Elements Of Data Science - F2025

Week 2: Python Intro/Review and Numpy

9/9/2025

TODOs

- **Review** Selections from PDSH Chapter 2
- **Read** Selections from PDSH Chapter 3
- **Skim** Selections from PDSH Chapter 4

- Complete Week 2 Quiz (Due 1 week from today, Sep 16 @11:59pm)

Announcements

Office hours

- Tuesday, 6pm-6:50pm, location: DSI Conference room
- Skype, by appointment only, Mondays 6pm+. (link to be sent once OH requested)

TODOs

- Ch 2. Introduction to NumPy
    - Understanding Data Types in Python
    - The Basics of NumPy Arrays
    - Skim: Computation on NumPy Arrays: Universal Functions
    - Aggregations: Min, Max, and Everything In Between
    - Skim: Computation on Arrays: Broadcasting
    - Comparisons, Masks, and Boolean Logic
    - Fancy Indexing
    - Sorting Arrays
    - Structured Data: NumPy's Structured Arrays

TODOs

- Ch 3. Data Manipulation with Pandas
    - Introducing Pandas Objects
    - Data Indexing and Selection
    - Operating on Data in Pandas
    - Handling Missing Data
    - Hierarchical Indexing
    - Combining Datasets: Concat and Append
    - Combining Datasets: Merge and Join
    - Aggregation and Grouping
    - Pivot Tables
    - Skim:Vectorized String Operations
    - Working with Time Series
    - High-Performance Pandas: eval() and query()

Getting Changes from Git

1. `cd` to the cloned class repostory
2. `git pull`

example:

```
$ cd ~/proj/eods-f25
$ git pull
```

Questions?

TODAY

- Tools Review
- Getting "Help" Documentation
- Python (Review?)
- Numpy

Tools Review

- Starting Jupyter
- Notebooks, Kernels and Virtual Environments

Getting ”Help” Documentation in Python

Getting "Help" Documentation in Python

In [1]:
```
1 help(print)
```

```
Help on built-in function print in module builtins:

print(...)
    print(value, ..., sep=' ', end='\n', file=sys.stdout, flush=False)

    Prints the values to a stream, or to sys.stdout by default.
    Optional keyword arguments:
    file:  a file-like object (stream); defaults to the current sys.stdout.
    sep:   string inserted between values, default a space.
    end:   string appended after the last value, default a newline.
    flush: whether to forcibly flush the stream.
```

Getting "Help" Documentation in Python

In [1]:
```
help(print)
```

```
Help on built-in function print in module builtins:

print(...)
    print(value, ..., sep=' ', end='\n', file=sys.stdout, flush=False)

    Prints the values to a stream, or to sys.stdout by default.
    Optional keyword arguments:
    file:  a file-like object (stream); defaults to the current sys.stdout.
    sep:   string inserted between values, default a space.
    end:   string appended after the last value, default a newline.
    flush: whether to forcibly flush the stream.
```

Also, in ipython/jupyter:

```
print?              # show docstring
print??             # show code as well
print([SHIFT+TAB] # get help in a popup
```

In [2]:
```
import pandas as pd

```

In [3]:
```
pd.DataFrame()
```

Out[3]:
```
_
```

Python (Review?)

- Whitespace Formatting
- Dynamic Typing
- Basic Data Types
- Functions
- String Formatting
- Exceptions and Try-Except
- Truthiness
- Comparisons and Logical Operators
- Control Flow
- Assert
- Sorting
- List/Dict Comprehensions
- Importing Modules
- collections Module
- Object Oriented Programming

Whitespace Formatting

- Instead of braces or brackets to delimit blocks, use whitespace

```
# The pound sign marks the start of a comment. Python itself
# ignores the comments, but they're helpful for anyone reading the code.
for i in [1, 2, 3, 4, 5]:
    print(i)                        # first line in "for i" block

    for j in [1, 2, 3, 4, 5]:

        print(j)                    # first line in "for j" block
        print(i + j)                # last line in "for j" block
    print(i)                        # last line in "for i" block
```

- 4 space indentations are conventional
- Style Guide : PEP 8 (https://www.python.org/dev/peps/pep-0008/)

In [4]:
```
for name in ['Matt', 'Jack', 'John']:
    print(name+ ' is a cool guy')
```

```
Matt is a cool guy
Jack is a cool guy
```

Dynamic Typing

- don't need to specify type at variable creation (though they'll get one at runtime)

In [5]:
```
1 x = 'apple'
```

In [6]:
```
1 type(x)
```

Out[6]:
```
str
```

Dynamic Typing

- don't need to specify type at variable creation (though they'll get one at runtime)

In [5]:
```
1 x = 'apple'
```

In [6]:
```
1 type(x)
```

Out[6]:
```
str
```

In [7]:
```
1 x = 3
2 x = 3.14
3 x = 'apple'
4 x
```

Out[7]:
```
'apple'
```

Dynamic Typing

- don't need to specify type at variable creation (though they'll get one at runtime)

In [5]:
```
x = 'apple'
```

In [6]:
```
type(x)
```

Out[6]: str

In [7]:
```
x = 3
x = 3.14
x = 'apple'
x
```

Out[7]: 'apple'

In [8]:
```
# to determine the current variable type
type(x)
```

Out[8]: str

Basic Python Data Types

- **int** (integer): `42`
- **float** : `4.2`, `4e2`
- **bool** (boolean): `True`, `False`
- **str** (string) : `'num 42'`, `"num 42"`, `"""multi-line string"""`
- **None** (null): `None`

- also `long`, `complex`, `bytes`, etc.

Functions

In [9]:
```
1 pd.DataFrame?
```

Functions

In [9]:
```
pd.DataFrame?
```

In [10]:
```
def add_two(x):
    """Adds 2 to the number passed in."""
    return x+2


add_two(2)
```

Out[10]: 4

Functions

In [9]:
```
pd.DataFrame?
```

In [10]:
```
def add_two(x):
    """Adds 2 to the number passed in."""
    return x+2


add_two(2)
```

Out[10]: 4

In [11]:
```
help(add_two)
```

```
Help on function add_two in module __main__:

add_two(x)
    Adds 2 to the number passed in.
```

Functions

In [9]:
```
pd.DataFrame?
```

In [10]:
```
def add_two(x):
    """Adds 2 to the number passed in."""
    return x+2


add_two(2)
```

Out[10]: 4

In [11]:
```
help(add_two)
```

```
Help on function add_two in module __main__:

add_two(x)
    Adds 2 to the number passed in.
```

Reminder, also in ipython/jupyter:

- `add_two?` # show docstring
- `add_two??` # show code as well
- `add_two([SHIFT+TAB]` # get help in a popup

Function Arguments

Function Arguments

- **positional arguments** must be entered in order

Function Arguments

- **positional arguments** must be entered in order

In [12]:
```
def subtract(x,y):
    return x-y

subtract(3,1)
```

Out[12]: 2

Function Arguments

- **positional arguments** must be entered in order

In [12]:
```
def subtract(x,y):
    return x-y

subtract(3,1)
```

Out[12]: 2

- **keyword arguments** must follow positional
- can be called in any order

Function Arguments

- **positional arguments** must be entered in order

In [12]:
```
def subtract(x,y):
    return x-y

subtract(3,1)
```

Out[12]: 2

- **keyword arguments** must follow positional
- can be called in any order

In [13]:
```
def proportion(numer,denom,precision=2):
    return round(numer/denom,precision)

proportion(2,precision=2,denom=3)
```

Out[13]: 0.67

String Formatting

String Formatting

In [14]:
```
x = 3.1415

'the value of x is ' + str(x)
```

Out[14]: 'the value of x is 3.1415'

In [15]:
```
type(x)
```

Out[15]: float

String Formatting

In [14]:
```
1 x = 3.1415
2
3 'the value of x is ' + str(x)
```

Out[14]: 'the value of x is 3.1415'

In [15]:
```
1 type(x)
```

Out[15]: float

In [16]:
```
1 'the value of x is %0.3f' % x
```

Out[16]: 'the value of x is 3.142'

String Formatting

In [14]:
```
x = 3.1415

'the value of x is ' + str(x)
```

Out[14]: 'the value of x is 3.1415'

In [15]:
```
type(x)
```

Out[15]: float

In [16]:
```
'the value of x is %0.3f' % x
```

Out[16]: 'the value of x is 3.142'

In [17]:
```
'the value of x is {:0.10f} and not {}'.format(x, 10)
```

Out[17]: 'the value of x is 3.1415000000 and not 10'

String Formatting

In [14]:
```
x = 3.1415

'the value of x is ' + str(x)
```

Out[14]: 'the value of x is 3.1415'

In [15]:
```
type(x)
```

Out[15]: float

In [16]:
```
'the value of x is %0.3f' % x
```

Out[16]: 'the value of x is 3.142'

In [17]:
```
'the value of x is {:0.10f} and not {}'.format(x, 10)
```

Out[17]: 'the value of x is 3.1415000000 and not 10'

In [18]:
```
f'the value of x is {x:0.2f}'
# note: f-string is a literal string, prefixed with 'f', which contains expressions inside braces.
```

Out[18]: 'the value of x is 3.14'

String Formatting

In [14]:
```
x = 3.1415

'the value of x is ' + str(x)
```

Out[14]: 'the value of x is 3.1415'

In [15]:
```
type(x)
```

Out[15]: float

In [16]:
```
'the value of x is %0.3f' % x
```

Out[16]: 'the value of x is 3.142'

In [17]:
```
'the value of x is {:0.10f} and not {}'.format(x, 10)
```

Out[17]: 'the value of x is 3.1415000000 and not 10'

In [18]:
```
f'the value of x is {x:0.2f}'
# note: f-string is a literal string, prefixed with 'f', which contains expressions inside braces.
```

Out[18]: 'the value of x is 3.14'

- often want to print variable values for debugging

String Formatting

In [14]:
```
x = 3.1415

'the value of x is ' + str(x)
```

Out[14]: 'the value of x is 3.1415'

In [15]:
```
type(x)
```

Out[15]: float

In [16]:
```
'the value of x is %0.3f' % x
```

Out[16]: 'the value of x is 3.142'

In [17]:
```
'the value of x is {:0.10f} and not {}'.format(x, 10)
```

Out[17]: 'the value of x is 3.1415000000 and not 10'

In [18]:
```
f'the value of x is {x:0.2f}'
# note: f-string is a literal string, prefixed with 'f', which contains expressions inside braces.
```

Out[18]: 'the value of x is 3.14'

- often want to print variable values for debugging

In [19]:
```
f'x = {x:0.2f}'
```

Out[19]: 'x = 3.14'

String Formatting

In [14]:
```
x = 3.1415

'the value of x is ' + str(x)
```

Out[14]: 'the value of x is 3.1415'

In [15]:
```
type(x)
```

Out[15]: float

In [16]:
```
'the value of x is %0.3f' % x
```

Out[16]: 'the value of x is 3.142'

In [17]:
```
'the value of x is {:0.10f} and not {}'.format(x, 10)
```

Out[17]: 'the value of x is 3.1415000000 and not 10'

In [18]:
```
f'the value of x is {x:0.2f}'
# note: f-string is a literal string, prefixed with 'f', which contains expressions inside braces.
```

Out[18]: 'the value of x is 3.14'

- often want to print variable values for debugging

In [19]:
```
f'x = {x:0.2f}'
```

Out[19]: 'x = 3.14'

In [20]:
```
f'{x = :0.2f}' # new in 3.8
```

String Formatting Cont.

