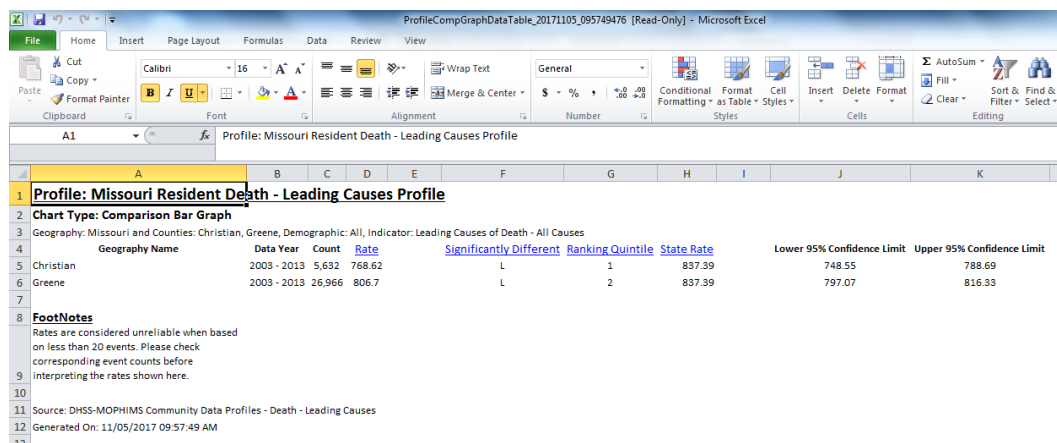## *Downloading Data into Microsoft Excel*

Both the Community Data Profiles and the MICAs provide the option of downloading images to insert into your report or downloading data for further manipulation in applications such as Microsoft Excel. If users are trying to custom create graphics to achieve a similar look throughout their entire report or simply prefer a different display for graphics than what is available on MOPHIMS, Excel's tools can be used to create charts that are customizable to the user's specifications. This handout provides some general tips regarding common issues that users may encounter when downloading data into Microsoft Excel. **The examples provided in this handout should be considered general guidelines only**. A complete description of every potential table type or every issue that may occur when using Profiles or MICA downloads to create Excel charts is not within the scope of this course. Additional documentation and tutorials are available on the Microsoft Excel website located at http://office.microsoft.com/en-us/excel.

Instructions for downloading data from the Profiles and MICA tools are included in the *MOPHIMS: Introduction to Profiles and MICA* and *Health Data Analysis* handbooks. The following examples look at creating various types of graphs from downloads using the **Leading Causes of Death Profile** for Greene County as the basis for selections.

Creating a Bar Chart in Microsoft Excel

A user has selected the Greene County Profile and wishes to create a Comparison Bar Chart showing Greene County's All Causes death rate compared to the neighboring county of Christian. The file below is downloaded from the bar chart Full Version for Greene with Christian County added from the drop down menu.


Several steps are needed to convert this data table into an Excel bar chart. First users must rearrange the table to only pull the relevant information. The county name (shown in Column A) and the rate statistic (shown in Column D) are needed. These data cells are selected using the cursor and then holding the control tab, as shown below. Users may also wish to add the state rate for comparison. Users have the option of either copying

the state rate over in a subsequent action or typing the state rate number into the worksheet. The state rate is located in Column H below.
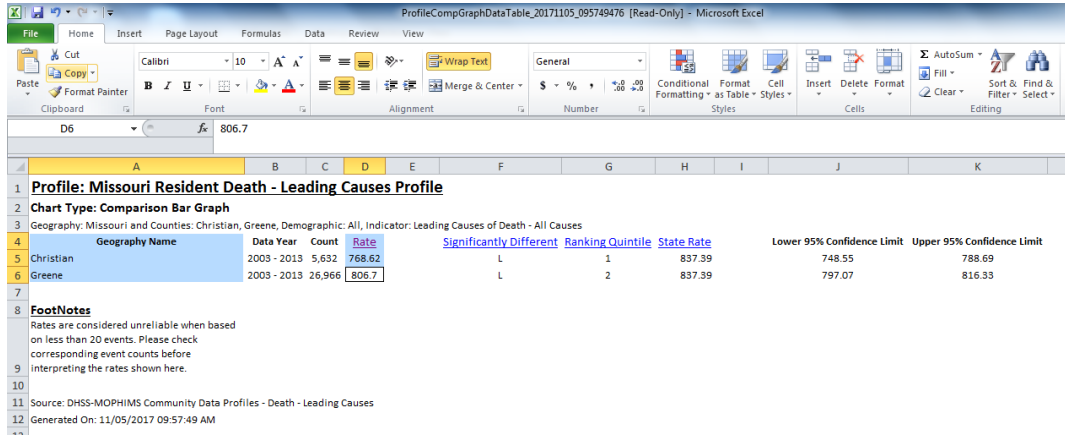

Users should then paste the information they wish to include in the bar chart into a different tab of the workbook. The image below shows what the new simplified table would look like. Users can now select the **Insert** tab in Excel circled below in red to create the chart in Excel.
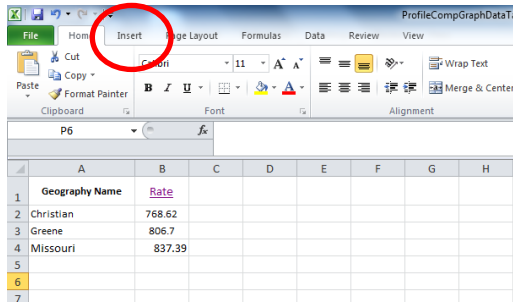

After selecting the **Insert** tab, click on one of the chart types shown below. A menu showing the options available for that chart type will appear. Hover the mouse over an option to see a description and suggestions for using that option. For this example, **Column Charts** will be used. Choose an option located in the first column of available chart types **(circled in the screenshot <u>on the left</u>)**. The second and third columns contain Stacked charts. Stacked charts function differently than the examples discussed in these courses.

