Beginner's Guide to
Getting Results with Westat's Data Xplorer

Table of Contents

Overview 2

How to Get Started Using Westat's Data Xplorer 2

How to Access Project Documents 3

How to Generate Project Reports 5

How to Create Your Own Report Tables and Charts Using Project Data 6

- One Variable Tables 7
- Using Results Outside of the Current WesDaX Session 17
- Saving Queries and Using Saved Queries 18
- Commonly Used Queries 19
- Restricting Tables to a Subset of the Data -- Filters 20
- Recoding Categorical Variables 25
- Two Variable Tables 30
- Three Variable Tables 34
- Comparisons 39
- Benchmarks 45

List of Figures 48

Overview

Westat's Data Xplorer (WesDaX) is an easy way to conduct simple analyses using project data.

This tutorial will show you how to:

- Get Started using Westat's Data Xplorer
- Access Project Documents
- Generate Project Reports
- Create Your Own Report Tables and Charts Using Project Data

The tutorial uses a demonstration project (or "demo" project) that you can access by going to http://www.dataxplorer.com. Because projects can customize Westat's Data Xplorer, some headings and titles shown for the demo project may be different from those used for your project. The functionality will be the same. Once you can do something on the demo project web site, you will know how to do it on your actual project web site.

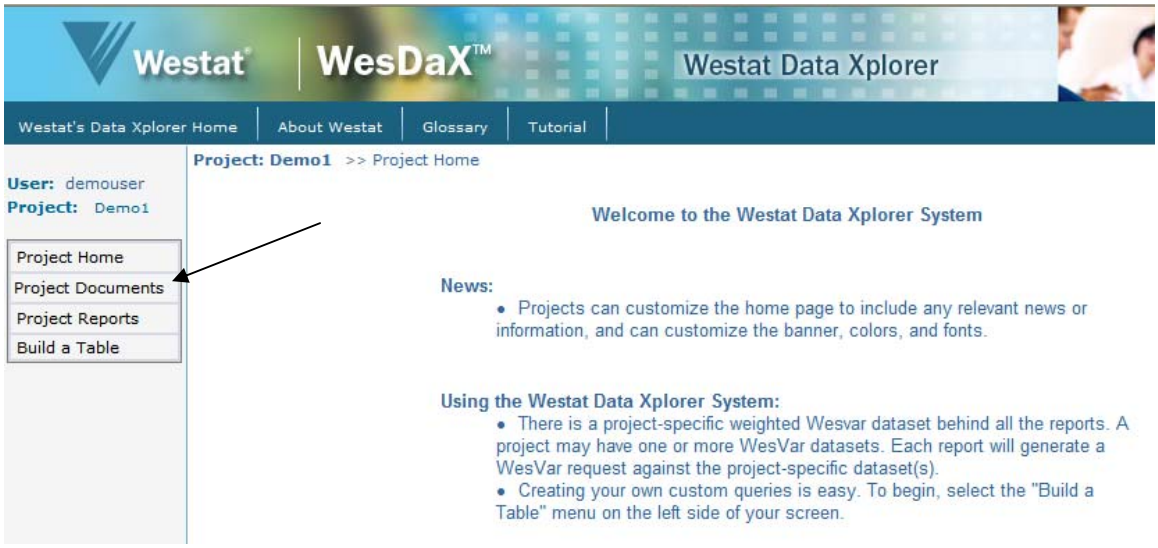
How to Get Started Using Westat's Data Xplorer

- Go to: http://www.dataxplorer.com
- Enter the word "demouser" (but don't include the quotation marks) for the User Name and the word "Westat!1" for the Password (but don't include the quotation marks)

How to Access Project Documents

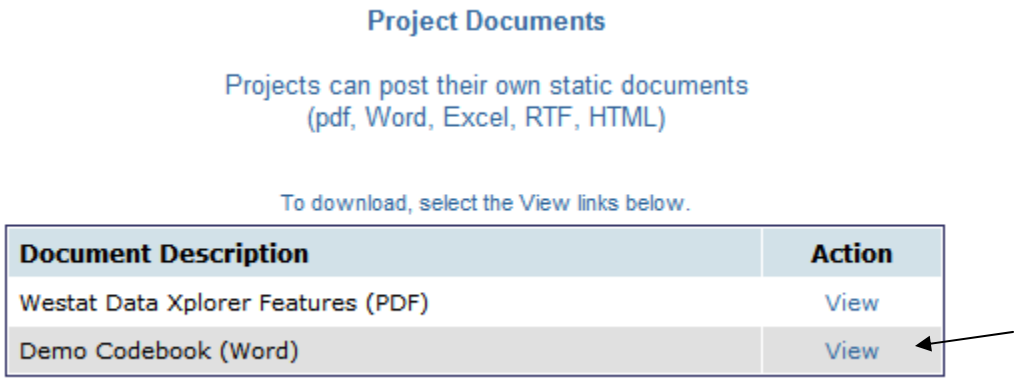
To see a list of project documents, click on the heading "Project Documents" on the left side of the screen, as shown in Figure 1.

Figure 1. Selecting Project Documents



You will see a list of documents (Figure 2). Simply click on the word "View" next to the name of a document to display it. Note that documents can be in a variety of different formats such as Microsoft Word, Portable Document Format (PDF), or Excel.

Figure 2. Document List



Any number of documents can be included in this list. For example, the list of documents may include a project overview. It could also include a description of the project methodology or a codebook that lists the variables in the data file.

If you choose to view the "Demo Codebook," you will be prompted to open or save the document. Select "Open" and you will see the codebook displayed in Figure 3:

Figure 3. Demo Codebook

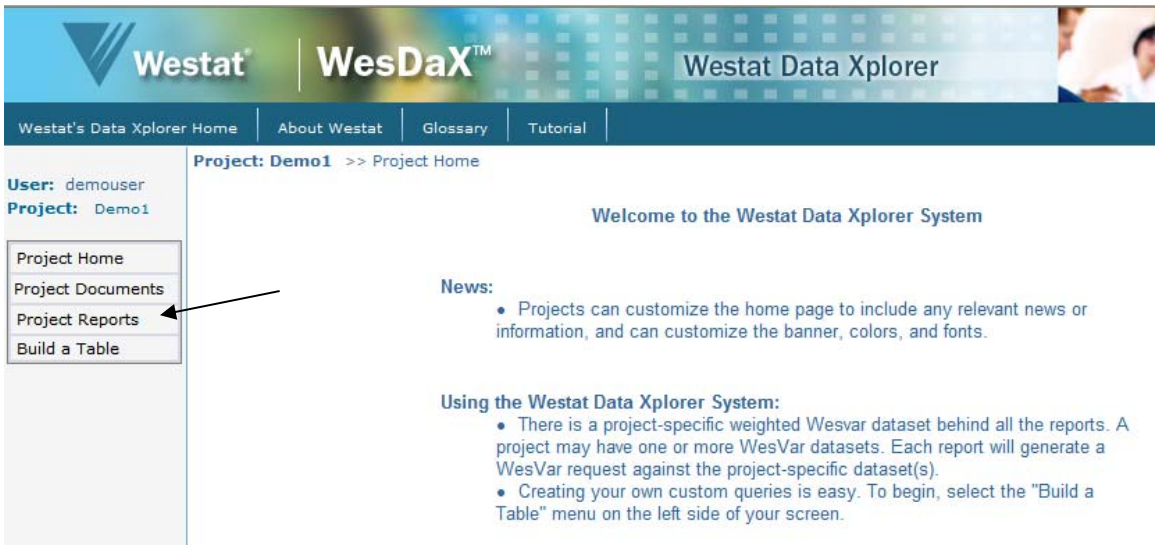
Variable Name	Description	Values
CIG100	Smoked within last 100 days	1=Yes 2=No
DIAB1	Diagnosis of Diabetes	1=Yes 2=No
HYPERTEN	Diagnosis of Hypertension	1=Yes 2=No
RACE	Coded Race	1=White 2=Black 3=Asian
RACEETHN	Coded Ethnicity	1=Asian 2=African 3=Latino 4=White
REC_AGE	Recoded Age	1=< 20 2=20 - 29 3=30 - 39 4=40 - 49 5=50 - 59 6=>= 60
REC_SYS	Systolic Indicator	0=Systolic OK 1=Systolic too high
SEX	Gender	1=M 2=F 3=U
STATUS	Interview Status	1=PARTIAL 2=COMPLETE 3=REFUSE

This table provides the user with the name, description, and value labels for each variable in the project data file.

How to Generate Project Reports

To see a list of reports that have already been created by the project staff, click on the heading "Project Reports" on the left side of the screen, as shown in Figure 4.

Figure 4. Selecting Project Reports



You will see a list of reports, as illustrated below in Figure 5. To display a report, simply click on the word "View", located next to the name of the report. Data Xplorer will run a program that looks at the current version of the project data file and will generate the report for you. For example, if you select the Predefined WesVar Report (Gender by Race) (see Figure 5.) WesDaX will produce the Sex by Race table shown in Figure 6.

Figure 5. Viewing List of Project Reports

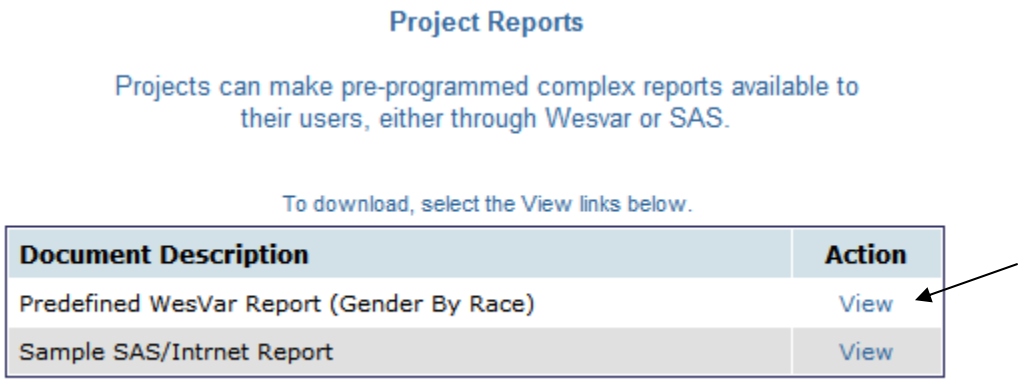


Figure 6. Predefined WesVar Report (Gender by Race)

Predefined WesVar Report (Gender By Race)
"Sex (SEX)" by "Race (RACE)"

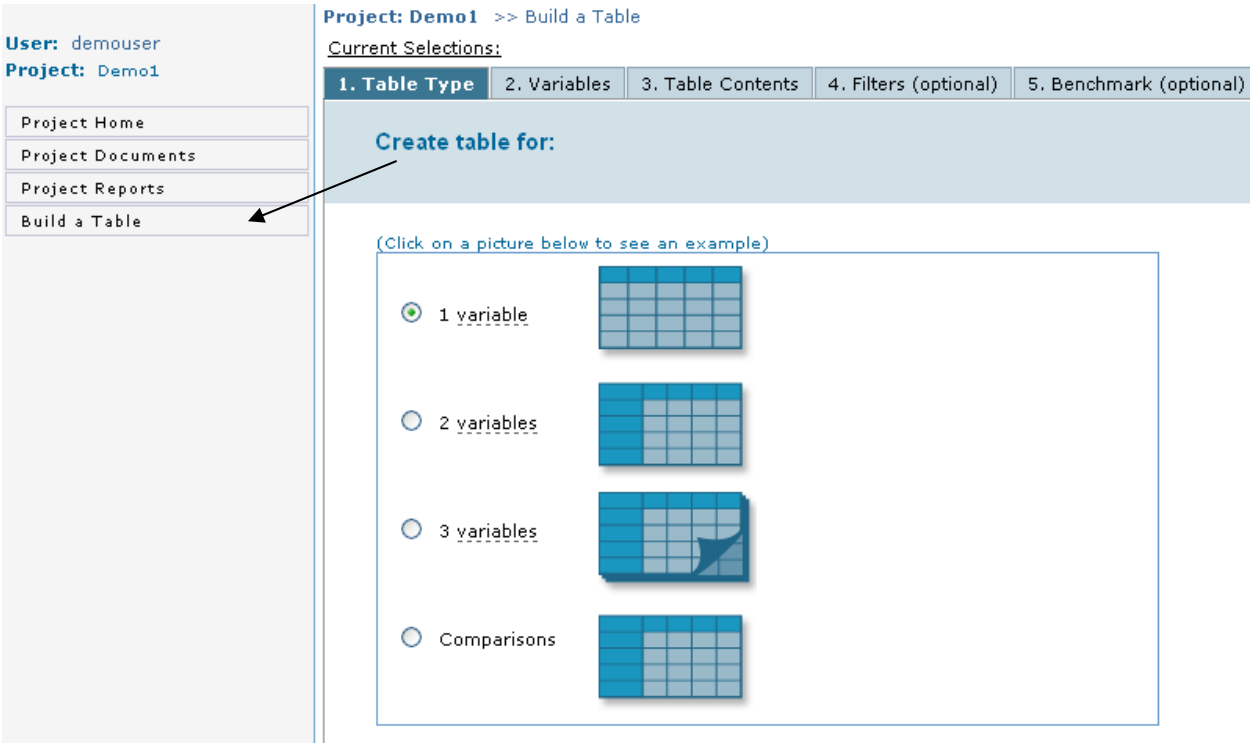
Number of Responses	Race (RACE)			
Sex (SEX)	White	Black	Other	Total
Male	76,012,386	9,671,224	3,945,928	89,629,538
Female	82,118,740	12,056,864	3,828,001	98,003,605
Total	158,131,126	21,728,088	7,773,929	187,633,143

How to Create Your Own Report Tables and Charts Using Project Data

WesDaX gives you the ability to easily summarize project data in ways that fit your needs. In addition to showing the summaries as tables, you can also display them as charts or export them to Excel for further analysis or use in documents or presentations.

To create your own table, click on "Build a Table" on the left side of the screen (see Figure 7.) You will see a display giving you the choice of creating a table with one, two, or three variables. Most of the time, you will probably want to use the one and two variable options.

Figure 7. Table Type Selection Screen



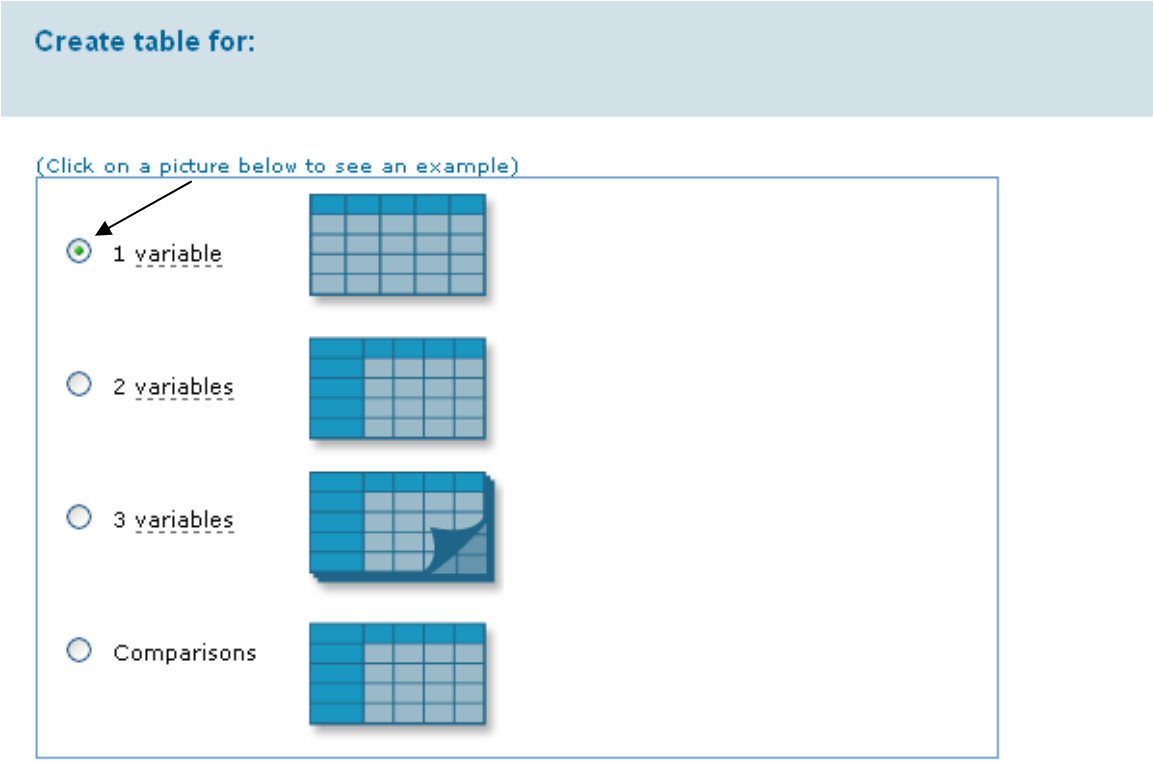
To build a table, you will need to provide information for each of the numbered tabs (e.g., 1. Table Type, 2. Variables, etc.) in the order that the tabs are laid out on the screen and then click on the SHOW RESULTS button in the right corner.

For example, to show the race distribution for respondents in the data set, you will need to first indicate that you would like to create a 1 variable table. Then you need to indicate that the variable you would like to see a distribution for is Race. To see the final results, simply click on the SHOW RESULTS button in the right upper corner. These steps are illustrated below.

