[220] Files and Directories

Relation to Recent Topics...

1. file objects are a type of iterator

2. exceptions occur a lot with files
   not found, out of space, no permissions, etc

Learning Objectives Today

Basic file interactions

- opening/closing
- reading/writing

OS module

- listdir, mkdir, exists, isdir, isfile, join

File exceptions

Encodings

Learning Objectives Today

Basic file interactions

- opening/closing
- reading/writing

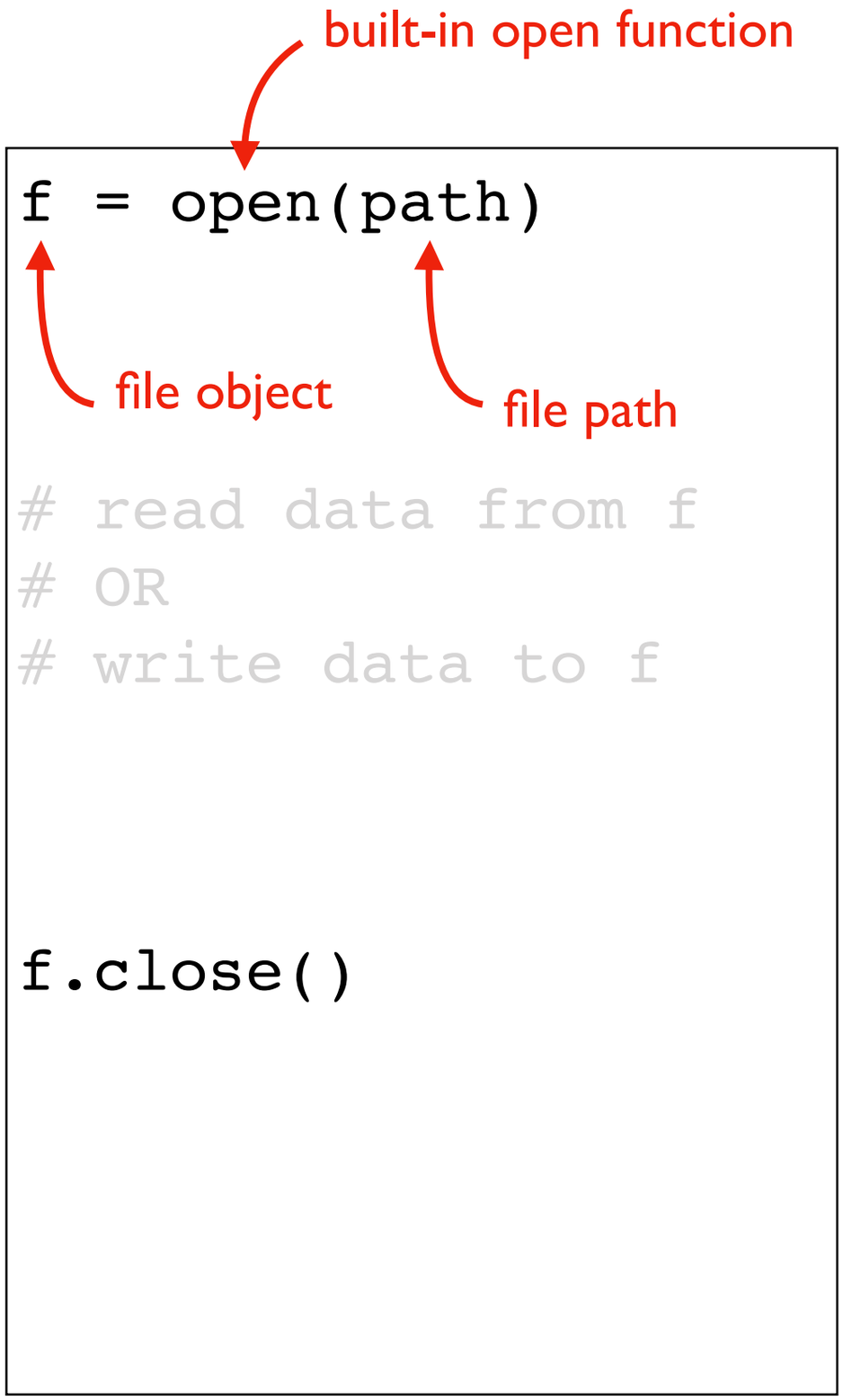
OS module

- listdir, mkdir, exists, isdir, isfile, join

File exceptions

Encodings

File objects



File objects

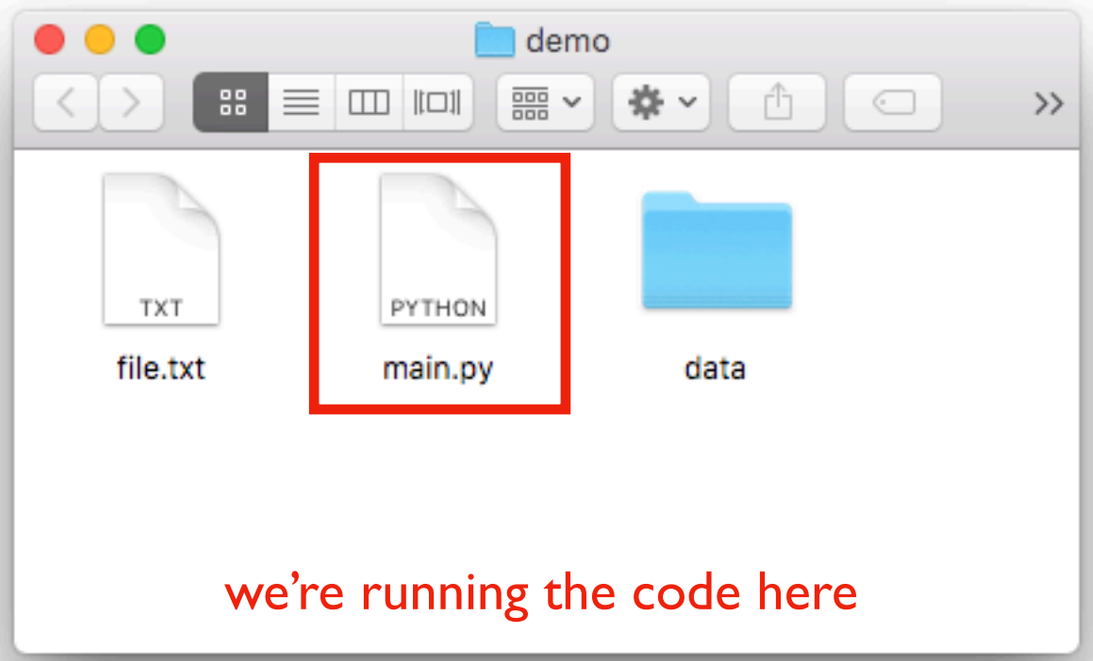
main.py:

built-in open function

```
f = open(path)

# read data from f
# OR
# write data to f



f.close()
```

file object

file path



File objects

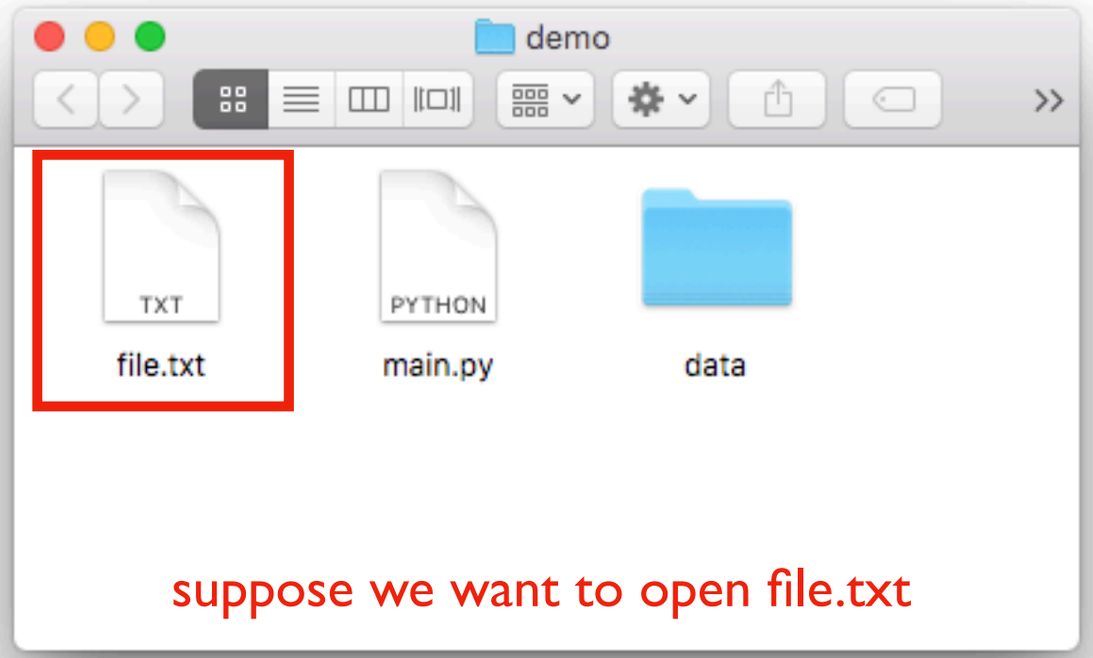
built-in open function

main.py:

```
f = open("file.txt")

# read data from f
# OR
# write data to f



f.close()
```

file object

file path



File objects

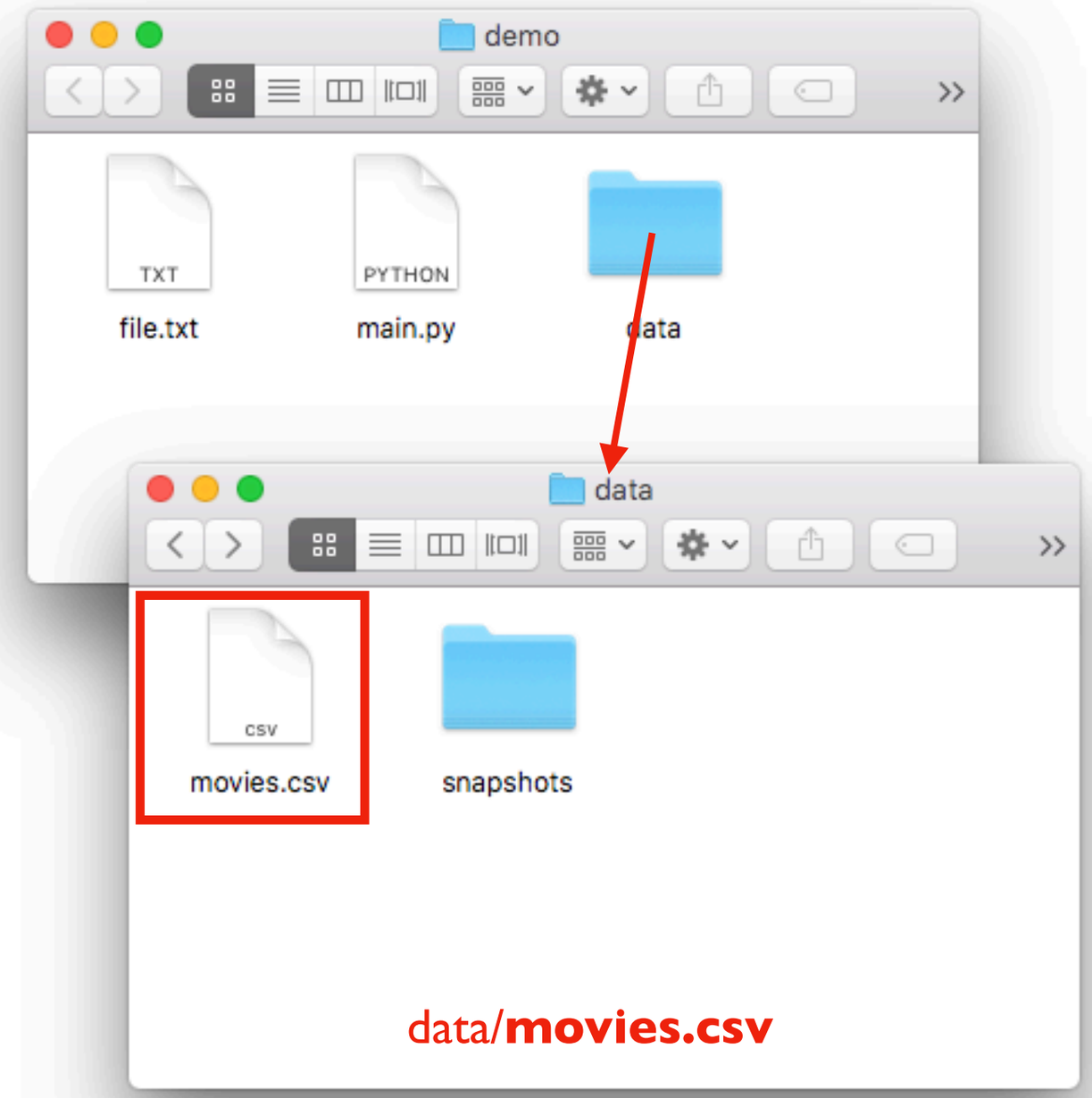
main.py:

built-in open function

```
f = open(
    "data/movies.csv")


# read data from f
# OR
# write data to f



f.close()
```

