics > Logistic Regression

In Epi Info™ 7, either the **M x N/ 2 x 2** table or **Logistic Regression** may be used when the outcome variable is dichotomous (for example, Ill–Yes/ Ill–No). However, an **M x N/ 2 x 2** analysis is useful only when there is one “risk factor,” or explanatory variable. **Logistic Regression** is useful when the number of explanatory variables (“risk factors”) is more than one. **Logistic Regression** shows the relationship between an outcome variable with two values and explanatory variables that can be categorical or continuous. To use **Logistic Regression**, the dependent (outcome) variable must be dichotomous (can be divided into only two parts) Yes/No or numeric 0/1. The independent (other variables) can be numeric, categorical, or Yes/No variables.

Records with missing values are excluded from **Logistic Regression** analyses. If the **Include Missing** option is used with missing values and Yes/No fields, dummy variables will generate automatically, contributing Yes vs. Missing and No vs. Missing. Independent variables of text type are automatically turned into dummy variables, which compare each value relative to the lowest value in the sort order. Date or numeric type independent variables are treated as continuous variables unless they are set as dummy variables, which compare each value relative to the lowest value.

The following example uses logistic regression to determine the odds ratio of six foods that could be the cause of a hypothetical foodborne illness. The dependent variable is Ill, and the independent variables are brownbread, cabbage, water, milk, chocolate, and vanilla.

1. Set the Data Source to the **Sample.PRJ** project. Open the **Oswego** form from the **Data Source Explorer** menu.
2. Click **OK**.
3. Right click on the canvas and select **Add Analysis Gadget > Advanced Statistics > Logistic Regression**. The **Logistic Regression** gadget configuration window opens to the Variables property panel. (See the subsequent **Logistic Regression Properties** section for a screen shot of the gadget.)
4. From the **Outcome** drop-down list, select **ILL**.
5. From the **Independent variables** drop-down list, select **BROWNBREAD**, **CABBAGESAL**, **WATER**, **MILK**, **CHOCOLATE**, and **VANILLA**.
6. Click **OK**. Results appear on the Visual Dashboard canvas.

Logistic Regression							
Term	Odds Ratio	95%	C.I.	Coefficient	S.E.	Z-Statistic	P-Value
BROWNBREAD (Yes/No)	1.7790	0.3929	8.0557	0.5761	0.7706	0.7476	0.4547
CABBAGESAL (Yes/No)	1.1341	0.2818	4.5643	0.1258	0.7104	0.1771	0.8594
MILK (Yes/No)	0.1343	0.0068	2.6659	-2.0077	1.5246	-1.3168	0.1879
WATER (Yes/No)	1.1120	0.2670	4.6319	0.1062	0.7280	0.1458	0.8841
VANILLA (Yes/No)	25.9752	5.4652	123.4566	3.2571	0.7953	4.0955	0.0000
CHOCOLATE (Yes/No)	1.0974	0.3024	3.9827	0.0930	0.6577	0.1414	0.8876
CONSTANT	*	*	*	-2.1266	0.9733	-2.1850	0.0289

Convergence: Converged
Iterations: 4
Final -2*Log-Likelihood: 69.2504
Cases Included: 74

Test	Statistic	D.F.	P-Value
Score	28.0180	6	0.0001
Likelihood Ratio	29.8484	6	0.0000

Figure 8.54 Logistic Regression results

The results show that vanilla has an Odds Ratio and Confidence Interval significantly greater than one. This indicates that consumption of vanilla was likely the cause of foodborne illness.

Logistic Regression Properties

Figure 8.55 Logistic Regression gadget

- The **Outcome variable** is the dependent variable for the regression. The outcome variable must be numeric or a Yes/No field.
- A **Weight** variable may be selected to use in weighted analyses.
- **Match variable** identifies the variable indicating the group membership of each record.
- The **Confidence limits** option specifies the probability level at which Epi Info computes confidence limits.
- If the user checks the **No intercept** box, the regression is performed without a constant term.

- The **Include missing values** setting controls independent variables. If **Include missing values** is used with missing values and true/false, dummy variables will be created automatically. This will contribute Yes vs. Missing and No vs. Missing. Independent variables of text type are automatically turned into dummy variables that compare each value relative to the lowest value in the sort order. Date or numeric type Independent variables are treated as continuous variables unless designated as dummy values using the **Make dummy** button. If that occurs, they automatically turn into dummy variables, which compare each value relative to the lowest value.
- The **Independent variables** drop-down list contains the predictor (independent) variables.
- When selected, the predictor variables will appear in the list box below the **Independent variables** drop-down. Double click on a variable to remove it from the **Other** variables list box. Select a variable from the **Independent variables** list box to activate the **Make Dummy** button. If you click **Make dummy**, the selected variable will then appear in the **Dummy variables** list box. Double click on a variable to remove it from the **Dummy variables** list box.
- **Interaction terms** are defined with the **Make interaction** button. The **Make dummy** button turns to **Make interaction** when two or more variables are selected at the same time from the **Fields** list box (hold the CTRL key to select more than one variable in this list). If you click **Make interaction**, the relationship populates the **Interaction** terms list box. **Make interaction** adds all possible combinations of the selected variables to the regression as interaction terms. Double click on a variable to remove it from the **Interaction terms** list box.
- The **Clear terms** button clears all variables from the **Dummy variables** and **Interaction terms** list boxes.

Advanced Statistics > Complex Sample Frequencies, Tables, and Means

The Frequency, Table, and Means options in Epi Info™ 7 perform statistical calculations assuming the data were collected using simple random sampling or unbiased systematic sampling. Many surveys also make use of more complex sampling strategies such as stratification, cluster sampling, and the use of unequal sampling fractions. Visual Dashboard provides three options to analyze complex sample data: **Complex Sample Frequencies**, **Complex Sample Means**, and **Complex Sample Tables**.

Generally, in complex sample analysis, there is a variable for the primary sampling unit (PSU) or Cluster from which a sample subject was selected. If the PSUs were chosen from different Strata (e.g., states or counties), there may be a stratification variable (Stratify by). The concept of sample stratification in complex sample design differs from the concept of stratification during epidemiologic analysis using the TABLES command, as the Strata are chosen in the sampling process before analysis. In addition, a weight variable (Weight) is used when sampling strategies result in unequal selection probabilities. The complex

sample commands in Epi Info™ 7 can compute proportions or means with standard errors and confidence limits. If a 2x2 table is requested, the odds ratio, risk ratio, and risk difference are provided.

Complex Sample Frequencies

The **Complex samples frequencies** option produces frequency tables for selected variables. The Sample project contains an Expanded Program for Immunization (EPI) cluster survey. Using the EPI method, a team selected 30 communities (i.e., clusters) from the chosen geographic area and visited each of the 30 communities. In each, they selected seven children in an appropriate age range and determined each child's immunization (VAC) status. The following example will determine the frequency of vaccinations using **Complex Sample Frequencies**.

1. Set the Data Source to the **Sample.prj** project.
2. Select the **Epi1** form from the **Data Source Explorer** menu.
3. Right click on the canvas and select **Add Analysis gadget > Advanced statistics > Complex sample frequencies**.
4. The **Complex Sample Frequencies** configuration window opens to the Variables property panel.

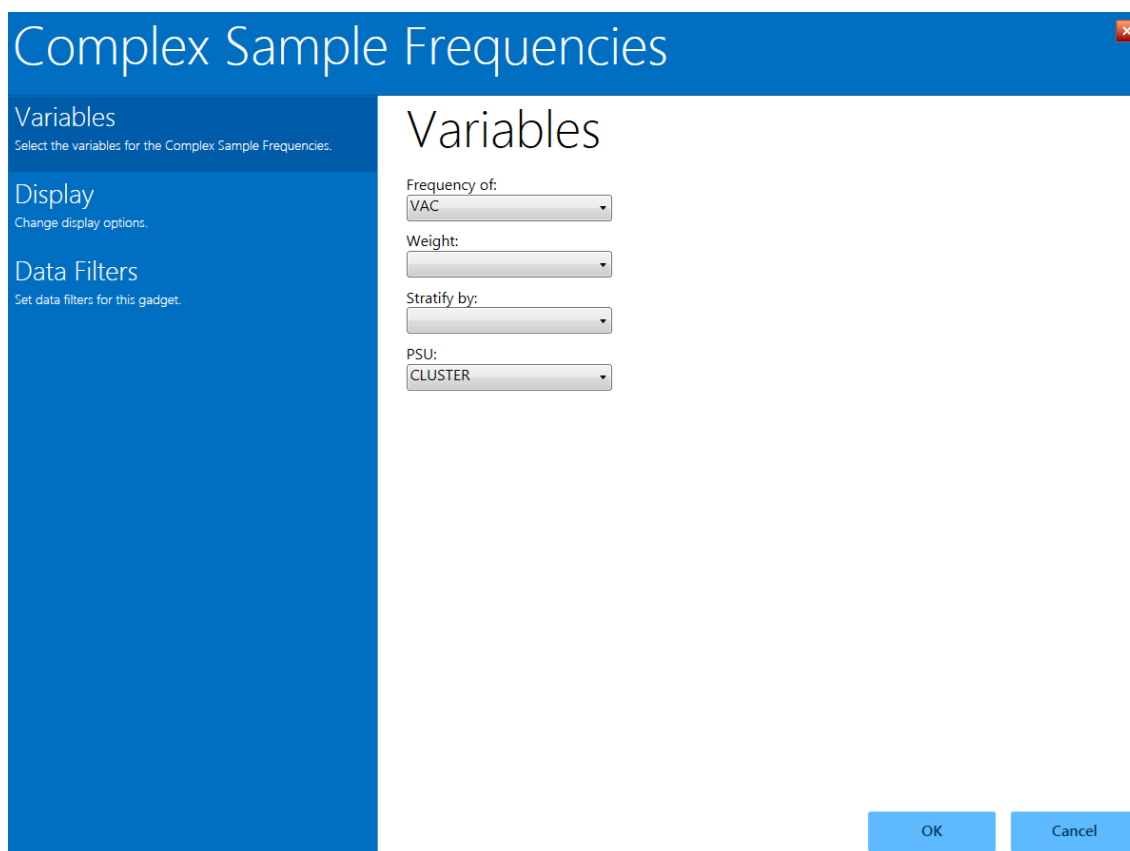


Figure 8.56: Complex Sample Frequencies gadget

Complex Sample Frequencies options are:

- **Frequency of** identifies the variable(s) whose frequency is computed.
 - A **Weight** variable is selected for use in weighted analyses.
 - **Stratify by** identifies the variable to be used to stratify or group the frequency data.
 - The **PSU** identifies the Primary Sampling Unit.
5. From the **Frequency Of** drop-down list, select **VAC (the number of vaccinations received)**. This will calculate the frequency of vaccinations in the data set.
 6. From the **PSU** drop-down list, select **CLUSTER (the population group for each case)**.
 7. Click **OK**.
 8. The **Complex Sample Frequencies** appear on the canvas.