One Variable Tables

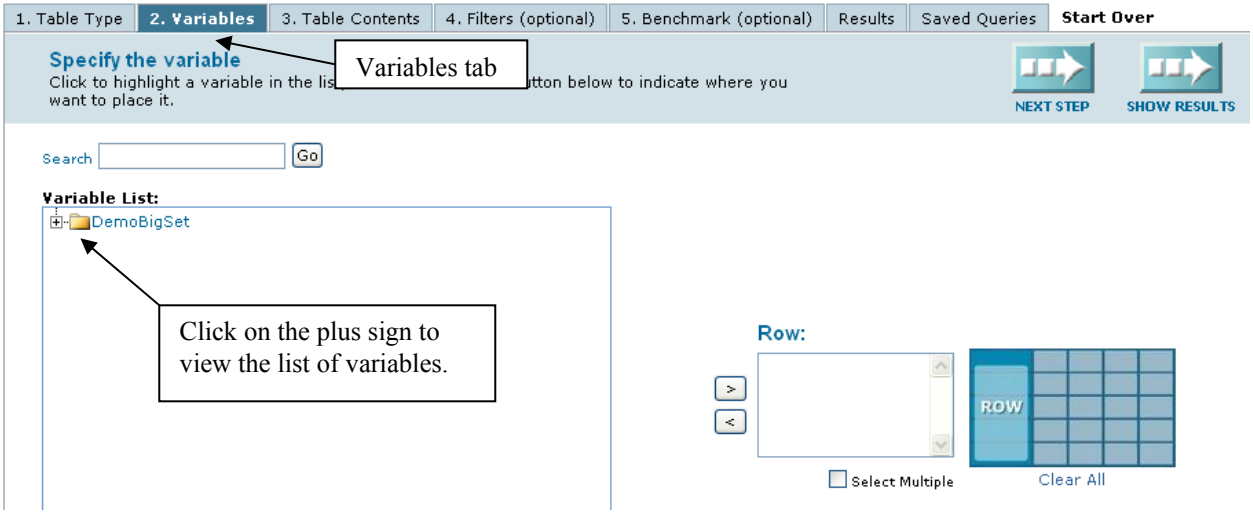
To tell WesDaX to create a one variable table, simply click on the button next to the words "1 variable" (see Figure 8).

Figure 8. Specifying One Variable Table



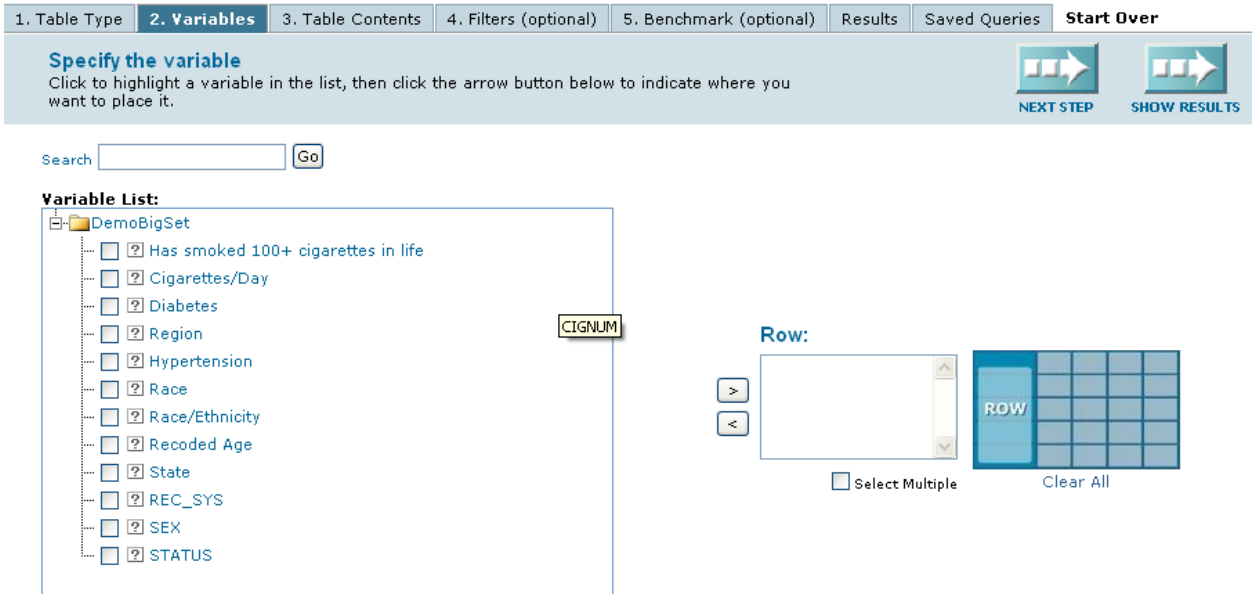
Next, click on the "Variables" tab. The screen will look like Figure 9. To see the list of variables in the data set, click on the plus sign (+) to expand the variable list.

Figure 9. Variable List before Being Expanded



After you click on the + next to the yellow folder (as shown in Figure 9), you will see:

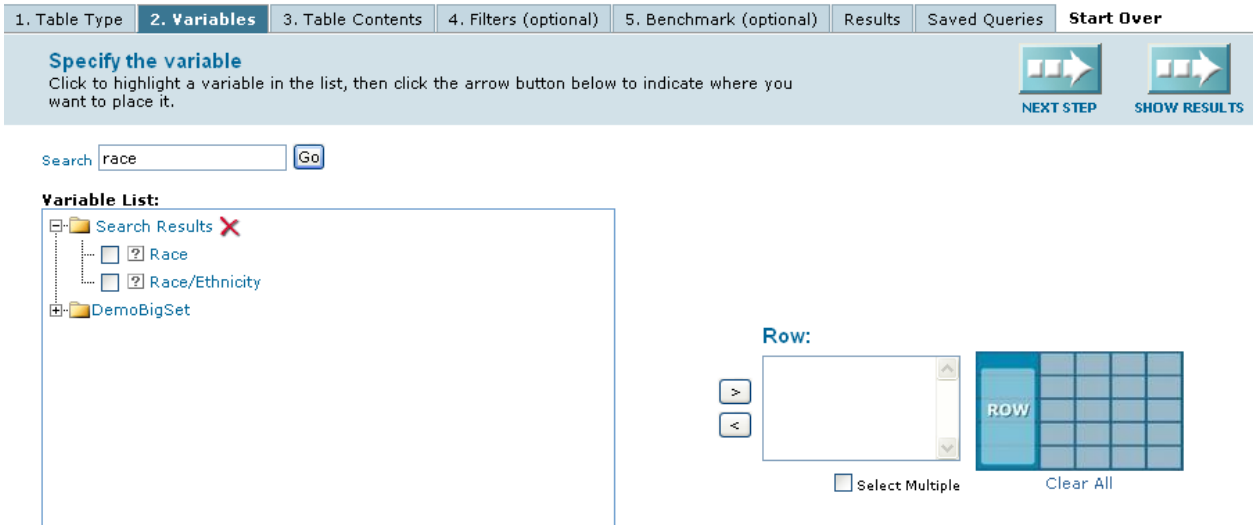
Figure 10. Variable List after Being Expanded



If the study has a large number of variables, it could be difficult to find a variable of interest. To make this task easier, it is possible to enter a search string in the "Search"

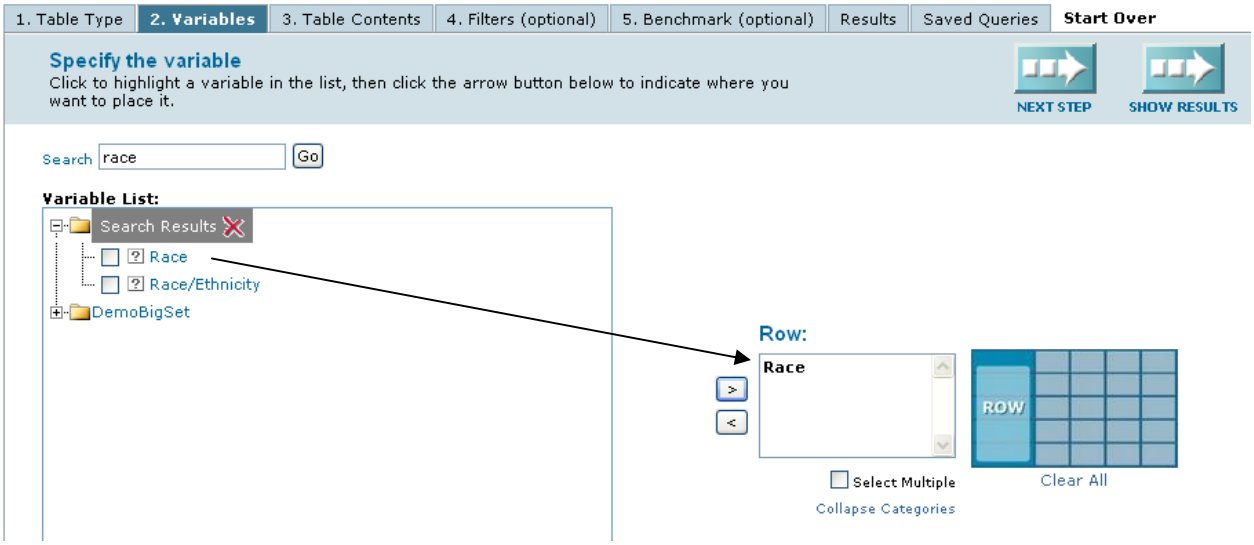
box immediately above the variable list. When you click on the "Go" button, a list of the variables matching your search string appears in the variable list as shown in Figure 11.

Figure 11. Variable List after Search for "race"



To select the variable "Race," left click on the word "Race" in the variable list and drag it into the box labeled "Row:" so that the screen looks like Figure 12. You can also accomplish the same thing by checking the box next to the variable name and then clicking the forward arrow symbol (>). The backward arrow symbol (<) removes the variable from the list. Note that this works the same in both the short version of the variable list and the full version. To remove the search results, click on the red "X" next to the words "Search Results."

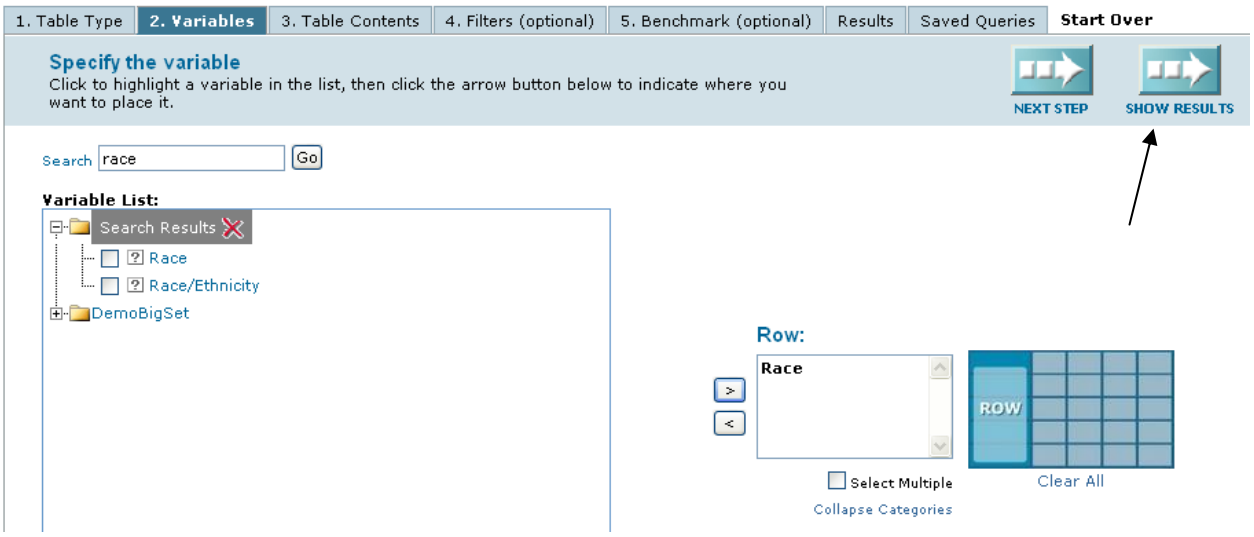
Figure 12. Selecting the Variable ("Race")



To change the variable in the "Row" box you can either click on the words "Clear All" beneath the grid and drag another variable into the "Row" box or you can simply drag the new variable into the box. The new variable will replace the previous variable.

When you have the variable you want in the "Row" box, click on the large button at the top right labeled SHOW RESULTS (see Figure 13).

Figure 13. Show Results Button



When you click on the SHOW RESULTS button, you will see the table shown in Figure 14.

Figure 14. Table Showing Frequency Distribution of a Variable (Race)

Race (RACE)

Race (RACE)	Number of Responses
(missing)	14,063
White	158,131,126
Black	21,728,088
Other	7,773,929
Total	187,647,206

An easy way to display these results graphically, is to export them to Excel by clicking on the Export to Excel button at the top of the screen (see Figure 15), highlighting the data table (including labels) in Excel, and then selecting one of the chart types on the Insert ribbon bar. Figure 16 is the result of choosing the first pie chart option in Excel.

Figure 15. Exporting to Excel

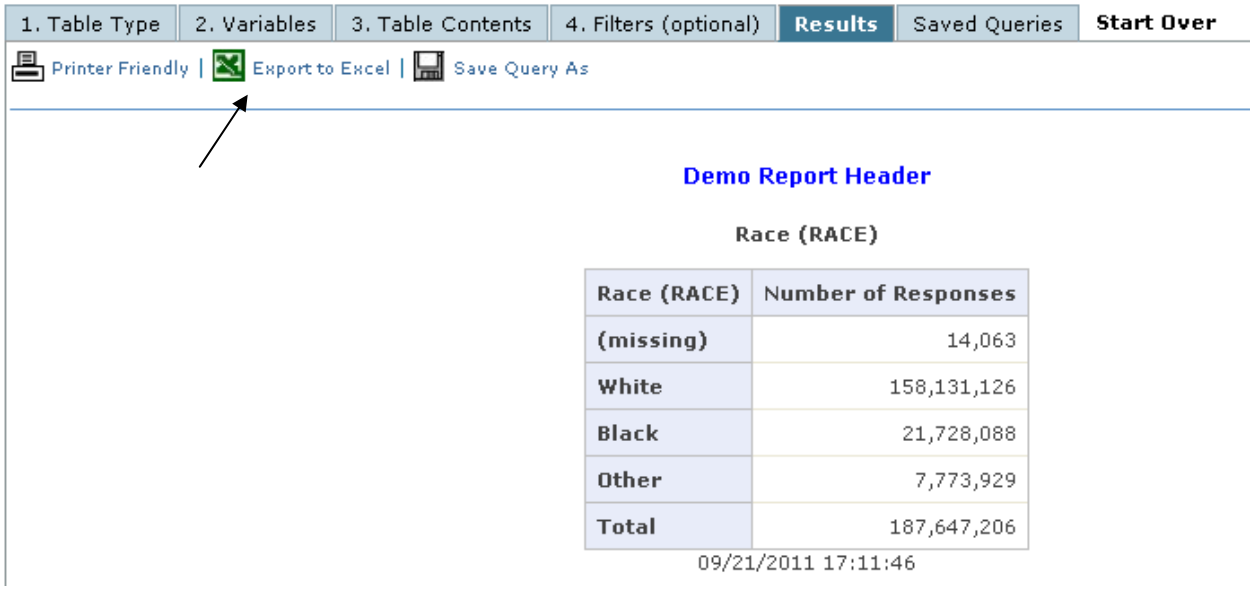
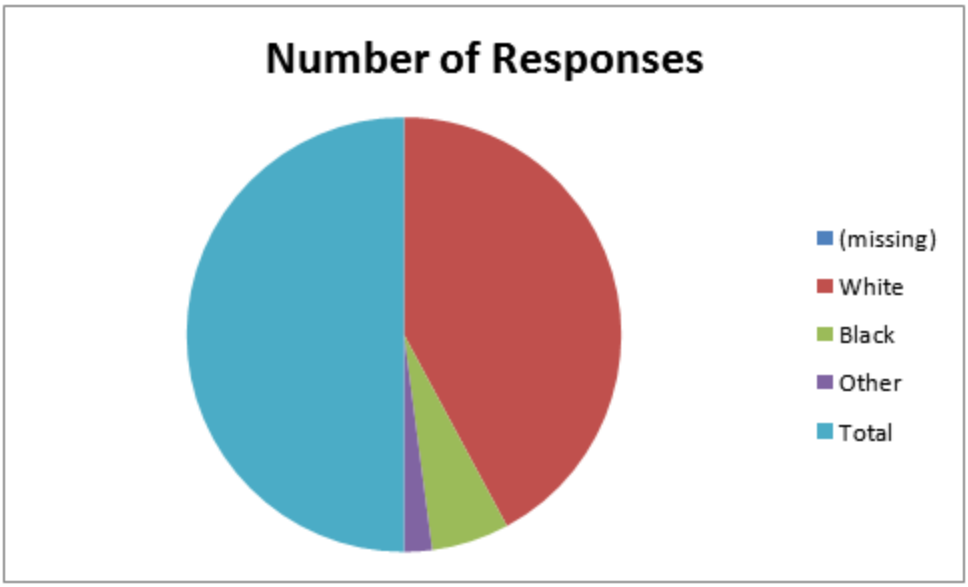
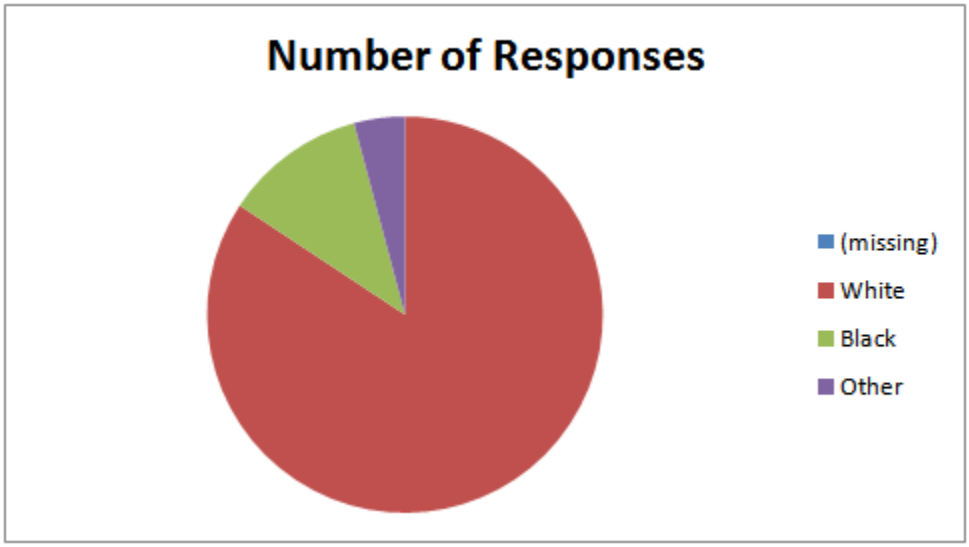


Figure 16. Pie Chart of a Variable (Race) – Including Total



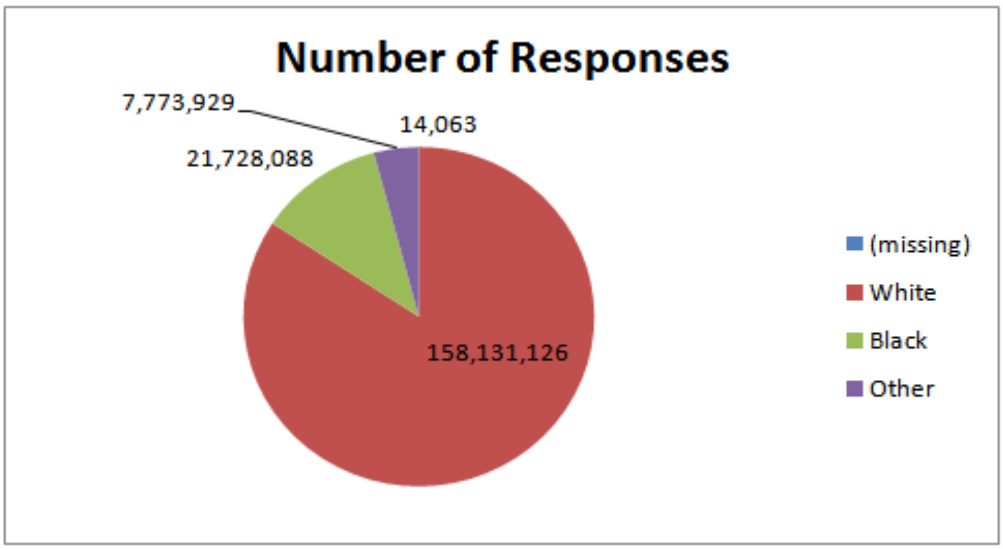
Since it does not make sense to include the total as well as all of the individual categories of the variable race in the chart, it is better to either modify the tabular values before creating the image or to select just the data values of interest. Figure 17, shows the results after deleting the Total row in the Excel table and then selecting the first pie chart option.

