epiinfo™
Guide Series

Epi Info™ Guide to Visual Dashboard

This page has been intentionally left blank.
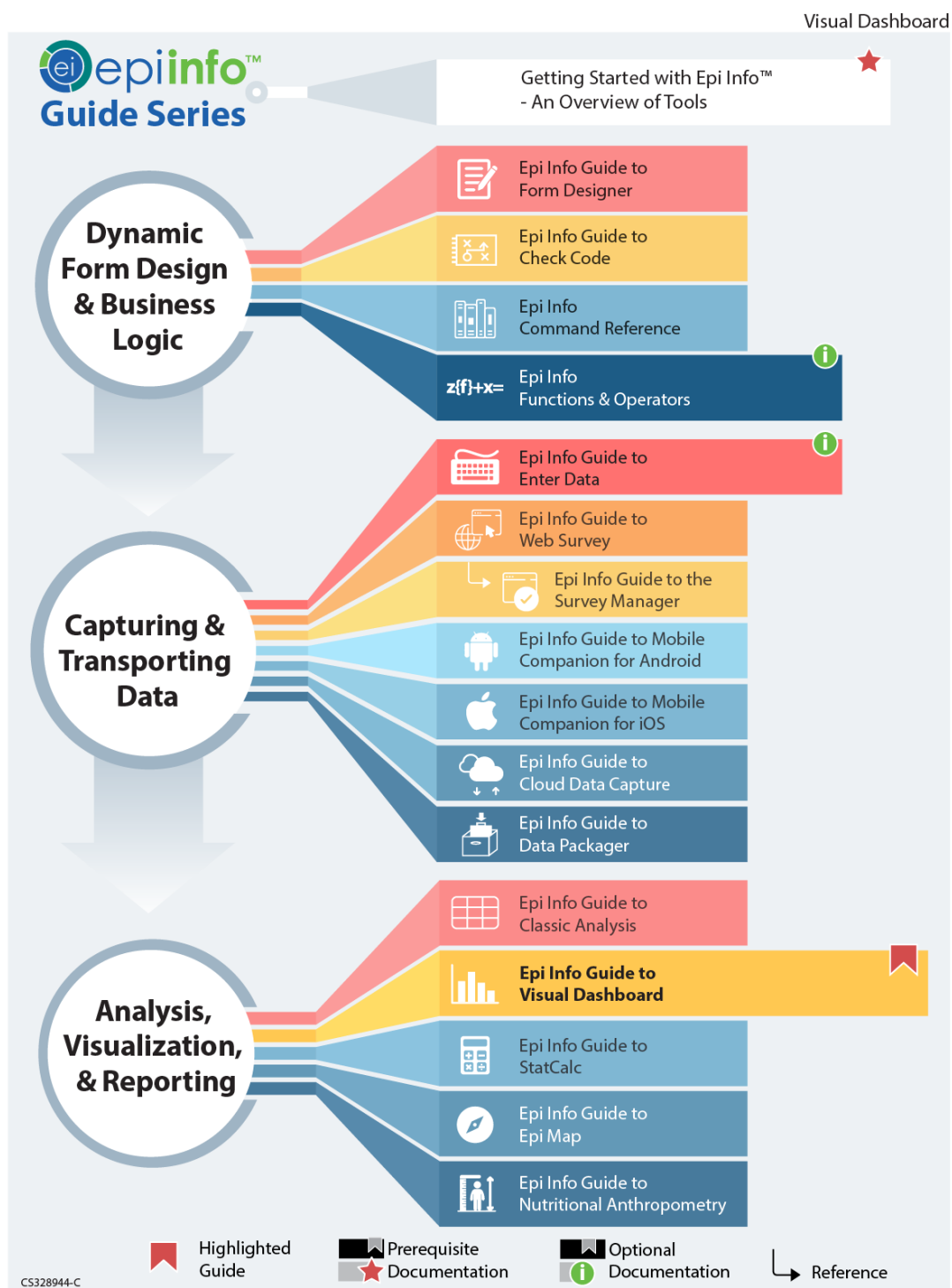
Epi Info™ Guide to Visual Dashboard

Epi Info™ 7

Preface

Epi Info™ is a public domain suite of interoperable software tools designed for the global community of public health practitioners and researchers. Perform data analysis with epidemiologic statistics, maps, and graphs. Build data entry forms, construct a database, and customize statistics applications. Physicians, epidemiologists, and public health officials without a technical background can easily work with critical data using Epi Info™ tools.

Use this guide alone or as a supplement to other Epi Info™ guides. On the next page, view the complete collection of helpful guides available for Epi Info™ tools. The diagram shows how this guide, highlighted with the bookmark icon ⚑ fits into the big picture of the Epi Info™ suite. Additional guides may offer prerequisite information for the tool you're working with. For example, if you're working with Epi Info™ Visual Dashboard, you'll find it helpful by referencing the Epi Info™ Guide to Classic Analysis, Stat Calc, and Epi Map. The image below highlights the Guide to Visual Studio with the prerequisite documentation icon (★). You may find helpful other optional guides that are highlighted by the optional documentation icon (ⓘ).

The image below shows the interrelationships among the guides in the series. This guide is **Getting Started with Epi Info™ Guide to Visual Dashboard**.

TABLE OF CONTENTS

Defining New Variables

Adding and Removing Data Filters

Displaying Statistics and Records

iv

Visual Dashboard: Performing Statistical Analyses with Visual Tools

Overview

Visual Dashboard is one of the Epi Info™ 7 analysis modules. Visual Dashboard is intuitive and simple to use. It simplifies analytic tasks by reducing traditional coding requirements. Instead, intuitive gadgets sort, list, or manipulate data. Statistical analysis gadgets in Visual Dashboard include frequencies, means, and more advanced statistical calculations (e.g., linear regression and logistic regression). Graphing functionality is native to Visual Dashboard. You can 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.
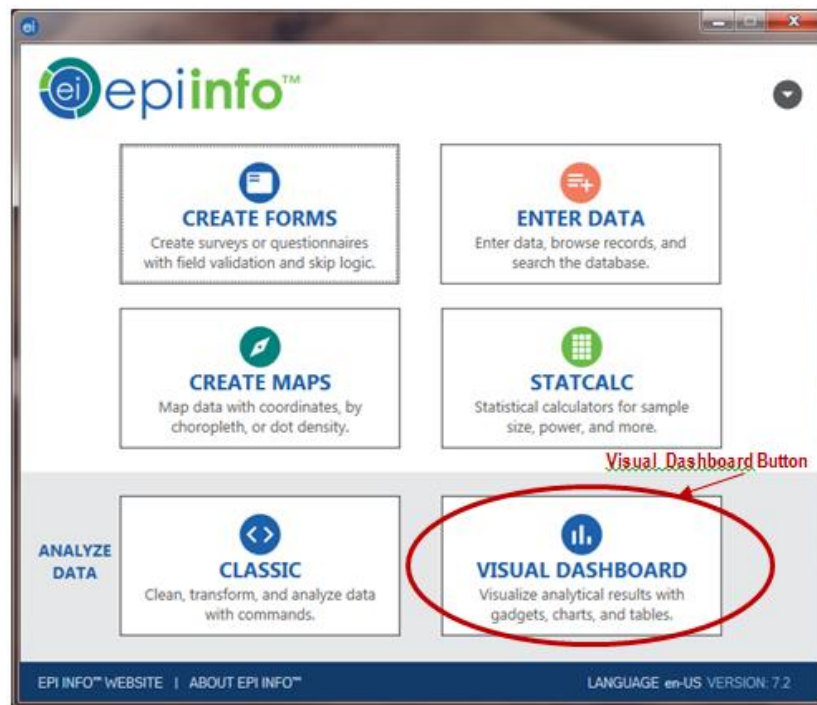
Main menu Navigation

Navigating the Canvas

Visual Dashboard Navigation

Visual Dashboard contains seven areas:

- Canvas
- Canvas Display Mode Selector
- Data Source and Record Count Display
- Canvas Output Controls
- Canvas Tools Toolbar
- Defined Variables Gadget
- Data Filters Gadget
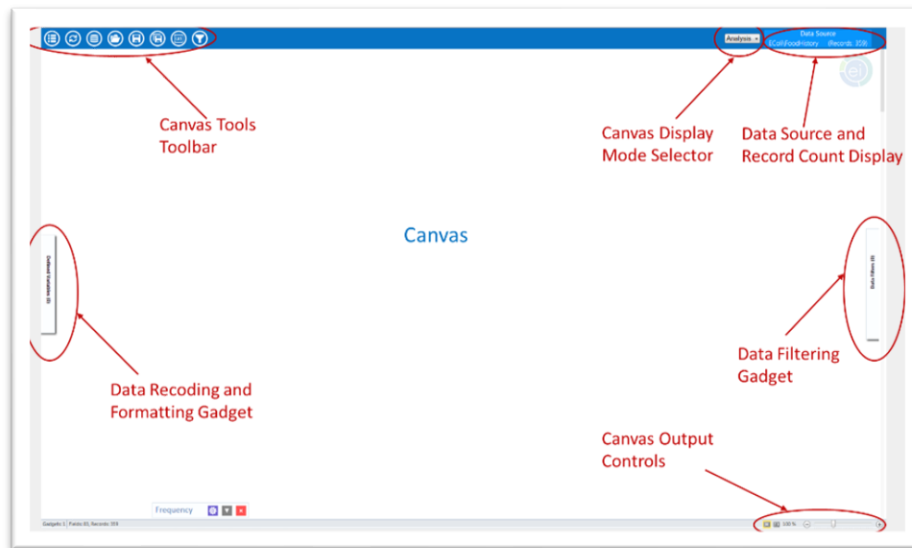
Visual Dashboard canvas

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. Right clicking on the canvas launches a function dialog box.

The **Canvas** Display Mode is in the upper right corner of Visual Dashboard.

Settings:

- Analysis (default) – Displays analysis gadgets.
- Data – Browse data in the loaded data set in grid format.
- Variables – displays a **Data Dictionary** for data set variables.

You can find the **Data Source and Record Count Display** at the top right of the screen, displays the source and record count of the data set. The number of fields (columns) in your data set displays on the bottom left of the Windows status bar.

The **Canvas Output Controls** are on the bottom right side of the Dashboard workspace. Canvas output controls contain three options.

- **Hide Borders button** – turns off external borders for each gadget, resulting in a report like appearance.
- **Auto-arrange button** – positions gadgets on the canvas, maximizing the display area.
- **The Zoom Percent and Zoom Slider** – zooms in and out over the canvas.

You can create, edit, and delete variables from the analysis session, using the **Defined Variables Gadget**. The Defined Variables Gadget is in the slide-in drawer on the left side of the canvas. Apply canvas wide filters using the **Data Filtering Gadget** (or **Data Filters**). Use it to select data for analysis. The Canvas Tools toolbar is at the top left corner of the canvas.

Features List:

- **Options** – opens options menu.
- **Refresh** – refreshes your screen with new data.
- **Set Data Source** – sets your data source.
- **Open** – opens a saved canvas.
- **Save** – saves your current canvas back to the original filename and location.
- **Save As** – saves your current canvas to a new name or new location.
- **Variables** – opens the Defined Variables Gadget.
- **Filters** – opens the Data Filtering Gadget.
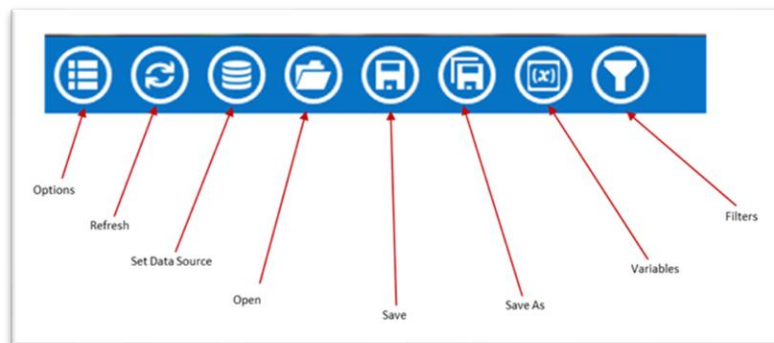
Canvas toolbar

Canvas Properties

Access Canvas Properties after you load a data set. Right clicking on the canvas launches the options menu. Select Canvas Properties.

Use these tools 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

Set the data source. Click the **Options menu** icon.

Toolbar highlighting Options menu

The **Options menu** has many features you can use with a set data source.
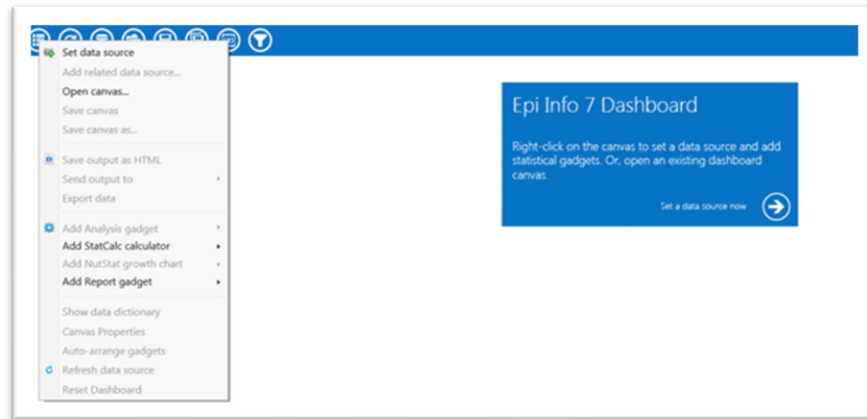
Options menu

Updating the Canvas with Refresh

The Refresh icon updates gadgets with fresh data.
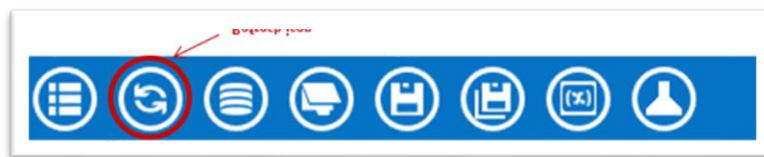
Visual Dashboard toolbar, with Refresh icon circled

Run from Command Line

Run Visual Dashboard from a batch program, custom menu, or command line. Associate the Visual Dashboard executable (visualdashboard.exe) with the CVS7 extension. Open your canvas using its saved file name.

If the .CSV7 file extension is not associated with the Visual Dashboard program, run the AnalysisDashboard.exe program followed by the name of the canvas using

the **/canvas:** switch like this**: `analysisdashboard.exe /canvas: c: "myfile.*"`**

These examples are given with the assumption that Epi Info™ 7 is in the **C:\Epi Info 7** folder on a C: drive. If your copy of Epi Info™ 7 is in another location, use your specific file location or a relative path.
You can launch Visual Dashboard by using the run command or through the command line (i.e.: AnalysisDashboard.exe). You can also click on the start icon or shortcut the within the Epi Info™ 7 folder.

Open a specific canvas using the **`/canvas:`** switch. In this example, open Visual Dashboard**.** The saved **Ecoli.cvs7** canvas is in **`C:\Epi Info 7\Projects\Ecoli`** folder.

Example:

C:\Epi Info 7\AnalysisDashboard.exe /canvas: "C:\Epi Info 7\Projects\Ecoli\Ecoli.cvs7"

Selecting a Data Source

Visual Dashboard works with Epi Info™ 7 project files, CSV files, Microsoft Excel, Microsoft Access, Microsoft SQL Server, MySQL, RedCap, and PostgreSQL. Select a data source by choosing any of the following steps.

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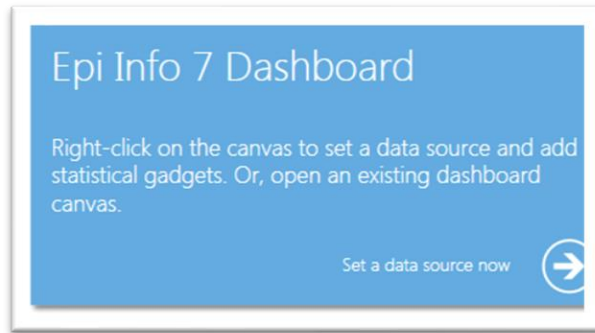
Set a data source

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

Set Data Source icon

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

After selecting the **Set Data Source** option, the Select Data Source dialog box displays. Select a **Database Type, Data Source**, and **form**, to open data in Visual Dashboard. View the following features:
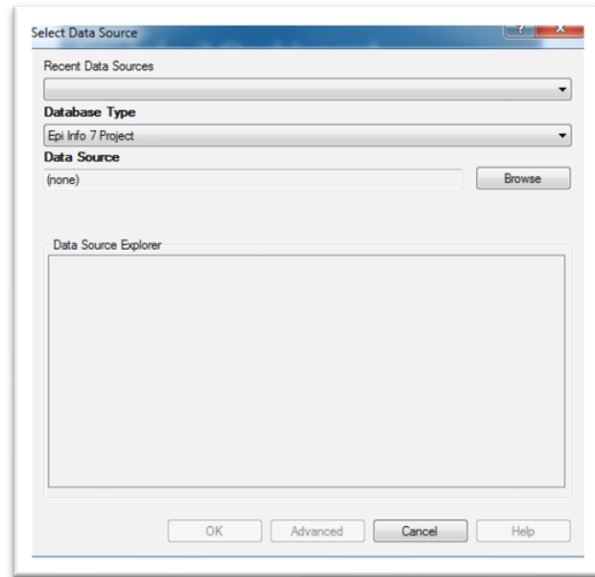
Select Data Source dialog box

Recent Data Sources – lists recently accessed databases from **Classic Analysis** or Visual Dashboard**.** You don't need to know the Database Type or Data Source Information. If you are reading an Epi Info™ 7 project file, all available forms appear in the Data Source Explorer box.

Database Type – field shows database files by database name (for some the file extension i.e. MDB, XLS are also provided). Specify the format for your data file. Locate output files in the default destination folder.

Data Source – field shows the file location/path. If you use a SQL Server database, the **Data Source** field requires server and database names.

Data Source Explorer – field allows you to select a project form. Explorer reads the form data table, and the Data Source field shows the file location and path. If

you use a SQL Server database, the Data Source field requires server and database names.

Advanced mode – allows SQL queries to select the data source. Currently, **Advanced Mode** supports **select** statements only.

Working with File Types

Read an Epi Info™ 7 Project File

1. Set a Data Source by clicking the **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 displays.
2. Select an Epi Info™ 7 Project from the Database Type drop-down list,
3. Click **browse** next to the **Data Source** field. The Epi Info™ 7 project folder displays.
4. Browse for your project file in Visual Dashboard. If your project is in a pre-selected folder, go to the folder and then open your project (PRJ) file. Epi Info™ 7 comes with some sample projects like Ecoli.prj. Click **Open.**
5. Select a **form**. The Data Source Explorer box populates with a list of all the forms that are associated with the project.
6. Click **OK**. The record count and data sources appear in the toolbar. Don't see it? Look at the top right of the canvas.
7. Click **OK**. The **Processing Finished** message displays, showing that the system has cached the data. Perform analysis on your data.

Read a Microsoft Excel File

1. Click **Set a Data Source Now** arrow. The Select Data Source dialog box displays.
2. Select either **MS Excel 97-2003 Workbook** or **MS Excel 2007 Workbook**, from the Database Type drop-down list,
3. Click **Browse**, next to the Data Source field. The **Open Existing File** dialog box displays.
4. Enter the **filename and path** to open your workbook or click **Browse** to search for the MS Excel workbook (XLS or XLSX extension).
5. Click **Open**. The First row contains header information. The checkbox is selected by default. If this option is unchecked, imported data won't have column headings.
6. Click **OK**, then select a worksheet from the list in the Data Source Explorer window.
7. Click **OK**. The **record coun**t and **data source** appear in the menu bar. Can't see it? Look for it at the top right of the Visual Dashboard canvas.

Read a Microsoft Access File

1. Click Set a Data Source now arrow. The Select Data Source dialog box displays.
2. Select either MS Access 2002-2003 (MDB) or MS Access 2007(ACCDB), from the Database Type drop-down list.
3. Click **Browse** next to the Data Source field. This displays the **Open Microsoft Access File** dialog box.
4. Enter the **filename** and path to open your file or click **Browse** to search for the MS Access file (MDB or ACCDB files). You can enter your **password** if one is set.

5. Click **OK** and select a data table from the list in the Data Source Explorer window.

6. Click **OK**. The record count and data source display 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 displays.

2. Select CSV File from the **Database Type** drop-down list,

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 and click **OK**.

Note: The First row contains header Information. The box is selected by default. If this option is unchecked, imported data won't have column headings.

6. Select the **filename** from the list in the Data Source Explorer window.

7. Click **OK**. The record counts and data source appear in the menu bar. Don't see it? Look at the top right of the 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 displays.

2. Select **Microsoft SQL Server Database**, from the Database Type drop-down list.

3. Click **Browse** next to the **Data Source** field.

4. The Connect to SQL Server database dialog box opens. Enter the **Server name and Database name**.

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. Select a **data table** from the list in the **Data Source Explorer** window,

8. Click **OK**. The record count and data source display in the menu bar. Don't see it? Find it at the top right of the 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 displays.

2. Select **MySQL Database** from the **Database Type** drop-down list,

3. Click Browse next to the **Data Source** field.

4. The **Connect to MySQL database** dialog box displays. Enter the **Server name** and **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 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 show in the menu bar at the top right of the Visual Dashboard canvas.

Add a Related Data Source

Perform analysis in Visual Dashboard data from multiple data sources. The **Add a Related Data Source** option, links tables or forms using a common identifier. Linking a form or table requires an important field which relates the records together. Currently, the link can only be done using one field. If your common identifier consists of multiple fields, then please review the **Classic Analysis** User Guide. Forms created in Form Designer or Visual Dashboard relate to each other using the Global Record ID (GLOBALRECORDID) of the first form and the Foreign Key or (FKEY) column of the related form. The demonstration below shows how adding a related data source works.

Epi Info™ 7 contains a related database in the **Sample.prj** file with two forms named Surveillance(parent). It has three child forms named RHepatitis, RLyme, and RMeningitis. In the example below, we relate Surveillance (parent) to RHepatitis (child).

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. Set a **Data Source** and select the **Sample.prj Data Source**, again.
3. Select **RHepatitis** from the list of forms in the **Data Source Explorer**. The **Child Key Field** drop-down list populates.
4. Select **GlobalRecordId**, from the **Parent Key Field** drop-down list.
5. Select **FKEY** from the **Child Key Field** drop-down list. **FKEY** is the foreign key in the child form. It matches **GlobalRecordId** in the parent form.
6. Check the **Use Unmatched (ALL)** checkbox. It includes all records from the primary form, regardless of matching records in related data source. Uncheck

the **Use Unmatched (ALL)** checkbox to include only records represented by keys that match in both tables.

