

# Collecting Data

Feb. 2 2002

# First, it's Groundhog Day



[http://www.cbsnews.com/8301-201\\_162-57370239/groundhog-day-6-more-weeks-of-winter/](http://www.cbsnews.com/8301-201_162-57370239/groundhog-day-6-more-weeks-of-winter/)

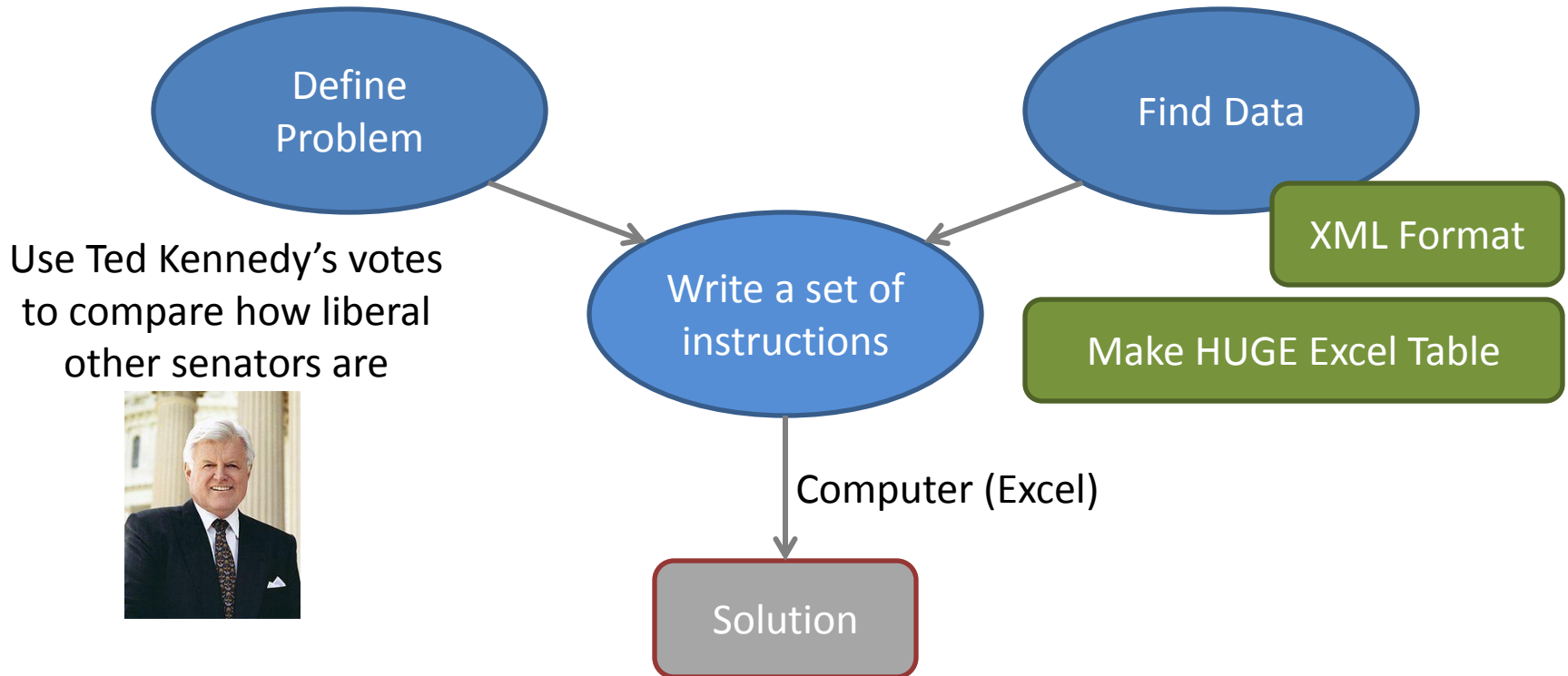
# Announcements

- TA & Faculty Hours will be posted soon
  - If you have a question *before* the hours are posted, ask [cs0931tas@cs.brown.edu](mailto:cs0931tas@cs.brown.edu).
  - Stock Exchange Question: [url](#)