Figure 17. Pie Chart of a Variable (Race) –after Deleting the Total Row



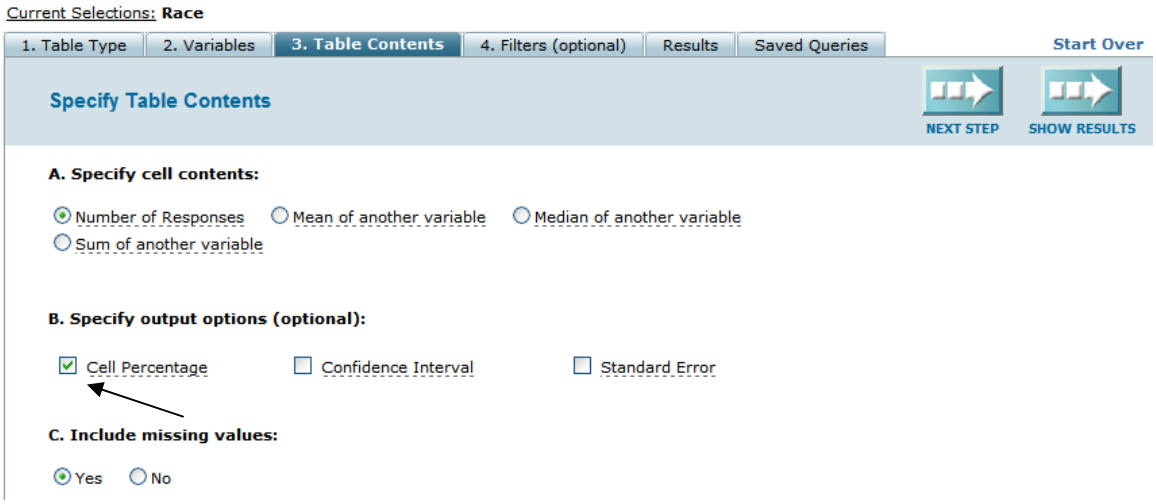
Note that by right clicking in the figure and selecting the "Add Data Labels" option, you can have Excel display the numbers associated with each wedge in the pie (Figure 18).

Figure 18. Pie Chart of a Variable (Race) –after Deleting the Total Row with Data Labels



If you are interested in seeing the percentage of respondents in each category, you can do so by going to the third tab labeled, "3. Table Contents" and checking the "Cell Percentage" box under "B. Specify output options (optional)" as shown in Figure 19 below.

Figure 19. Selecting Cell Percentage Display Option



If you now click on SHOW RESULTS in the right corner you will see that a new column with percentages has been added to your table. It should look like Figure 20:

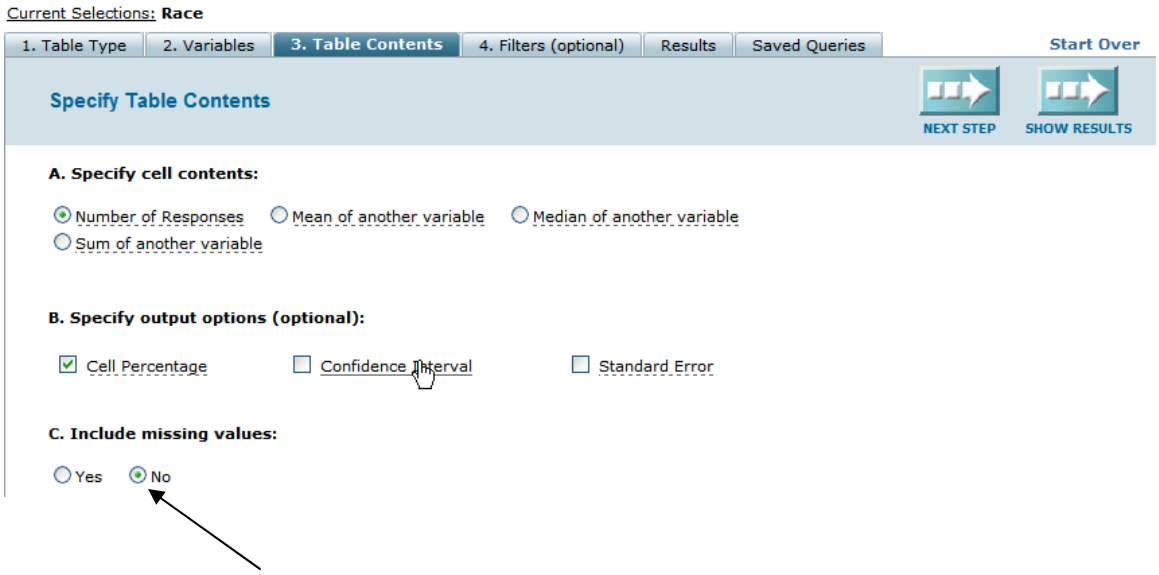
Figure 20. Table with Both Counts and Percentages Displayed (Race Variable)

Race (RACE)

Race (RACE)	Number of Responses	Cell Pct
(missing)	14,063	0.01%
White	158,131,126	84.27%
Black	21,728,088	11.58%
Other	7,773,929	4.14%
Total	187,647,206	100.00%

If you want to look at only those people for whom the project has collected race data, you can go back to the tab labeled, "3. Table Contents" and exclude the missing values by checking the "No" button in Section C, "Include missing values." (See Figure 21.)

Figure 21. Suppressing the Display of Missing Values



After you click the "No" button for missing values, you need to click on the SHOW RESULTS button in the right corner. Your new table will look like Figure 22.

Figure 22. Table with Missing Values Excluded (Race Variable)

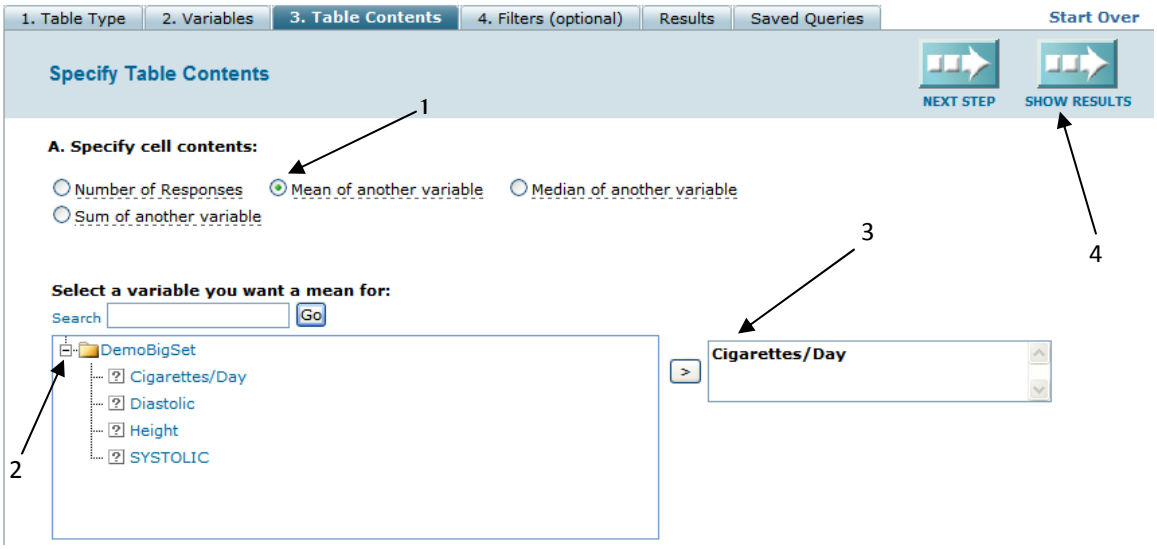
Race (RACE)

Race (RACE)	Number of Responses	Cell Pct
White	158,131,126	84.28%
Black	21,728,088	11.58%
Other	7,773,929	4.14%
Total	187,633,143	100.00%

If you are interested in displaying a summary statistic about each group in the table, for example, the mean number of cigarettes smoked per day for each race group you can indicate that in Section A of the Table Contents tab. To do this, you will need to follow the steps outlined below and illustrated in Figure 23:

1. Go to tab 3. Table Contents and check the statistic you want (in this case "Mean of another variable")
2. Expand the variable list by clicking on the plus (+) sign next at the top of the box that displays variable names or by using the search box to search for a specific variable name,
3. Drag the variable name of interest to the box on the right, and
4. Click on SHOW RESULTS.

Figure 23. Specifying a Summary Statistic (in this Case, the Mean) for Each Value of the Primary Variable



You will then see a table with the average number of cigarettes smoked per day by race instead of counts or percents. (See Figure 24.)

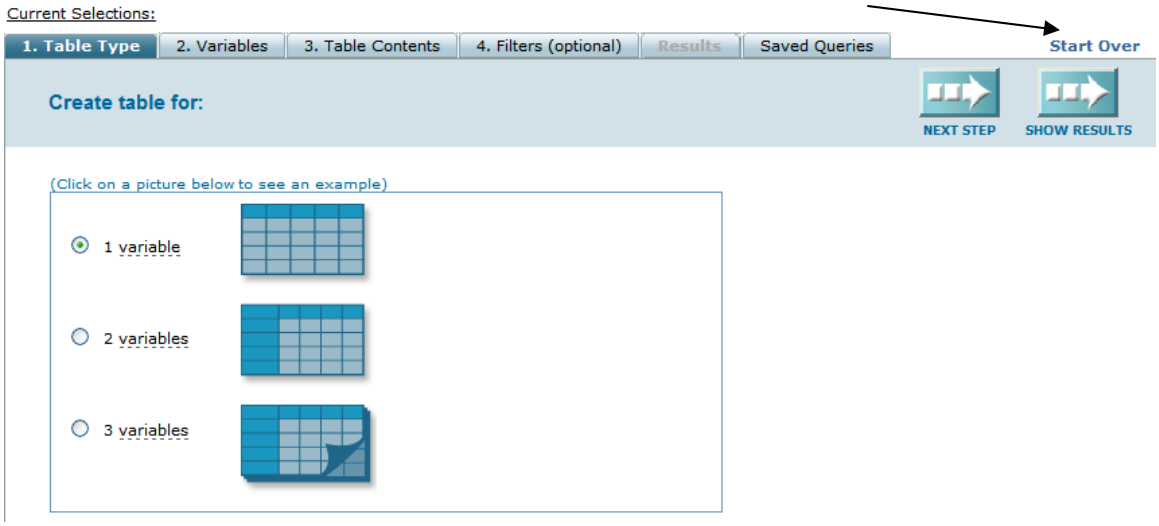
Figure 24. Mean Number of a Variable (Number of Cigarettes for Each Race)

Race (RACE)

Race (RACE)	Mean of # cigarettes smoked per day (CIGNUM)
White	5.55
Black	3.96
Other	2.68
Total	5.25

Changes made on each of the first four tabs (Table Type, Variables, Table Contents, and Filters) persist until you change them individually or reset them all to the defaults by clicking the "Start Over" button in the upper right-hand corner of the screen (see Figure 25) or by selecting the "Build a Table button on the left side of the screen.

Figure 25. Location of the "Start Over" Button



IMPORTANT

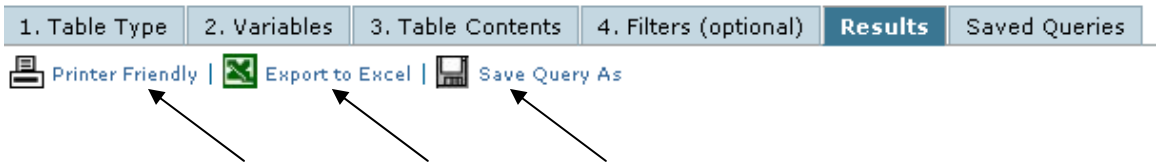
WesDax remembers your settings during a session. You can substitute another categorical variable for Race or another continuous variable for Cigarettes/Day and generate a report directly comparable to the one shown above. The decisions you previously made on such things as, whether or not to include missing values, will be kept.

If you want to start over with new variables and the default settings, you should click on the "**Start Over**" button near the top right side of the screen. (See Figure 25.)

Using Results Outside of the Current WesDaX Session

Up to this point, you have seen how to display results in table and chart format on your screen. You will often want to do other things with these results outside of WesDaX or to recreate the same table in a later session. After displaying results, three option buttons give you these abilities (see Figure 26).

Figure 26. Option Buttons Available after Displaying Results



- **The Printer Friendly** button prepares a version of your table or chart in a form that is ready for printing.

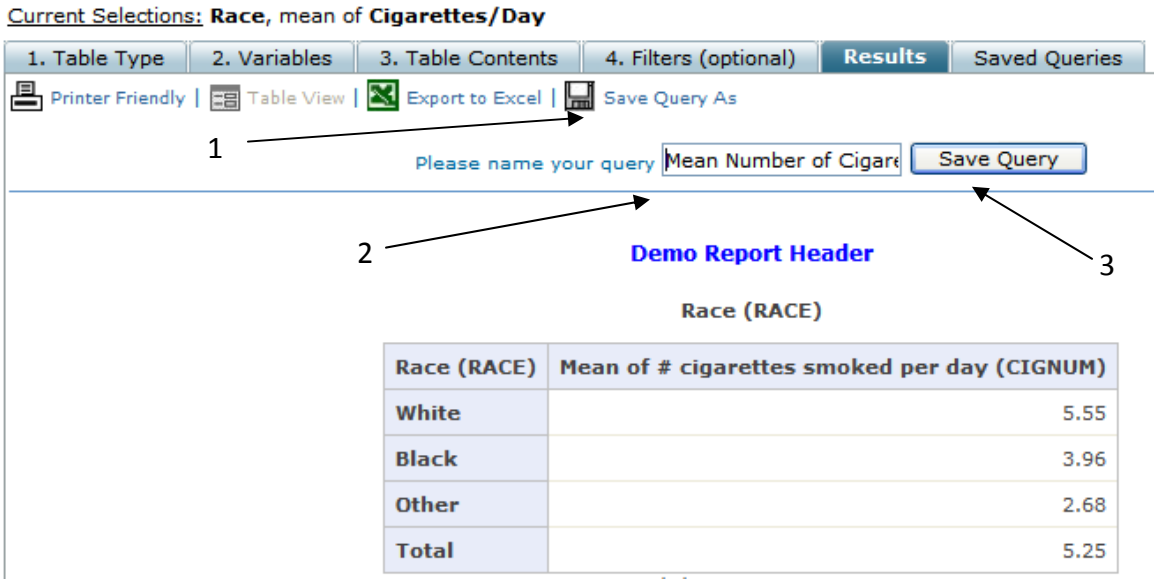
- The **Export to Excel** button creates a new Excel spreadsheet that contains the data in your table or chart. You can work with it as you would any other Excel spreadsheet. When you click on the Export to Excel button, the system gives you the option of opening or saving the Excel file.

- The **Save Query As** button lets you save a copy of your current query with a name you choose. This query will be available the next time you log in so that you can produce the same table or chart as before without having to specify your choices again. This is useful if the data are being updated on a regular basis and you need to create tables or charts using the most recent version of the data.

Saving Queries and Using Saved Queries

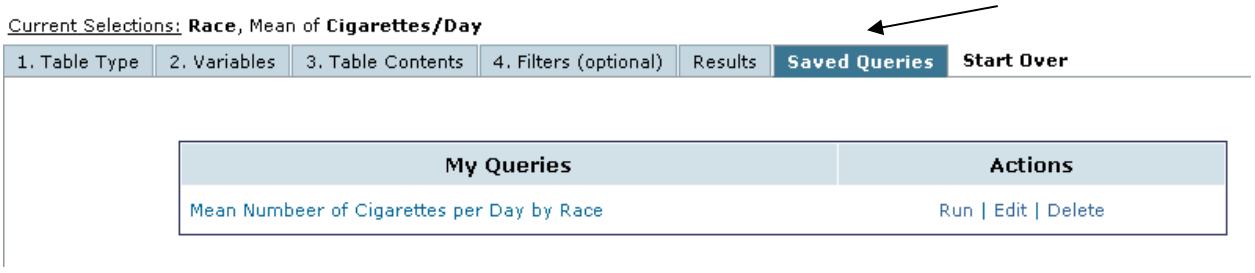
To save the query that produced Figure 24 (Mean Number of Cigarettes per Day by Race), (1) first click on the Save Query As button, then (2) type in a name for your query, and finally (3) click on "Save Query" (see Figure 27). Note that the query name can be longer than the data entry box. The system will display "Query saved successfully" in red just below the Save Query As button.

Figure 27. Steps to Save a Query



Before you have saved any queries, if you click on the "Saved Queries" tab near the top right corner of the screen, you will see the headings for saved queries, but no entries below them. After you have at least one saved query, you will see a list of queries as in Figure 28.

