ECS 10

11/28

Database

- Program that answers questions based on data.

Select a Congressperson: Harman
Jane Harman is the Representative from District 36

- Questions to a database known as **queries.**
- Different data structures for different kinds of queries.

List of Lists

- Last time, we considered using a list of lists.
- Have to read through whole list to find the item you're looking for.
- Possibly save time by stopping as soon as you find an answer.
- Stop a for loop early with the **break** command.

Break

- Immediately jumps out of the loop.
- Usually used with **for** loops. Avoid with **while**.

```
import random
num = random.randint(1,10)
print "I'm thinking of a number between 1 and 10"
print "You have three guesses."
for i in range(3):
    guess = raw_input("Guess: ")
    if int(guess) == num:
        print "Right!"
        break
    else:
        print "Wrong!"
print "My number was",num
```

Efficient Solution

- Better to "look up" names, like in a phone book.
- Use a **dictionary.**
- A dictionary can be accessed by indexing, like a list, but any kind of data can be used as the index.

Dictionary

- Words are **keys**.
- Definitions are **values**.

Dictionary Entry

myDict = {"ferret":"Ferdinand"}

key value

- Curly braces {} indicate a dictionary
- {} is the empty dictionary.

Looking up an Item

- Look up a value, use key as index.

name = myDict["ferret"]

value key

- Use square brackets for indexing, just like a list.

Structure of Database Program

- Part 1: Reads in data, builds a data structure.
  - Usually has to read whole file. Can be very slow, but only done once.
- Part 2: gets and answers queries
  - Tries to go right to the data it needs.
  - Hopefully does NOT have to look at all the data to answer the query.
  - This is very important for huge data, eg. Google, airline reservations, myUCDavis.

Using a Dictionary

- Initialize the empty dictionary:
  dataBase = {}
- Add values, use key as "index"
  dataBase[name] = district
- Look up values using key as "index"
  district = dataBase[name]
- Test if key is used in dictionary
  name in dataBase

How Could We…

- Loop to ask many queries? What part of the program goes into the loop?
- Loop up both first and last names?
- What if there are two Representatives with the same first name?

Program 6

- Second chance to learn about reading files.
- Second chance to use graphics
- One and only chance to use a dictionary.
- Database program:
  - Step 1: Read in file of name popularity, store in dictionary
  - Step 2: Answer queries by getting data out of dictionary, drawing graphs.

Name Data

A 83 140 228 286 426 612 486 577 836 0 0
Aaliyah 0 0 0 0 0 0 0 0 0 380 215
Aaron 193 208 218 274 279 232 132 36 32 31 41
Abagail 0 0 0 0 0 0 0 0 0 0 958

- 195 means 195'th most popular
- NOT percent of population
- 0 means not in top 1000

Name Data

- Babies born in US, with a Social Security Number
- Average popularity over each decade