file object

file path



File objects

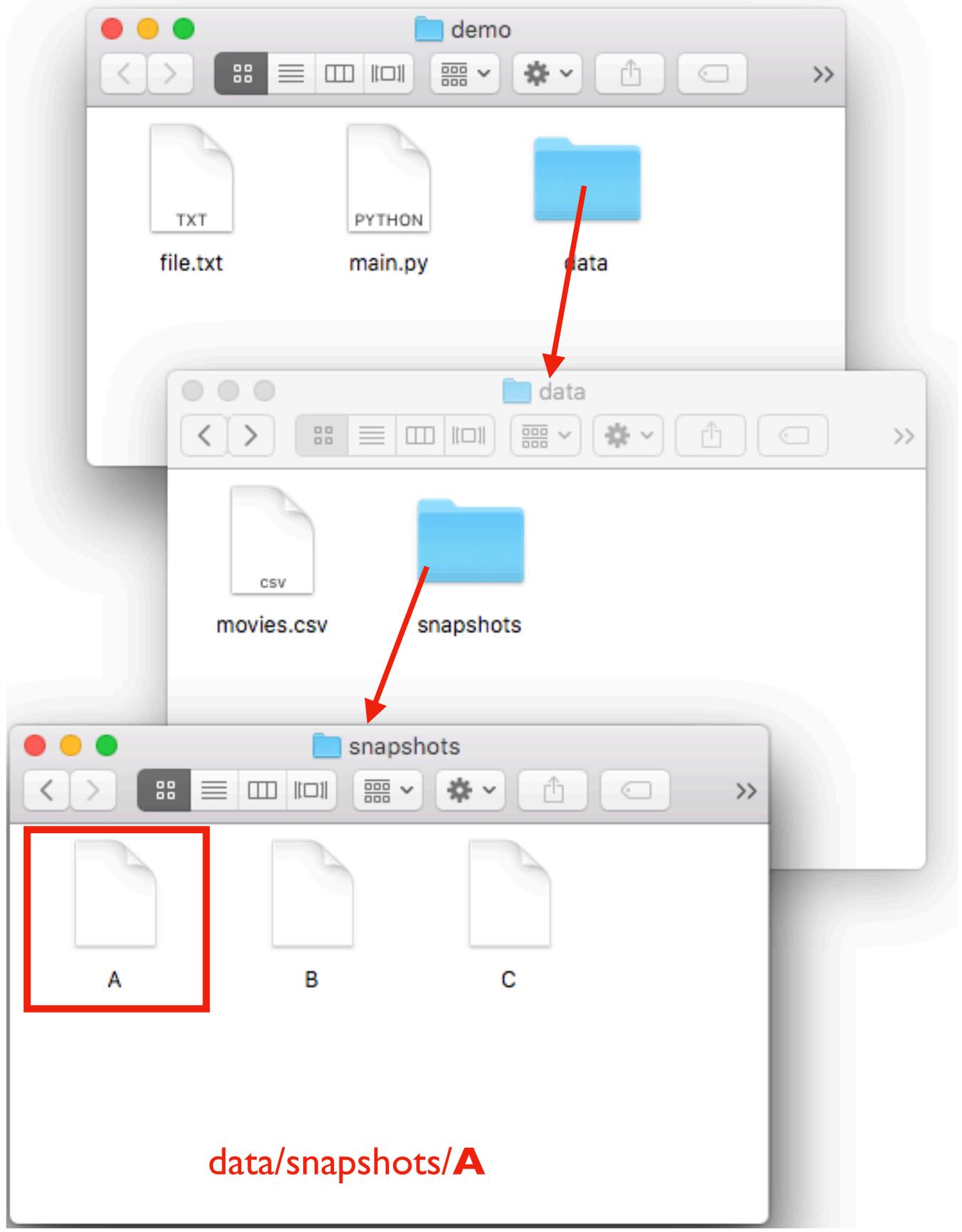
main.py:

built-in open function

```
f = open(
"data/snapshots/A")

# read data from f
# OR
# write data to f


f.close()
```

file object

file path



File objects

main.py:

```
f = open("file.txt")



# read data from f
# OR
# write data to f



f.close()
```

File objects

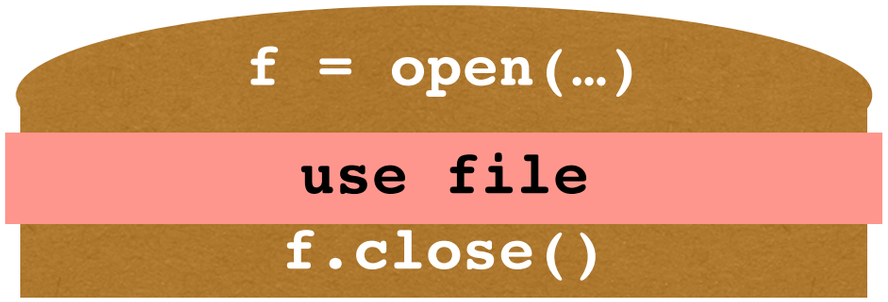
main.py:

```
f = open("file.txt")

# read data from f
# OR
# write data to f

f.close()
```

using file

File objects

imagine a file object as a sandwich…



main.py:

```
f = open("file.txt")

# read data from f
# OR
# write data to f

f.close()
```

using file

cleanup

Reasons for closing

- avoid data loss
- limited number of open files

Learning Objectives Today

Basic file interactions

- opening/closing
- reading/writing

OS module

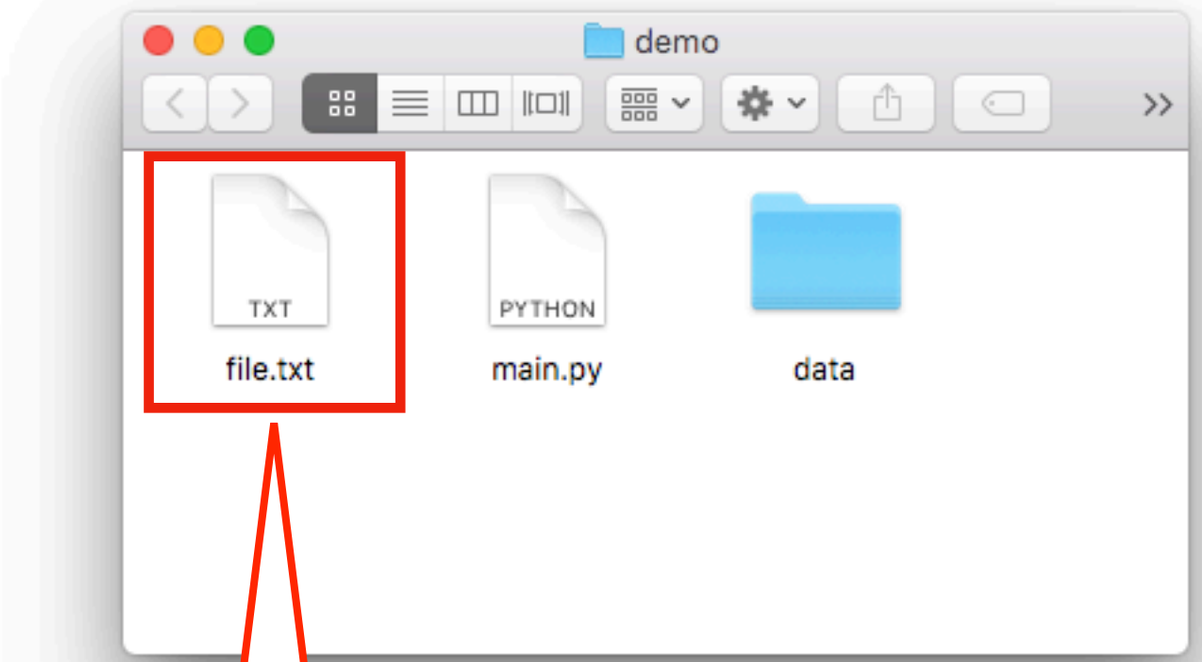
- listdir, mkdir, exists, isdir, isfile, join

File exceptions

Encodings

Reading a file

```
f = open("file.txt")



# read data from f
# OR
# write data to f



f.close()
```



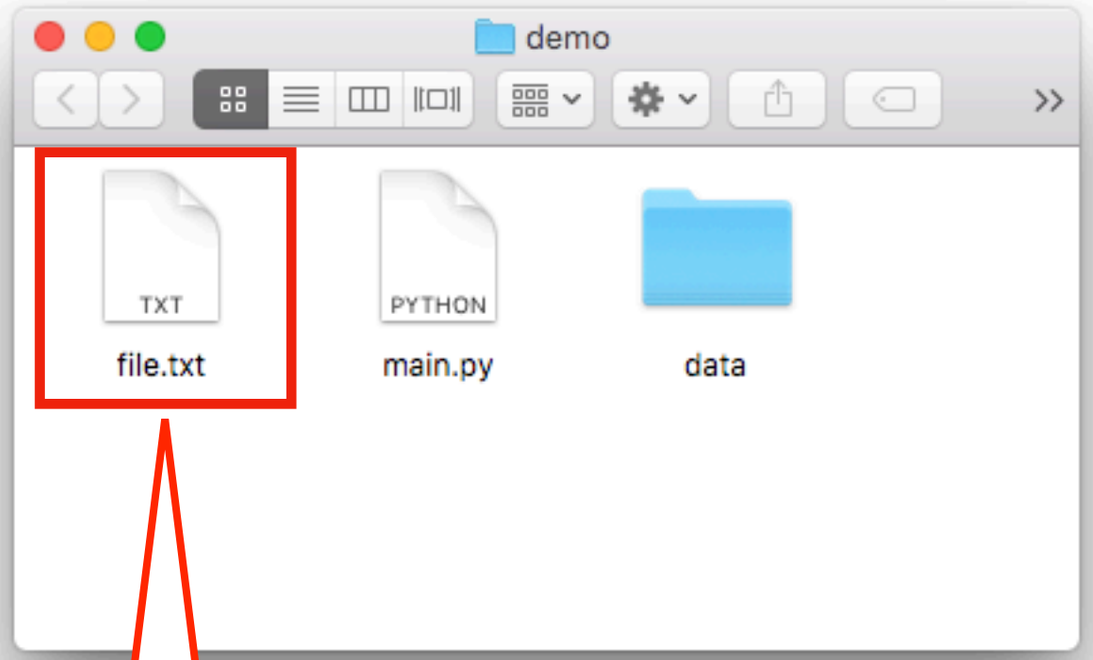
I promise
to always
close my files

Reading a file

```
f = open("file.txt")


data = f.read()

print(data)


f.close()
```

Option I

data is: "I promise\nto always\nclose my files"



I promise
to always
close my files

`read()` method

- fetch entire file contents
- return as a string

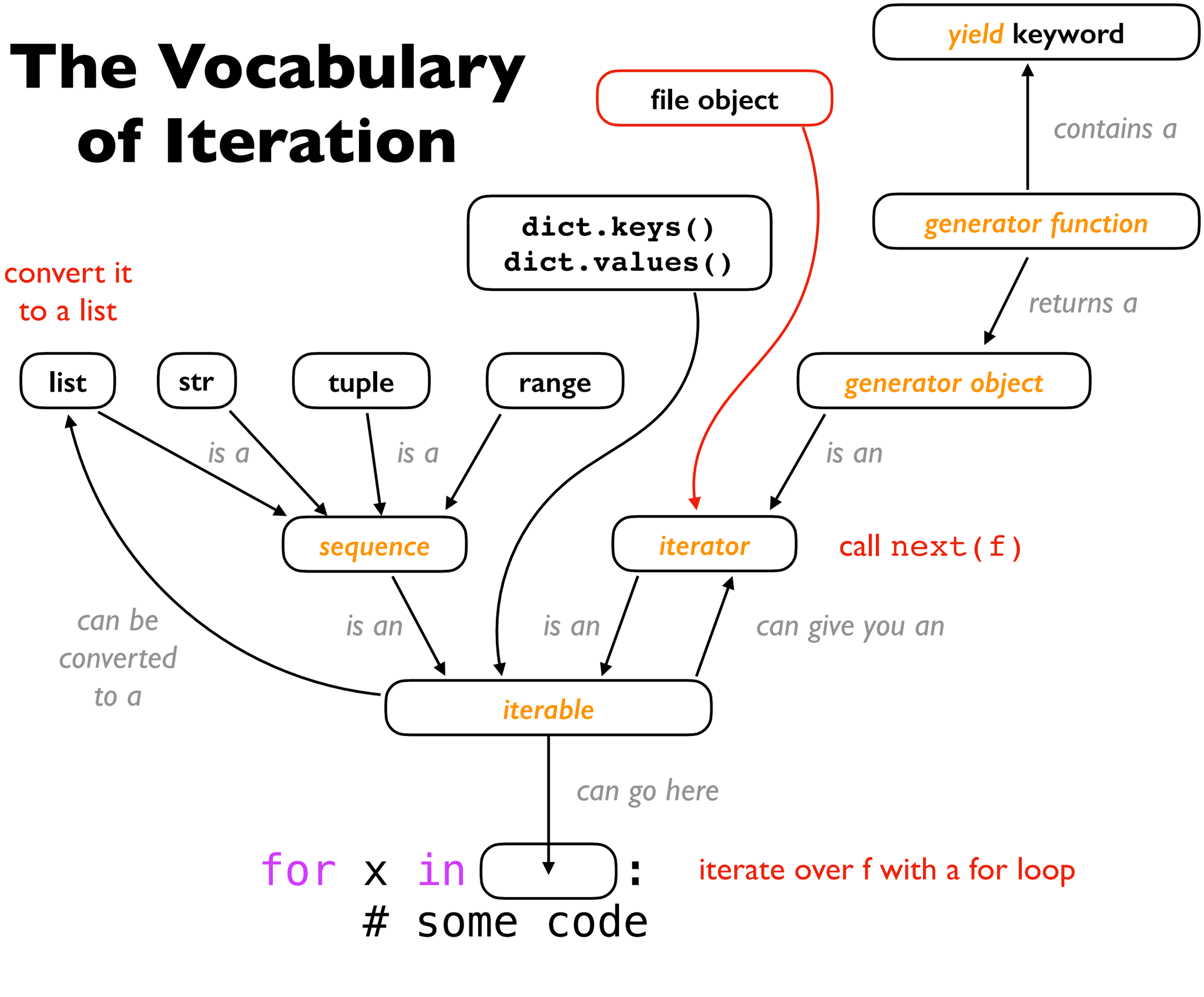
Reading a file

```
f = open("file.txt")



# read data from f
# OR
# write data to f



f.close()
```