Click on one of the options to insert a chart into the spreadsheet. The **Clustered Column Chart** was selected in the following graphic. To move the chart to another location within the same worksheet, position the mouse pointer over the chart border. The mouse pointer will change from a white plus sign to a black 4-headed arrow. Press the left mouse button and use the mouse to drag the chart to the desired location.
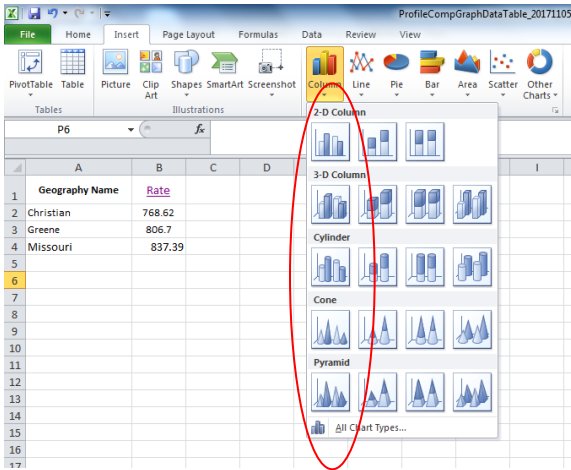
The ensuing table is shown below. The rates for Greene and Christian Counties and the state of Missouri are displayed in the chart. Note that the cells used for creating the graph are still identified by the blue, purple and green boxes in the data table in the upper left corner. These will be identified as long as the user has clicked the cursor in the chart area. In this case, users can select the rate legend on the right side of the graph and delete it. In other situations (such as when choosing to show differences by age or gender for multiple counties) the legend key would need to remain in the graph to help distinguish categories.
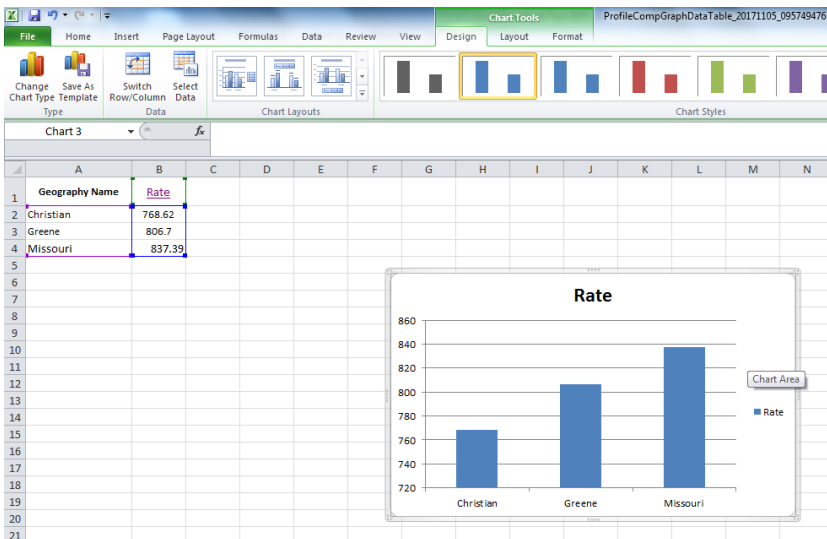
There are still a few additional formatting options that are needed. The default Excel scale is probably not appropriate (although this is more a stylistic choice than object fact) for this data table. In most situations, charts created should have 0 as the vertical axis. In order to change, users should right click on the vertical axis (where the values are listed) and select the **Format Axis** option as shown in the following screen capture.
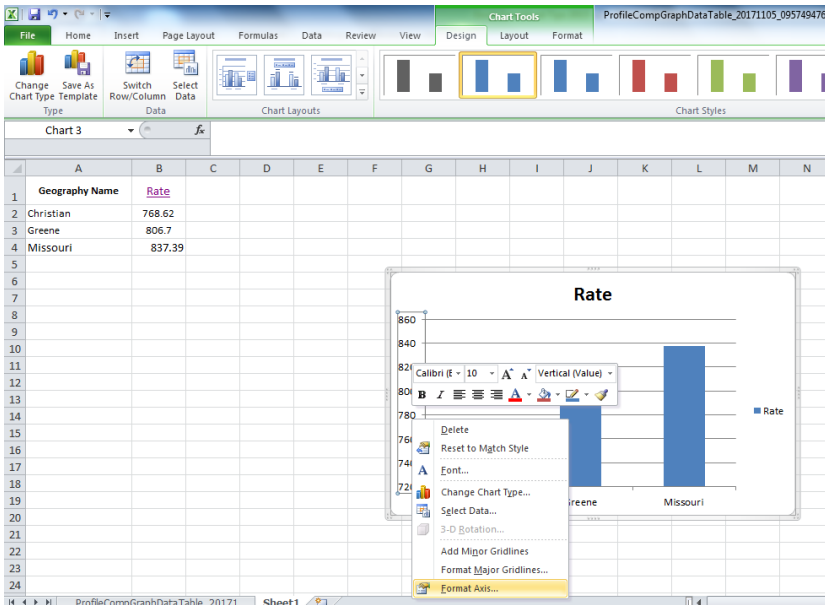
The **Format Axis** box allows users to modify the way the chart is displayed. To change the minimum value, users should change the Minimum field to fixed and type in 0 in the box to the right of the Auto/Fixed box options and then hit close.
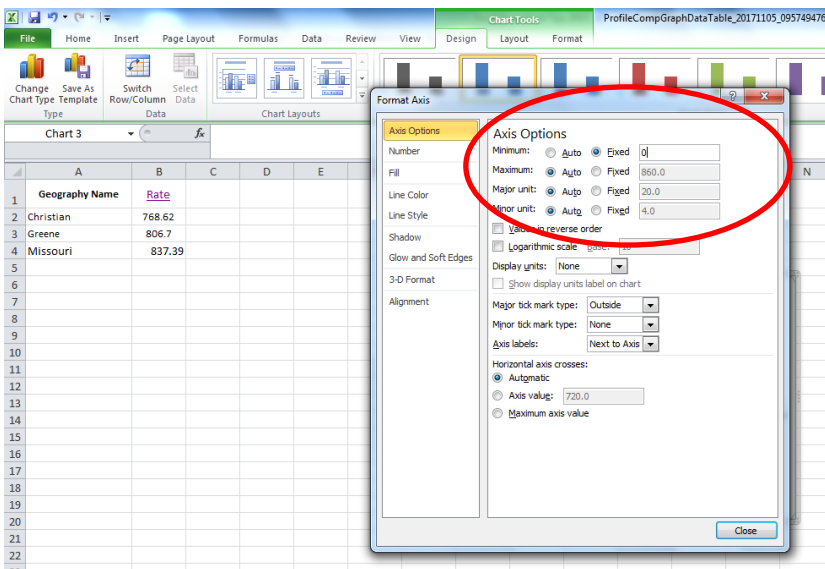

The chart is now modified and the graph now displays the full range of values on the vertical axis. This is preferred because if a graph that doesn't display the full range of values is created, it has the potential to mislead viewers of the graph into thinking that relative differences between bars on the chart are greater than is reality.

The title also needs to be modified. This can be accomplished by clicking on the legend title and highlighting the words selected, in this case 'Rate'. Users should then select an appropriate new label, as shown below in the table to the right.
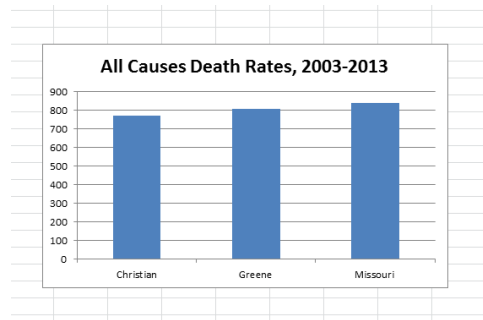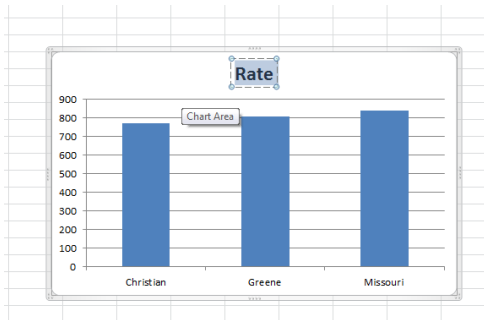
There are a handful of other optional items which could be added to the chart for additional clarity. A few of these choices are described below.