7. Click **OK**. The **Processing Finished** message displays. The data is now local. Use the data from two tables 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 analyses, please review the **Classic Analysis** User Guide.
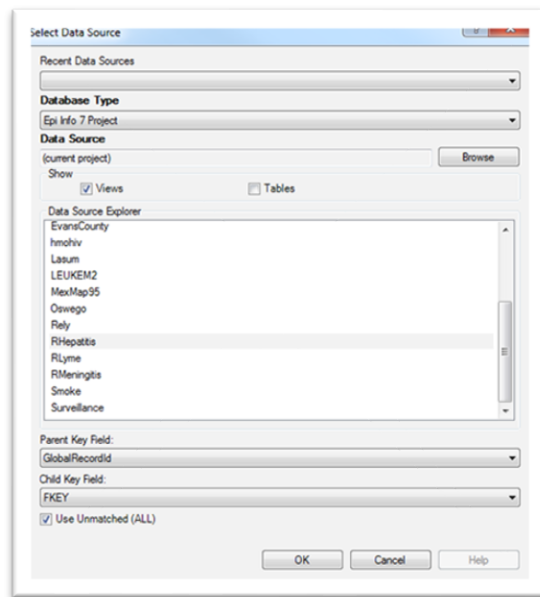
Add Related Data Source dialog box

Using Canvas Controls

Open Canvas

Open an existing Visual Dashboard canvas by selecting the **Open icon** on the **Canvas Tools** toolbar. You can also right-click on the canvas to open the **Canvas Options** dialog. Epi Info™ 7 includes several sample canvas files including **Oswego.cvs7**.

Toolbar highlighting Open icon

If you don't see your saved canvas in the **Quick Access List**, navigate to a saved canvas file using Windows Explorer.
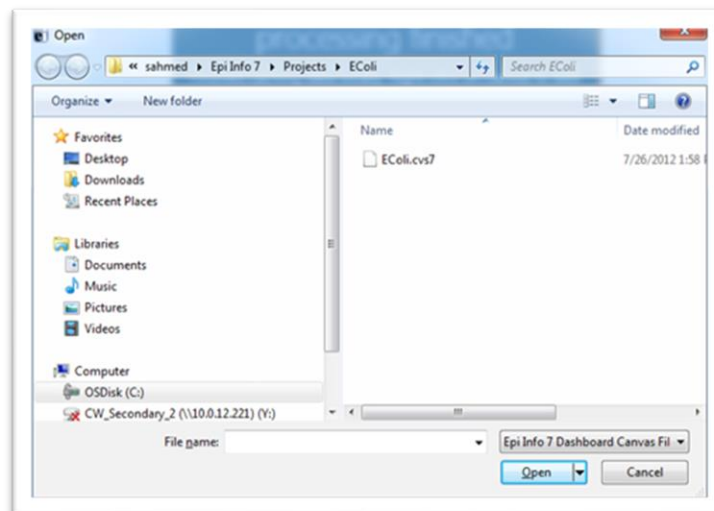
Open a Canvas

When your canvas loads, you'll see the confirmation message at the bottom of the screen. You can also click the **Open file** folder icon in the top left corner of the canvas for quick access to saved canvas files.

1. Select the location of the canvas file. The file can be identified by the file name extension CVS7.
2. Click Open. The canvas displays in Visual Dashboard, connects to the data source, and displays the latest results.
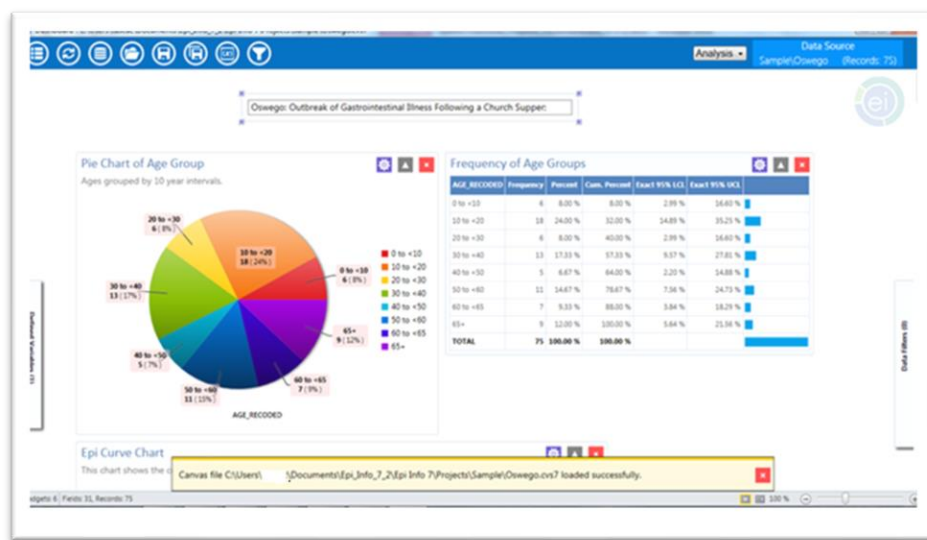
Open Canvas

The Canvas filename displays in the yellow message field at the bottom of the screen. You can also click the open file folder icon in the top left corner of the canvas for quick access to saved canvas files.

Save Canvas

Save the current canvas to include the data source connection, analysis gadgets, data filters, and user-defined variables for reuse. You won't need to recreate the work previously done during your analysis.

Save the canvas by right-clicking on the canvas. Select **Save Canvas**. You can also click the **Save icons** at the top left corner of the canvas.

⚠️ Warning! This action will overwrite an existing saved canvas.

- If this is the first time saving your canvas, assign a filename, and choose a file location.

Toolbar with highlighting Save icon

Save Canvas As

The **Save Canvas As** option allows the canvas to be saved for the first time or saved to a new name or 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 displays.
2. Select a location to save the canvas and enter a **filename**.
3. Click **Save**.

Toolbar highlighting Save As icon

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

Save Visual Dashboard HTML

You can save your Visual Dashboard output from gadgets as an HTML document. Save your output by right-clicking on the canvas and select **Save Output** as HTML. The **Save As** dialog box displays.

Select a location to save the canvas output and enter a filename. The default location for saving output is to the **Epi Info™ 7 Output** folder.
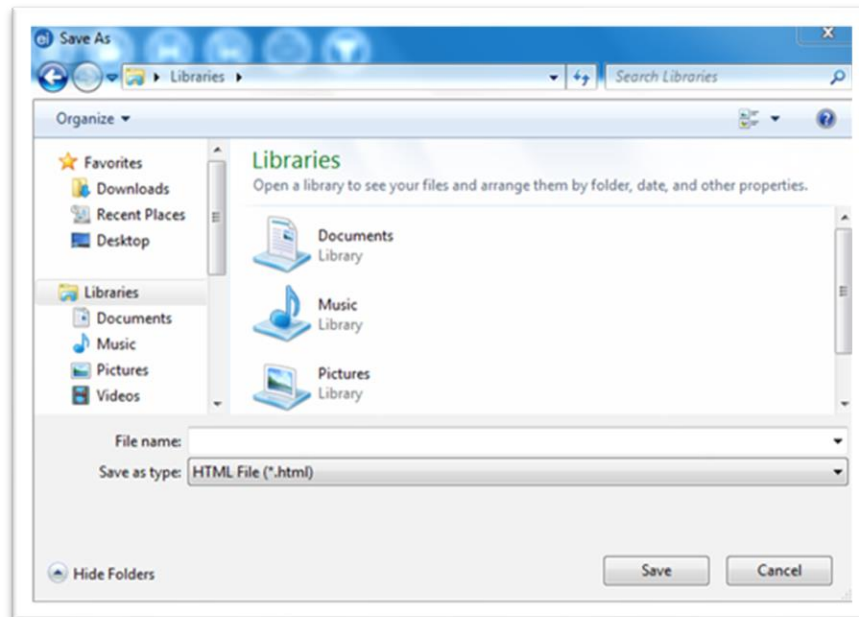
Save output as HTML

Click the **Save** button. The output opens in the web browser.

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

Send Visual Dashboard Output

If you have MS Word and Excel, you can view outputs from Visual Dashboard with the **Send Output To** option.

Microsoft Excel

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

Microsoft Word

Right-click on the canvas, select **Send Output To > Microsoft Word**. The Visual Dashboard output opens in Microsoft Word as an .HTML formatted web document. Saving in this format your charts reside in a sub-folder. When sharing this HTML file, add the sub-folder as well. Alternatively, change the format to a Microsoft Word DOCX format. Then, your charts display with the DOCX file as embedded images.

Defining New Variables

You to create, edit, or delete variables with no coding in the Defined Variables gadget in Visual Dashboard. 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.

Toolbar, with Variable button

Defined Variables Gadget Controls:

- Create New Variable
- Recoded Value
- Simple Assignment
- Conditional Assignment
- Formatted Value

- Assigned Expression
- Create Variable Group
- Edit Variable
- Delete Variable

Creating a New Variable

Data isn't always in the proper format for analysis. As an example, consider analyses for patients by age, using age groups. Some data may contain county codes, but descriptive county names might be more useful. Creating a new variable is a function of the **Defined Variables** gadget. Create new variables that make your data more useful for analysis. For example, recoding patient ages to age groups in five-year increments or assign county names to county codes. Create new variables and change the formatting of existing data with **Recoded Value**, **Simple Assignment**, **Formatted Value**, **Assigned Expression**. Then, try grouping existing variables as you wish.

With Recoded Value

The Recoded Value option assigns a value from an existing variable and places the results into a new variable. Click **New Variable** in the gadget and select **With Recoded Value**. The **Add Recoded Variable** dialog box displays.

View the field definitions below the following image.

Add Recoded Variable dialog box

- **Source field** – drop-down list showing the column names in the current data that you can recode.
- **Destination field** – a text box for assigning the name of a new temporary column in the current data source to store recoded values.
- **Destination field type** – representing data types for new columns that store recoded values. Options include Text, Numeric, or Yes/No.
- **Maintain Sort Order Where Appropriate** – Output in Visual Dashboard sorts alphabetically or sequentially when the box is checked. Leave this box checked if you want the output to sort in the same order as entered.
- **Use Wildcards** – option for text searches in the **From** column. This option is unchecked by default. Identify the asterisk (*) represents a wildcard character.

- **From** – value points to a range or single value. To recode a single value instead of a range, use the **From** column and leave the **To** column blank.

When recoding a range of values, the recode range includes **From** column values that **are less than but not equal** (Recode range ≤ **To** column value) values in the **To** column. To recode all unspecified values, leave both **From** and **To** columns blank.

- **To** – value (if any) identifying the upper limit of a range of values to be recoded. The recode range does not include the value in the **To** column. Select a numeric variable and this option displays.
- **Else Value** – represents an optional value assignable to records that do not match the recoding rules above. For non-matching records or recoding rules, the assigned value is null, blank, or missing.
- **Representation Value** – identifies an assigned value to a destination variable, with specified values from the source field variable.
- **Fill Ranges** – Use **Fill Ranges** to populate range values on a numeric source field. **Fill Ranges** quickly create numeric groupings like age categories based on ranges of equal distribution.
- **OK** – accepts your settings and closes the dialog.
- **Cancel** – exits the window without saving or executing a command.

The following example shows how to add a recoded variable:

In the **E.coli Food History** table, there are over 60 different values for age. Analysis may be challenging when comparing various age values. Grouping age values in 10-year increments, simplifies analysis using the age variable.
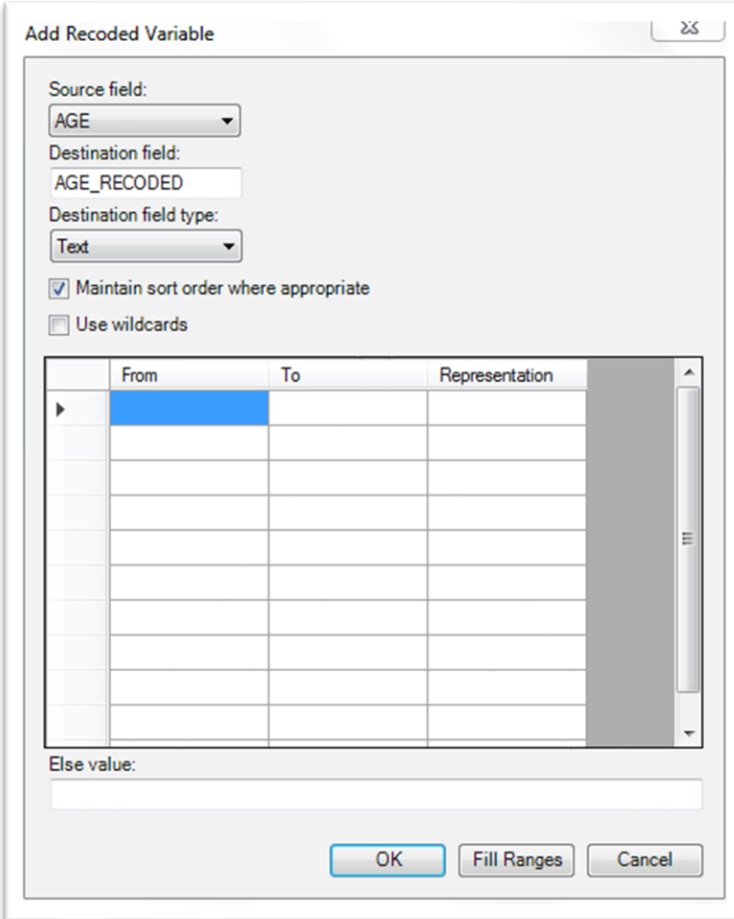
Group Age values in categories Part 1:
1. Set the Data Source to the **EColi.prj** project located in the **Epi Info 7\Projects\Ecoli** folder.

2. Click **FoodHistory**, then Click **OK**.

Move the mouse cursor over the **Defined Variables gadget** located at the left-hand side of the Visual Dashboard canvas. The gadget fully expands. Also, try clicking the **Variables** icon on the top left of the canvass.

3. Click the **New Variable** button.
4. Select **With Recoded Value**. The **Add Recoded Variable** dialog box displays.

Add Recoded Variable dialog box

Group Age values in categories continued:

5. Select **Age** from the Source field drop-down list,

6. You can modify the name of the new variable, if desired, in the Destination field box. Otherwise, the variable name assigned will be based on the Source field name followed by the suffix **_RECODED** format (i.e. AGE_RECODED).

7. Click the **Fill Ranges** button. The Fill Ranges dialog box displays.

8. Enter **0** for **Start value**, **65** for **End value**, and **10** for **By**. Click **OK**.
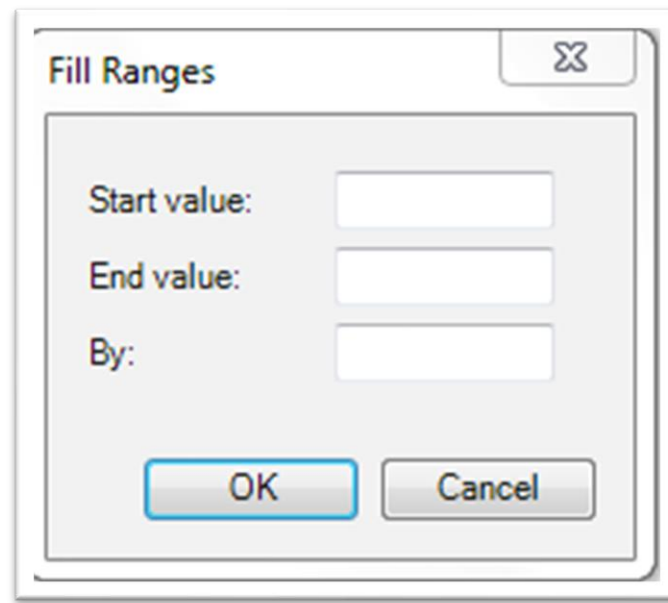
Add Recoded Variable/Fill Ranges

The Add Recoded Variable dialog box displays with the **From, To,** and **Representation columns.** Click **OK** and the recoded variables and conditions show in the **Variable Definitions text box**.

Add recoded variable ranges
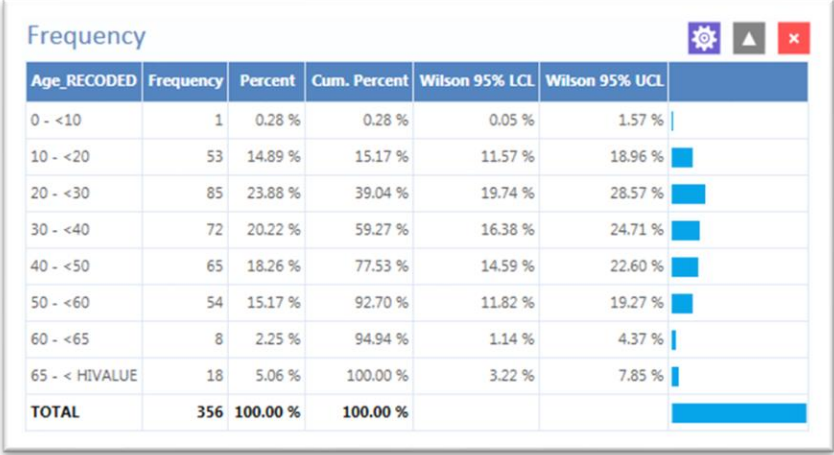
Variable Definitions dialog box

View the new frequency of **Age_RECODED** values by right-clicking on the canvas. Next, go to **Add Analysis Gadget > Frequency**. Select **Age_RECODED** from the **Frequency of** list 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 %**			

Age_Recoded Frequency list

 The **Age_Recoded** variable contains eight distinct values, making analysis easier than the original data set.

With Simple Assignment

Simple Assignment assigns values from existing variables to newly defined variables.

Simple Assignment steps Part 1:
1. Set the **Data Source** to the **Sample.prj** project located in the **Epi Info 7\Projects\Sample** folder.
2. Select **Surveillance** from the **Data Source Explorer** menu and click **OK**.
3. Move the mouse cursor over from the left-hand side of the Visual Dashboard canvas, hovering over the **Defined Variables** gadget. Also, you can click the variables icon in the top left. The gadget fully expands.

4. Click the **New Variable** button.

5. Select With **Simple Assignment**. The **Simple Assignment** dialog box displays.
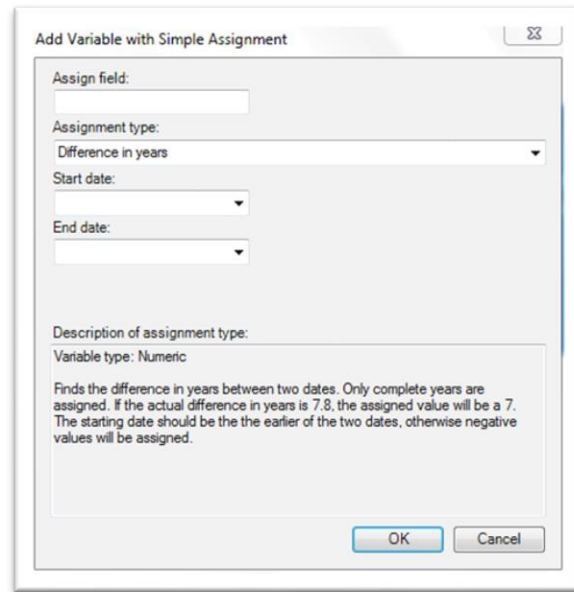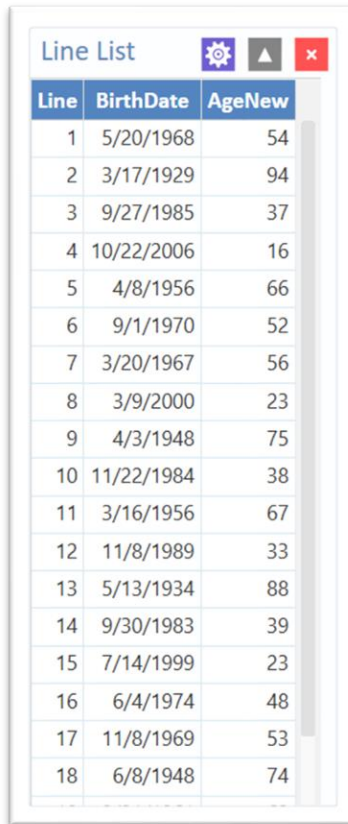
Add Variable with the Simple Assignment

Simple Assignment steps part 2:

1. Enter **Age_New, i**n the **Assign** field.

2. Select **Difference in years** from the **Assignment type** drop-down list.

3. Select **BirthDate** from the **Start Date** drop-down list.

4. Select **SYSTEMDATE** from the **End Date** drop-down list. The system date is a system-level variable representing the current date and time on a host computer.

5. Click **OK**. A description of the assignment displays in the **Variable Definitions** text box within the **Defined Variables** gadget.

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

Simple Assignment Line List

The example shows the current age. The age is the difference between the patient birth date and the current date (on the host computer). **Analysis Gadget > Line List** view displays the new variable created via Simple Assignment.

With Conditional Assignment

Conditional Assignment defines one or more criterion for a new variable. Execution of the Conditional Assignment occurs if the criteria is true. Alternative assignments occur after the **Else statement**. The **Else statement** executes when conditions are not true. When a condition is not true, and the **Else statement** has not been selected, no assignments are added.

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. Also, try clicking the **Variables** icon on the top left of the canvass. The gadget fully expands. Click the **New Variable** button and select **With Conditional Assignment**.

Add New Variable with Conditional Assignment

Field descriptions of the Add New Variable with Conditional Assignment box are listed below:

- **Assign Field** – add the new variable name.
- **Assign Field Type** – shows the format of the new variable. The options are Text, Numeric, or Yes/No.
- **Create/Edit Condition** – allows the creation of criteria to be executed. Applies assigned values when specified conditions for the text box are true.

31

- **Assign Value** – states the value of a new variable when the **Assign Condition** is true.
- **Use Else** – checkbox shows a specified **Else Value**. This checkbox is unmarked by default.
- **Else Value** – occurs when the statement in the **Assign Condition** text box is false.

View the following example. if the **Age** variable has a value between 17 and 30, **YoungAdult** is assigned the **Yes** variable. The **No** variable is assigned when the **Age** value falls outside of this range.