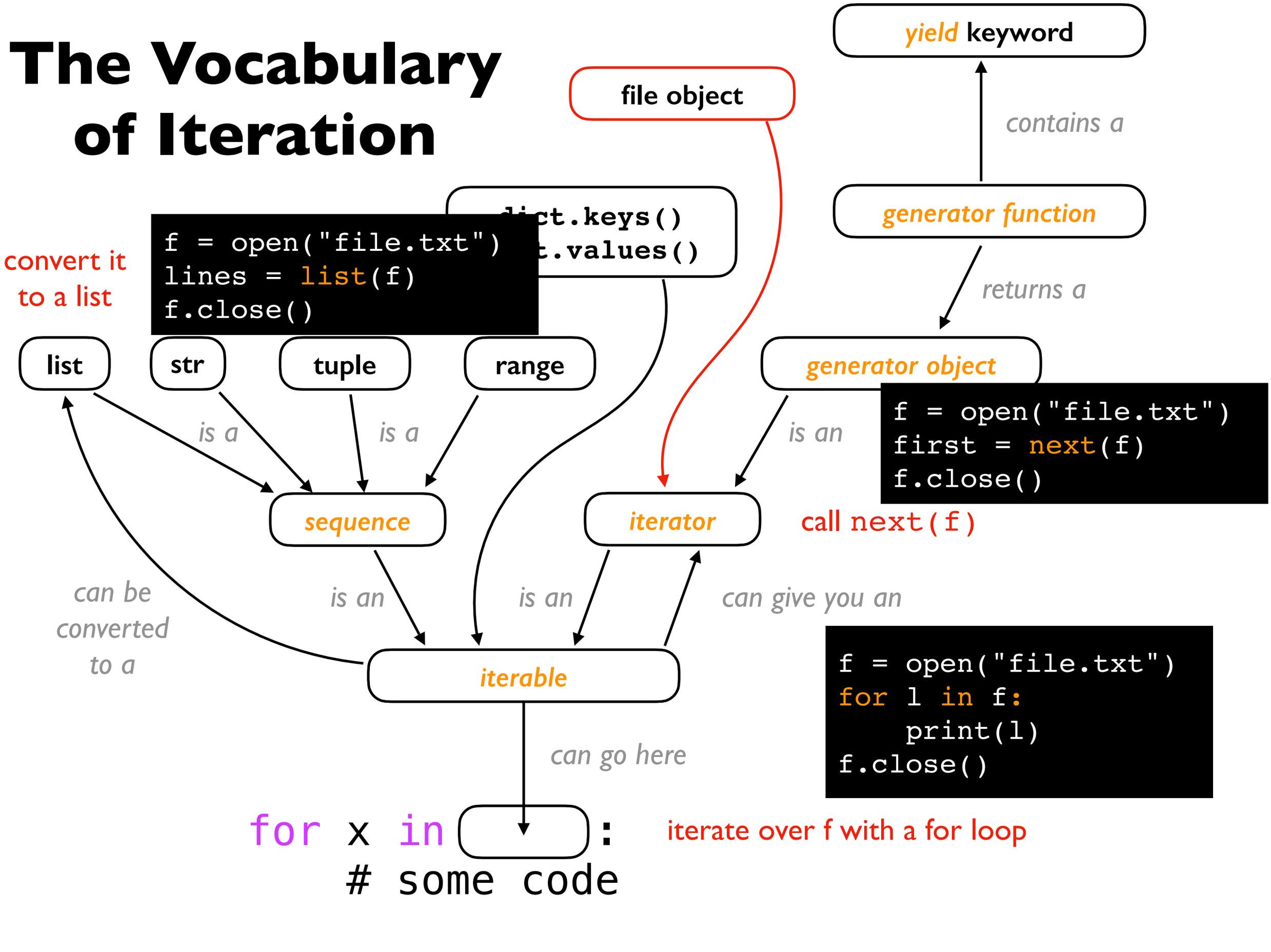
Option 2

file objects are iterators!

The Vocabulary of Iteration

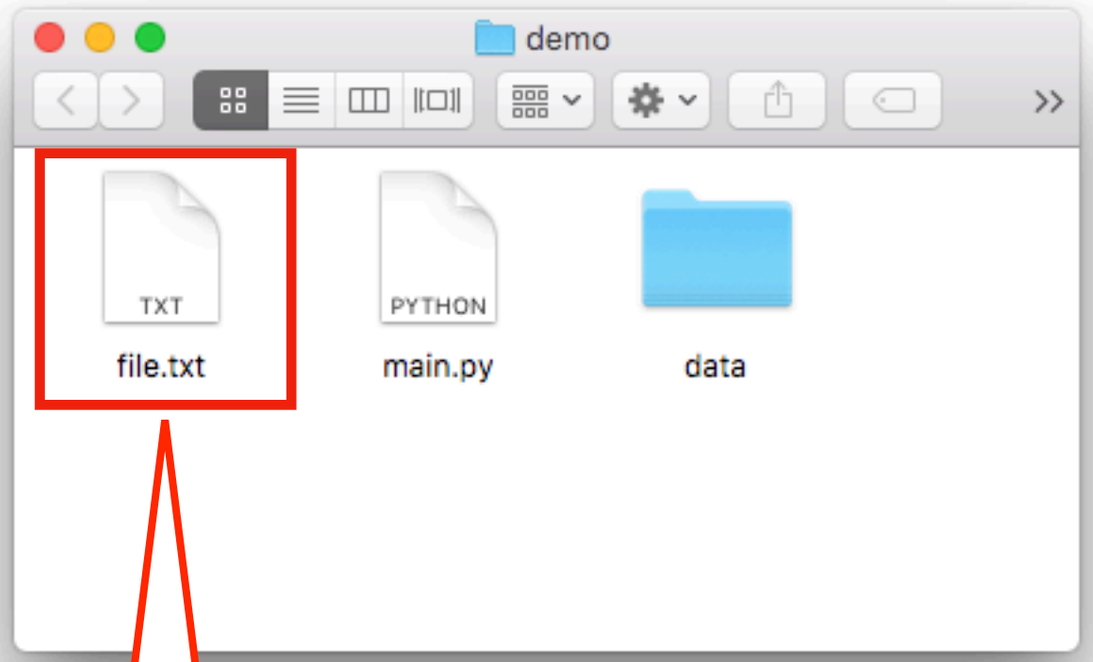


The Vocabulary of Iteration



Write a file

```
f = open("file.txt")



# read data from f
# OR
# write data to f



f.close()
```

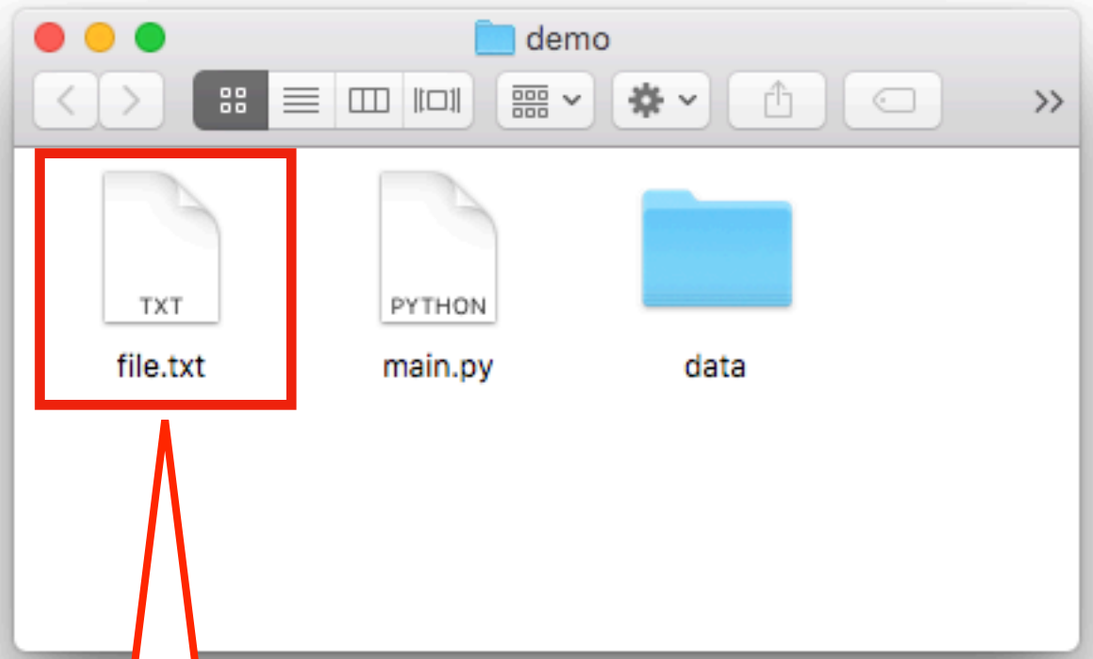


I promise
to always
close my files

Write a file

"w" mode indicates we want to write to this file

```
f = open("file.txt", "w")



# read data from f
# OR
# write data to f



f.close()
```

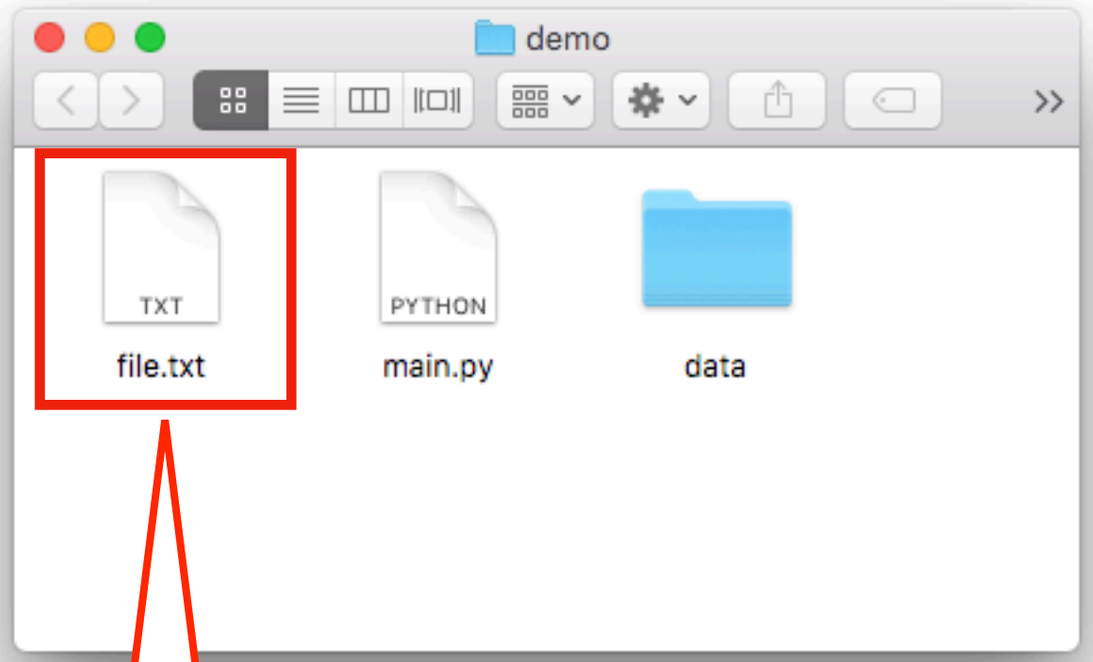


I promise to always close my files

Write a file

"w" mode indicates we want to write to this file

```
f = open("file.txt", "w")



f.write("hello")
f.write(" world\n")
f.write("!!!!!\n")



f.close()
```



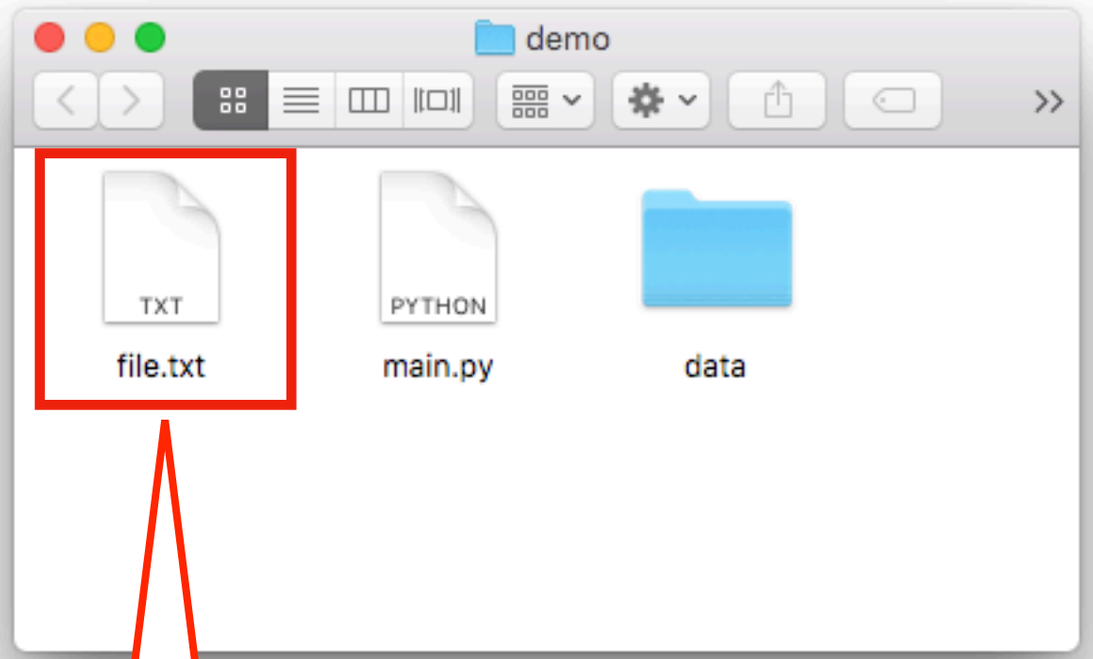
I promise to always close my files

let's run it!

Write a file

"w" mode indicates we want to write to this file

```
f = open("file.txt", "w")



f.write("hello")
f.write(" world\n")
f.write("!!!!!\n")



f.close()
```



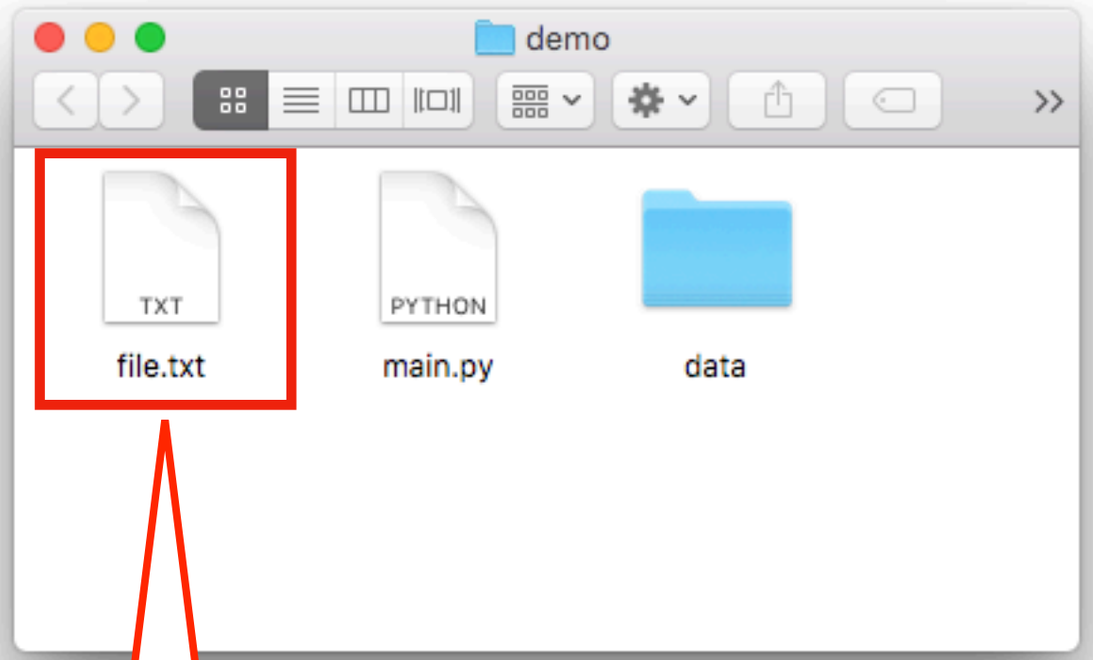
opening with "w" is dangerous. It immediately wipes out your file.

