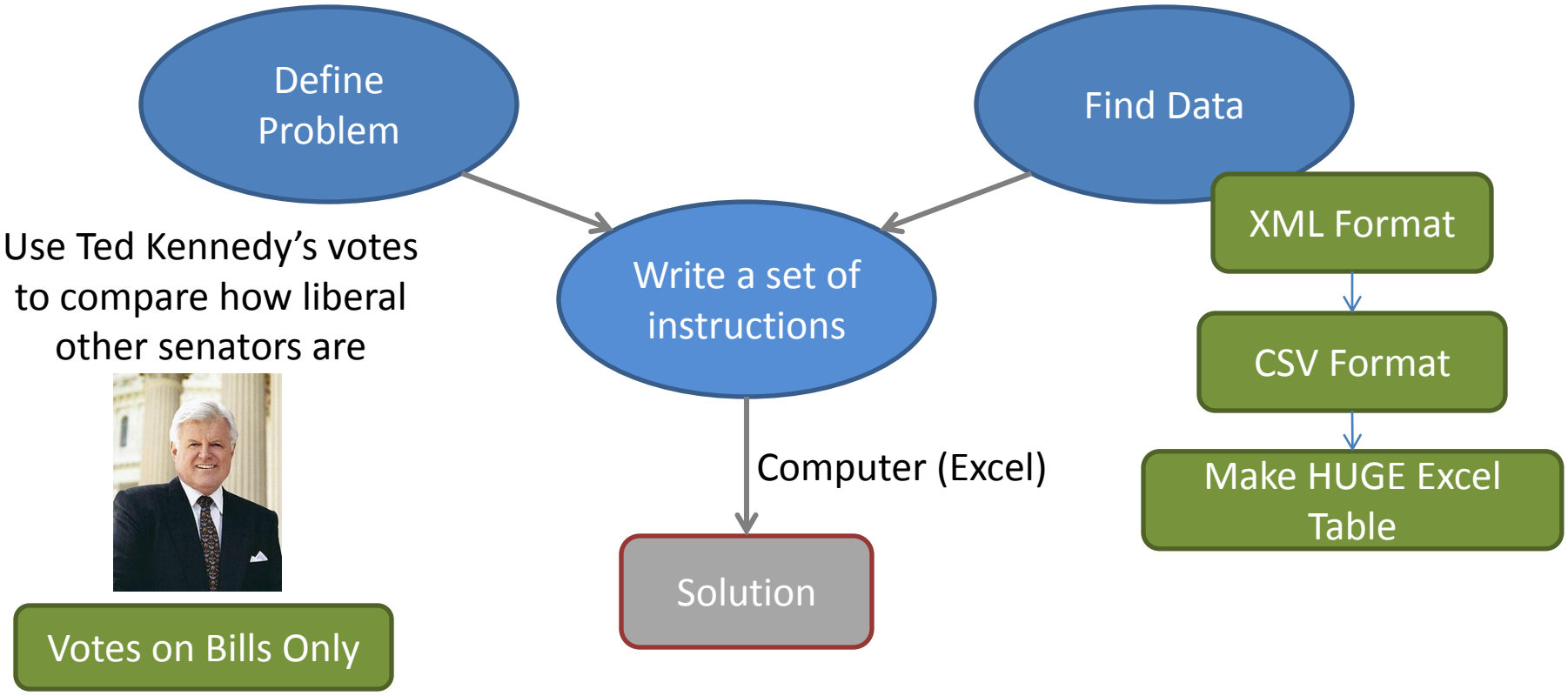
Ranking Senators with Senator X's Votes

Feb 9, 2012

Last Class


Define Problem
Find Data
Use Ted Kennedy's votes to compare how liberal other senators are
Write a set of instructions
XML Format
CSV Format
Make HUGE Excel Table
Computer (Excel)
Solution
Votes on Bills Only

Last Class: Completed Task 1

Compare Votes to Ted Kennedy

In cell `B2`:
`1` if Akaka's vote is same as Kennedy's
`-1` if Akaka's vote is different from Kennedy's
`0` if either Akaka or Kennedy did not vote