Conditional Assignment Part 1:
1. Set the **Data Source** to the **Sample.prj** project located in the **Epi Info 7\Projects\Sample** folder.
2. Select **Surveillance** from the **Data Source Explorer** menu and click **OK**.
3. Move the mouse cursor over the **Defined Variables gadget**. Find it on the left-hand side of the Visual Dashboard canvas. You can expand the gadget for full visibility.
4. Click the **New Variable** button.
5. Select **With Conditional Assignment**. The Add Variable with Conditional Assignment dialog box displays.
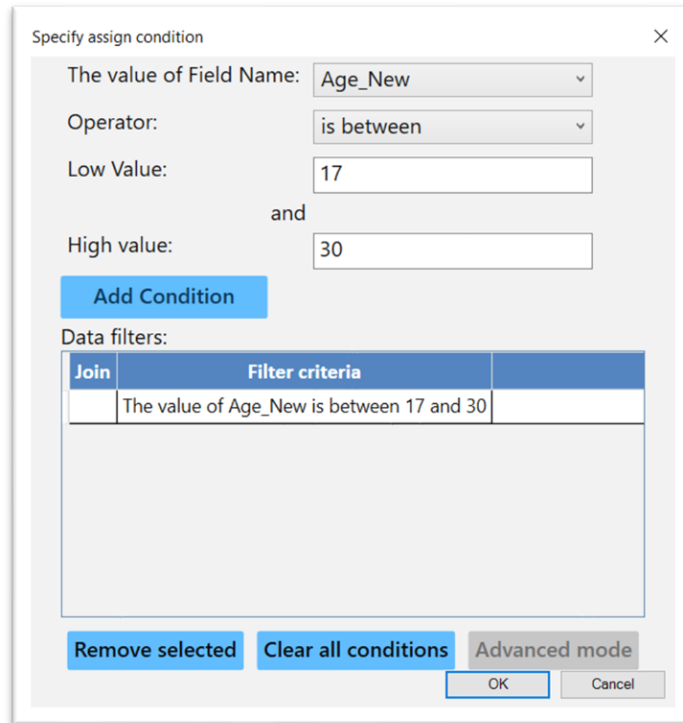
Conditional Assignment / Create Edit Condition

6. Enter **YoungAdult**, in the **Assign** field.

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 displays.

Add Condition dialog box

Conditional Assignment Part 2:

1. Select **Age_New** from the **Field Name** drop-down list.

2. Select **is between** from the **Operator** drop-down list.

3. Enter **17** for the Low Value: parameter.

4. Enter **30** for the High value: parameter

5. Click **Add Condition**.

6. Click **OK**. The filter criteria appear in the **Data Filters** list.

7. Click **OK** in the **Specify Assign Condition** dialog box.
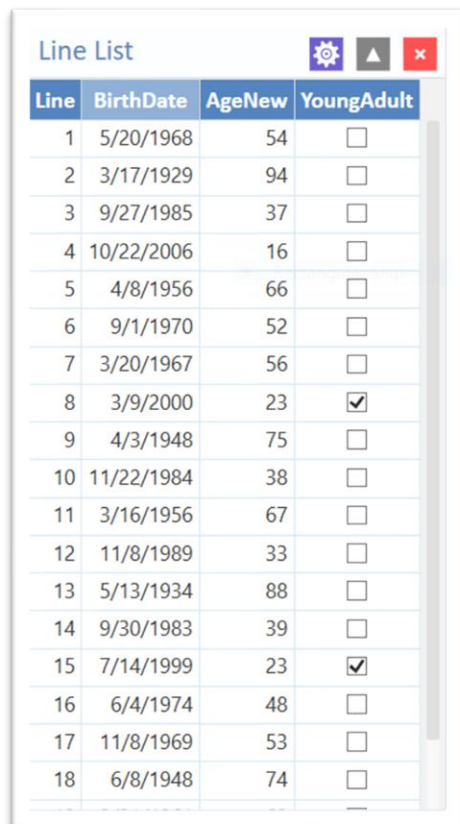
The condition now displays in the **Assign Condition** text box within the **Add Variable with Conditional Assignment** dialog box.

8. Select **Yes** from the **Assign Value** drop-down list.

9. Check the **Use Else** checkbox.

10. Select **No** from the **Else Value** drop-down list and click **OK**.

A description of the assignment displays in the **Variable Definitions** text box within the **Defined Variables** gadget.

You can see the **YoungAdult** variable in line list view. Right-click on the canvas and then select **Add Analysis Gadget > Line List**. Select **YoungAdult** and **AgeNew** from **List of Variables to Display** and click **OK**. Check marks highlight lines for which the conditional assignment is **true**.

Line	BirthDate	AgeNew	YoungAdult
1	5/20/1968	54	☐
2	3/17/1929	94	☐
3	9/27/1985	37	☐
4	10/22/2006	16	☐
5	4/8/1956	66	☐
6	9/1/1970	52	☐
7	3/20/1967	56	☐
8	3/9/2000	23	☑
9	4/3/1948	75	☐
10	11/22/1984	38	☐
11	3/16/1956	67	☐
12	11/8/1989	33	☐
13	5/13/1934	88	☐
14	9/30/1983	39	☐
15	7/14/1999	23	☑
16	6/4/1974	48	☐
17	11/8/1969	53	☐
18	6/8/1948	74	☐

YoungAdult Conditional Assignment Line List

Note: The Else clause, found at the end of IF statements, specifies actions for unmet conditions. Without the Else clause, YoungAdult would be null or missing for any values falling outside the 17 to 30 range.

With Formatted Value

You can change the format and name of existing variables with the **Formatted Value** option. The option is effective with date fields. View the example below.

1. Set the **Data Source** to the **Sample.prj** project.
2. Select **Surveillance** from the **Data Source Explorer** menu and click **OK**.
3. Move the mouse cursor over the **Defined Variables gadget**. You can also click the **Variables** icon in the top left of the screen. The gadget expands to full visibility. Click the **New Variable** button.
4. Select With **Formatted Value**. The Add **Formatted Value** dialog box displays.
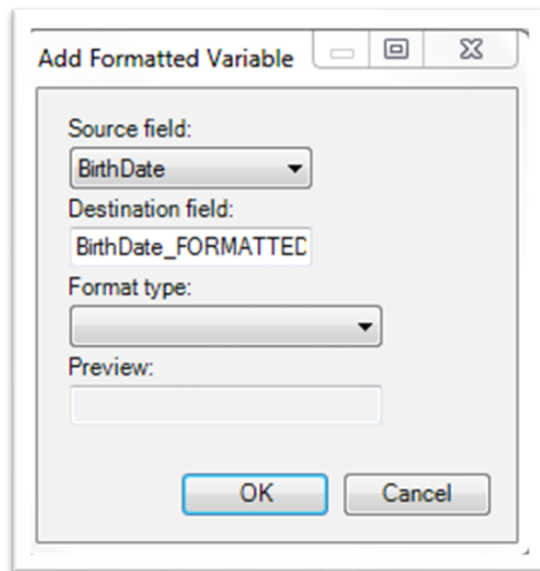
Add Formatted Variable dialog box

5. Select **BirthDate** from the **Source field** drop-down list.
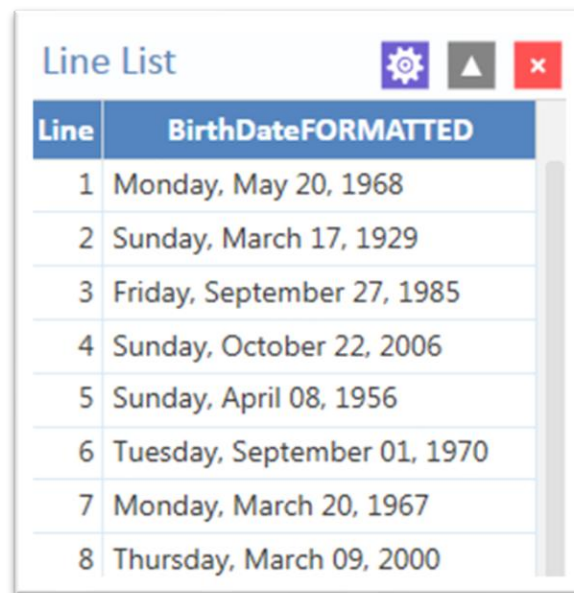
The Destination field automatically shows the source field with a suffix of "FORMATTED" (i.e., **Birthdate FORMATTED**). You can change the name of the Destination field.

6. Select the **long date** from the **Format type** drop-down list. Using the current system date, a sample long date value displays in the **Preview** field. Click **OK**.

A description of the assignment displays in the **Variable Definitions** text box within the **Defined Variables gadget**.

7. Right-click on the canvas and select, **Add Analysis Gadget > Line List**. Now you can see the **BirthDate_FORMATTED** variable in line list view. Select **BirthDate_FORMATTED** from **List of Variables** and then click **OK.**

Line List	
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

Formatted Variable line list

The **Analysis Gadget > Line List** displays the new variable created in **Formatted Variable**.

With Assigned Expression

Assigns the result of an expression. Also, it assigns a variable field value, as shown in the example below. In the example, the zip code field is a number. Use the zip code in a map. Define a new zip code variable and assign text values.

Steps:
1. Set the **Data Source** to the **Sample.prj** project.
2. Select **Surveillance** from the Data Source Explorer menu and click **OK**.

Move the mouse cursor over the **Defined Variables gadget** on the left-hand side of the Visual Dashboard canvas. Also, you can click the **Variables** icon at the top left of the screen. The gadget fully expands.

3. Click the **New Variable** button and select **Assigned Expression**. The Add Variable with Expression dialog box displays.
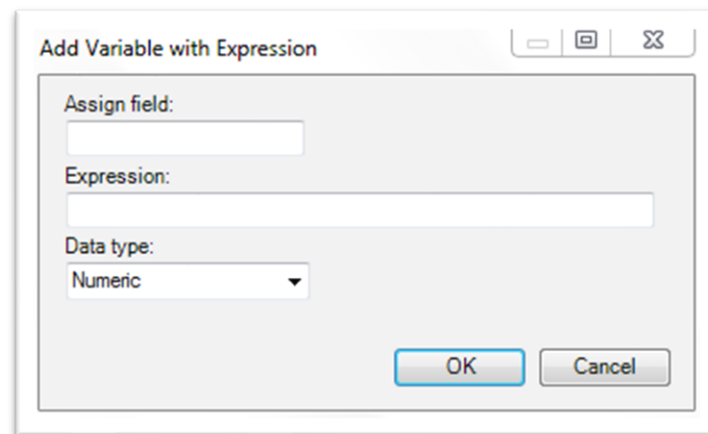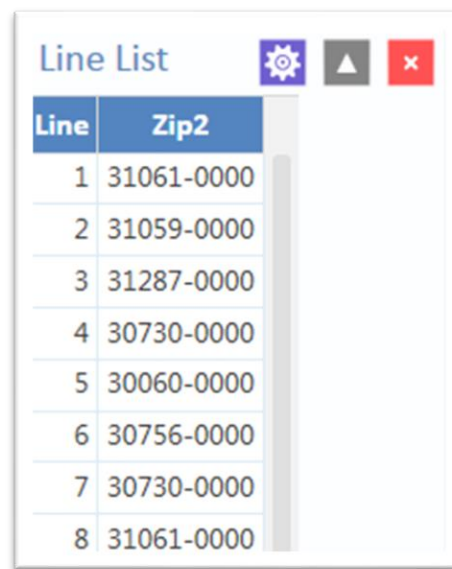
Add Variable with Expression

4. In the **Assign Field**, enter **Zip2**.

5. Type the syntax: `<ZipCode + '-0000'>` in the **Expression** field.

6. Select **Text** from the **Data type** drop-down list and click **OK.**

The assignment displays in the **Variable Definition** text box within the **Defined Variables gadget**. Right-click on the canvas and then select **Add Analysis Gadget > Line List**. Now you can see the **Zip2** variable in line list view.

7. Select **Zip2** from **List of Variables to display**. Then click **Generate Line List**.

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 multiple variables of the same format and analyzing each is impractical. Use Group variables with

Combined Frequency and **MxN/2×2 gadgets**. The following example shows how to create a variable group.

1. Set the **Data Source** to the **Sample.prj** project.
2. Select **Oswego** from the **Data Source Explorer** menu Click **OK**.
3. Move the mouse cursor over the **Defined Variables gadget** on the left-hand side of the canvas. You can also click the **Variable** icon in the top left of the screen. The gadget fully expands.
4. Click the **New Variable** button.
5. Select **Create Variable Group**. The **Create Group** dialog box displays.
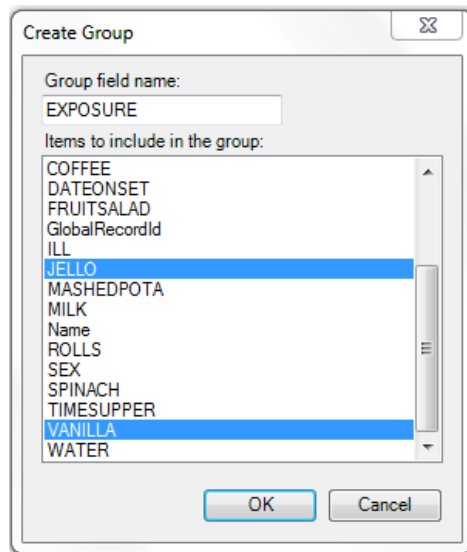
Create Group variable options

6. Enter **"Exposure"** in the **Group field name** and select **JELLO** and **VANILLA** from the choices for the group list. Hold down the **Ctrl key** on your keyboard to select more than one variable at a time. Click **OK**.
7. The **Defined Variables gadget** displays the variable group statement in the **Variable Definitions** text box.

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. Move the mouse cursor over the **Defined Variables gadget** on the left-hand side of the Visual Dashboard canvas. Also, you can click the **Variables icon** in the top left of the screen. The gadget expands to full visibility.
2. Click the **Edit Variable** button.
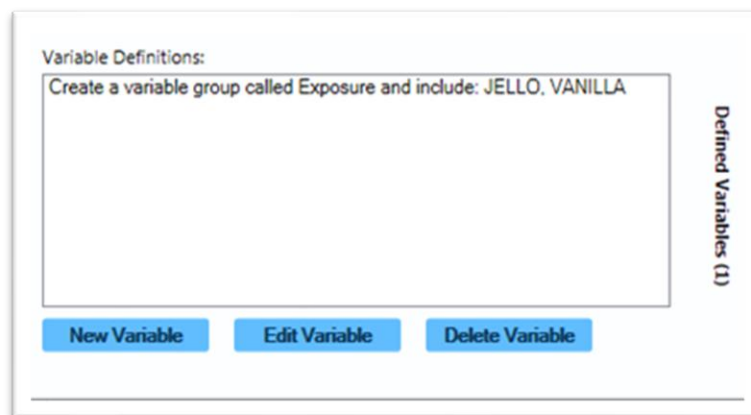
Variable group defined

3. The Create Group dialog box displays, with **JELLO** and **VANILLA** highlighted.
4. Press and hold the **Ctrl key** on your keyboard and select **SPINACH** to add it to the variable. Click **OK**.

The Exposure group variable now includes **Spinach**. It displays in the Variable Definitions text box 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. 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. Click on the **Exposure variable** definition within the textbox.
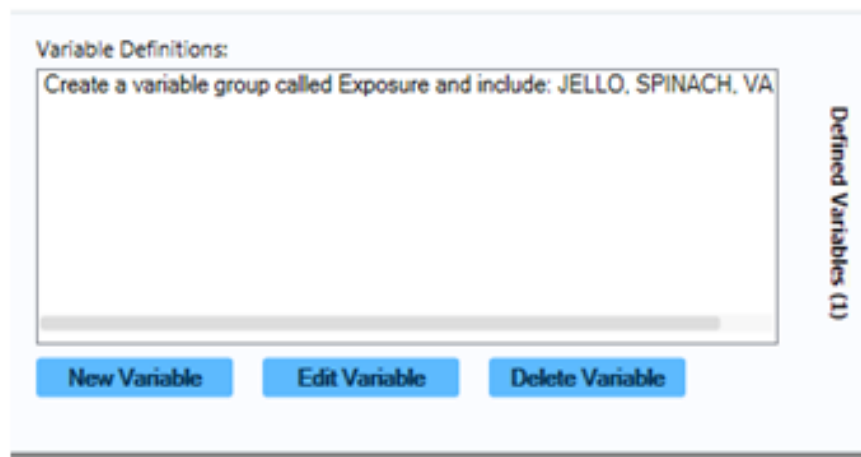
Variable Definitions box

3. Click on the **Delete Variable** button.

The variable is deleted from the **Variable Definitions** text box.

Adding and Removing Data Filters

Add a Filter

Visual Dashboard typically loads 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. Follow the steps below and add a filter in Guided Mode. Advanced mode is discussed in a later section.

1. Select the **EColi.prj** data source.
2. Click **FoodHistory** from the **Data Source Explorer** menu. Click **OK**.
3. Move the mouse over the **Data Filter gadget** from the right-hand side of the Visual Dashboard canvas. You can also click the **Filters gadget** (funnel) icon in the top left of the screen. The gadget expands to full visibility.
4. Select **ILL** from the **Field Name** drop-down list.
5. Select **equal to** from the **Operator** drop-down list.
6. Select **Yes** from the **Value** drop-down list.

Data Filter Gadget/Add Filter

Now you can filter record results with the Data Filter gadget.

7. Click the **Add Filter** button. The filter condition is added to the **Data filters** grid within the gadget. View the screenshot of the Data Filter gadget.

Data Filter Gadget/Add Filter with Condition

All gadgets on the canvas refresh when adding or removing a filter condition. The record count now reads 276, showing that there are 276 out of 359 records in which "The value of [ILL] is Yes."

Add another filter:
Move the mouse over the **Data Filter** gadget, from the right-hand side of the canvas. You can also click the **Filters** gadget at the top left corner. The gadget fully expands.

1. Select **Age** from the **Field Name** drop-down list.
2. Select **is between** from the **Operation** drop-down list.
3. Type **20** for the first box of the two **Value** text boxes.
4. Type **29** for the second box from the Value text boxes.
5. Click **Add Filter**. A context menu displays asking if you want to add this condition using an **AND** or **OR** operator. Choose the appropriate operator.

Adding a second condition

The filter condition is added to the **Data filters** grid. You can find it beneath the first condition.

The record count shows 64 records. This shows only 64 records match two established filter criteria.

- Criteria 1: the patient was ill.
- Criteria 2: The patient is between 20 and 29 years of age.

All gadgets added to the 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
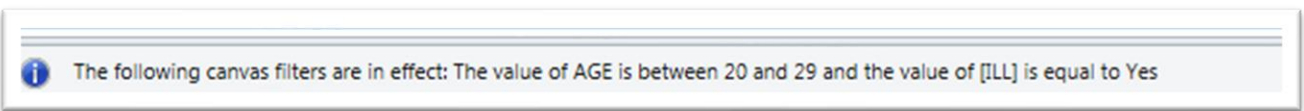
Data Filter setting

Remove Selected Filter

Follow the steps below to remove individual filters. Move the mouse over the **Data Filter gadget**, on the right-hand side of the Visual Dashboard canvas. You can also click the **Filters gadget** at the top left corner. The gadget expands to full visibility.

1. Remove the second condition, "The value of Age is between 20 and 29."
   Click on it to select it from the list.
2. Click the **Remove Selected** button. The condition displays.
3. Select the remaining **condition**.
4. Click the **Remove Selected** button. The condition displays, leaving any additional conditions. The system returns to the full dataset when all filters are removed.

Clear All Filters

The **Clear All** button removes all existing filters. Use this quick and easy step rather than removing each filter.

1. Move the mouse over the **Data Filter** gadget or click the **Filters** gadget in the top left corner of the right-hand side of the Visual Dashboard canvas. The gadget expands to full visibility.
2. Click the **Clear All** button. All existing filter conditions disappear.

Filter with Two Conditions
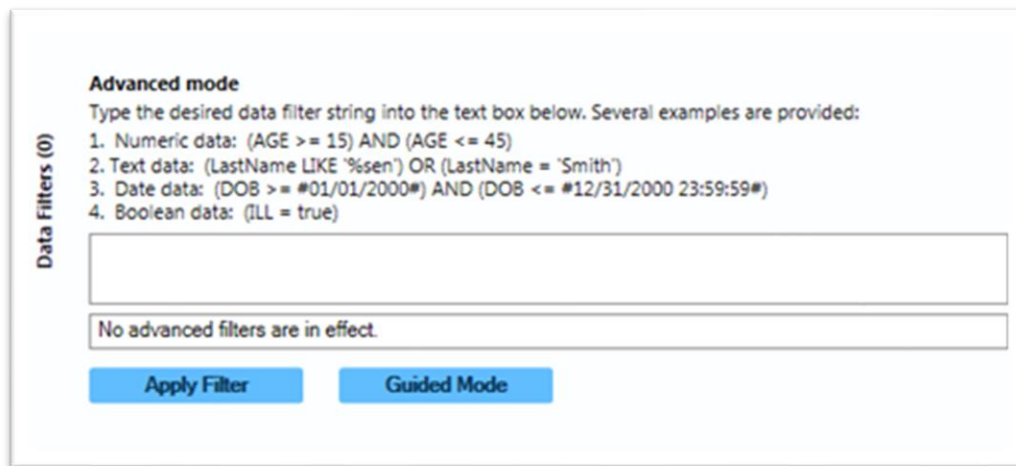
Advanced Filter Mode

Type the desired data filter string into the text box using Advanced Filter Mode. Apply the filter and the canvas shows records that meet specified conditions.

Note: Advanced mode uses the. NET DataView.RowFilter syntax. For more Information, refer to Microsoft's website for DataView.RowFilter Property. https://docs.microsoft.com/en-us/dotnet/api/system.data.dataview.rowfilter?view=net-5.0

The following example demonstrates how to use advanced filters.

1. Move the mouse over the **Data Filter** gadget or click the **Filters** gadget in the top left corner of the right-hand side of the Visual Dashboard canvas. The gadget expands to full visibility.

2. Select **Advanced Mode**. The Advanced mode dialog box displays.

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

Advanced Mode

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).

3. Click the **Apply Filter** button.

Instead of "No advanced filters are in effect," you will see, "The filter is valid and is now in effect." **Apply filters and conditions** displays in green font.
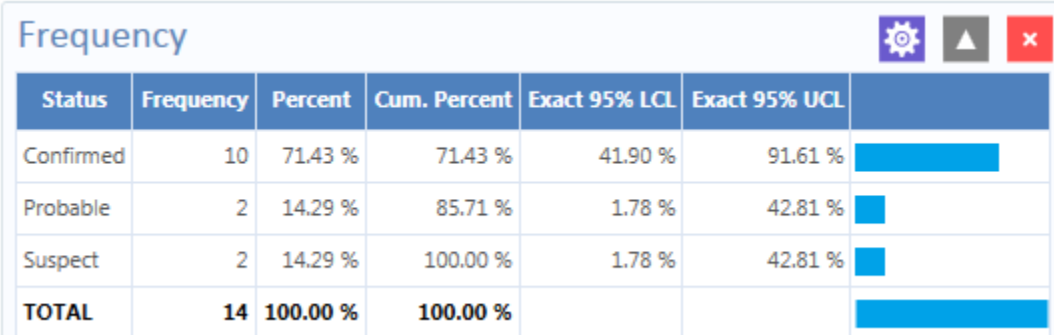
The record count in the top right corner of the Visual Dashboard canvas reads 64, showing that there are 64 out of 359 records where the value of **ILL** is **Yes** and the value for **AGE** is between 20 and 29.