(or creates a new one if there isn't already a file.txt)

Write a file

"w" mode indicates we want to write to this file

```
f = open("file.txt", "w")



f.write("hello")
f.write(" world\n")
f.write("!!!!!\n")



f.close()
```

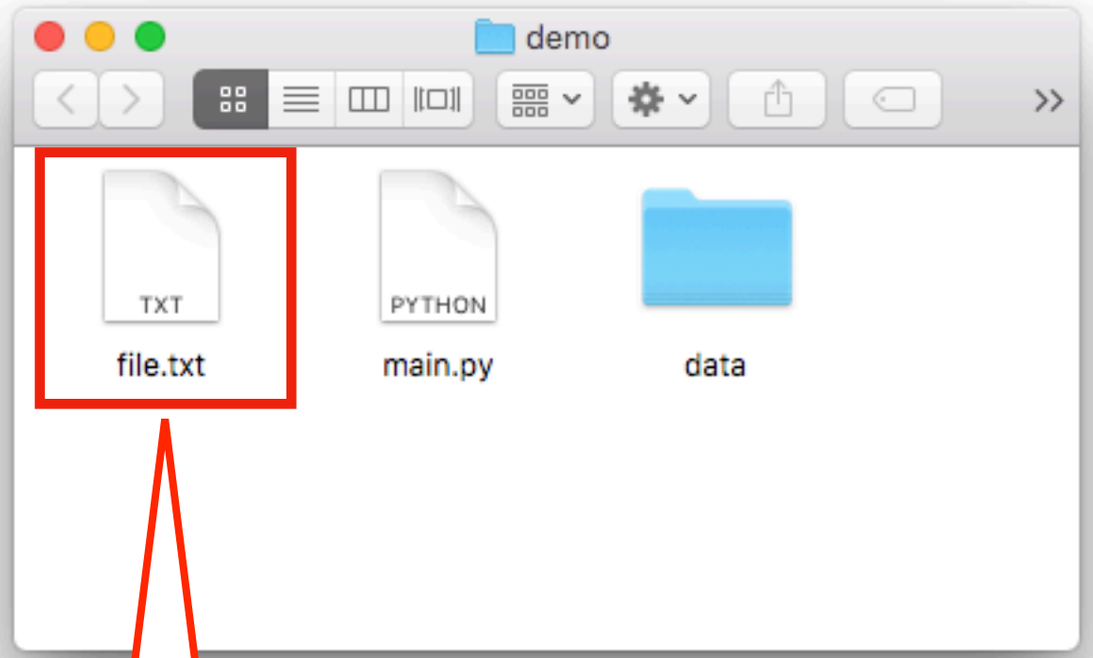



```
hello
```

Write a file

"w" mode indicates we want to write to this file

```
f = open("file.txt", "w")



f.write("hello")
f.write(" world\n")
f.write("!!!!!\n")



f.close()
```

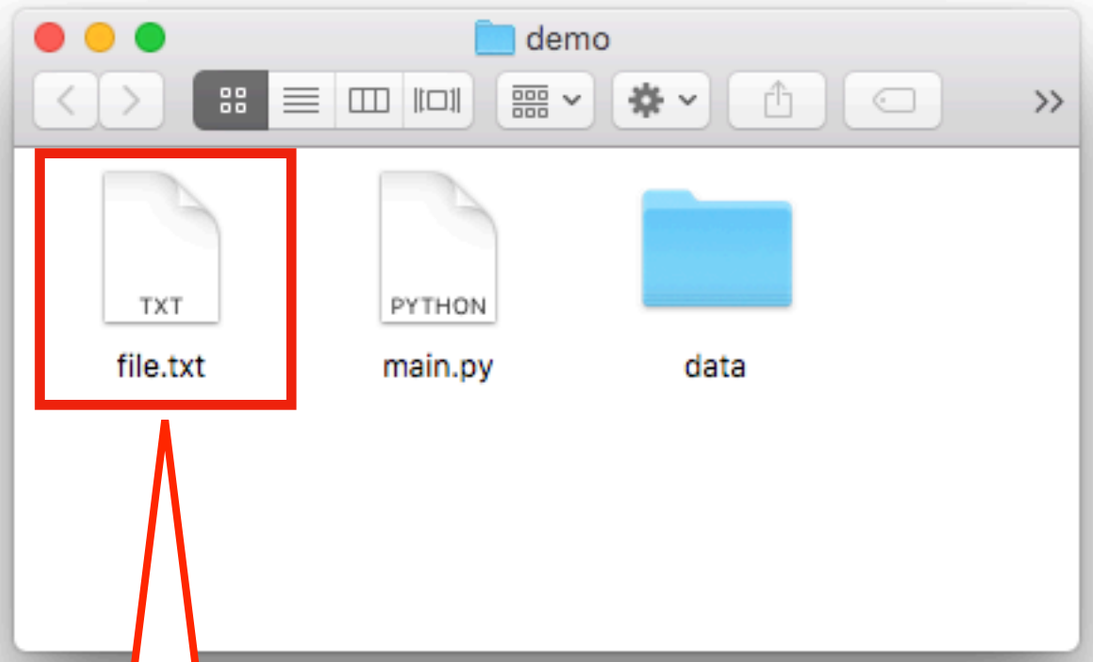



```
hello world
```

Write a file

"w" mode indicates we want to write to this file

```
f = open("file.txt", "w")



f.write("hello")
f.write(" world\n")
f.write("!!!!!\n")



f.close()
```

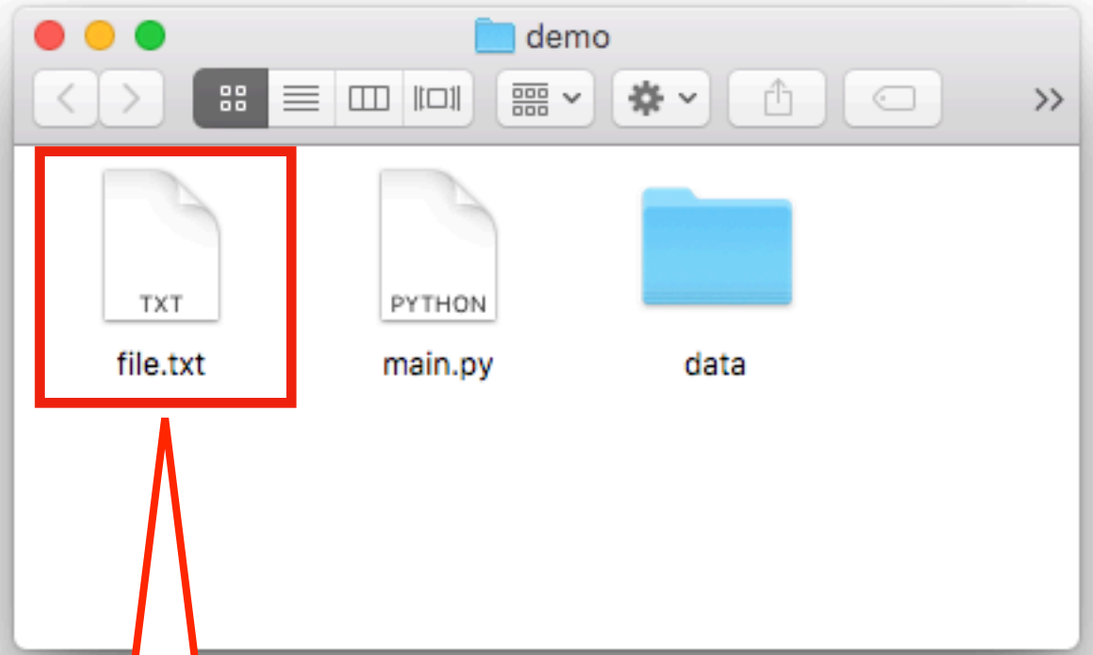



```
hello world
!!!!!!
```

Write a file

"w" mode indicates we want to write to this file

```
f = open("file.txt", "w")



f.write("hello")
f.write(" world\n")
f.write("!!!!!\n")



f.close()
```



be careful with newlines

(write doesn't add them like print does)

Learning Objectives Today

Basic file interactions

- opening/closing
- reading/writing

OS module

- listdir, mkdir, exists, isdir, isfile, join

File exceptions

Encodings

OS Module (Operating System)

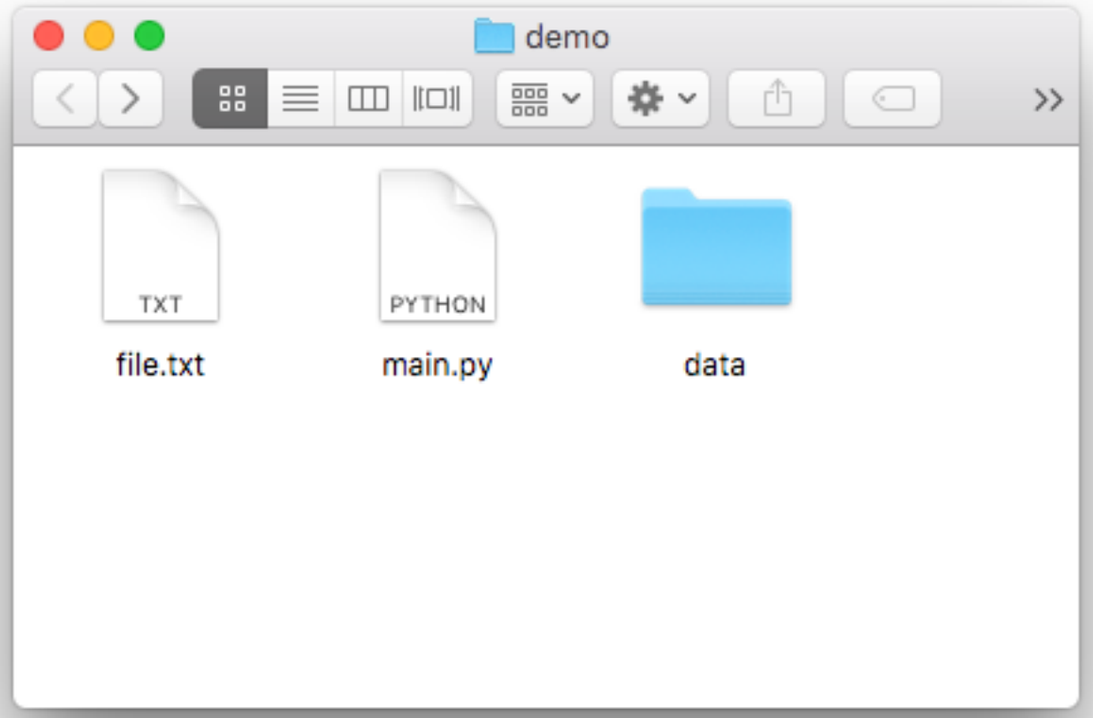
Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- os.path.exists
- os.path.isfile
- os.path.isdir
- os.path.join

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- **os.listdir**
- os.mkdir
- os.path.exists
- os.path.isfile
- os.path.isdir
- os.path.join



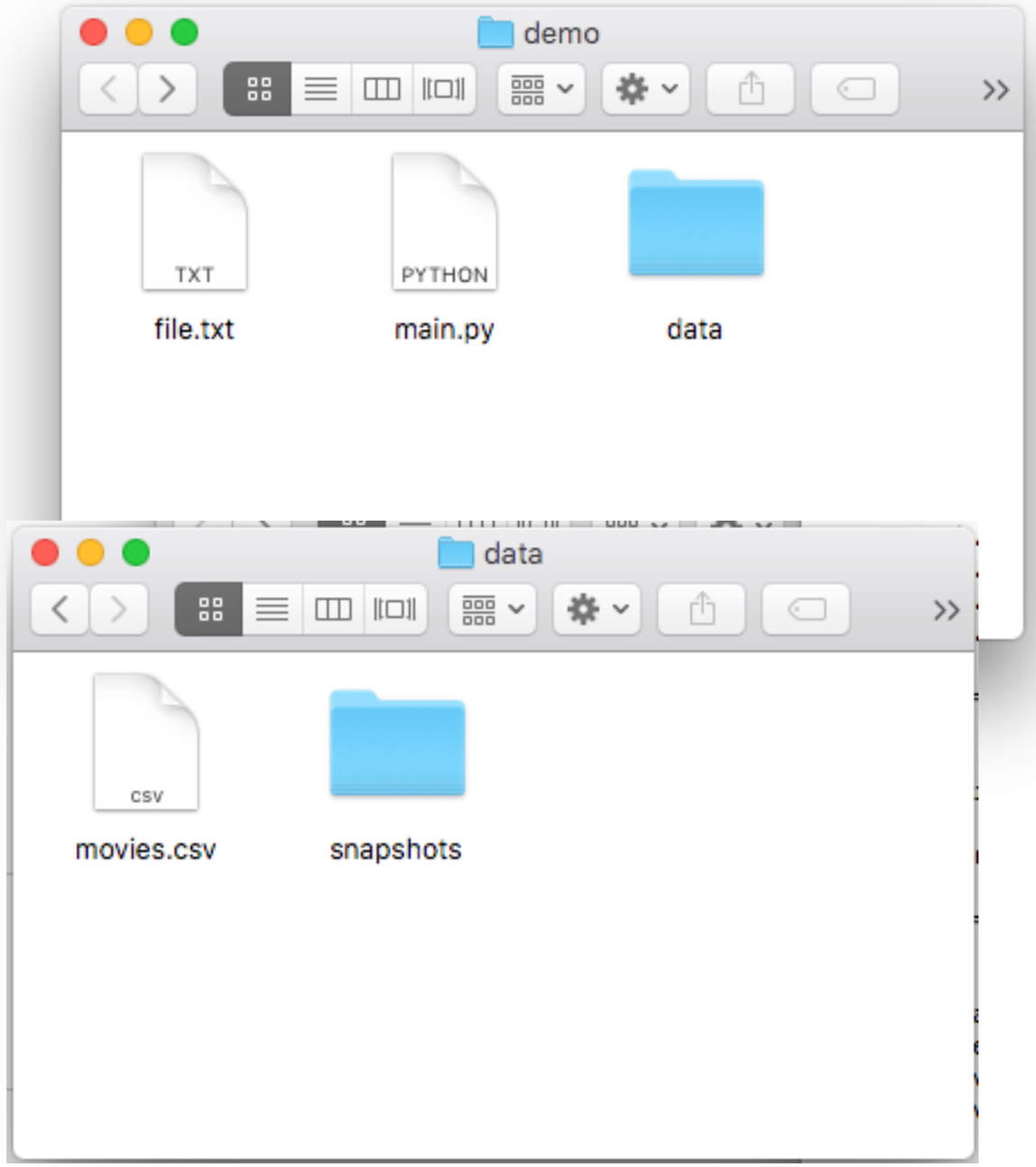

```
>>> import os
>>> os.listdir(“.”)
[“file.txt”, “main.py”, “data”]
```