In [21]:
```
"""This is a multiline string.
The value of x is {}."""
```

Out[21]:
```
'This is a multiline string.\nThe value of x is {}.'
```

String Formatting Cont.

In [21]:
```
"""This is a multiline string.
The value of x is {}."""
```

Out[21]:
```
'This is a multiline string.\nThe value of x is {}.'
```

In [22]:
```
"""This is a multiline string.
The value of x is {}.""".format(x)
```

Out[22]:
```
'This is a multiline string.\nThe value of x is 3.1415.'
```

String Formatting Cont.

In [21]:
```
"""This is a multiline string.
The value of x is {}."""
```

Out[21]:
```
'This is a multiline string.\nThe value of x is {}.'
```

In [22]:
```
"""This is a multiline string.
The value of x is {}.""".format(x)
```

Out[22]:
```
'This is a multiline string.\nThe value of x is 3.1415.'
```

In [23]:
```
print("""This is a multiline string.
The value of x is {}.""".format(x))
```

```
This is a multiline string.
The value of x is 3.1415.
```

String Formatting Cont.

In [21]:
```
"""This is a multiline string.
The value of x is {}."""
```

Out[21]:
```
'This is a multiline string.\nThe value of x is {}.'
```

In [22]:
```
"""This is a multiline string.
The value of x is {}.""".format(x)
```

Out[22]:
```
'This is a multiline string.\nThe value of x is 3.1415.'
```

In [23]:
```
print("""This is a multiline string.
The value of x is {}.""".format(x))
```

```
This is a multiline string.
The value of x is 3.1415.
```

- to learn more https://realpython.com/python-string-formatting/

Python Data Types Continued: `list`

Python Data Types Continued: `list`

In [24]:
```
# elements of a python list do not all have to be of the same type
x = [42,'e',2.0]
x
```

Out[24]:
```
[42, 'e', 2.0]
```

Python Data Types Continued: `list`

In [24]:
```
# elements of a python list do not all have to be of the same type
x = [42,'e',2.0]
x
```

Out[24]:
```
[42, 'e', 2.0]
```

In [25]:
```
x[0] # indexing
```

Out[25]:
```
42
```

Python Data Types Continued: `list`

In [24]:
```
# elements of a python list do not all have to be of the same type
x = [42,'e',2.0]
x
```

Out[24]: [42, 'e', 2.0]

In [25]:
```
x[0] # indexing
```

Out[25]: 42

In [26]:
```
x[-1] # reverse indexing
```

Out[26]: 2.0

Python Data Types Continued: `list`

In [24]:
```
# elements of a python list do not all have to be of the same type
x = [42,'e',2.0]
x
```

Out[24]: [42, 'e', 2.0]

In [25]:
```
x[0] # indexing
```

Out[25]: 42

In [26]:
```
x[-1] # reverse indexing
```

Out[26]: 2.0

In [27]:
```
x[2] = 4 # assignment
x
```

Out[27]: [42, 'e', 4]

Python Data Types Continued: `list`

In [24]:
```
# elements of a python list do not all have to be of the same type
x = [42,'e',2.0]
x
```

Out[24]:
```
[42, 'e', 2.0]
```

In [25]:
```
x[0] # indexing
```

Out[25]:
```
42
```

In [26]:
```
x[-1] # reverse indexing
```

Out[26]:
```
2.0
```

In [27]:
```
x[2] = 4 # assignment
x
```

Out[27]:
```
[42, 'e', 4]
```

In [28]:
```
x.append('a') # add a value to list
x
```

Out[28]:
```
[42, 'e', 4, 'a']
```

Python Data Types Continued: `list`

In [24]:
```
# elements of a python list do not all have to be of the same type
x = [42,'e',2.0]
x
```

Out[24]: [42, 'e', 2.0]

In [25]:
```
x[0] # indexing
```

Out[25]: 42

In [26]:
```
x[-1] # reverse indexing
```

Out[26]: 2.0

In [27]:
```
x[2] = 4 # assignment
x
```

Out[27]: [42, 'e', 4]

In [28]:
```
x.append('a') # add a value to list
x
```

Out[28]: [42, 'e', 4, 'a']

In [29]:
```
value_at_1 = x.pop(1) # remove/delete at index
x
```

Out[29]: [42, 4, 'a']

Python Data Types Continued: `dict` (Dictionary)

- Stores key:value pairs

Python Data Types Continued: `dict` (Dictionary)

- Stores key:value pairs

In [30]:
```
x = {'b':[2,1],'a':1,'c':4}
# or x = dict(b=2,a=1,c=4)
x
```

Out[30]:
```
{'b': [2, 1], 'a': 1, 'c': 4}
```

Python Data Types Continued: `dict` (Dictionary)

- Stores key:value pairs

In [30]:
```
x = {'b':[2,1],'a':1,'c':4}
# or x = dict(b=2,a=1,c=4)
x
```

Out[30]:
```
{'b': [2, 1], 'a': 1, 'c': 4}
```

In [31]:
```
# index into dictionary using key
x['b']
```

Out[31]:
```
[2, 1]
```

In [32]:
```
grades = {'John':{90: 100},
          'Mike':[100, 90],
          'Julia':[100,100]}
```

In [33]:
```
grades = {'John':{'midterm': 90,
                  'final':100},
          'Mike':[100, 90],
          'Julia':[100,100]}
```

In [34]:
```
grades['John']['midterm']
```

Out[34]:
```
90
```

Python Data Types Continued: `dict` (Dictionary)

- Stores key:value pairs

In [30]:
```
x = {'b':[2,1],'a':1,'c':4}
# or x = dict(b=2,a=1,c=4)
x
```

Out[30]:
```
{'b': [2, 1], 'a': 1, 'c': 4}
```

In [31]:
```
# index into dictionary using key
x['b']
```

Out[31]:
```
[2, 1]
```

In [32]:
```
grades = {'John':{90: 100},
          'Mike':[100, 90],
          'Julia':[100,100]}
```

In [33]:
```
grades = {'John':{'midterm': 90,
                  'final':100},
          'Mike':[100, 90],
          'Julia':[100,100]}
```

In [34]:
```
grades['John']['midterm']
```

Out[34]:
```
90
```

In [35]:
```
# assign a value to a (new or existing) key
x['a'] = 3
x
```

Python Data Types Continued: `dict` (Dictionary)

- Stores key:value pairs

In [30]:
```
x = {'b':[2,1],'a':1,'c':4}
# or x = dict(b=2,a=1,c=4)
x
```

Out[30]:
```
{'b': [2, 1], 'a': 1, 'c': 4}
```

In [31]:
```
# index into dictionary using key
x['b']
```

Out[31]:
```
[2, 1]
```

In [32]:
```
grades = {'John':{90: 100},
          'Mike':[100, 90],
          'Julia':[100,100]}
```

In [33]:
```
grades = {'John':{'midterm': 90,
                  'final':100},
          'Mike':[100, 90],
          'Julia':[100,100]}
```

In [34]:
```
grades['John']['midterm']
```

Out[34]:
```
90
```

In [35]:
```
# assign a value to a (new or existing) key
x['a'] = 3
x
```

Python Data Types Continued: `dict` (Dictionary)

- Stores key:value pairs

In [30]:
```
x = {'b':[2,1],'a':1,'c':4}
# or x = dict(b=2,a=1,c=4)
x
```

Out[30]:
```
{'b': [2, 1], 'a': 1, 'c': 4}
```

In [31]:
```
# index into dictionary using key
x['b']
```

Out[31]:
```
[2, 1]
```

In [32]:
```
grades = {'John':{90: 100},
          'Mike':[100, 90],
          'Julia':[100,100]}
```

In [33]:
```
grades = {'John':{'midterm': 90,
                  'final':100},
          'Mike':[100, 90],
          'Julia':[100,100]}
```

In [34]:
```
grades['John']['midterm']
```

Out[34]:
```
90
```

In [35]:
```
# assign a value to a (new or existing) key
x['a'] = 3
x
```

Python Data Types Continued: `dict` Cont.

Python Data Types Continued: `dict` Cont.

In [39]:
```
# using the same dictionary
x
```

Out[39]:
```
{'b': [2, 1], 'a': 3, 'c': 4}
```

Python Data Types Continued: `dict` Cont.

In [39]:
```
# using the same dictionary
x
```

Out[39]:
```
{'b': [2, 1], 'a': 3, 'c': 4}
```

In [40]:
```
# get a set of keys
x.keys()
```

Out[40]:
```
dict_keys(['b', 'a', 'c'])
```

Python Data Types Continued: `dict` Cont.

In [39]:
```
# using the same dictionary
x
```

Out[39]:
```
{'b': [2, 1], 'a': 3, 'c': 4}
```

In [40]:
```
# get a set of keys
x.keys()
```

Out[40]:
```
dict_keys(['b', 'a', 'c'])
```

In [41]:
```
# get a set of values
x.values()
```

Out[41]:
```
dict_values([[2, 1], 3, 4])
```

Python Data Types Continued: `dict` Cont.

In [39]:
```
# using the same dictionary
x
```

Out[39]:
```
{'b': [2, 1], 'a': 3, 'c': 4}
```

In [40]:
```
# get a set of keys
x.keys()
```

Out[40]:
```
dict_keys(['b', 'a', 'c'])
```

In [41]:
```
# get a set of values
x.values()
```

Out[41]:
```
dict_values([[2, 1], 3, 4])
```

In [42]:
```
# get a set of (key,value) tuples
x.items()
```

Out[42]:
```
dict_items([('b', [2, 1]), ('a', 3), ('c', 4)])
```

Python Data Types Continued: `dict` Cont.