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

Analysis is carried out in the Dashboard using Gadgets. Gadgets are moveable screen elements that perform specific statistical functions. For example, there is a Logistic Regression gadget that can carry out conditional and unconditional logistic regression. Shown below is an example of a frequency gadget. Multiple gadgets can be incorporated in the Visual Dashboard.

Status	Frequency	Percent	Cum. Percent	Exact 95% LCL	Exact 95% UCL	
Confirmed	10	71.43 %	71.43 %	41.90 %	91.61 %	
Probable	2	14.29 %	85.71 %	1.78 %	42.81 %	
Suspect	2	14.29 %	100.00 %	1.78 %	42.81 %	
TOTAL	**14**	**100.00 %**	**100.00 %**			

Frequency gadget output

Summary of Analysis Gadgets

Epi Info™ 7 Visual Dashboard includes a collection of analysis gadgets designed to view, interpret, and display data on the canvas. Customize your gadgets to highlight important data. Arrange and resize any combination of gadgets so you

can correlate information and draw conclusions. View the list of Analysis Gadgets below.

- **Line List** – View data in tabular format.
- **Frequency** – Shows number of records with field values.
- **Word Cloud** – Displays the most common words used in a font size proportional to word frequency.
- **Combined Frequency** – Displays frequency and percentage of grouped variables.
- **M x N / 2 x 2 Table** – Examine relationships between two or more categorical values.

- **Matched Pair Case - Control** – Compares an exposure variable, outcome variable like case vs. control. Match variables, linking each case to its control.
- **Means** – Calculates the average for a continuous numeric variable.
- **Duplicates List** – 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+
- o Pie Chart
- o Aberration Detection Chart
- o Pareto Chart
- o Scatter Chart
- o 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

Line List Introduction

The **Line List** option creates a list of the current dataset, so you can view details in your current data source. You can display, sort and group selected variables. **Line List** supports adding groups or pages. Select this option from the bottom of the **Variables to display** in 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**. The Line List gadget configuration window opens to the **Variables property panel**.
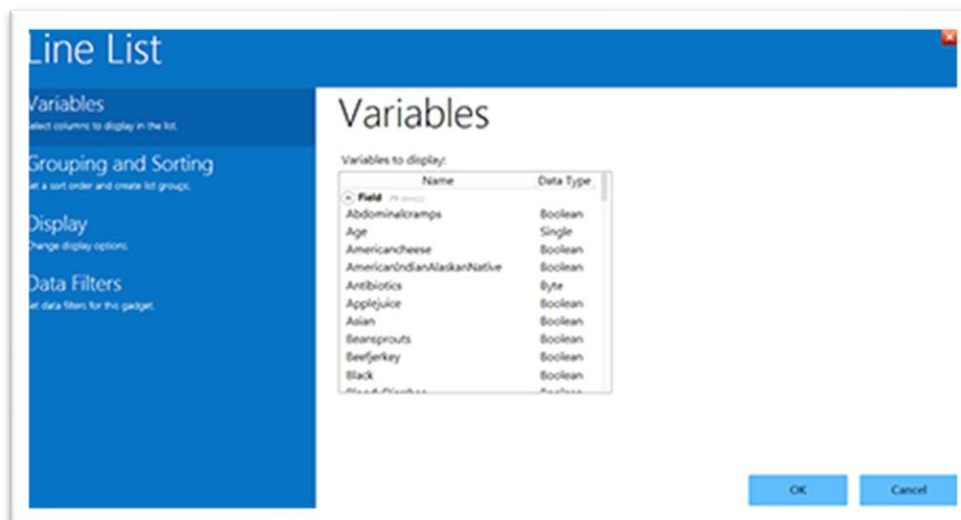
Line List gadget with options

3. Select multiple variables from **Variables to display**. Hold down
    the **Ctrl** key and click on the variables you want to display. Use **Ctrl + A** to
    select all variables at once.

4. Click **OK**. The line list appears on the canvas. The graphic below shows the
    list with all variables selected.

Generated Line List of all variables

Note: View the 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

Arrange file records in a particular order, such as alphabetically by name, or by age. Sort records by clicking on the column name in the line list. You can also go to the **Sorting and Grouping** property panel. Use **Sort options** in the Line List configuration window.
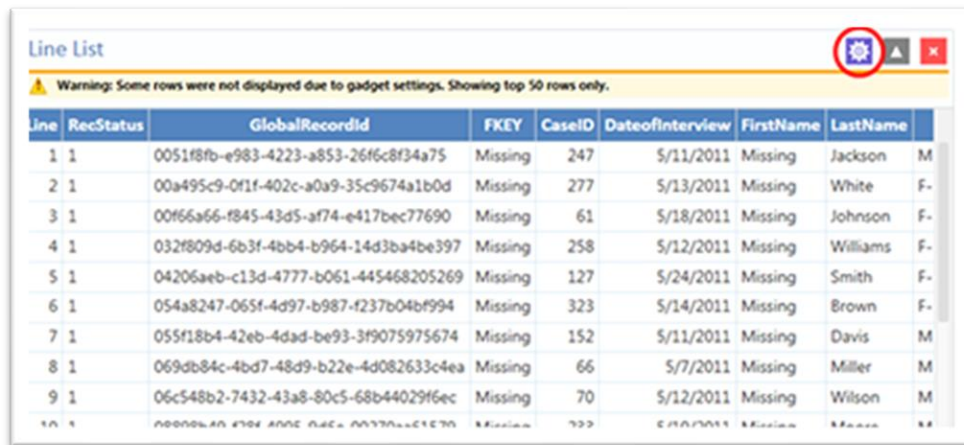
Sorting a List by Column Name

Follow the steps below to sort the list by Age.

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**. The Line List gadget configuration window opens to the **Variables property panel.**
3. 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.
4. 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 show records will sort by the **Age** variable.
5. Click the **Age** column header again to sort in descending order. The **Age** column header turns green to show it is sorted in descending order.

Sorting a List Using Line List Grouping and Sorting Properties

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**. The Line List gadget configuration window opens to the **Variables property panel.**
3. Select all variables from **List Variables to Display,** by pressing **Ctrl + A** on your keyboard and Click **OK**. The **Line List** results appear on the Visual Dashboard canvas.
4. Click on the **gadget gear icon** at the top of Line List to sort variables.

Line List with Gadget Configuration Icon Noted

6. Select **Grouping and Sorting** in the configuration window and then select Sorting Method of Ascending or Descending.

The following steps require a mouse double-click action.

Double-click:

- a variable from the **Available Variables** list.
- on each variable will arrange multiple variables.
- the selected variable movies it to the **Sort Order** box.

- on a variable in the **Sort Order** box to remove the variable from the list. If you want to group the list by a particular field, select the field from the **Group By** drop-down list. Click **OK.**
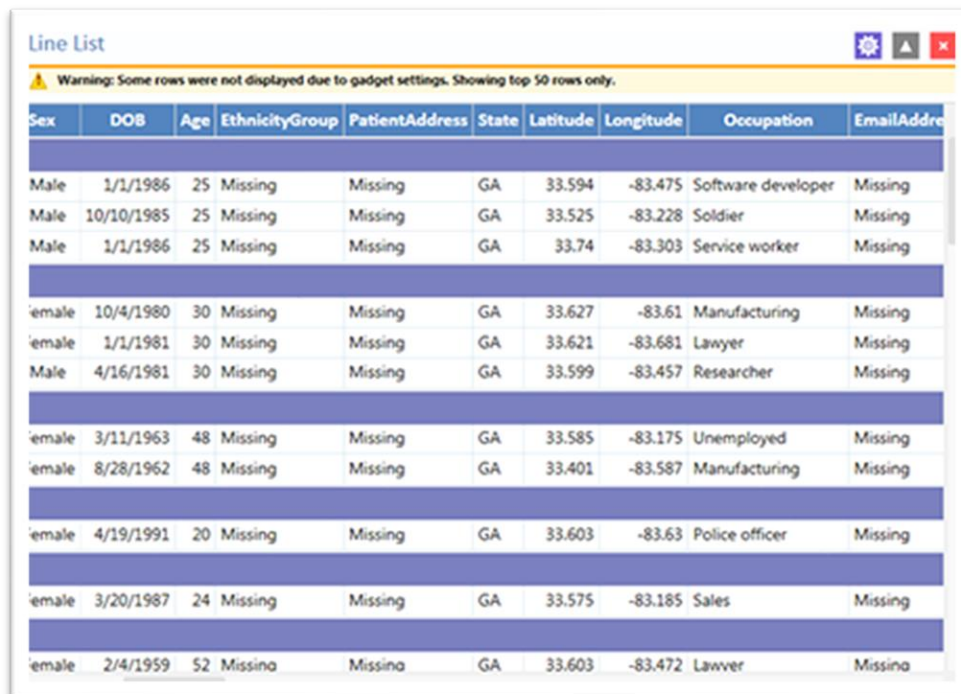
Line List Grouping and Sorting

Visual Dashboard generates and displays a separate line list, grouped by the variable you select. You can hide a list from view by clicking on the arrow to the left of the list heading.

Line List with Grouping results

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**.

The Line List will appear with the sort criteria removed.

Other Line List Properties
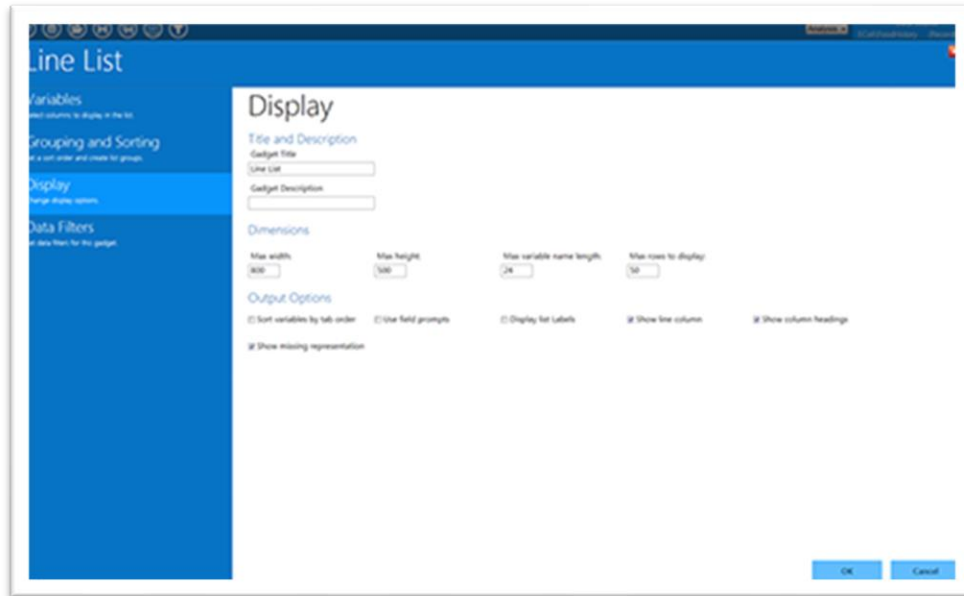
Line List Display options

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

- **Max width** — indicates the maximum Line List display width in pixels.
- **Max height** — indicates the maximum Line List display height in pixels.
- **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.
- **Max rows to display** — text box allows the user to change the maximum number of rows that the Line List gadget displays. It is set to 50 by default. Display up to 2,000 rows.

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

- **Sort variables by tab order** — The checkbox forces the columns in the output to be sorted by their tab order. It is not checked by default.

- **Use field prompts** — The checkbox uses the field prompt as the column heading, rather than the field name. It is not checked by default.
- **Display list labels** — The checkbox uses the label value for option fields and comment legal fields rather than the 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 from something else like an Excel spreadsheet, these options have no effect.

- **Show line column** — option that adds the Line column to the list and displays the line number when the checkbox is selected.
- **Show column headings** — option that shows the line list column headings when the checkbox is selected.
- **Show missing representation** — option that shows 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, an empty field shows a blank space.

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 the list data to the clipboard.

1. Right-click on the **Line List**.
2. Select **Copy List Data to Clipboard**.
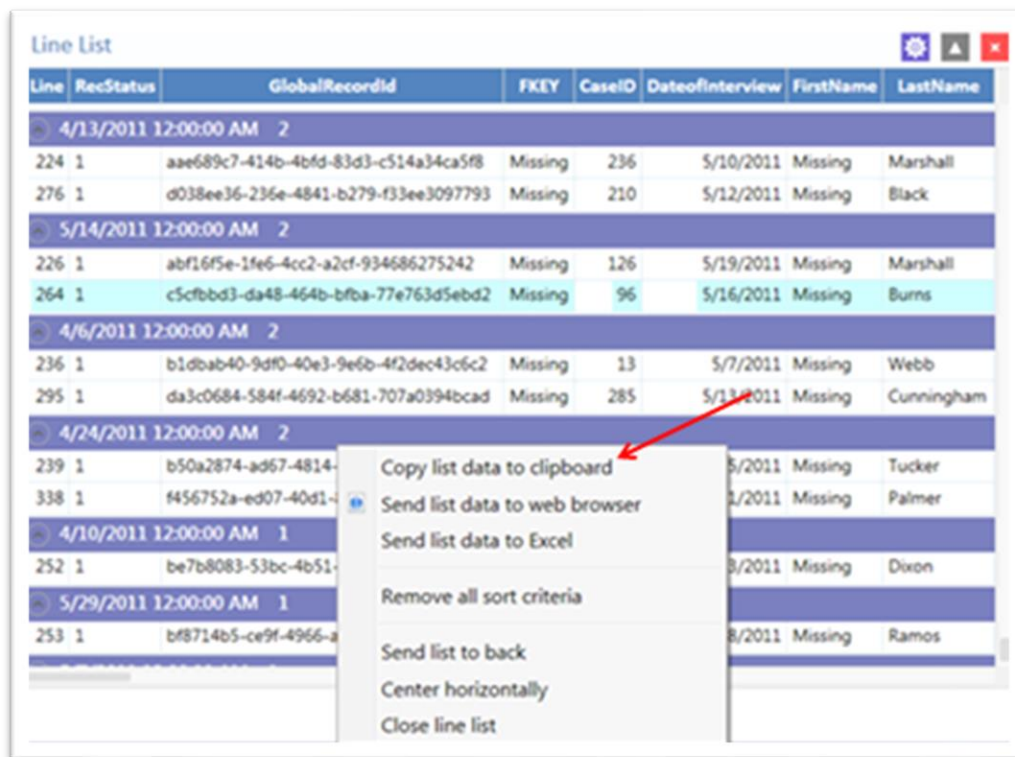
Copy list data to clipboard

The data is 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**.
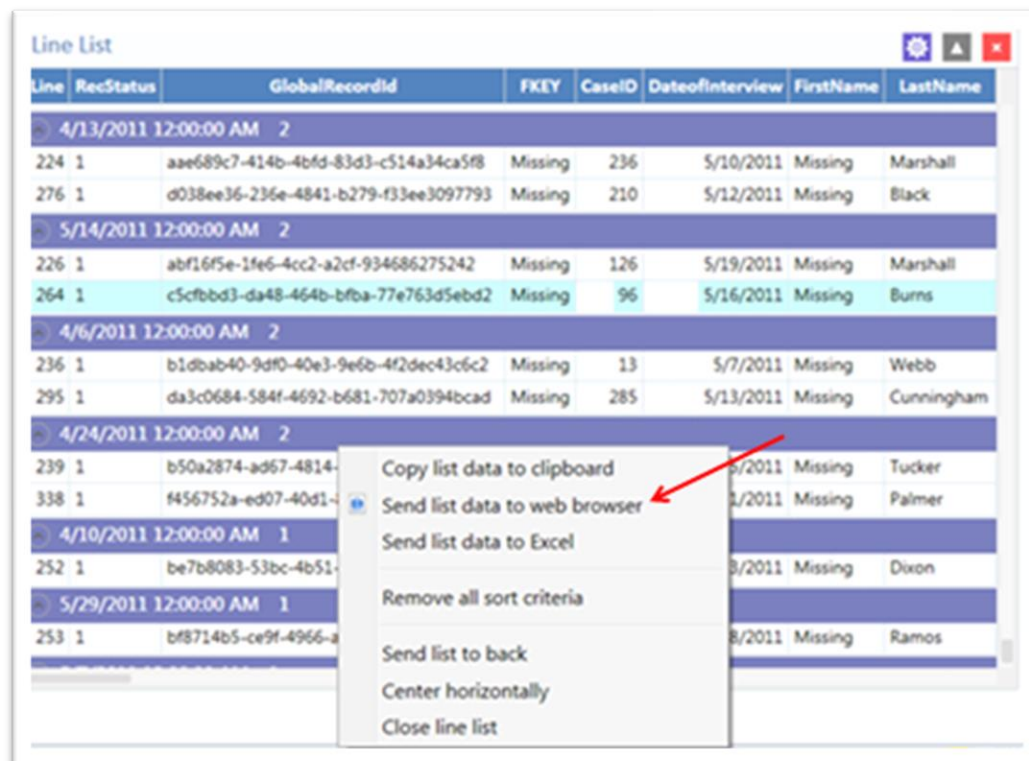
Send list data to web browser

The data appears in your default web browser.

List data in web browser

The next three steps explain how to send list data to Microsoft Excel.

1. Right-click on the **Line List**.

2. Select **Send List Data to Excel**.

Send list data to Excel

The list data appear in a Microsoft Excel spreadsheet.

List data in Excel

Frequency

The Frequency gadget counts each occurrence of a value for a specified variable and gives the absolute and relative frequencies for each value. The resulting output is a frequency table showing how many records have values 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.**
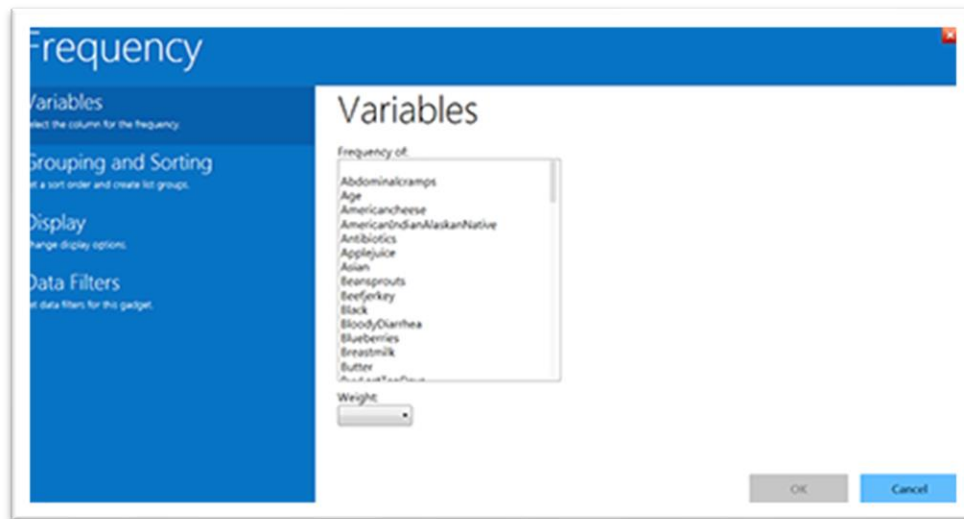
<div align="center">Frequency gadget</div>

3. Select a **variable** from the data table in the **Frequency of** list. For this example, select **sex** as the desired variable.

Selecting frequency of variable

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

Sex	Frequency	Percent	Cum. Percent	Wilson 95% LCL	Wilson 95% UCL	
F-Female	186	51.81 %	51.81 %	46.65 %	56.93 %	
M-Male	173	48.19 %	100.00 %	43.07 %	53.35 %	
TOTAL	359	100.00 %	100.00 %			

Frequency results

The Frequency column counts individuals that are either female or male. The Percent column shows 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:

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**.**
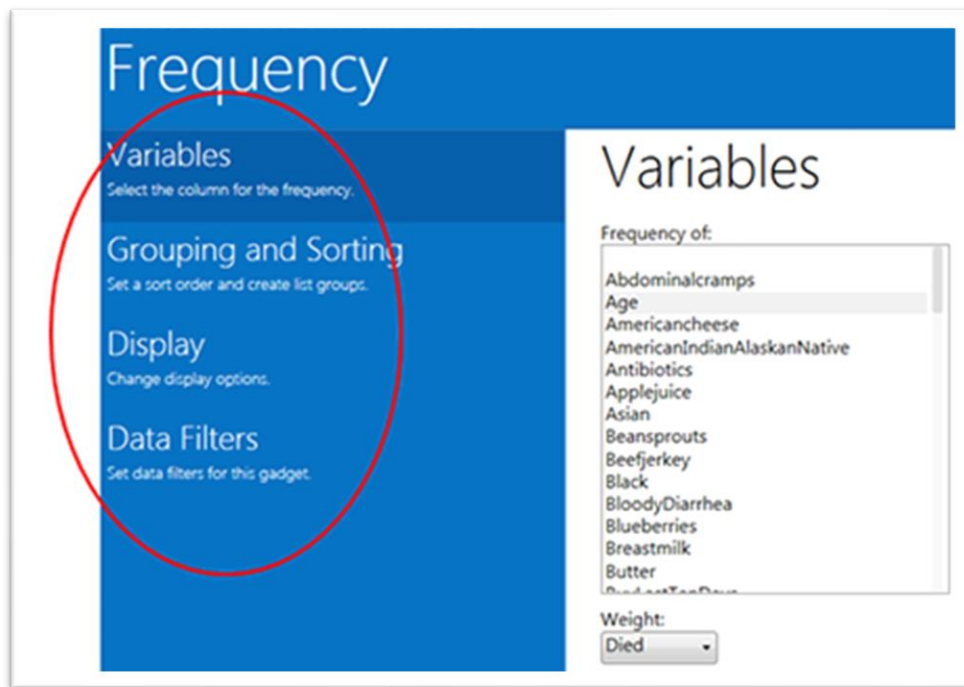
Frequency gadget functions menu

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

      o **Weight** (available under **Variables**) – field for the weight variable. (e.g., aggregate data analysis.)

- o **Stratify By** (available under **Grouping and Sorting**) – Use this field for stratifying (grouping) the output.
- o **Sort high to low** (available under **Grouping and Sorting**) – Displays the results sorted from high to low based on the selected frequency of variables.
- o **Display all list values** (available under **Display**) – Applies when the frequency variable is a drop-down list field in an Epi Info™ 7 project. When selected, the output will show all the drop-down list values even if they have a count of zero.
- o **Display list labels** (available under **Display**) – Applies 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 label for **Comment Legal fields** and the **option button label** for **Option fields**.
- o **Include Missing Values** (available under **Display**) – Displays the count of rows having missing values. By default, it is unchecked.