A footnote at the bottom of the chart may be added to provide the source of the data. To add a footnote, select the **Insert** tab from the Excel header list and then select **Text Box** (circled in red below). Then choose the location for your box and click and drag to obtain the proper size of the box. Type the source information and click out of the box to see how it displays. Some experimentation in terms of text box size/shape, and font/size will likely be necessary.
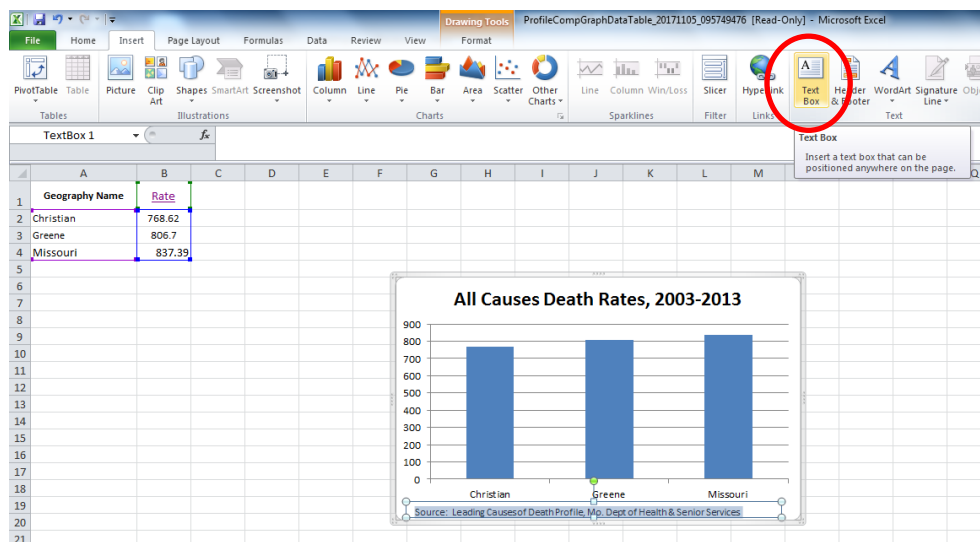
A user may also wish to add the specific values to the bar chart. This can be done by selecting **Chart Tools** (note that the chart must be selected for **Chart Tools** to display in the Excel ribbon). Select **Data Labels** and choose among the various options. For the example below, select the Outside End option with the values displaying just above the bars (also reflected in the chart below).
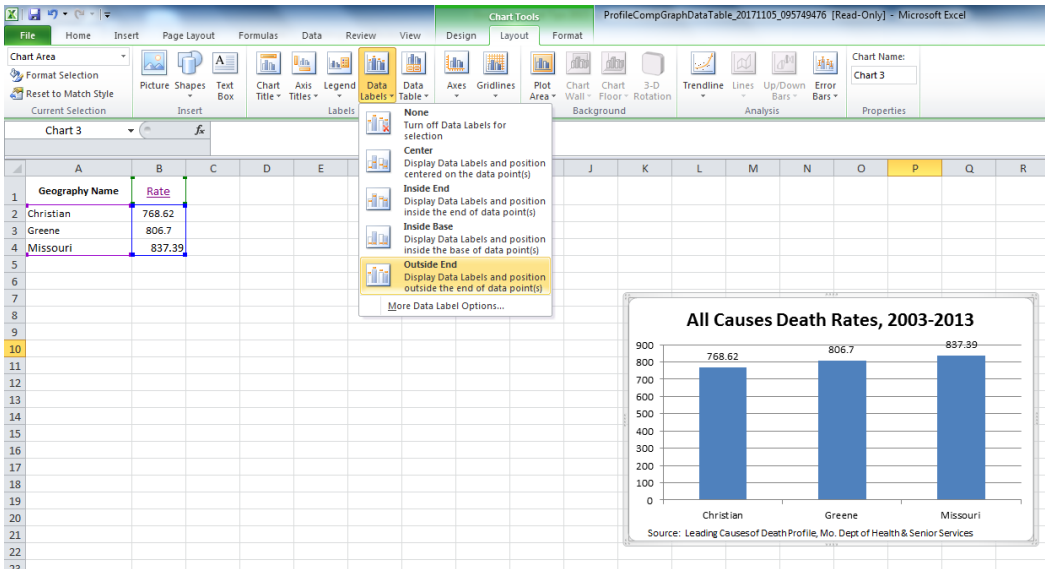
Including information about the constant and whether the death rates are age-adjusted is also important. This information could be displayed either at the bottom of the graph, just above the Source by adding an additional text box or it could be displayed next to the vertical axis, as is shown below. Users would again select the **Chart Tools** and **Layout** options from the Excel ribbon. Then select the **Axis Titles** and the **Primary Vertical Axis Title** from the drop down menu. For the example below, select the Rotated Title and type a title that provides information about the constant and age-adjustment information.
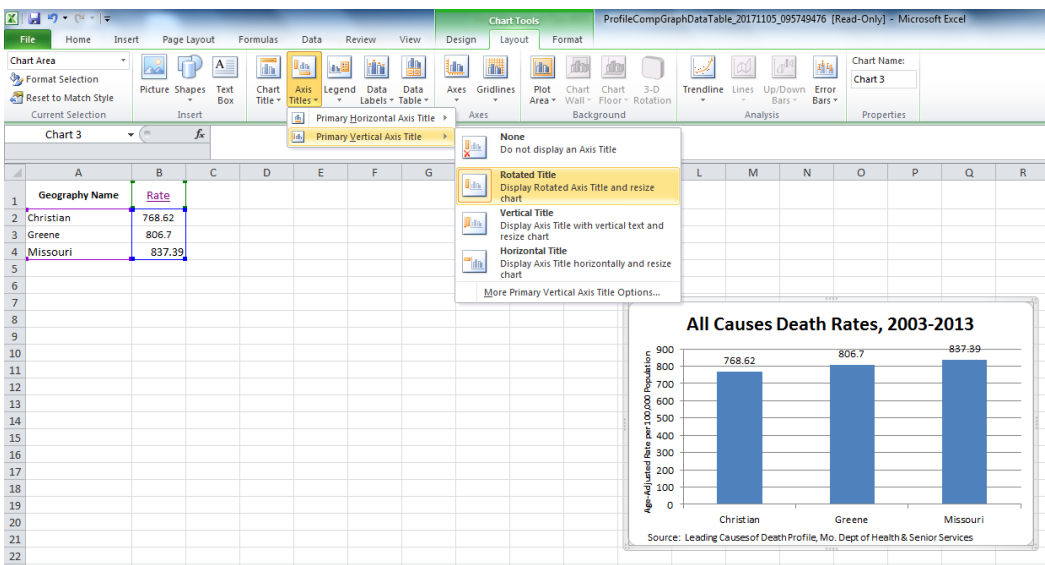
Creating a Line Chart in Microsoft Excel

Next, a user wishes to use the Greene County **Leading Causes of Death Profile** to create a trend line chart showing Greene County's death rate for the top four causes based on the most recent 3-year time period. After selecting Heart Disease, Cancer, Chronic Lower Respiratory Disease, and Unintentional Injuries from the drop down list for Causes on the Full Version Trend Line drop down list, the table shown on the left hand side of the page below (shown in part) is downloaded into Excel.

