Class Instructions: Working with Data in Excel

Sean Hribal | Athens-Clarke County Library

Overview

In this class, you'll learn how to download U.S. census data from the Internet, open it in an Excel spreadsheet, and then format, sort, use formulas, and insert a chart.

Review Basics

- Excel structure: ribbon, groups, tools, scroll bar, and zoom
- Navigating: Pointer, arrow keys, [Enter], and [Tab]
- Selecting: cells, non-contiguous cells, rows, columns, all

Exercise 1: Format, sort, and use formulas with data

Download Census Data

- Go to census.gov.
- Click Data > Data Tools & Apps > American FactFinder (http://factfinder.census.gov).
- Under Community Facts, type "Georgia" in the search bar and click Go.
- Under 2010 Census, click Compare Counties for Population, Housing, Area, and Density > Download (button above data)
- In the Download window, select "Data and annotations in separate files" and click OK.
- Click Download
- A new window opens. Select "Open with Windows Explorer" and click OK.
- Click "Extract" tab at top of File Explorer window and then "Extract all". Browse to the Desktop folder and click Extract.

Clean Up Data

- Double-click "DEC_10_SF1_GCTPH1.ST05.csv". An Excel spreadsheet opens containing census data for Georgia by county.
- Rows A-F contain data that we don't need so let's delete those rows. Click on the column A header, hold down the shift key, and then click on column F header. This selects Columns A-F. Right-click in the selected area and then click "Delete".
- Row 1 also contains data we don't need so click on row header 1. Then press the [Delete] key to clear out that data.
  - Note: The delete key deletes data but not the row, whereas clicking "Delete" in the right-click menu deletes the data *and* the row.
- Row 3 contains totals (for Georgia) of all the county data. Delete that row.
  - These totals could interfere with sorting or calculating county data.

Format Data

Add title

- In cell A1, type in "Georgia County Census Data 2010".

- Select cells A1-H1 by dragging across those cells. In the Alignment group on the ribbon, click the "Merge & Center" down arrow and then click "Merge Across".
  - Now we can auto re-size columns below the title.
- Select the title cell. Make bold and font size 14.

Change spreadsheet file type
- Click the "File" tab, click "Save As", change the File name to "GA Census Data", in the "Save as type" drop down menu select "Excel Workbook (*.xlsx), and click "Save."
  - We started with a .csv file which stands for comma separated values. This file format does not let us make changes to formatting so we will have to save our spreadsheet as an .xlsx file (the default file format for Excel 2007 and later versions).

Format column headings
- Select row 2 (column headings) by clicking the row 2 heading. Make bold.
- Auto-adjust columns widths: Click square at intersection of row header 1 and column A header to select all cells. Place pointer between any two column headers (when you see the horizontal arrow) and double click.
- Give column headings to shorter names. Make following replacements.
  - Area in square miles -> Area (sq mi)
  - Area in square miles - Water area -> Water Area (sq mi)
  - Area in square miles – Land area -> Land Area (sq mi)
  - Density per square mile of land area – Population -> Density (pop/land)
  - Density per square mile of land area – Housing Units -> Density (housing/land)
- Auto-adjust column widths

Format numerical data
- Before we sort or calculate data, we want to make sure it's in the right format.
- Select the first cell of numerical data (B3). Hold down the [Shift] and [Ctrl] keys and then tap the [→] key. This selects data to the last column. Hold down the [Shift] and [Ctrl] keys and then tap the [↓] key. This selects data to the last row.
- In the Excel ribbon, there's a group of icons called "Number". In the Number group, find the drop down menu that says "General". Click the menu's down arrow and select "Number".
- Click "comma style" button.
- Click the "Increase decimal" button a few times. Then click the "Decrease Decimal" button until there are no decimal points.

Sort data
- As we can see, the census data we downloaded is sorted alphabetically by county name. What if we want to sort by something else like population or land area?
- Select **ALL** your data (INCLUDE row headings and column headings) using the technique we used in previous steps.
  - If you select one column or a few before you sort, the sort will re-order only those columns but not adjacent columns, which will scramble your data. Selecting all data ensures that your rows remain intact.
- Click on the "Data" tab. Click the "Sort" button.
- A Sort window opens. Select the column you want to sort by… let's pick "Population", Sort on "Values", Order "Largest to Smallest", and click "OK".
- Now we can easily see which counties have the highest populations.