VAC	TOTAL
1	155
Row %	100.000
Col %	73.810
SE %	4.599
LCL %	64.403
UCL %	83.216
2	55
Row %	100.000
Col %	26.190
SE %	4.599
LCL %	16.784
UCL %	35.597
TOTAL	210
Design effect	2.2868

Figure 8.57: Complex Sample Frequencies results

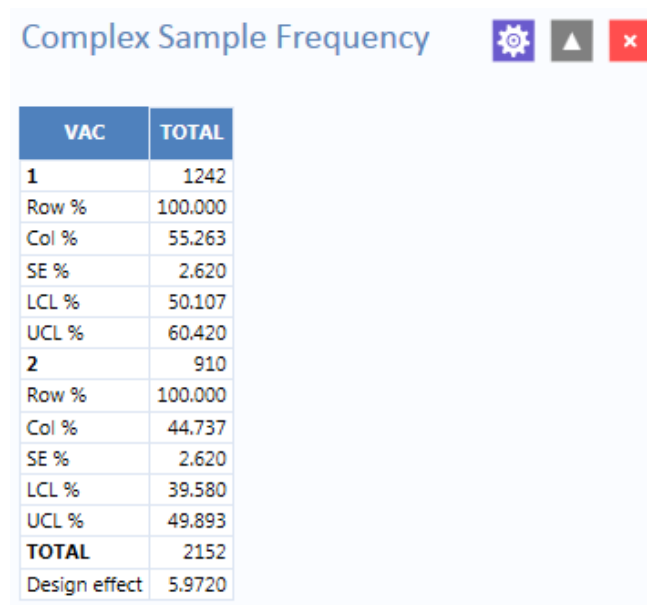
Information provided in the output includes:

- **Row %** – the row percentage; frequency will always be 100% because the total number of cases will always be 100%.
- **Col %** – the column percentage. This number represents the frequency computation; in this case, children with 1 vaccination represent 73.810 % and children with 2 vaccinations represent 26.190% out of the 210 total cases.
- **SE %** – the standard error, which takes into account the complex sample design.
- **LCL %** – Lower Confidence Limit.
- **UCL %** – Upper Confidence Limit.
- **Total** – total number of individuals/elements surveyed.

The results indicate that 73.8% of the 210 children surveyed are vaccinated, with a 95% confidence interval range from 64.4% to 83.2%.

The following example is similar to the one above, except that this is a stratified cluster survey with a separate 30-cluster survey completed in each of 10 strata. To analyze this dataset correctly, we will take into account where each child lives (LOCATION). We will also use a weight variable to account for the differences in population sizes between the different locations.

1. Set the Data Source to the **Sample.prj** project.
2. Select the **Epi10** form from the **Data Source Explorer** menu.
3. Right click on the canvas and select **Add Analysis gadget > Advanced statistics > Complex sample frequencies**.
4. The **Complex Sample Frequencies** configuration window opens to the Variables property panel.
5. From the **Frequency Of** drop-down list, select **VAC**.
6. From the **Weight** drop-down list, select **POPW (the percent of the total population in each cluster)**.
7. From the **Stratify By** drop-down list, select **LOCATION**.
8. From the **PSU** drop-down list, select **CLUSTER**.
9. Click **OK**.
10. The **Complex Sample Frequencies** results appear on the canvas.



VAC	TOTAL
1	1242
Row %	100.000
Col %	55.263
SE %	2.620
LCL %	50.107
UCL %	60.420
2	910
Row %	100.000
Col %	44.737
SE %	2.620
LCL %	39.580
UCL %	49.893
TOTAL	2152
Design effect	5.9720

Figure 8.58: Complex Sample Frequencies stratified and weighted results

The results indicate that 55.3% of the 2,152 children surveyed are vaccinated, with a 95% confidence interval range from 50.1% to 60.4%.

Advanced Statistics > Complex Sample Means

The **Complex Sample Means** command can be used when the outcome variable is continuous, such as age, cholesterol level, etc. You can either calculate an overall mean with its measures of variation or compare means across a grouping variable.

As an example of calculating means with a grouping variable, use the data from the Smoke form located in the Sample project folder. In this example, the investigator is interested in determining if, among smokers, there is a difference in the average number of cigarettes smoked between males and females. In these data, the variable SEX is coded as 1=male and 2=female.

1. Set the Data Source to the **Sample.prj** project.
2. Select the **Smoke** form from the **Data Source Explorer** menu.
3. Right click on the canvas and select **Add Analysis gadget > Advanced statistics > Complex sample means**.
4. The **Complex Sample Means** configuration window opens to the Variables property panel.

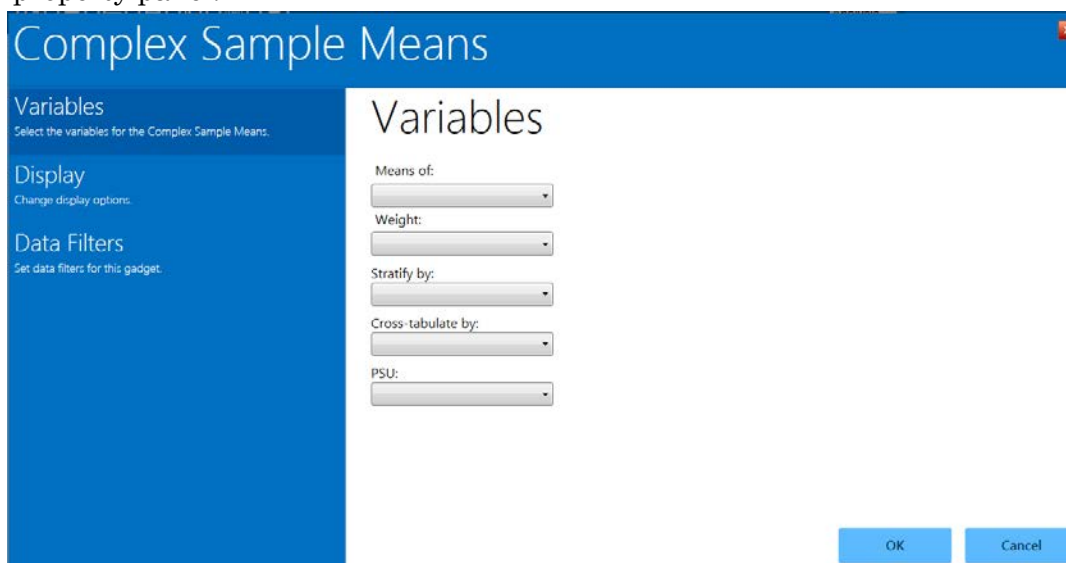
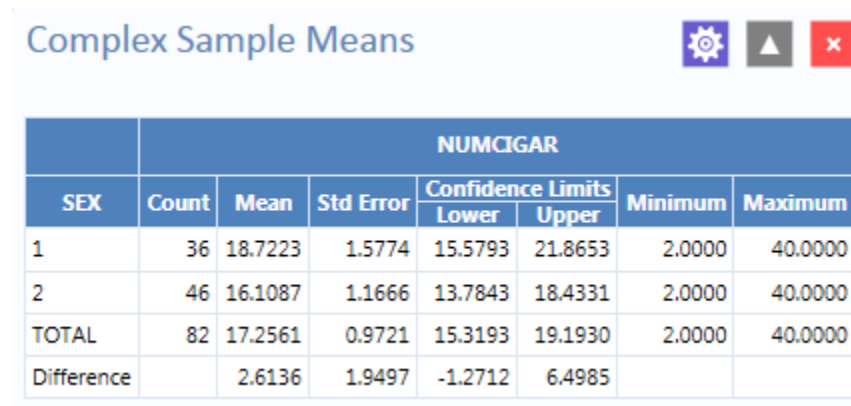


Figure 8.59: Complex Sample Means gadget

Complex Sample Means options:

- **Means of** identifies the variable for which the mean is being computed.
 - A **Weight** variable is selected for use in weighted analyses.
 - **Stratify by** identifies the variable to be used to stratify or group the frequency data.
 - **Cross-tabulate by** identifies the variable to be used to cross-tabulate the main variable.
 - The **PSU** identifies the Primary Sampling Unit.
5. From the **Means Of** drop-down list, select **NUMCIGAR (number of Cigarettes smoked)**.

6. From the **Weight** drop-down list, select **SAMPW (calculated based on sample size and population size)**.
7. From the **Stratify By** drop-down list, select **Strata**.
8. From the **Cross-tabulate By** drop-down list, select **Sex**.
9. From the **PSU** drop-down list, select **PSUID**.
10. Click **OK**.
11. The **Complex Sample Means** results appear on the canvas.



Complex Sample Means							
NUMCIGAR							
SEX	Count	Mean	Std Error	Confidence Limits		Minimum	Maximum
				Lower	Upper		
1	36	18.7223	1.5774	15.5793	21.8653	2.0000	40.0000
2	46	16.1087	1.1666	13.7843	18.4331	2.0000	40.0000
TOTAL	82	17.2561	0.9721	15.3193	19.1930	2.0000	40.0000
Difference		2.6136	1.9497	-1.2712	6.4985		

Figure 8.60: Complex Sample Means results

The results indicate that among the 82 individuals who smoked cigarettes, the average number of cigarettes smoked per day (NUMCIGAR) for men was 18.7, and 16.1 for women, with a 95% confidence interval range of 15.6 to 21.8 for men and 13.7 to 18.4 for women. Note that the Smoke file has 337 individuals (this value is shown in the upper right hand corner of the canvas). However, the number of cigarettes smoked per day has data for only the 82 smokers. For nonsmokers, this variable was left blank and is therefore treated as missing data and excluded from analysis.

Advanced Statistics > Complex Sample Tables

The **Complex Sample Tables** option in Visual Dashboard allows you to specify an Exposure Variable and an Outcome Variable. The following example uses **Complex sample tables** to show whether the mother receiving prenatal care (PRENATAL) has an effect on the child's vaccination status. If the mother had received prenatal care, PRENATAL=1, else PRENATAL=2.

1. Set the Data Source to the **Sample.prj** project.
2. Select the **Epi10** form from the **Data Source Explorer** menu.
3. Right click on the canvas and select **Add Analysis gadget > Advanced statistics > Complex sample tables**.
4. The **Complex Sample Tables** configuration window opens to the Variables property panel.