This chart has to be modified substantially to generate a single trend line chart showing trends for the four causes listed above. Users will need to copy and paste the columns showing the years and the rates (Columns B and C respectively) into a new Excel tab. In addition, the label describing the cause of death will need to be typed in Row 1 at the top of the column. It will take a series of copy and paste selections to move the information into the new table layout. The ensuing worksheet should look something like the right hand image shown above.

To create the trend line, users should choose the **Insert** tab and select the first option, labeled in Excel as **Line**. The graph is then displayed on the right hand side of the Excel sheet as shown below.
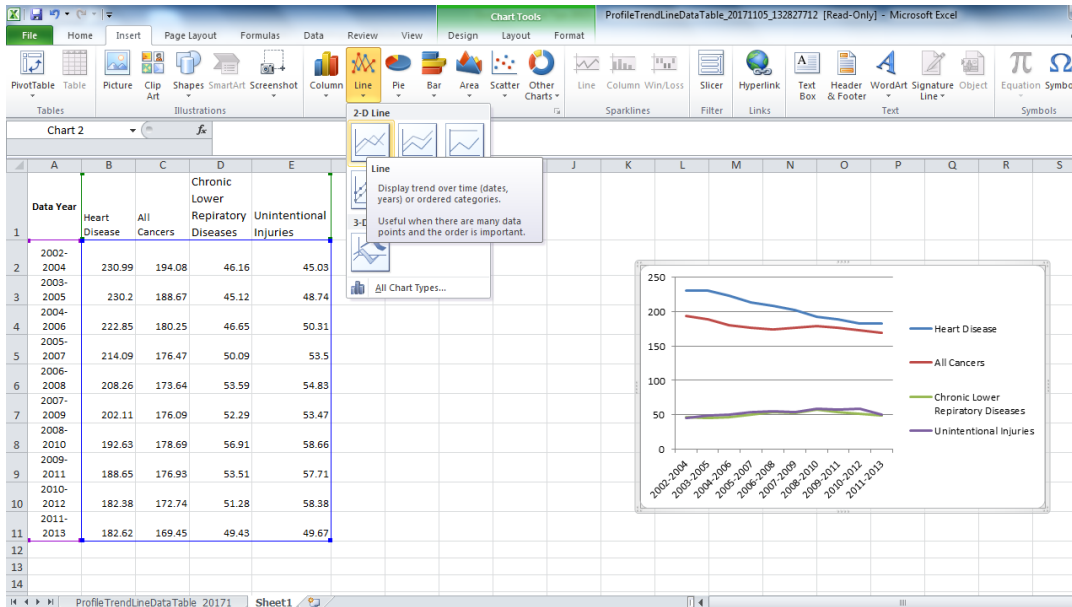
Users will next want to add a chart title. In this case, a default title is not added automatically. Instead, users can use the **Chart Tools** tab and select the **Layout** option from the Excel ribbon. Next, select **Chart Title** and choose the Above Chart option. The words 'Chart Title' will be used as the default by Excel, but users can select and type in a more appropriate title for the chart. Unlike with bar charts, the legend should remain as it delineates the meaning of the colors of the lines on the graph.
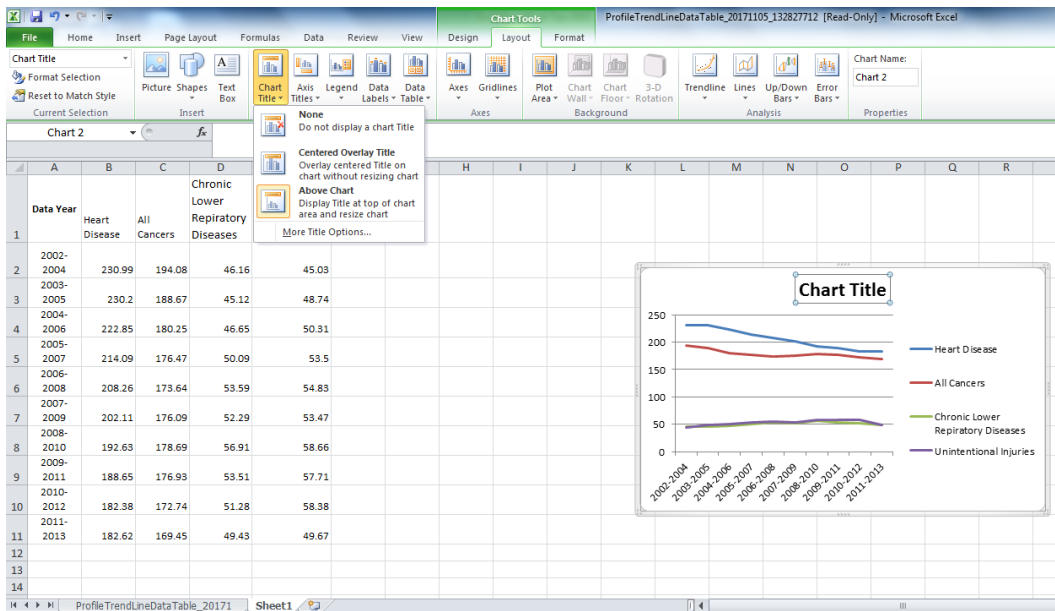
The chart below adds appropriate footnotes similar to what was described in the previous Bar Chart section of this document.

Greene County: Four Leading Causes of Death

A line chart showing Age-Adjusted Rate per 100,000 Population over 3-year moving averages (2002-2004 through 2011-2013) for Heart Disease, All Cancers, Chronic Lower Repiratory Diseases, and Unintentional Injuries.

Source: Leading Causes of Death Profile, Mo. Dept of Health & Senior Services
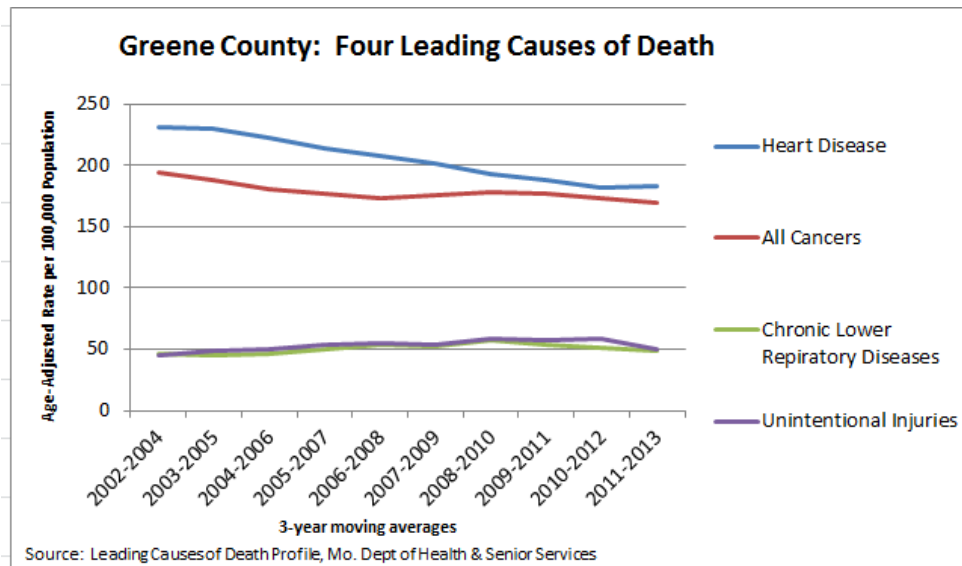
Formatting Microsoft Excel Charts

The following information goes into more complete detail about the various formatting options described in the examples listed above. The examples given below fluctuate between geographic and indicator variables.

Once a chart has been inserted into the spreadsheet, three additional tabs (**Design**, **Layout**, and **Format**) will appear at the top of the window, under the label **Chart Tools**. (If a chart has been inserted but these tabs do not appear, click anywhere on the chart to select it.) The **Design**, **Layout**, and **Format** tabs contain many options for customizing the chart. The most commonly used options are described in the following sections. Some options may not be available, depending on the type of chart that is selected. (Options that are rarely used have been omitted from this handout due to space limitations.)

Design Tab Options

Change Chart Type – Allows a user to quickly convert the chart to a different type (e.g., line, column, pie). Any labels or other information already added to the chart will be retained.

Save As Template – Saves the formatting and layout of the current chart as a template to be applied to future charts.

Switch Row/Column – Data being charted on the x-axis will switch to the y-axis and vice versa.

Select Data – Allows the user to modify the data cells included in a chart. This feature provides an alternative to deleting a chart and starting over. For example, data for the 2005-2007 time period are found in column G of a data table. Because of the study period in question, these data need to be removed from a chart based on this table.
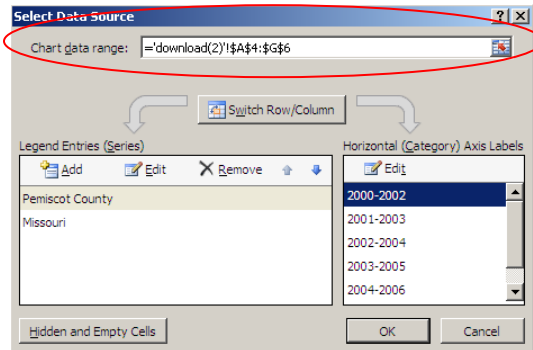
When the **Select Data** option is chosen, a Select Data Source box opens. The Chart data range listed is $A$4:$G$6.
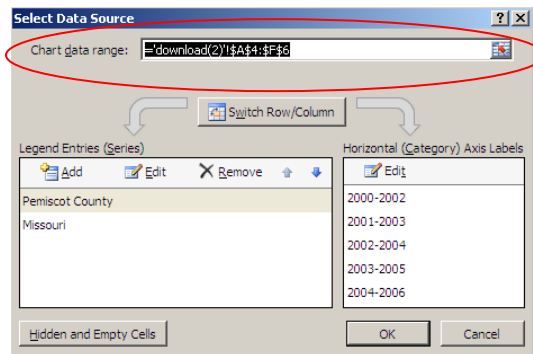
Changing $G:$6 to $F:$6 removes the 2005-2007 data from the chart, as shown below.

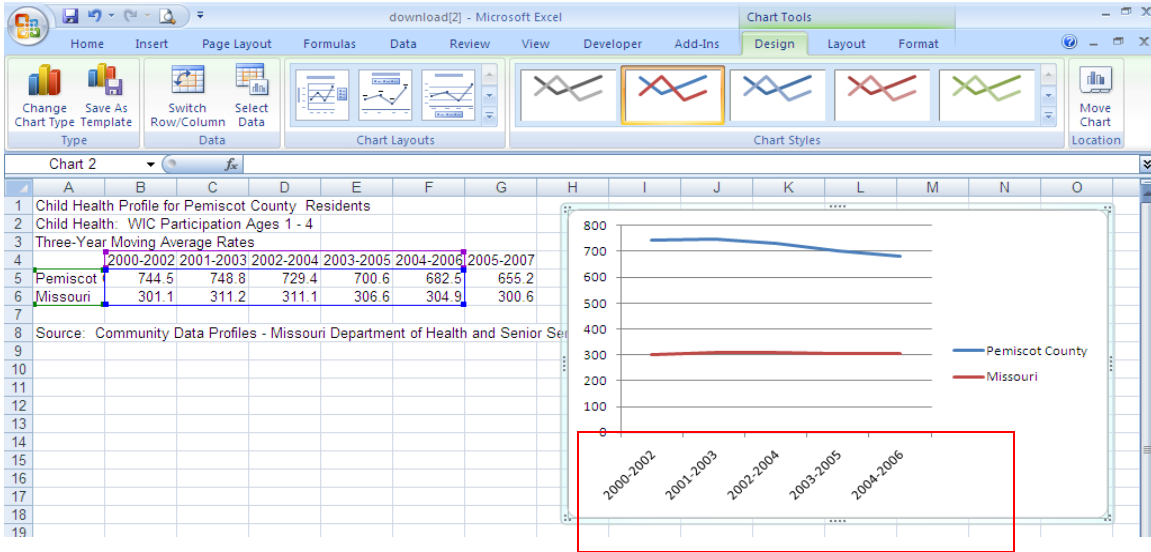The 2005-2007 data have been removed from the chart using the **Select Data** option.

Chart Layouts and *Chart Styles* – Allow the user to quickly select predefined layouts and styles (colors). The first layout option adds placeholders for the chart title and axis labels. Click the downward-pointing arrow with the bar over it (circled below) to see all options available in each section.
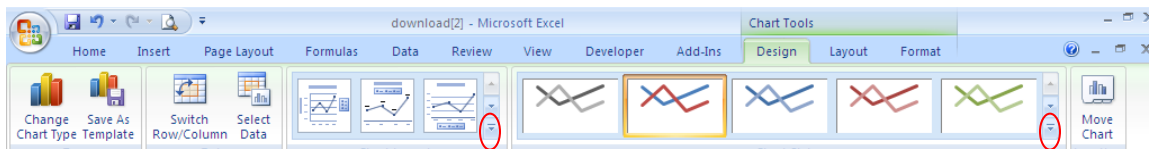
Move Chart – Allows a user to move the chart within the workbook. If the first option (New Sheet) is selected, the chart replaces the entire sheet specified. If the second option (Object in) is selected, the chart is simply moved (in its current format) to another worksheet.

