

## Third Activity

*Sept. 15, 2011*

1. Close the Excel spreadsheet you were looking at.  
Instead of having you load in all the votes from the 109th congress, which takes about 20 minutes, we've done that for you.
2. Right-click and save the `AllVotes.xlsx` file on the Class 2 website into your flash drive. If it got saved as a `.zip` file (it'll look like a folder with a belt around it), change the extension to `.xlsx`. [This problem arises when you use Internet Explorer; if you use Firefox, it should not occur.] Then open that file in Excel. Opening it will take a minute or so.
3. Get familiar with the data. How many rows are there? How many columns?
4. Spend some time deleting columns from the data (unfortunately, you have to do this manually). We'll need the vote number, session, member\_full, vote\_question\_text, and vote\_cast.
5. Save the much-reduced spreadsheet with the name `Class2Work.xlsx`
6. Now figure out how to delete all the rows of the spreadsheet that correspond to votes that are *not* on the passage of a bill. Hint: this should only take a few minutes; it does *not* involve deleting the non-bill-vote rows one by one! Save the resulting spreadsheet, which should be a lot smaller – perhaps about 4800 rows instead of 65000. If you are having trouble with this, don't hesitate to ask a TA for help.
7. Add another column to the table by entering “voteID” as the title for the column in cell `F1`. Notice that Excel automatically infers that this defined a new part of the table. In row 2 of the `voteID` column, write a formula to combine the “session” for this row with the “vote\_number” for this row, placing a colon between them. So if the session is 2 and the vote number is 104, your cell should end up containing `2:104`. Use “fill down”, if necessary, to apply this formula to all the other rows. (Excel may do this automatically because you're working on a table; whether it does so or not depends on how it's been set up.)

8. Add another column to the table, just to the right of `voteID`, and title it `numericalVote`. In the second row of that column, enter a formula that has a 0 if the senator did not vote, a 1 if s/he votes “Nay,” and a 2 if s/he voted “Yea.” If s/he has some other vote than these three, your cell should contain the word “ERROR”. (This will not happen at all in this data.)
9. Click on the **Insert** tab, and insert a pivot table, using the suggested “Table 1” (that’s what Excel automatically named our vote data). Make the rows be labelled by the senator’s full name, the columns be labelled by the `voteID` that you just created, and the value in the cells be the `numericalVote` that you just created. Quickly scan this newly added column to confirm that there are no errors.
10. Change the value in the cells to be the maximum rather than the sum of the votes; do so by left-clicking on the top-left corner of the pivot table and selecting **Value Field Settings**; choose **Max** from the list that appears, and click **OK**.
11. Delete the **grand Total** row and column by using **Option ...PivotTable ...Options ...Option** and selecting the **Totals & Filters** tab, and uncheck the “show grand totals” checkbox. Alternatively, you can simply delete the relevant row and column.
12. Save your work and email it to yourself; we’ll be using it again for the next class, when we start analyzing the data.