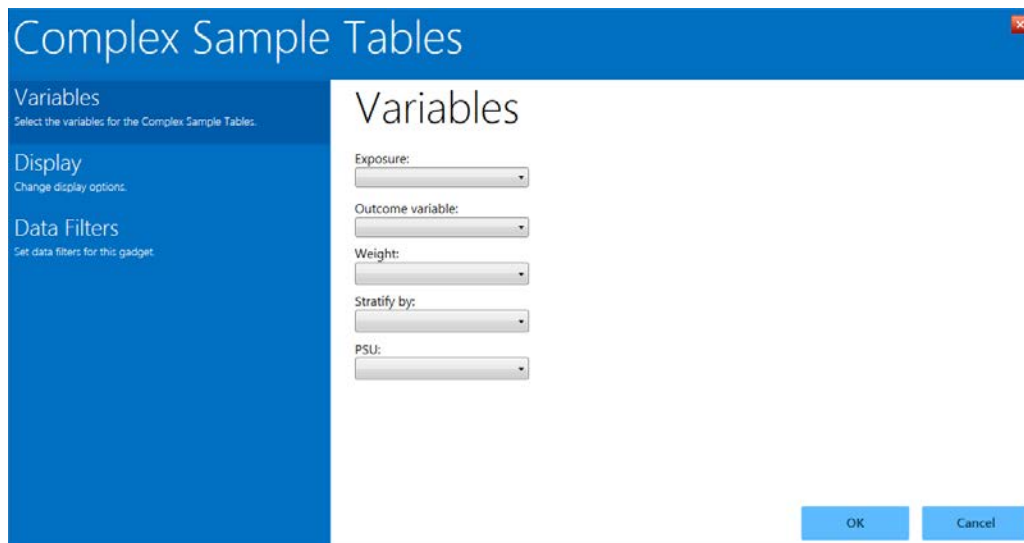





Figure 8.61: Complex Sample Tables gadget

Complex Sample Tables options:

- **Exposure** identifies the variable that will appear on the horizontal axis of the table. It is considered to be the risk factor.
 - **Outcome Variable** identifies the variable that will appear on the vertical axis of the table.
 - A **Weight** variable is selected for use in weighted analyses.
 - **Stratify by** identifies the variable to be used to stratify or group the frequency data.
 - The **PSU** identifies the Primary Sampling Unit.
5. From the **Exposure** drop-down list, select **PRENATAL**.
 6. From the **Outcome** drop-down list, select **VAC**.
 7. From the **Weight** drop-down list, select **POPW (the percent of the total population in each cluster)**.
 8. From the **Stratify By** drop-down list, select **Location**.
 9. From the **PSU** drop-down list, select **Cluster**.
 10. Click **OK**.
 11. The **Complex Sample Tables** results appear on the canvas. You may need to scroll down to see the full results.

Complex Sample Tables   

PRENATAL	VAC		TOTAL
	1	2	
1	675	413	1088
Row %	60.734	39.266	100.00 %
Col %	76.817	61.349	69.90 %
SE %	3.375	3.375	
LCL %	54.091	32.622	
UCL %	67.378	45.909	
Design Effect	5.193	5.193	
2	567	497	1064
Row %	42.560	57.440	100.00 %
Col %	23.183	38.651	30.10 %
SE %	2.414	2.414	
LCL %	37.808	52.689	
UCL %	47.311	62.192	
Design Effect	2.535	2.535	
TOTAL	1242	910	2152
Row %	55.263	44.737	100.00 %
Col %	100.000	100.000	100.00 %
SE %	2.620	2.620	
LCL %	50.107	39.580	
UCL %	60.420	49.893	
Design Effect	5.972	5.972	

**Complex Sample Design
Analysis of 2x2 Table**

Odds Ratio (OR)	2.0875
Standard Error (SE)	0.3074
95% Conf. Limits	(1.502, 2.901)
Risk Ratio (RR)	1.4270
Standard Error (SE)	0.1096
95% Conf. Limits	(1.227, 1.660)
Risk Difference (RD%)	18.1744
Standard Error (SE)	4.0213
95% Conf. Limits	(10.260, 26.089)

Figure 8.62: Complex Sample Tables results

The results show that 60.7% of children whose mothers received prenatal care were vaccinated, compared to 42.6% of those children whose mothers did not receive prenatal care.

The 2 x 2 data shows that the odds ratio in the data was 2.088, the risk ratio was 1.427 and this risk difference is 18.2%. The prevalence (risk) ratio indicates that 1.427 times as many children of women who received prenatal care were vaccinated (60.734% divided by 42.560% = 1.427), compared to children born to women who had not received prenatal care—a 40% difference.

Open Canvas

Open an existing Visual Dashboard canvas by selecting the Open icon on the Canvas Tools Toolbar or right clicking on the canvas to open the canvas Options dialog. Epi Info 7 includes several sample canvas files such as the Oswego.cvs7 canvas.



Figure 8.63: Visual Dashboard toolbar, with Open icon circled

The browse arrow opens a dialog to navigate to a saved canvas file, if it is not in the Quick Access List.

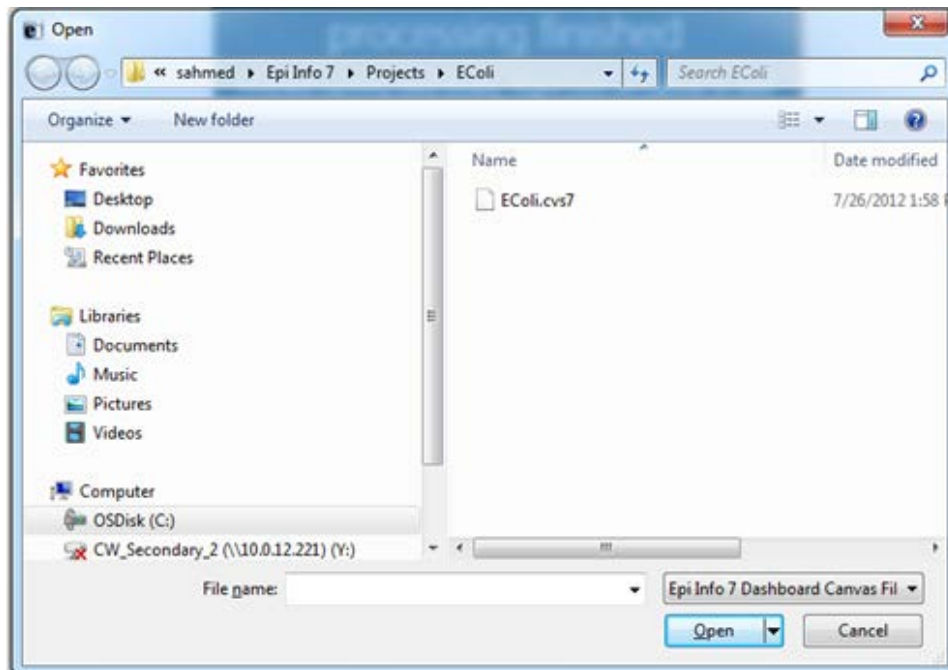


Figure 8.64: Open a Canvas

1. Select the location of the canvas file. The file can be identified by the file name extension (.cvs7).
2. Click **Open**. The canvas appears in Visual Dashboard, connects to the data source, and displays the latest results.

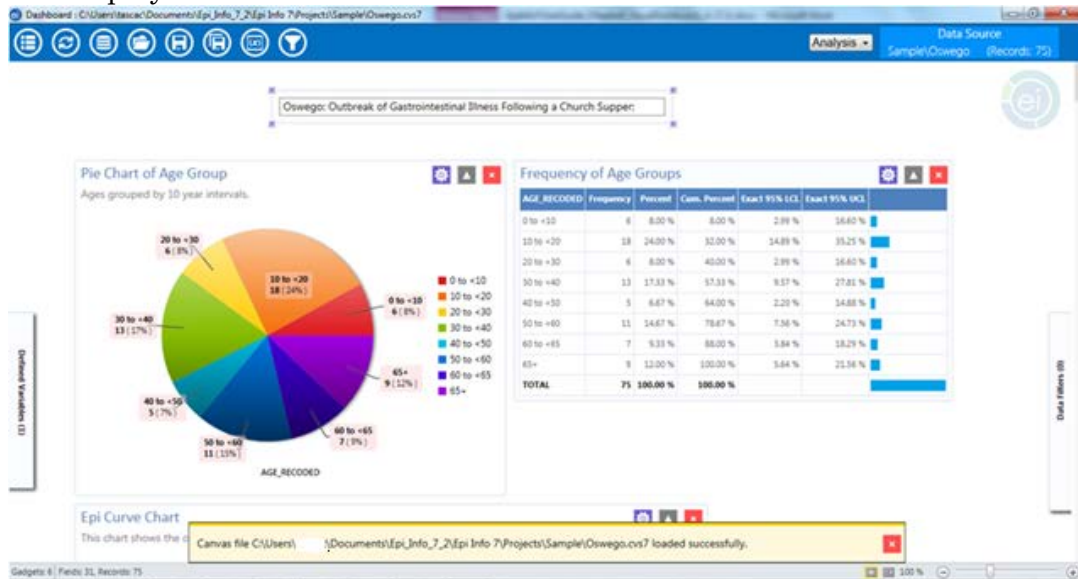


Figure 8.65: Opened canvas

The yellow message at the bottom of the screen confirms that the canvas loaded successfully. You can also click the open folder icon in the top left corner of the canvas for quick access to saved canvas files.

Save Canvas

The current canvas including data source, gadgets, filters, and user-defined variables can be saved and reused later. Saved canvases may be opened later with the Visual Dashboard tool. This will allow you to reconnect to the data source without having to manually apply the filters again.

To save a canvas, right click on the canvas and select **Save Canvas**. You can also click the Save icons at the top left corner of the canvas.

- If the canvas was previously saved and named, this will overwrite the existing file.
- If the canvas has never been saved, you will be prompted for a file name and have the opportunity to change the file location.



Figure 8.66: Visual Dashboard toolbar, with Save icon circled

Save Canvas As

The **Save Canvas As** option allows the canvas to be saved for the first time, or saved to a new location.

1. Right click on the canvas and select **Save Canvas As**, or click the Save As icon in the top left corner of the canvas. The **Save As** dialog box appears.
2. Select a location to save the canvas to and enter a File Name.
3. Click **Save**.



Figure 8.67: Visual Dashboard toolbar, with Save As icon circled

Note: If a file with the same name already exists, a confirmation message will appear, asking if you would like to replace the existing file.

Save Visual Dashboard HTML

Output from the gadgets may be saved as an HTML document. To save the current output:

1. Right click on the canvas, select **Save Output as HTML**. The **Save As** dialog box appears.
2. Select a location to save the canvas output and enter a File Name. The default location for saving output is to the Epi Info 7\Output folder.