Last Class: Completed Task 1

Compare Votes to Ted Kennedy

In cell `B2`:
`1` if Akaka's vote is same as Kennedy's
`-1` if Akaka's vote is different from Kennedy's
`0` if either Akaka or Kennedy did not vote

Akaka's Cell: PivotTable!B5=0
Kennedy's Cell: PivotTable!B$60

```
=IF(OR(Akaka=0,Kennedy=0),0,IF(Akaka=Kennedy,1,-1))
```

Last Class: Completed Task 1

Compare Votes to Ted Kennedy

In cell `B2`:
`1` if Akaka's vote is same as Kennedy's
`-1` if Akaka's vote is different from Kennedy's
`0` if either Akaka or Kennedy did not vote

Akaka's Cell: PivotTable!B5=0
Kennedy's Cell: PivotTable!B$60

```
=IF(OR(Akaka=0,Kennedy=0),0,IF(Akaka=Kennedy,1,-1))
```

```
=IF(OR(PivotTable!B5=0,PivotTable!B$60=0),0,IF(PivotTable!B5=PivotTable!B$60,1,-1))
```

Task 2: Rank Each Senator

1. In `KennedyCompare`, add three columns on the right-hand side (use the `CountIf` function):

   `num_agree` counts agreements with Kennedy's votes

   `num_disagree` counts disagreements with Kennedy's votes

   `num_not_voting` counts the votes where either senator was absent.

Task 2: Rank Each Senator

1. In `KennedyCompare`, add three columns on the right-hand side (use the `CountIf` function):

   `num_agree` counts agreements with Kennedy's votes

   `num_disagree` counts disagreements with Kennedy's votes

   `num_not_voting` counts the votes where either senator was absent.

num_agree: `=COUNTIF(B2:AW2,1)`

num_disagree : `=COUNTIF(B2:AW2,-1)`

num_not_voting: `=COUNTIF(B2:AW2,0)`

Task 2: Rank Each Senator

2. Then add a column - `rank` - that reports

$$\frac{\text{num\_agree} - \text{num\_disagree}}{\text{num\_agree} + \text{num\_disagree}}.$$

What are the minimum and maximum values of this function?

Task 2: Rank Each Senator

2. Then add a column - `rank` - that reports

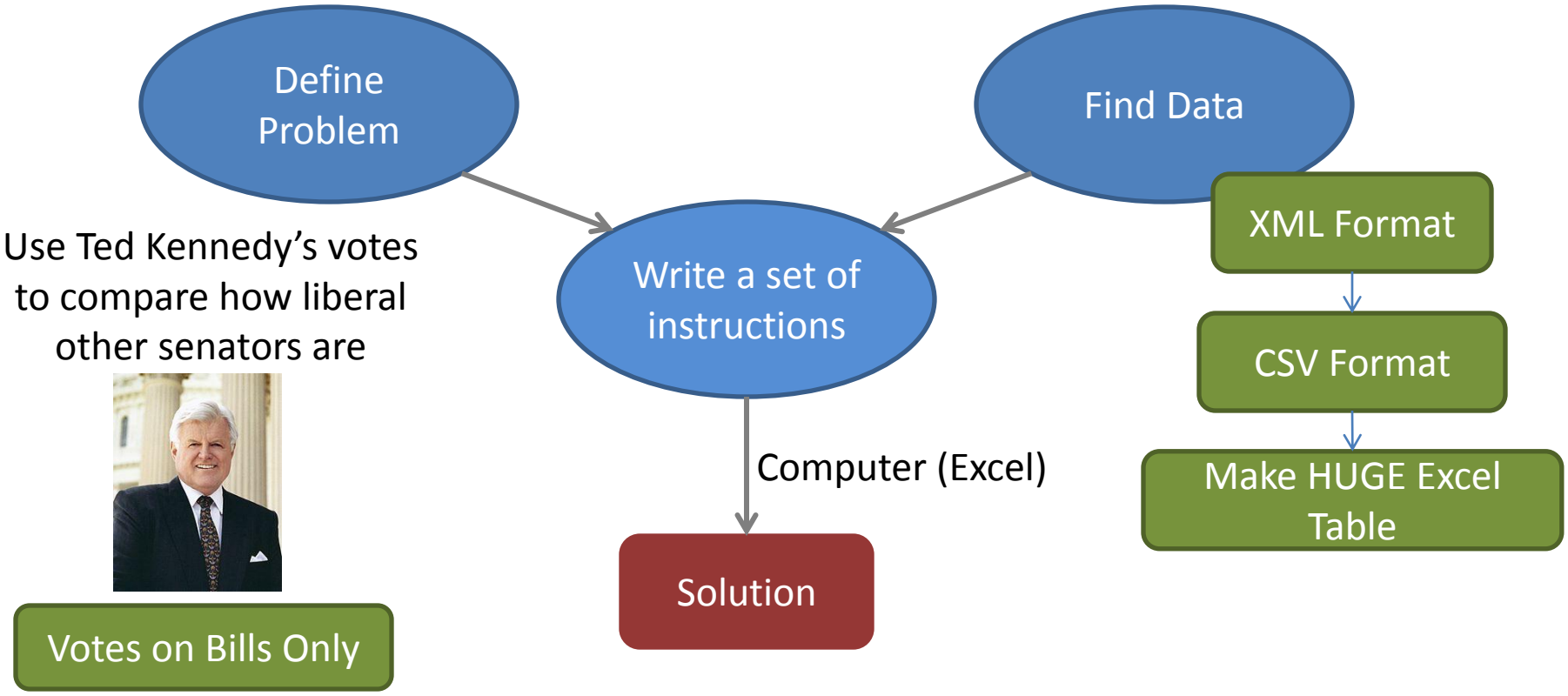
$$\frac{\text{num\_agree} - \text{num\_disagree}}{\text{num\_agree} + \text{num\_disagree}}.$$