In [39]:
```
# using the same dictionary
x
```

Out[39]:
```
{'b': [2, 1], 'a': 3, 'c': 4}
```

In [40]:
```
# get a set of keys
x.keys()
```

Out[40]:
```
dict_keys(['b', 'a', 'c'])
```

In [41]:
```
# get a set of values
x.values()
```

Out[41]:
```
dict_values([[2, 1], 3, 4])
```

In [42]:
```
# get a set of (key,value) tuples
x.items()
```

Out[42]:
```
dict_items([('b', [2, 1]), ('a', 3), ('c', 4)])
```

In [43]:
```
# get a list of (key,value) pairs
list(x.items())
```

Out[43]:
```
[('b', [2, 1]), ('a', 3), ('c', 4)]
```

Python Data Types Continued: `tuple`

- like a list, but **immutable**

Python Data Types Continued: `tuple`

- like a list, but **immutable**

In [44]:
```
1 x = (2,'e',3,4)
2 x
```

Out[44]:
```
(2, 'e', 3, 4)
```

Python Data Types Continued: `tuple`

- like a list, but **immutable**

In [44]:
```
1 x = (2,'e',3,4)
2 x
```

Out[44]: (2, 'e', 3, 4)

In [45]:
```
1 x[0] # indexing
```

Out[45]: 2

Python Data Types Continued: `tuple`

- like a list, but **immutable**

In [44]:
```
1 x = (2,'e',3,4)
2 x
```

Out[44]:
```
(2, 'e', 3, 4)
```

In [45]:
```
1 x[0] # indexing
```

Out[45]:
```
2
```

In [46]:
```
1 x[0] = 3 # assignment? Nope, error: immutable`
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
Cell In[46], line 1
----> 1 x[0] = 3

TypeError: 'tuple' object does not support item assignment
```

In [47]:
```
1 a = (1,2,[10,12])
```

In [48]:
```
1 a[2]
```

Out[48]:
```
[10, 12]
```

In [49]:
```
1 a[2] =15
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
```

Python Data Types Continued: `set`

In [54]:
```
lst = [1,1,1,1,1,1,5,5,'Jack']
# lst = list(set(lst))
```

In [55]:
```
lst
```

Out[55]:
```
[1, 1, 1, 1, 1, 1, 5, 5, 'Jack']
```

In [56]:
```
list(set(lst))
```

Out[56]:
```
[1, 'Jack', 5]
```

Python Data Types Continued: `set`

In [54]:
```
lst = [1,1,1,1,1,1,5,5,'Jack']
# lst = list(set(lst))
```

In [55]:
```
lst
```

Out[55]:
```
[1, 1, 1, 1, 1, 1, 5, 5, 'Jack']
```

In [56]:
```
list(set(lst))
```

Out[56]:
```
[1, 'Jack', 5]
```

In [57]:
```
x = {2,'e','e'} # or set([2,'e','e'])
x
```

Out[57]:
```
{2, 'e'}
```

Python Data Types Continued: `set`

In [54]:
```
lst = [1,1,1,1,1,1,5,5,'Jack']
# lst = list(set(lst))
```

In [55]:
```
lst
```

Out[55]: [1, 1, 1, 1, 1, 1, 5, 5, 'Jack']

In [56]:
```
list(set(lst))
```

Out[56]: [1, 'Jack', 5]

In [57]:
```
x = {2,'e','e'} # or set([2,'e','e'])
x
```

Out[57]: {2, 'e'}

In [58]:
```
x.add(1) # insert
x
```

Out[58]: {1, 2, 'e'}

Python Data Types Continued: `set`

In [54]:
```
lst = [1,1,1,1,1,1,5,5,'Jack']
# lst = list(set(lst))
```

In [55]:
```
lst
```

Out[55]: [1, 1, 1, 1, 1, 1, 5, 5, 'Jack']

In [56]:
```
list(set(lst))
```

Out[56]: [1, 'Jack', 5]

In [57]:
```
x = {2,'e','e'} # or set([2,'e','e'])
x
```

Out[57]: {2, 'e'}

In [58]:
```
x.add(1) # insert
x
```

Out[58]: {1, 2, 'e'}

In [59]:
```
x.remove('e') # remove/delete
x
```

Out[59]: {1, 2}

Python Data Types Continued: `set`

In [54]:
```
lst = [1,1,1,1,1,1,5,5,'Jack']
# lst = list(set(lst))
```

In [55]:
```
lst
```

Out[55]: [1, 1, 1, 1, 1, 1, 5, 5, 'Jack']

In [56]:
```
list(set(lst))
```

Out[56]: [1, 'Jack', 5]

In [57]:
```
x = {2,'e','e'} # or set([2,'e','e'])
x
```

Out[57]: {2, 'e'}

In [58]:
```
x.add(1) # insert
x
```

Out[58]: {1, 2, 'e'}

In [59]:
```
x.remove('e') # remove/delete
x
```

Out[59]: {1, 2}

In [60]:
```
x.intersection({2,3})
```

Out[60]: {2}

Python Data Types Continued: `set`

In [54]:
```
lst = [1,1,1,1,1,1,5,5,'Jack']
# lst = list(set(lst))
```

In [55]:
```
lst
```

Out[55]: [1, 1, 1, 1, 1, 1, 5, 5, 'Jack']

In [56]:
```
list(set(lst))
```

Out[56]: [1, 'Jack', 5]

In [57]:
```
x = {2,'e','e'} # or set([2,'e','e'])
x
```

Out[57]: {2, 'e'}

In [58]:
```
x.add(1) # insert
x
```

Out[58]: {1, 2, 'e'}

In [59]:
```
x.remove('e') # remove/delete
x
```

Out[59]: {1, 2}

In [60]:
```
x.intersection({2,3})
```

Out[60]: {2}

In [61]:
```
x.difference({2,3})
```

Python Data Types Continued: `set`

In [54]:
```
lst = [1,1,1,1,1,1,5,5,'Jack']
# lst = list(set(lst))
```

In [55]:
```
lst
```

Out[55]: [1, 1, 1, 1, 1, 1, 5, 5, 'Jack']

In [56]:
```
list(set(lst))
```

Out[56]: [1, 'Jack', 5]

In [57]:
```
x = {2,'e','e'} # or set([2,'e','e'])
x
```

Out[57]: {2, 'e'}

In [58]:
```
x.add(1) # insert
x
```

Out[58]: {1, 2, 'e'}

In [59]:
```
x.remove('e') # remove/delete
x
```

Out[59]: {1, 2}

In [60]:
```
x.intersection({2,3})
```

Out[60]: {2}

In [61]:
```
x.difference({2,3})
```

Determining Length with `len`

Determining Length with `len`

In [63]:
```
1 len([1,2,3])
```

Out[63]: 3

Determining Length with `len`

In [63]:
```
1 len([1,2,3])
```

Out[63]: 3

In [64]:
```
1 len({'a':1,'b':2,'c':3})
```

Out[64]: 3

In [65]:
```
1 ['a', 'p','p','l','e']
```

Out[65]: ['a', 'p', 'p', 'l', 'e']

Determining Length with `len`

In [63]:
```
1 len([1,2,3])
```

Out[63]: 3

In [64]:
```
1 len({'a':1,'b':2,'c':3})
```

Out[64]: 3

In [65]:
```
1 ['a', 'p','p','l','e']
```

Out[65]: ['a', 'p', 'p', 'l', 'e']

In [66]:
```
1 len('apple')
```

Out[66]: 5

In [67]:
```
1 'john'[0]
```

Out[67]: 'j'

Exceptions

Exceptions

In [68]:
```
1 'a' + 2
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
Cell In[68], line 1
----> 1 'a' + 2

TypeError: can only concatenate str (not "int") to str
```

Exceptions

```
In [68]: 1 'a' + 2
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
Cell In[68], line 1
----> 1 'a' + 2

TypeError: can only concatenate str (not "int") to str
```

Common exceptions:

- SyntaxError
- IndentationError
- ValueError
- TypeError
- IndexError
- KeyError
- and many more https://docs.python.org/3/library/exceptions.html

Catching Exceptions with `try-except`

Catching Exceptions with `try-except`

In [69]:

```
try:
    'a' + 2
except TypeError as e:
    print(f"We did this on purpose, and here's what's wrong:\n{e}")
```

```
We did this on purpose, and here's what's wrong:
can only concatenate str (not "int") to str
```

Catching Exceptions with `try-except`

In [69]:
```
try:
    'a' + 2
except TypeError as e:
    print(f"We did this on purpose, and here's what's wrong:\n{e}")
```

```
We did this on purpose, and here's what's wrong:
can only concatenate str (not "int") to str
```

In [70]:
```
try:
    set([1,2,3])[0]
except SyntaxError as e:
    print(f"Print this if there's a syntax error")
except Exception as e:
    print(f"Print this for any other error")
```

```
Print this for any other error
```

Truthiness

In [71]:
```
1 False /0
```

```
---------------------------------------------------------------------------
ZeroDivisionError                         Traceback (most recent call last)
Cell In[71], line 1
----> 1 False /0

ZeroDivisionError: division by zero
```

Truthiness

```
In [71]: 1 False /0
```

```
---------------------------------------------------------------------------
ZeroDivisionError                         Traceback (most recent call last)
Cell In[71], line 1
----> 1 False /0

ZeroDivisionError: division by zero
```

- boolean: `True`, `False`

- These all translate to `False`:
  - `None`
  - `[]` (empty list)
  - `{}` (empty dictionary)
  - `''` (empty string)
  - `set()`
  - `0`
  - `0.0`

Comparison Operators

Comparison Operators

- equality: `==`
- inequality: `!=`

Comparison Operators

- equality: `==`
- inequality: `!=`

```
In [72]: 1 3 == 3

Out[72]: True
```

Comparison Operators

- equality: `==`
- inequality: `!=`

In [72]:
```
1 3 == 3
```

Out[72]: True

In [73]:
```
1 3 != 4
```

Out[73]: True

In [74]:
```
1 1!= False
```

Out[74]: True

Comparison Operators

- equality: `==`
- inequality: `!=`

```
In [72]: 1 3 == 3
```

```
Out[72]: True
```

```
In [73]: 1 3 != 4
```

```
Out[73]: True
```

```
In [74]: 1 1!= False
```

```
Out[74]: True
```

- less than: `<`
- greater than: `>`
- '(less than/greater than) or equal to: `<=` , `>=`

Comparison Operators

- equality: `==`
- inequality: `!=`

In [72]:
```
1 3 == 3
```

Out[72]: True

In [73]:
```
1 3 != 4
```

Out[73]: True

In [74]:
```
1 1!= False
```

Out[74]: True

- less than: `<`
- greater than: `>`
- '(less than/greater than) or equal to: `<=` , `>=`