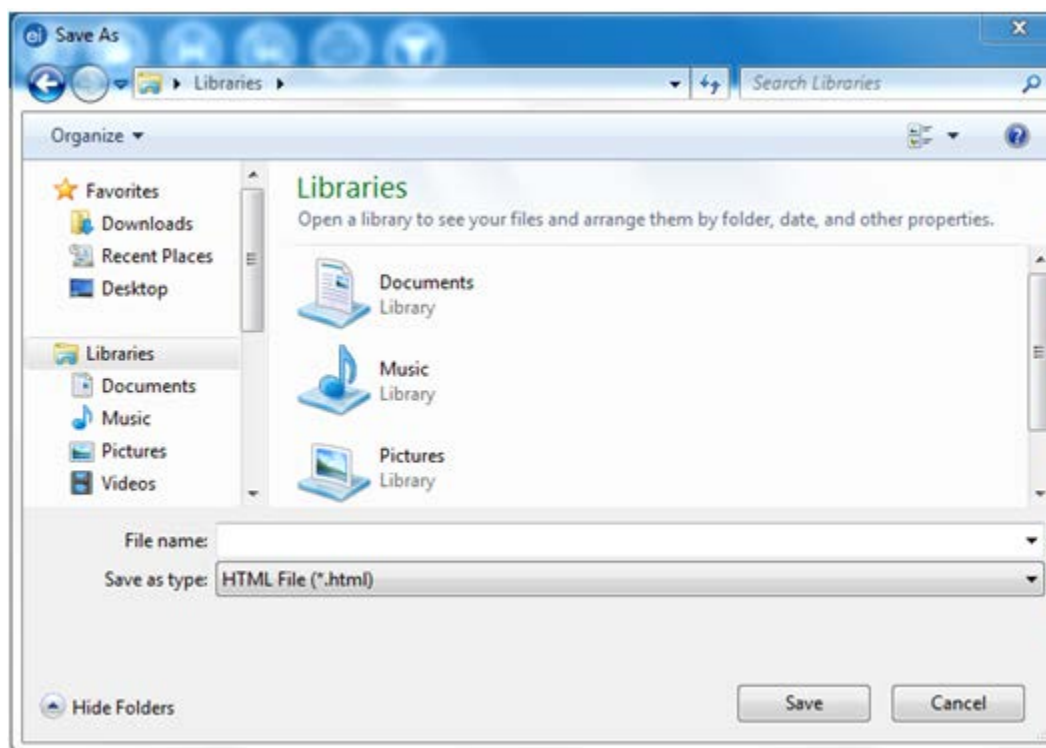


Figure 8.68: Save output as HTML

3. Click the **Save** button. The output automatically opens in the Web browser.

Note: *HTML documents can only be opened in a Web browser and cannot be used to re-create the Visual Dashboard canvas.*

Send Visual Dashboard Output To

Outputs from Visual Dashboard may be viewed in either Microsoft Word or Microsoft Excel with the **Send Output To** option. You must have Microsoft Word or Microsoft Excel installed to use either option.

Microsoft Excel

1. Right click on the canvas, select **Send Output To > Microsoft Excel**.
2. The **Visual Dashboard** output opens in **Microsoft Excel**.

Note: *Charts are not included with the output when sent to Excel.*

Microsoft Word

1. Right click on the canvas, select **Send Output To > Microsoft Word**.
2. The **Visual Dashboard** output opens in **Microsoft Word** as an .html formatted Web document. If saved in this format, charts are saved separately to a sub-folder. To share this .html file, remember to share the subfolder as well. Alternatively, change the format to a Microsoft Word .docx format so the charts are included with the file as embedded images.

Defining New Variables

The Defined Variables Gadget in Visual Dashboard allows you to create, edit, or delete variables with no coding. To access the Defined Variables gadget, move your mouse over Defined Variables on the left-hand side of the screen or by clicking the variable symbol (x) in the Visual Dashboard toolbar.



Figure 8.69: Visual Dashboard toolbar, with the Variable button circled

Creating a New Variable

Project data are not always in the proper format needed for analysis. As an example, a patient's age may be entered in a questionnaire but the ages may need to be in five-year increments for proper analysis. The data may contain county codes, but descriptive county names may be more useful for analysis. Creating a new variable is one of the functions of the Defined Variables gadget. By creating a new variable, the ages of patients can be recoded to five-year increments, or county names may be assigned to the county codes making the data useful for analysis. A new variable may be created to change the format of existing data with recoded value, simple assignment, formatted value, assigned expression, or by grouping existing variables.

Recoded Value

The New Variable with Recoded Value option assigns a new value to existing data and places the results in a defined variable. Click **New Variable** in the gadget and select **With Recoded Value**. The **Add Recoded Variable** dialog box appears.

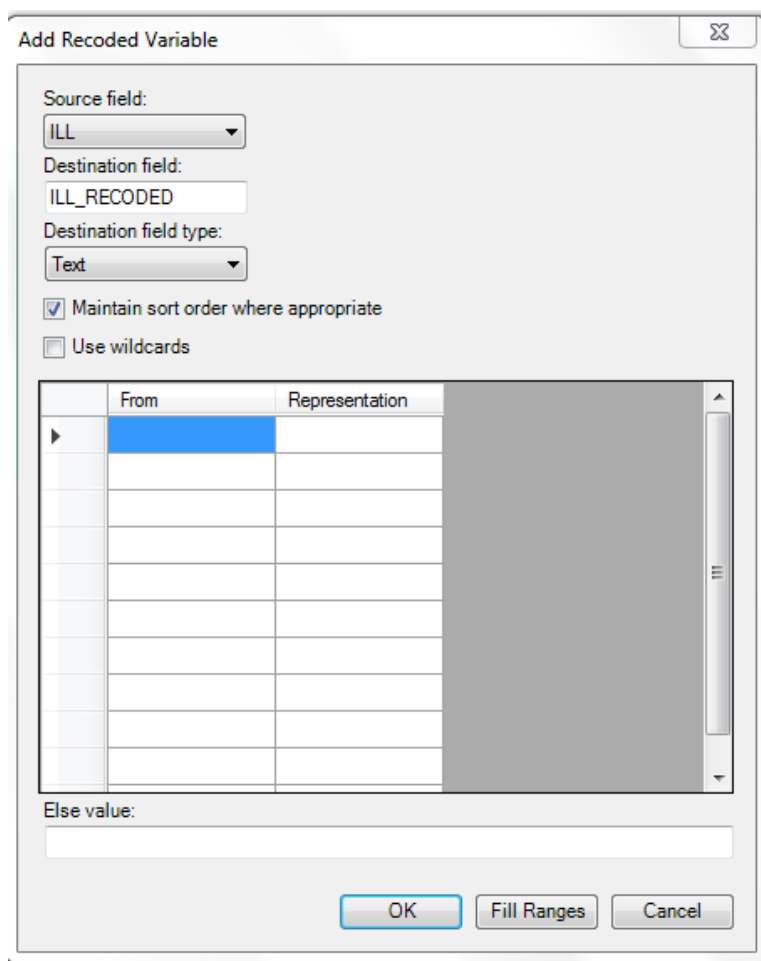


Figure 8.70: Add Recoded Variable dialog box

- The **Source field** drop-down list represents the column in the current data source that will have its values recoded.
- The **Destination field** drop-down list represents a new column that will be temporarily created in the current data source to store the recoded values.
- The **Destination field type** represents the data type of the new column that will be created to store the recoded values. The available options are text, numeric, or Yes/No.
- The **Maintain Sort Order Where Appropriate** box is checked by default. Usually output in Visual Dashboard is sorted alphabetically or sequentially. Leave this box checked if you would like the output to remain sorted exactly in the sequence that the data were entered.
- The **Use Wildcards** option can be used for text searches in the FROM column. This option is unchecked by default. The asterisk (*) is the wildcard character.

- The **From** column identifies the bottom of a range of values or a single value. To recode a single value instead of a range, use this From column only and leave the To column blank. When used to recode a range of values, the recode range includes the value in the From column and extends up to, but not including the value in the To column. To recode all unspecified values, leave both From and To columns blank.
- **To** value (if any) identifies the upper limit of a range of values to be recoded. The recode range does not include the value in the To column. This option only appears if a numeric variable is selected.
- The **Else Value** represents an optional value to assign to any records that do not match the recoding rules above. If left blank, the value of the new variable on records that don't match the recoding rules will be blank, null, or "Missing".
- **Representation Value** identifies the value to be assigned to the destination variable for the specified values of the source variable.
- **Fill Ranges** can be used to automatically populate range values on a numeric source field and is a very quick way to create numeric groupings such as age categories.
- **OK** accepts the current settings and data, and closes the form or window.
- **Cancel** exits the window without saving or executing a command.

The following example demonstrates how to add a recoded variable:

In the **E. coli Food History** table, there are more than 60 different values for Age. When performing an analysis, it is difficult to compare all of the different age values. By grouping the age values in 10-year increments, analysis on this variable is simplified. To group the values in Age groupings or categories:

1. Set the Data Source to the **EColi.PRJ** project.
2. From the Form section, click **FoodHistory**. Click **OK**.
3. On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget. The gadget expands and becomes fully visible.
4. Click the **New Variable** button.
5. Select **With Recoded Value**. The **Add Recoded Variable** dialog box appears.

Source field:
AGE

Destination field:
AGE_RECODED

Destination field type:
Text

Maintain sort order where appropriate

Use wildcards

	From	To	Representation
▶			

Else value:

OK Fill Ranges Cancel

Figure 8.71: Add Recoded Variable dialog box

6. From the **Source field** drop-down list, select **Age**.
7. Click the **Fill Ranges** button. The **Fill Ranges** dialog box appears.
8. Enter **0** for the **Start value**, **65** for the **End value**, and **10** for **By**.

Fill Ranges

Start value:

End value:

By:

OK Cancel

Figure 8.72: Add Recoded Variable/Fill Ranges

9. Click **OK**. The **Add Recoded Variable** dialog box appears with the **From**, **To**, and **Representation** columns populated.

The dialog box 'Add Recoded Variable' contains the following fields and options:

- Source field: AGE
- Destination field: AGE_RECODED
- Destination field type: Text
- Maintain sort order where appropriate
- Use wildcards

	From	To	Representation
▶	LOVALUE	0	LOVALUE - <0
	0	10	0 - <10
	10	20	10 - <20
	20	30	20 - <30
	30	40	30 - <40
	40	50	40 - <50
	50	60	50 - <60
	60	65	60 - <65
	65	HIVALUE	65 - < HIVALUE
*			

Buttons: OK, Fill Ranges, Cancel

Figure 8.73: Add Recoded Variable Ranges populated

10. Click **OK**. The variables are recoded and the conditions appear in the **Variable Definitions** textbox.

The 'Variable Definitions' dialog box shows:

- Variable Definitions: Recode the values in Age to Age_RECODED
- Defined Variables (1):
- Buttons: New Variable, Edit Variable, Delete Variable

Figure 8.74: Variable Definitions dialog box

To see the frequency of the new Age_RECoded values, right click on the canvas, and then select **Add Analysis Gadget > Frequency**. From the **Frequency of** list, select **Age_RECoded** and click **OK**.

Age_RECoded	Frequency	Percent	Cum. Percent	Wilson 95% LCL	Wilson 95% UCL	
0 - <10	1	0.28 %	0.28 %	0.05 %	1.57 %	
10 - <20	53	14.89 %	15.17 %	11.57 %	18.96 %	■
20 - <30	85	23.88 %	39.04 %	19.74 %	28.57 %	■
30 - <40	72	20.22 %	59.27 %	16.38 %	24.71 %	■
40 - <50	65	18.26 %	77.53 %	14.59 %	22.60 %	■
50 - <60	54	15.17 %	92.70 %	11.82 %	19.27 %	■
60 - <65	8	2.25 %	94.94 %	1.14 %	4.37 %	■
65 - < HIVALUE	18	5.06 %	100.00 %	3.22 %	7.85 %	■
TOTAL	356	100.00 %	100.00 %			■

Figure 8.75: Age_Recoded Frequency list

The Age_Recoded variable contains eight distinct values, which makes analysis easier than with the original data set.