3. Select **Age** from the data table, from the **Frequency of** list in **Variables**.
4. Select **Died** from the data table, from the **Weight** drop-down list below.
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% LCL	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 %	

Frequency line list with variables

Was the patient hospitalized? = No

Age	Frequency	Percent	Cum. Percent	Exact 95% LCL	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 %	

Frequency Line List stratified view #2

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

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

- **Gadget Title** – Enter a title for the current gadget. This title will appear at the top of the gadget results on the canvas.

68

- **Gadget Description** – 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** – Use the field question or prompt instead of the field name. This property is checked by default.
- **Draw borders** – Use this setting for drawing frequency border.
- **Draw header row** – Control the appearance of frequency headers.
- **Draw total row** – Manipulate the frequency totals.
- **Decimals to display** – Set 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** – Choose the number of rows to display.
- **Max width of percent bar** – Select the width for the percent bar column using a % value.

Word Cloud

The Word Cloud gadget shows a visual snapshot of frequently used terms in the current data set. Word Cloud returns a graphic of words and terms, with font sizes illustrating the number of times a word or term appears, in comparison to others. This gadget is most valuable for analyzing text fields and multi-line text fields. For example, it can condense a field with patient complaints from verbose to commonly 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 such as: 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, their, this, to, us, was, we, were, will, with, would, you, and, your, from the analysis.

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**.
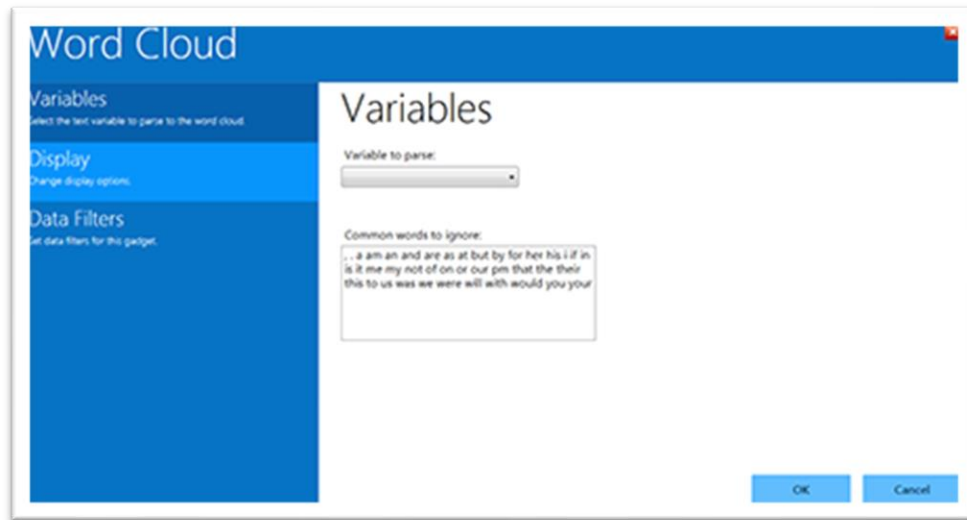
Word Cloud gadget

3. Choose the desired variable on which to parse words from the drop-down list. For this example, select **Race**. **Common words to ignore** 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 others in the group. In this example, the most common color in the selection is "White," followed by "Blue."

Word Cloud results

Combined Frequency

The Combined Frequency gadget creates a table that shows the frequency and percentage of affirmative responses to the 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 several check box fields for **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.

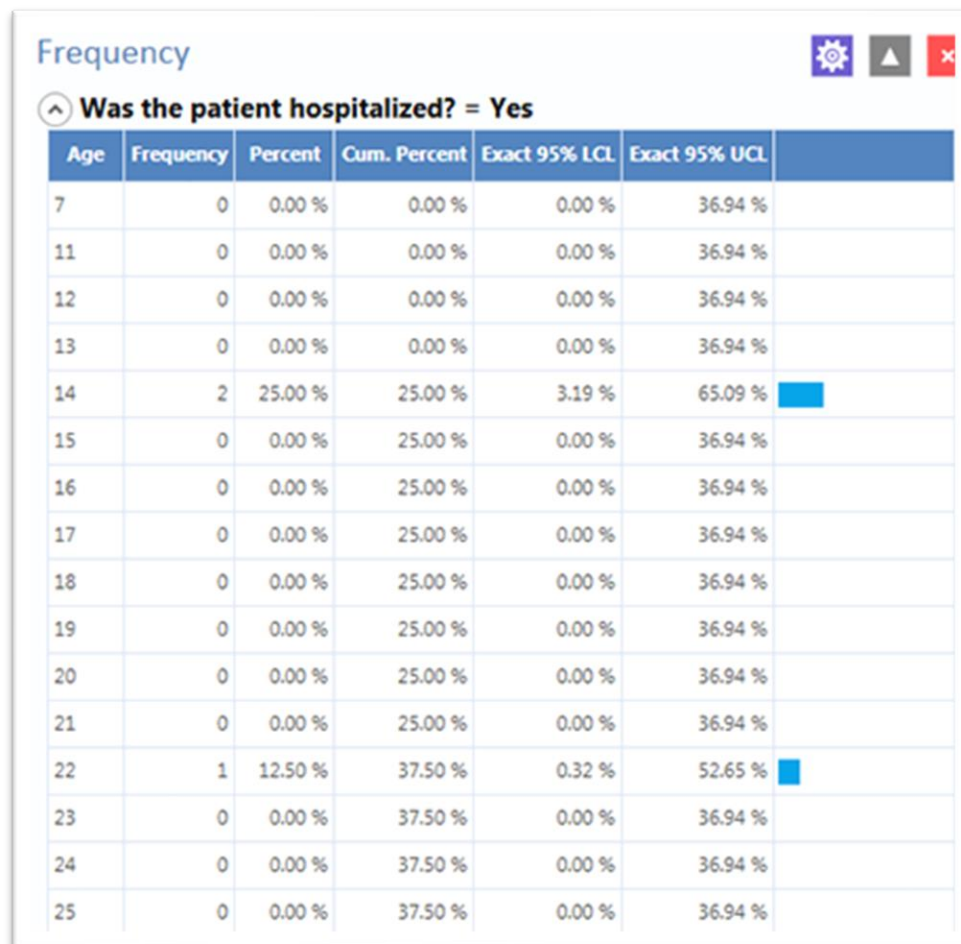2. Right-click on the canvas and select **Add Analysis Gadget > Combined Frequency**. The Combined Frequency gadget configuration window opens to the **Variables** property panel. The **Variables to display** list only shows Group variables.
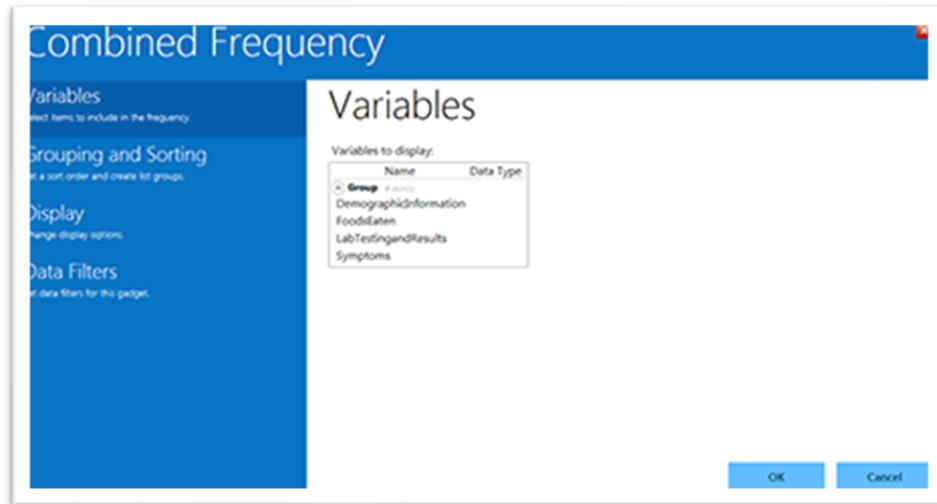
Combined Frequency gadget

3. Select **Symptoms** from the Variables to Display list under Group,
4. Click **OK**. The 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

Combined Frequency results

Combined Frequency Advanced Options

Using 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. Select **DemographicInformation** in the **Variables** options from
   the **Group** field.

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 is 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.

The **Sort High to Low** checkbox is checked by default. This option displays the results sorted from highest group field frequency to the lowest.

5. 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.

6. Click **OK**. The Combined Frequency 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

Combined Frequency results with advanced options

The result displays the frequency and percentage of each symptom selected check box and each Yes-No field with a value of "Yes". Note the denominator included at the bottom of the results window.

M x M/ 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 **Cross Tabulation 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.
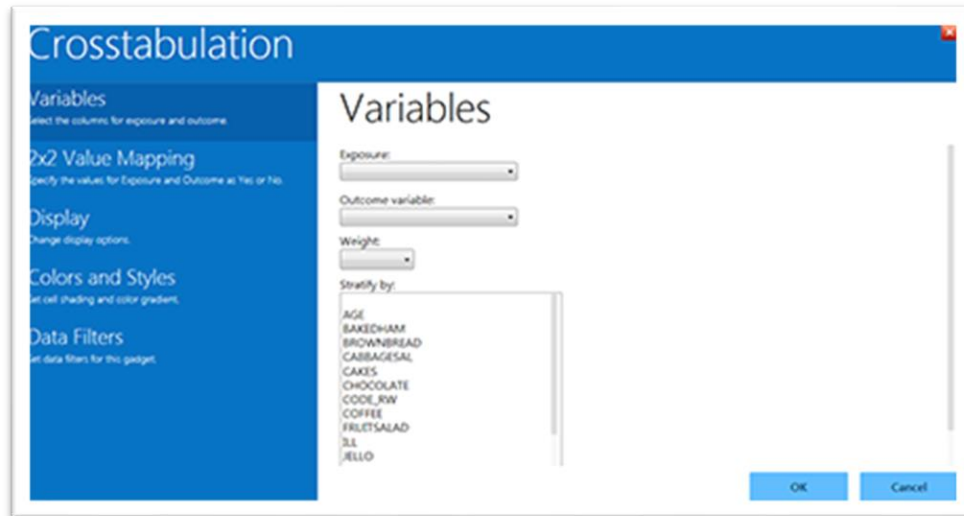
Cross tabulation gadget

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

Cross tabulation results

The results display a 2 x 2 table and a single table analysis of all records cross tabulated by exposure to vanilla where the outcome was ill. Out of the 75 records in this data set, 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.

Matched Pair Case-Control

The Matched pair case-control 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 2×2 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 Pair 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**. The Matched Pair Case-Control Properties gadget appears.
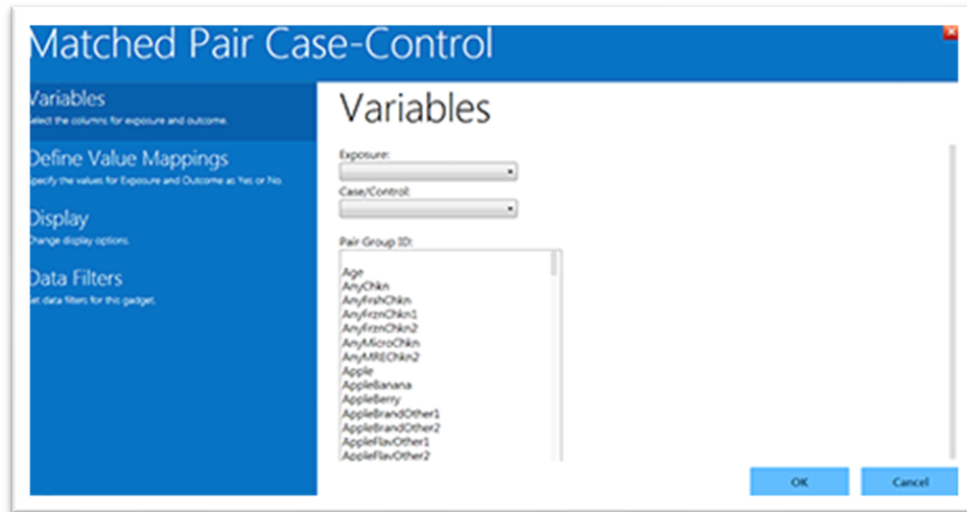
Matched Pair Case-Control gadget

Select:

- **AnyChkn** from the **Exposure** drop-down list.
- **CaCo**, from the **Case/Control** drop-down list,
- **Matched pairs**, from the Pair Group ID list.

In the following order, click on:

- **Define Value Mappings**.
- **1**.
- the **>** button on the **Yes values:** section.
- **0**.
- the **>** button on the **No values:** section.
- **OK**.

The Matched Pair Case-Control results displays on the canvas.
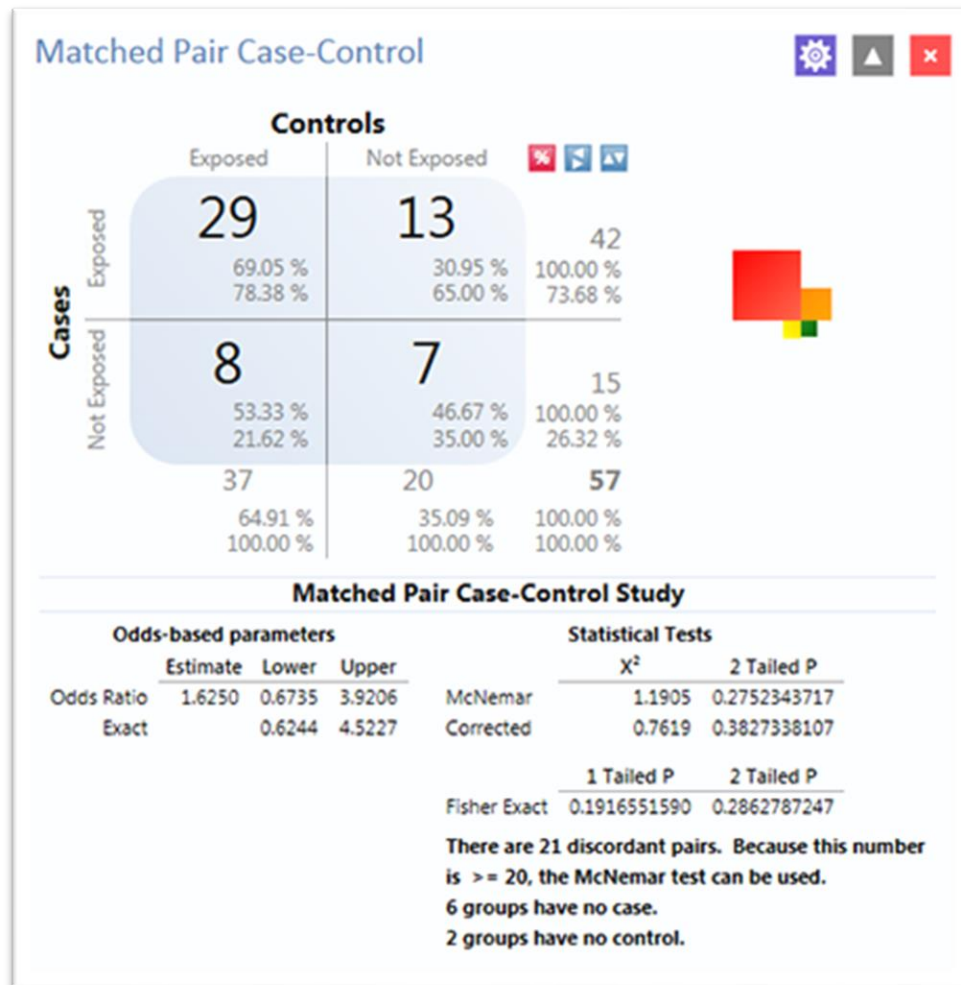
Matched Pair Case-Control results

The results show that out of 57 matched pairs, 29 had both the Case and the Controls exposed to **AnyChkn**, 13 had the Case but not the Controls exposed, eight pairs had the Controls exposed but the Case unexposed, and seven had neither the Case nor the Controls exposed.

The Odds-Ratio is greater than 1, which suggests that eating chicken may more likely result in illness. 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.

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 like 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 like 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.
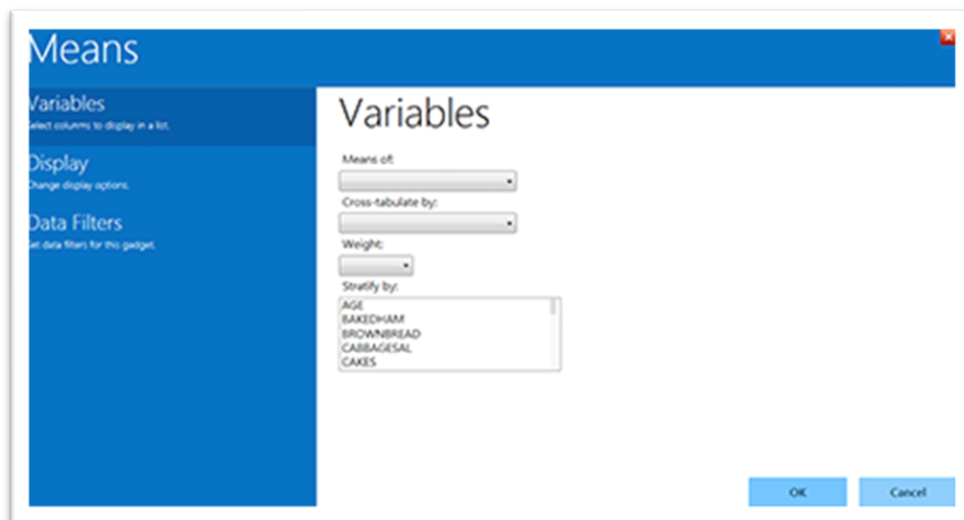
From the **Means Of** drop-down list, select **Age**.

4. Other choices that can be used in the **Variables** option include:

   o The **Cross-tabulate by** drop-down list contains a variable to help determine if the means of a group are equal.

   o The **Weight** drop-down list contains a variable for weighted analysis.

   o The **Stratify by** list contains a variable to act as a grouping variable.

Note: These options are not used in this example.

5. 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 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 using a cross-tabulate variable. The descriptive statistics always show 4 decimal places.

   - The **Display ANOVA statistics** checkbox specifies whether to display the analysis of variance statistics, set by default. If you are working with large data sets without ANOVA statistics, you can improve performance by turning off (un-checking) this feature.

   The mean age of individuals in the data set is 36.8.

Means	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

Means Output

Duplicates List

The Duplicates List gadget lists potentially duplicate records. Like a patient registry, everyone must have only one record. The Duplicates List gadget allows you to select a few fields that should be unique and display a few other non-unique fields too. 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, let's 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 view additional fields, such as **CaseID**, **DateofInterview**, and **OnsetDate**. Then we can make a determination that an individual is actually duplicated in the registry.

Follow these steps:
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.
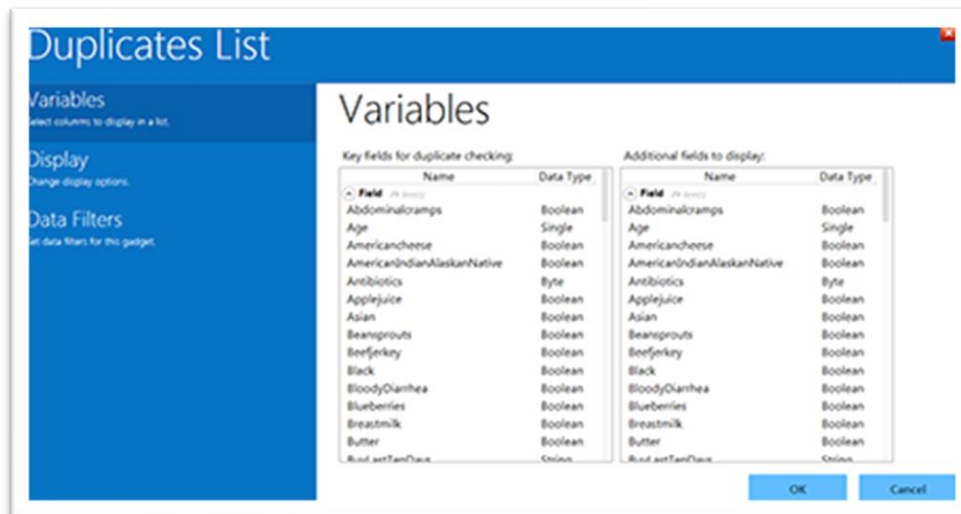
Duplicates list gadget

1. Hold down the **CTRL key** and click on the respective field names.
2. Select DOB and LastName in the Key fields for duplicate checking list.
3. Select other fields like

   CaseID, DateofInterview, DoctorVisit, DoctorVisitDate, and OnsetDate, in the **Additional fields to display** list. This will help you determine if the record is a true duplicate.
4. Hold down the **CTRL key** and click of the **field names**.
5. Click **OK**.

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 results

Analysis Gadget Charts

- Column Chart
- Line Chart
- Area Chart
- Pie Chart

- Aberration Detection
- Pareto Chart
- Scatter Chart
- Epi Curve

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 generates 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. A slice of the circle is 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 you can create each 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. Select the variable **Age** from the **Main variable** drop-down list, to use for the independent variable.
4. Maintain the **Weight** and select the **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.
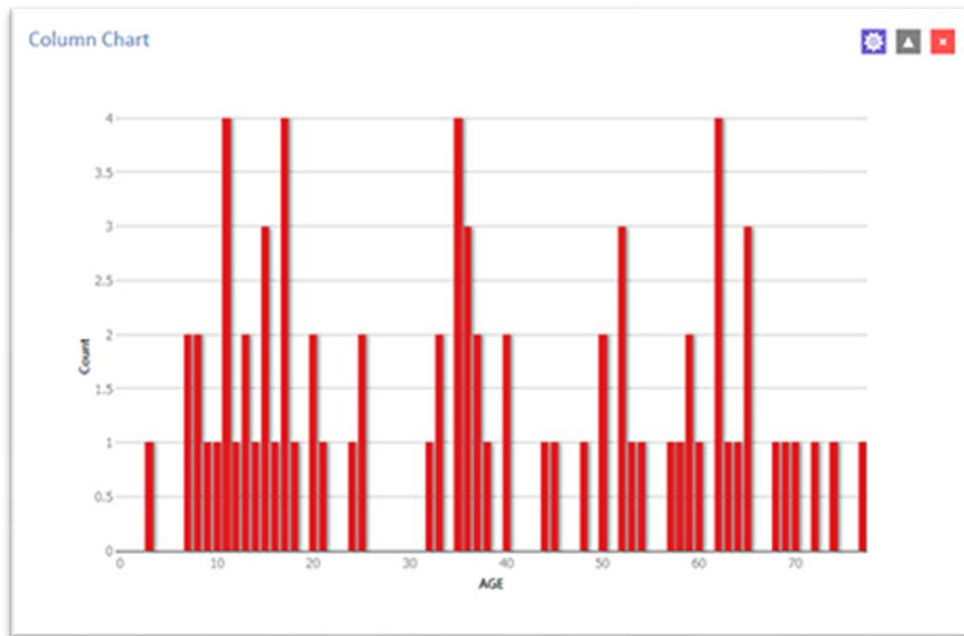
Column chart

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.