In [75]:
```
1 3 < 4
```

Out[75]: True

Logical Operators

Logical Operators

- logical operators: `and`, `or`, `not`

In [76]:
```
(False or True and not False)
```

Out[76]: True

In [77]:
```
(False or (True and True ) and not False)
```

Out[77]: True

Logical Operators

- logical operators: `and`, `or`, `not`

In [76]:
```
(False or True and not False)
```

Out[76]: True

In [77]:
```
(False or (True and True ) and not False)
```

Out[77]: True

In [78]:
```
( (3 > 5) or ((3 < 4) and (5 > 4)) ) and not (3 == 5)
```

Out[78]: True

Logical Operators

- logical operators: `and`, `or`, `not`

In [76]:
```
(False or True and not False)
```

Out[76]: True

In [77]:
```
(False or (True and True ) and not False)
```

Out[77]: True

In [78]:
```
( (3 > 5) or ((3 < 4) and (5 > 4)) ) and not (3 == 5)
```

Out[78]: True

- `any()`: at least one element is true

Logical Operators

- logical operators: `and`, `or`, `not`

In [76]:
```
(False or True and not False)
```

Out[76]: True

In [77]:
```
(False or (True and True ) and not False)
```

Out[77]: True

In [78]:
```
( (3 > 5) or ((3 < 4) and (5 > 4)) ) and not (3 == 5)
```

Out[78]: True

- `any()`: at least one element is true

In [79]:
```
any([0,0,0])
```

Out[79]: False

Logical Operators

- logical operators: `and`, `or`, `not`

In [76]:
```
(False or True and not False)
```

Out[76]: True

In [77]:
```
(False or (True and True ) and not False)
```

Out[77]: True

In [78]:
```
( (3 > 5) or ((3 < 4) and (5 > 4)) ) and not (3 == 5)
```

Out[78]: True

- `any()`: at least one element is true

In [79]:
```
any([0,0,0])
```

Out[79]: False

- `all()`: all elements are true

Logical Operators

- logical operators: `and`, `or`, `not`

In [76]:
```
(False or True and not False)
```

Out[76]: True

In [77]:
```
(False or (True and True ) and not False)
```

Out[77]: True

In [78]:
```
( (3 > 5) or ((3 < 4) and (5 > 4)) ) and not (3 == 5)
```

Out[78]: True

- `any()`: at least one element is true

In [79]:
```
any([0,0,0])
```

Out[79]: False

- `all()`: all elements are true

In [80]:
```
all([True,1,False])
```

Out[80]: False

Logical Operators

- logical operators: `and`, `or`, `not`

In [76]:
```
(False or True and not False)
```

Out[76]: True

In [77]:
```
(False or (True and True ) and not False)
```

Out[77]: True

In [78]:
```
( (3 > 5) or ((3 < 4) and (5 > 4)) ) and not (3 == 5)
```

Out[78]: True

- `any()`: at least one element is true

In [79]:
```
any([0,0,0])
```

Out[79]: False

- `all()`: all elements are true

In [80]:
```
all([True,1,False])
```

Out[80]: False

Assert

Assert

- use `assert` to test anything we know should be true
- simple unit test
- raises exception when assertion is false, otherwise nothing

Assert

- use `assert` to test anything we know should be true
- simple unit test
- raises exception when assertion is false, otherwise nothing

In [81]:
```
1 assert 2+2 != 4
```

```
---------------------------------------------------------------------------
AssertionError                            Traceback (most recent call last)
Cell In[81], line 1
----> 1 assert 2+2 != 4

AssertionError:
```

Assert

- use `assert` to test anything we know should be true
- simple unit test
- raises exception when assertion is false, otherwise nothing

In [81]:
```
1 assert 2+2 != 4
```

```
---------------------------------------------------------------------------
AssertionError                            Traceback (most recent call last)
Cell In[81], line 1
----> 1 assert 2+2 != 4

AssertionError:
```

In [82]:
```
1 assert 1 == 0, 'if this is wrong what you want to display'
```

```
---------------------------------------------------------------------------
AssertionError                            Traceback (most recent call last)
Cell In[82], line 1
----> 1 assert 1 == 0, 'if this is wrong what you want to display'

AssertionError: if this is wrong what you want to display
```

Assert

- use `assert` to test anything we know should be true
- simple unit test
- raises exception when assertion is false, otherwise nothing

In [81]:
```
1 assert 2+2 != 4
```

```
---------------------------------------------------------------------------
AssertionError                            Traceback (most recent call last)
Cell In[81], line 1
----> 1 assert 2+2 != 4

AssertionError:
```

In [82]:
```
1 assert 1 == 0, 'if this is wrong what you want to display'
```

```
---------------------------------------------------------------------------
AssertionError                            Traceback (most recent call last)
Cell In[82], line 1
----> 1 assert 1 == 0, 'if this is wrong what you want to display'

AssertionError: if this is wrong what you want to display
```

In [83]:
```
1 # can add an error message
2 assert 1 == 0, "1 does not equal 0"
```

```
---------------------------------------------------------------------------
AssertionError                            Traceback (most recent call last)
Cell In[83], line 2
```

Control Flow: `if`, `elif`, `else`

Control Flow: `if`, `elif`, `else`

- `if` then `elif` then `else`

Control Flow: `if`, `elif`, `else`

- `if` then `elif` then `else`

In [84]:

```
x = 3
if x >= 2:
    print('x > 2')
elif 0 < x < 1:
    print('0< x < 1')
elif 1 <= x < 2:
    print('0< x < 1')
else:
    print('x == 0')
```

```
x > 2
```

Control Flow: `if`, `elif`, `else`

- `if` then `elif` then `else`

In [84]:
```
x = 3
if x >= 2:
    print('x > 2')
elif 0 < x < 1:
    print('0< x < 1')
elif 1 <= x < 2:
    print('0< x < 1')
else:
    print('x == 0')
```

```
x > 2
```

- single-line `if` then `else`

Control Flow: `if`, `elif`, `else`

- `if` then `elif` then `else`

In [84]:
```
x = 3
if x >= 2:
    print('x > 2')
elif 0 < x < 1:
    print('0< x < 1')
elif 1 <= x < 2:
    print('0< x < 1')
else:
    print('x == 0')
```

```
x > 2
```

- single-line `if` then `else`

In [85]:
```
print("x < 0") if (x < 0) else print("x = 0")
```

```
x = 0
```

More Control Flow: `for` and `while`

More Control Flow: `for` and `while`

- `for` each element of an iterable: do something

More Control Flow: `for` and `while`

- `for` each element of an iterable: do something

In [86]:
```
a = []
for x in [0,1,2]:
    a.append(2*x**2)
a
```

Out[86]:
```
[0, 2, 8]
```

More Control Flow: `for` and `while`

- `for` each element of an iterable: do something

In [86]:
```
a = []
for x in [0,1,2]:
    a.append(2*x**2)
a
```

Out[86]:
```
[0, 2, 8]
```

- `while` something is true

More Control Flow: `for` and `while`

- `for` each element of an iterable: do something

In [86]:
```
a = []
for x in [0,1,2]:
    a.append(2*x**2)
a
```

Out[86]:
```
[0, 2, 8]
```

- `while` something is true

In [87]:
```
x = 0
while x < 3:
    x += 1
x
```

Out[87]:
```
3
```

More Control Flow: `break` and `continue`

More Control Flow: `break` and `continue`

- `break` : break out of current loop

More Control Flow: `break` and `continue`

- `break` : break out of current loop

In [88]:

```
x = 0
while True:
    x += 1
    if x == 10:
        print(x)
        break
```

```
10
```

More Control Flow: `break` and `continue`

- `break` : break out of current loop

In [88]:
```
x = 0
while True:
    x += 1
    if x == 10:
        print(x)
        break
```

```
10
```

- `continue` : continue immediately to next iteration of loop

In [89]:
```
for name in ['John', 'Ben', 'Julia']:
    if name == 'Ben':
        continue
    print(name)
```

```
John
Julia
```

More Control Flow: `break` and `continue`

- `break` : break out of current loop

In [88]:
```
x = 0
while True:
    x += 1
    if x == 10:
        print(x)
        break
```

```
10
```

- `continue` : continue immediately to next iteration of loop

In [89]:
```
for name in ['John', 'Ben', 'Julia']:
    if name == 'Ben':
        continue
    print(name)
```

```
John
Julia
```

In [90]:
```
for x in range(5):
    if x == 1:
        continue
    print(x)
```

```
0
2
3
```

Generate a Range of Numbers: `range`

Generate a Range of Numbers: `range`

In [91]:
```
# create list of integers from 0 up to but not including 4
a = []
for x in range(4):
    a.append(x)
a
```

Out[91]:
```
[0, 1, 2, 3]
```

In [92]:
```
range(10)
```

Out[92]:
```
range(0, 10)
```

Generate a Range of Numbers: `range`

In [91]:
```
# create list of integers from 0 up to but not including 4
a = []
for x in range(4):
    a.append(x)
a
```

Out[91]: [0, 1, 2, 3]

In [92]:
```
range(10)
```

Out[92]: range(0, 10)

In [93]:
```
list(range(10))
```

Out[93]: [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]

Generate a Range of Numbers: `range`

In [91]:
```
# create list of integers from 0 up to but not including 4
a = []
for x in range(4):
    a.append(x)
a
```

Out[91]: [0, 1, 2, 3]

In [92]:
```
range(10)
```

Out[92]: range(0, 10)

In [93]:
```
list(range(10))
```

Out[93]: [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]

In [94]:
```
list(range(1,5)) # with a start and end+1
```

Out[94]: [1, 2, 3, 4]

Generate a Range of Numbers: `range`

In [91]:
```
# create list of integers from 0 up to but not including 4
a = []
for x in range(4):
    a.append(x)
a
```

Out[91]: [0, 1, 2, 3]

In [92]:
```
range(10)
```

Out[92]: range(0, 10)

In [93]:
```
list(range(10))
```

Out[93]: [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]

In [94]:
```
list(range(1,5)) # with a start and end+1
```

Out[94]: [1, 2, 3, 4]

In [95]:
```
list(range(0,100,5)) # with start, end+1 and step-size
```

Out[95]: [0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95]

Keep track of list index or for-loop iteration: `enumerate`

Keep track of list index or for-loop iteration: `enumerate`

In [96]:

```
for i,x in enumerate(['a','b','c']):
    print(i,x)
```

```
0 a
1 b
2 c
```

Keep track of list index or for-loop iteration: `enumerate`

In [96]:
```
for i,x in enumerate(['a','b','c']):
    print(i,x)
```

```
0 a
1 b
2 c
```

In [97]:
```
list(enumerate(['a','b','c']))
```

Out[97]:
```
[(0, 'a'), (1, 'b'), (2, 'c')]
```

Sorting

Sorting

Two ways to sort a list:

Sorting

Two ways to sort a list:

1. by changing the list itself: `list.sort()`

Sorting

Two ways to sort a list:

1. by changing the list itself: `list.sort()`

In [98]:
```
x = [4,1,2,3]
x.sort()
assert x == [1,2,3,4], 'Not same lists'
```

Sorting

Two ways to sort a list:

1. by changing the list itself: `list.sort()`

In [98]:
```
x = [4,1,2,3]
x.sort()
assert x == [1,2,3,4], 'Not same lists'
```

2. without changing the list: `sorted()`

Sorting

Two ways to sort a list:

1. by changing the list itself: `list.sort()`

In [98]:
```
x = [4,1,2,3]
x.sort()
assert x == [1,2,3,4], 'Not same lists'
```

2. without changing the list: `sorted()`

In [99]:
```
x = [4,1,2,3]
y = sorted(x)
assert x == [4,1,2,3]
assert y == [1,2,3,4]
```

Sorting Cont.