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- **os.listdir**
- os.mkdir
- os.path.exists
- os.path.isfile
- os.path.isdir
- os.path.join

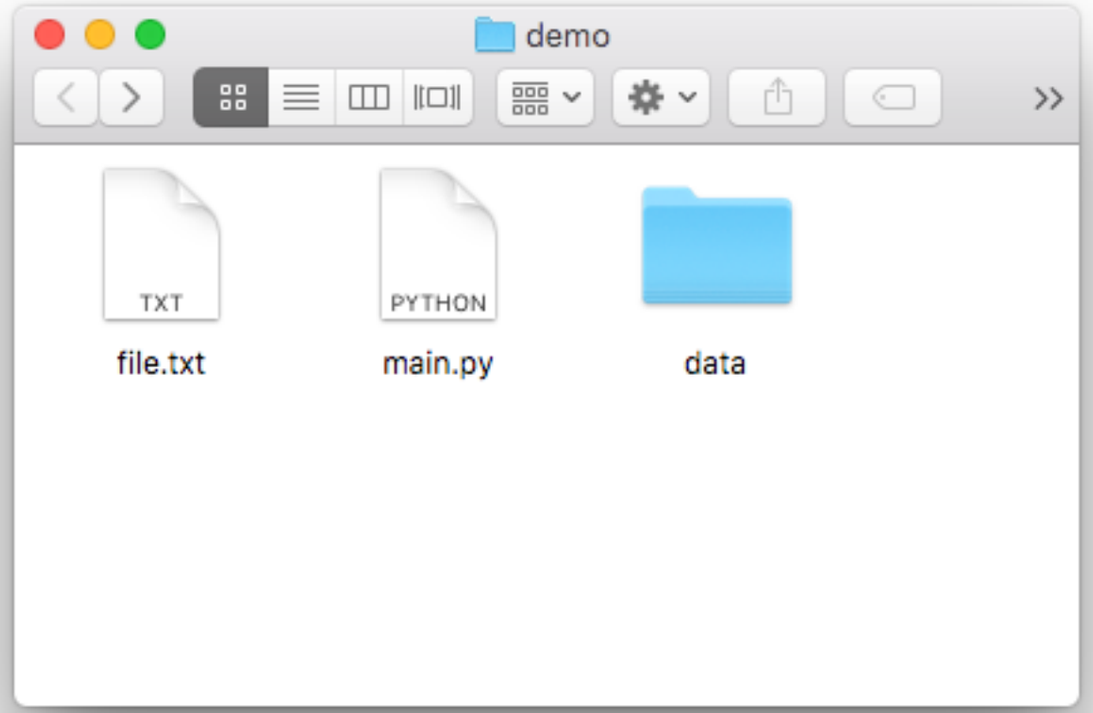
```
>>> import os
>>> os.listdir("data")
["movies.csv", "snapshots"]
```



OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- **os.mkdir**
- os.path.exists
- os.path.isfile
- os.path.isdir
- os.path.join



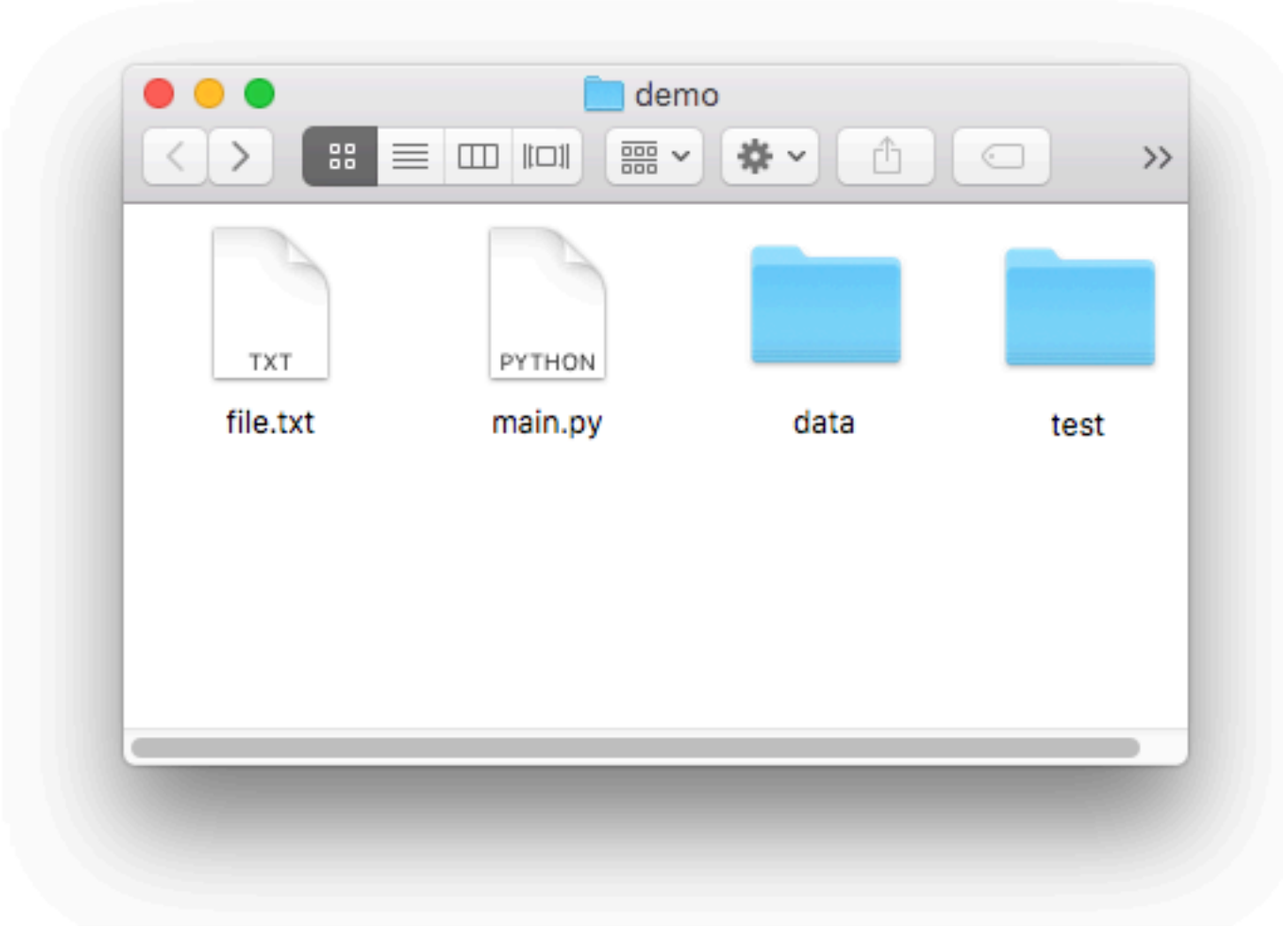

```
>>> import os
>>> os.mkdir(“test”)
```

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- **os.mkdir**
- os.path.exists
- os.path.isfile
- os.path.isdir
- os.path.join

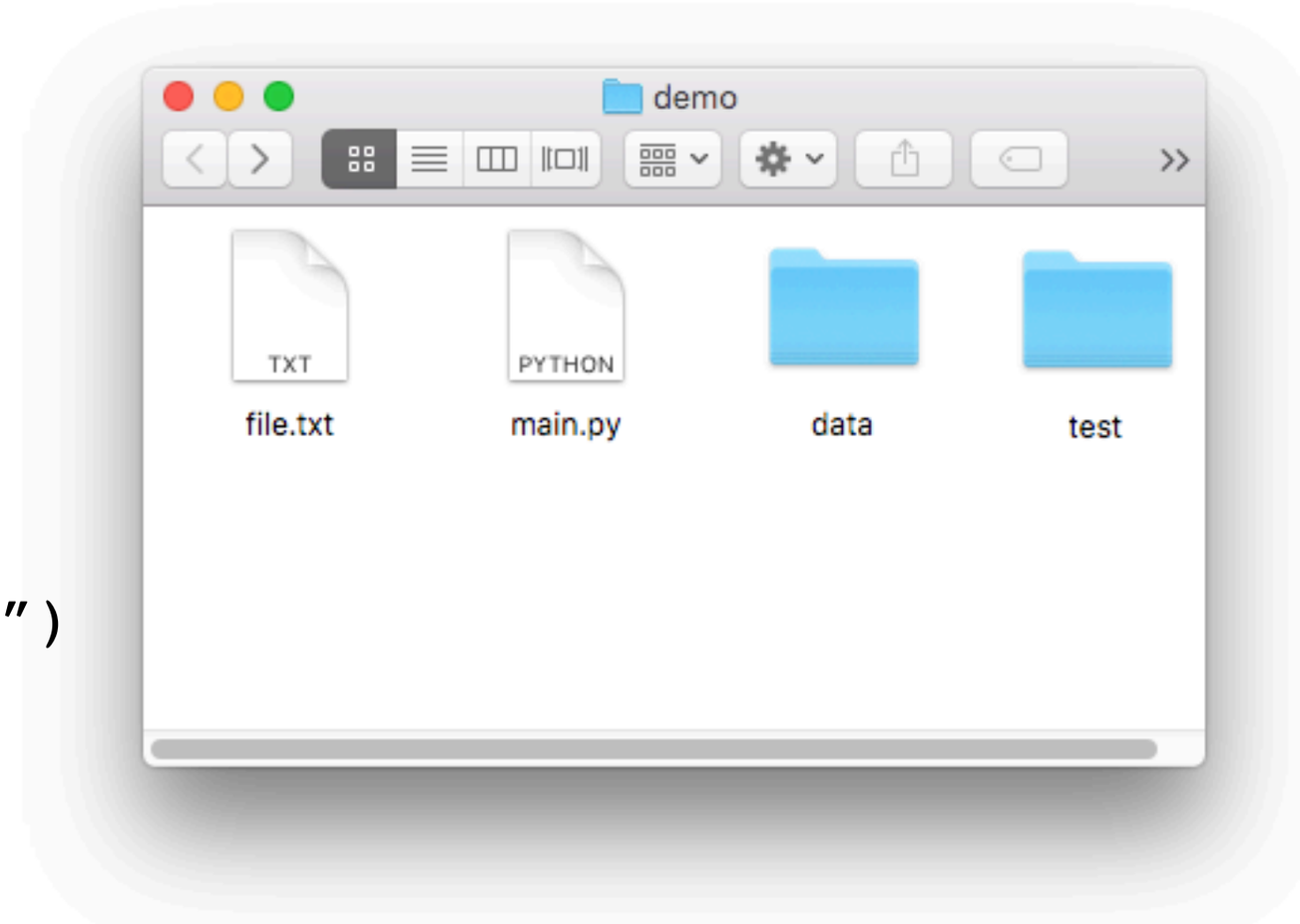
```
>>> import os
>>> os.mkdir("test")
```



OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- **os.path.exists**
- os.path.isfile
- os.path.isdir
- os.path.join

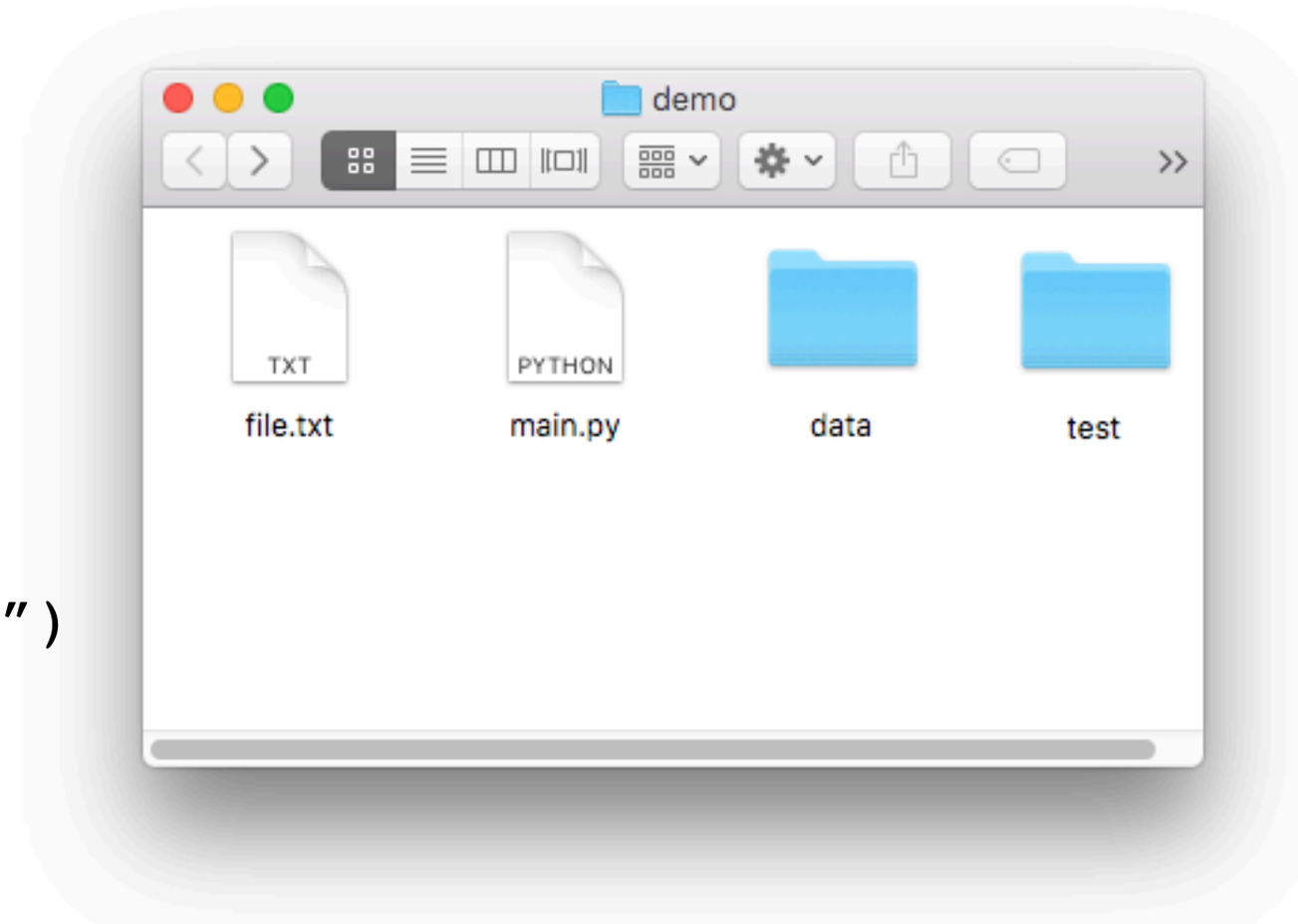



```
>>> import os
>>> os.path.exists(“file.txt”)
True
```