Simple Assignment

Simple Assignment will assign the value of an existing variable to a newly defined variable. The following example demonstrates **Simple Assignment**.

1. Set the Data Source to the **Sample.prj** project.
2. From the Data Source Explorer menu, select **Surveillance**. Click **OK**.
3. On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget, or click the **Variables** icon in the top left. The gadget expands and becomes fully visible.
4. Click the **New Variable** button.
5. Select **With Simple Assignment**. The **Simple Assignment** dialog box appears.

Add Variable with Simple Assignment

Assign field:

Assignment type:
Difference in years

Start date:

End date:

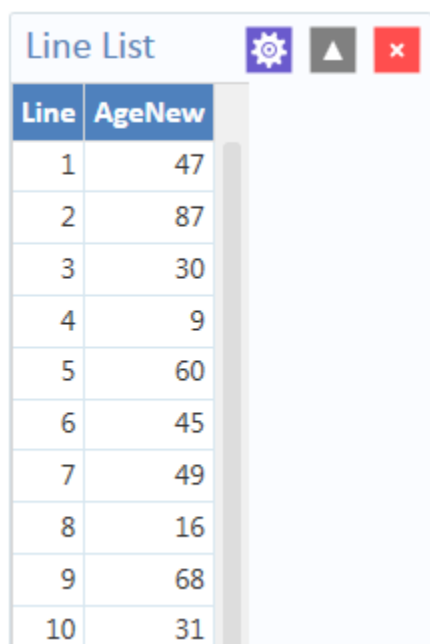
Description of assignment type:
Variable type: Numeric
Finds the difference in years between two dates. Only complete years are assigned. If the actual difference in years is 7.8, the assigned value will be a 7. The starting date should be the the earlier of the two dates, otherwise negative values will be assigned.

OK Cancel

Figure 8.76: Add Variable with Simple Assignment dialog box

6. In the Assign field, enter **Age_New**.
7. Select **Difference in years** from the **Assignment type** drop-down list.
8. Select **BirthDate** from the **Start Date** drop-down list.
9. Select **SYSTEMDATE** from the **End Date** drop-down list. The System Date is a system-level variable that represents the current date and time on your computer.
10. Click **OK**. A description of the assignment appears in the **Variable Definitions** textbox within the **Defined Variables** gadget.

To see the **Age_New** variable in line list view, right click on the canvas and then select **Add Analysis Gadget > Line List**. Select **Age_New** from **List of Variables to display** and then click **OK**.



Line	AgeNew
1	47
2	87
3	30
4	9
5	60
6	45
7	49
8	16
9	68
10	31

Figure 8.77: Simple Assignment Line List

In this example, the current age of the patient is calculated from the difference between the birth date in the data source and the date the variable is created (current date in your computer). The **Analysis Gadget > Line List** view displays the new variable created via **Simple Assignment**.

Conditional Assignment

The **Conditional Assignment** option defines conditions and one or more events that occur if the specified conditions are met. If the condition is true, the first event is executed. An alternative event can be given after the **ELSE** statement. The **ELSE** statement will be executed if the first set of conditions is not true. If the statement is false and no **ELSE** statement is selected, the statement is bypassed. To add a new variable with a Conditional Assignment, move the cursor over the **Defined Variables** gadget on the left-hand side of the Visual Dashboard canvas or click the Variables icon in the top left. The gadget expands and becomes fully visible. Click the **New Variable** button and select **With Conditional Assignment**.

Figure 8.78: Add New Variable with Conditional Assignment

- **Assign Field** indicates the new variable name.
- **Assign Field Type** indicates the format of the new variable. The available options are Text, Numeric, or Yes/No.
- **Assign Condition** determines whether subsequent commands will be run. If the condition specified in this textbox is true, the Assign Value will be applied.
- **Assign Value** states the value of the new variable when the Assign Condition is true.
- **Use Else** checkbox indicates whether an Else Value is specified. This checkbox is unmarked by default.
- The **Else Value** is applied when the statement in the Assign Condition textbox is false.

In the following example, the **YoungAdult** variable will be conditionally assigned the value of **Yes** if the Age variable has a value between 17 and 30 inclusive. If the Age value falls outside of this range, a value of **No** will be assigned to YoungAdult.

1. Set the Data Source to the **Sample.prj** project.
2. From the **Data Source Explorer** menu, select **Surveillance**. Click **OK**.
3. On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget. The gadget expands and becomes fully visible.
4. Click the **New Variable** button.
5. Select **With Conditional Assignment**. The **Add Variable with Conditional Assignment** dialog box appears.

Add Variable with Conditional Assignment

Assign field:

Assign field type:

Assign condition:

Assign value:

Use else

Else value:

Figure 8.79: Conditional Assignment / Create Edit Condition

6. In the Assign field, enter **YoungAdult**.
7. Select **Yes/No** from the **Assign field type** drop-down list.
8. Click on the **Create/Edit Condition** button. The **Specify Assign Condition** dialog box appears.

Specify assign condition

The value of Field Name:

Operator:

Value:

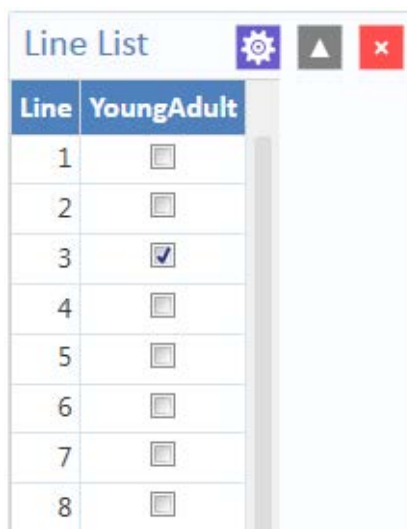
Data filters:

Join	Filter criteria

Figure 8.80: Specify Conditional Assignment Condition dialog box

9. Select **Age** from the **Field Name** drop-down list.
10. Select **is between** from the **Operator** drop-down list.
11. Enter **17** and **30** for Values. Click **Add Condition**.
12. Click **OK**. The filter criteria appear in the **Data Filters** list.
13. Click **OK** in the **Specify Assign Condition** dialog box.
14. The condition now appears in the **Assign Condition** textbox within the **Add Variable with Conditional Assignment** dialog box.
15. Select **Yes** from the **Assign Value** drop-down list.
16. Check the **Use Else** checkbox.
17. Select **No** from the **Else Value** drop-down list.
18. Click **OK**. A description of the assignment appears in the **Variable Definitions** textbox within the **Defined Variables** gadget.

To see the **YoungAdult** variable in line list view, right click on the canvas and then select **Add Analysis Gadget > Line List**. Select **YoungAdult** from **List of Variables to display** and then click **OK**. Check marks illustrate the lines for which the conditional assignment is **true**.



Line	YoungAdult
1	<input type="checkbox"/>
2	<input type="checkbox"/>
3	<input checked="" type="checkbox"/>
4	<input type="checkbox"/>
5	<input type="checkbox"/>
6	<input type="checkbox"/>
7	<input type="checkbox"/>
8	<input type="checkbox"/>

Figure 8.81: YoungAdult Conditional Assignment Line List

Note: The Else clause at the end of the IF statement is important to specify what should happen if the conditions are not met. Without the Else clause, YoungAdult would be null or missing for any values falling outside the 17 to 30 range.

Formatted Value

The format and name of an existing variable may be changed with the Formatted Value option. This is currently only useful with date fields as demonstrated by the example below.

1. Set the Data Source to the **Sample.prj** project.
2. From the **Data Source Explorer** menu, select **Surveillance**. Click **OK**.

3. On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget or click the **Variables** icon in the top left of the screen. The gadget expands and becomes fully visible.
4. Click the **New Variable** button.
5. Select **With Formatted Value**. The **Add Formatted Value** dialog box appears.

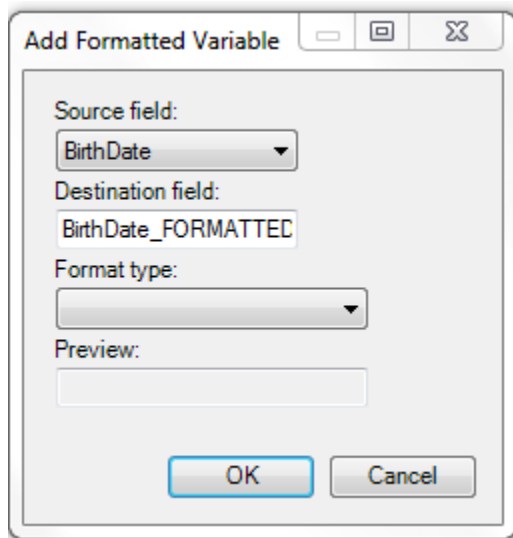
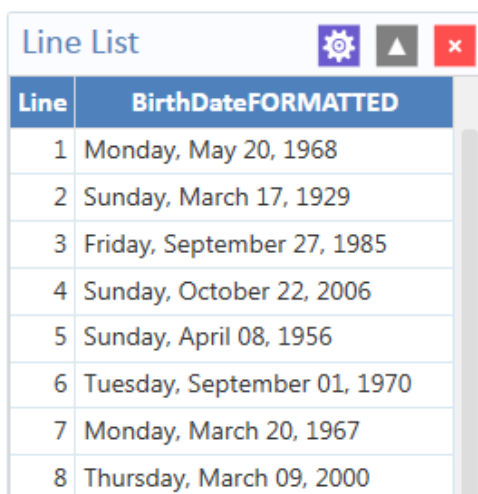


Figure 8.82: Add Formatted Variable dialog box

6. Select **BirthDate** from the **Source field** drop-down list.
7. The Destination field automatically shows the source field with a suffix of “FORMATTED” (i.e. “BirthDate_FORMATTED”). Modify the name of the Destination field as desired.
8. Select **the long date** from the **Format type** drop-down list. Using the current system date, a sample long date value appears in the Preview field.
9. Click **OK**. A description of the assignment appears in the **Variable Definitions** textbox within the **Defined Variables** gadget.

To see the **BirthDate_FORMATTED** variable in line list view, right click on the canvas and then select **Add Analysis Gadget > Line List**. Select **BirthDate_FORMATTED** from **List of Variables to display** and then click **OK**.



Line	BirthDateFORMATTED
1	Monday, May 20, 1968
2	Sunday, March 17, 1929
3	Friday, September 27, 1985
4	Sunday, October 22, 2006
5	Sunday, April 08, 1956
6	Tuesday, September 01, 1970
7	Monday, March 20, 1967
8	Thursday, March 09, 2000

Figure 8.83: Formatted Variable line list

The **Analysis Gadget > Line List** view displays the new variable created via Formatted Value.