Sorting Cont.

- To sort descending, use `reverse=True`:

Sorting Cont.

- To sort descending, use `reverse=True`:

In [100]:
```
1 assert sorted([1,2,3,4], reverse=True) == [4,3,2,1]
```

Sorting Cont.

- To sort descending, use `reverse=True`:

In [100]:
```
1 assert sorted([1,2,3,4], reverse=True) == [4,3,2,1]
```

- Pass a `lambda` function to 'key=' to specify what to sort by:

Sorting Cont.

- To sort descending, use `reverse=True`:

In [100]:
```
assert sorted([1,2,3,4], reverse=True) == [4,3,2,1]
```

- Pass a `lambda` function to 'key=' to specify what to sort by:

In [101]:
```
# for example, to sort a dictionary by value
d = {'a':3,'b':5,'c':1}

# recall that .items() returns a set of key,value tuples
s = sorted(d.items(), key=lambda x: x[1])

assert s == [('c', 1), ('a', 3), ('b', 5)]
```

In [102]:
```
sorted(d.items(), key=lambda x: x[0])
```

Out[102]:
```
[('a', 3), ('b', 5), ('c', 1)]
```

List Comprehensions

List Comprehensions

- Like a single line for loop over a list or other iterable

List Comprehensions

- Like a single line for loop over a list or other iterable

In [103]:
```
# which integers between 0 and 3 inclusive are divisible by 2?
is_even = []
for x in range(0,4):
    is_even.append(x%2 == 0)
is_even
```

Out[103]:
```
[True, False, True, False]
```

List Comprehensions

- Like a single line for loop over a list or other iterable

In [103]:
```
# which integers between 0 and 3 inclusive are divisible by 2?
is_even = []
for x in range(0,4):
    is_even.append(x%2 == 0)
is_even
```

Out[103]: [True, False, True, False]

In [104]:
```
[x%2 == 0 for x in range(0,4)] # using a list comprehension
```

Out[104]: [True, False, True, False]

List Comprehensions

- Like a single line for loop over a list or other iterable

In [103]:
```
# which integers between 0 and 3 inclusive are divisible by 2?
is_even = []
for x in range(0,4):
    is_even.append(x%2 == 0)
is_even
```

Out[103]: [True, False, True, False]

In [104]:
```
[x%2 == 0 for x in range(0,4)] # using a list comprehension
```

Out[104]: [True, False, True, False]

In [105]:
```
# what are the indices of the vowels in 'apple'?
vowels = ['a','e','i','o','u']
[i for i,x in enumerate('apple') if x in vowels]
```

Out[105]: [0, 4]

Dictionary Comprehension

Dictionary Comprehension

- list comprehension but for (key,value) pairs
- can add logic to dictionary creation

Dictionary Comprehension

- list comprehension but for (key,value) pairs
- can add logic to dictionary creation

In [106]:
```
pairs = [(1,'e'),(2,'f'),(3,'g')]
```

Dictionary Comprehension

- list comprehension but for (key,value) pairs
- can add logic to dictionary creation

In [106]:
```
1 pairs = [(1,'e'),(2,'f'),(3,'g')]
```

In [107]:
```
1 dict(pairs)
```

Out[107]:
```
{1: 'e', 2: 'f', 3: 'g'}
```

Dictionary Comprehension

- list comprehension but for (key,value) pairs
- can add logic to dictionary creation

In [106]:
```
pairs = [(1,'e'),(2,'f'),(3,'g')]
```

In [107]:
```
dict(pairs)
```

Out[107]:
```
{1: 'e', 2: 'f', 3: 'g'}
```

In [108]:
```
# modify value and only include odd keys
{key:'value_'+ val for key,val in pairs if key%2 == 1}
```

Out[108]:
```
{1: 'value_e', 3: 'value_g'}
```

In [110]:
```
next(iter_obj)
```

Out[110]:
```
1
```

Object Oriented

Object Oriented

In [112]:

```python
class MyClass:
    """A descriptive docstring."""

    # constructor
    def __init__(self,myvalue = 0):  # what happens when created
        # attributes
        self.myvalue = myvalue

    def __repr__(self):   # what gets printed out (string repr.)
        return f'MyClass(myvalue={self.myvalue})'

    # any other methods

    def get_value(self):
        """Return the value in myvalue."""
        return self.myvalue
```

Object Oriented

In [112]:

```python
class MyClass:
    """A descriptive docstring."""

    # constructor
    def __init__(self,myvalue = 0):  # what happens when created
        # attributes
        self.myvalue = myvalue

    def __repr__(self):   # what gets printed out (string repr.)
        return f'MyClass(myvalue={self.myvalue})'

    # any other methods

    def get_value(self):
        """Return the value in myvalue."""
        return self.myvalue
```

In [113]:

```python
x = MyClass(100)            # instantiate object

assert x.myvalue == 100     # access object attribute

assert x.get_value() == 100 # use object method
```

Importing Modules

- Want to import a module/library? Use `import`

Importing Modules

- Want to import a module/library? Use `import`

In [114]:
```
1 import math
2
3 math.sqrt(2)
```

Out[114]:
```
1.4142135623730951
```

Importing Modules

- Want to import a module/library? Use `import`

In [114]:
```
1 import math
2
3 math.sqrt(2)
```

Out[114]: 1.4142135623730951

- Want to import a submodule or function from a module? Use `from`

Importing Modules

- Want to import a module/library? Use `import`

In [114]:
```
1 import math
2
3 math.sqrt(2)
```

Out[114]:
```
1.4142135623730951
```

- Want to import a submodule or function from a module? Use `from`

In [115]:
```
1 from math import sqrt,floor
2
3 print(sqrt(2))
4 print(floor(sqrt(2)))
```

```
1.4142135623730951
1
```

Importing Modules Cont.

- Want to import a module using an alias? Use 'as'

Importing Modules Cont.

- Want to import a module using an alias? Use 'as'

In [116]:
```
1 import math as m
2 m.sqrt(2)
```

Out[116]:
```
1.4142135623730951
```

Importing Modules Cont.

- Want to import a module using an alias? Use 'as'

In [116]:
```
import math as m
m.sqrt(2)
```

Out[116]:
```
1.4142135623730951
```

- Don't do: `import *`

```
from math import *
# for example, what if there is a math.print() function?
# what happens when we then call print()?
```

`collections` Module

`collections` Module

In [117]:
```
1 from collections import Counter, defaultdict
```

`collections` Module

In [117]:
```
1 from collections import Counter, defaultdict
```

- `Counter` : useful for counting hashable objects
- `defaultdict` : create dictionaries without checking keys
- `OrderedDict` : key,value pairs returned in order added

- others : https://docs.python.org/3.7/library/collections.html

`collections` Module: `Counter`

`collections` Module: `Counter`

In [118]:
```
1 c = Counter(['red', 'blue', 'red', 'green', 'blue', 'blue'])
2 c
```

Out[118]:
```
Counter({'red': 2, 'blue': 3, 'green': 1})
```

`collections` Module: `Counter`

In [118]:
```
c = Counter(['red', 'blue', 'red', 'green', 'blue', 'blue'])
c
```

Out[118]:
```
Counter({'red': 2, 'blue': 3, 'green': 1})
```

In [119]:
```
c = Counter()
for word in ['red', 'blue', 'red', 'green', 'blue', 'blue']:
    c[word] += 1
```

`collections` Module: `Counter`

In [118]:
```
c = Counter(['red', 'blue', 'red', 'green', 'blue', 'blue'])
c
```

Out[118]:
```
Counter({'red': 2, 'blue': 3, 'green': 1})
```

In [119]:
```
c = Counter()
for word in ['red', 'blue', 'red', 'green', 'blue', 'blue']:
    c[word] += 1
```

In [120]:
```
c.most_common()
```

Out[120]:
```
[('blue', 3), ('red', 2), ('green', 1)]
```

`collections` Module Cont. : `defaultdict`

`collections` Module Cont. : `defaultdict`

In [121]:

```
%xmode Minimal
# reduce the amount printed when an exception is thrown
```

```
Exception reporting mode: Minimal
```

`collections` Module Cont.: `defaultdict`

In [121]:
```
%xmode Minimal
# reduce the amount printed when an exception is thrown
```

```
Exception reporting mode: Minimal
```

In [122]:
```
# create mapping from length of word to list of words
colors = ['red', 'blue', 'purple', 'gold', 'orange']
d = {}
for word in colors:
    d[len(word)].append(word)
```

```
KeyError: 3
```

`collections` Module Cont.: `defaultdict`

In [121]:
```
%xmode Minimal
# reduce the amount printed when an exception is thrown
```

```
Exception reporting mode: Minimal
```

In [122]:
```
# create mapping from length of word to list of words
colors = ['red', 'blue', 'purple', 'gold', 'orange']
d = {}
for word in colors:
    d[len(word)].append(word)
```

```
KeyError: 3
```

In [123]:
```
d = {}
for word in colors:
    if len(word) in d:
        d[len(word)].append(word)
    else:
        d[len(word)] = [word]
d
```

Out[123]:
```
{3: ['red'], 4: ['blue', 'gold'], 6: ['purple', 'orange']}
```

`collections` Module Cont.: `defaultdict`

In [121]:
```
%xmode Minimal
# reduce the amount printed when an exception is thrown
```

```
Exception reporting mode: Minimal
```

In [122]:
```
# create mapping from length of word to list of words
colors = ['red', 'blue', 'purple', 'gold', 'orange']
d = {}
for word in colors:
    d[len(word)].append(word)
```

```
KeyError: 3
```

In [123]:
```
d = {}
for word in colors:
    if len(word) in d:
        d[len(word)].append(word)
    else:
        d[len(word)] = [word]
d
```

Out[123]:
```
{3: ['red'], 4: ['blue', 'gold'], 6: ['purple', 'orange']}
```

In [124]:
```
d = defaultdict(list)
for word in colors:
     d[len(word)].append(word)
d
```

Out[124]:
```
defaultdict(list, {3: ['red'], 4: ['blue', 'gold'], 6: ['purple', 'orange']})
```

`collections` Module Cont.: `defaultdict`

In [121]:
```
%xmode Minimal
# reduce the amount printed when an exception is thrown
```

```
Exception reporting mode: Minimal
```