Layout Tab Options
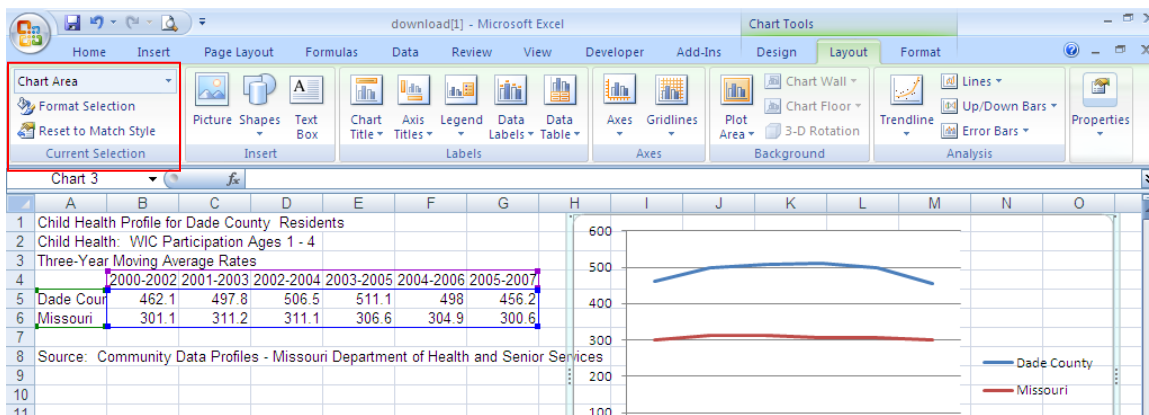
Current Selection – The first section of the **Layout** tab contains current selection tools (outlined in the above screenshot). The **Chart Area** drop-down list allows the user to select a section of the chart to be formatted. Clicking on the **Format Selection** option then opens a Format box specific to the section selected on the drop-down list. The **Format Selection** menus can also be accessed by right clicking on each area of the chart. A pop-up menu that includes a Format option will appear.

Reset to Match Style – Clears custom formatting and applies default formats.

Insert – Three options allow the user to insert various elements onto the chart.
- *Picture* – Allows the user to browse any available drives and select a picture for insertion.
- *Shapes* – Inserts a predesigned shape.
- *Text Box* – Inserts a box that can be positioned anywhere on the page. Text can then be entered. This will overlay any other elements on the page.

Labels – Allow the user to add, remove, or position the Chart Title, Axis Titles, Legend, Data Labels (which indicate actual values on lines/bars), and Data Table (which contains data used to build the chart).

Axes – Allow the user to change characteristics of the horizontal and vertical axes.
- *Axes* – Change axis formatting and layout.
- *Gridlines* – Turn gridlines on or off.

Format Tab Options

Current Selection – See description under Layout Tab Options.

Shape Styles – Allow the user to apply various colors, fills, outlines, and effects to the currently selected area of the chart.

WordArt Styles – Allow the user to apply different styles, fills, outlines, and effects to the numbers and text in the currently selected area of the chart.

Size – Increases or decreases the size of the entire chart. Charts can also be resized by clicking and dragging any of the dotted areas on the chart border (circled in the screenshot on the next page).
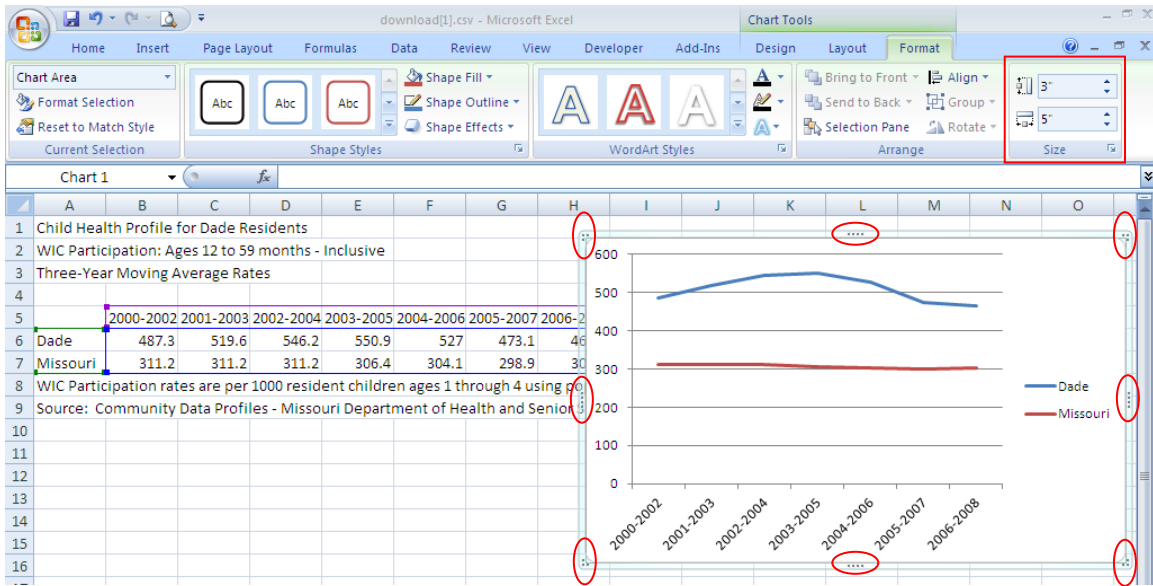
Troubleshooting

Some common problems that users may encounter when formatting charts are described in this section.

Problem: *Series labels appear on the chart instead of the actual data labels.*
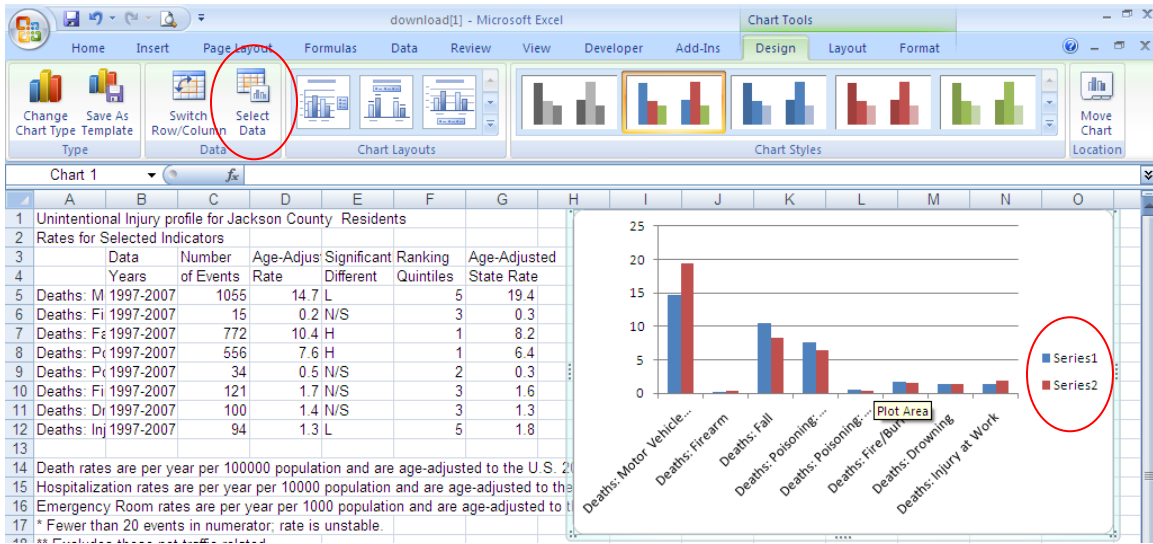
Solution: In the example above, the Jackson County rates in Column D and the Missouri rates in Column G were used to create the chart. However, the county and state labels are not listed with the data, so Excel applied the default labels of Series1 and Series2. To customize these labels, choose **Select Data** (circled in the screenshot on the previous page) from the **Design** tab. The Select Data Source box will open.
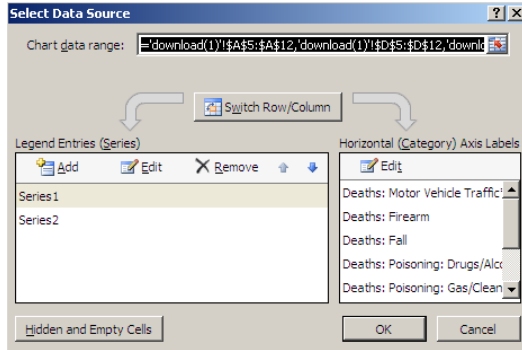

