Ranking Senators
Project 1

https://cs.brown.edu/courses/csci0030/projects/project1.html
Project Description

• Use the Senate data to answer a different question than we explored in class, or

• Find another dataset to which you can extend the analytical methods we used on the Senate data
Follow our Problem Solving Workflow

1. Define question
2. Select appropriate dataset
3. Design a valid method
4. Analyze your results
5. Communicate your findings
Question Needs to be Clear

• Use clear precise terms

• Questions need to be answerable

• “I will rank the senators according to their political leanings” is not a valid question. It’s just a task they may provide some insight

• “Do Blue Dog Democrats (Democrats from southern states) vote more similarly to Democrats or Republicans?” is a good question
Dataset Needs to be Relevant

• Data should be comprehensive but manageable
• We limited the amount of data we collected for the Senate in class
• For the project we would expect all the votes of a single year, or the votes on bills from multiple years
• Ensure that the data selection is relevant to the question being asked
Method Needs to be Clearly Defined

• In the proposal, outline the steps that you will take

• The steps you list in your proposal should have enough definition that another student could follow them and answer your question

• The steps are likely to change after the proposal, and that is ok
Proposal

• Background: Describe why we should care about your problem
• Claim: What question will you plan to answer
• Dataset: What the data represents, the data location, and what it looks like as it is now

• Steps of Methodology
  • Importing and formatting your data
  • Formulas you will use
  • Presentation and visualization of your results
  • How will the end result validate or invalidate your question

• Potential Roadblocks: Describe problems that you will likely face and might struggle to solve
Negative results are OK

• If your methods and data do not support your hypothesis, that is OK

• Negative results happen all the time, and though they are rarely published, they are still significant

• A negative result still needs to be supportable by your data and methods

• A negative result that occurs due to insufficient or sparse data or a poor method is not sufficient for this project
Build a Website

- To communicate results for your project, you will be building a simple, single-page website

- Help for building your website is available on the resources tab of the course page

- Your website can include information you put in your proposal

- Should also include effective visualizations

- Discuss your findings and reflect on the project as a whole
Follow the Rubric

https://cs.brown.edu/courses/csci0030/projects/PROJ1_Rubric.pdf
Spreadsheet Components
SUMPRODUCT(range1, range2,...)

- Multiplies to cell ranges together, element-wise
- Sums the resulting product

$$\begin{array}{ccc}
1 & 2 & 2 \\
2 & 4 & 8 \\
3 & 8 & 24 \\
4 & 16 & 64 \\
\end{array}$$

SUMPRODUCT = 98
## Financial

<table>
<thead>
<tr>
<th>Item</th>
<th>Apples</th>
<th>Orange</th>
<th>Bananas</th>
<th>Plums</th>
<th>Grapefruit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Cost</td>
<td>$1.00</td>
<td>$0.75</td>
<td>$0.50</td>
<td>$1.25</td>
<td>$1.50</td>
</tr>
</tbody>
</table>

Total cost = SUMPRODUCT(B2:F2, B3:F3) = $24.75

<table>
<thead>
<tr>
<th>Symbol</th>
<th>AAPL</th>
<th>GOOG</th>
<th>FB</th>
<th>V</th>
<th>MSFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Price</td>
<td>$100</td>
<td>$800</td>
<td>$130</td>
<td>$80</td>
<td>$60</td>
</tr>
</tbody>
</table>

Total Portfolio Value = SUMPRODUCT(B2:F2, B3:F3) = $2790
### Change of Base Examples

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Digits</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Place Values</td>
<td>10000</td>
<td>1000</td>
<td>100</td>
<td>10</td>
</tr>
</tbody>
</table>

\[
\text{SUMPRODUCT}(B1:F1, B2:F2) = 12477
\]

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Digits</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Place Values</td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

\[
\text{SUMPRODUCT}(B1:F1, B2:F2) = 22
\]
Matrices

• 2-dimensional data structures with a specified number of rows and columns

• In a spreadsheet, a cell range can be used directly as a matrix

• When describing a matrix’s size, rows are given first then column. A 5x4 matrix is 5 rows and 4 columns

• A matrix’s data type should be consistent throughout the entire matrix
Matrix addition and subtraction

• We’ve already been using addition and subtraction

• For addition/subtraction, the matrices need to be the same size
Matrix Operations

- Component-wise multiplication and division, applies to the components at the same row and column index
- Matrices need to be the same size

\[
\begin{array}{ccc}
1 & 2 & \circ \\
2 & 3 & 2 \\
3 & 4 & 2 \\
\end{array}
\begin{array}{ccc}
2 & 2 & \\
2 & 2 & \\
2 & 2 & \\
\end{array}
= 
\begin{array}{ccc}
2 & 4 & \\
4 & 6 & \\
6 & 8 & \\
\end{array}
\]
Component-Wise operations in Spreadsheets