In [122]:
```
# create mapping from length of word to list of words
colors = ['red', 'blue', 'purple', 'gold', 'orange']
d = {}
for word in colors:
    d[len(word)].append(word)
```

```
KeyError: 3
```

In [123]:
```
d = {}
for word in colors:
    if len(word) in d:
        d[len(word)].append(word)
    else:
        d[len(word)] = [word]
d
```

Out[123]:
```
{3: ['red'], 4: ['blue', 'gold'], 6: ['purple', 'orange']}
```

In [124]:
```
d = defaultdict(list)
for word in colors:
     d[len(word)].append(word)
d
```

Out[124]:
```
defaultdict(list, {3: ['red'], 4: ['blue', 'gold'], 6: ['purple', 'orange']})
```

Contexts

Contexts

- a context is like applying a scope with helper functions
- For example: open and write to a file

Contexts

- a context is like applying a scope with helper functions
- For example: open and write to a file

In [126]:
```
with open('tmp_context_example.txt','w') as f:
    f.write('test')
```

Contexts

- a context is like applying a scope with helper functions
- For example: open and write to a file

In [126]:
```
with open('tmp_context_example.txt','w') as f:
    f.write('test')
```

In [127]:
```
# instead of
f = open('tmp_context_example.txt','w')
f.write('test')
f.close() # this is easy to forget to do
```

Contexts

- a context is like applying a scope with helper functions
- For example: open and write to a file

In [126]:
```
with open('tmp_context_example.txt','w') as f:
    f.write('test')
```

In [127]:
```
# instead of
f = open('tmp_context_example.txt','w')
f.write('test')
f.close() # this is easy to forget to do
```

In [128]:
```
# remove the example file we just created
%rm tmp_context_example.txt
```

Python (Review?)

- Dynamic Typing
- Whitespace Formatting
- Basic Data Types
- Functions
- String Formatting
- Exceptions and Try-Except
- Truthiness
- Comparisons and Logical Operators
- Control Flow
- Assert
- Sorting
- List/Dict Comprehensions
- Importing Modules
- collections Module
- Object Oriented Programming

Questions?

Working with Data

Working with Data

Want to:

- transform and select data quickly (numpy)
- manipulate datasets: load, save, group, join, etc. (pandas)
- keep things organized (pandas)

Intro to NumPy

Intro to NumPy

Provides (from numpy.org):

- a powerful N-dimensional array object
- sophisticated (broadcasting) functions
- linear algebra and random number capabilities
- (Fourier transform, tools for integrating C/C++ and Fortran code, etc.)

Python Dynamic Typing

Python Dynamic Typing

In [129]:

```
x = 5
x = 'five'
x = 5.0
```

Python Dynamic Typing

In [129]:
```
1 x = 5
2 x = 'five'
3 x = 5.0
```

- Note: still *strongly* typed

- Python is both a strongly typed and a dynamically typed language.
    - Strong typing means that variables do have a type and that the type matters when performing operations on a variable.
    - Dynamic typing means that the type of the variable is determined only during runtime.

Python Dynamic Typing

In [129]:
```
x = 5
x = 'five'
x = 5.0
```

- Note: still *strongly* typed

- Python is both a strongly typed and a dynamically typed language.
  - Strong typing means that variables do have a type and that the type matters when performing operations on a variable.
  - Dynamic typing means that the type of the variable is determined only during runtime.

In [130]:
```
x,y = 5,'five'
x+y
```

```
TypeError: unsupported operand type(s) for +: 'int' and 'str'
```

Importing NumPy

Importing NumPy

Often imported as alias `np`

Importing NumPy

Often imported as alias `np`

In [131]:
```
1 import numpy as np
2 #Create an numpy array with random integers
3 np.random.randint(10,size=5)
```

Out[131]:
```
array([0, 1, 0, 5, 6])
```

NumPy Datatypes

NumPy Datatypes

```
bool_       Boolean (True or False) stored as a byte
int_        Default integer type (same as C long; normally either int64 or int32)
intc        Identical to C int (normally int32 or int64)
intp        Integer used for indexing (same as C ssize_t; normally either int32 or int64)
int8        Byte (-128 to 127)
int16       Integer (-32768 to 32767)
int32       Integer (-2147483648 to 2147483647)
int64       Integer (-9223372036854775808 to 9223372036854775807)
uint8       Unsigned integer (0 to 255)
uint16      Unsigned integer (0 to 65535)
uint32      Unsigned integer (0 to 4294967295)
uint64      Unsigned integer (0 to 18446744073709551615)
float_      Shorthand for float64.
float16     Half precision float: sign bit, 5 bits exponent, 10 bits mantissa
float32     Single precision float: sign bit, 8 bits exponent, 23 bits mantissa
float64     Double precision float: sign bit, 11 bits exponent, 52 bits mantissa
complex_    Shorthand for complex128.
complex64   Complex number, represented by two 32-bit floats
complex128  Complex number, represented by two 64-bit floats
```

NumPy Arrays

NumPy Arrays

In [132]:
```
1 x = np.array([1,2,3])
2 x
```

Out[132]:
```
array([1, 2, 3])
```

NumPy Arrays

In [132]:
```
1 x = np.array([1,2,3])
2 x
```

Out[132]:
```
array([1, 2, 3])
```

In [133]:
```
1 type(x[0])
```

Out[133]:
```
numpy.int64
```

NumPy Arrays

In [132]:
```
x = np.array([1,2,3])
x
```

Out[132]:
```
array([1, 2, 3])
```

In [133]:
```
type(x[0])
```

Out[133]:
```
numpy.int64
```

In [134]:
```
# use dtype to show the datatype of the array
x.dtype
```

Out[134]:
```
dtype('int64')
```

NumPy Arrays

In [132]:
```
x = np.array([1,2,3])
x
```

Out[132]:
```
array([1, 2, 3])
```

In [133]:
```
type(x[0])
```

Out[133]:
```
numpy.int64
```

In [134]:
```
# use dtype to show the datatype of the array
x.dtype
```

Out[134]:
```
dtype('int64')
```

In [135]:
```
# np arrays can only contain one datatype and default to the most flexible type
x = np.array([1,'two',3])
x
```

Out[135]:
```
array(['1', 'two', '3'], dtype='<U21')
```

NumPy Arrays

In [132]:
```
x = np.array([1,2,3])
x
```

Out[132]:
```
array([1, 2, 3])
```

In [133]:
```
type(x[0])
```

Out[133]:
```
numpy.int64
```

In [134]:
```
# use dtype to show the datatype of the array
x.dtype
```

Out[134]:
```
dtype('int64')
```

In [135]:
```
# np arrays can only contain one datatype and default to the most flexible type
x = np.array([1,'two',3])
x
```

Out[135]:
```
array(['1', 'two', '3'], dtype='<U21')
```

In [136]:
```
x.dtype
```

Out[136]:
```
dtype('<U21')
```

NumPy Arrays

In [132]:
```
x = np.array([1,2,3])
x
```

Out[132]:
```
array([1, 2, 3])
```

In [133]:
```
type(x[0])
```

Out[133]:
```
numpy.int64
```

In [134]:
```
# use dtype to show the datatype of the array
x.dtype
```

Out[134]:
```
dtype('int64')
```

In [135]:
```
# np arrays can only contain one datatype and default to the most flexible type
x = np.array([1,'two',3])
x
```

Out[135]:
```
array(['1', 'two', '3'], dtype='<U21')
```

In [136]:
```
x.dtype
```

Out[136]:
```
dtype('<U21')
```

In [137]:
```
# many different ways to create numpy arrays
np.zeros(5,dtype=int)
```

Out[137]:
```
array([0, 0, 0, 0, 0])
```

In [138]:
```
np.ones(5, dtype=float)
```

Out[138]:
```
array([1., 1., 1., 1., 1.])
```

NumPy Array Indexing

NumPy Array Indexing

- For single indices, works the same as list

NumPy Array Indexing

- For single indices, works the same as list

In [139]:
```
1 x = np.arange(6)
2 x
```

Out[139]:
```
array([0, 1, 2, 3, 4, 5])
```

NumPy Array Indexing

- For single indices, works the same as list

In [139]:
```
1 x = np.arange(6)
2 x
```

Out[139]:
```
array([0, 1, 2, 3, 4, 5])
```

In [140]:
```
1 x[0], x[-1], x[-2], x[-4]
```

Out[140]:
```
(0, 5, 4, 2)
```

NumPy Array Slicing

NumPy Array Slicing

In [141]:
```
x = np.arange(5) # note that in numpy it's arange instead of range
x
```

Out[141]:
```
array([0, 1, 2, 3, 4])
```

NumPy Array Slicing

In [141]:
```
x = np.arange(5) # note that in numpy it's arange instead of range
x
```

Out[141]:
```
array([0, 1, 2, 3, 4])
```

In [142]:
```
# return first two items, start:end (exclusive)
x[0:3]
```

Out[142]:
```
array([0, 1, 2])
```

NumPy Array Slicing

In [141]:
```
x = np.arange(5) # note that in numpy it's arange instead of range
x
```

Out[141]: array([0, 1, 2, 3, 4])

In [142]:
```
# return first two items, start:end (exclusive)
x[0:3]
```

Out[142]: array([0, 1, 2])

In [143]:
```
# missing start implies position 0
x[:4]
```

Out[143]: array([0, 1, 2, 3])

NumPy Array Slicing

In [141]:
```
x = np.arange(5) # note that in numpy it's arange instead of range
x
```

Out[141]: array([0, 1, 2, 3, 4])

In [142]:
```
# return first two items, start:end (exclusive)
x[0:3]
```

Out[142]: array([0, 1, 2])

In [143]:
```
# missing start implies position 0
x[:4]
```

Out[143]: array([0, 1, 2, 3])

In [144]:
```
# missing end implies length of array
x[2:]
```

Out[144]: array([2, 3, 4])

NumPy Array Slicing

In [141]:
```
x = np.arange(5) # note that in numpy it's arange instead of range
x
```

Out[141]: array([0, 1, 2, 3, 4])

In [142]:
```
# return first two items, start:end (exclusive)
x[0:3]
```

Out[142]: array([0, 1, 2])

In [143]:
```
# missing start implies position 0
x[:4]
```

Out[143]: array([0, 1, 2, 3])

In [144]:
```
# missing end implies length of array
x[2:]
```

Out[144]: array([2, 3, 4])

In [145]:
```
# return last two items
x[-2:]
```

Out[145]: array([3, 4])

NumPy Array Slicing with Steps

NumPy Array Slicing with Steps

In [146]:
```
1 x
```

Out[146]:
```
array([0, 1, 2, 3, 4])
```

NumPy Array Slicing with Steps

In [146]:
```
1 x
```

Out[146]:
```
array([0, 1, 2, 3, 4])
```

In [147]:
```
1 # return every other item from position 1 to 4 exclusive
2 # start:end:step_size
3 x[1:4:2]
```