- Sort a few different columns.
- Then return to the data to its original order by sorting by Geographic area A->Z.

Use Formulas
- Right-click the row heading 3 and then click "Insert". Repeat until there are 5 new rows between the title and the column headings.
- In A3, type "GA Total", in A4 "GA Average", A5 "GA Median", A6 "GA Max", A7 "GA Min"
- To find the total population of Georgia. We need to take the sum of the population from all counties. So we will use the Sum formula.
  - In cell B3, type "=sum(" to start the Sum formula
- Now we tell the formula the range of cells we want to sum.
  - Click on B8. Hold down [Ctrl] and [Shift] and press [↓]
  - Notice how this enters B8:B166 into the formula. We are finding the sum of the range of cells from B8 to B166.
  - Press [Enter] to run the Sum formula which adds in the ")"
  - The sum is now shown in the cell where we entered the formula.
- How do we find the totals for all our different columns?
  - Click on cell B3. Move your pointer over the black square in the bottom right of that cell. Your pointer will become a black cross. Drag over to H3 and let go.
  - Double click (or click and [F2]) on any of the new sum values and Excel will show the colored border containing the range of data used by that formula.
- Let's find the average population of all Georgia counties.
  - Go to cell B4. Now repeat the same steps we followed for doing the sum except type in "=average(" and then select the range.
  - Then find the average for all the columns.
- Now calculate median, max, and min values by using these formulas
  - =median(
  - =max(
  - =min(
- Freeze panes: Select row 8 (Appling County). Click the "View" tab. Click the "Freeze Panes" button. Click "Freeze Panes".
  - Now when we scroll down through our data the column headings remain visible and we can easily compare each county's values to the state total, average, median, max, and min.
- Highlight the title and change the background color to a medium blue
- Highlight column headings and change background color to light grey
- Highlight formulas and change background color to light blue
- Save your spreadsheet.

Worksheets
- At the bottom of your spreadsheet, you'll see a tab named "DEC_10_SF1_GCTPH1.ST05".
- Right-click this tab, click "Rename", and type in "2010 GA County Data".
- To the right of this tab is an "Insert Worksheet" button. Click it. Rename the new worksheet "2000-2009 State Census Data"

Exercise 2: Creating a time series graph

Download Census Data
- Go to census.gov.
- Click <u>Topics</u> > <u>Population</u> > <u>Population Main</u> > <u>Population Estimates</u> > <u>Current Estimates Data</u> > Intercensal estimates are available <u>here</u> > <u>State</u>
- Where it says "Intercensal Estimates of the Resident Population for the United States, Regions, States, and Puerto Rico: April 1, 2000 to July 1, 2010 [<u>XLS</u> - 21k] | [<u>CSV</u> - 8k]", right-click "CSV", click "Save Link As…", and save to Desktop.

Format Data
- Pull up your GA Census Data.xlsx spreadsheet.
- Select the 2000-2009 State Census Data worksheet
- Click on the "Data" tab, click "From Text", and double-click "ST-EST00INT-01.csv"
- Make sure "Delimited" is selected.
  - Delimited means that the data values are separated by characters. We're importing a .csv so we know the data are separated by commas.
- Click "Next". Unselect "Tab" and select "Comma".
  - Notice in the Data preview how the values are now separated into rows.
- Click "Next". Make sure "General is selected". Click "Finish".
- A box pops up asking where to put the data. Make sure "Existing worksheet" is selected and click the button with the small red arrow. Click cell A1. Press [Enter]. Click "OK".
- Let's clean up the data. Delete rows 1, 2, and 3. Delete rows 59-70. Delete Columns B, M, and N. Delete cell B1.
- Auto re-size columns.
- Make column headings bold
- In A1, type "State Intercensal Census Data 2000-2009"
- Make title bold, Calibri, size 14, and merge columns across data.
- Check to see if data is in number format. It is.

Time Series Graph
- Select the column headings (years)
- Hold down [Ctrl] and highlight Georgia and its data, as well as two other states.
- Click "Insert" tab. Click "Scatter" and "Scatter with smooth lines and markers".
- Move pointer over graph until you get 4-way arrow. Drag graph off to the right. Make graph larger by dragging corner.
- Click on 1st Chart Layout. Add title and axis labels.
- Click on grid lines and press [Delete].
- Show how to add and remove data.