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- **os.path.exists**
- os.path.isfile
- os.path.isdir
- os.path.join

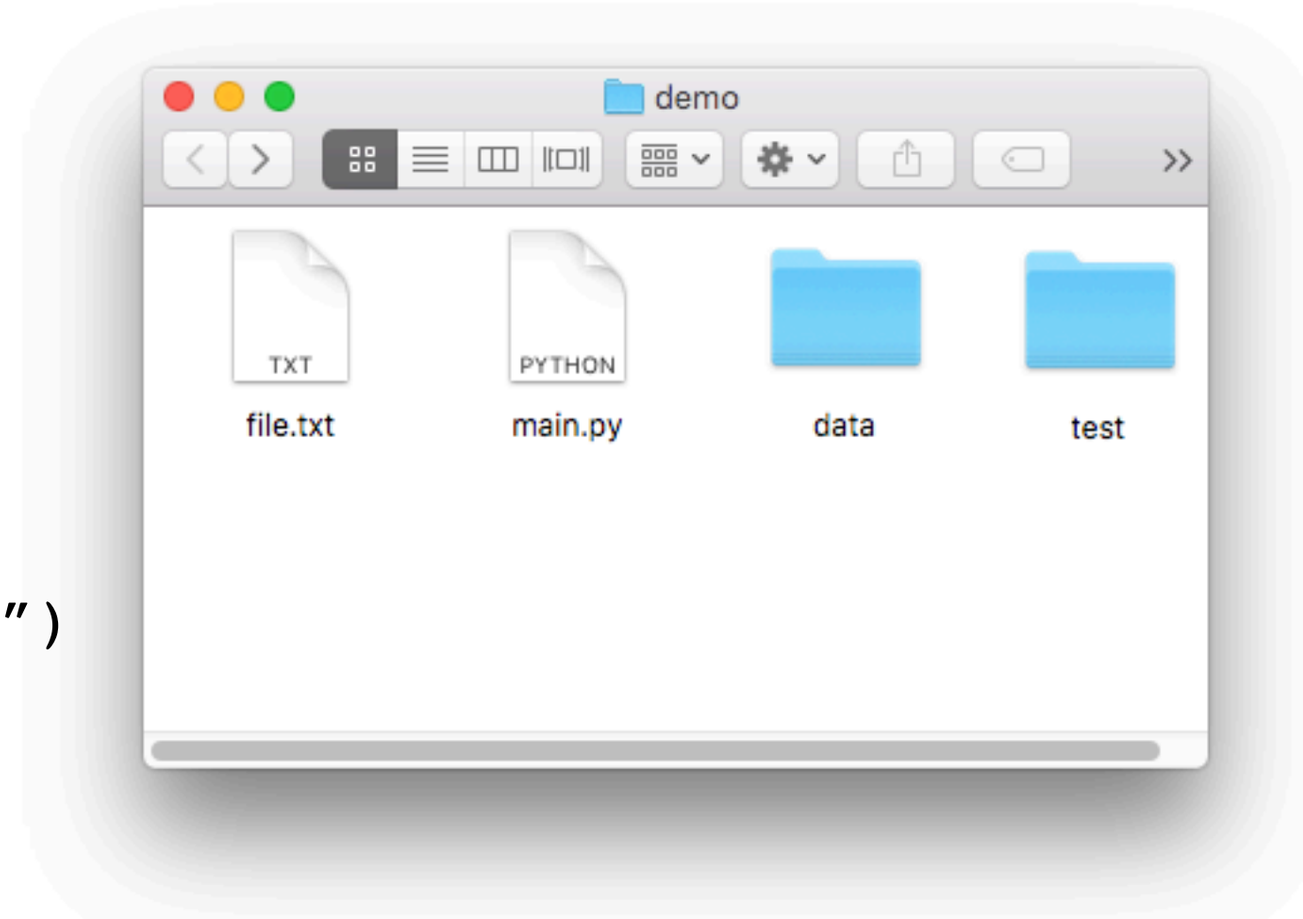



```
>>> import os
>>> os.path.exists(“haha.txt”)
False
```

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- os.path.exists
- **os.path.isfile**
- os.path.isdir
- os.path.join

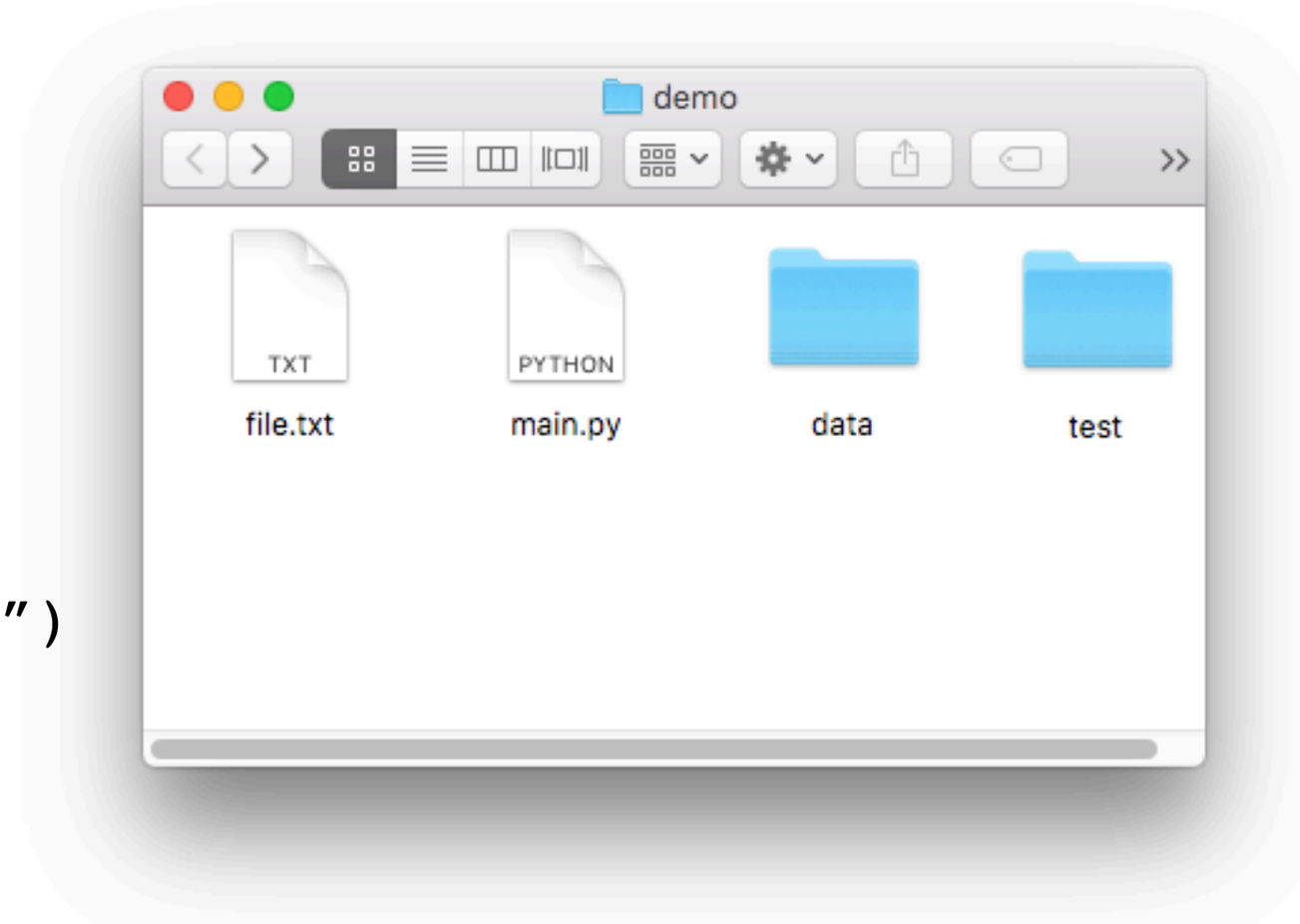



```
>>> import os
>>> os.path.isfile(“haha.txt”)
False
```

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- os.path.exists
- **os.path.isfile**
- os.path.isdir
- os.path.join



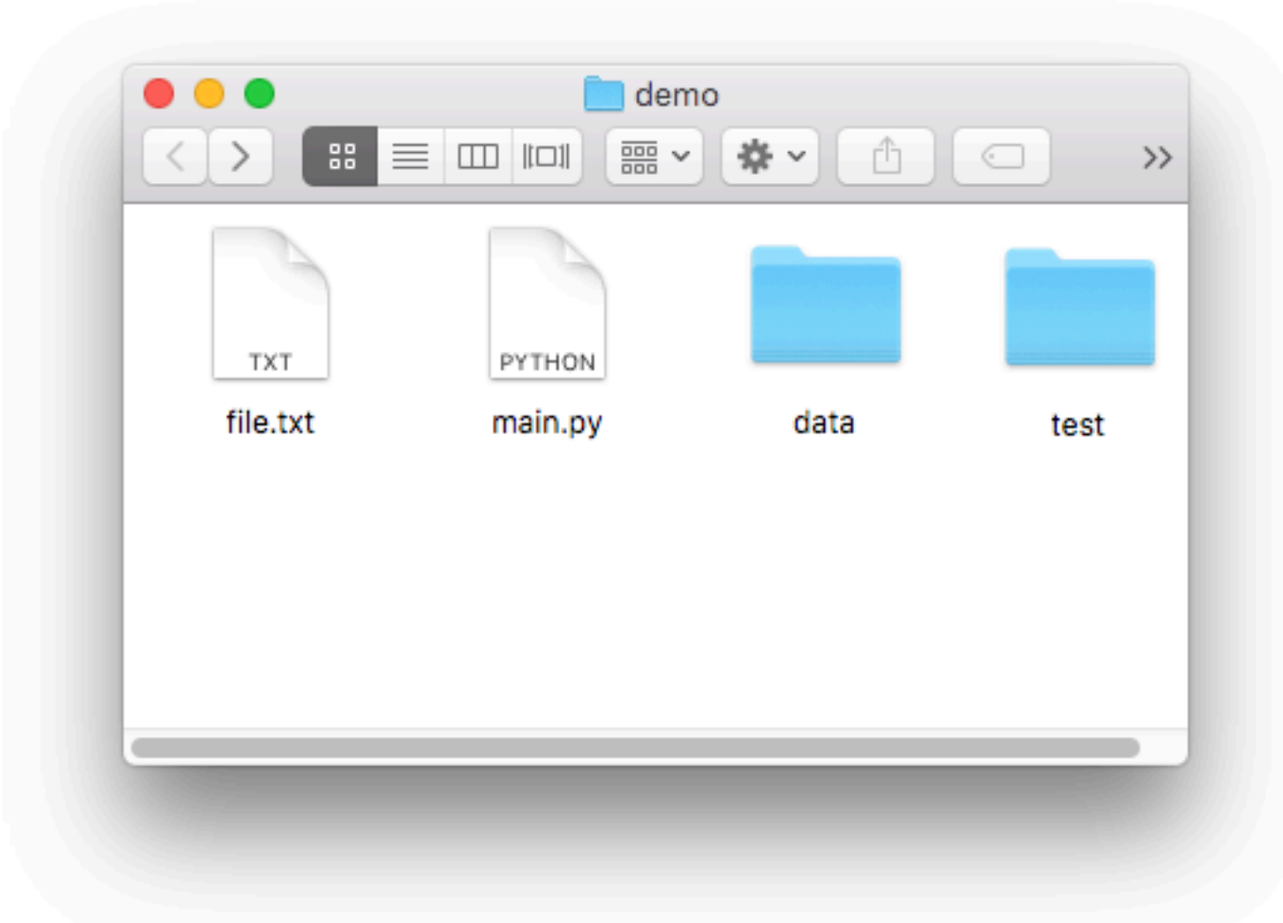

```
>>> import os
>>> os.path.isfile(“file.txt”)
True
```

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- os.path.exists
- **os.path.isfile**
- os.path.isdir
- os.path.join