Out[147]:
```
array([1, 3])
```

Reverse array with step-size of -1

Reverse array with step-size of -1

In [148]:
```
1 x
```

Out[148]:
```
array([0, 1, 2, 3, 4])
```

Reverse array with step-size of -1

In [148]:
```
1 x
```

Out[148]: array([0, 1, 2, 3, 4])

In [149]:
```
1 x[::-1]
```

Out[149]: array([4, 3, 2, 1, 0])

NumPy Fancy Indexing

NumPy Fancy Indexing

- Accessing multiple, non-consecutive indices at once using a list

NumPy Fancy Indexing

- Accessing multiple, non-consecutive indices at once using a list

In [150]:
```
1 x = np.arange(5,10)
2 x
```

Out[150]:
```
array([5, 6, 7, 8, 9])
```

NumPy Fancy Indexing

- Accessing multiple, non-consecutive indices at once using a list

In [150]:
```
1 x = np.arange(5,10)
2 x
```

Out[150]:
```
array([5, 6, 7, 8, 9])
```

In [151]:
```
1 x[[0,4]]
```

Out[151]:
```
array([5, 9])
```

NumPy Fancy Indexing

- Accessing multiple, non-consecutive indices at once using a list

In [150]:
```
x = np.arange(5,10)
x
```

Out[150]:
```
array([5, 6, 7, 8, 9])
```

In [151]:
```
x[[0,4]]
```

Out[151]:
```
array([5, 9])
```

In [152]:
```
x[[0,2,-1]]

```

Out[152]:
```
array([5, 7, 9])
```

Boolean Indexing using a Boolean Mask

Boolean Indexing using a Boolean Mask

In [153]:
```
1 x
```

Out[153]:
```
array([5, 6, 7, 8, 9])
```

Boolean Indexing using a Boolean Mask

In [153]:
```
x
```

Out[153]:
```
array([5, 6, 7, 8, 9])
```

In [154]:
```
# Which indices have a value divisible by 2?
# mod operator % returns remainder of division
x%2 == 0
```

Out[154]:
```
array([False,  True, False,  True, False])
```

In [155]:
```
x[[False,  True, False,  True, False]]
```

Out[155]:
```
array([6, 8])
```

Boolean Indexing using a Boolean Mask

In [153]:
```
x
```

Out[153]:
```
array([5, 6, 7, 8, 9])
```

In [154]:
```
# Which indices have a value divisible by 2?
# mod operator % returns remainder of division
x%2 == 0
```

Out[154]:
```
array([False,  True, False,  True, False])
```

In [155]:
```
x[[False,  True, False,  True, False]]
```

Out[155]:
```
array([6, 8])
```

In [156]:
```
# Which values are divisible by 2?
x[x%2 == 0]
```

Out[156]:
```
array([6, 8])
```

Boolean Indexing using a Boolean Mask

In [153]:
```
x
```

Out[153]:
```
array([5, 6, 7, 8, 9])
```

In [154]:
```
# Which indices have a value divisible by 2?
# mod operator % returns remainder of division
x%2 == 0
```

Out[154]:
```
array([False,  True, False,  True, False])
```

In [155]:
```
x[[False,  True, False,  True, False]]
```

Out[155]:
```
array([6, 8])
```

In [156]:
```
# Which values are divisible by 2?
x[x%2 == 0]
```

Out[156]:
```
array([6, 8])
```

In [157]:
```
# Which values are greater than 6?
x[x <= 8]
```

Out[157]:
```
array([5, 6, 7, 8])
```

Boolean Indexing And Bitwise Operators

Boolean Indexing And Bitwise Operators

In [158]:
```
1 x
```

Out[158]:
```
array([5, 6, 7, 8, 9])
```

Boolean Indexing And Bitwise Operators

In [158]:
```
1 x
```

Out[158]:
```
array([5, 6, 7, 8, 9])
```

In [159]:
```
1 (x%2 == 0)
```

Out[159]:
```
array([False,  True, False,  True, False])
```

Boolean Indexing And Bitwise Operators

In [158]:
```
1 x
```

Out[158]:
```
array([5, 6, 7, 8, 9])
```

In [159]:
```
1 (x%2 == 0)
```

Out[159]:
```
array([False,  True, False,  True, False])
```

In [160]:
```
1 (x > 6)
```

Out[160]:
```
array([False, False,  True,  True,  True])
```

Boolean Indexing And Bitwise Operators

In [158]:
```
x
```

Out[158]:
```
array([5, 6, 7, 8, 9])
```

In [159]:
```
(x%2 == 0)
```

Out[159]:
```
array([False,  True, False,  True, False])
```

In [160]:
```
(x > 6)
```

Out[160]:
```
array([False, False,  True,  True,  True])
```

```
# Which values are divisible by 2 AND greater than 6?
# 'and' expexts both elements to be boolean, not arrays of booleans!
(x%2 == 0) and (x > 6)
```

Boolean Indexing And Bitwise Operators

In [158]:
```
x
```

Out[158]:
```
array([5, 6, 7, 8, 9])
```

In [159]:
```
(x%2 == 0)
```

Out[159]:
```
array([False,  True, False,  True, False])
```

In [160]:
```
(x > 6)
```

Out[160]:
```
array([False, False,  True,  True,  True])
```

```
# Which values are divisible by 2 AND greater than 6?
# 'and' expexts both elements to be boolean, not arrays of booleans!
(x%2 == 0) and (x > 6)
```

In [161]:
```
# & compares each element pairwise
x[(x%3 == 0) & (x > 5) & (x%2==0) ]
```

Out[161]:
```
array([6])
```

In [162]:
```
x[np.array([False, False, False,  True, False])]
```

Out[162]:
```
array([8])
```

Boolean Indexing And Bitwise Operators

In [158]:
```
x
```

Out[158]: array([5, 6, 7, 8, 9])

In [159]:
```
(x%2 == 0)
```

Out[159]: array([False, True, False, True, False])

In [160]:
```
(x > 6)
```

Out[160]: array([False, False, True, True, True])

```
# Which values are divisible by 2 AND greater than 6?
# 'and' expexts both elements to be boolean, not arrays of booleans!
(x%2 == 0) and (x > 6)
```

In [161]:
```
# & compares each element pairwise
x[(x%3 == 0) & (x > 5) & (x%2==0) ]
```

Out[161]: array([6])

In [162]:
```
x[np.array([False, False, False,  True, False])]
```

Out[162]: array([8])

In [163]:
```
x[(x%2 == 0) & (x > 6)]
```

Out[163]: array([8])

Boolean Indexing And Bitwise Operators

Boolean Indexing And Bitwise Operators

- and : `&` (ampersand)

Boolean Indexing And Bitwise Operators

- and : `&` (ampersand)

In [164]:
```
# Which values are even AND greater than 6?
x[(x%2 == 0) & (x > 6)]
```

Out[164]:
```
array([8])
```

Boolean Indexing And Bitwise Operators

- and : `&` (ampersand)

In [164]:
```
# Which values are even AND greater than 6?
x[(x%2 == 0) & (x > 6)]
```

Out[164]:
```
array([8])
```

- or : `|` (pipe)

Boolean Indexing And Bitwise Operators

- and : `&` (ampersand)

In [164]:
```
# Which values are even AND greater than 6?
x[(x%2 == 0) & (x > 6)]
```

Out[164]:
```
array([8])
```

- or : `|` (pipe)

In [165]:
```
# which values are even OR greater than 6?
x[(x%3 == 0) | (x > 8)]
```

Out[165]:
```
array([6, 9])
```

Boolean Indexing And Bitwise Operators

- and : `&` (ampersand)

In [164]:
```
# Which values are even AND greater than 6?
x[(x%2 == 0) & (x > 6)]
```

Out[164]:
```
array([8])
```

- or : `|` (pipe)

In [165]:
```
# which values are even OR greater than 6?
x[(x%3 == 0) | (x > 8)]
```

Out[165]:
```
array([6, 9])
```

- not : `~` (tilde)

In [166]:
```
( (x%2 == 0) | (x > 6) )
```

Out[166]:
```
array([False,  True,  True,  True,  True])
```

In [167]:
```
~( (x%2 == 0) | (x > 6) )
```

Out[167]:
```
array([ True, False, False, False, False])
```

Boolean Indexing And Bitwise Operators

- and : `&` (ampersand)

In [164]:
```
# Which values are even AND greater than 6?
x[(x%2 == 0) & (x > 6)]
```

Out[164]:
```
array([8])
```

- or : `|` (pipe)

In [165]:
```
# which values are even OR greater than 6?
x[(x%3 == 0) | (x > 8)]
```

Out[165]:
```
array([6, 9])
```

- not : `~` (tilde)

In [166]:
```
( (x%2 == 0) | (x > 6) )
```

Out[166]:
```
array([False,  True,  True,  True,  True])
```

In [167]:
```
~( (x%2 == 0) | (x > 6) )
```

Out[167]:
```
array([ True, False, False, False, False])
```

In [168]:
```
# which values are NOT (even OR greater than 6)
x[~( (x%2 == 0) | (x > 6) )]
```

Boolean Indexing And Bitwise Operators

- and : `&` (ampersand)

In [164]:
```
# Which values are even AND greater than 6?
x[(x%2 == 0) & (x > 6)]
```

Out[164]:
```
array([8])
```

- or : `|` (pipe)

In [165]:
```
# which values are even OR greater than 6?
x[(x%3 == 0) | (x > 8)]
```

Out[165]:
```
array([6, 9])
```

- not : `~` (tilde)

In [166]:
```
( (x%2 == 0) | (x > 6) )
```

Out[166]:
```
array([False,  True,  True,  True,  True])
```

In [167]:
```
~( (x%2 == 0) | (x > 6) )
```

Out[167]:
```
array([ True, False, False, False, False])
```

In [168]:
```
# which values are NOT (even OR greater than 6)
x[~( (x%2 == 0) | (x > 6) )]
```

Indexing Review

Indexing Review

- standard array indexing (including reverse/negative)
- slicing [start:end:step-size]
- fancy indexing (list/array of indices)
- boolean indexing (list/array of booleans)