Figure 28. Saved Queries List



To the right of each query in the list is a choice of three actions: Run, Edit, and Delete. If you run the query, it will set all the options the same as they were when you saved the query. If the underlying data have not changed, the system will generate the same table

as before. If the data have changed, the rows and column headings will be the same, but the values in the table will reflect the new data.

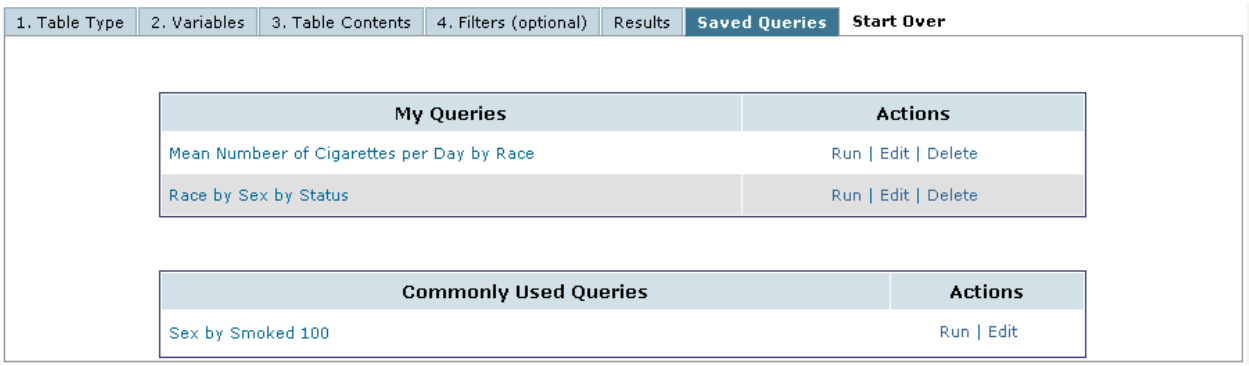
If you select "Edit," you will be able to make any of the changes to the table specifications that you could have when you originally created the table.

You can, of course, select "Delete" to remove any unneeded queries from the list of saved queries.[1]

Commonly Used Queries

Queries which you have saved appear in the My Queries list. Some projects will also have a list of Commonly Used Queries that can be seen and used by everyone on the project. Figure 29, for example, shows both a My Queries list with two queries that the user has saved and a Commonly Used Queries list with one query (Sex by Smoked 100), which was added by a project manager.

Figure 29. Sample Saved Queries Display Showing both My Queries and Commonly Used Queries Lists



You can run and edit queries in the Commonly Used Queries list just as you would queries in the My Queries list. The only difference is that if you save a query from the Commonly Used Queries list that you have run or edited, the saved version will appear in your My Queries list and the one in the Commonly Used Queries list will not be changed. Note that you do not have the ability to delete a query in the Commonly Used Queries list.

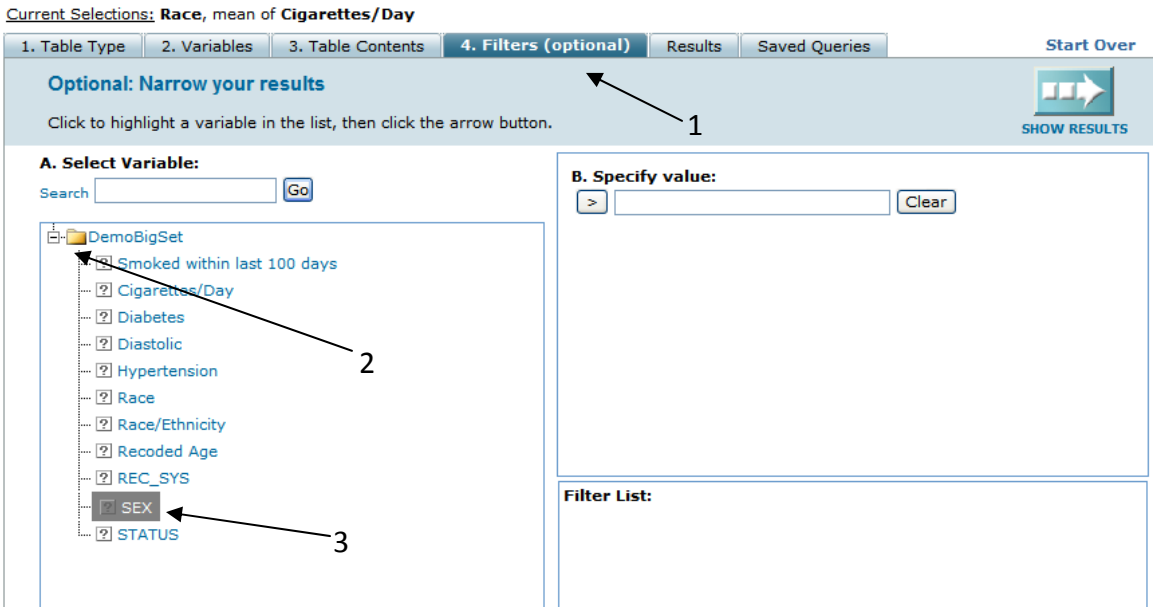
[1] If other users are logged onto the demonstration site as demouser, it is possible for them to create, edit, or delete queries on the list. When you are using a unique user ID and password, you will see only queries that you have created.

Restricting Tables to a Subset of the Data -- Filters

If you want to calculate and display data for a subset of your data, you can do so by using the optional Filters tab.

Suppose that you would like to limit the results displayed by the Mean Number of Cigarettes per day by Race query shown in Figure 24 to display the results for men only. To do so, (1) first click on the Filters tab and (2) then expand the variable list in Section A of the dialog box as shown in Figure 30 or use the search box to find the variable you want, and (3) highlight the variable that you will use for the filter criterion (in this case the variable "Sex").

Figure 30. Filters Tab



Next, specify the value of the variable to use as the filter by (4) clicking on the arrow in Section B of the dialog box. This copies the variable name into the box and displays the available values. Next, (5) select the value(s) you want to use by clicking on the corresponding box(es) as shown in Figure 31. When you next (6) click on the Add to Filter List button, the box in Section B is cleared and the selected variable and condition are added to the Filter List at the lower right of the screen as shown in Figure 32.

Figure 31. Specifying Variable Values in Section B of the Filters Tab

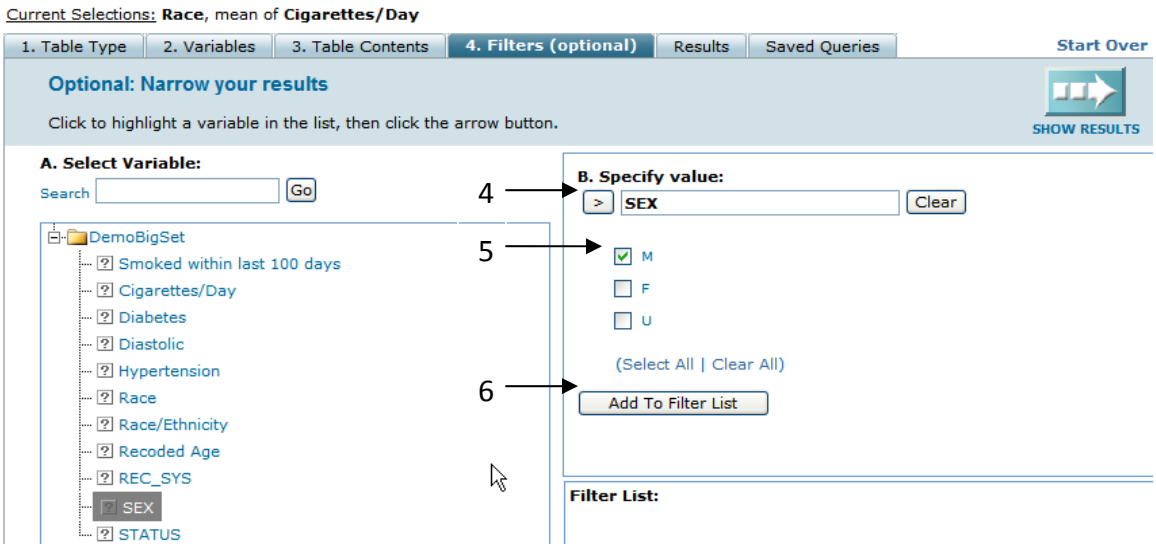
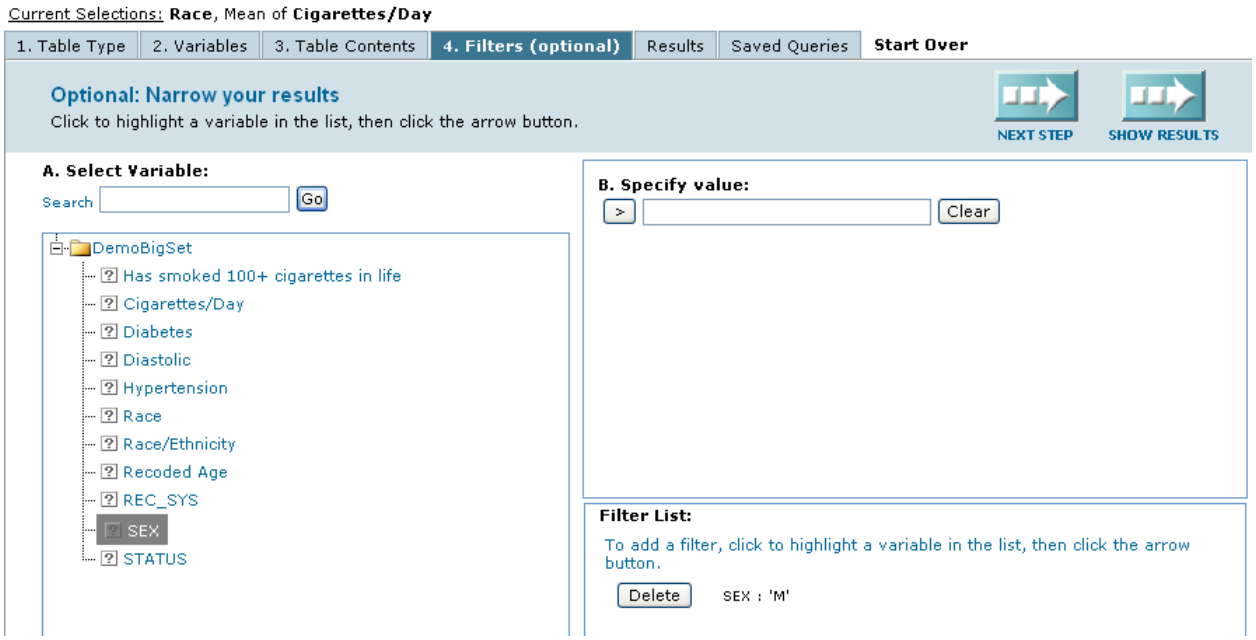


Figure 32. Filters Tab after Specifying Filter Variable and Value (Variable "Sex" with Value "M")



If you now run the query by clicking on either the Results tab or the SHOW RESULTS button, you will generate the table shown in Figure 33.

Figure 33. Mean Number of a Filtered Variable (Cigarettes for Each Race -- Males Only)

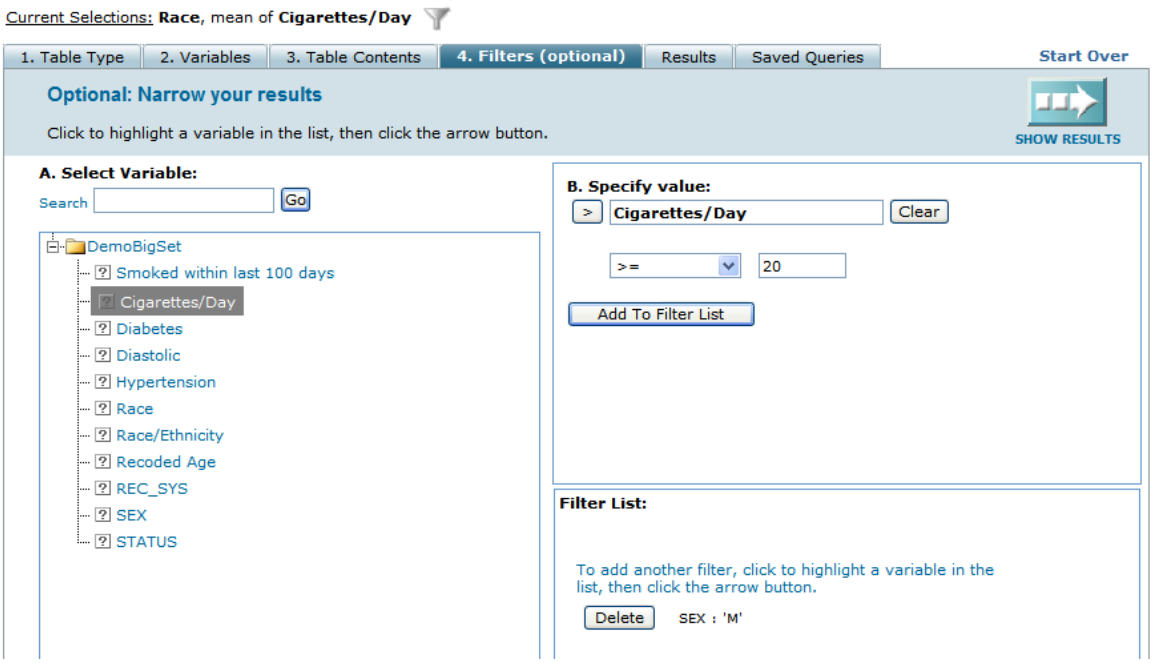
Data in table for:	
SEX:	M

Race (RACE)

Race (RACE)	Mean of # cigarettes smoked per day (CIGNUM)
White	6.51
Black	5.03
Other	4.33
Total	6.25

You can use multiple filters simultaneously. If you are interested in seeing results only for men who smoke at least a pack (20 cigarettes) per day, you would go back to the Filters tab and complete it as shown in Figure 34 and then click on the Add to Filter List button.

Figure 34. Specifying a Lower Bound for a Variable (20 Cigarettes per Day as Lower Bound)



After you click on SHOW RESULTS, the system will generate the table shown in Figure 35.

Figure 35. Mean Number of a Variable Using Multiple Filters (Number of Cigarettes for Each Race -- Men Who Smoke at Least One Pack per Day)

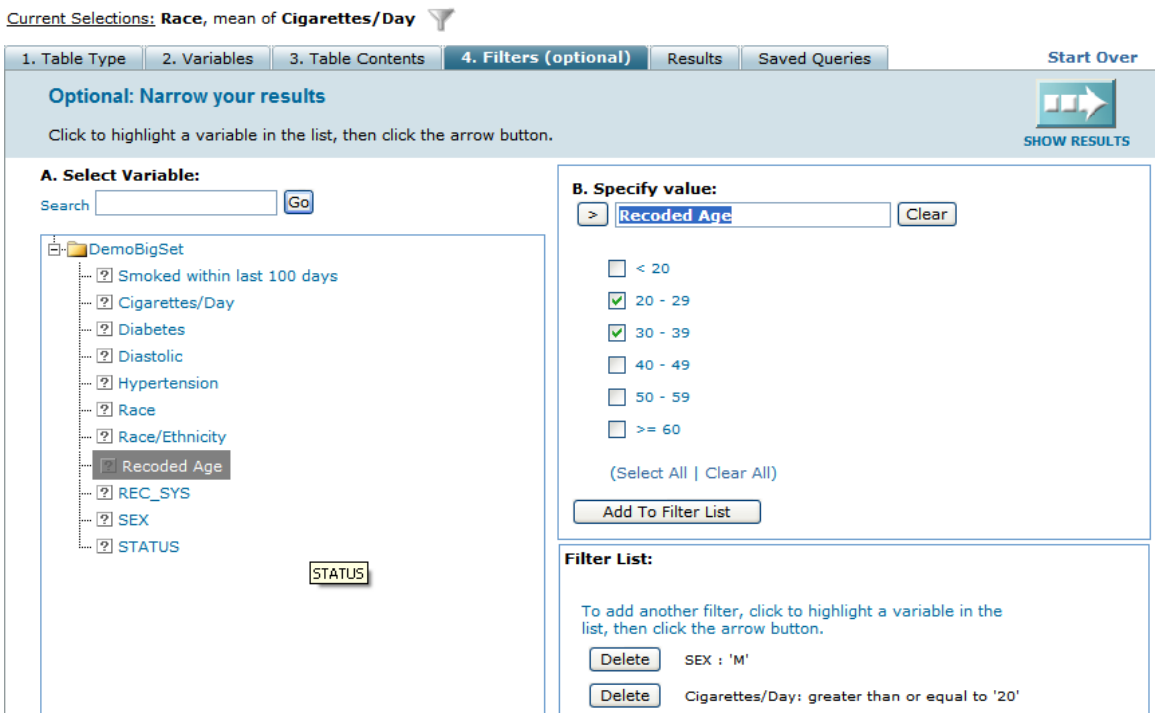
Data in table for:	
SEX:	M
Cigarettes/Day:	greater than or equal to 20

Race (RACE)

Race (RACE)	Mean of # cigarettes smoked per day (CIGNUM)
White	28.58
Black	24.03
Other	23.99
Total	28.12

Some variables allow you to select multiple values to use when filtering. In the sample data, for example, the variable Recoded Age takes the values: <20, 20-29, 30-39, 40-49, 50-59, and >= 60. You can use any combination of these values in a filter. For instance, if you want to limit the data in Figure 35 to men in the age ranges 20-29 and 30-39, select the Recoded Age variable and check both "20-29" and "30-39" as in Figure 36 and then click the "Add to Filter List" button.

Figure 36. Specifying Multiple Values for a Filter Variable (Variable Recoded Age: Values 20-29 or 30-39)



If you then click on SHOW RESULTS, WesDaX produces the table in Figure 37

Figure 37. Mean Number of a Variable Using Multiple Filters Including One with Multiple Values Selected (Variable Recoded Age: Values 20-29 or 30-39)

Data in table for:	
SEX:	M
Cigarettes/Day:	greater than or equal to 20
Recoded Age:	20 - 29 or 30 - 39

Race (RACE)

Race (RACE)	Mean of # cigarettes smoked per day (CIGNUM)
White	29.13
Black	24.81
Other	25.70
Total	28.65

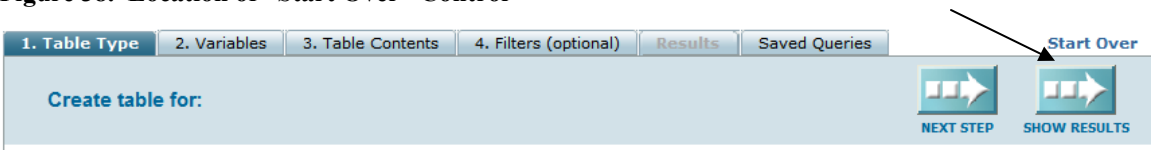
Note that the legend above the table clearly specifies the filter variables and the values used in the query that produced the table.