The Series labels are listed under Legend Entries. Click on Series1, then choose Edit from the menu just above it.
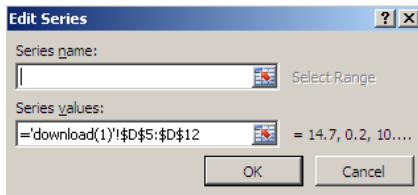
The Edit Series box will appear. Notice that the Series name is blank (which is why Excel assigned a default name). Examination of the chart reveals that the Series1 data represent Jackson County. Type Jackson County into the Series name field, then click OK to close the Edit Series box. Click OK again to close the Select Data Source box.

Jackson County is now listed on the legend (circled below). Follow the same steps to replace the Series2 label with State of Missouri.
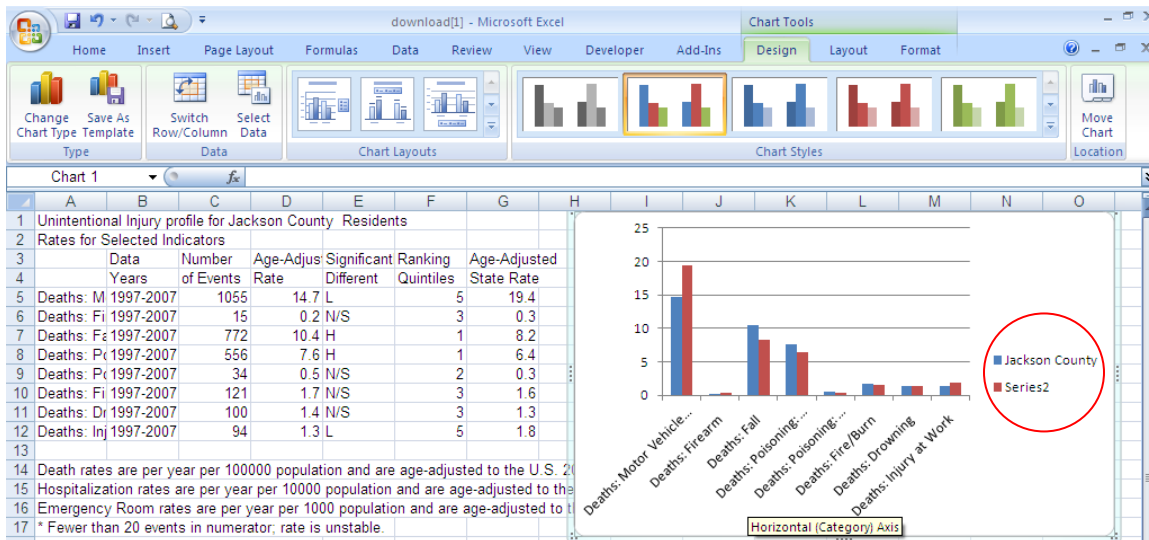
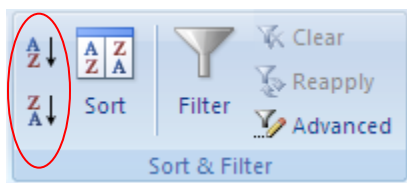
Problem: *Labels are too long.*

Solution: In the example above, the ellipses at the end of many of the horizontal axis labels indicate that portions of those labels have been truncated. Enlarging the table may provide enough space for the full labels to appear. In this case, the **Chart Type** is Column, but it could be converted to a Horizontal Bar Chart to accommodate the long labels.

Problem: *Labels need to be changed.*

Solution: The labels that automatically appear on a chart are pulled directly from the data table. The only way to change these automatic labels is to change the labels on the table. In the previous example, the data labels were rather long. Since all labels began with "Deaths," a user might wish to remove "Deaths" from the individual labels and instead include "Deaths" in the chart title. In that case, the user must delete the term "Deaths" from each label cell (A5 through A12).

Problem: *Data should be sorted differently.*

Solution: In the example on page 14, the data categories (types of deaths) do not appear to be sorted in any particular order. However, a user may wish to chart the categories in either ascending or descending order of the county or state rates. To sort chart data, simply click on the table column that is to be used for sorting. (For example, if sorting by the Jackson County rates, click on any cell from D5 through D12.) Then choose the **Data** tab at the top of the window to view the sort options. Choose ascending or descending order. The chart will automatically sort based on the new table order.

Creating Confidence Interval Charts

Confidence intervals are discussed in both *MOPHIMS: Introduction to Profiles and MICA* and *MOPHIMS: Health Data Analysis.* Examples of confidence interval charts are presented in *MOPHIMS: Health Data Analysis*. These confidence interval charts were created using Excel's Other Charts, Stock, High-Low-Close option. The following instructions explain how to create this type of chart.

1. The table used to create the confidence interval chart must include only four columns: the data label, the rate, the lower confidence level, and the upper confidence level. The different categories to be compared must be listed on the rows. In this screenshot, Table 1 represents the original download, while Table 2 illustrates the necessary format for the confidence interval chart.