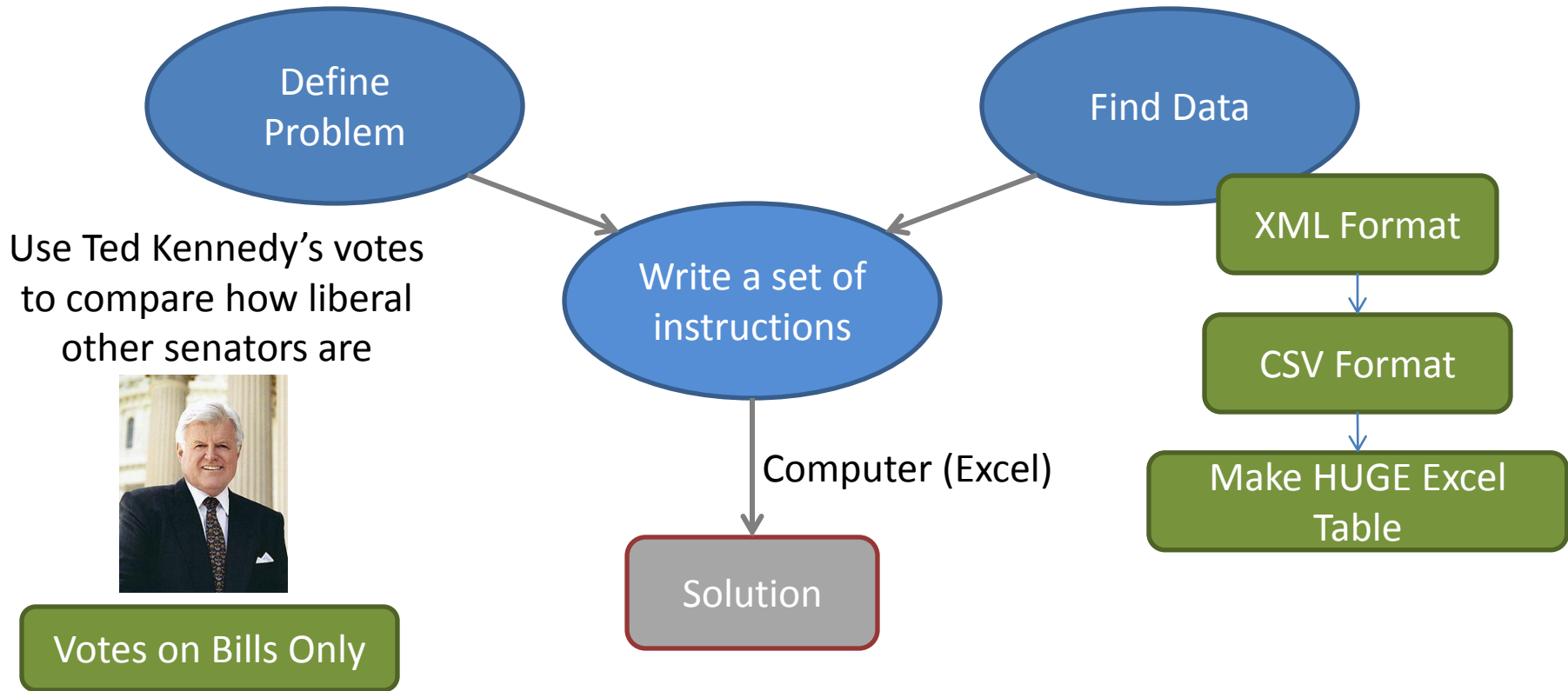
# Announcements

- TA & Faculty Hours will be posted soon
  - If you have a question *before* the hours are posted, ask [cs0931tas@cs.brown.edu](mailto:cs0931tas@cs.brown.edu).
  - Stock Exchange Question: [url](#)
- Lab time has been reserved in this room (265):
  - Sundays 7pm-9pm
  - Mondays 7pm-10pm
  - (Feb. 6 & Feb. 13 in 267 instead)