Recoding Categorical Variables

You can recode categorical variables by collapsing any combinations of the categories and placing the collapsed variables in any order needed.

To demonstrate recoding, first click on the words "Start Over" near the top right corner of your screen as shown in Figure 38. This resets the system to the default state.

Figure 38. Location of "Start Over" Control



Next, generate a table of mean number of cigarettes smoked by age by performing the following steps:

- Select the 1 Variable option on the Table Type tab.
- Go to the Variables tab and
  - Expand the variable list by clicking on the plus (+) sign
  - Drag the "Recoded Age" variable to the box labeled "Row"
- Go to the Table Contents tab and
  - Specify the cell contents by clicking on the words "Mean of another variable"
  - Expand the variable list by clicking on the plus (+) sign
  - Drag the "Cigarettes/Day" variable to the box to the right of the list

Click on SHOW RESULTS to generate the table shown in Figure 39.

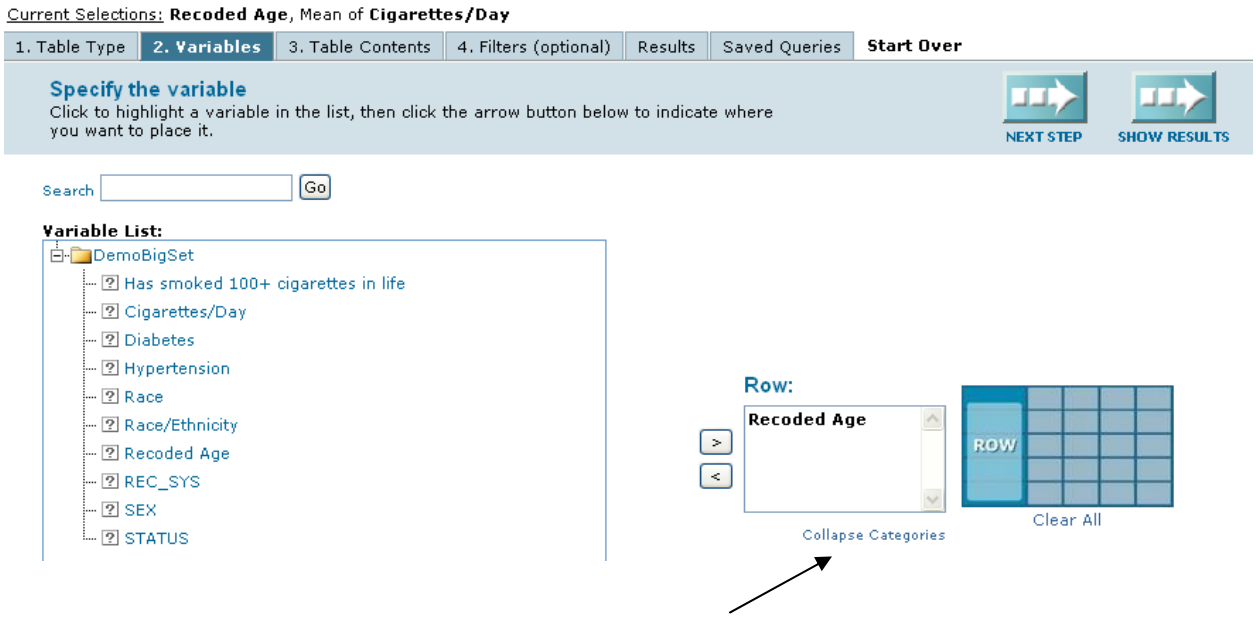
Figure 39. Mean Number of Cigarettes per Day by Age Category

Rec_age (REC_AGE)

Rec_age (REC_AGE)	Mean of # cigarettes smoked per day (CIGNUM)
< 20	4.98
20 - 29	6.35
30 - 39	6.47
40 - 49	5.95
50 - 59	4.33
>= 60	1.66
Total	5.25

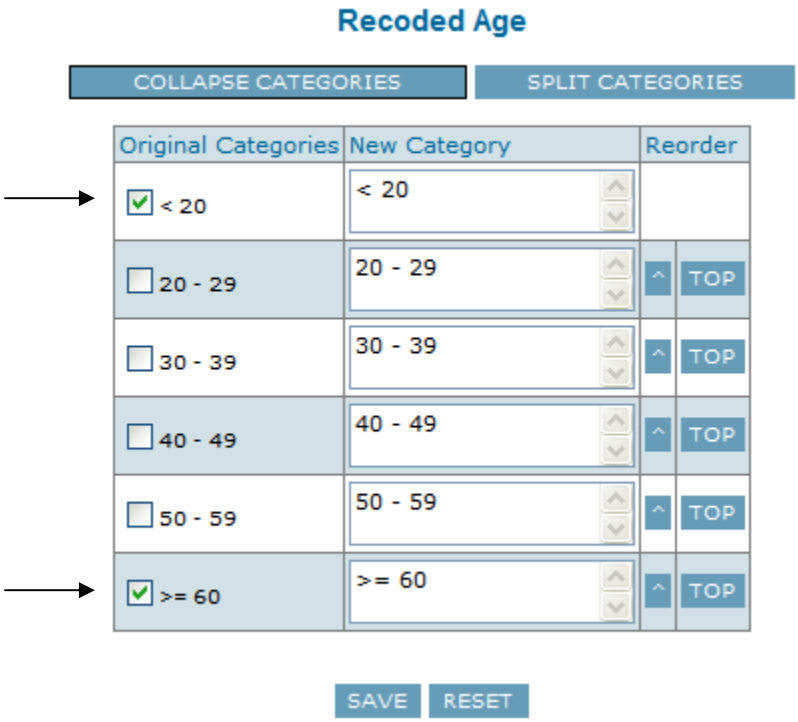
If you wanted to combine the oldest and youngest groups into one group and place everyone else in another group, you could do so on the "Variables" tab by clicking on the words "Collapse Categories" below the Row box as shown in Figure 40.

Figure 40. Location of "Collapse Category" Control on the Variables Tab



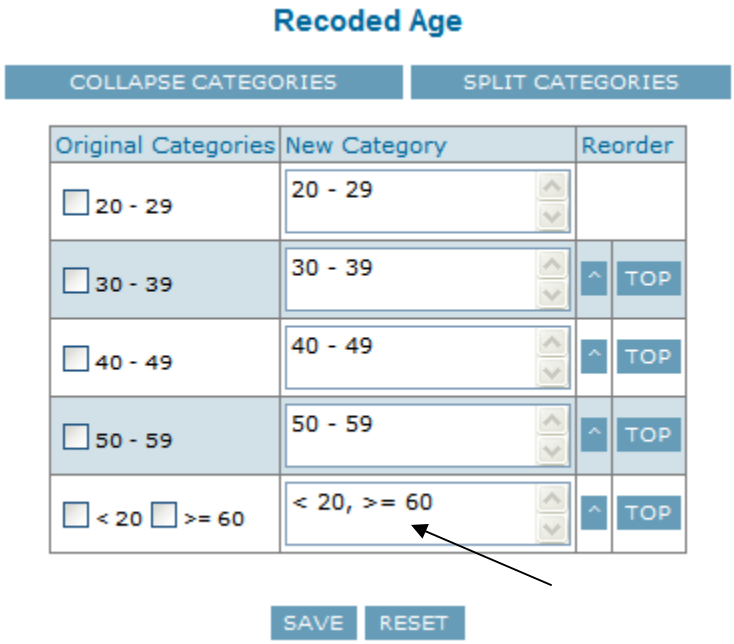
A pop-up window will open that that displays the categories of the Recoded Age variable. You can click on the check boxes next to the categories to be combined as shown in Figure 41.

Figure 41. Collapse Category Pop-Up Window with 2 Categories to be Combined Marked



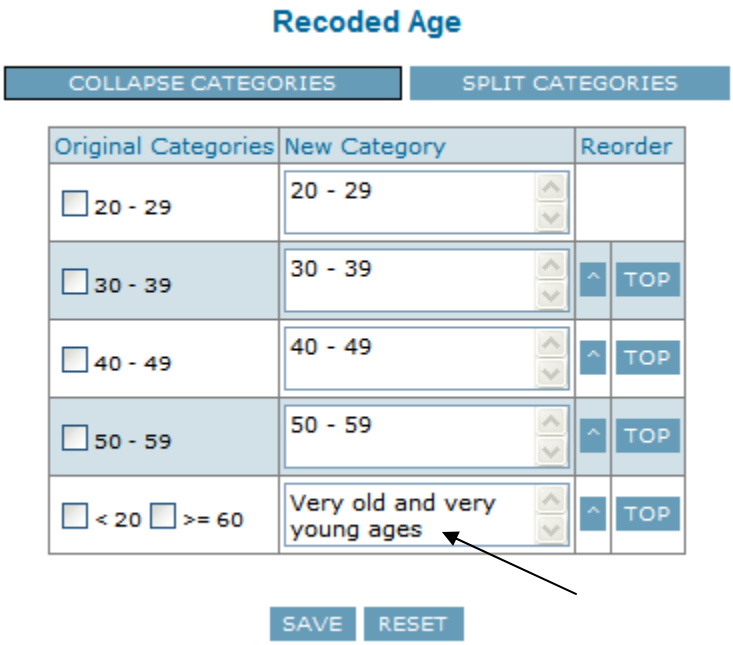
If you click on the "Collapse Categories" button, the categories you selected are combined and given a new category name made up of the original names (Figure 42).

Figure 42. Collapse Category Pop-Up Window with 2 Categories to be Combined Marked



The system automatically creates a default name for the new category based on the original categories that were combined. You can then rename the newly created category by typing over the default name assigned by the system with new text as in Figure 43.

Figure 43. Renaming New Category



When you click on the SAVE button, you will see a message in red at the bottom of the box reading "Saved successfully." If you then click on the Results tab, you will generate a table that uses your new category and category label (see Figure 44).

Figure 44. Table Generated Using New Category

Rec_age (REC_AGE)

Rec_age (REC_AGE)	Mean of # cigarettes smoked per day (CIGNUM)
20 - 29	6.35
30 - 39	6.47
40 - 49	5.95
50 - 59	4.33
Very old and very young ages	4.03
Total	5.25

You can, of course, combine other values of the same variable. For example, if you go back to the Variables Tab and again click on the words "Collapse Categories," you could check all the remaining original categories, click on the COLLAPSE CATEGORIES button, and name the new category "20 – 59." (Be sure to click the Save button.) When you then go to the Results tab or click on SHOW RESULTS, you will generate the table shown in Figure 45.

Figure 45. Revised Table after Collapsing All Categories from Age 20 thru Age 59

Rec_age (REC_AGE)

Rec_age (REC_AGE)	Mean of # cigarettes smoked per day (CIGNUM)
Very old and very young ages	4.03
20 - 59	5.97
Total	5.25

At this point, you can go back to the Variables tab and click on "Collapse Categories" under the Row box. You can now use the buttons in the "Reorder" column to adjust the order in which your recoded variables are displayed. You can also use the "SPLIT CATEGORIES" button to remove any or all of the original categories from a new category created by recoding. Simply check the items to be separated, click on SPLIT CATEGORIES, and then Save the changes.

Note that you can rename one or more categories without collapsing them by typing the new name in the "New Category" box. Figure 46 shows the table "Mean Number of Cigarettes per Day by Age Category" after relabeling "<20" to "Under 20" and ">=60" to "60 or older."

Figure 46. Mean Number of Cigarettes per Day by Age Category after Relabeling First and Last Categories

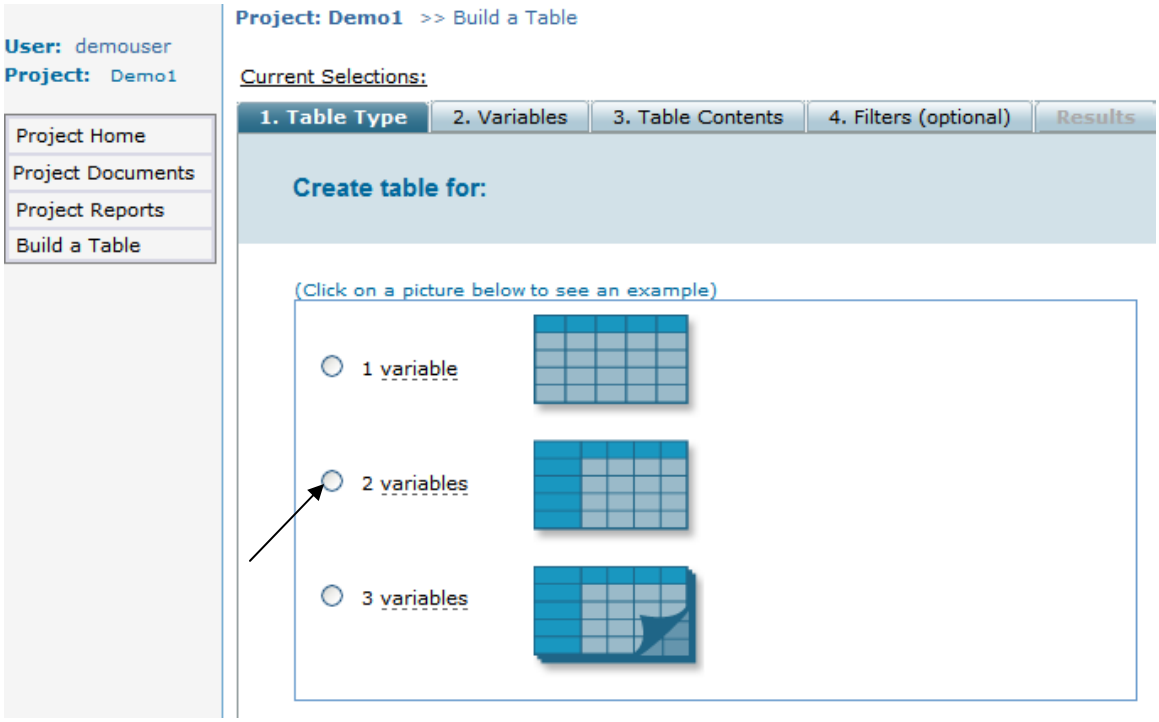
Rec_age (REC_AGE)

Rec_age (REC_AGE)	Mean of # cigarettes smoked per day (CIGNUM)
Under 20	4.98
20 - 29	6.35
30 - 39	6.47
40 - 49	5.95
50 - 59	4.33
60 or older	1.66
Total	5.25

Two Variable Tables

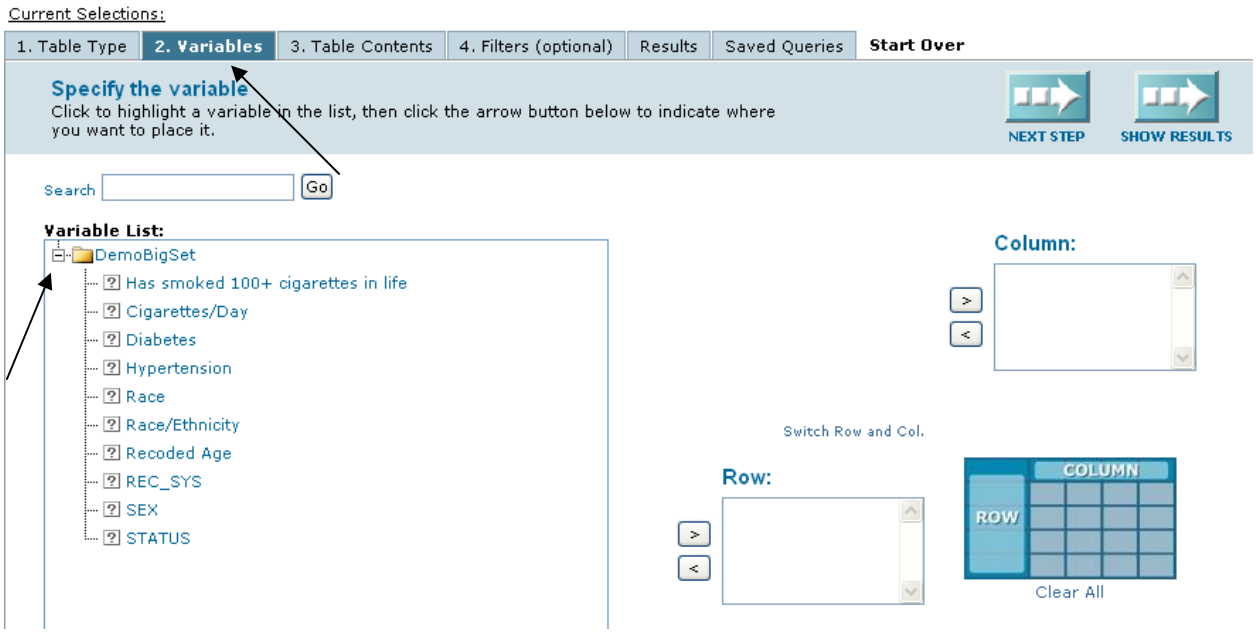
To create a table with two variables, first click on "Start Over" to ensure that you are starting with the default settings. Then select the 2-variable button as shown in Figure 47.

Figure 47. Selecting Two Variable Table



Next, go to Tab 2, "Variables," and expand the variable list as shown in Figure 48. The search function is available if you prefer to use it rather than seeing the full list of variables.