   What are the minimum and maximum values of this function?

3. For readability, copy the row headings to a column after `rank`.

4. Select the entire table and sort in decreasing `rank`. Look at the resulting data. Do you see anything surprising? Try to explain any surprises.

What We've Accomplished



Break (Index Cards)

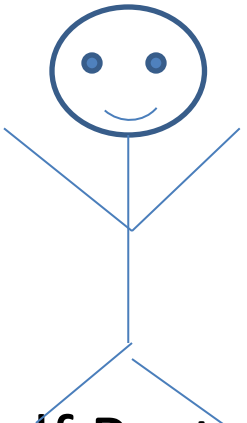
Name

Class Year

Concentration

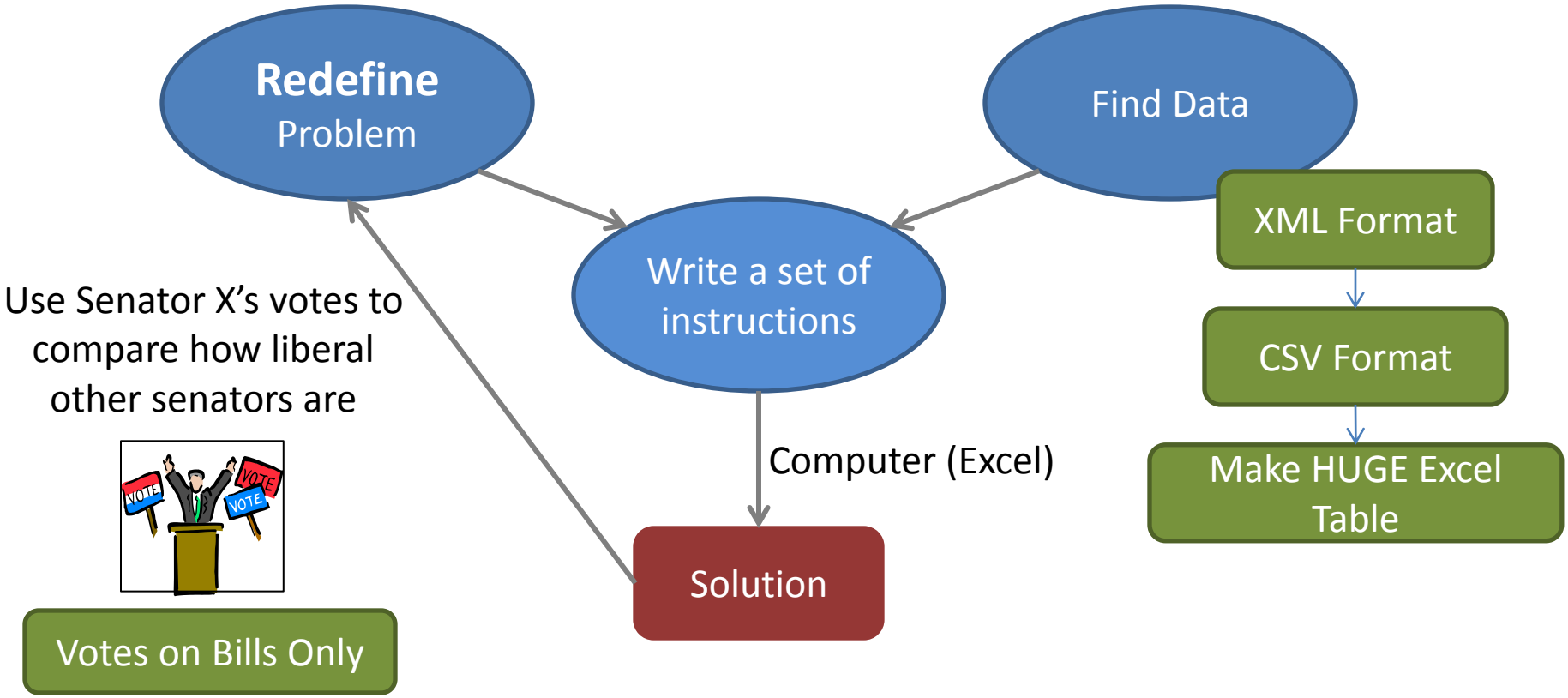
Where You're From

Something Interesting About Yourself

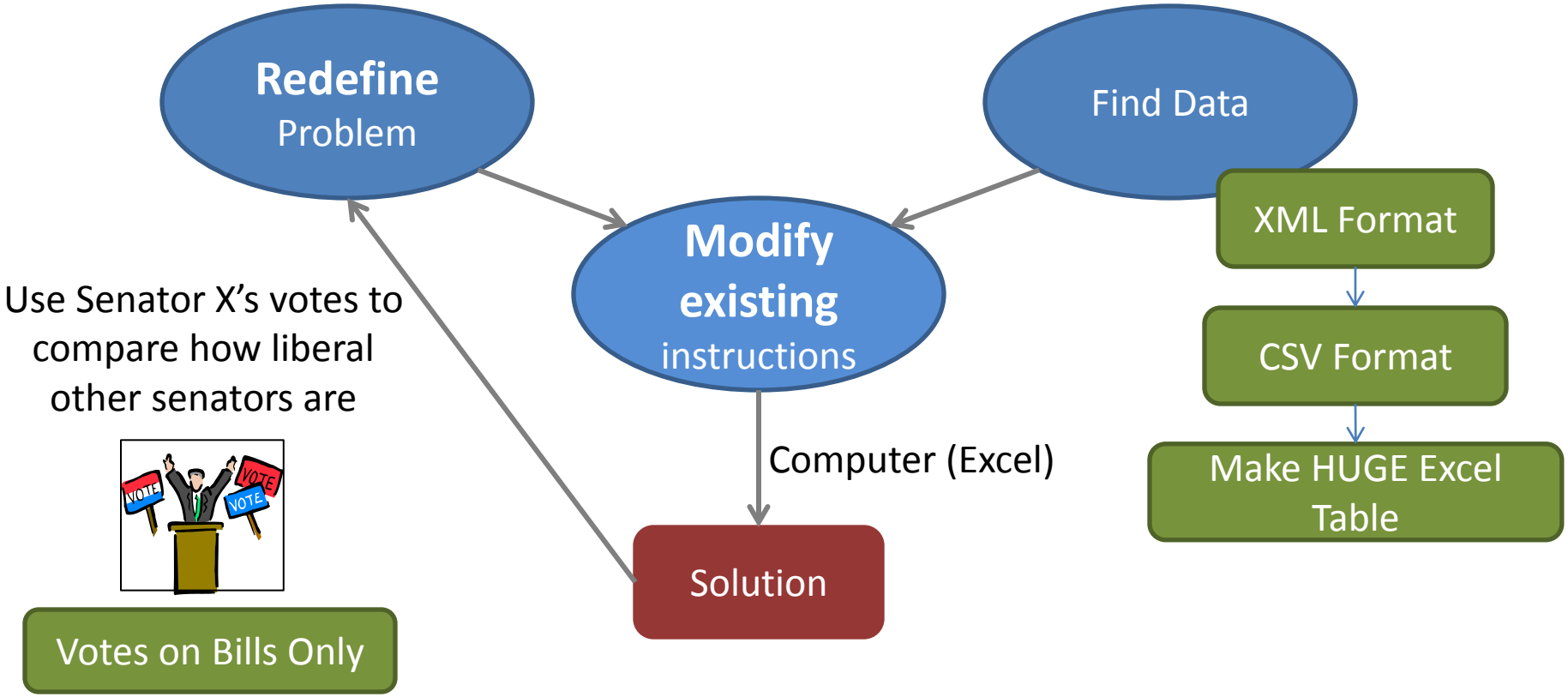


Self Portrait

New Problem


Redefine Problem
Find Data
XML Format
Write a set of instructions
CSV Format
Use Senator X's votes to compare how liberal other senators are
Computer (Excel)
Make HUGE Excel Table
VOTE
VOTE
VOTE
Solution
Votes on Bills Only

New Problem


Redefine Problem
Find Data
XML Format
Modify existing instructions
CSV Format
Use Senator X's votes to compare how liberal other senators are
Computer (Excel)
Make HUGE Excel Table
VOTE
VOTE
VOTE
Solution
Votes on Bills Only

Activity 1-3

- Do Task 1 with a partner.

OFFSET: Returns a reference to a range that is a specified number of rows and columns from a cell or range of cells. The reference that is returned can be a single cell or a range of cells.

```
