Intro to Programming for Public Policy Week 4
Sets

Eric Potash

April 19, 2018

Copying a list

To copy an existing list, pass it to the `list()` function:

```
>>> ls = [1,2,3,4]
>>> ls_copy = list(ls)
>>> ls.pop()
4
>>> ls
[1,2,3]
>>> ls_copy
[1,2,3,4]
```

Data structures

So far we have seen two main data structures:

- Lists
  - Useful for storing an ordered sequence of elements.
- Dictionaries
  - Dictionaries are useful for storing a mapping for keys to values. The keys are unique, the values need not be.

Sets

- Sets are another python data structure.
- They store a collection of unique elements
  - They are like the keys of a dictionary
  - Without the values

Constructing a set

- Construct a set using the `set()` function:

```
>>> s = set()
set()
```

- Add elements with the `add` function:

```
>>> s.add('a')
>>> s
{'a'}
```

Adding the same element twice

Since a set only stores unique elements, adding the same element twice has no effect:

```
>>> s = set()
>>> s.add('a')
>>> s.add('b')
>>> s
{'a', 'b'}
>>> s.add('a')
>>> s
{'a', 'b'}
```

Construct set from a list

Or construct it using a list:

```
>>> set(['a', 'b', 'c', 'd'])
{'a', 'b', 'c', 'd'}
```

Again duplicate elements are ignored:

```
>>> set(['a', 'b', 'b', 'c'])
{'a', 'b', 'c'}
```

Unique list

Thus if a list and its corresponding set have the same length, then all elements in the list were unique:

```
>>> ls = ['a', 'b', 'b', 'c']
>>> len(ls)
4
>>> len(set(ls))
3
```

Comparing unique elements

Sets are very useful for checking whether two lists have the same elements:

```
>>> ls1 = ['a','b','c','c','d','e','a']
>>> ls2 = ['b','c','d','a']
>>> set(ls1)
{'a','b','c','d','e'}
>>> set(ls2)
{'a','b','c','d'}
>>> set(ls1) == set(ls2)
False
```

Another example

```
students = ['A', 'B', 'C']
hospital_prefs = {
    'X': ['B', 'A', 'C'],
    'Y': ['A', 'B', 'C'],
    'Z': ['A', 'B', 'C']
}
```

Want to check whether hospital X ranked every student:

```
set(hospital_prefs['X']) == set(students):
```

And that X's rankings contain no duplicates:

```
len(set(hospital_prefs['X'])) == len(hospital_prefs['X']):
```