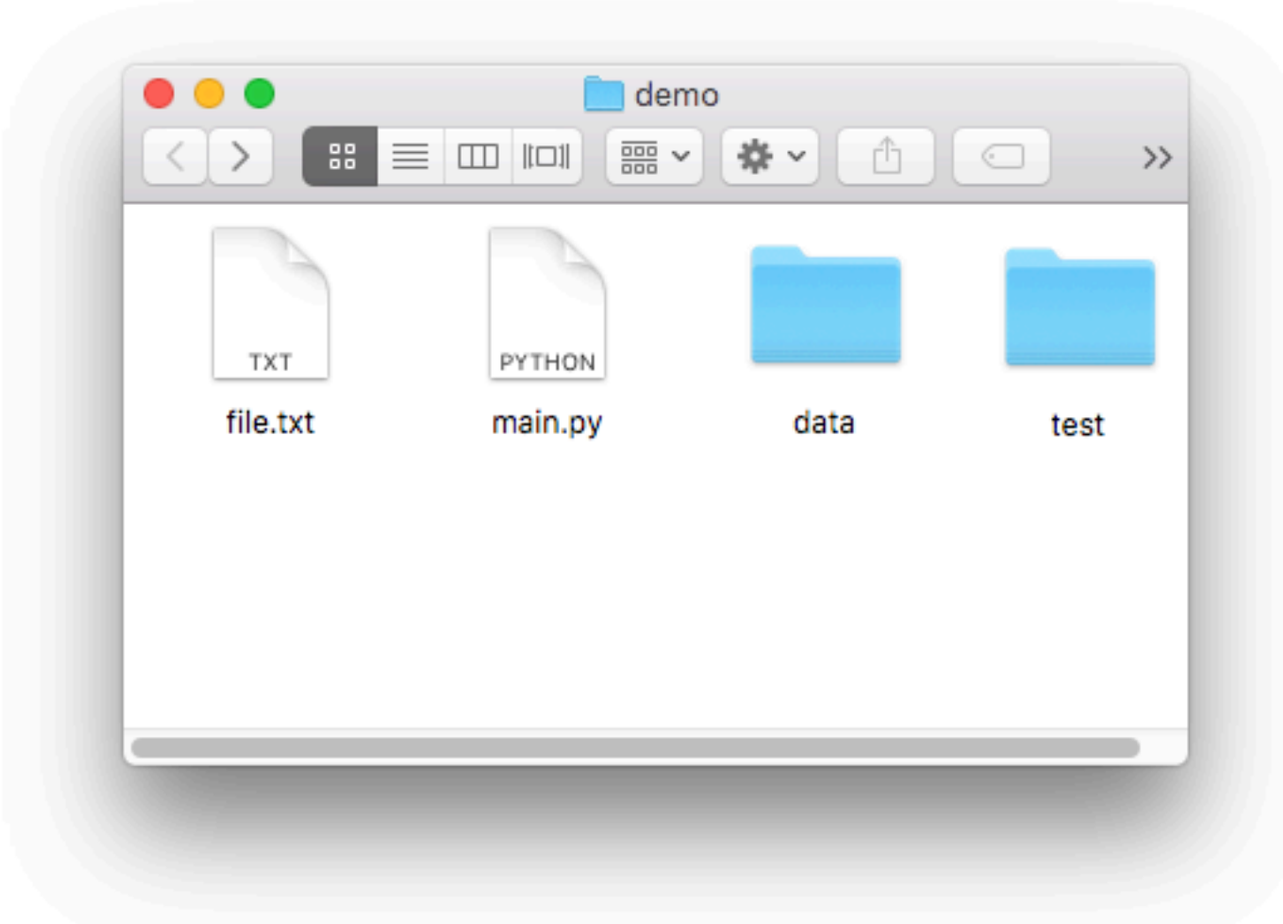
```
>>> import os
>>> os.path.isfile(“data”)
False
```



OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- os.path.exists
- os.path.isfile
- **os.path.isdir**
- os.path.join

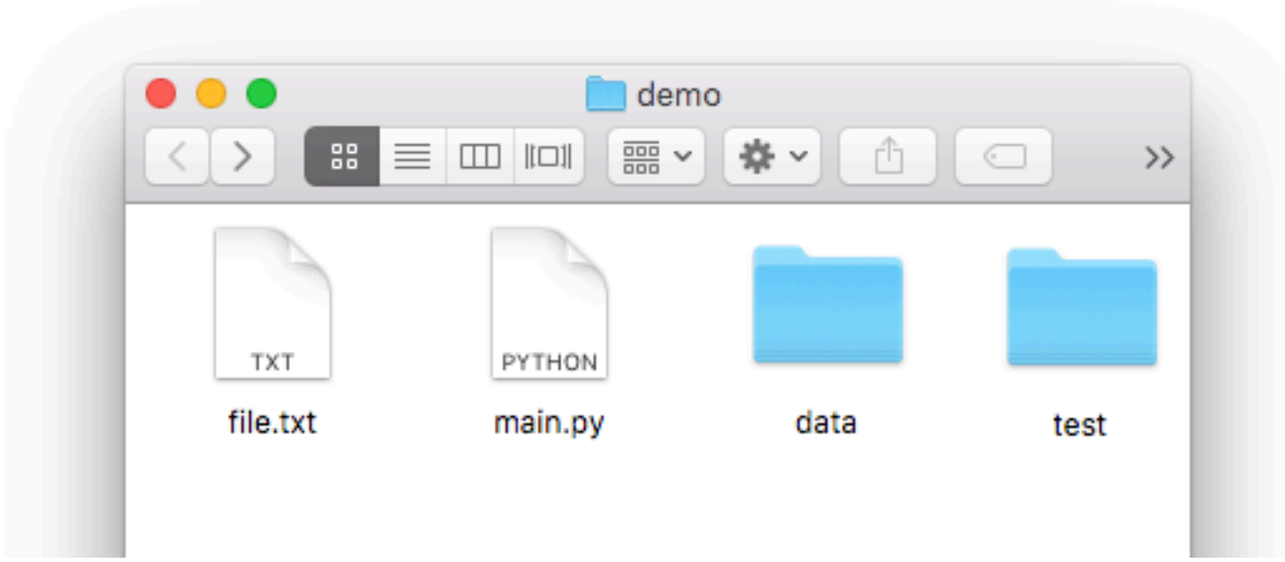



```
>>> import os
>>> os.path.isdir(“data”)
True
```

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- os.path.exists
- os.path.isfile
- os.path.isdir
- **os.path.join**



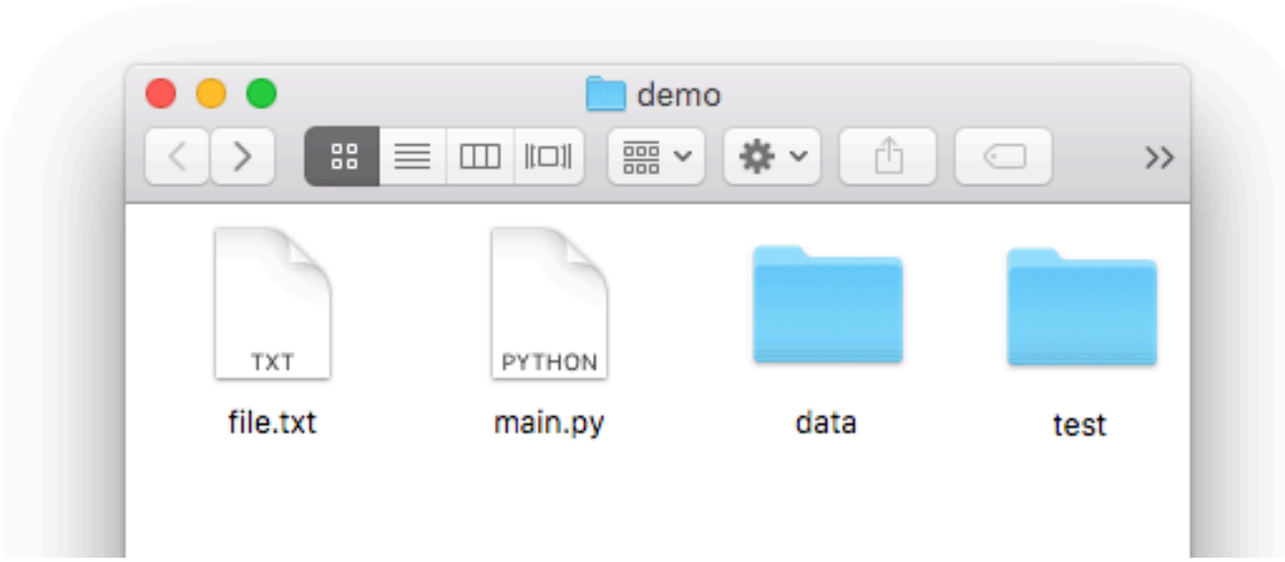

```
>>> import os
>>> os.path.join("data", "movies.csv")
data/movies.csv
```

on Mac/Linux

OS Module (Operating System)

Many functions in os and os.path for working w/ files

- os.listdir
- os.mkdir
- os.path.exists
- os.path.isfile
- os.path.isdir
- **os.path.join**



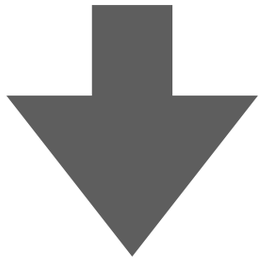

```
>>> import os
>>> os.path.join("data", "movies.csv")
data\movies.csv
```

on Windows

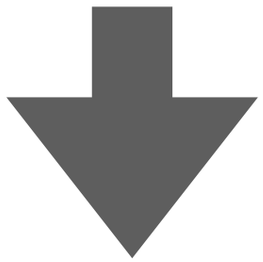
Windows

Your project:

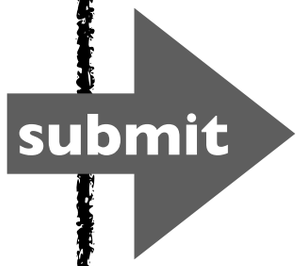
```
path = "\".join("data", "movies.csv")
f = open(path)
...
```



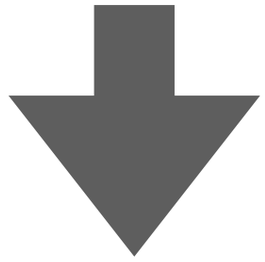
you run test.py



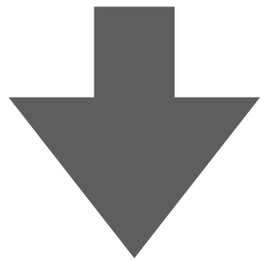
Linux




```
...
```



we run test.py



Windows

Linux

Your project:

```
path = os.path.join("data","movies.csv")
f = open(path)
...
```

submit

```
...
```

you run test.py

we run test.py

Learning Objectives Today

Basic file interactions

- opening/closing
- reading/writing

OS module

- listdir, mkdir, exists, isdir, isfile, join

File exceptions

Encodings

Exceptions

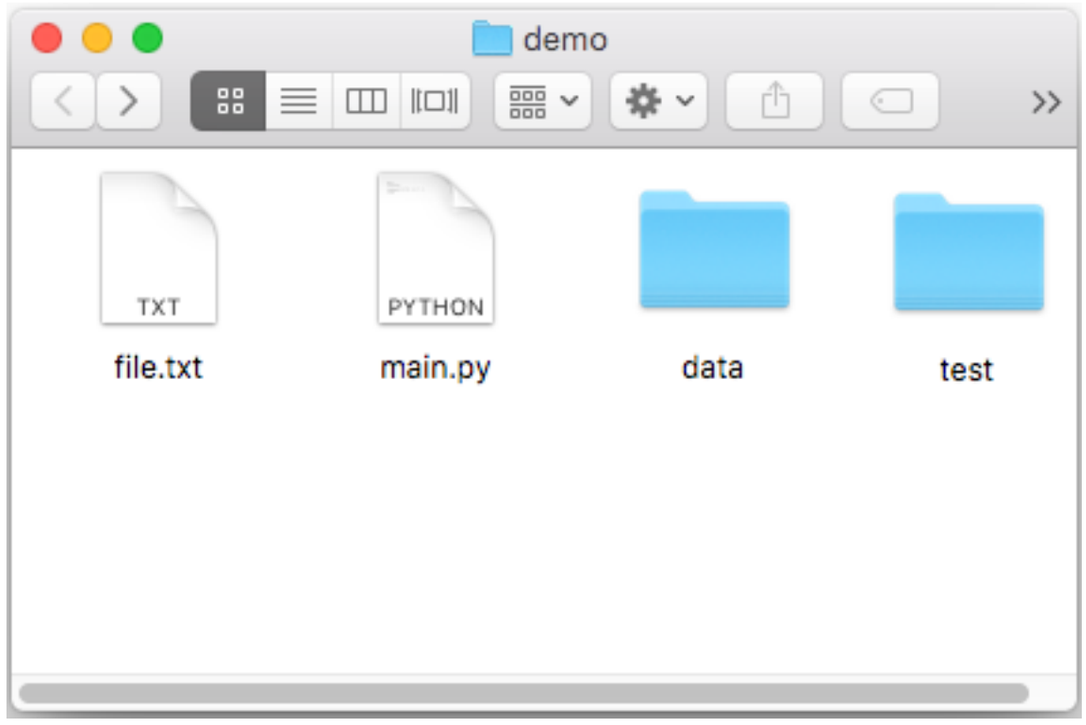
Working with files leads to many exceptions

- missing files
- lacking permissions
- not enough space
- mixing up directories and files
- corrupt formats
- etc, etc

Exceptions

Working with files leads to many exceptions

- missing files
- lacking permissions
- not enough space
- mixing up directories and files
- corrupt formats
- etc, etc

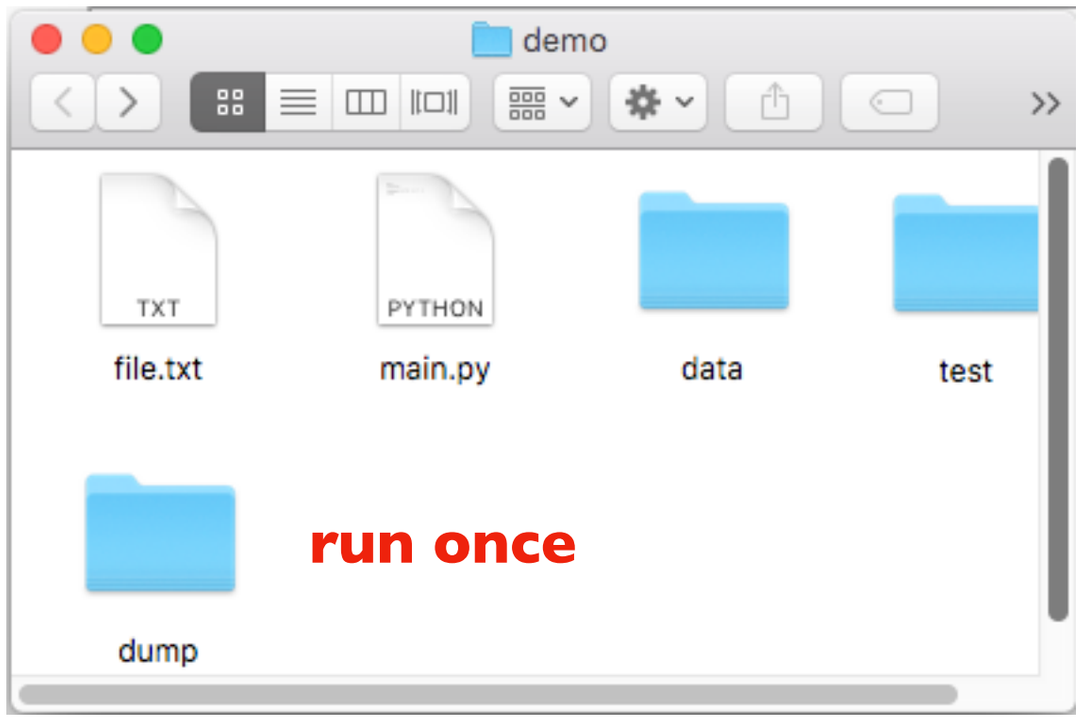



```
import os

os.mkdir('dump')

f = open(os.path.join('dump', 'out.txt'), 'w')
f.write('hi')
f.close()
```

