How liberal is your senator?

Sept 9, 2014
Last Class: Liberal Media Bias
Last Class: Liberal Media Bias

“On the Bias.” Geoffrey Nunberg, on NPR’s Fresh Air
Today’s Class

Define Problem

Find Data

Write a set of instructions

Computer (spreadsheet)

Solution
Discuss

• If someone had asked you, in August, “How liberal is your senator?”, how would you have answered?
  – If you’re not really politically aware, what would you have done to help you give a confident answer?
Google knows all

• Google to quickly find out who was the most liberal senator in 2003 and 2007.

• Results just happen to be the democratic frontrunner. See this article for a snarky review of the details.
More serious approaches?
A few approaches to defining the problem
Problem Definition

Conservatives

Liberals
Problem Definition

Conservatives

How Liberal/Conservative is

? 

Liberals
Problem Definition

• Defining a computational problem must sometimes be influenced by the availability of data.

• More on data sources later in today’s class.
Problem Definition

• Consider a set of politicians
• Rank them somehow
Problem Definition

• Consider a set of politicians
• Rank them somehow

Let’s use the Senate:
• 100 members (two per state)
• Senators serve 6-year terms
• 112th congress (2-yr meeting)
How to rank senator liberal-ness?

• Look at the votes!
• But how do we know which bills are “liberal”???
• Sidestep the issue: “judge a man by the company he keeps” (attributed to Aesop, Euripides, Jesus, Saki, ...)
  – Doesn’t this just push the problem down to labelling the company?
Problem Definition

- They vote on the same bills

| Senator | Bill 1 | Bill 2 | Bill 3 | Bill 4 | ...
|---------|--------|--------|--------|--------|--------
| A       | Yes    | No     | Yes    | Yes    |        |
| B       | No     | Yes    | No     | Yes    |        |
| C       | Yes    | Yes    | Yes    | Yes    |        |
| D       | Yes    | No     | Yes    | Yes    |        |
| E       | No     | Yes    | No     | No     |        |
| F       | Yes    | No     | Yes    | Yes    |        |
| G       | No     | No     | No     | No     |        |
| H       | No     | Yes    | No     | No     |        |
| I       | No     | No     | No     | No     |        |
| ...     |        |        |        |        |        |
### Problem Definition

- They vote on the same bills

<table>
<thead>
<tr>
<th>Senator</th>
<th>Bill 1</th>
<th>Bill 2</th>
<th>Bill 3</th>
<th>Bill 4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>
Problem Definition
Problem Definition

• Difficult to determine whether a single senator’s voting record is “liberal” or “conservative”
Problem Definition

• Difficult to determine whether a single senator’s voting record is “liberal” or “conservative”
• But it’s easy to compare two voting records
Problem Definition

• Difficult to determine whether a single senator’s voting record is “liberal” or “conservative”
• But, easy to compare two voting records
• So, pick a senator known to be on the end of the liberal/conservative spectrum
Problem Definition

- Idea: Pick one senator and rank relative to that person

<table>
<thead>
<tr>
<th>Senator</th>
<th>Bill 1</th>
<th>Bill 2</th>
<th>Bill 3</th>
<th>Bill 4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Problem Definition

• Idea: Pick one senator and rank relative to that person.

<table>
<thead>
<tr>
<th>Senator</th>
<th>Bill 1</th>
<th>Bill 2</th>
<th>Bill 3</th>
<th>Bill 4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Same</td>
<td>Diff</td>
<td>Same</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Same</td>
<td>Same</td>
<td>Same</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Problem Definition**

- **Idea:** Pick one senator and rank relative *to that person*

<table>
<thead>
<tr>
<th>Senator</th>
<th>Bill 1</th>
<th>Bill 2</th>
<th>Bill 3</th>
<th>Bill 4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Same</td>
<td>Diff</td>
<td>Same</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Same</td>
<td>Same</td>
<td>Same</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

[Most Liberal] A D C B E [Most Conservative]
**Problem Definition**

- **Idea:** Pick one senator and rank relative *to that person*

<table>
<thead>
<tr>
<th>Senator</th>
<th>Bill 1</th>
<th>Bill 2</th>
<th>Bill 3</th>
<th>Bill 4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Same</td>
<td>Diff</td>
<td>Same</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Same</td>
<td>Same</td>
<td>Same</td>
<td>Same</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

Problem Definition
Problem Definition

Claim: We can tell how liberal or conservative a senator is compared to his/her peers using their voting record.
Problem Definition

Claim: We can tell how liberal or conservative a senator is compared to his/her peers using their voting record.

Test: Rank all senators by how similar they vote compared to a particular liberal senator.
Problem Definition

Claim: We can tell how liberal or conservative a senator is compared to his/her peers using their voting record.

Test: Rank all senators by how similar they vote compared to a particular liberal senator.