Assigned Expression

Assigned Expression assigns the result of an expression or the field value to a variable as demonstrated in the example below. In the following example, the zip code field is a number. To use the zip code in a map, a new zip code variable must be defined and assigned text values.

1. Set the Data Source to the **Sample.prj** project.
2. From the **Data Source Explorer** menu, select **Surveillance**. Click **OK**.
3. On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget or click the **Variables** icon in the top left of the screen. The gadget expands and becomes fully visible.
4. Click the **New Variable** button.
5. Select **With Assigned Expression**. The **Add Variable with Expression** dialog box appears.

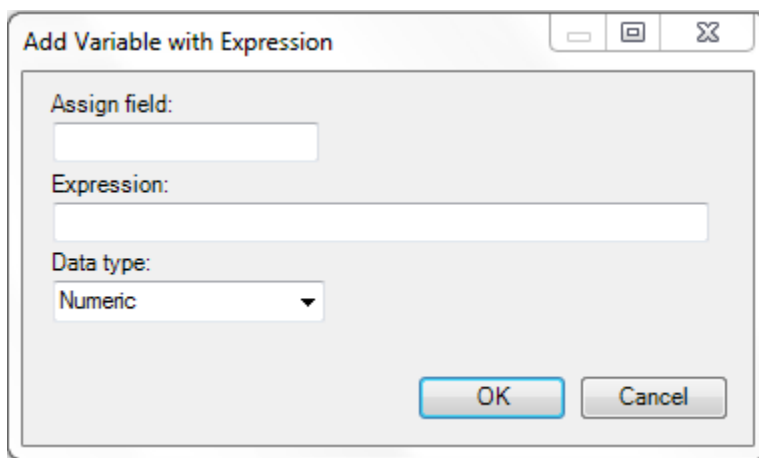


Figure 8.84: Add Variable with Expression

6. In the **Assign Field**, enter **Zip2**.
7. In the **Expression** field, type the syntax: **ZipCode + '-0000'**.
8. Select **Text** from the **Data type** drop-down list.
9. Click **OK**. The assignment appears in the Variable Definition textbox within the **Defined Variables** gadget.

To see the **Zip2** variable in line list view, right click on the canvas and then select **Add Analysis Gadget > Line List**. Select **Zip2** from **List of Variables to display** and then click **Generate Line List**.

Line	Zip2
1	31061-0000
2	31059-0000
3	31287-0000
4	30730-0000
5	30060-0000
6	30756-0000
7	30730-0000
8	31061-0000

Figure 8.85: Add Variable with Expression Line List

Creating a Variable Group

The **Create Variable Group** option in Visual Dashboard allows for the temporary grouping of variables. This option is useful when there are many variables and analyzing each is impractical. Group variables can be used in the Combined Frequency and MxN/2x2 gadgets. The following example demonstrates how to create a variable group.

1. Set the Data Source to the **Sample.prj** project.
2. From the **Data Source Explorer** menu, select **Oswego**. Click **OK**.
3. On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget or click the **Variable** icon in the top left of the screen. The gadget expands and becomes fully visible.
4. Click the **New Variable** button.
5. Select **Create Variable Group**. The **Create Group** dialog box appears.

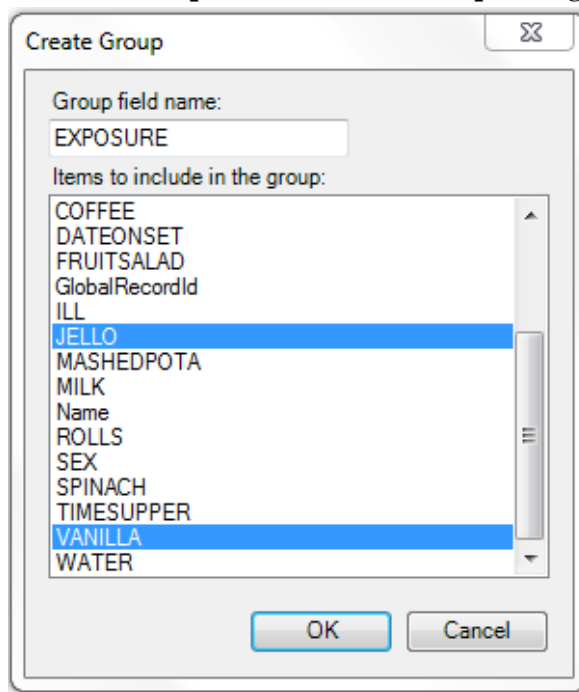


Figure 8.86: Create Group variable options

6. In the Group field name, enter **Exposure**.
7. Select **JELLO** and **VANILLA** from the **Items to include in the group** list. Hold down the **Ctrl** key on your keyboard to select more than one variable at a time.
8. Click **OK**.
9. The variable group statement now appears in the **Variable Definitions** textbox within the Defined Variables gadget.

Edit Variable

The **Edit Variable** option allows an existing variable definition to be changed. The following example will add a new variable to an existing Exposure group variable such as the one created in the section **Creating a Variable Group**.

1. On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget, or click the **Variables** icon in the top left of the screen. The gadget expands and becomes fully visible.
2. Click the **Edit Variable** button.

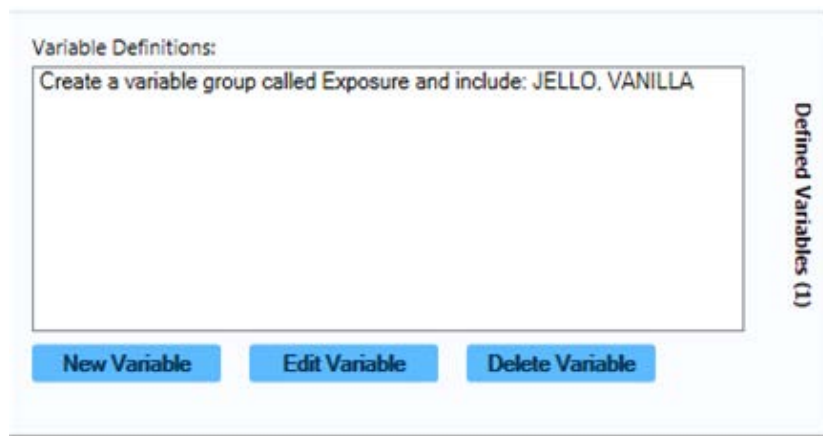


Figure 8.87: Variable group defined

3. The **Create Group** dialog box appears, with JELLO and VANILLA highlighted.
4. Press and hold the **Ctrl** key on your keyboard and select **Spinach** to add it to the variable.
5. Click **OK**.
6. The Exposure variable group now includes Spinach, which appears in the Variable Definitions textbox within the **Defined Variables** gadget.

Delete Variable

The **Delete Variable** option in the Defined Variables gadget removes a defined variable from the system. In this example, the Exposure group variable will be deleted.

1. On the left-hand side of the Visual Dashboard canvas, move the mouse cursor over the **Defined Variables** gadget, or click the **Variables** icon at the top left of the screen. The gadget expands and becomes fully visible.
2. Select the **Exposure** variable definition within the textbox.

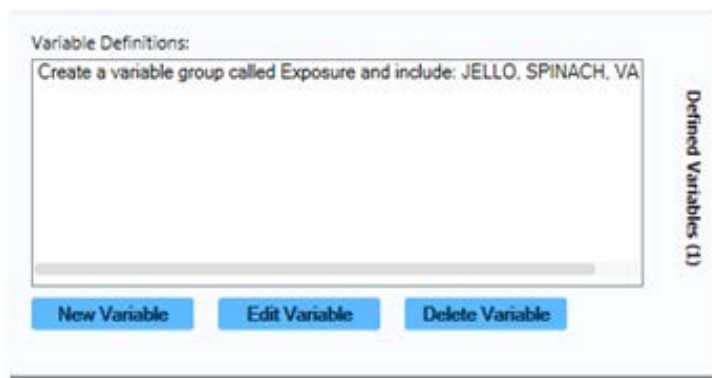


Figure 8.88: Variable Definitions box, illustrating Delete Variable option

3. Click on the **Delete Variable** button.
4. The variable is deleted from the **Variable Definitions** text box.

Adding and Removing Data Filters

Add a Filter

Visual Dashboard typically opens all records in a file. The **Data Filters** gadget in Visual Dashboard allows you to select a subset of data by specifying certain conditions. This is useful when trying to analyze a specific group of records. There are two modes in this gadget; guided and advanced. The **Guided Mode** allows you to apply conditions by selecting values from Field Name, Operator, and Value drop-down lists. Filters may be applied in the **Advanced Mode** by entering the filtering conditions in the free form text box. To add a filter in **Guided Mode**, follow the steps below. **Advanced Mode** is discussed in a later section.

1. Select the **EColi.PRJ** Data Source.
2. From the **Data Source Explorer** menu, click **FoodHistory**.
3. Click **OK**.
4. From the right-hand side of the Visual Dashboard canvas, move the mouse over the **Data Filter** gadget, or click the **Filters** gadget (funnel) icon in the top left of the screen. The gadget expands and becomes fully visible.
5. From the **Field Name** drop-down list, select **ILL**.
6. From the **Operator** drop-down list, select **is equal to**.
7. From the **Value** drop-down list, select **Yes**.

The value of Field Name: Operator: Value:

Add Filter

Data filters:

Join	Filter criteria

Figure 8.89: Data Filter Gadget/Add Filter

- Click the **Add Filter** button. The filter condition is added to the Data filters grid within the gadget. Screen shot of the data filter gadget, which can be used to filter record results

The value of Field Name: Operator: Value:

Add Filter

Data filters:

Join	Filter criteria
	The value of [ILL] is equal to Yes

Figure 8.90: Data Filter Gadget/Add Filter with Condition

If a filter condition is added or removed, all gadgets on the canvas will automatically refresh. The record count at the top now reads **276**, indicating that there are 276 out of 359 records in which the value of ILL is Yes.

To add another filter:

- On the right-hand side of the Visual Dashboard canvas, move the mouse over the **Data Filter** gadget, or click the **Filters** gadget in the top left corner. The gadget expands and becomes fully visible.
- From the **Field Name** drop-down list, select **Age**.
- From the **Operator** drop-down list, select **is between**.

- From the **Value** text boxes, type **20** for the first box.
- From the **Value** text boxes, type **29** for the second box.
- Click **Add Filter**. A context menu appears asking if you want to add this condition using an AND or OR operator. Select your desired operator.