# This Class



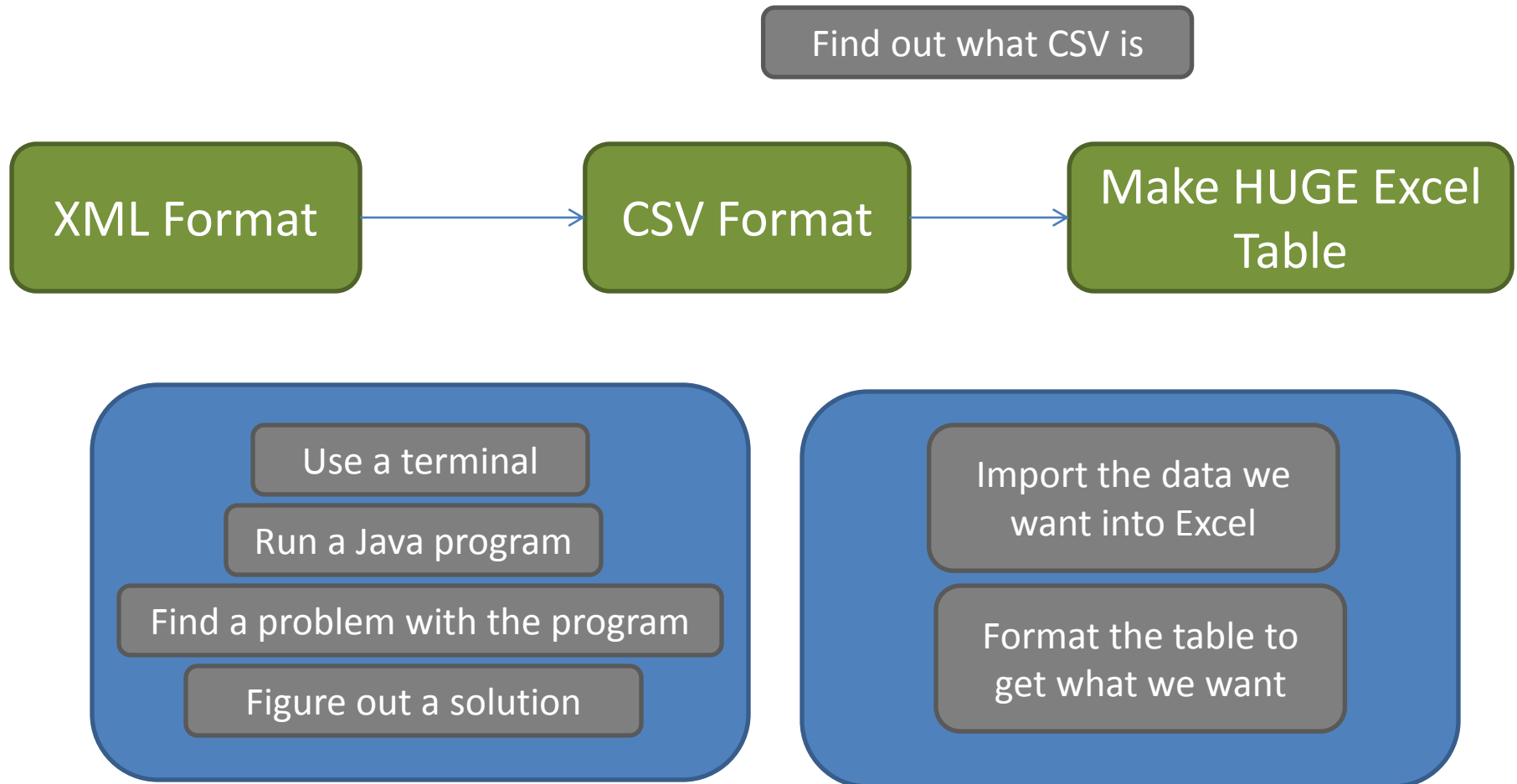
# This Class



# This is going to be a lab day

- Remember to save & email your work!
- Start bringing a flashdrive if you want.
- Confusion might arise
  - Different versions of Excel
  - Different operating systems
  - I might not have explained things clearly.
- Ask for help/clarification *at any point*.

# By the end of the day, you will have done these things



# Activity 1-1

- Do Task 0 (Get the Files We Need)

# Activity 1-1

- Do Task 0 (Get the Files We Need)
  - XML file from senate website
  - JAR file from  
<http://code.google.com/p/xml2csv-conv/downloads/detail?name=xml2csv-conv-0.0.1.zip&can=2&q=>
  - Excel Versions might be different
- Note: I call files by their file *extension*: an XML file ends with ``.xml``, a JAR file ends with ``.jar``, etc.

# Activity 1-1

- Do Task 1 (Convert an XML File to a CSV File)

# Activity 1-1

- Do Task 1 (Convert an XML File to a CSV File)
- Command Prompt Questions
  - What does `dir` or `ls` do?
  - What does `cd` do?

# Activity 1-1

- Do Task 1 (Convert an XML File to a CSV File)
- Command Prompt Questions
  - What does `dir` or `ls` do?
  - What does `cd` do?
- CSV Questions
  - What does CSV stand for?
  - Why did we put the files in the Desktop directory?

