

Lists of Lists & Hypothesis Testing

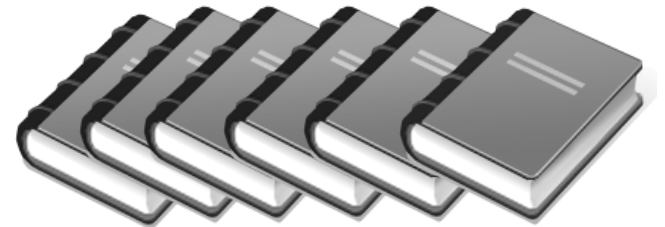
April 5, 2012

Determining Authorship

Define Problem

Find Data

Write a set of instructions



1 Book from Alice in Wonderland
5 Books from Wizard of Oz



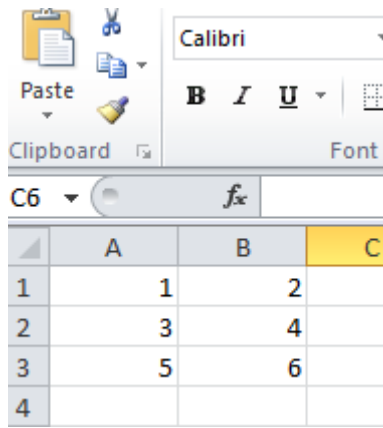
Python

Solution

Discern the **Outlier**:
The one book that is
NOT in the series of
the others.

Table Representations

Excel



The image shows a screenshot of the Microsoft Excel interface. At the top, the 'Clipboard' and 'Font' tabs are visible. The 'Clipboard' tab shows 'Paste' and 'Clipboard' options. The 'Font' tab shows 'Calibri' as the selected font, with options for Bold (B), Italic (I), and Underline (U). Below the ribbon, the active cell is C6, and the formula bar is empty. The spreadsheet grid shows a table with 4 rows and 3 columns. The columns are labeled A, B, and C. The rows are labeled 1, 2, 3, and 4. The data in the table is as follows:

	A	B	C
1	1	2	
2	3	4	
3	5	6	
4			

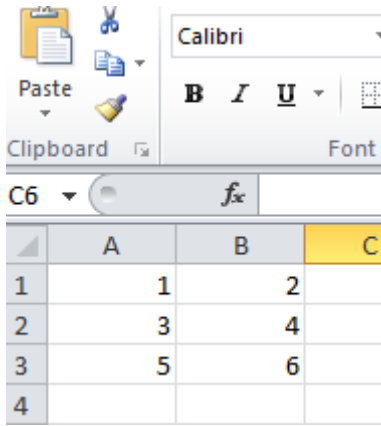
Rows: Indexed by 1,2,3

Cols: Indexed by A,B

Cell B4 contains 4

Table Representations

Excel



	A	B	C
1	1	2	
2	3	4	
3	5	6	
4			

Rows: Indexed by 1,2,3

Cols: Indexed by A,B

Cell B4 contains 4

Python

```
myList = [[1,2], [3,4], [5,6]]
```

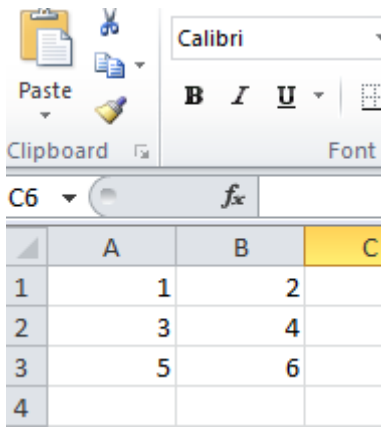
Rows: Indexed by ?

Cols: Indexed by ?

How do you
index to get 4?

Table Representations

Excel



	A	B	C
1	1	2	
2	3	4	
3	5	6	
4			

Rows: Indexed by 1,2,3

Cols: Indexed by A,B

Cell B4 contains 4

Python

```
myList = [[1,2], [3,4], [5,6]]
```

Rows: Indexed by 0,1,2

Cols: Indexed by 0,1

How do you

index to get 4?

```
myList[1][1]
```

Activity 3-1

- Task 1: build a list of lists

```
[ [0, 1, 2], [1, 2, 3], [2, 3, 4], [3, 4, 5], [4, 5, 6] ]
```

Activity 3-1

- Task 1: build a list of lists

```
[ [0, 1, 2], [1, 2, 3], [2, 3, 4], [3, 4, 5], [4, 5, 6] ]
```