The system returns a line chart of onset dates between April and May from the dataset.
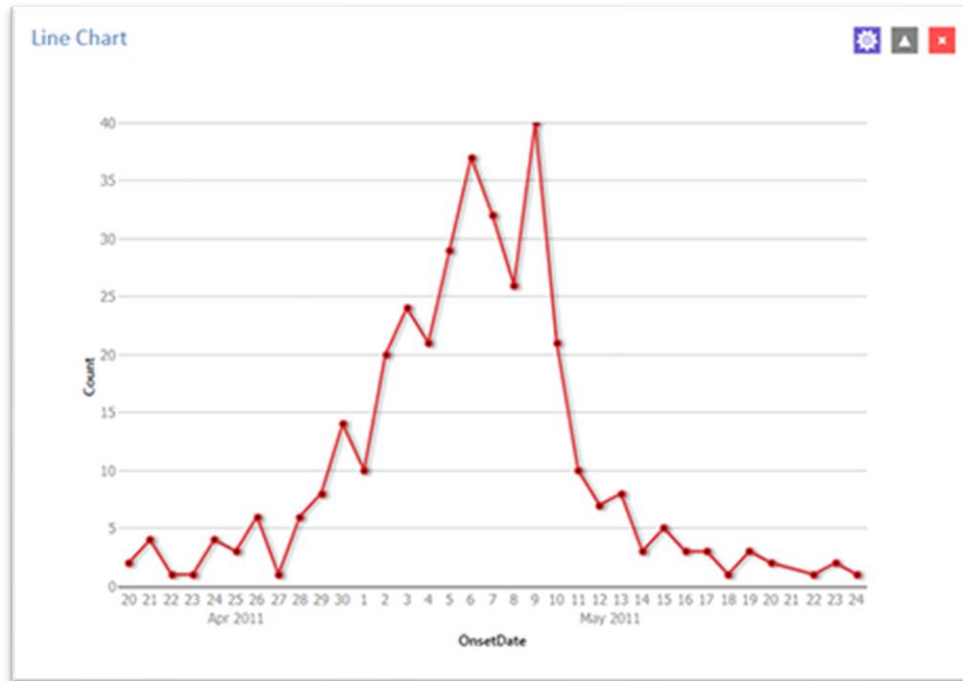
Line chart

Area Chart

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

- Open the **EColi.prj** project and select the **Food History** form.
- 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.
- Select **OnsetDate** from the **Main variables** list.

The system returns a line chart of onset dates between April and May from the dataset.
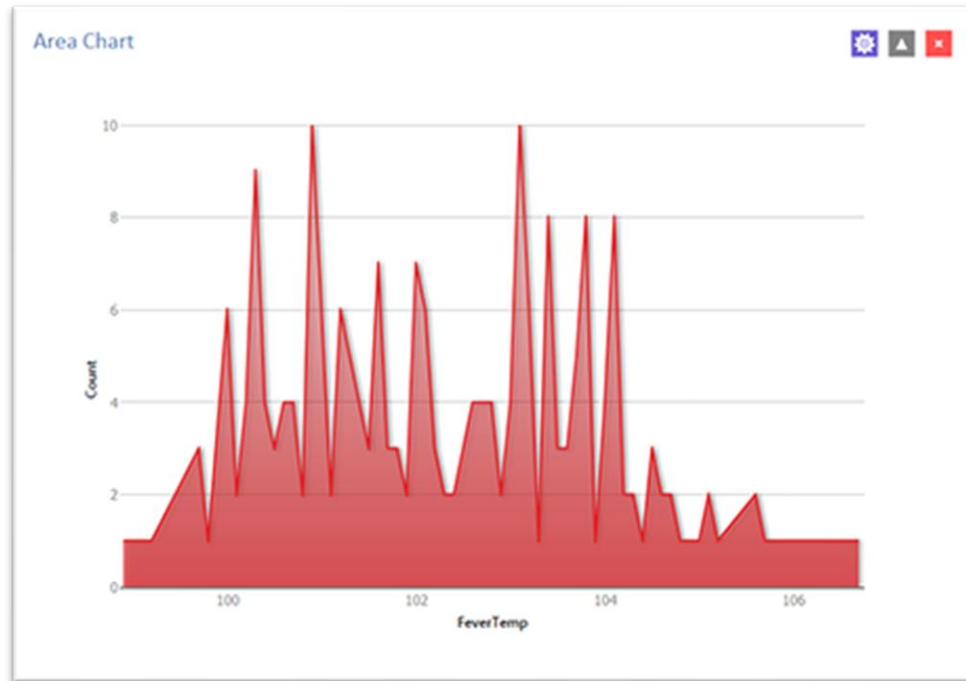
Area chart

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 displays.

- 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. Select the variable **AGE_RECODED** from the **Main variable** drop-down list.

4. Maintain the **Weight** and **One graph for each value of**, option at the default setting.

5. Click **OK**. The graph appears on the canvas.

The chart shows the age groups of church supper attendees.
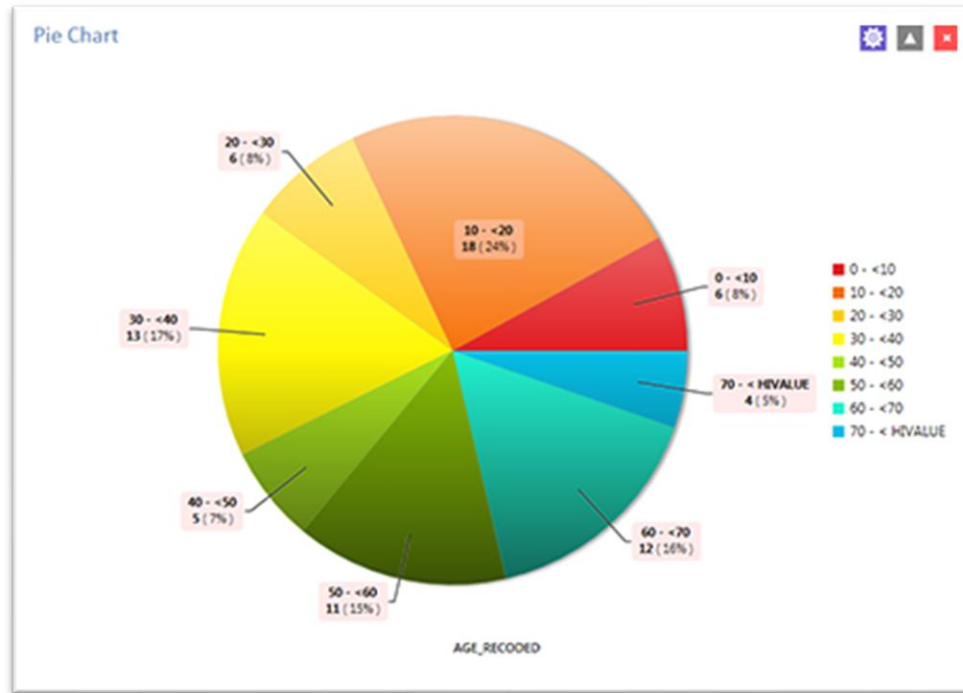
Pie chart

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 the CDC Division of Bioterrorism Preparedness and Response. The division assists local-level health offices with early detection of bioterrorism events or possible outbreaks.

The following example shows an Excel spreadsheet named **EARS.xls** that comes with Epi Info™ 7. Find it in the **ProjectsSample** project folder.

1. Right-click on the canvas and choose **Set data source** or click the **data source icon** in the toolbar.
2. Click the **Database Type** drop-down to select Microsoft Excel 97-2003.
3. Click **Browse** and then **Browse** again to search for the EARS.xls file.
4. Browse to the Epi Info™ 7 **ProjectsSample folder** and select the **EARS.xls** file.
5. Click on **Sheet 1$**, then click **OK**.
6. Right-click on the canvas and select **Add Analysis Gadget**. Scroll to highlight **Charts**, then select **Aberration Detection**.
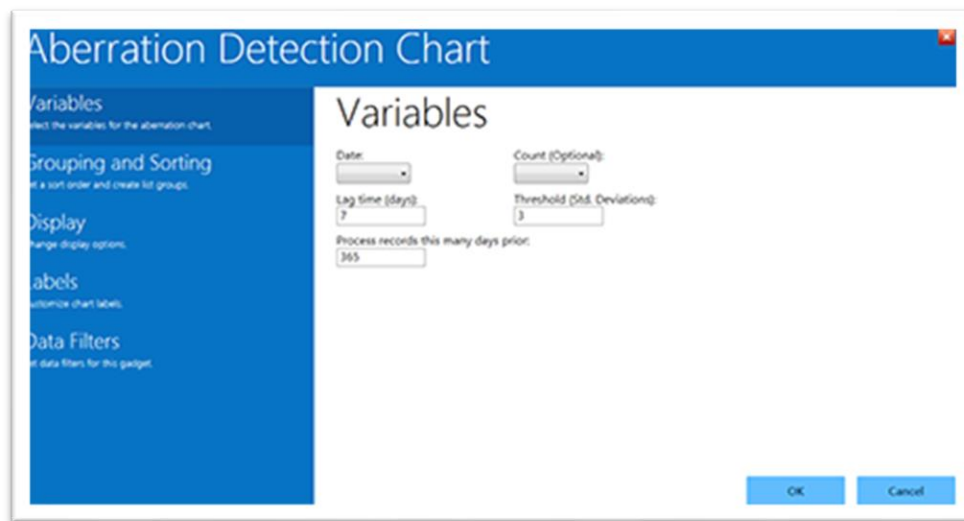
The **Aberration Detection** gadget configuration window opens to the Variables property panel.

Begin by Selecting:

- the **DATE variable f**rom the **Date** drop-down list,
- **COUNT**, from the **Count (Optional)** drop-down list.

Maintain the default settings for Lag Time, Threshold, and Process Records This Many Days Prior.

Continue by clicking:

- the Grouping and Sorting option for the property panel.
- the variable **SYNDROME**, from the **Indicators list,**
- **OK**. The Aberration Detection line graph displays on the canvas.
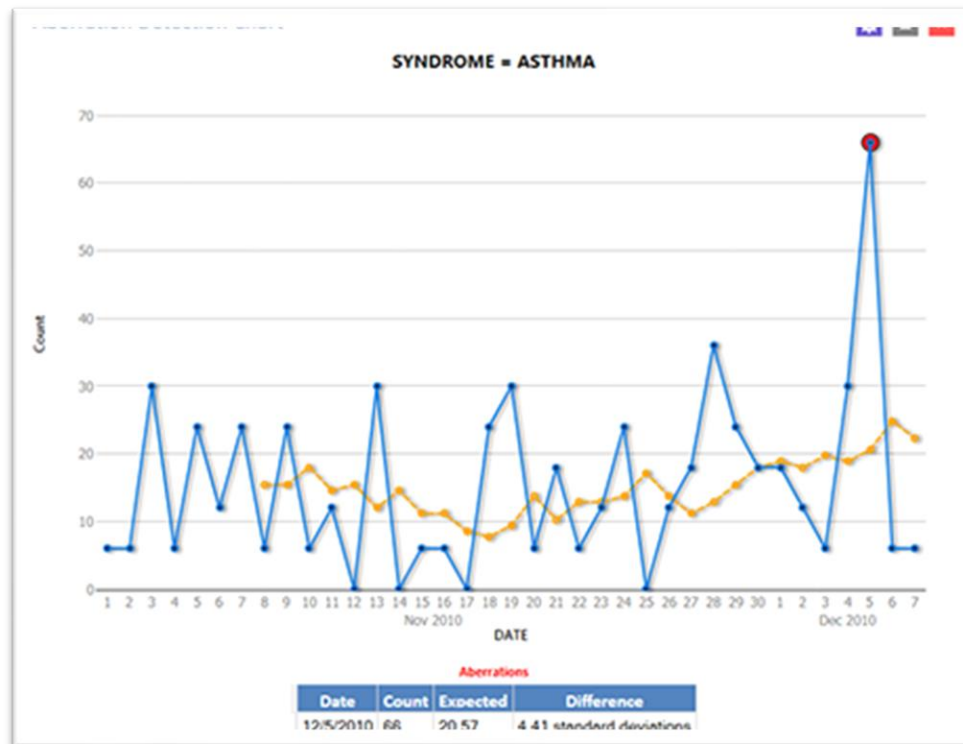
Aberration Detection line graph

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.

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.

Begin by recoding the age variable. Open the **Ecoli.prj** project and select the **FoodHistory** form. You can recode the Age variable using the **Defined Variables** gadget.

1. Move the mouse cursor over the **Defined Variables** gadget on the left-hand side of the Visual Dashboard canvas. The gadget expands and becomes fully visible.
2. Click the **New Variable** button.
3. Select **With Recoded Value**. The Add Recoded Variable dialog box appears.
4. Select **Age** from the **Source** field drop-down list. Notice the Destination field populates with **Age_RECODED**.
5. Click the **Fill Ranges** button. The Fill Ranges dialog box appears.
6. 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. Click **OK** again.

Build your Pareto Chart. 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.

1. Select the new variable, **Age_RECODED**, from the **Main variable** drop-down list.

2. Click the **Labels** option to customize chart labels and axes.

3. Update the **X-axis angle** setting to -90. This step will permit the date values to display fully on the X-axis.

4. Click the **Display** option.

5. 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.

6. Click **OK**. The Pareto chart appears on the canvas.
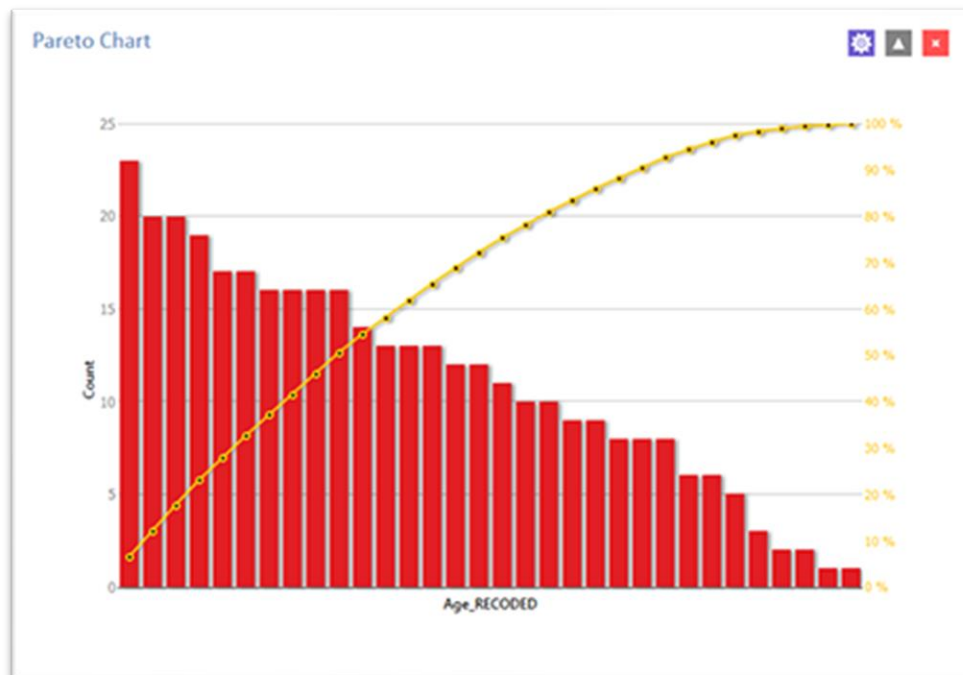
Pareto chart

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.
3. Select the variable **Estriol** from the **Main variable** drop-down list. Use it for
   the independent variable.
4. Select the variable **Birthweight** to use for the independent variable, from
   the **Outcome variable** drop-down list,
5. 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.

Y = (0.6082)X + 21.5234
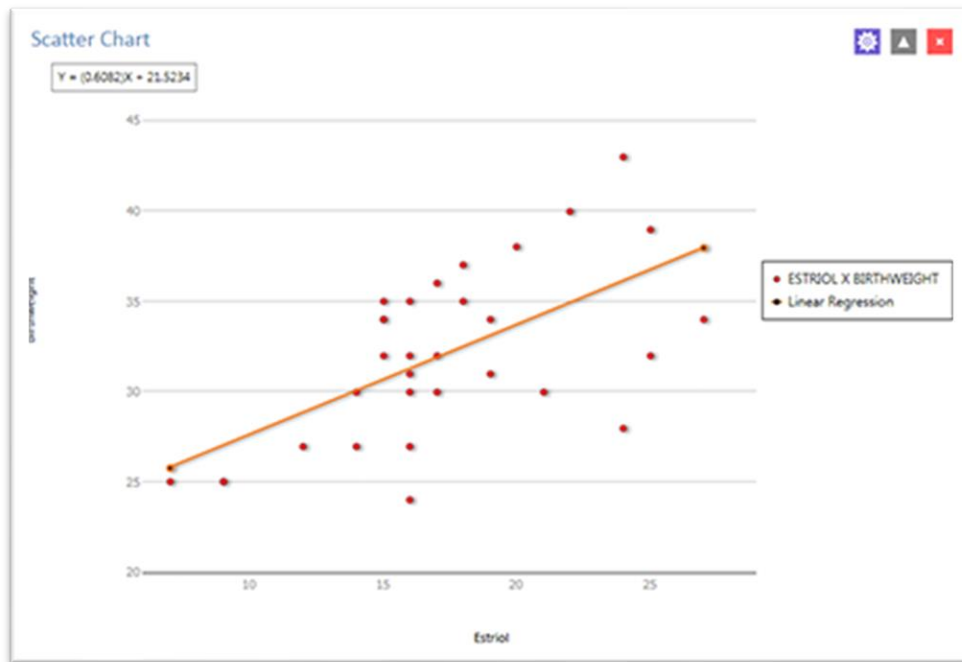
Scatter chart

Epi Curve Chart

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

1. Open the **Ecoli.prj** project and select the **FoodHistory** form.
2. 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.
3. Select **OnsetDate** from the **Main variable** drop-down list,
4. Click the **Labels** option to customize chart labels and axes.
5. Update the **X-axis angle** setting to -90. This step will permit the date values to display fully on the X-axis.

Control how the Epi Curve Chart displays on the screen.

1. Click the **Display** button. The Display property panel appears.
2. 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.
3. Click the **Legend** button. The Legend property panel appears.
4. Click the checkmark to select **Show Legend**.
5. Click the **Grouping and Sorting** button. The Grouping and Sorting property panels appears.
6. In the **Stratify by** list, scroll down, and select the variable Sex.
7. 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.
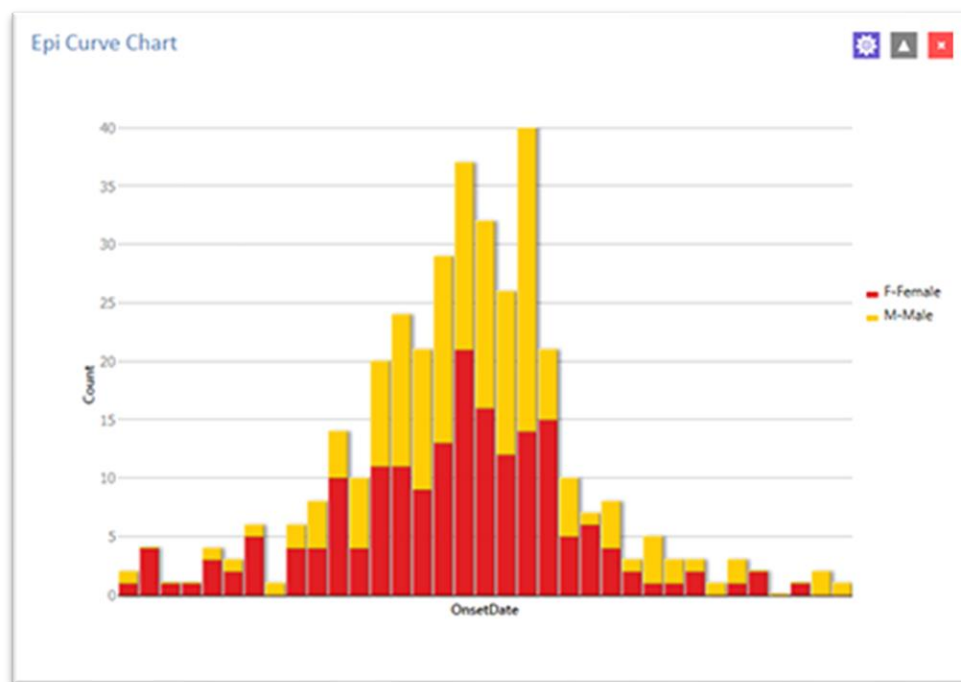
Epi Curve chart

Advanced Statistics

- [Logistic Regression](#)
- [Logistic Regression Properties](#)

- [Complex Sample FTMs](#)
- [Complex Sample Frequencies](#)
- [Complex Sample Means](#)
- [Complex Sample Tables](#)

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 infant 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. Select **SystolicBlood** from the **Outcome Variable** drop-down list,

5. Select **AgeInDays** and **Birthweight** from the Independent variables drop-down list.

6. Click **OK**. The linear regression results appear on the Visual Dashboard canvas.

Linear Regression

Variable	Coefficient	95% Confidence	Limits	Std Error	F-test	P-value
AgeInDays	5.888	4.418	7.357	0.680	74.9229	0.000001
Birthweight	0.126	0.051	0.200	0.034	13.3770	0.002896
CONSTANT	53.450	43.660	63.241	4.532	139.1042	0.000000

Correlation Coefficient: $r^2 = 0.88$

Source	df	Sum of Squares	Mean Square	F-statistic	p-value
Regression	2	591.0356	295.5178	48.0806	0.0000
Residuals	13	79.9019	6.1463		
Total	15	670.9375			

Linear Regression results

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

Linear Regression Properties

Linear Regression gadget

Linear Regression Gadget settings and elements:

- The **Outcome variable** — the dependent variable for the regression. The outcome variable must be numeric or a Yes/No field.

- **Weight** variable — use in weighted analyses.

- **Confidence Limits** option — specifies the probability level at which Epi Info™ computes confidence limits.