OFFSET(reference, rows, cols, [height],[width])
```

Activity 1-3

- Do Task 2

Activity 1-3

- Do Task 2
  - We broke this task down into manageable pieces.
  - Nice formatting is useful
  - This spreadsheet is now useful for other people

Project 1

Due Dates
Proposal: Th 2/16
Project: Th 2/23

- Choose a problem related to the work we've done so far in the course
  - Choose a different dataset and perform a similar analysis
  - Perform a different analysis on the Senate dataset to answer a different political question

Project 1

Due Dates
Proposal: Th 2/16
Project: Th 2/23

- Choose a problem related to the work we've done so far in the course
  - Choose a different dataset and perform a similar analysis
  - Perform a different analysis on the Senate dataset to answer a different political question
- You must have a **testable** hypothesis!
  - "Senators who vote similarly to Kennedy are Democrats and senators who vote differently from Kennedy are Republicans"

Project 1

Due Dates
Proposal: Th 2/16
Project: Th 2/23

- Proposal
  - Background
  - Claim
  - Data
  - Analysis Steps
  - Potential Roadblocks

Meet with Tas/Faculty! Ask Questions!

Project 1

Due Dates
Proposal: Th 2/16
Project: Th 2/23

- Proposal
  - Background
  - Claim
  - Data
  - Analysis Steps
  - Potential Roadblocks
- Project
  - All Excel Spreadsheets and a **website**
- Interactive Grading

Meet with Tas/Faculty! Ask Questions!

Project 1

Due Dates
Proposal: Th 2/16
Project: Th 2/23

Category	# Points Earned
Proposal	15
Degree of Difficulty	25
Execution	25
Code Quality	10
Website	15
Interactive Grading	10
TOTAL	100

Proposal Rubric

Proposal (15%)

_____ (2 points) Problem is placed in context and the overview is concisely stated.

_____ (3 points) The claim poses a problem that can be answered with computation. "I will rank all senators" is not a claim.

_____ (1 point) The data source is specified

_____ (2 points) There is a good description of what the original data looks like.

_____ (5 points) The steps of the analysis are clear, specfic, and manageable. "I will import the data and cluster according to votes" is not clear. "I will import all the data for all 2,012 congressional meetings" is not manageable.

_____ (2 points) There are some roadblocks listed, and no obvious roadblocks are missed. Obvious roadblocks are things like "I have never seen this type of data format before" or "I want to perform a particular statistical test, but I don't know how to write that formula."

_____ Total

Purchase Patterns of Political Books

- Click on the Book URL link on the class website