By the end of this activity, you will be able to:

1. Identify the key features in CSV data
2. Import CSV data to a spreadsheet and plot values

Step 1. **Open a terminal shell.** Open a terminal shell by clicking on the square black box on the top left of the screen.



Run *cd Downloads/big-data-2/csv*to change into the directory containing the csv file. (This was downloaded previously : <https://www.coursera.org/learn/big-data-management/supplement/YVDPj/instructions-for-downloading-hands-on-datasets>)

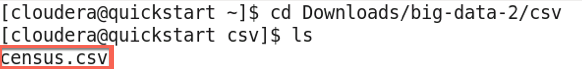
1

cd Downloads/big-data-2/csv

Step 2. **Look at CSV file.** The CSV file contains census data for the United States. Run *ls*to see the name of the CSV file.

1

ls

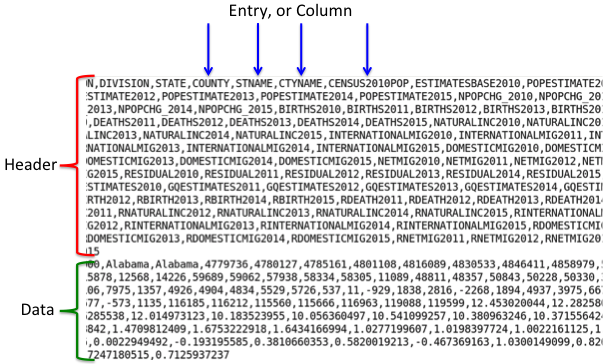


Run *more census.csv*to look at the contents of the CSV file.

1

more census.csv

The first line of the file is the head and the remaining lines are the data. Each entry in the file is separated by a comma.

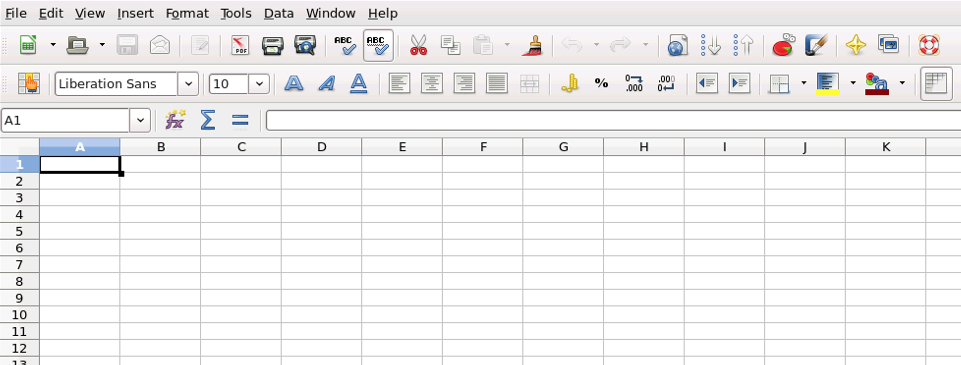


Hit the spacebar to scroll down, and *q* to quit more.

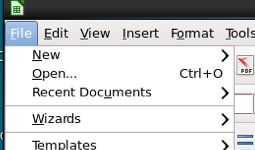
Step 3. **Open spreadsheet application.** Run *oocalc* to start the spreadsheet application.

1

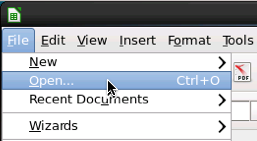
oocalc



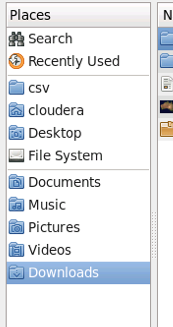
Step 4. **Import CSV to spreadsheet.** Let's import the CSV file to the spreadsheet by clicking on *File:*



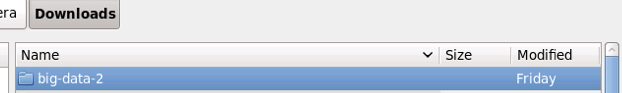
Next, click *Open:*



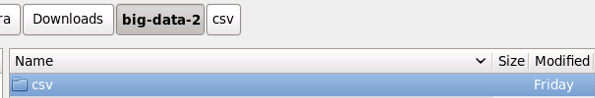
Next, click*Downloads*in the Places pane:



Next, double-click *big-data-2*in the file pane:



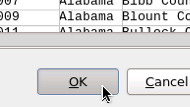
Next, double-click*csv*:



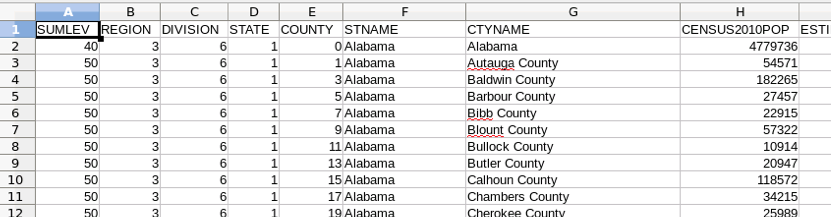
Next, double-click *census.csv*:



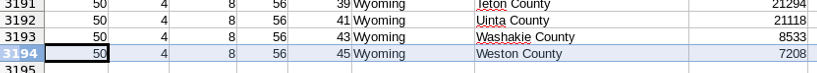
In the Text Import dialog, click *OK:*



The CSV data is now loaded into the spreadsheet.



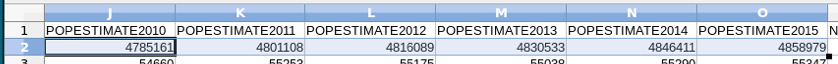
Step 5. **See size of CSV.** Scroll to the bottom of the spreadsheet to see the size of the CSV file.



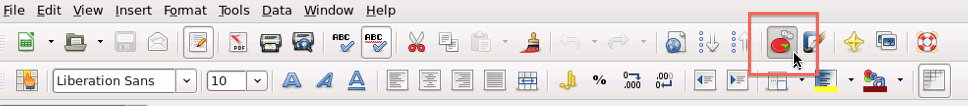
There are 3194 rows. If the CSV file had millions or more rows, then we could not import it into a spreadsheet. In this case, we would need a Big Data system such as Hadoop to analyze the data.

Scroll back to the top.

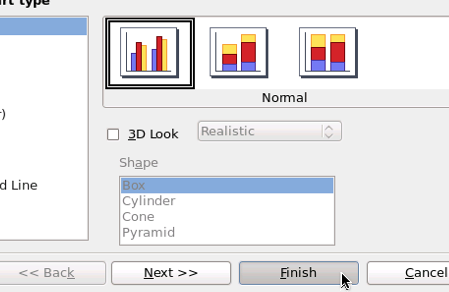
Step 6. **Create chart.** Let's create a chart of the estimated population of the state of Alabama. Row 2 contains the data for Alabama. Select cells in row 2 and columns J through O to get the estimated population for 2010 through 2015.



Click on the chart button:



Click *Finish* to display the chart:



The chart should be displayed in the spreadsheet:

