



DATA 8
Fall 2016

Lecture 5, September 2

Tables

Slides created by Ani Adhikari and John DeNero

Announcements

- Office hours are in Weekly Schedule in the top menu bar of data8.org
 - Thanks for turning in HW 1. Lab 2 is due today by 7 pm.
 - The waiting list is being processed automatically but it doesn't seem to be moving very fast. There is nothing you or I can do about that, except be patient.
-

Different types of data

| Longitude | Latitude | City | Direction | Survivors |
|-----------|----------|-------------|-----------|-----------|
| 32 | 54.8 | Smolensk | Advance | 145000 |
| 33.2 | 54.9 | Dorogobouge | Advance | 140000 |
| 34.4 | 55.5 | Chjat | Advance | 127100 |
| 37.6 | 55.8 | Moscou | Advance | 100000 |
| 34.3 | 55.2 | Wixma | Retreat | 55000 |
| 32 | 54.6 | Smolensk | Retreat | 24000 |
| 30.4 | 54.4 | Orscha | Retreat | 20000 |
| 26.8 | 54.3 | Moiodexno | Retreat | 12000 |

float:
decimal number

string:
text

int:
integer

Arrays

An array is a sequence that can be manipulated easily.

- All elements of an array should have the same type.
 - Arithmetic is applied to each element individually.
 - Elementwise operations can be done on arrays of the same size.
-

Methods

- Functions that operate on a particular type of object
`object.method(maybe some arguments)`
- String methods operate on strings
 - `'APPLE'.lower` has the value `'apple'`
 - `'APPLE'.replace('A', 'SNA')` has the value `'SNAPPLE'`
- Array methods operate on arrays

(Demo)

Table

- A sequence of named columns that each describe one feature of all the entries in the dataset

also

- A sequence of rows that each describe all the features of one entry in the dataset
 - Columns are arrays.
-

A table

Each entry has 5 features. Each column corresponds to one feature.

The “entries” are locations on the map of Europe.

Each row corresponds to one entry. There are 8 entries in all. (Demo)

| Longitude | Latitude | City | Direction | Survivors |
|------------------|-----------------|-------------|------------------|------------------|
| 32 | 54.8 | Smolensk | Advance | 145000 |
| 33.2 | 54.9 | Dorogobouge | Advance | 140000 |
| 34.4 | 55.5 | Chjat | Advance | 127100 |
| 37.6 | 55.8 | Moscou | Advance | 100000 |
| 34.3 | 55.2 | Wixma | Retreat | 55000 |
| 32 | 54.6 | Smolensk | Retreat | 24000 |
| 30.4 | 54.4 | Orscha | Retreat | 20000 |
| 26.8 | 54.3 | Moiodexno | Retreat | 12000 |

Table methods thus far

- Creating and extending tables
 - Referring to columns: labels, indices, and relabeling
 - Accessing data in a column
 - Using array methods to work with data in columns
 - Creating new tables containing some but not all of the columns of the original
-