Exceptions

Working with files leads to many exceptions

- missing files
- lacking permissions
- not enough space
- mixing up directories and files
- corrupt formats
- etc, etc

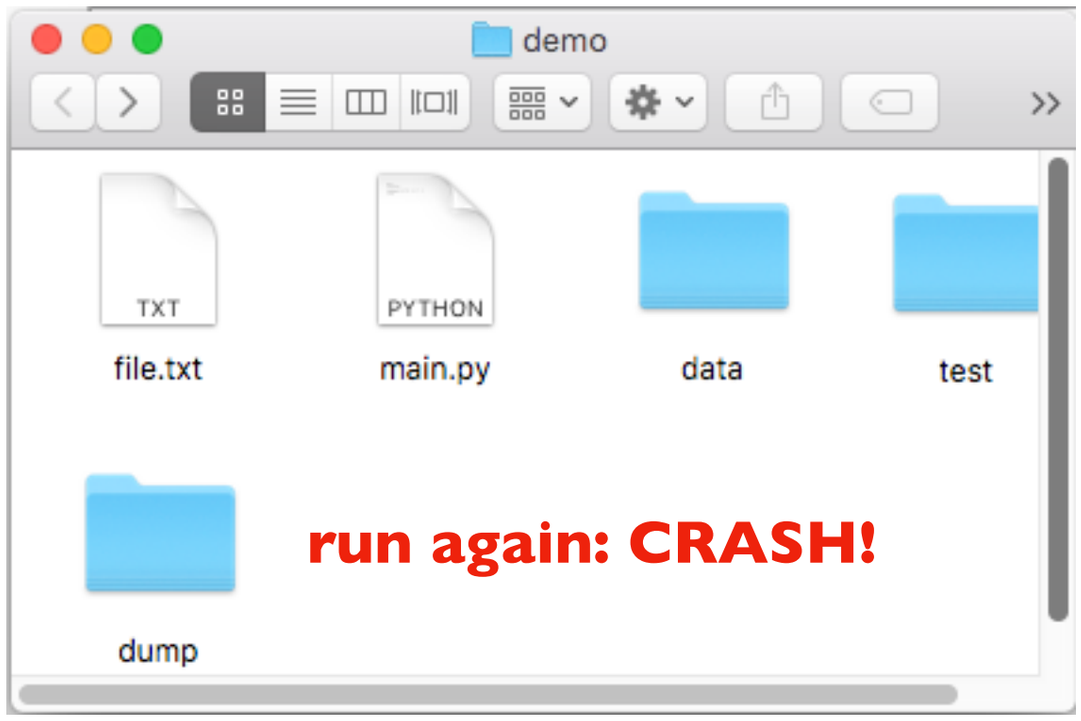



```
import os

os.mkdir('dump')

f = open(os.path.join('dump', 'out.txt'), 'w')
f.write('hi')
f.close()
```

Exceptions

Working with files leads to many exceptions

- missing files
- lacking permissions
- not enough space
- mixing up directories and files
- corrupt formats
- etc, etc



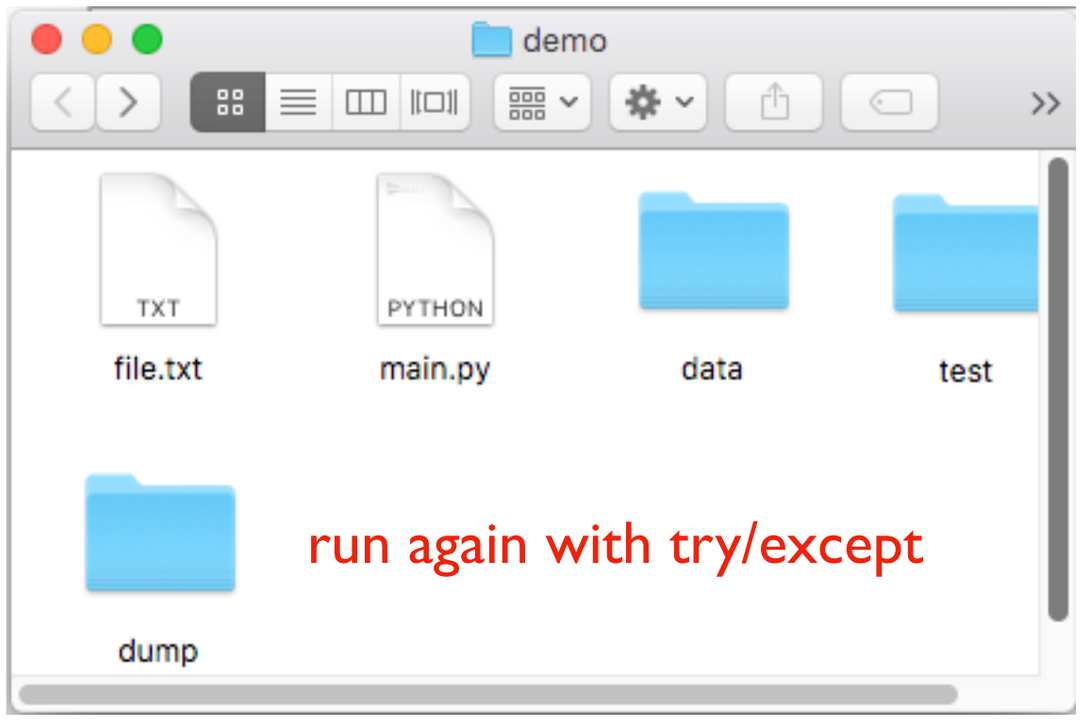

```
import os

os.mkdir('dump')

f = open(os.path.join('dump', 'out.txt'), 'w')
f.write('hi')
f.close()
```

```
Traceback (most recent call last):
  File "test2.py", line 3, in <module>
    os.mkdir('dump')
FileExistsError: [Errno 17] File exists: 'dump'
```

Exceptions

Working with files leads to many exceptions

- missing files
- lacking permissions
- not enough space
- mixing up directories and files
- corrupt formats
- etc, etc




```
import os

try:
    os.mkdir('dump')
except FileExistsError:
    pass # ignore it if dump already existed

f = open(os.path.join('dump', 'out.txt'), 'w')
f.write('hi')
f.close()
```

Learning Objectives Today

Basic file interactions

- opening/closing
- reading/writing

OS module

- listdir, mkdir, exists, isdir, isfile, join

File exceptions

Encodings

Exercise: person 1 encodes a word with encoding 1,
person 2 decodes with encoding 2

Word: ukulele

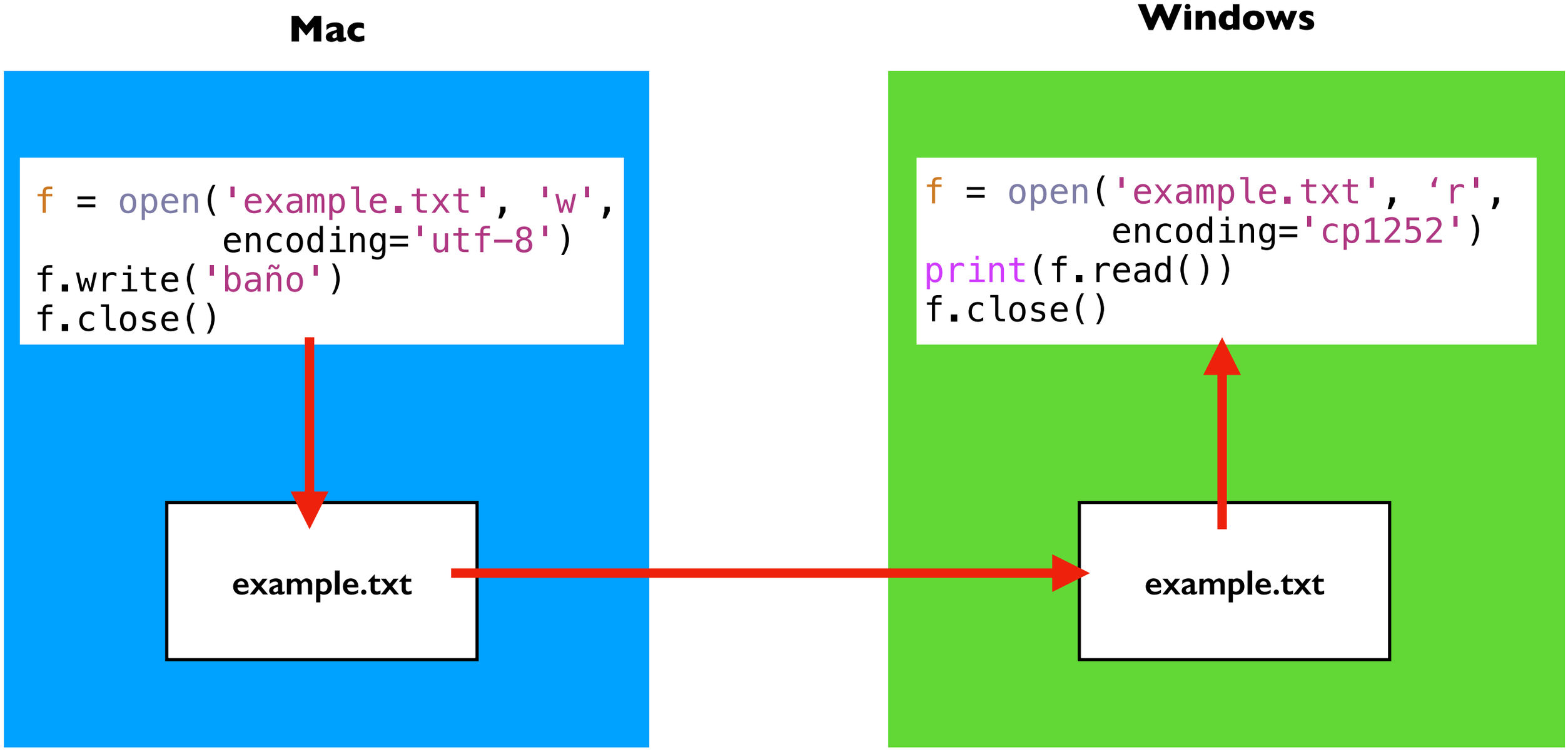
A	00001	N	01110
B	00010	O	01111
C	00011	P	10000
D	00100	Q	10001
E	00000	R	10010
F	00110	S	10011
G	00111	T	10100
H	01000	U	10101
I	01001	V	10110
J	01010	W	10111
K	01011	X	11000
L	11111	Y	11001
M	01101	Z	11010

encoding 1

A	00001	N	01110
B	00010	O	01011
C	00011	P	10000
D	00100	Q	10001
E	00000	R	10010
F	00110	S	10011
G	00111	T	10100
H	01000	U	01100
I	01001	V	10110
J	01010	W	10111
K	01111	X	11000
L	10101	Y	11001
M	01101	Z	11010

encoding 2

Encoding Defaults Done Wrong



Windows computer prints **“baÃ±o”** instead of **“baño”**

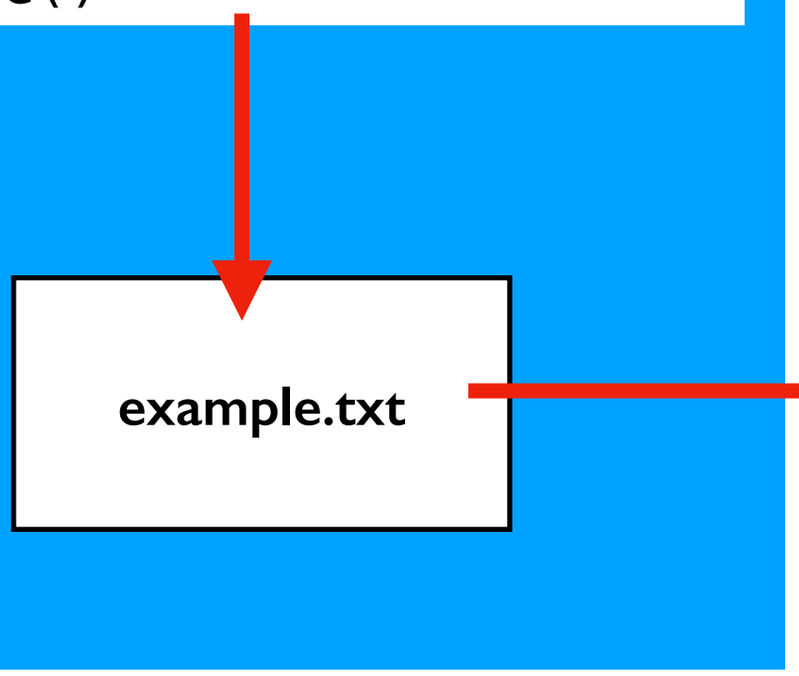
Encoding Defaults Done Wrong

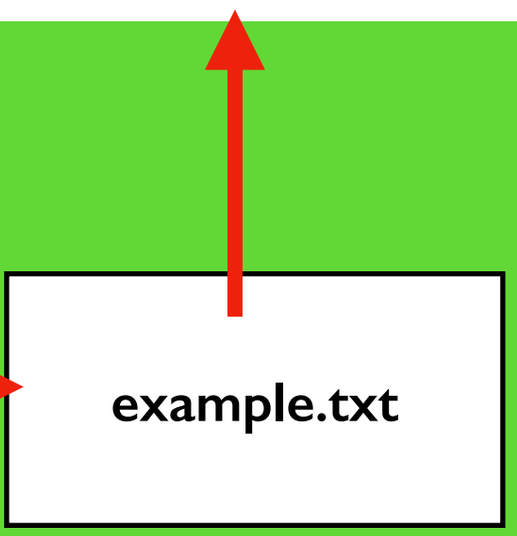
Mac

```
f = open('example.txt', 'w',
         encoding='utf-8')
f.write('baño')
f.close()
```

Windows

```
f = open('example.txt', ‘r',
         encoding='cp1252')
print(f.read())
f.close()
```





Takeaway: if you see weird characters printed by your program, it’s a good time to learn more about encodings

Coding Demos

Demo: How Many Unique Weather Stations are in WI?

Goal: count NOAA weather stations

Input:

- emshr_lite file from here: https://www.ncdc.noaa.gov/homr/reports
- name of a state

Output:

- number of station entries
- number of unique stations

```
NCDC      BEG_DT   END_DT   COOP   WBAN  ICAO FAA     ...
--------- -------- -------- ------ ----- ---- -----   ...
10000001 19490713 19501115 356032 24285               ...
10000001 19590701 19621001        24285               ...
...
```