- **No Intercept box** — the regression is performed without a constant term, which forces the regression line through the origin.

- **Fields** drop-down list — contains predictor (independent) variables.

- **Predictor variables** — appear in the list box below the **Fields** drop-down. Double-clicking on a variable removes it from the list box.

- **Make Dummy** button — If you click **Make Dummy**, the variable appears in the **Dummy Variable** list box. Double-click on a variable to remove it from the list box.

- **Make Interaction button** — defines Interaction Terms.

Note: **Make Dummy** button turns to the **Make Interaction** button 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 combinations of the selected variables to the regression as interaction terms.

Logistic Regression

In Epi Info™ 7, either the **M x N/ 2 x 2** table or **Logistic Regression** is usable when outcome variables are 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 variable 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 brown bread, 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. Select **ILL** from the **Outcome** drop-down list.
5. Select BROWNBREAD, CABBAGESAL, WATER, MILK, CHOCOLATE, and VANILLA, from the Independent variables drop-down list.
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
WATER (Yes/No)	1.1120	0.2670	4.6319	0.1062	0.7280	0.1458	0.8841
MILK (Yes/No)	0.1343	0.0068	2.6659	-2.0077	1.5246	-1.3168	0.1879
CHOCOLATE (Yes/No)	1.0974	0.3024	3.9827	0.0930	0.6577	0.1414	0.8876
VANILLA (Yes/No)	25.9752	5.4652	123.4566	3.2571	0.7953	4.0955	0.0000
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

Logistic Regression gadget

- **Outcome variable** — is the dependent variable for the regression. The
  outcome variable must be numeric or a Yes/No field.

- **Weight variable** — may be selected to use in weighted analyses.

- **Match variable** — identifies the variable indicating the group membership
  of each record.

- **Confidence limits** — option specifies the probability level at which Epi
  Info™ computes confidence limits.

- **No intercept** box (checked) — the regression is performed without a
  constant term.

- **Include missing values** — 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.

104

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.

Note: The **Independent variables** drop-down list contains the predictor (independent) variables.

Note: 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. Doubl-click on a variable to remove it from the **Dummy variables** list box.

- **Make interaction button** — defines Interaction terms list box.
- **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 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.
- **Clear terms button** — clears all variables from the **Dummy variables** and **Interaction terms** list boxes.

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 2×2 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**.

The **Complex Sample Frequencies** configuration window opens to the Variables property panel.
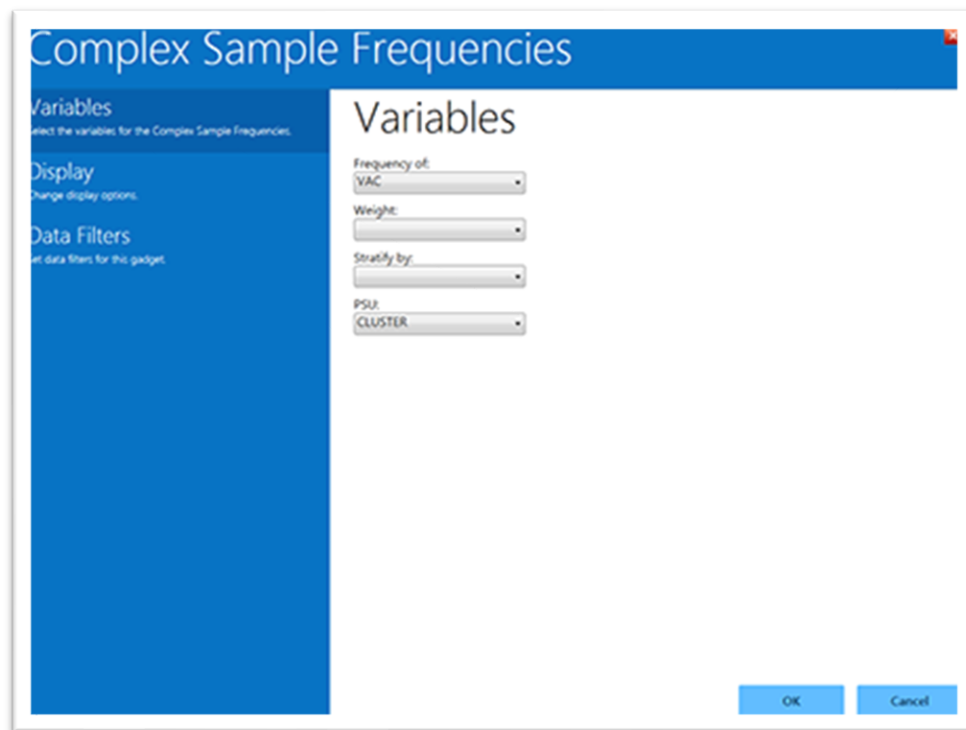
Complex Sample Frequencies gadget

Complex Sample Frequencies options are:

- **Frequency of** — identifies the variable(s) whose frequency is computed.

- **Weight** variable — 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.

4. Select **VAC (the number of vaccinations received)** from the **Frequency Of** drop-down list. This will calculate the frequency of vaccinations in the data set.

5. Select CLUSTER (the population group for each case) from the **PSU drop-down list**.

6. Click **OK**.

The **Complex Sample Frequencies** appear on the canvas.

Complex Sample Frequency	⚙ ▲ ✕

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

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 considers 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 like 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 consider 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**.

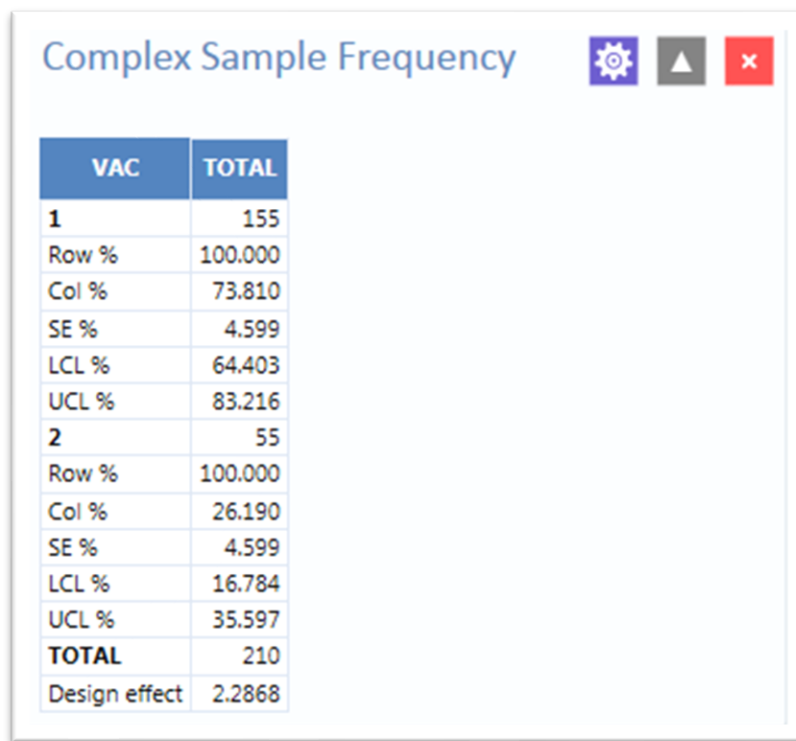The **Complex Sample Frequencies** configuration window opens to the Variables property panel.

1. Select **VAC** from the **Frequency Of** drop-down list.
2. Select **POPW** (the percent of the total population in each cluster) from the **Weight** drop-down list.
3. Select **LOCATION** from the **Stratify By** drop-down list.

4. Select **CLUSTER** from the **PSU** drop-down list.

5. Click **OK**.

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

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%.

Complex Sample Means

The Complex Sample Means command can be used when the outcome variable is continuous, such as age, cholesterol level, and so on. 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

110

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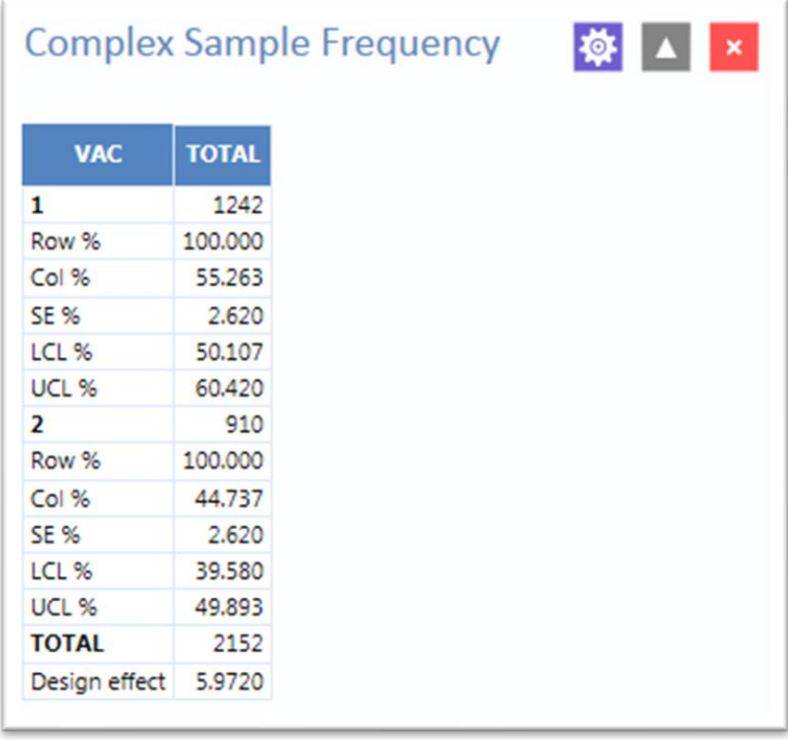
1. Set the Data Source to the **Sample.prj** project.
2. Select the **Smoke form** from the **Data Source Explorer** menu.
3. Right-click on the canvas and select **Add Analysis gadget > Advanced statistics > Complex sample means**.
4. The Complex Sample Means configuration window opens to the Variables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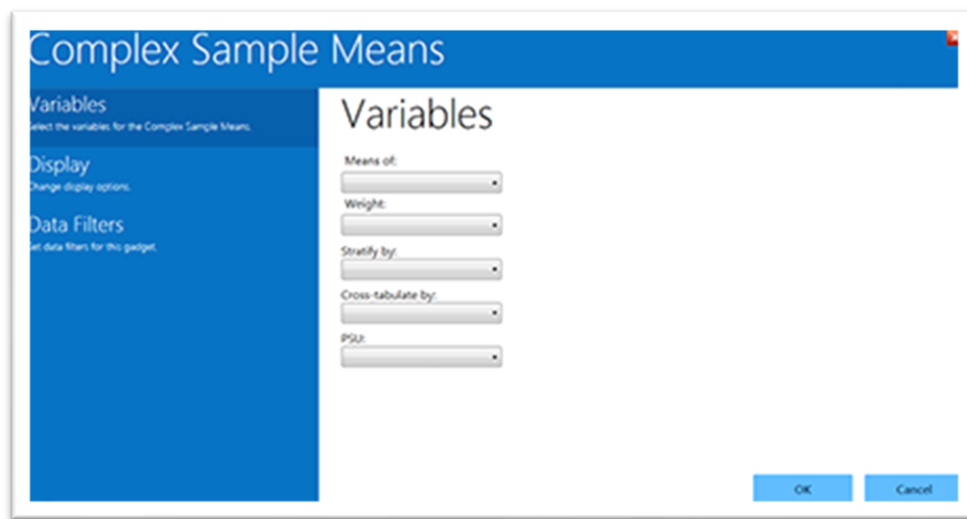
Complex Sample Means gadget

Complex Sample Means options:

- **Means of** identifies the variable for which the mean is computed.
- A **Weight** variable is selected for use in weighted analyses.
- **Stratify by** identifies the variable for grouping the frequency data.
- **Cross-tabulate by** identifies the variable to cross-tabulate the main variable.
- The **PSU** identifies the Primary Sampling Unit.

1. Select **NUMCIGAR** (number of Cigarettes smoked), from the **Means Of** drop-down list.

2. Select **SAMPW** (calculated based on sample size and population size) from the **Weight** drop-down list.

3. Select **Strata**, from the **Stratify By** drop-down list.

4. Select **Sex** from the **Cross-tabulate By** drop-down list.

5. Select **PSUID**, from the **PSU** drop-down list.

6. Click **OK**.

7. The **Complex Sample Means** results appear on the canvas.

				Confidence Limits			
SEX	Count	Mean	Std Error	Lower	Upper	Minimum	Maximum
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		

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.

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) influences 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**.

The Complex Sample Tables configuration window opens to the Variables property panel.
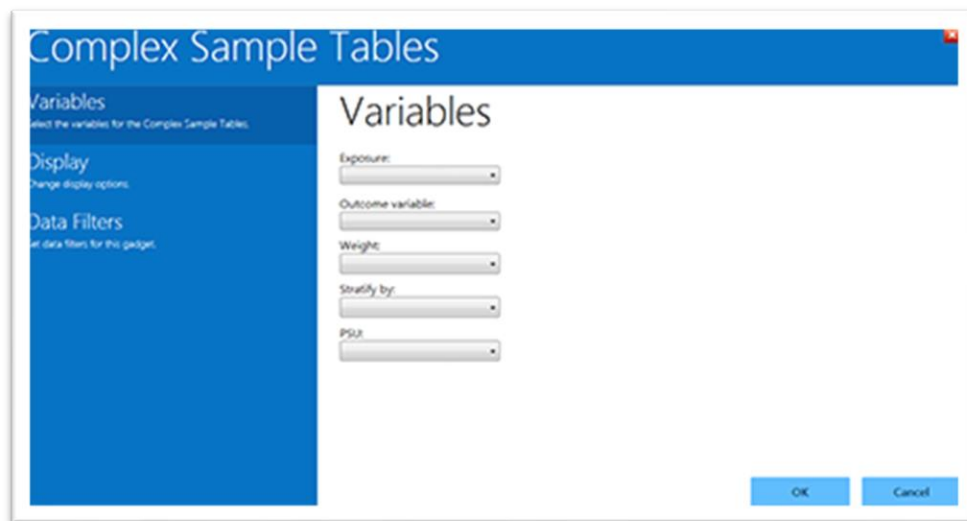
Complex Sample Tables gadget

Complex Sample Tables options:

- **Exposure** identifies the variable that displays on the horizontal axis of the table. It is considered a risk factor.

- **Outcome Variable** identifies the variable that displays on the vertical axis of the table.

- A **Weight variable** is selected for weighted analyses.

- **Stratify by** identifies the variable for grouping the frequency data.

- The **PSU** identifies the Primary Sampling Unit.

Select the following from the sequential drop-down lists:

- **PRENATAL** from the **Exposure** drop-down list,

- **VAC** from the **Outcome** drop-down list.

- **POPW** (the percent of the total population in each cluster) from the **Weight** drop-down list.

- **Location** from the **Stratify By** drop-down list.

- **Cluster** from the **PSU** drop-down list.

- Click **OK**.

- 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 1	VAC 2	TOTAL
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)

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.

Gadget Configuration Settings

You can move gadgets around the screen wherever you like. Each gadget contains three buttons located in the top-right corner. These icons are:
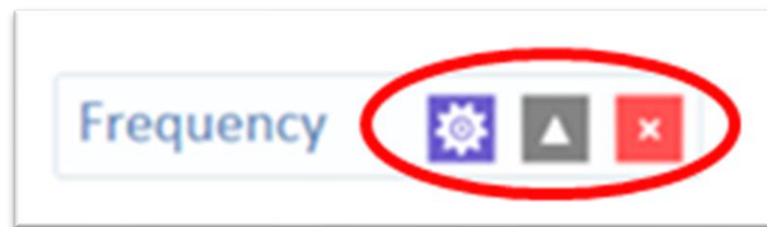
Gadget widgets

Change Gadget Properties: Use to review configuration set which fields the gadget will process.

: **Collapse/Expand Gadget Output**: Use to collapse and expand the gadget's output.

: **Close**: Removes the gadget.

The Change Gadget Properties button opens the gadget's configuration window including gadget filters. The Collapse/Expand button gray up/down arrow on the gadget's header panel enlarges and reduces the gadget's output. This button

117

removes the output from each gadget. Clicking on the red **X,** closes the gadget. Confirm the removal of the gadget from your canvas.

Line List

The Line List gadget displays the variables of the current dataset with their corresponding values. You can select specific variables from the variables box to narrow the list or press CTRL+A to select all the variables. Selecting a group or a page will show the variables contained in the group or page, respectively.

Follow the steps below to create a line list.

1. Load the Sample.PRJ project located in the **Epi Info™ 7\Projects\Sample** folder.
2. Select Oswego. Seventy-five (75) records should be listed in the **Record Gadget** window.
3. Right-click a blank portion of the **Dashboard canvas**. A context menu appears.
4. Select **Add Analysis Gadget > Line list** from the context menu. The line list gadget appears.
5. From the **Variables section**, use CTRL+Left Click to multi-select the following fields:
   a. Age
   b. DateOnset
   c. Sex
   d. Timesupper
   e. Ill
   f. Water

Line	AGE	SEX	TIMESUPPER	DATEONSET	ILL	WATER
1	20	Male	4/18/1940 10:00:00 PM	Missing	No	No
2	33	Female	Missing	4/18/1940 10:00:00 PM	Yes	Yes
3	62	Female	Missing	Missing	No	Yes
4	8	Female	Missing	4/19/1940 12:30:00 AM	Yes	No
5	11	Male	Missing	Missing	No	Yes
6	9	Female	Missing	Missing	No	No
7	14	Female	4/18/1940 10:00:00 PM	Missing	No	No
8	59	Female	4/18/1940 6:30:00 PM	4/19/1940 12:30:00 AM	Yes	No
9	52	Female	4/18/1940 8:00:00 PM	4/19/1940 12:30:00 AM	Yes	No
10	77	Male	Missing	4/19/1940 2:30:00 AM	Yes	No
11	45	Female	Missing	4/18/1940 11:00:00 PM	Yes	No
12	11	Male	Missing	Missing	No	No
13	72	Female	Missing	4/19/1940 2:00:00 AM	Yes	No
14	15	Female	4/18/1940 10:00:00 PM	4/19/1940 1:00:00 AM	Yes	No
15	53	Female	4/18/1940 7:30:00 PM	4/18/1940 11:30:00 PM	Yes	Yes
16	17	Male	4/18/1940 7:30:00 PM	Missing	No	No
17	64	Male	Missing	Missing	No	No
18	35	Male	Missing	4/18/1940 9:00:00 PM	Yes	No

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

Line List Fields

Notice the **warning** message showing at the top of the output. The warning states that the output is limited to 50 rows due to a gadget setting. The row limit prevents you from accidentally generating a line list of hundreds of thousands of records (which would greatly impair performance).

To modify this setting, follow the steps below:

1. Click the gear icon to show the **Line List Properties** panel.

2. Click on **Display**.

119

3. Change the **Max rows** to display from 50 to 100

4. Click **OK.**

The line list frequency output is re-generated with all the records displayed.

For example, if you want to sort the line list by Age, complete the following steps:

1. Click the gear icon to show the **Line List Properties** panel.

2. Click on the **Grouping and Sorting** section.

3. Select Ascending from the **Method drop-down** list.

4. Double-click on Age in the **Available Variables** box.

5. Click **OK**. The line list is sorted by Age.

Line	AGE	SEX	TIMESUPPER	DATEONSET	ILL	WATER
1	3	Male	Missing	4/18/1940 9:45:00 PM	Yes	Yes
2	7	Male	Missing	4/18/1940 11:00:00 PM	Yes	Yes
3	7	Male	Missing	Missing	No	Yes
4	8	Female	Missing	4/19/1940 12:30:00 AM	Yes	No
5	8	Male	4/18/1940 11:00:00 AM	4/18/1940 3:00:00 PM	Yes	No
6	9	Female	Missing	Missing	No	No
7	10	Male	4/18/1940 7:30:00 PM	4/19/1940 2:00:00 AM	Yes	No
8	11	Male	Missing	Missing	No	Yes
9	11	Male	Missing	Missing	No	No
10	11	Female	4/18/1940 7:30:00 PM	Missing	No	Yes
11	11	Female	Missing	Missing	No	No
12	12	Female	4/18/1940 10:00:00 PM	4/19/1940 1:00:00 AM	Yes	No
13	13	Female	Missing	Missing	No	No
14	13	Female	4/18/1940 10:00:00 PM	4/19/1940 1:00:00 AM	Yes	No
15	14	Female	4/18/1940 10:00:00 PM	Missing	No	No
16	15	Female	4/18/1940 10:00:00 PM	4/19/1940 1:00:00 AM	Yes	No
17	15	Male	4/18/1940 10:00:00 PM	4/19/1940 1:00:00 AM	Yes	No
18	15	Female	4/18/1940 10:00:00 PM	4/19/1940 1:00:00 AM	Yes	No

Line List

Moving and collapsing the Line List

The Line List gadget can be dragged around the canvas by left clicking on it, holding the left mouse button down, and then dragging the mouse. Gadgets can be placed anywhere, including on top of other gadgets.

Clicking the ▲ icon will collapse the contents of the line list output without close the line list. The benefit to collapsing output is more apparent when many gadgets are on the canvas at once.

Sending line list output to other programs

You can send the line list output to other programs or copy it to the clipboard. To send the line list data to Microsoft Excel, follow the steps below:

1. Right-click on the **output for the line list gadget**. A gadget-specific context menu appears.

2. Select **Send list data to Excel**. The line list data appears in Microsoft Excel after several moments.

If your computer does not have Microsoft Excel installed, you can select **Send list data to web browser**. Another option on the context menu, Copy list data to clipboard, allows you to paste the list output into other Windows applications. For most gadgets, you can send gadget data to the web browser, to Excel, and add it to the clipboard.

Frequency

The Frequency gadget produces a table that shows how many records have a value for each variable, the percentage of the total, and a cumulative percentage. You can select multiple variables with a single frequency gadget by pressing CTRL-click.