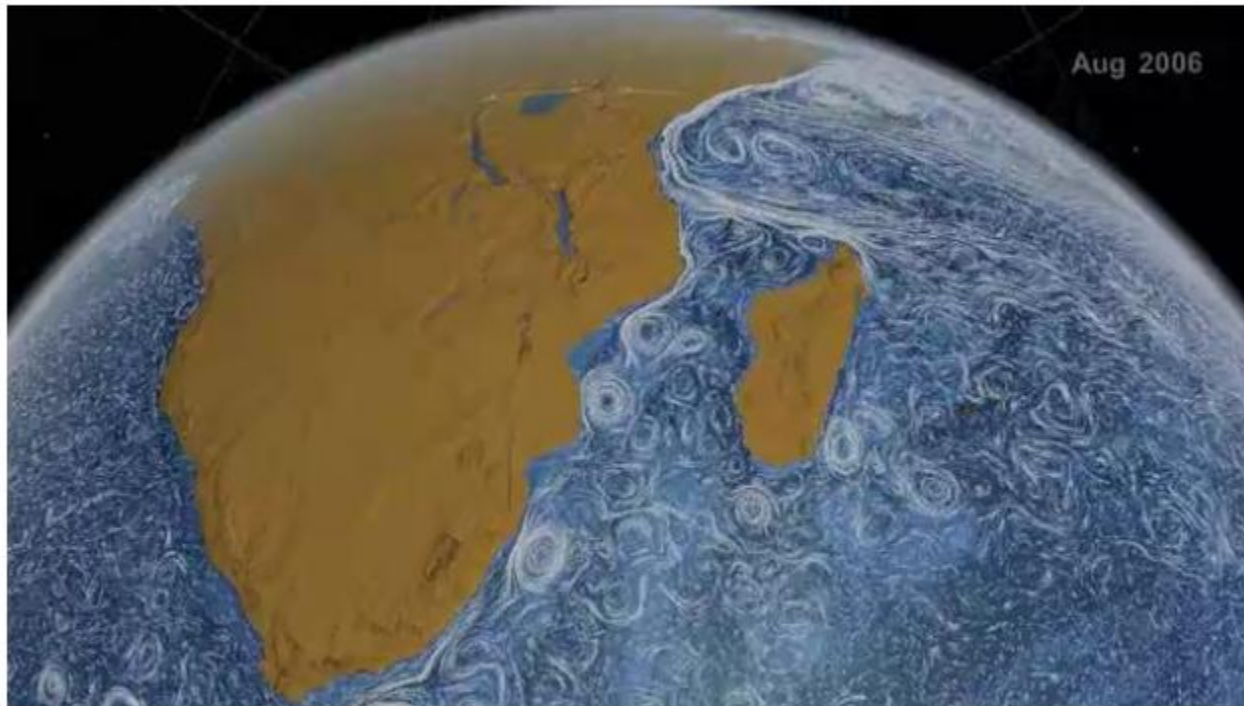
- Task 2: Write list of lists to a CSV file
 - Separate columns by commas
 - Separate rows by newlines (`'\n'`)

Break

Perpetual Ocean

[PerpetualOcean](#)

March 27, 2012 to Mapping · Add Comment · Share on Twitter



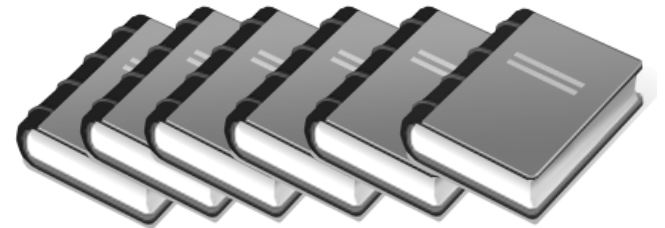
Using a computational model called Estimating the Circulation and Climate of the Ocean, Phase II (ECCO2), the NASA Goddard Space Flight Center Scientific Visualization Studio (I think NASA has a thing for long names.) [visualizes surface currents around the world](#). This is beautiful science here. Make sure you turn on high-def and go full screen.

Determining Authorship

Define Problem

Find Data

Write a set of instructions



1 Book from Alice in Wonderland
5 Books from Wizard of Oz



Python

Solution

Discern the **Outlier**:
The one book that is
NOT in the series of
the others.