2. Select the cells to be included in the chart. In this example, the appropriate cells would be A14 through D17. (At least three comparison categories must be selected. Unnecessary categories, such as All Races, can be removed from the chart using the process described in step 4.) Then choose the **Insert** tab, Other Charts, Stock, High-Low-Close option.
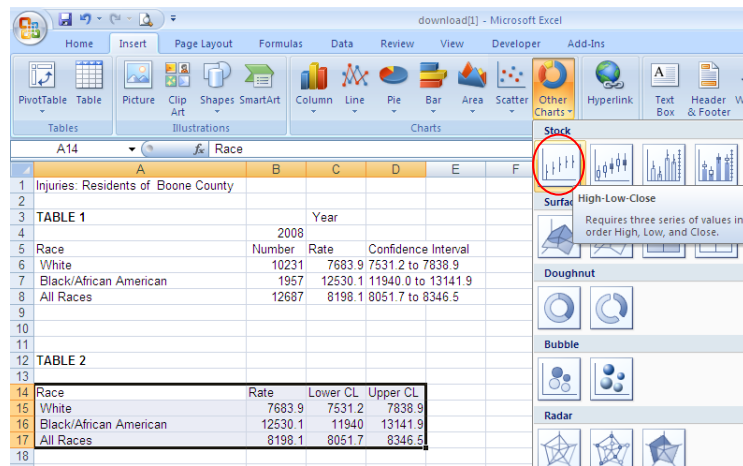

3. The resulting chart lists the statistics (rate, lower CL, and upper CL) on the horizontal axis and the comparison categories (in this example, White and Black/African-American) on the legend. Click the **Design** tab, Switch Row/Column to reverse the placement of the statistics and the comparison categories.
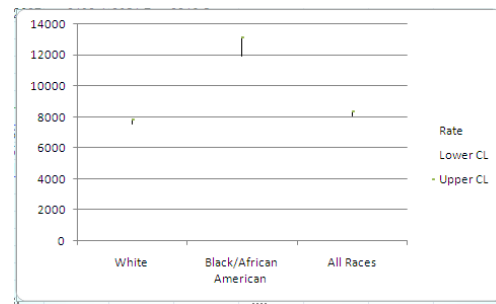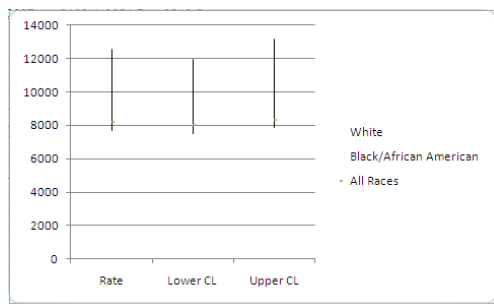
4. *Optional:* To remove All Races from the chart, choose the Select Data option on the **Design** tab. In the Select Data Source box, change the Chart data range so that it ends with Row 16 instead of Row 17.
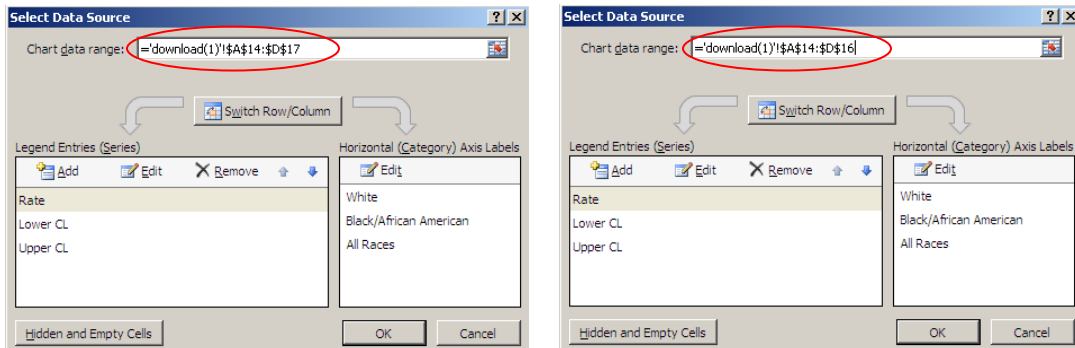

5. *Optional:* Notice that the confidence interval lines are very small due to the large scale range (0 to 14,000). Since the smallest rate on the table is 7,531.2, the scale can be adjusted to start at 7,000. Select the **Layout** tab, choose Vertical Axis on the Current Selection list, and click on Format Selection. Under Axis Options, change Minimum to Fixed, then type 7,000 in the Minimum box. The confidence interval lines are now somewhat longer. The chart could be resized to further lengthen the lines.
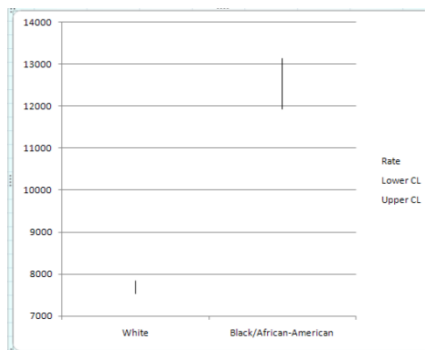
6. *Optional:* To add dots to represent the actual rates, choose Series "Rate" on the Current Selection list of the **Layout** tab, then Format Selection. Choose Marker Options, Built-in. Select a dot from the Type drop-down list. Increase or decrease the size as needed. Choose Marker Fill, Solid Fill, then select a color that matches the line. Close the Format Data Series box.

7. *Optional:* To add bars at the bottom of the confidence interval lines, choose Series "Lower CL" on the Current Selection list of the **Layout** tab, then Format Selection. Choose Marker Options, Built-in. Select a dash and increase the size if needed. Choose Marker Line Color, Solid line, then select a color that matches the confidence interval line. Choose Marker Line Style and increase the width if needed. Close the Format Data Series box. Choose Series "Upper CL" on the

Current Selection list of the Layout tab, then Format Selection and repeat these steps to add bars at the top of the confidence interval lines.
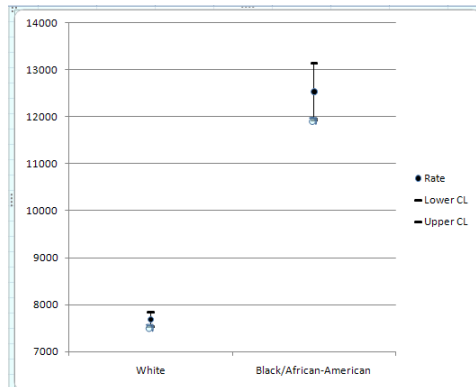
Copying and Pasting Microsoft Excel Charts into Microsoft Word

Select the entire chart, then right click. On the menu that appears, choose Copy. In a Microsoft Word document, position the cursor on the point at which the chart should be inserted. On the Home tab in Word, click the drop-down arrow under Paste. Choose Paste Special, Bitmap, then click OK. Using the bitmap option reduces the file size and helps to prevent corrupted files. To compress the Word document further, select the chart image. A **Picture Tools** tab will appear at the top of the window. Choose Compress Pictures, then click OK.

If a Microsoft Word 2007 (.docx) file does corrupt, close the file immediately. Open My Documents and go to the folder where the corrupted file is stored. Right click on the corrupted file. Choose Save As. . . on the pop-up menu. In the Save as type: box (just below the file name), select Microsoft Office Word 97 - 2003 Document (*.doc). Often the file can be opened using the older file format.

These instructions were created using the 2007 version of Microsoft Excel. Screens and functionalities may differ in other versions of Microsoft Excel.