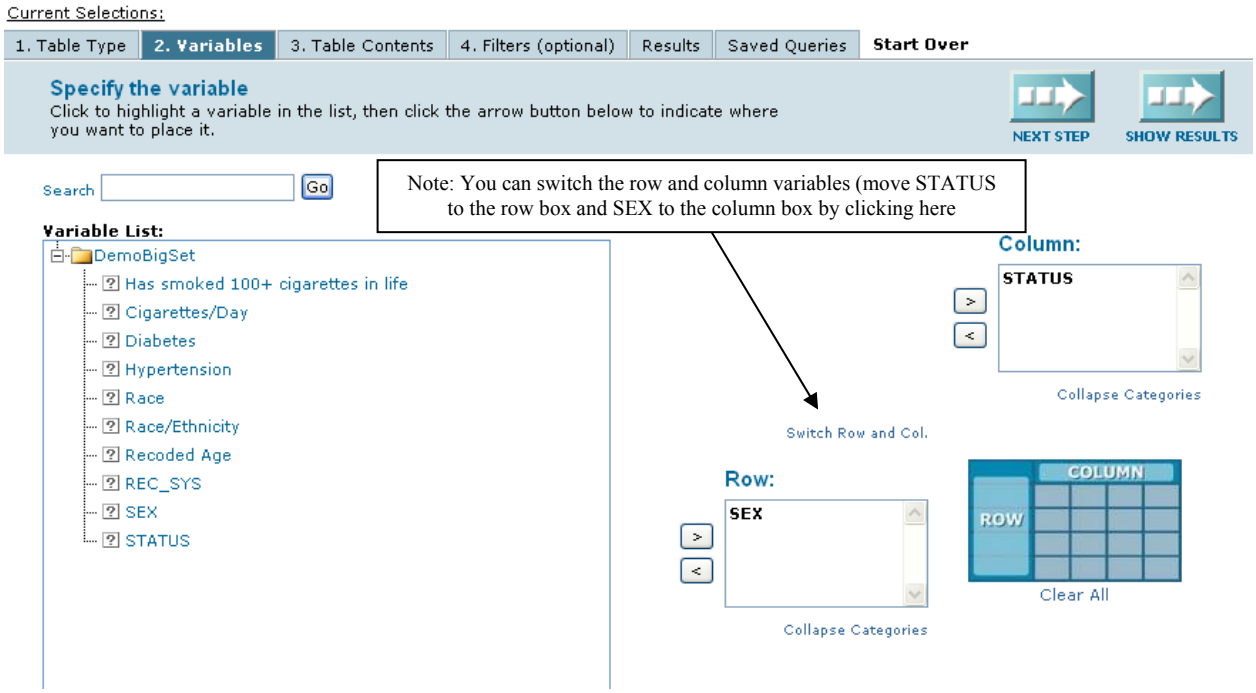
Figure 48. Expanded Variable List When Specifying a Two Variable Table



You now need to decide which variable you want to display in the table row and which variable you want to display in the table columns. If you are not sure, try it different ways and see which display format is the most informative to you.

To create a table that shows the interview completion status for males and females, drag the SEX variable to the box labeled "Row" and drag the STATUS variable to the box labeled "Column." The screen will now look like Figure 49:

Figure 49. Specifying a Two Variable Table (Sex by Status)



Click on the SHOW RESULTS button in the top right corner to generate your table. It should look like Figure 50.

Figure 50. Two Variable Table (Sex by Status)

"Sex (SEX)" by "Examination/interview status (STATUS)"

Number of Responses	Examination/interview status (STATUS)			
Sex (SEX)	Complete	Incomplete	Other	Total
Male	9,331,259	79,442,886	863,396	89,637,541
Female	9,674,902	86,634,910	1,699,853	98,009,665
Total	19,006,161	166,077,795	2,563,250	187,647,206

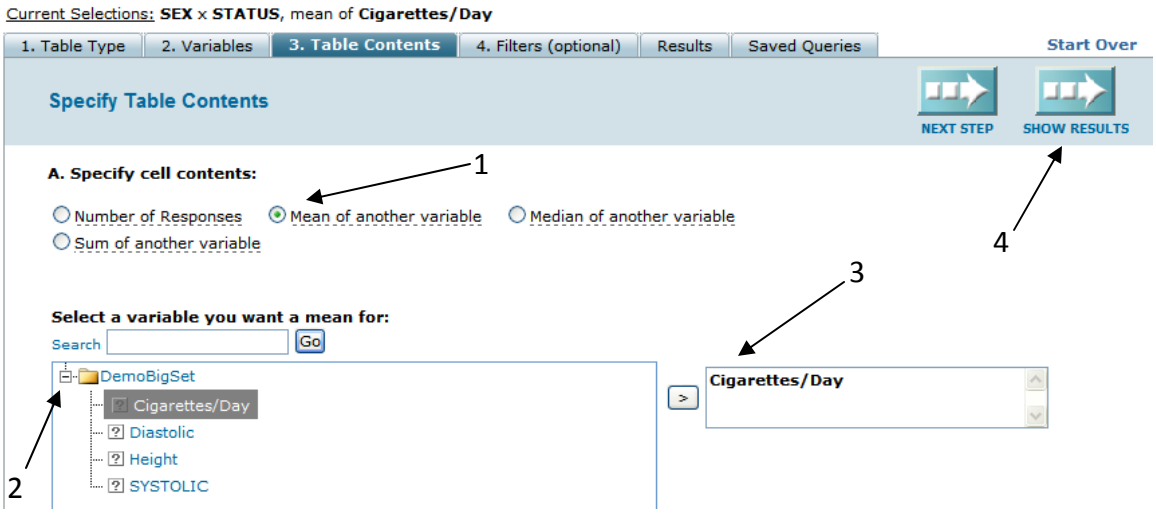
You can choose to display a summary statistic for some variable corresponding to each cell in the table instead of displaying the number of people in a cell or instead of displaying a percentage in each cell.

To display a summary statistic, you do the same steps as for tables with one variable. For example, to display values for the variable "Cigarettes per Day," follow the steps below and as shown in Figure 51.

1. Go to the Table Contents tab. In Section A check the statistic you want (in this case "Mean of another variable").
2. Expand the variable list by clicking on the plus (+) sign next at the top of the box that displays variable names

3. Drag the variable name of interest to the box on the right; and
4. Click on SHOW RESULTS.

Figure 51. Specifying a Summary Statistic in a Two-Way Table



After clicking on SHOW RESULTS, you will get a table that looks like Figure 52.

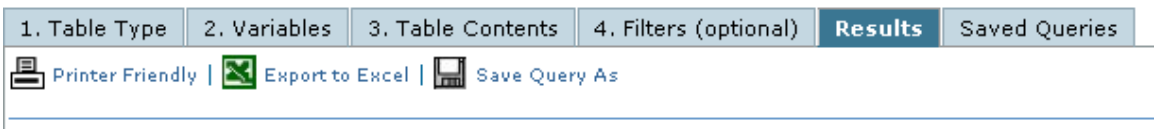
Figure 52. Summary Statistics in a Two-Way Table (Mean Number of Cigarettes Smoked per Day for Each Sex by Status)

"Sex (SEX)" by "Examination/interview status (STATUS)"

Mean of # cigarettes smoked per day (CIGNUM)	Examination/interview status (STATUS)			
Sex (SEX)	Complete	Incomplete	Other	Total
Male	6.43	6.24	5.81	6.25
Female	4.10	4.38	2.81	4.32
Total	5.24	5.27	3.82	5.25

The "Printer Friendly" button and the "Export to Excel" button shown at the top of the Results tab (see Figure 53) work the same way for tables with two or three variables as they do for tables with one variable.

Figure 53. Display and Output Options Buttons

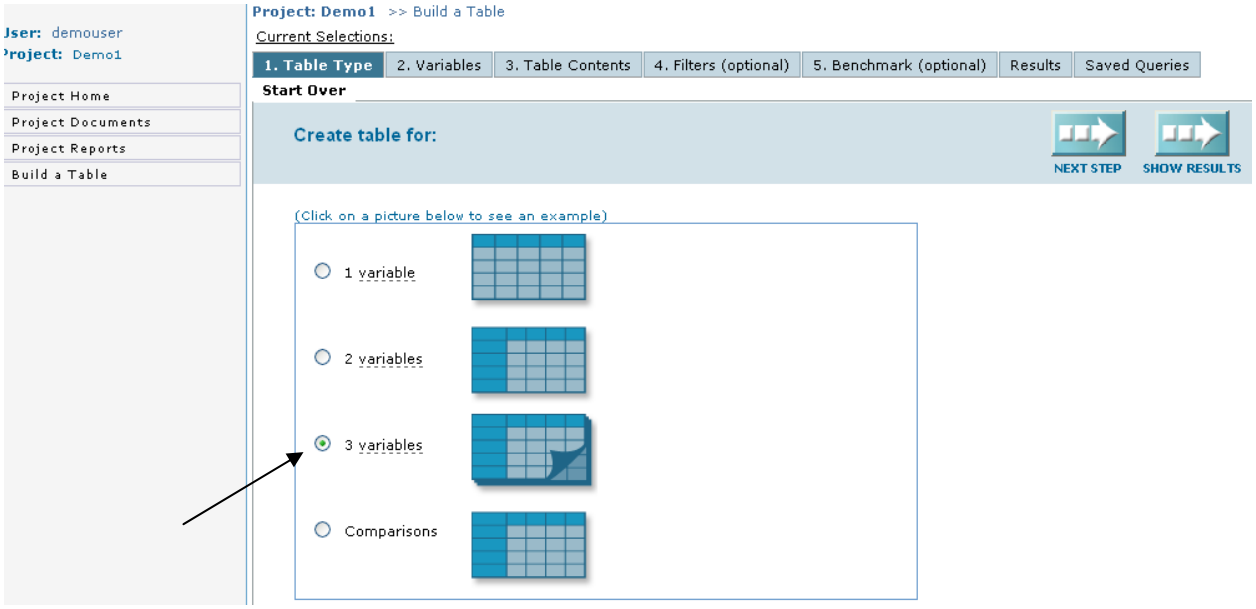


Three Variable Tables

Creating tables with three variables is similar to creating them for one and two variables. The output is displayed as a series of two-way tables, with one table for each of the values of the third variable. If you choose to print this output, each table is printed on a separate page. When choosing the "Page variable," it is usually best to select one that has a relatively small number of distinct values.

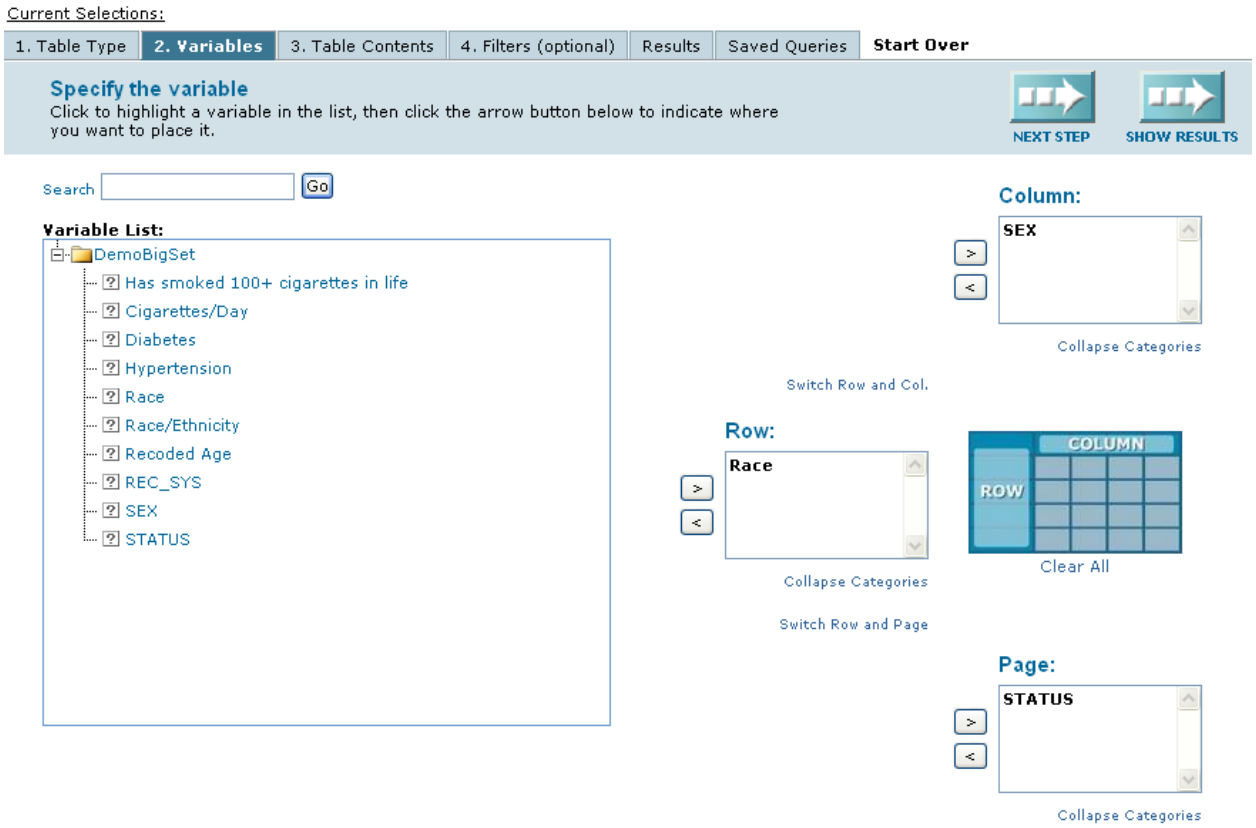
For example, assume that you are interested reviewing the distribution of race by sex for people in different stages of study completion. The sample database variable "Status" has 3 values: Complete Incomplete, and Other. To generate a Race by Sex table for each of these three status values, first click on "Start Over" to ensure that you are starting with the default settings. Then select the 3-variable button as shown in Figure 54.

Figure 54. Selecting Three Variable Table



Next, click on the Variables tab and specify the three variables to use as shown in Figure 55.

Figure 55. Specifying Variables for a Three-Way Table



Note that since each of the variables is categorical, a separate Collapse Categories option is available for each. Also note that in addition to being able to switch the row and column variables in the same manner as for two-way tables, you can also switch the row and page variables by clicking on the words "Switch Row and page."

Clicking on SHOW RESULTS at this point will produce the output shown in Figure 56. As noted above, if you select the Printer Friendly option and print this output, each of the two-way tables will be printed on a separate sheet of paper.

Figure 56. Three-Way Table (Race by Sex by Status)

Demo Report Header

Controlling for "Examination/interview status (STATUS)" = "Complete"
"Race (RACE)" by "Sex (SEX)"

Number of Responses	Sex (SEX)		
Race (RACE)	Male	Female	Total
(missing)	2,035	2,604	4,640
White	8,231,386	8,660,595	16,891,981
Black	777,769	756,544	1,534,312
Other	320,069	255,159	575,228
Total	9,331,259	9,674,902	19,006,161

4/1/2010 11:28:10 AM

Demo Report Footer

Demo Report Header

Controlling for "Examination/interview status (STATUS)" = "Incomplete"
"Race (RACE)" by "Sex (SEX)"

Number of Responses	Sex (SEX)		
Race (RACE)	Male	Female	Total
(missing)	5,968	3,456	9,424
White	66,992,003	71,921,526	138,913,528
Black	8,837,979	11,163,634	20,001,613
Other	3,606,936	3,546,294	7,153,231
Total	79,442,886	86,634,910	166,077,795

4/1/2010 11:28:10 AM

Demo Report Footer

Demo Report Header

Controlling for "Examination/interview status (STATUS)" = "Other"
"Race (RACE)" by "Sex (SEX)"

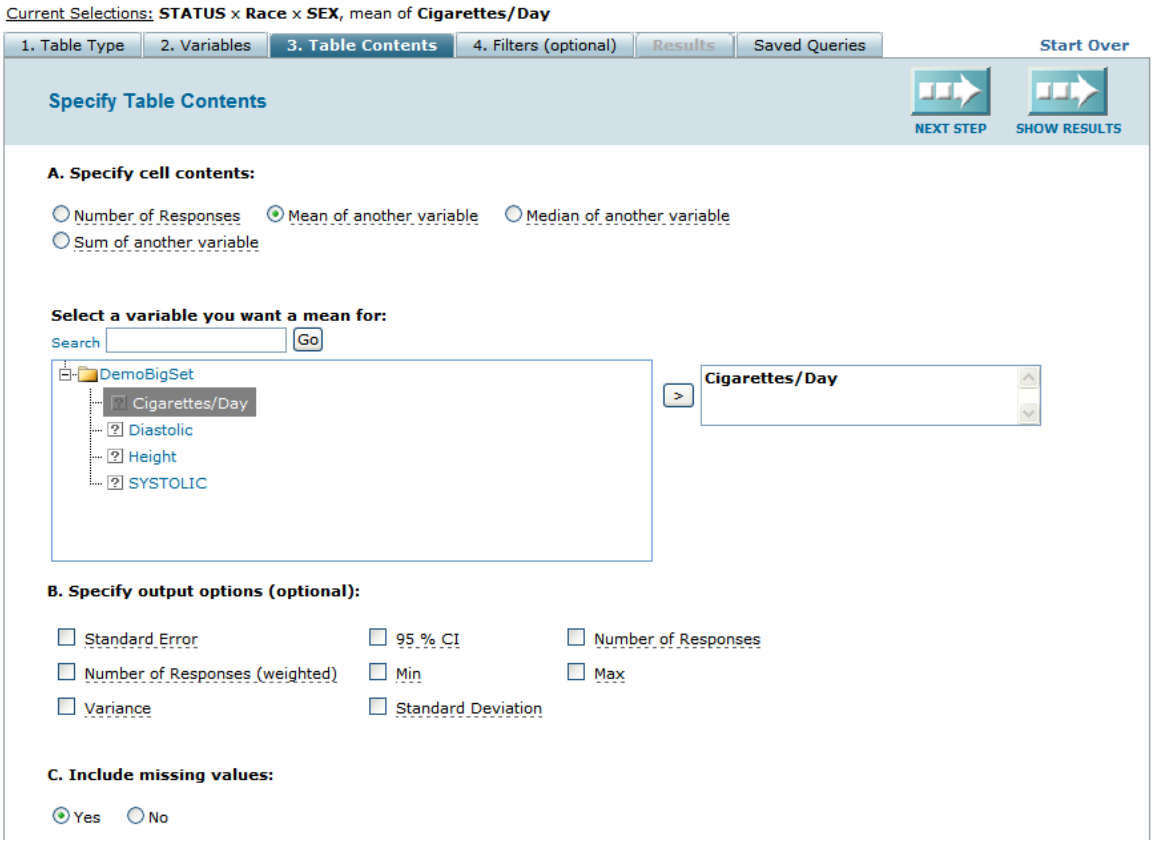
Number of Responses	Sex (SEX)		
Race (RACE)	Male	Female	Total
(missing)	.	.	.
White	788,997	1,536,620	2,325,617
Black	55,476	136,686	192,162
Other	18,923	26,548	45,471
Total	863,396	1,699,853	2,563,250

4/1/2010 11:28:10 AM

Demo Report Footer

Other WesDaX options are also available for three way (variable) tables. If instead of displaying the counts of the number of people in each cell, you want to display the mean number of cigarettes smoked per day, select the Table Contents tab and complete it as shown in Figure 57.

