Announcements

- Reading and RQ13 due next time
- Assignment 5 due next Thursday
- APT 5 due Tuesday

- Today:
  - Solving a problem with list comprehensions
  - more with sets

Grace Hopper Celebration of Women in Computing Conference
Apply now for scholarships to attend!
Due March 8
Process Exam Scores
Total number, Average, Median

Start with grades
Create histogram

76
68
74
72
64
77
46
81
...

86 86 86 86
85 85 85 85 85 85
84 84 84 84 84 84 84
83 83 83 83 83 83 83 83 83
82 82 82 82 82 82 82 82 82 82 82
81 81 81 81 81 81 81 81 81
80 80 80 80 80 80 80 80 80 80 80 80
79 79 79 79 79 79 79 79 79 79 79 79
78 78 78 78 78 78 78 78 78 78
77 77 77 77 77 77 77
76 76 76 76 76 76 76 76 76 76

Process Exam Scores
bit.ly/101s17-0302-1

• Calculate
  – Total number of scores
  – Number of unique scores
  – Average score
  – Median score
• Print a visualization of the grades
• Get snarf file

Latanya Sweeney
Former Chief Technologist at FTC. I am a computer scientist with a long history of weaving technology and policy together to remove stakeholder barriers to technology adoption. My focus is on "computational policy" and I term myself a "computer (cross) policy" scientist. I have enjoyed success at creating technology that weaves with policy to resolve real-world technology-privacy clashes.

http://latanyasweeney.org/
Identify 87% of US population using (dob,zip,gender). Director of Harvard Data Privacy Lab, instrumental in HIPAA because of de-identification work

aboutmyinfo.org

• Entered my data

How Unique are You?
Enter your ZIP code, date of birth, and gender to see how unique you are (and therefore how easy it is to identify you from these values).

Date of Birth
Gender
Male
Female
5-digit ZIP
Submit
aboutmyinfo.org

How Unique are You?
Enter your ZIP code, date of birth, and gender to see how unique you are (and therefore how easy it is to identify you from those values).

- Entered my data
- Easily identifiable by birth date (about 1)
- Lots with my birth year (about 273)
- Lots of people in my age range (of four years) – (1,365)

Set Operations from pictures
bit.ly/101s17-0302-2

Question: Which operation does the red represent?

A)  
B)  
C)  
D)  
E)  

APT SandwichBar

Problem Statement
It's time to get something to eat and I've come across a sandwich bar. Like most people, I prefer certain types of sandwiches. In fact, I keep a list of the types of sandwiches I like.

The sandwich bar has certain ingredients available. I will list the types of sandwiches I like in order of preference and buy the first sandwich the bar can make for me. In order for the bar to make a sandwich for me, it must include all of the ingredients I desire.

Given available, a list of Strings/ingredients the sandwich bar can use, and a orders, a list of Strings that represent the types of sandwiches I like, in order of preference (most preferred first), return the 0-based index of the sandwich I will buy. Each element of orders represents one type of sandwich I like as a space-separated list of ingredients in the sandwich. If the bar can make no sandwiches I like, return -1.

APT SandwichBar

2.  
[ "cheese", "mustard", "lettuce" ]

[ "cheese ham", "cheese mustard lettuce", "ketchup", "beer" ]

Returns: 1

They've run out of ham, but I'll consider other options now.

3.  
[ "cheese", "cheese", "cheese", "tomato" ]

[ "ham ham ham", "water", "pork", "bread", "cheese tomato cheese", "beef" ]

Returns: 4

Ignore any duplicate elements in the lists.
Problems — snarf setExample.py

- Given a list of strings that have the name of a course (one word), followed by last names (one word each) of people in the course:
  1. Find total number of people taking any course
  2. Find number of people taking just one course

```
["econ101 Abroms Curtson Williams Smith",
"history230 Black Wrigley Smith", ... ]
```

Process data — create lists of strings of names for each course

Data for example

```
["compsci101 Smith Ye Li Lin Abroms Black",
"math101 Green Wei Lin Williams DeLong Noell Ye Smith",
"econ101 Abroms Curtson Williams Smith",
"french1 Wills Wrigley Olson Lee",
"history230 Black Wrigley Smith"]
```

To easier format to work with:

```
[['Smith', 'Ye', 'Li', 'Lin', 'Abroms', 'Black'],
['Green', 'Wei', 'Lin', 'Williams', 'DeLong', 'Noell', 'Ye', 'Smith'],
['Abroms', 'Curtson', 'Williams', 'Smith'], ...
]
Part 1 — processList
bit.ly/101s17-0302-4
• Given a list of strings that have the name of a course (one word), followed by last names of people in the course:
  – Convert list into lists of strings of names for each course

  ["econ101 Abroms Curtson Williams Smith", "history230 Black Wrigley Smith", ... ]
  [ ['Abroms', 'Curtson', 'Williams', 'Smith'], ['Black', 'Wrigley', 'Smith'], ... ]

Part 2 — peopleTakingCourses
bit.ly/101s17-0302-5
• Given a list of lists of names, each list represents the people in one course:
  – Find total number of people taking any course
  – peopleTakingCourses should return unique list of names

• Small Example

  Answer is 6 unique names
Next, find the number of people taking just one course

To solve this problem

- First let’s write a helper function
Part 3 — unionAllSetsButMe

• Given example, a list of sets of strings, and the index of one of the sets, return the union of all the sets but that one

example = [set(["a", "b", "c"]), set(["b", "c", "d", "g")], set(["e", "d", "a"])]
unionAllSetsButMe(example, 1) is set(["a", "b", "c", "e", "d"])

Part 4 — peopleTakingOnlyOneCourse

• Given a list of lists of strings of names representing people from courses
  – Find number of people taking just one course

[[‘Abroms’, ‘Curtson’, ‘Williams’, ‘Smith’],
4