- Addition
  \[ =\text{ARRAYFORMULA}(A1:C10 + D1:F10) \]
- Subtraction
  \[ =\text{ARRAYFORMULA}(A1:C10 - D1:F10) \]
- Multiplication
  \[ =\text{ARRAYFORMULA}(A1:C10 \times D1:F10) \]
- Division
  \[ =\text{ARRAYFORMULA}(A1:C10 - D1:F10) \]
Matrix multiplication

• For each output cell, its value is the sum product of its corresponding row from the first matrix and its corresponding column in the second matrix.
Matrix multiplication

- For each output cell, its value is the sum product of its corresponding **row** from the first matrix and its corresponding **column** in the second matrix.
Matrix multiplication

• For each output cell, its value is the sum product of its corresponding row from the first matrix and its corresponding column in the second matrix.
Rules for matrix multiplication

- The number of columns of the first matrix must equal the number of rows in the second matrix.
- The size of the output matrix is the number of rows from the first matrix and the number of columns from the second matrix.
- Example: $M^{3x4} \times N^{4x5} = O^{3x5}$
Remember from ACT1-2

• Each vote in the senate was categorized as 1, 0, -1

• Agreement or Disagreement was then just the product of two senators votes

• To compute the difference between the agreements and disagreements (A - D), we summed the agreements and disagreements
Using Matrix Multiplication with Senators

Coded Votes  100 rows x N columns

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexander</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ayotte</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Baldwin</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>Barrasso</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Coded Votes Transposed  N rows x 100 columns

<table>
<thead>
<tr>
<th></th>
<th>Alexander</th>
<th>Ayotte</th>
<th>Baldwin</th>
<th>Barrasso</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\text{Coded Votes} & \quad \begin{pmatrix}
\text{Alexander} & \text{Ayotte} & \text{Baldwin} & \text{Barrasso} \\
1 & -1 & -1 & 1 \\
2 & 1 & 1 & 1 \\
3 & 1 & 1 & -1 \\
\end{pmatrix} \\
\times \\
\begin{pmatrix}
\text{Alexander} \\
\text{Ayotte} \\
\text{Baldwin} \\
\text{Barrasso} \\
\end{pmatrix}
\end{align*}
\]
Using Matrix Multiplication with Senators

Coded Votes  100 rows x N columns

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexander</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ayotte</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Baldwin</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>Barrasso</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Coded Votes Transposed  N rows x 100 columns

<table>
<thead>
<tr>
<th>Alexander</th>
<th>Ayotte</th>
<th>Baldwin</th>
<th>Barrasso</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

\[ \begin{bmatrix}
    1 & -1 & 1 \\
    2 & 1 & 1 \\
    3 & 1 & -1
\end{bmatrix}
\times
\begin{bmatrix}
    -1 & -1 & 1 & -1 \\
    1 & 1 & 1 & 1 \\
    1 & -1 & 1 & 1
\end{bmatrix}
\]

= 

<table>
<thead>
<tr>
<th></th>
<th>Alexander</th>
<th>Ayotte</th>
<th>Baldwin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexander</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ayotte</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baldwin</td>
<td></td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Coded Votes Transposed   N rows x 100 columns
Using Matrix Multiplication with Senators

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alexander</strong></td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Ayotte</strong></td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Baldwin</strong></td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td><strong>Barrasso</strong></td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Coded Votes Transposed

<table>
<thead>
<tr>
<th></th>
<th>Alexander</th>
<th>Ayotte</th>
<th>Baldwin</th>
<th>Barrasso</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

\[
\begin{array}{ccc}
  & \text{Alexander} & \text{Ayotte} & \text{Baldwin} & \text{Barrasso} \\
\hline
\text{Alexander} & 1 & -1 & 1 & -1 \\
\text{Ayotte} & 1 & 1 & 1 & 1 \\
\text{Baldwin} & 1 & 1 & -1 & 1 \\
\end{array}
\]

\[
\begin{array}{ccc}
  & \text{Alexander} & \text{Ayotte} & \text{Baldwin} \\
\hline
\text{Alexander} & 3 & 3 & -1 \\
\text{Ayotte} & 3 & 3 & -1 \\
\text{Baldwin} & -1 & -1 & 3 \\
\end{array}
\]

Coded Votes

\[=\]
mmult(range1,range2)

- Computes the matrix multiplication of two matrices
- Number of columns in the first matrix must equal the number of rows in the second matrix
Reminders for Homework

• You’ll be asked to import a lot of votes with importXML

• Import only 10 columns at a time

• Copy the values to a new sheet, using ‘Paste special’ -> ‘values only’

• There are several possible methods to solving task 5, instead of trying to solve it using a single formula, break down the task into several smaller problems.