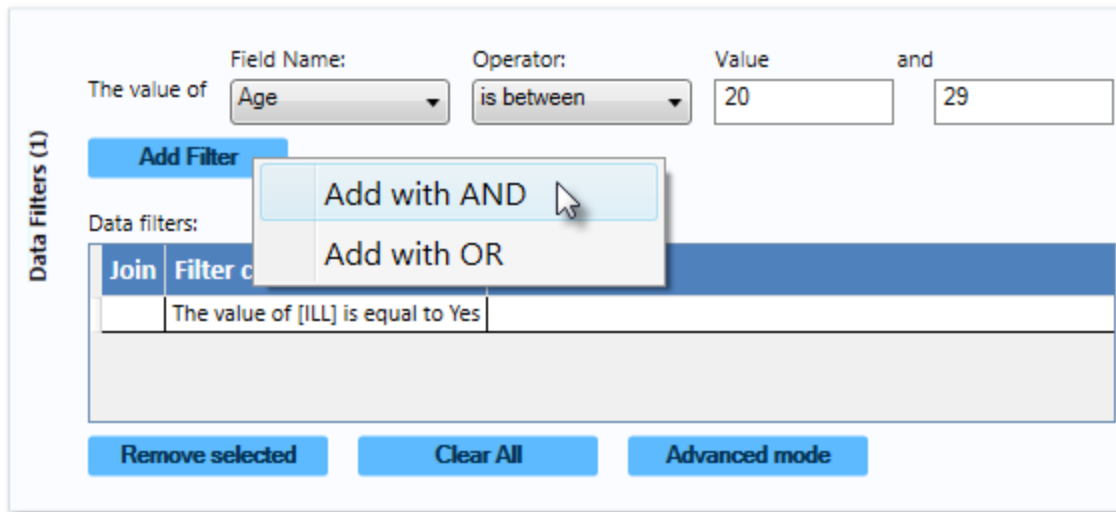


Figure 8.91: Data Filter Gadget illustrating addition of a second condition

- The filter condition is added to the Data filters grid view beneath the first condition.

The record count shows 64 records. This indicates that only 64 records match the two established filter criterion that the patient was ill **and** between 20 and 29 years of age. All gadgets added to the Visual Dashboard canvas will only display data from these 64 records. The filter settings appear at the top of the canvas.


 The following canvas filters are in effect: The value of AGE is between 20 and 29 and the value of [ILL] is equal to Yes

Figure 8.92: Data Filter setting

Remove Selected Filter

If you desire to work with only one or some of the existing filters, the unwanted data filters must be removed. Follow the steps below to remove individual filters.

- On the right-hand side of the Visual Dashboard canvas, move the mouse over the **Data Filter** gadget or click the **Filters** gadget in the top left corner. The gadget expands and becomes fully visible.
- To remove the **second condition**, which should appear as “**The value of Age is between 20 and 29,**” select it in the list.
- Click the **Remove Selected** button. The condition disappears.
- Select the remaining **condition**.

5. Click the **Remove Selected** button. The condition disappears. If you had other filters in the list, they would remain. If all filters are removed, the system would return to working with the full dataset.

Clear All Filters

The **Clear All** button removes all existing filters at once. If you wanted to clear all filters, you could use this step rather than removing filters individually as described in the topic Remove Selected Filter.

1. On the right-hand side of the Visual Dashboard canvas, move the mouse over the **Data Filter** gadget or click the **Filters** gadget in the top left corner. The gadget expands and becomes fully visible.
2. Click the **Clear All** button. All existing filter conditions disappear.

The screenshot shows the 'Data Filters (2)' gadget interface. At the top, there is a form for adding a filter: 'The value of' followed by a dropdown menu with 'Age' selected, an 'Operator:' dropdown with 'is between' selected, and two 'Value' input boxes containing '20' and '29'. Below this is an 'Add Filter' button. Underneath is a table labeled 'Data filters:' with two columns: 'Join' and 'Filter criteria'. The table contains two rows: the first row has 'and' in the 'Join' column and 'The value of [ILL] is equal to Yes' in the 'Filter criteria' column; the second row has 'and' in the 'Join' column and 'The value of Age is between 20 and 29' in the 'Filter criteria' column. At the bottom of the gadget are three buttons: 'Remove selected', 'Clear All', and 'Advanced mode'.

Figure 8.93: Data Filter Gadget/Add Filter with Two Conditions

Advanced Filter Mode

The **Advanced mode** allows you to type the desired data filter string into the text box provided. Once the filter is applied, the canvas will only display the records that meet the specified condition(s).

Note: *Advanced mode uses the .NET DataView.RowFilter syntax. For more information about this syntax, refer to Microsoft's website describing the DataView.RowFilter Property.*

The following example demonstrates how to use advanced filters.

1. On the right-hand side of the Visual Dashboard canvas, move the mouse over the **Data Filter** gadget or click the **Filters** gadget in the top left corner. The gadget expands and becomes fully visible
2. Select **Advanced Mode**. The Advanced mode dialog box appears.

Advanced mode
Type the desired data filter string into the text box below. Several examples are provided:

Data Filters (0)

1. Numeric data: (AGE >= 15) AND (AGE <= 45)
2. Text data: (LastName LIKE '%sen') OR (LastName = 'Smith')
3. Date data: (DOB >= #01/01/2000#) AND (DOB <= #12/31/2000 23:59:59#)
4. Boolean data: (ILL = true)

No advanced filters are in effect.

Apply Filter **Guided Mode**

Figure 8.94: Advanced Mode

3. In the text box, enter the following condition “**(([ILL] = 1) and (Age >= 20 and Age <= 29)) AND RECSTATUS > 0**” (do not include the quotation marks).
4. Click the **Apply Filter** button.
5. The verbiage “**No advanced filters are in effect**” will be replaced with “**The filter is valid and is now in effect.**” When filters are applied, the conditions appear in green font.
6. The record count in the top right corner of the Visual Dashboard canvas now reads **64**, indicating that there are 64 out of 359 records where the value of ILL is Yes and the value for age is between 20 and 29.

Note: You can toggle between the Guided Mode and the Advanced Mode by clicking on the respective button on each screen.

Displaying Statistics and Records

Gadget Filters and Settings

In addition to setting canvas wide filters, you can also set a gadget-specific filter using the widgets at the top of the gadget. The widgets and borders disappear when the “Hide Borders” button is pressed. If the widgets are not visible, click the “Hide Borders” button at the bottom right corner of the canvas as described in **Navigate the Visual Dashboard Workspace**.



Figure 8.95: Gadget widgets

1. The purple **gear** button collapses and expands the gadget's configuration window.
2. Gadget filters are accessed by clicking the purple gear icon and opening the **Data Filters** property panel. Gadget filters are combined with canvas filters, e.g., if the canvas filter is 'Age > 20' and the gadget filter is 'ILL = Yes', then the filter applied to the gadget is 'Age >20 AND ILL = Yes'. Some gadgets do not have filtering capabilities, such as the advanced statistics.
3. Gadget **Display** options, such as title and description, are accessed by clicking the purple gear icon and selecting the **Display** property panel.
4. The gray up/down arrow on the gadget's header panel collapses and expands the gadget's output, which is useful when you want to minimize the amount of gadgets that are displayed on the canvas.
5. You may close the gadget by clicking on the red "X" at the top right of the header panel. Doing so results in a dialog box asking you to confirm that you want to remove the gadget from your canvas.

StatCalc Calculator

Refer to the StatCalc Calculator section of this user guide.

NutStat Growth Chart

Refer to the NutStat section of this user guide.

Data Dictionary

Show Data Dictionary

The Data Dictionary displays form(s) and defined variables for the current Data Source. Fields or variables are sorted and displayed by page number in the form. Defined variables appear at the end of the listing. Information retrieved from the form includes Field Name, Prompt, Form, Page, Tab, Data Type, Epi Info Field Type, and Table.

1. Right click on the canvas and select **Show data dictionary**.
2. The **Data Dictionary** is displayed on the canvas.

Column Name	Prompt	Form Name	Page	Tab
UniqueKey		Epi10		
RecStatus		Epi10		
FKEY		Epi10		
GlobalRecordId		Epi10		
VAC	VAC	Epi10	1	5
PRENATAL	PRENATAL	Epi10	1	4
CLUSTER	CLUSTER	Epi10	1	3
POPW	POPW	Epi10	1	2
LOCATION	LOCATION	Epi10	1	1
OldUniqueKey				
SYSTEMDATE				

Figure 8.96: Data Dictionary view

The scroll bar at the bottom of the **Data Dictionary** can be used to view the rest of the window. Column values for **Prompt**, as well as **Data Type** and **Epi Field Type** (both of which are viewable when the user scrolls right) are developed when fields are created from the **Field Definition** dialog box discussed in Chapter 2: Form Designer.

1. To open the **Data Dictionary** as an **HTML** page inside a web browser, hover over the **Data Dictionary** in the dictionary window, right click, and select **Send Data Dictionary to Web Browser**.

Column Name	Prompt	Form Name	Page	Tab
UniqueKey		Epi10		
RecStatus		Epi10		
FKEY		Epi10		
GlobalRecordId		Epi10		
VAC	VAC	Epi10	1	5
PRENATAL	PRENATAL	Epi10	1	4
CLUSTER	CLUSTER	Epi10	1	3
POPW	POPW	Epi10	1	2
LOCATION	LOCATION	Epi10	1	1
OldUniqueKey				
SYSTEMDATE				

Figure 8.97: Option to send Data Dictionary to Web

From the browser, the data can be printed using the print tool within the respective web browser, or saved with the file tool.

2. You can export the data directly to an Excel spreadsheet by right clicking over the table in the web page to show the pop-up menu and selecting **Export to Microsoft Excel**.

Data Dictionary

columnname	Prompt	formname	Page	Tab	datatype	FieldType	tablename	Items	epifieldtype
UniqueKey		Epi10			Int32		Epi10		
RecStatus		Epi10			Int32		Epi10		
FKEY		Epi10			String		Epi10		
GlobalRecordId		Epi10			String		Epi105		Epi.Fields.GlobalRecordIdField
VAC	VAC	Epi10	1	5	Double	NumberField	Epi105		Epi.Fields.NumberField
PRENATAL	PRENATAL	Epi10	1	4	Double	NumberField	Epi105		Epi.Fields.NumberField
CLUSTER	CLUSTER	Epi10	1	3	Double	NumberField	Epi105		Epi.Fields.NumberField
POPW	POPW	Epi10	1	2	Double	NumberField	Epi105		Epi.Fields.NumberField
LOCATION	LOCATION	Epi10	1	1	Double	NumberField	Epi105		Epi.Fields.NumberField
OldUniqueKey					Int32				
SYSTEMDATE					DateTime				

Figure 8.98: Data Dictionary Export to HTML

3. Close the browser window to exit the **Data Dictionary**.

Note: The Data Dictionary can be refreshed at any time to display the most current data. Right click on the Data Dictionary and select Refresh Data Dictionary.

Report Gadgets

There are several report gadgets available to add text and images to your reports. The **Simple Text box** option allows you to specify several lines of text to appear on the canvas. You can use these to annotate the canvas above or beside the various gadgets. This is particularly useful for adding a title to the canvas. To add a title to your canvas:

1. Right click on the **Visual Dashboard canvas**. Select **Add Report gadget** from the menu options, and then select **Simple Text box**.
2. The text box appears.
3. Enter a title or other text. You can move the text box around by clicking on it and dragging your mouse to the desired location.