The Big Question: Who do we use?
Bernie Sanders

- Co-founder of Congressional Progressive Caucus
- Self-described democratic socialist
- More liberal than the most conservative conservatives are conservative?

Use Bernie Sanders’ votes to compare how liberal other senators are
Use Bernie Sanders' votes to compare how liberal other senators are

Define Problem

Find Data

Write a set of instructions

Computer (spreadsheet)

Solution
Let’s talk about data
Types of Data

VS.

VS.

VS.
Types of Data

Human

Computer

0101010
1010000
0101010
1001010
0101010
0101010
1001010
0100000
Hard for Computers, Easy for Humans
Hard for Computers, Easy for Humans

Who is it?
Donald Rumsfeld
Tony Blair
Colin Powell
George W. Bush

http://scikit-learn.org/0.9/auto_examples/applications/face_recognition.html
Easy for Computers, Hard for Humans
Easy for Computers, Hard for Humans

1. Divide
2. Multiply
3. Subtract
4. Bring down

\[
\begin{align*}
118 \\
6 \div 7 & \quad 08 \\
- 6 & \\
10 & \quad 48 \\
- 6 & \\
48 & \quad 0
\end{align*}
\]
Easy for Computers, Hard for Humans
Easy for Computers, Hard for Humans

Confusion at Palm Beach County polls
Some Al Gore supporters may have mistakenly voted for Pat Buchanan because of the ballot's design.

Although the Democrats are listed second in the column on the left, they are the third hole on the ballot.

Punching the second hole casts a vote for the Reform party.

http://www.mit.edu
Horribly hard for computers

- But closed-caption data may be available...
Today's Class

Use Bernie Sanders' votes to compare how liberal other senators are.

Define Problem

Find Data

Write a set of instructions

Computer (spreadsheet)

Solution
Break
Finding Data

This data is freely accessible and online.

<table>
<thead>
<tr>
<th>Senator</th>
<th>Bill 1</th>
<th>Bill 2</th>
<th>Bill 3</th>
<th>Bill 4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
URL

• “Universal Resource Locator”
• The stuff in the browser’s “Address bar”
• In our case:

http://www.senate.gov/legislative/LIS/roll_call_lists/roll_call_votete_cfm.cfm?congress=113&session=2&vote=00008
Examining the URL

http://www.senate.gov/legislative/

... It’s from the senate

LIS/roll_call_lists/

... It’s probably listed by senator

roll_call_vote_cfm.cfm?

... It’s from the senate. “cfm” = “coldFusion markup”, an Adobe product

congress=113&session=2&vote=00008

... Stuff after “?” is a “query” to a database

113\textsuperscript{th} congress, 2\textsuperscript{nd} session, 8\textsuperscript{th} vote
How could we fill this table out by hand? What are the instructions?
How long will it take to make this table?

<table>
<thead>
<tr>
<th>Senator</th>
<th>Bill 1</th>
<th>Bill 2</th>
<th>Bill 3</th>
<th>Bill 4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• **For Each** Bill:
  • Navigate to the webpage (using the URL trick)
  • **For Each** Senator:
    • **Record** the senator’s vote
In practice...

• Could write a program to do this
  – But you don’t have that skill yet
• Could get lucky and find data in another form
Fortunately, this data is already recorded online
Fortunately, this data is already recorded online

(How do you think your web browser is able to show you the vote page?)
XML: Extensible Markup Language
OK for Humans, OK for Computers

First look at the structure
XML: Extensible Markup Language
OK for Humans, OK for Computers

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<roll_call_vote>
  <congress>113</congress>
  <session>2</session>
  ...
  <members>
    <member>
      <member_full>Alexander (R-TN)</member_full>
      ...
      <vote_cast>Yea</vote_cast>
      ...
    </member>
    ...
  </members>
</roll_call_vote>
```
XML: Extensible Markup Language
OK for Humans, OK for Computers

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<roll_call_vote>
  <congress>113</congress>
  <session>2</session>
  ...
  <members>
    <member>
      <member_full>Alexander (R-TN)</member_full>
      ...
      <vote_cast>Yea</vote_cast>
      ...
    </member>
    ...
  </members>
</roll_call_vote>
```
XML very widely used

All recent microsoft office programs (.pptx, .docx, ...) use it (and hide that from you).
Use Bernie Sanders’ votes to compare how liberal other senators are.
Now

• On the course webpage, open up activity 1-1 ("ACT 1-1"), and do what it says.
• Work with your neighbors as you get stuck
• Ask us questions
• Don’t expect to finish!
Next Class

Define Problem

Use Bernie Sanders’ votes to compare how liberal other senators are

Find Data

XML Format

Write a set of instructions

Make a HUGE spreadsheet table

Solution

Computer (spreadsheet)