Figure 57. Specifying Cell Contents as Mean Number of a Variable (Cigarettes per Day)



The results generated by running this modified query are shown in Figure 58.

Figure 58. Three-Way Table with Cells Containing the Mean Number of a Variable (Cigarettes Smoked per Day)

Demo Report Header

Controlling for "Examination/interview status (STATUS)" = "Complete"
"Race (RACE)" by "Sex (SEX)"

Mean of # cigarettes smoked per day (CIGNUM)	Sex (SEX)		
Race (RACE)	Male	Female	Total
(missing)	0.00	0.00	0.00
White	6.56	4.33	5.41
Black	4.73	2.16	3.46
Other	7.21	2.00	4.90
Total	6.43	4.10	5.24

4/1/2010 4:59:20 PM

Demo Report Footer

Demo Report Header

Controlling for "Examination/interview status (STATUS)" = "Incomplete"
"Race (RACE)" by "Sex (SEX)"

Mean of # cigarettes smoked per day (CIGNUM)	Sex (SEX)		
Race (RACE)	Male	Female	Total
(missing)	2.64	0.00	1.67
White	6.51	4.74	5.59
Black	5.06	3.18	4.01
Other	4.00	0.90	2.47
Total	6.24	4.38	5.27

4/1/2010 4:59:20 PM

Demo Report Footer

Demo Report Header

Controlling for "Examination/interview status (STATUS)" = "Other"
"Race (RACE)" by "Sex (SEX)"

Mean of # cigarettes smoked per day (CIGNUM)	Sex (SEX)		
Race (RACE)	Male	Female	Total
(missing)	.	.	.
White	5.70	2.92	3.86
Black	3.15	2.14	2.43
Other	18.05	0.00	7.51
Total	5.81	2.81	3.82

4/1/2010 4:59:20 PM

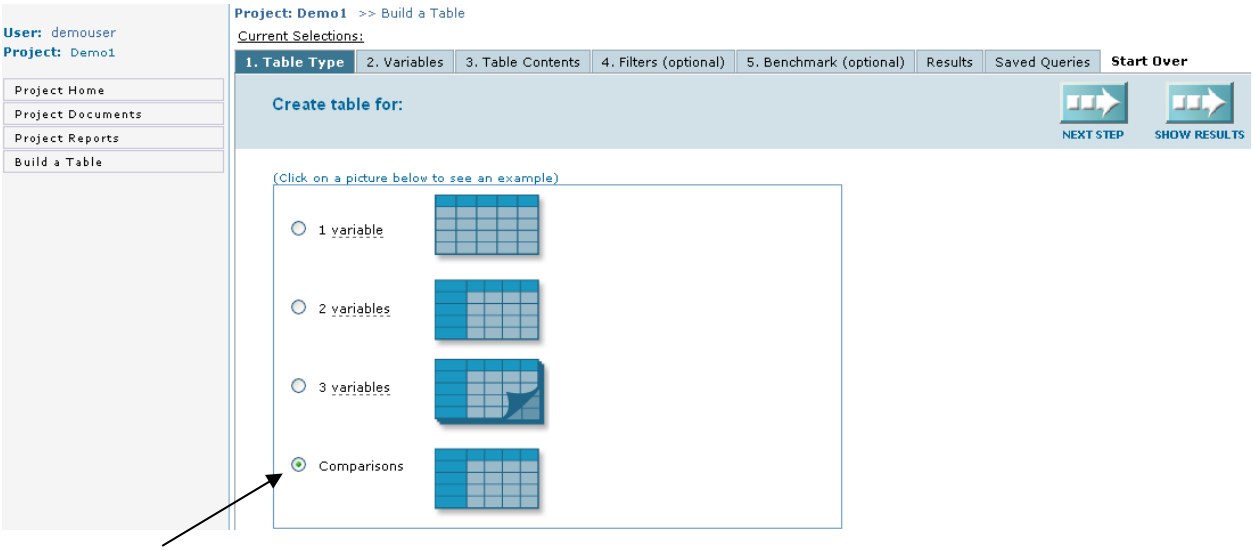
Demo Report Footer

Comparisons

Comparison tables let you display the mean values of continuous variables over different values of a categorical variable. You can also use them to perform t-tests of the difference of the means of pairs of values that you select.

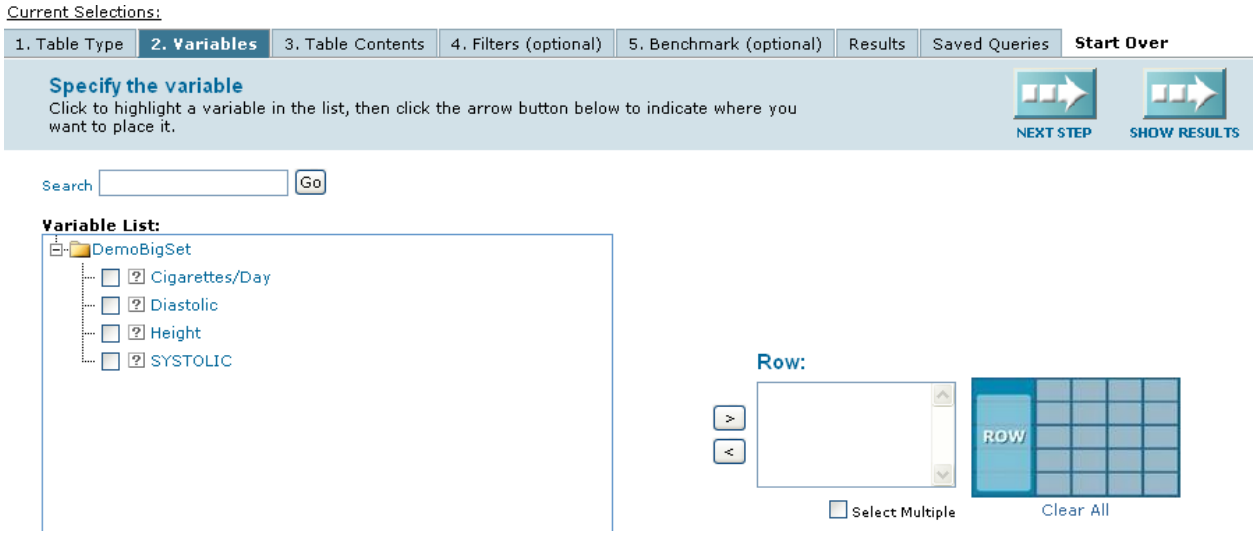
For example, assume that you are interested reviewing the distribution of race by sex for people in different stages of study completion. The sample database variable "Status" has 3 values: Complete Incomplete, and Other. To generate a Race by Sex table for each of these three status values, first click on "Build a Table" or "Start Over" to ensure that you are starting with the default settings. Then select the Comparisons button as shown in Figure 59.

Figure 59. Selecting Comparisons Table



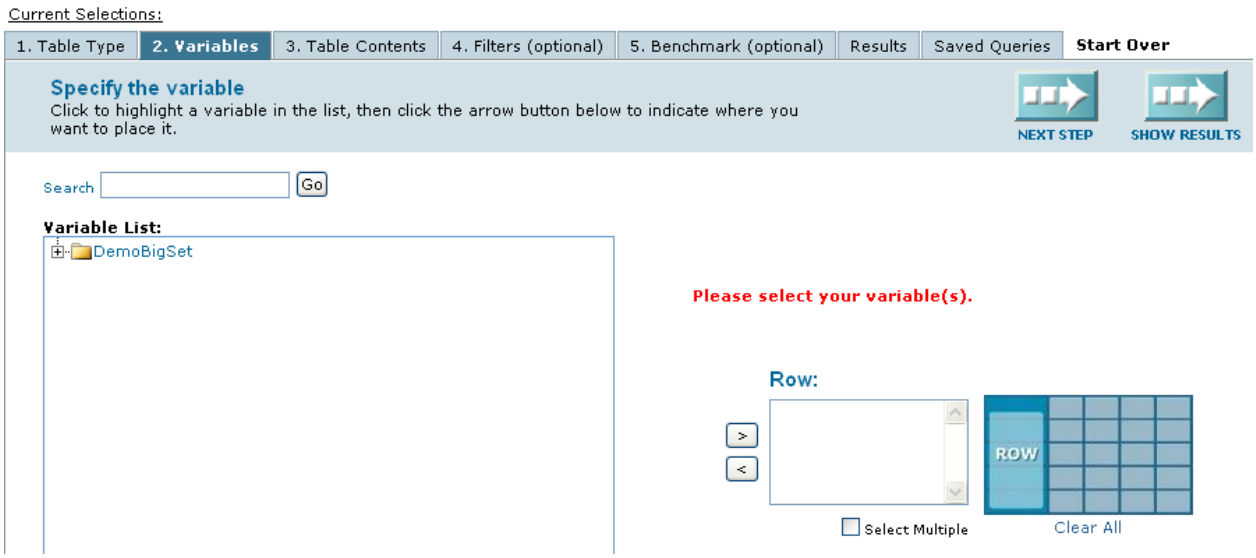
Next seledct the Variables tab and expand the list of variables. Note that only continuous variables are displayed as these are the only items that can be selected for the rows of a comparsions table.

Figure 60. Variables Available for Comparisons Table



If you select Cigarettes per Day and click the "Show Results" button, you will be taken to the Contents tab and prompted to select the categorical variable to use for the columns as shown in Figure 61.

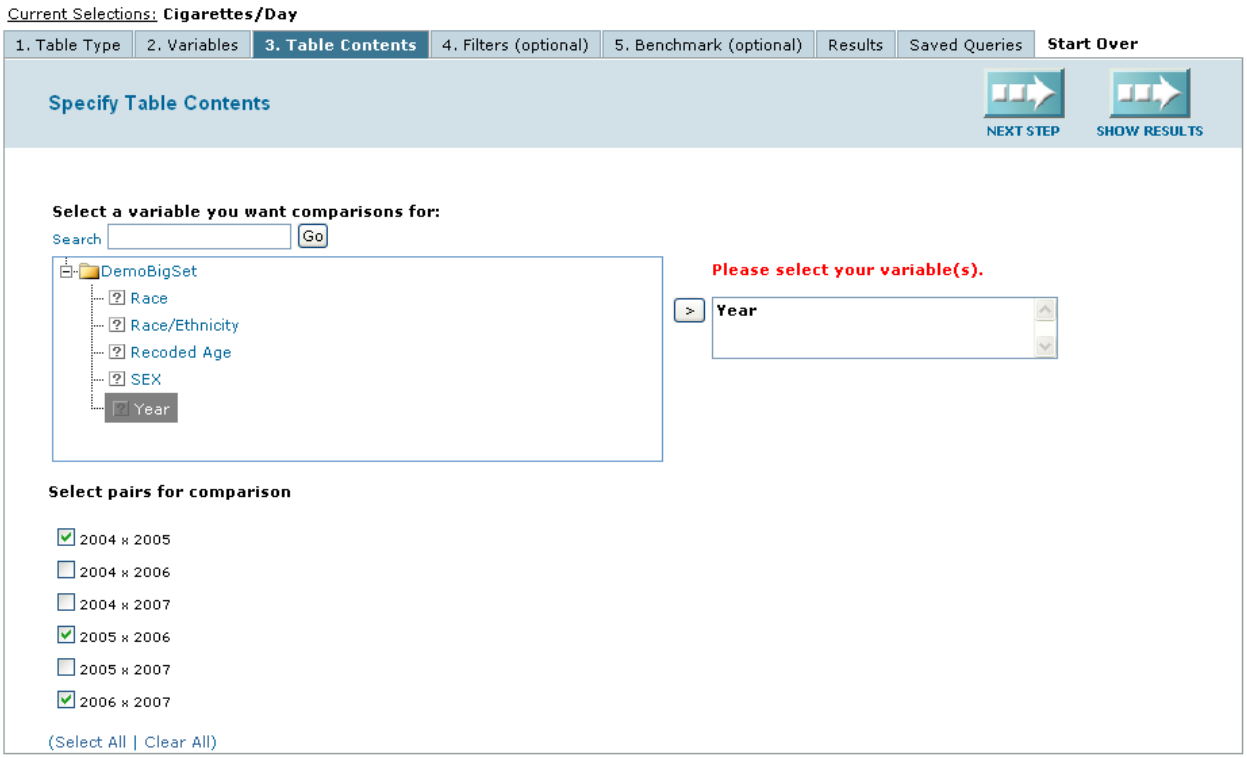
Figure 61. Prompt for Column Variable after Selecting Cigarettes/Day



After expanding the variable list and selecting a variable, the pairs of values for which you can run t-tests are displayed. You can pick one or more of these pairs for display in the final table. In

Figure 62, the Year variable has been selected and the year pairs 2004-2005, 2005-2006, and 2006-2007 have been selected. You, of course, go directly to the Contents tab after selecting a variable on the Variables tab.

Figure 62. Selection of Pairs of Values for Comparisons



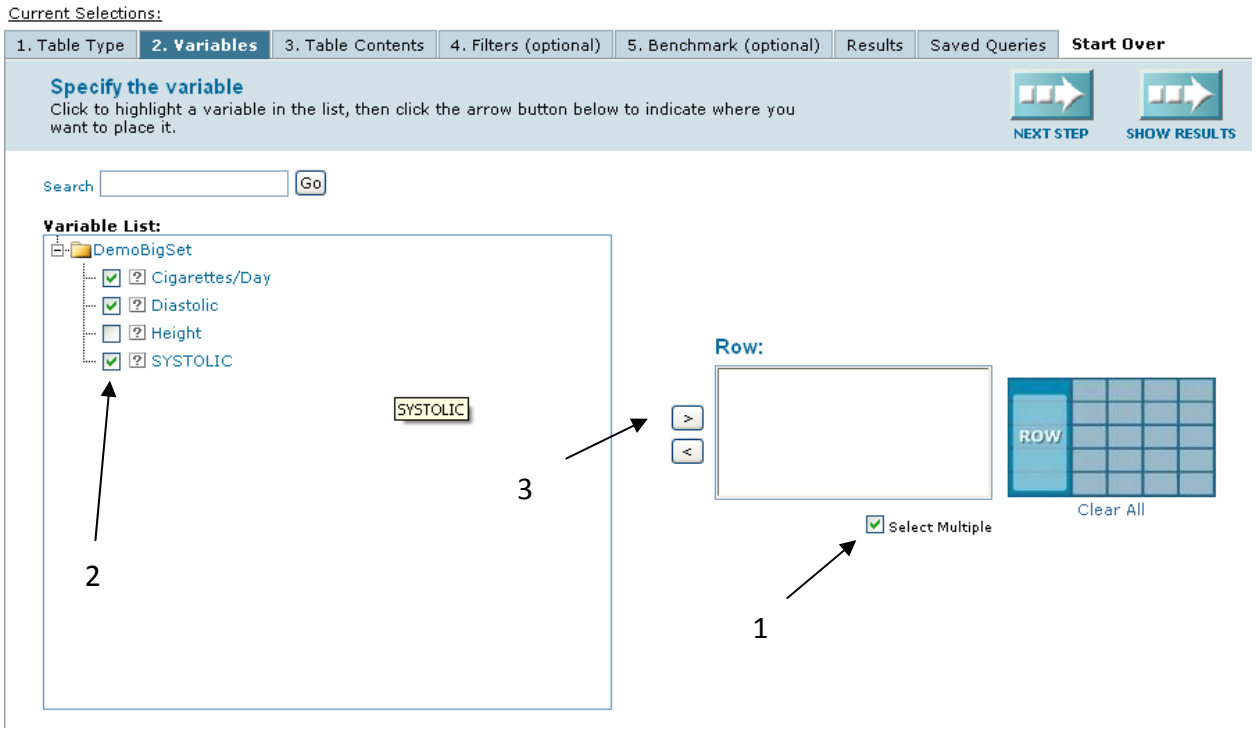
The table resulting from these selections is shown in Figure 63, which shows that none of the yearly changes in the mean number of cigarettes smoked per day were significant in this sample.

Figure 63. Comparison Table of Cigarettes per Day over 4 Years

	Year (YEAR)						
	Mean				p-value		
Question	2004	2005	2006	2007	2004 x 2005	2005 x 2006	2006 x 2007
# cigarettes smoked per day	5.41	5.14	5.38	5.05	0.4251	0.5235	0.3240

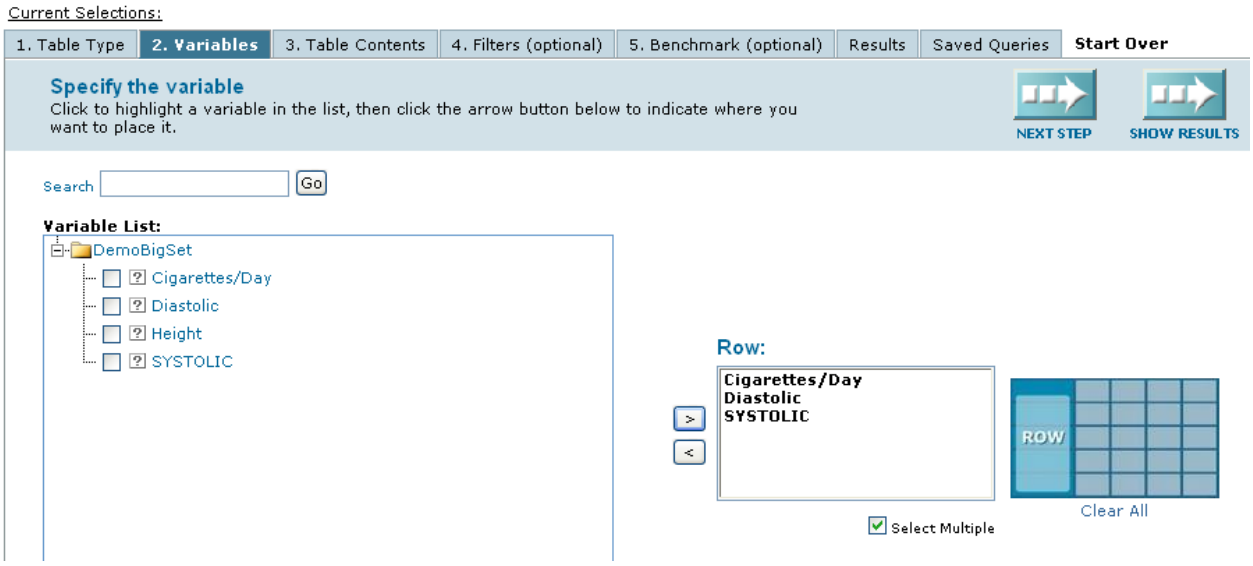
You can choose to examine multiple continuous variables at the same time. For example, if you are interested in comparing smoking measures, diabetes, and hypertension for Hispanics as compared to those classified as “Other”, you can click on “Start Over,” select the Comparisons table, and go to the Variables tab. If you then (1) check the “Select Multiple” box below the area for specifying row variables, when you expand the variable list, a set of check boxes appears next to each variable name. You can (2) choose which of these continuous measures you want to use. In Figure 64, three of the available variables have been checked.