Multidimensional Lists

Multidimensional Lists

In [169]:
```
1 x = [[1,2,3],[4,5,6]] # list of lists
2 x
```

Out[169]:
```
[[1, 2, 3], [4, 5, 6]]
```

Multidimensional Lists

In [169]:
```
x = [[1,2,3],[4,5,6]] # list of lists
x
```

Out[169]:
```
[[1, 2, 3], [4, 5, 6]]
```

In [170]:
```
# return first row
x[0]
```

Out[170]:
```
[1, 2, 3]
```

Multidimensional Lists

In [169]:
```
x = [[1,2,3],[4,5,6]] # list of lists
x
```

Out[169]:
```
[[1, 2, 3], [4, 5, 6]]
```

In [170]:
```
# return first row
x[0]
```

Out[170]:
```
[1, 2, 3]
```

In [171]:
```
# return first row, second column
x[1][2]
```

Out[171]:
```
6
```

Multidimensional Lists

In [169]:
```
x = [[1,2,3],[4,5,6]] # list of lists
x
```

Out[169]:
```
[[1, 2, 3], [4, 5, 6]]
```

In [170]:
```
# return first row
x[0]
```

Out[170]:
```
[1, 2, 3]
```

In [171]:
```
# return first row, second column
x[1][2]
```

Out[171]:
```
6
```

In [172]:
```
# return second column?
[row[1] for row in x]
```

Out[172]:
```
[2, 5]
```

NumPy Multidimensional Arrays

NumPy Multidimensional Arrays

In [173]:
```
1 x = np.array([[1,2,3],[4,5,6]])
2 x
```

Out[173]:
```
array([[1, 2, 3],
       [4, 5, 6]])
```

NumPy Multidimensional Arrays

In [173]:
```
x = np.array([[1,2,3],[4,5,6]])
x
```

Out[173]:
```
array([[1, 2, 3],
       [4, 5, 6]])
```

In [174]:
```
x[0,1] # first row, second column
```

Out[174]:
```
2
```

NumPy Multidimensional Arrays

In [173]:
```
x = np.array([[1,2,3],[4,5,6]])
x
```

Out[173]:
```
array([[1, 2, 3],
       [4, 5, 6]])
```

In [174]:
```
x[0,1] # first row, second column
```

Out[174]:
```
2
```

In [175]:
```
x[0,:] # first row
```

Out[175]:
```
array([1, 2, 3])
```

NumPy Multidimensional Arrays

In [173]:
```
x = np.array([[1,2,3],[4,5,6]])
x
```

Out[173]:
```
array([[1, 2, 3],
       [4, 5, 6]])
```

In [174]:
```
x[0,1] # first row, second column
```

Out[174]:
```
2
```

In [175]:
```
x[0,:] # first row
```

Out[175]:
```
array([1, 2, 3])
```

In [176]:
```
x[0,:] # first row (first to last column)
```

Out[176]:
```
array([1, 2, 3])
```

NumPy Multidimensional Arrays

In [173]:
```
x = np.array([[1,2,3],[4,5,6]])
x
```

Out[173]:
```
array([[1, 2, 3],
       [4, 5, 6]])
```

In [174]:
```
x[0,1] # first row, second column
```

Out[174]:
```
2
```

In [175]:
```
x[0,:] # first row
```

Out[175]:
```
array([1, 2, 3])
```

In [176]:
```
x[0,:] # first row (first to last column)
```

Out[176]:
```
array([1, 2, 3])
```

In [177]:
```
x[:,1] # second column (first to last row)
```

Out[177]:
```
array([2, 5])
```

NumPy Array Attributes

NumPy Array Attributes

In [178]:
```
1 x = np.array([[1,2,3],[4,5,6]])
```

NumPy Array Attributes

In [178]:
```
x = np.array([[1,2,3],[4,5,6]])
```

In [179]:
```
x.ndim # number of dimensions
```

Out[179]: 2

NumPy Array Attributes

In [178]:
```
1 x = np.array([[1,2,3],[4,5,6]])
```

In [179]:
```
1 x.ndim # number of dimensions
```

Out[179]: 2

In [180]:
```
1 x.shape # shape in each dimension
```

Out[180]: (2, 3)

NumPy Array Attributes

In [178]:
```
1 x = np.array([[1,2,3],[4,5,6]])
```

In [179]:
```
1 x.ndim # number of dimensions
```

Out[179]: 2

In [180]:
```
1 x.shape # shape in each dimension
```

Out[180]: (2, 3)

In [181]:
```
1 x.size # total number of elements
```

Out[181]: 6

NumPy Operations (UFuncs)

NumPy Operations (UFuncs)

In [182]:

```
1 x = [1,2,3]
2 y = [4,5,6]
```

NumPy Operations (UFuncs)

In [182]:
```
x = [1,2,3]
y = [4,5,6]
```

In [183]:
```
x+y
```

Out[183]:
```
[1, 2, 3, 4, 5, 6]
```

NumPy Operations (UFuncs)

In [182]:
```
x = [1,2,3]
y = [4,5,6]
```

In [183]:
```
x+y
```

Out[183]:
```
[1, 2, 3, 4, 5, 6]
```

In [184]:
```
x = np.array([1,2,3])
y = np.array([4,5,6])
```

NumPy Operations (UFuncs)

In [182]:
```
1 x = [1,2,3]
2 y = [4,5,6]
```

In [183]:
```
1 x+y
```

Out[183]:
```
[1, 2, 3, 4, 5, 6]
```

In [184]:
```
1 x = np.array([1,2,3])
2 y = np.array([4,5,6])
```

In [185]:
```
1 x+y
```

Out[185]:
```
array([5, 7, 9])
```

NumPy Operations (UFuncs)

In [182]:
```
x = [1,2,3]
y = [4,5,6]
```

In [183]:
```
x+y
```

Out[183]:
```
[1, 2, 3, 4, 5, 6]
```

In [184]:
```
x = np.array([1,2,3])
y = np.array([4,5,6])
```

In [185]:
```
x+y
```

Out[185]:
```
array([5, 7, 9])
```

In [186]:
```
%time sum(range(0,int(1e8)))
```

```
CPU times: user 1.56 s, sys: 8.05 ms, total: 1.57 s
Wall time: 1.56 s
```

Out[186]:
```
4999999950000000
```

NumPy Operations (UFuncs)

In [182]:
```
x = [1,2,3]
y = [4,5,6]
```

In [183]:
```
x+y
```

Out[183]:
```
[1, 2, 3, 4, 5, 6]
```

In [184]:
```
x = np.array([1,2,3])
y = np.array([4,5,6])
```

In [185]:
```
x+y
```

Out[185]:
```
array([5, 7, 9])
```

In [186]:
```
%time sum(range(0,int(1e8)))
```

```
CPU times: user 1.56 s, sys: 8.05 ms, total: 1.57 s
Wall time: 1.56 s
```

Out[186]:
```
4999999950000000
```

In [187]:
```
%time np.arange(0,int(1e8)).sum()
```

```
CPU times: user 293 ms, sys: 267 ms, total: 560 ms
Wall time: 574 ms
```

Out[187]:
```
4999999950000000
```

NumPy Broadcasting

NumPy Broadcasting

Allows for vectorized computation on arrays of different sizes

NumPy Broadcasting

Allows for vectorized computation on arrays of different sizes

In [188]:
```
# square every element in a list
x = [1,2,3]
```

NumPy Broadcasting

Allows for vectorized computation on arrays of different sizes

In [188]:
```
# square every element in a list
x = [1,2,3]
```

In [189]:
```
x**2
```

```
TypeError: unsupported operand type(s) for ** or pow(): 'list' and 'int'
```

NumPy Broadcasting

Allows for vectorized computation on arrays of different sizes

In [188]:
```
# square every element in a list
x = [1,2,3]
```

In [189]:
```
x**2
```

```
TypeError: unsupported operand type(s) for ** or pow(): 'list' and 'int'
```

In [190]:
```
# square every element in a numpy array
x = np.array([1,2,3])
```

NumPy Broadcasting

Allows for vectorized computation on arrays of different sizes

In [188]:
```
# square every element in a list
x = [1,2,3]
```

In [189]:
```
x**2
```

```
TypeError: unsupported operand type(s) for ** or pow(): 'list' and 'int'
```

In [190]:
```
# square every element in a numpy array
x = np.array([1,2,3])
```

In [191]:
```
x**2
```

Out[191]:
```
array([1, 4, 9])
```

NumPy Broadcasting

Allows for vectorized computation on arrays of different sizes

In [188]:
```
# square every element in a list
x = [1,2,3]
```

In [189]:
```
x**2
```

```
TypeError: unsupported operand type(s) for ** or pow(): 'list' and 'int'
```

In [190]:
```
# square every element in a numpy array
x = np.array([1,2,3])
```

In [191]:
```
x**2
```

Out[191]:
```
array([1, 4, 9])
```

In [192]:
```
a = np.array([1.0, 2.0, 3.0])
b = 2.0
a * b
```

Out[192]:
```
array([2., 4., 6.])
```

NumPy Broadcasting

Allows for vectorized computation on arrays of different sizes

In [188]:
```
# square every element in a list
x = [1,2,3]
```

In [189]:
```
x**2
```

```
TypeError: unsupported operand type(s) for ** or pow(): 'list' and 'int'
```

In [190]:
```
# square every element in a numpy array
x = np.array([1,2,3])
```

In [191]:
```
x**2
```

Out[191]:
```
array([1, 4, 9])
```

In [192]:
```
a = np.array([1.0, 2.0, 3.0])
b = 2.0
a * b
```

Out[192]:
```
array([2., 4., 6.])
```

is equivalent to

NumPy Broadcasting

Allows for vectorized computation on arrays of different sizes

In [188]:
```
# square every element in a list
x = [1,2,3]
```

In [189]:
```
x**2
```

```
TypeError: unsupported operand type(s) for ** or pow(): 'list' and 'int'
```

In [190]:
```
# square every element in a numpy array
x = np.array([1,2,3])
```

In [191]:
```
x**2
```

Out[191]:
```
array([1, 4, 9])
```

In [192]:
```
a = np.array([1.0, 2.0, 3.0])
b = 2.0
a * b
```

Out[192]:
```
array([2., 4., 6.])
```

is equivalent to

NumPy `random` Submodule

NumPy `random` Submodule

Provides many random sampling functions

In [194]:
```
1 from numpy.random import seed, randn
```

In [195]:
```
1 seed(1990)
```

In [196]:
```
1 randn(10)
```

Out[196]:
```
array([-0.38232581,  0.20089394,  1.56176025, -0.2637672 , -0.58363805,
        1.22412869, -0.32044072, -1.06974806, -0.31455216, -0.41181193])
```

NumPy `random` Submodule

Provides many random sampling functions

```
In [194]:
1 from numpy.random import seed, randn
```

```
In [195]:
1 seed(1990)
```

```
In [196]:
1 randn(10)
```

```
Out[196]: array([-0.38232581,  0.20089394,  1.56176025, -0.2637672 , -0.58363805,
                 1.22412869, -0.32044072, -1.06974806, -0.31455216, -0.41181193])
```

`from numpy.random import ...`

- `rand` : random floats
- `randint` : random integers
- `randn` : standard normal distribution
- `permutation` : random permutation

Questions?