# Activity 1-1

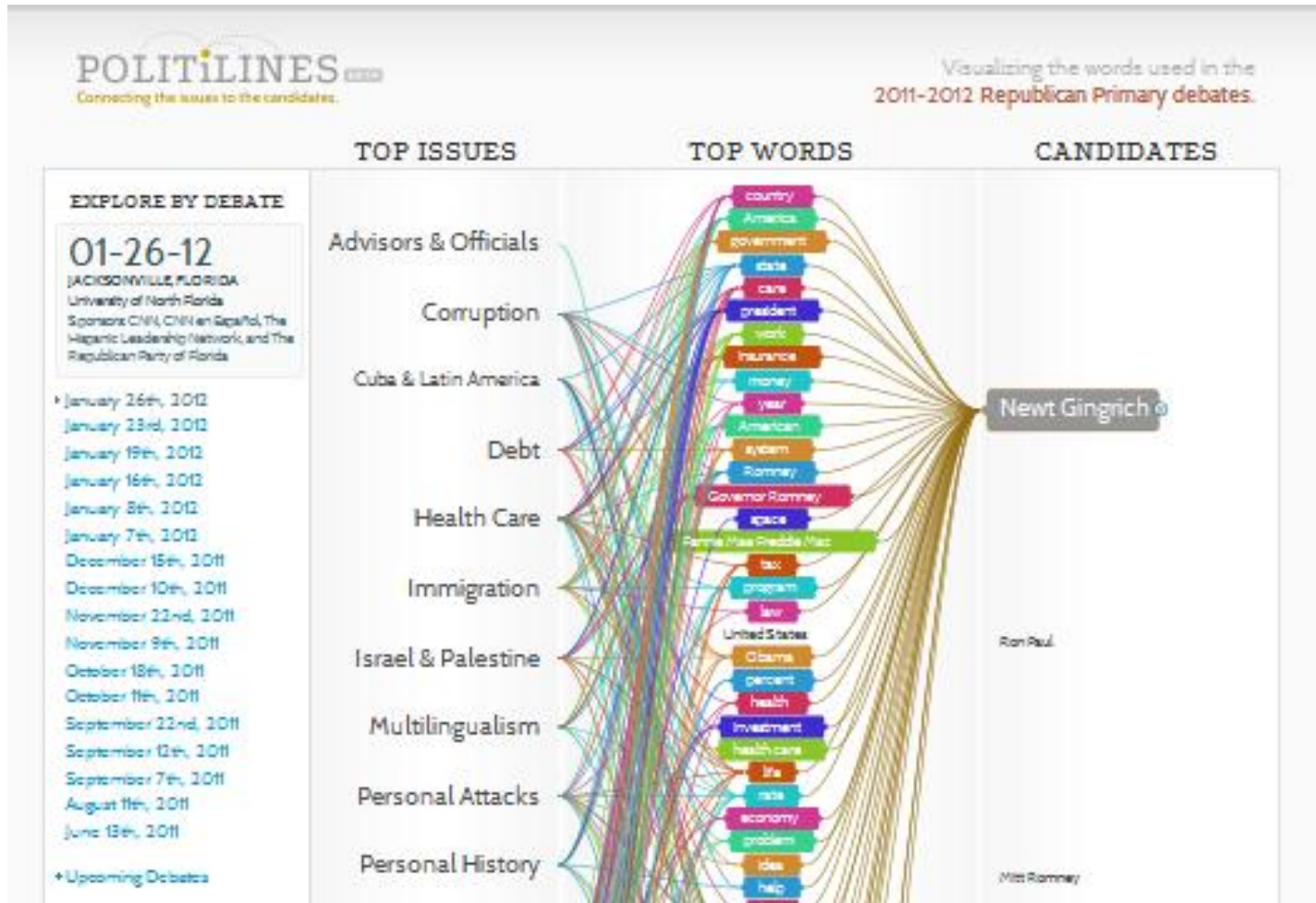
- Do Task 1 (Convert an XML File to a CSV File)
- Command Prompt Questions
  - What does `dir` or `ls` do?
  - What does `cd` do?
- CSV Questions
  - What does CSV stand for?
  - Why did we put the files in the Desktop directory?
    - *Full path name vs. local path name*