Figure 64. Specifying Multiple Row Variables – Part 1



After you click on the right arrow (item 3 in Figure 64), the selected variables are copied to the box labeled "Row" as shown in Figure 65.

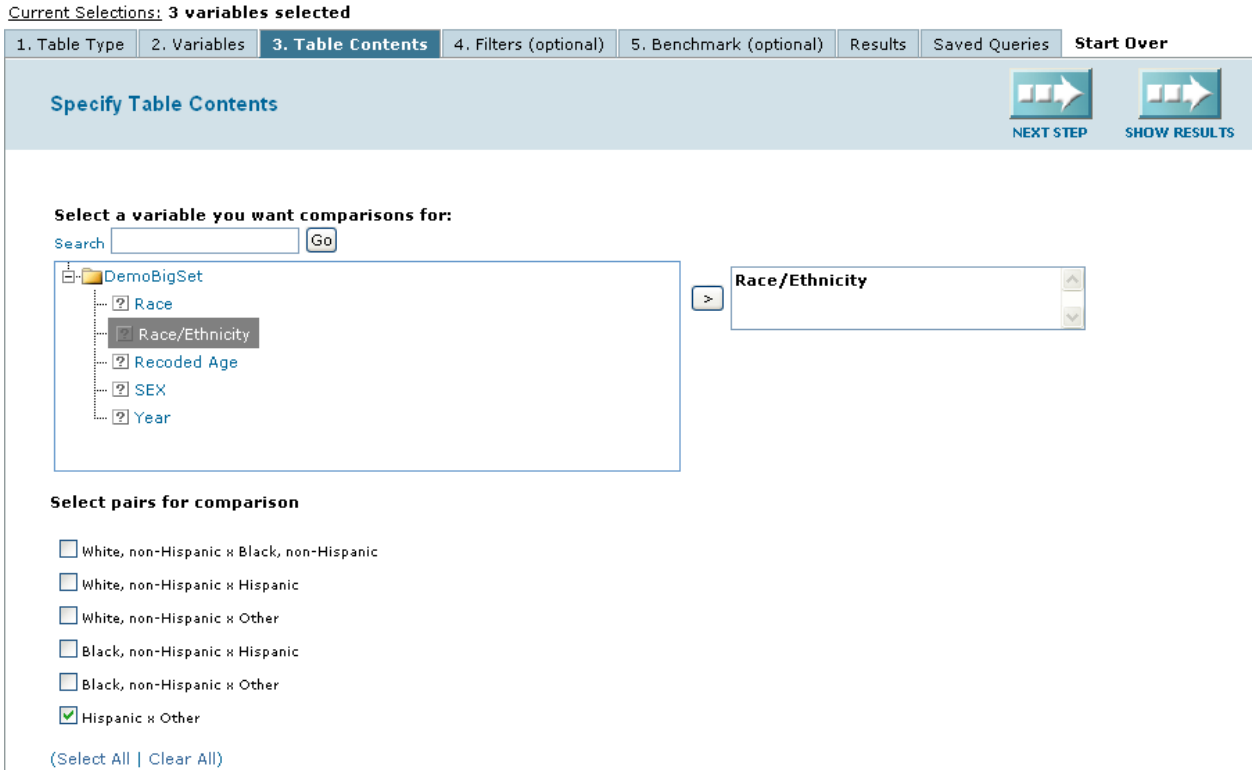
Figure 65. Specifying Multiple Row Variables – Part 2



Note that the order that the variables appear in this list is the order in which they will appear in the results table. You can control the order by adding the variables one at a time in the "Row" box.

After selecting the row variables, go to the Table Contents tab, select the Race/Ethnicity variable, and check the comparison to test, in this case Hispanic X Other (see Figure 66).

Figure 66. Specifying Comparison to Test



The resulting table (Figure 67) indicates that the Hispanic group has significantly fewer diagnoses of diabetes than the "Other" group and that Hispanics smoked fewer cigarettes per day than the "Other" group ($p < 0.05$). The mean blood pressure readings were not significantly different between the two groups ($p = 0.51$ and $p = 0.15$).

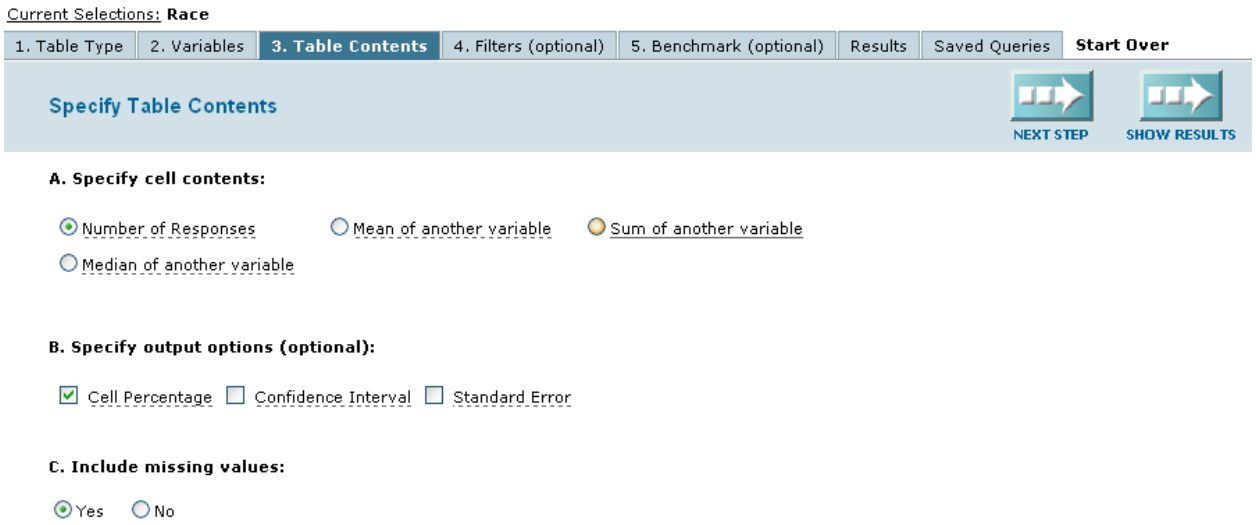
Figure 67. Comparison of Hispanic to Other on Four Measures

	Race-ethnicity (RACEETHN)				
	Mean				p-value
Question	White, non-Hispanic	Black, non-Hispanic	Hispanic	Other	Hispanic x Other
# cigarettes smoked per day	5.91	4.00	1.86	2.79	0.0136
Average Diastolic BP	75.25	76.21	73.62	74.16	0.5062
Average Systolic BP	124.23	124.15	121.03	119.48	0.1514

Benchmarks

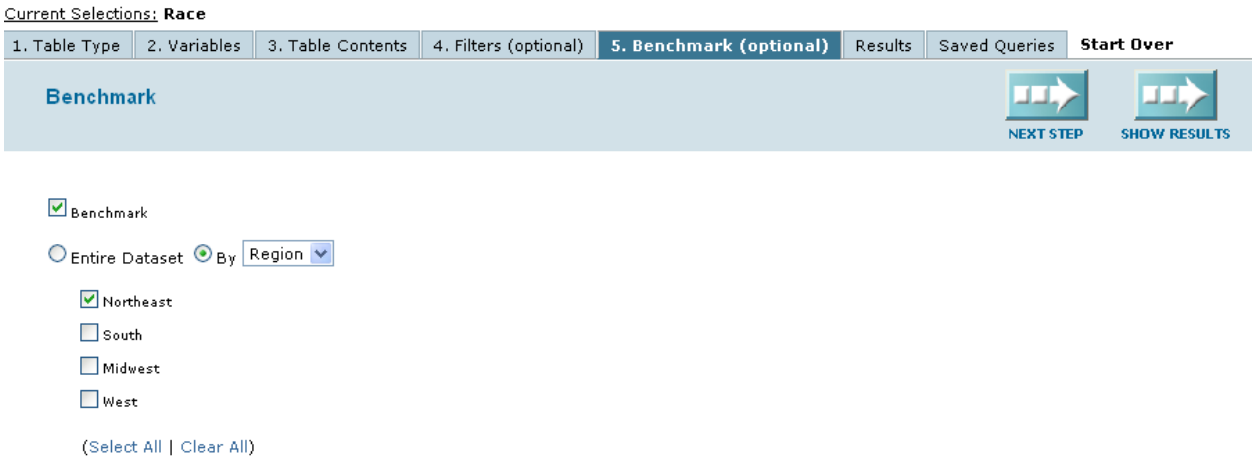
Benchmarking allows you to display values from the full set of records satisfying the filtering conditions to a selected subset of these records. For example, if you would like to compare the distribution of races in the Northeast to that of the entire country you could select a 1-variable table, specify Race as the row variable, and then on the Table Contents tab check the "Number of Responses" button and the "Cell Percentages" check box as shown in Figure 68.

Figure 68. Specifying Number of Responses and Cell Percentages as Table Contents.



Next, go to the Benchmark tab, check the "Benchmark" check box and the "By" radio button. Then select Region from the drop down list and check the "Northeast" check box. The tab should appear as in Figure 69.

Figure 69. Benchmark Tab Configured for Comparing the Northeast to the Entire Dataset



Clicking on the "Results" tab or the "Show Results" button will produce the table shown in Figure 70.

Figure 70. Distribution of Races for Northeast Benchmarked to Entire Dataset

Data in table for:	
Your Dataset:	
No filter(s) selected.	
Benchmark Dataset:	
Region:	Northeast

		Your Dataset		*Benchmark Dataset*	
		Number of Responses	Cell Pct	Number of Responses	Cell Pct
Race (RACE)	(missing)	14,063.3	0.0%	8,596.0	0.0%
	White	158,131,126.1	84.3%	36,966,624.5	85.1%
	Black	21,728,087.5	11.6%	4,936,884.1	11.4%
	Other	7,773,929.4	4.1%	1,528,530.4	3.5%
	Total	187,647,206.3	100.0%	43,440,635.0	100.0%

Benchmarks can also be used to produce tables with any of the other types of cell contents that you can specify on the Table Contents tab. For example,
Figure 71 is a benchmark table showing the mean number of cigarettes smoked per day for the full dataset versus the Northeast region by age groups.

Figure 71. Benchmark of Mean Number of Cigarettes Smoked per Day for the Full Dataset versus the Northeast Region by Age Groups

Data in table for:	
Your Dataset:	
No filter(s) selected.	
Benchmark Dataset:	
Region:	Northeast

		Your Dataset	*Benchmark Dataset*
		Mean of # cigarettes smoked per day (CIGNUM)	Mean of # cigarettes smoked per day (CIGNUM)
Recoded Age (REC_AGE)	< 20	5.0	5.3
	20 - 29	6.3	6.5
	30 - 39	6.5	6.1
	40 - 49	5.9	5.8
	50 - 59	4.3	4.8
	>= 60	1.7	1.5
	Total	5.2	5.3

You can also create benchmark tables showing values for any number of row variables. For example, Figure 72 is a benchmark table showing the mean of the average systolic blood pressures for the full dataset versus the South for subjects with and without diabetes, with and without hypertension, of various races, and by gender.

Figure 72. Benchmark of Average Systolic Blood Pressures for the Full Dataset versus the South for Subjects with Various Characteristics.

Data in table for:	
Your Dataset:	
No filter(s) selected.	
Benchmark Dataset:	
Region:	South

		Your Dataset	*Benchmark Dataset*
		Mean of Average Systolic BP (SYSTOLIC)	Mean of Average Systolic BP (SYSTOLIC)
Has diabetes (DIAB1)	(missing)	126.5	126.1
	Yes	135.6	137.4
	No	123.0	123.0
	Total	123.7	123.7
H hypertension/HBP (HYPERTEN)	(missing)	117.8	117.7
	Yes	138.7	138.6
	No	119.3	119.5
	Total	123.7	123.7
Race (RACE)	(missing)	128.7	130.0
	White	123.9	124.2
	Black	124.0	123.7
	Other	119.0	116.3
	Total	123.7	123.7
Sex (SEX)	Male	126.2	127.3
	Female	121.4	120.6
	Total	123.7	123.7

List of Figures

Figure 1. Selecting Project Documents 3
Figure 2. Document List 3
Figure 3. Demo Codebook 4
Figure 4. Selecting Project Reports 5
Figure 5. Viewing List of Project Reports 5
Figure 6. Predefined WesVar Report (Gender by Race) 6
Figure 7. Table Type Selection Screen 6
Figure 8. Specifying One Variable Table 7
Figure 9. Variable List before Being Expanded 8
Figure 10. Variable List after Being Expanded 8
Figure 11. Variable List after Search for "race" 9
Figure 12. Selecting the Variable ("Race") 9
Figure 13. Show Results Button 10
Figure 14. Table Showing Frequency Distribution of a Variable (Race) 10
Figure 15. Exporting to Excel 11
Figure 16. Pie Chart of a Variable (Race) – including Total 11
Figure 17. Pie Chart of a Variable (Race) –after Deleting the Total Row 12
Figure 18. Pie Chart of a Variable (Race) –after Deleting the Total Row with Data Labels 12
Figure 19. Selecting Cell Percentage Display Option 13
Figure 20. Table with Both Counts and Percentages Displayed (Race Variable) 13
Figure 21. Suppressing the Display of Missing Values 14
Figure 22. Table with Missing Values Excluded (Race Variable) 14
Figure 23. Specifying a Summary Statistic (in this Case, the Mean) for Each Value of the Primary Variable 15
Figure 24. Mean Number of a Variable (Number of Cigarettes for Each Race) 15
Figure 25. Location of the "Start Over" Button 16
Figure 26. Option Buttons Available after Displaying Results 17
Figure 27. Steps to Save a Query 18
Figure 28. Saved Queries List 18
Figure 29. Sample Saved Queries Display Showing both My Queries and Commonly Used Queries Lists 19
Figure 30. Filters Tab 20
Figure 31. Specifying Variable Values in Section B of the Filters Tab 21
Figure 32. Filters Tab after Specifying Filter Variable and Value (Variable "Sex" with Value "M") 21
Figure 33. Mean Number of a Filtered Variable (Cigarettes for Each Race -- Males Only) 22
Figure 34. Specifying a Lower Bound for a Variable (20 Cigarettes per Day as Lower Bound) 22
Figure 35. Mean Number of a Variable Using Multiple Filters (Number of Cigarettes for Each Race -- Men Who Smoke at Least One Pack per Day) 23

Figure 36. Specifying Multiple Values for a Filter Variable (Variable Recoded Age: Values 20-29 or 30-39) 23
Figure 37. Mean Number of a Variable Using Multiple Filters Including One with Multiple Values Selected (Variable Recoded Age: Values 20-29 or 30-39) 24
Figure 38. Location of "Start Over" Control 25
Figure 39. Mean Number of Cigarettes per Day by Age Category 26
Figure 40. Location of "Collapse Category" Control on the Variables Tab 26
Figure 41. Collapse Category Pop-Up Window with 2 Categories to be Combined Marked 27
Figure 42. Collapse Category Pop-Up Window with 2 Categories to be Combined Marked 27
Figure 43. Renaming New Category 28
Figure 44. Table Generated Using New Category 28
Figure 45. Revised Table after Collapsing All Categories from Age 20 thru Age 59 29
Figure 46. Mean Number of Cigarettes per Day by Age Category after Relabeling First and Last Categories 29
Figure 47. Selecting Two Variable Table 30
Figure 48. Expanded Variable List When Specifying a Two Variable Table 31
Figure 49. Specifying a Two Variable Table (Sex by Status) 31
Figure 50. Two Variable Table (Sex by Status) 32
Figure 51. Specifying a Summary Statistic in a Two-Way Table 33
Figure 52. Summary Statistics in a Two-Way Table (Mean Number of Cigarettes Smoked per Day for Each Sex by Status) 33
Figure 53. Display and Output Options Buttons 33
Figure 54. Selecting Three Variable Table 34
Figure 55. Specifying Variables for a Three-Way Table 35
Figure 56. Three-Way Table (Race by Sex by Status) 36
Figure 57. Specifying Cell Contents as Mean Number of a Variable (Cigarettes per Day) 37
Figure 58. Three-Way Table with Cells Containing the Mean Number of a Variable (Cigarettes Smoked per Day) 38
Figure 59. Selecting Comparisons Table 39
Figure 60. Variables Available for Comparisons Table 40
Figure 61. Prompt for Column Variable after Selecting Cigarettes/Day 40
Figure 62. Selection of Pairs of Values for Comparisons 42
Figure 63. Comparison Table of Cigarettes per Day over 4 Years 42
Figure 64. Specifying Multiple Row Variables – Part 1 43
Figure 65. Specifying Multiple Row Variables – Part 2 43
Figure 66. Specifying Comparison to Test 44
Figure 67. Comparison of Hispanic to Other on Four Measures 44
Figure 68. Specifying Number of Responses and Cell Percentages as Table Contents. .. 45
Figure 69. Benchmark Tab Configured for Comparing the Northeast to the Entire Dataset 45
Figure 70. Distribution of Races for Northeast Benchmarked to Entire Dataset 46
Figure 71. Benchmark of Mean Number of Cigarettes Smoked per Day for the Full Dataset versus the Northeast Region by Age Groups 46

Figure 72. Benchmark of Average Systolic Blood Pressures for the Full Dataset versus the South for Subjects with Various Characteristics..... 47