Determining Authorship

1		file1.txt	file2.txt	file3.txt	file4.txt	file5.txt	file6.txt		Row Sum
2	file1.txt	0	0.001414	0.000738	0.000742	0.001173	0.000881		0.004949
3	file2.txt	0.001414	0	0.001419	0.001364	0.001521	0.001454		0.007172
4	file3.txt	0.000738	0.001419	0	0.000605	0.000933	0.000667		0.004362
5	file4.txt	0.000742	0.001364	0.000605	0	0.001112	0.000779		0.004603
6	file5.txt	0.001173	0.001521	0.000933	0.001112	0	0.001236		0.005975
7	file6.txt	0.000881	0.001454	0.000667	0.000779	0.001236	0		0.005018

File	Title	Series	Author
file1.txt			
file2.txt			
file3.txt			
file4.txt			
file5.txt			
file6.txt			

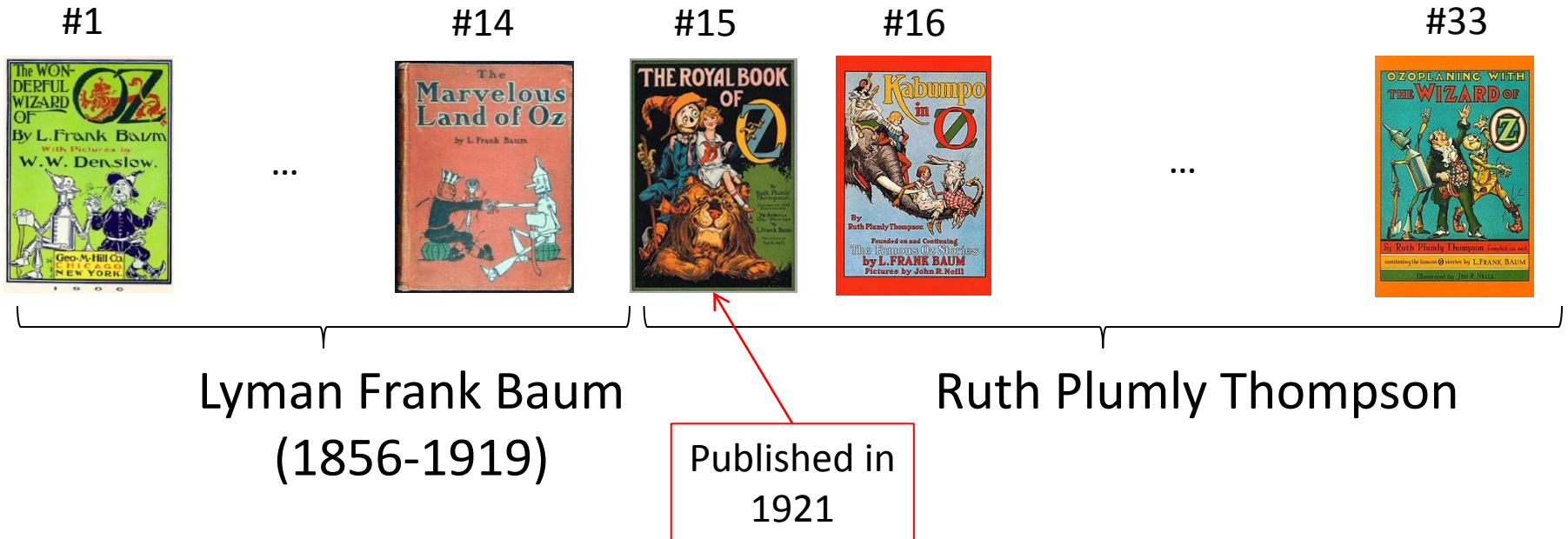
Determining Authorship

1		file1.txt	file2.txt	file3.txt	file4.txt	file5.txt	file6.txt		Row Sum
2	file1.txt	0	0.001414	0.000738	0.000742	0.001173	0.000881		0.004949
3	file2.txt	0.001414	0	0.001419	0.001364	0.001521	0.001454		0.007172
4	file3.txt	0.000738	0.001419	0	0.000605	0.000933	0.000667		0.004362
5	file4.txt	0.000742	0.001364	0.000605	0	0.001112	0.000779		0.004603
6	file5.txt	0.001173	0.001521	0.000933	0.001112	0	0.001236		0.005975
7	file6.txt	0.000881	0.001454	0.000667	0.000779	0.001236	0		0.005018

File	Title	Series	Author
file1.txt	Wonder Wizard of Oz	Oz	
file2.txt	Alice's Adventures in Wonderland	Alice in Wonderland	
file3.txt	Dorothy and the Wizard in Oz	Oz	
file4.txt	Emerald City of Oz	Oz	
file5.txt	Royal Book of Oz	Oz	
file6.txt	Glinda of Oz	Oz	