# Activity 1-1

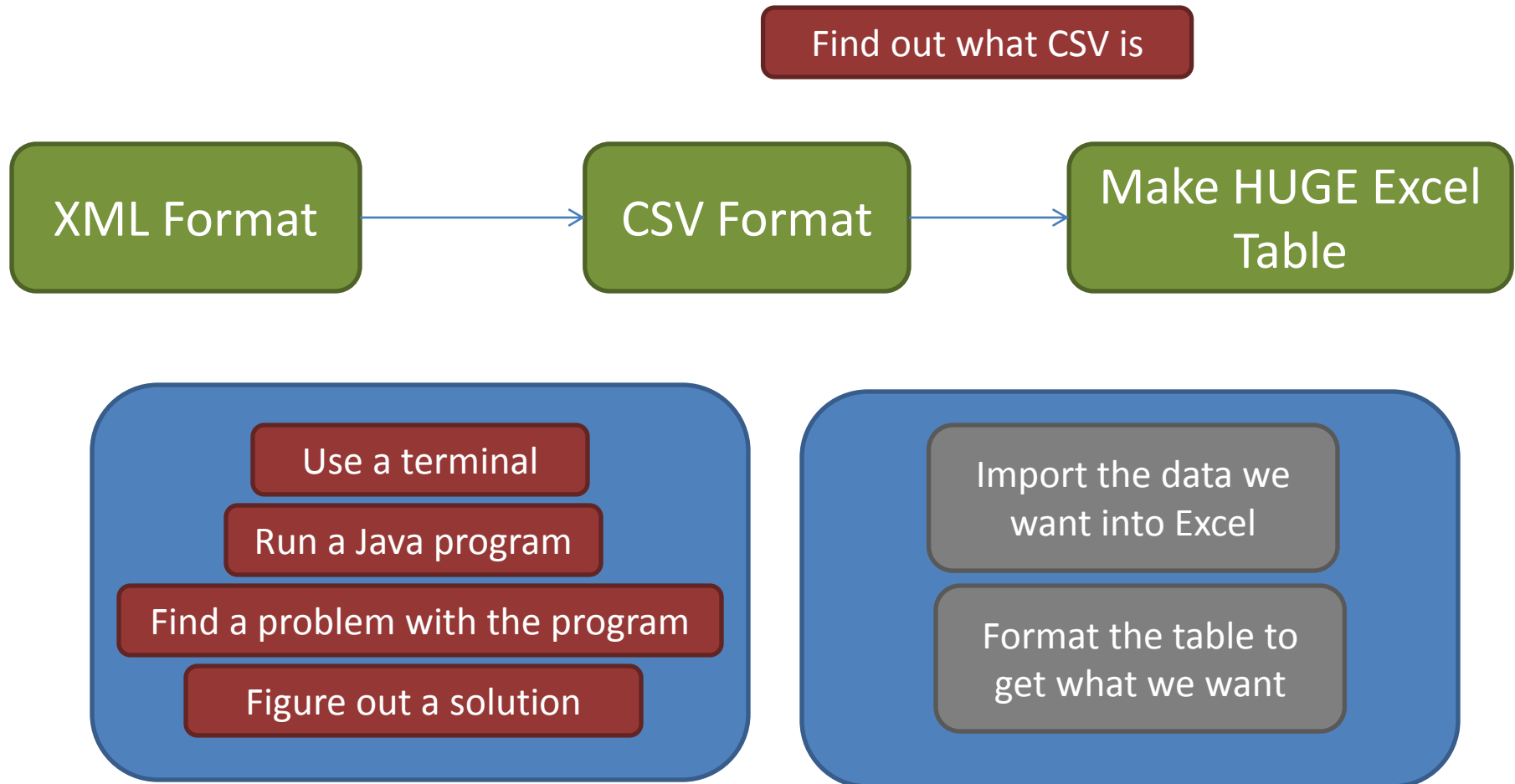
- Do Task 1 (Convert an XML File to a CSV File)
- Command Prompt Questions
  - What does `dir` or `ls` do?
  - What does `cd` do?
- CSV Questions
  - What does CSV stand for?
  - Why did we put the files in the Desktop directory?
    - *Full path name vs. local path name*
- How can we fix the bug?

# Break

# PolitiLines



# By the end of the day, you will have done these things



# Activity 1-1

- Do Task 2 (Import an XML/CSV File into Excel)

# Activity 1-1

- Do Task 3 (Format the Data)

Tip: Press 'Ctrl' and an arrow to go ALL THE WAY to the beginning/end of a row/column.

# Activity 1-1

- Do Task 3 (Format the Data)
- Data Questions:
  - # of Rows (before Trimming)?
  - # of Votes in Session 1? Session 2?

Tip: Press 'Ctrl' and an arrow to go ALL THE WAY to the beginning/end of a row/column.

# Activity 1-1

- Do Task 3 (Format the Data)
- Data Questions:
  - # of Rows (before Trimming)?
  - # of Votes in Session 1? Session 2?
- Trimming Questions:
  - # of Rows now?

Tip: Press 'Ctrl' and an arrow to go ALL THE WAY to the beginning/end of a row/column.

Tip: To get back to original sorting: Sort by both 'session' and then by 'vote\_number'

# Activity 1-1

- Task 3 (cont'd).
- Concatenation Questions
  - What are the two columns that we want?
  - What do you type in  $F2$ ?
- Numerical Vote Questions
  - What do you type in  $G2$ ? (this is kind of tricky)

# Activity 1-2

- Do the LAST Task (Make a Pivot Table)
  - Make sure you save!

# You have done these things