Follow the steps below to show a frequency of sex:

1. Load the Sample.PRJ project located in the **Epi Info™ 7\Projects\Sample** folder.

2. Select Oswego. Seventy-five (75) records should be listed in the **Record Gadget window**.

3. Right-click on a blank portion of the **Dashboard canvas**. A context menu appears.

4. Select **Add Analysis Gadget > Frequency** from the context menu. The frequency gadget appears.

5. Click on **ILL** from the list of variables.

6. Click **OK**. The frequency is displayed.

Ill?	Frequency	Percent	Cum. Percent	Exact 95% LCL	Exact 95% UCL	
Yes	46	61.33%	61.33%	49.38%	72.36%	
No	29	38.67%	100.00%	27.64%	50.62%	
TOTAL	**75**	**100.00%**	**100.00%**			
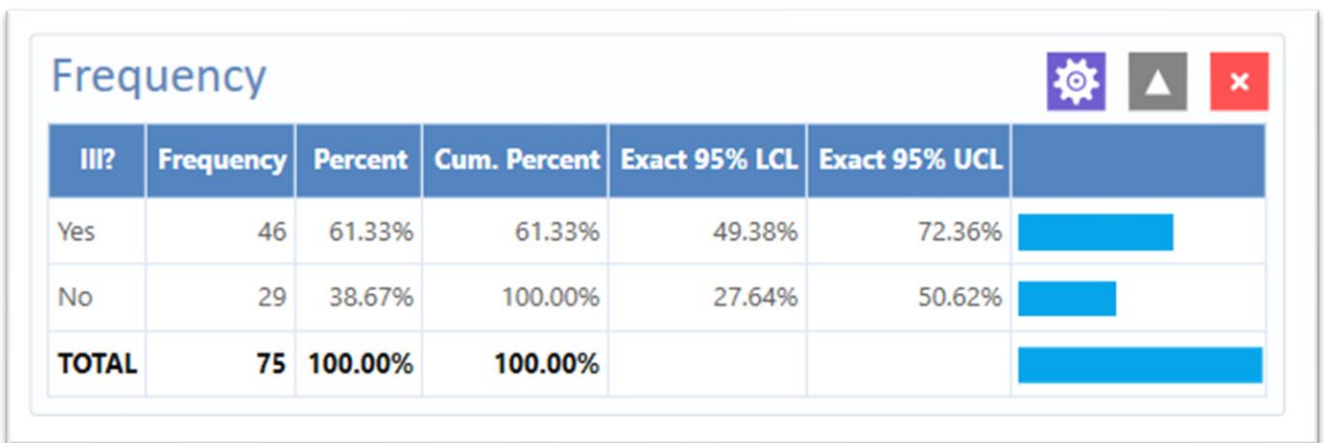
Frequency Fields

The Frequency column provides the count of individuals that were ill and not ill.

The Percent column indicates the percentage of who were ill or not ill.
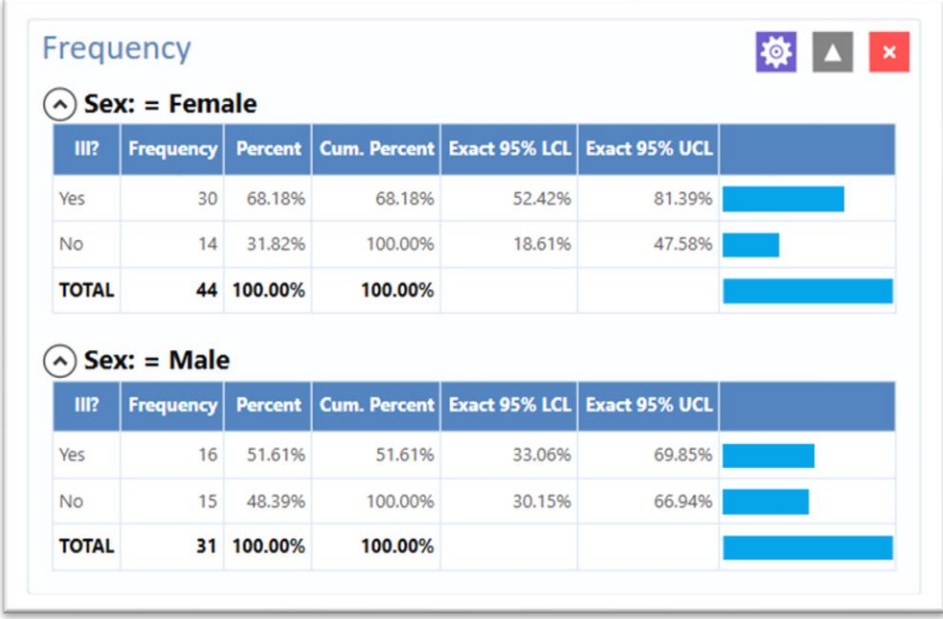
The 95% Confidence Limits are a range of values that indicates the likely location of the true value of a measure, meaning (in this instance) that the number of ill could be as low as 49.38%, or as high as 72.36%.

Note: The Yes Ill percentage of 61.33% falls within the 95% Confidence Limits range of values based on the data.

You can stratify frequency output. To stratify the sex frequency by Ill, follow the steps below:

1. Click the gear icon ⚙ to show the **Frequency Properties** panel.

2. Click on **Grouping and Sorting** section.

3. Select Sex from the **Stratify by list**.

Click **OK**. The frequency output is re-generated with stratified output.

Frequency Output by Sex

Additional Frequency Gadget Properties

The following options are available in the Frequency gadget properties panel shown in Figure 68.

Variables

- Frequency of: The field on which to run the frequency.
- Weight: The field to use as the weight variable, for example if aggregate data is being analyzed.

Grouping And Sorting

- Stratify by: The field to use for stratifying the output.
- Sort high to low: Sorts the output from highest frequency to lowest when selected. When cleared, the output is generally sorted alphabetically.

Display

- Gadget Title: Allows specific text to be displayed in the header of the gadget.

- Gadget Description: Allows specific text to be displayed as a description of the gadget.

- Display all list values: 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: Only applicable when the frequency variable is a Comment Legal field or Option Field in an Epi Info™™ 7 project. When

125

selected, the output will show the code and the code description for Comment Legal fields. For Option fields, the label is displayed.

- Include missing values: If selected, missing values are included in the output.

- Use field prompt: Only applicable when the frequency variable is a field in an Epi Info™ 7 project.

If checked, the **field prompt** will appear as the column heading instead of the field name.

- Draw borders: Setting for drawing the frequency border. Draw header row: Setting for displaying the frequency headers.
- Draw total row: Setting for displaying the frequency totals.
- Decimals to display: Specifies number of decimal places to display.
  - Maximum rows to display: Specify number of rows to display.
  - Max width of percent bar: Width for percent bar column using a % value.
  - Output columns to display: Determines which columns in the output are shown. By default, all columns are selected for display.

Data Filters

- Allows setting the data filters for a gadget.

Frequency Title and Description

Means
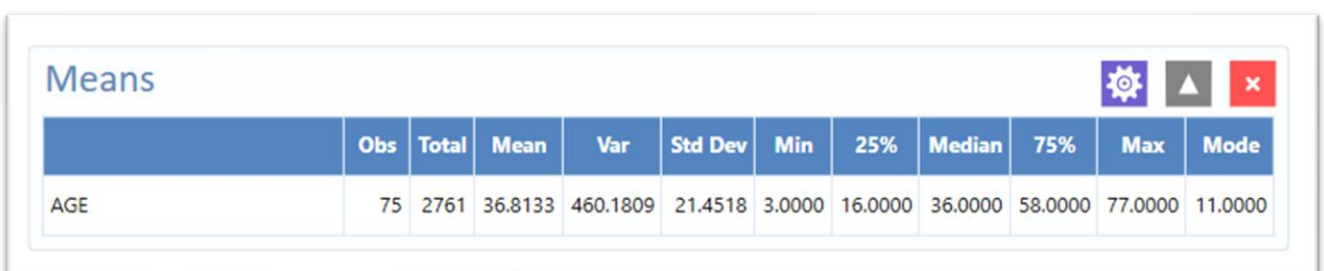
The **Means** gadget can obtain the Sum, Total, and/or the average of values for a continuous numeric variable. Since it only works with numbers, the variable must be numeric, date, or time. The results include the total number of observations (**Obs**); the sum of all observations (**Total**); the mean, variance, and standard deviation (**Std Dev**) of the observations; minimum and maximum values; 25th, 50th (**median**), and 75th percentiles; and the mode. If the variable has two or more values (**a tie**) for the mode (**most frequent**) value, the smallest will be presented in the Epi Info™ output. If only one variable is supplied, the **Means** gadget produces

a table with descriptive statistics. If two variables are supplied, the first is numeric containing data to be analyzed. The second indicates how groups will be distinguished. The output of this format is a table with descriptive statistics of the numeric variable for each group variable value.

Follow the steps below to show descriptive statistics for the Age variable:

1. Load the Sample.PRJ project located in the **Epi Info™ 7\Projects\Sample folder**.
2. Select Oswego. Seventy five (75) records should be listed in the **Record Gadget** window.
3. Right-click on a blank portion of the **Dashboard canvas**. A context menu appears.
4. Select Add Analysis Gadget > Means from the **context menu**. The means gadget appears.
5. Select Age in the **Means of: drop-down** list.
6. Click **OK**.

The summary statistics for Age are displayed. Notice the mean age of individuals in the dataset is 36.8.

	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

Stats for Age

For the next example, use a numeric variable or a Yes/No variable to determine the Means for a group.

1. Right-click on a blank portion of the **Dashboard canvas**. A context menu appears.

2. Select Add Analysis Gadget > Means from the **context menu**. The means gadget appears.

3. Select Age in the **Means of: drop-down** list.

4. Select ILL from the **Cross-Tabulate by Value of drop-down** list,.

5. Click OK.

Notice the Mean Age of those answering Yes to ill is 39.2. Those answering No to ill is 32.9.

Means

AGE * ILL	Obs	Total	Mean	Var	Std Dev	Min	25%	Median	75%	Max	Mode
Yes	46	1806	39.2609	477.2638	21.8464	3.0000	17.0000	38.5000	59.0000	77.0000	15.0000
No	29	955	32.9310	423.7094	20.5842	7.0000	14.0000	35.0000	50.0000	69.0000	11.0000

⌃ T-Test

	Method	Mean	95% CL		StdDev
Diff (Group 1 - Group 2)	Pooled	6.3298	-3.7693	16.4290	21.3711
Diff (Group 1 - Group 2)	Satterthwaite	6.3298	-3.6601	16.3198	

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	73	1.2491	0.2156
Satterthwaite	Unequal	62.3293	1.2663	0.2101

⌃ ANOVA, a Parametric Test for Inequality of Population Means

(For normally distributed data only)

Variation	SS	dF	MS	F-Statistic
Between	712.6550	1	712.6550	1.5604
Within	33,340.7316	73	456.7224	
Total	34,053.3867	74		

P-Value	0.2156

Bartlett's Test for Inequality of Population Variances

Chi Square	0.1193
Degrees of freedom	1
P-Value	0.7298

A small p-value (e.g., less than 0.05) suggests that the variances are not homogeneous and that the ANOVA may not be appropriate.

Mann-Whitney/Wilcoxon Two-Sample Test (Kruskal-Wallis test for two groups)

Kruskal-Wallis H	1.1612
Degrees of freedom	1
P-Value	0.2812

Means for Age

When comparing two means an independent t-test is performed, and for comparing more than two means, a one-way analysis of variance (ANOVA) is performed. You can use the independent t-test and ANOVA if the following assumptions are met:

- The outcome variable is normally distributed in each group.
- The underlying variances are the same in each group.

Epi Info™ does not provide a statistical test to determine the normal distribution of data; graph the data to visually detect normal data distribution. Epi Info™ does perform a test to determine if the second assumption above is met called "Bartlett's Test"; if the p-value is large (say > 0.05) this would suggest the variances are approximately equal; if the p-value from Bartlett's test is small (say < 0.05), this suggests that the underlying variances are not the same and therefore the t-test and ANOVA results may not be appropriate for the data.

MxN/2x2 Table

The MxN/2x2 Table gadget examines the relationship between two or more categorical values. For 2×2 tables, the gadget produces odds and risk ratios. A 2×2 table is created when each selected variable has a Boolean or Yes/No answer value.

Create a 2×2 table using data from the Sample.MDB project.

1. Read in the **Sample.PRJ** project located in the Epi Info™ 7\Projects\Sample.
2. Select **EvansCounty**.
3. Six hundred (600) records should be listed in the **Output window**.

4. Right-click on a blank portion of the **Dashboard canvas**. A context menu appears.

5. Select Add **Analysis Gadget > MxN/2x2 Table**. The gadget appears.

6. Select CAT From the **Exposure Variable drop-down** list.

7. Select CHD From the **Outcome Variable drop-down** list.

8. Click **OK**. Results are generated.

The variable CAT corresponds to catecholamine while the variable CHD corresponds to coronary heart disease. The output for two Yes/No variables, shown below, is a 2-by-2 table that contains: the number of observations in each cell, the row percent (**Row %**), and the column percent (**Col %**). You can adjust Row and Column percents to on or off from the **SETTINGS** command, the last one in the command list.

View the graphic at the right of the table. It depicts how the distribution of observations within the table – the larger the box, the larger the number of observations.

Crosstabulation

CHD

CAT	CHD Yes	CHD No	
Yes	27 22.13% 38.03%	95 77.87% 17.66%	122 100.00 % 20.03%
No	44 9.03% 61.97%	443 90.97% 82.34%	487 100.00 % 79.97%
	71 11.66% 100.00 %	538 88.34% 100.00 %	609 100.00% 100.00 %

Single Table Analysis

Odds- and Risk-based parameters	Estimate	Lower	Upper	Statistical Tests	X²	2 Tailed P
Odds ratio	2.8615	1.6878	4.8514	Uncorrected	16.2465	0.0000556135
MLE Odds ratio (Mid-P)	2.8554	1.6690	4.8350	Mantel-Haenszel	16.2198	0.0000564021
Fisher-Exact		1.6148	4.9853	Corrected	14.9998	0.0001075244
Risk ratio	2.4495	1.5837	3.7887		1 Tailed P	2 Tailed P
Risk difference	13.0962	5.3021	20.8903	Mid-P Exact	0.0000911051	
				Fisher-Exact	0.0001374257	0.0002048979

Crosstabulation and Analysis

The small red box represents the upper left cell of the table, the ill persons with exposure to the risk factor. The large green box on the lower right represents the persons who were not ill and did not have the specified exposure. Both boxes are in favor of the hypothesis that there is an association between exposure and illness. The other two boxes in the upper right and lower left are discordant with the hypothesis.

The odds ratio, one of the measures of association, can be thought of as red x green divided by brown x yellow, or the product of boxes in favor divided by the product of boxes against the association. If the boxes against predominate, then the exposure

may be protective (odds ratio less than 1.0). If the red and green boxes (upper left and lower right) win the day, the odds ratio exceeds 1.0, and there is a positive association. If all the boxes are the same size or the two cross-products are equal, the odds ratio is 1.0, suggesting no association.

Beneath the table is the Single Table Analysis that shows parameter estimates with confidence intervals and statistical tests. The first set of parameter estimates are based on the odds, with an odds ratio based on the cross product [i.e., (a x d) / (b x c)], and one based on the maximum likelihood estimation approach (MLE). View the three different confidence intervals; the Taylor series, mid-P exact, and the Fisher exact.

Next are the Risk-based estimates - the risk ratio and the risk difference with their confidence intervals.

Note: If the outcome variable is based on prevalent disease, then substitute the terms "prevalence ratio" and "prevalence difference" for "risk ratio" and "risk difference", respectively.

Finally, several statistical test results are provided: three different chi square tests and two exact tests. The chi square tests are presented as two-tailed p-values (although you could divide the two-tailed p-value by 2 to calculate a one-tailed p-value), and the exact tests are presented as one-tailed p-values (you could multiply the one-tailed p-values by 2 to get a two-tailed p-value).
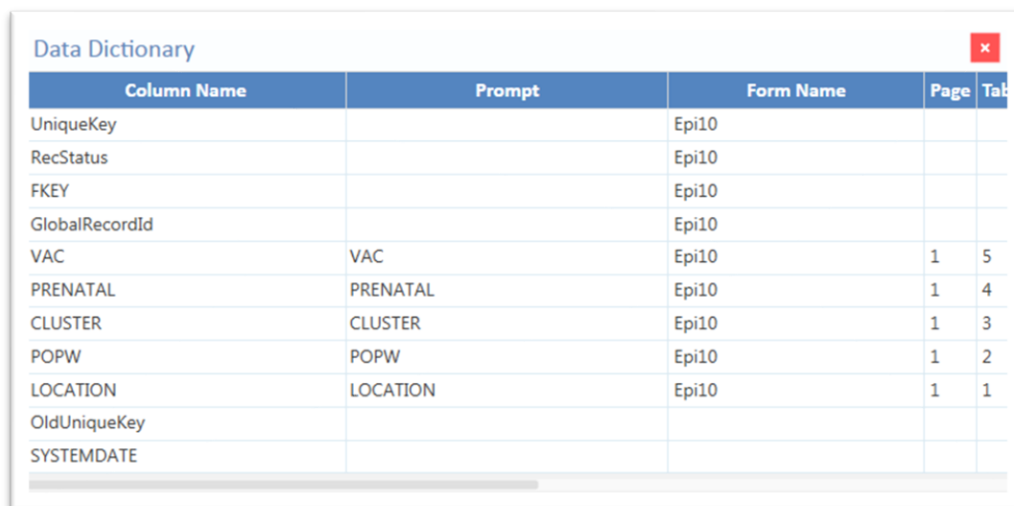
Data Dictionary

Show Data Dictionary

The Data Dictionary displays form(s) and defined variables for the current Data Source. Fields or variables sort and display by page number in the form (if using an Epi Info project as the data source). 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** displays 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				

Data Dictionary view

Use the scroll bar at the bottom of the **Data Dictionary** 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) populate when fields are

created from the **Field Definition** dialog box. Review the Form Designer Guide from the collection of Epi Info™ 7 help documentation for more Information.

## Open Data Dictionary as an HTML page inside a web browser

Steps:

1. Hover over the dictionary window

2. Right-click and select **Send Data Dictionary to Web Browser**.
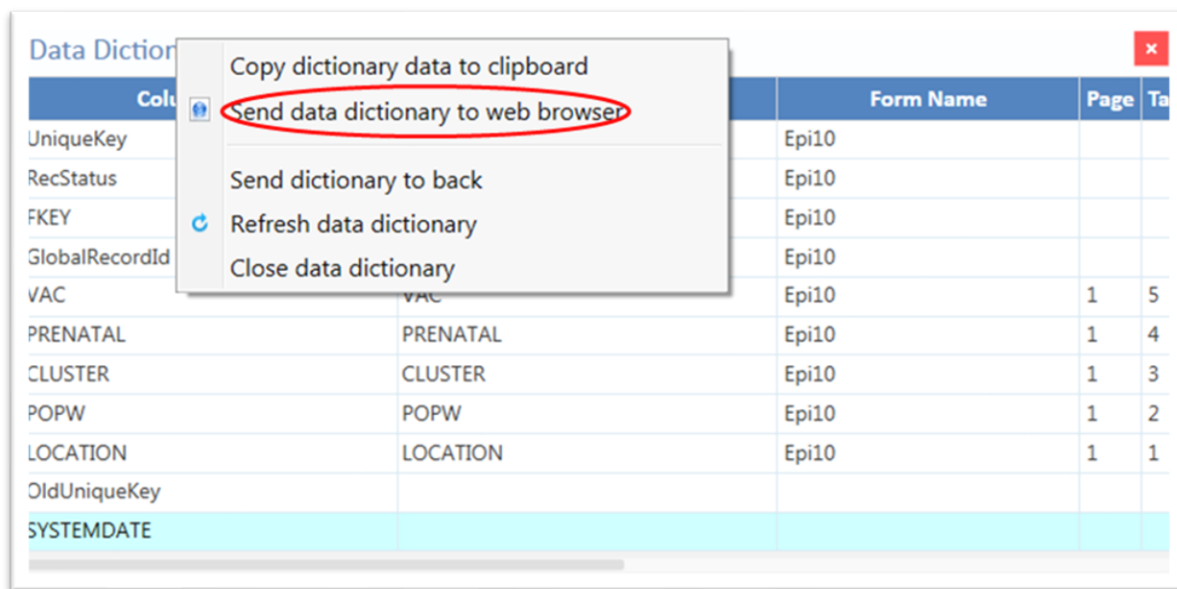
Data Dictionary to Web

1. Print web browser content by right-clicking on the page window and selecting **print** or use the browser **options menu**. You can also save content from **save as** in the browser **options menu**.

2. You can export data to an Excel spreadsheet by right-clicking over the table in the web page. The pop-up menu shows **Export to Microsoft Excel**.
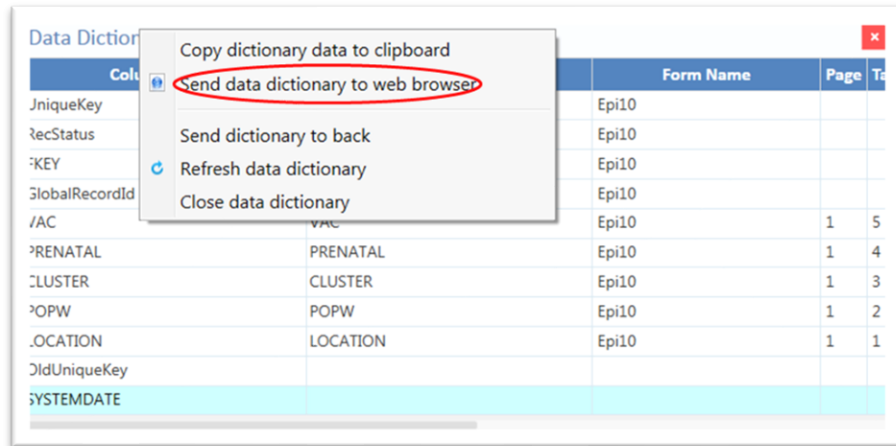
Data Dictionary Export to HTML

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

Note: Refresh **Data Dictionary** at any time so you see the most current data. Right
click on **Data Dictionary** and select **Refresh Data Dictionary**.

Additional Gadgets for Reports

Simple Text Box and the Image Gadget

These gadgets add text and images to your reports.

Simple Text Box

You to specify several lines of text displaying on the canvas using Simple Text Box.
Use these lines of text for annotations, above or beside gadgets on the canvas.
Annotations are great for adding titles to the canvas.

Steps:

1. Right-click on the Visual Dashboard canvas. Select the **Add Report gadget** from menu options.
2. Select **Simple Text box**. The text box displays.
3. Enter a title or other text. You can move the text box around by clicking on it and dragging the text box to where you want it.

Removing your title:
1. Right-click on the **text box border**.
2. Click **Close** to close this gadget.

Image Gadget

Image Gadget is another report gadget. You can add an image or logo to the canvas.

Steps:
1. Right-click on the Visual Dashboard **canvas**.
2. Select **Add Report** gadget from the menu options, and then select **Image gadget**. A **light blue box** displays near your mouse pointer location.
3. Click on the **light blue box** and the **Open** dialog box displays. Select the **location** of your image or enter the **filename** of your image.

Note: You may only select a PNG file for the Image gadget.

4. Double-click the **image**. You can also select it and click **Open**.
5. The image displays on the canvas.

Removing your image:
1. Right-click on your image.
2. Click **Close** to c

This page was intentionally left blank.