The Wizard of Oz

- About 40 Books, written by 7 different authors



<http://www.ssc.wisc.edu/~zzeng/soc357/OZ.pdf>

Determining Authorship

1		file1.txt	file2.txt	file3.txt	file4.txt	file5.txt	file6.txt		Row Sum
2	file1.txt	0	0.001414	0.000738	0.000742	0.001173	0.000881		0.004949
3	file2.txt	0.001414	0	0.001419	0.001364	0.001521	0.001454		0.007172
4	file3.txt	0.000738	0.001419	0	0.000605	0.000933	0.000667		0.004362
5	file4.txt	0.000742	0.001364	0.000605	0	0.001112	0.000779		0.004603
6	file5.txt	0.001173	0.001521	0.000933	0.001112	0	0.001236		0.005975
7	file6.txt	0.000881	0.001454	0.000667	0.000779	0.001236	0		0.005018

File	Title	Series	Author
file1.txt	Wonder Wizard of Oz	Oz	Lyman Frank Baum
file2.txt	Alice's Adventures in Wonderland	Alice in Wonderland	Lewis Carroll
file3.txt	Dorothy and the Wizard in Oz	Oz	Lyman Frank Baum
file4.txt	Emerald City of Oz	Oz	Lyman Frank Baum
file5.txt	Royal Book of Oz	Oz	Ruth Plumly Thompson
file6.txt	Glinda of Oz	Oz	Lyman Frank Baum

Determining Authorship

1		file1.txt	file2.txt	file3.txt	file4.txt	file5.txt	file6.txt		Row Sum
2	file1.txt	0	0.001414	0.000738	0.000742	0.001173	0.000881		0.004949
3	file2.txt	0.001414	0	0.001419	0.001364	0.001521	0.001454		0.007172
4	file3.txt	0.000738	0.001419	0	0.000605	0.000933	0.000667		0.004362
5	file4.txt	0.000742	0.001364	0.000605	0	0.001112	0.000779		0.004603
6	file5.txt	0.001173	0.001521	0.000933	0.001112	0	0.001236		0.005975
7	file6.txt	0.000881	0.001454	0.000667	0.000779	0.001236	0		0.005018

Are these values surprising?

Coin Flipping

I flip a coin 25 times.

- I get 12 heads. Is this surprising?
- What if I get 3 heads?



Coin Flipping

I flip a coin 25 times.

- I get 12 heads. Is this surprising?
- What if I get 3 heads?

My Hypothesis: This coin is unfair.



Coin Flipping

I flip a coin 25 times.

- I get 12 heads. Is this surprising?
- What if I get 3 heads?



My Hypothesis: This coin is unfair.



How do I prove this?

Coin Flipping

- What if this penny was a person?



?



Coin Flipping

- What if this penny was a person?



?



Innocent until proven guilty

Coin Flipping

My Hypothesis:

- This coin is unfair.



The “Innocent” Hypothesis:



Coin Flipping

My Hypothesis:

- This coin is unfair.



The “Innocent” Hypothesis:

- The coin is heads with probability 0.5



Coin Flipping

My Hypothesis:

- This coin is unfair.

Alternate hypothesis



The “Innocent” Hypothesis:

- The coin is heads with probability 0.5

Null hypothesis



Hypothesis Testing



Alternate hypothesis: The coin is unfair.

Null hypothesis: The coin is heads with prob. 0.5



- We want to *reject* the null hypothesis by showing that it is **unlikely** to get only 3 heads with a fair coin.

Hypothesis Testing



Alternate hypothesis: The coin is unfair.

Null hypothesis: The coin is heads with prob. 0.5



- We want to *reject* the null hypothesis by showing that it is **unlikely** to get only 3 heads with a fair coin.

Prove that the coin is NOT innocent.



Hypothesis Testing



Alternate hypothesis: The coin is unfair.

Null hypothesis: The coin is heads with prob. 0.5



- We want to *reject* the null hypothesis by showing that it is **unlikely** to get only 3 heads with a fair coin.

Prove that the coin is NOT innocent.

- Open coinflip.xls



Project 2 Notes

- Feedback over the weekend
 - Not going to be as in-depth as the first project

Sun	Mon	Tues	Wed	Thurs	Fri	Sat
4/1	4/2	4/3	4/4	4/5	4/6	4/7
				Proposal Due		
4/8	4/9	4/10	4/11	4/12	4/13	4/14
				Project 2 Due (Code)		
4/15	4/16	4/17	4/18	4/19	4/20	4/21
		Project 2 Due (Website)				