To remove your title:

1. Right click on the **textbox** border.
2. Click **Close this gadget**.

Image Gadget is another report gadget. It allows you to add an image or logo to your canvas.

1. Right click on the **Visual Dashboard canvas**. Select **Add Report gadget** from the menu options, and then select **Image gadget**.
2. A light blue box appears on the canvas near where you right clicked. Click on the **box**.

The **Open** dialog box appears. Select the location of your image or enter the **File name** of your image. **Note: You may only select a *.png file for the Image gadget.**

3. Double-click your image or select it and click **Open**.
4. The image appears on the canvas.

To remove your image:

1. Right click on your image.
2. Click **Close this gadget**.

Read a Microsoft Excel File

1. Click on the **Set a Data Source now** arrow. The **Select Data Source** dialog box appears.
2. From the **Database type** drop-down list, select either **MS Excel 97-2003 Workbook** or **MS Excel 2007 Workbook**.
3. Click **Browse** next to the **Data Source** field. The **Open Existing File** dialog box will appear.
4. Enter the filename and path to open your workbook or click **Browse** to search for the **MS Excel workbook** (.xls or .xlsx).
5. Click **Open**. The *First row contains header information* checkbox is selected by default. If this option is unchecked, data will import without column headings.
6. Click **OK**. Select a **worksheet** from the list provided in the **Data Source Explorer** window.
7. Click **OK**. The record count and data source appear in the menu bar at the top right of the Visual Dashboard canvas.

Read a Microsoft Access File

1. Click on the **Set a Data Source now** arrow. The **Select Data Source** dialog box appears.
2. From the **Database Type** drop-down list, select either **MS Access 2002-2003 (.mdb)** or **MS Access 2007(.accdb)**.
3. Click **Browse** next to the **Data Source** field. The **Open Microsoft Access File** dialog box will appear.
4. Enter the filename and path to open your file or click **Browse** to search for the **MS Access file (.mdb or .accdb)**. If the file is password protected, enter the password.
5. Click **OK**.
6. Select a **data table** from the list provided in the **Data Source Explorer** window.

7. Click **OK**. The record count and data source appear in the menu bar at the top right of the Visual Dashboard canvas.

Read a CSV (ASCII) File

1. Click on the **Set a Data Source now** arrow. The **Select Data Source** dialog box appears.
2. From the **Database Type** drop-down list, select **CSV File**.
3. Click **Browse** next to the **Data Source** field.
4. Click **Browse** in the **Flat File location** dialog box to search the directory.
5. Select the directory containing the CSV file to be read.
6. Click **OK**. Note that the *First row contains header information* box is checked by default. If the first row of the file does not contain header information, deselect the checkbox. If this option is unchecked, data will be imported without column headings.
7. Select the **file name** from the list provided in the **Data Source Explorer** window.
8. Click **OK**. The record count and data source appear in the menu bar at the top right of the Visual Dashboard canvas.

Read a Microsoft SQL Server Database

Note: You must have Microsoft SQL Server installed on your computer or you must have access to a server over a network to perform this function.

1. Click on the **Set a Data Source now** arrow. The **Select Data Source** dialog box appears.
2. From the **Database Type** drop-down list, select **Microsoft SQL Server Database**.
3. Click **Browse** next to the **Data Source** field.
4. The **Connect to SQL Server** database dialog box opens. Specify the **Server name** and **Database name** to connect to a SQL Server database.
5. The **Use Windows Authentication** radio button is selected by default. You can use **SQL Server Authentication** by checking the corresponding radio button and then entering Login and Password credentials.
6. Click **OK**.
7. From the list provided in the **Data Source Explorer**, select a **data table**.
8. Click **OK**. The record count and data source appear in the menu bar at the top right of the Visual Dashboard canvas.

Read a MySQL Server Database

Note: You must have MySQL Server installed on your computer or you must have access to a server over a network to perform this function.

1. Click on the **Set Data Source** arrow. The **Select Data Source** dialog box appears.

2. From the **Database Type** drop-down list, select **MySQL Database**.
3. Click **Browse** next to the **Data Source** field.
4. The **Connect to MySQL** database dialog box appears. Enter the **Server name**, **Database name**, **Username** and **Password** to connect to the MySQL database. You can select **<Browse for more>** from the **Server name** drop-down list to select a MySQL server instance in your network.
5. Click **OK**.
6. From the list provided in the Data Source Explorer, select a **data table**.
7. Click **OK**. The record count and data source appear in the menu bar at the top right of the Visual Dashboard canvas.

Add a Related Data Source

Analysis in Visual Dashboard may be performed on data from multiple data sources. The **Add a Related Data Source** option links tables or forms using a common identifier. The form/table to be linked requires a key field to relate the records of the forms/tables together. If the forms were created in Form Designer, Visual Dashboard can establish a relationship by using the Global Record ID (GLOBALRECORDID) of the first form and the Foreign Key (or [FKEY] () column of the related form as the key fields. The example below demonstrates how to add a related data source.

Epi Info™ 7 contains a related database in the Sample.PRJ file with two forms: The form named Surveillance is the primary or parent form. It has three related or child forms named RHepatitis, R Lyme, and RMeningitis. In the example below, we relate the Surveillance parent form to its child form, RHepatitis.

1. Follow the steps to **Set a Data Source** and select the **Sample.PRJ** Data Source. Open **Surveillance**.
2. Right click on the canvas and select **Add related data source**. The **Select Data Source** dialog box opens.
3. Follow the steps to **Set a Data Source** and select the **Sample.PRJ** Data Source again.
4. Select **RHepatitis** from the list of forms in the Data Source Explorer. The **Child Key Field** drop-down list populates.
5. From the **Parent Key Field** drop-down list, select **GlobalRecordId**.
9. Select **FKEY** from the **Child Key Field** drop-down list. **FKEY** is the Foreign Key in the child form that matches the GlobalRecordId of the record in the parent form.
10. Check the “Use Unmatched (ALL)” checkbox to include all records from the primary form whether or not there are matching records in the related data source. Uncheck the “Use Unmatched (ALL)” checkbox to force the resulting data to include only those records represented by keys that match in both tables.
11. Click **OK**. The “Processing Finished” message appears, indicating that the system has cached the data locally. Data from the two tables can now be used for analysis.

Note: You can only perform a one-to-one analysis with a related table in Visual Dashboard. To perform a one-to-many analysis, please see the Classic Analysis section of this user guide.

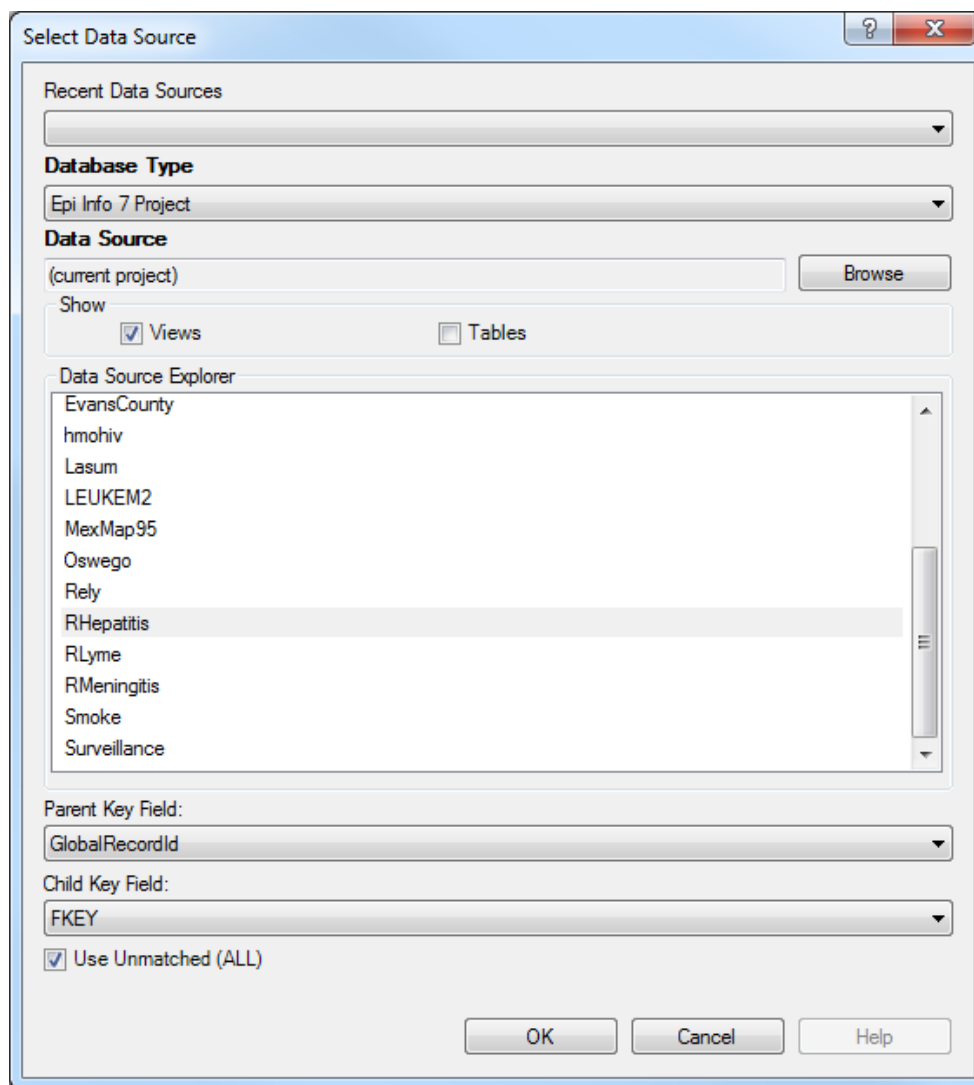


Figure 8.99: Add Related Data Source dialog box

Run Visual Dashboard from the Command Line

The Visual Dashboard can be run from a batch program, custom menu, or the operating system command line. If the VisualDashboard.exe program is associated with files having the .cvs7 extension, then opening Visual Dashboard to show a saved canvas can be as simple as executing the canvas file name.

If the .csv7 file extension is not associated with the Visual Dashboard program, then you will need to call the AnalysisDashboard.exe program followed by the name of the canvas using the /canvas: switch.

These examples are given with the assumption that Epi Info™ 7 is in the C:\Epi_Info_7 folder on the C:\ drive. If your copy of Epi Info™ 7 is in another location, you may need to use your location or referential pathing.

If run from within the C:\Epi_Info_7\ folder, a batch program can open Visual Dashboard with simply the statement:
AnalysisDashboard.exe .

To open a specific canvas, use the /canvas: switch. In this example, we will open Visual Dashboard to show the saved Ecoli.csv7 canvas which may be found in the C:\Epi_Info_7\Projects\Ecoli folder.

```
C:\Epi_Info_7\AnalysisDashboard.exe  
/canvas:"C:\Epi_Info_7\Projects\Ecoli\Ecoli.csv7"
```

In the batch program, the statement above should appear all on one line.