**Plain Text Files**

eg - Opening of

**Table Data**

eg - Titanic Dataset

filename = ‘huck\_finn.txt’

file = open(filename, mode=’r’) # ‘r’ is to read

file.close()

with open(‘huck\_finn’, ‘r’) as file:

print(file.read())

**Importance of Flat Files**

Basic text files containing records without structured relationships

Record: row of fields or attributes

First row can be header

**How to import flat files?**

Most of the time using NumPy or pandas

**Working with APIs and JSON files**

Connect to omdb API

import requests

url = ‘<http://ww.omdbapi.com/?t=hackers>’

r = requests.get(url)

json\_data = r.json() # using json reader / interpreter

for key, value in json\_data.items():

print(key + ‘:’, value)

What was that URL?

‘http’ - making an HTTP request

‘[www.omdbapi.com](http://www.omdbapi.com)’ - querying the OMDB API

‘?t=hackers’ - Query string

Querying returning data for a movie with title (t) ‘hackers’

**Using Tweepy**

st\_class.py:

class MyStreamListener(tweepy.StreamListener):

def \_\_init\_\_(self, api=None):

super(MyStreamListener, self).\_\_init\_\_()

self.num\_tweets = 0

slef.file = open(‘tweets.txt’, ‘w’)

def on\_status(self, status):

tweet = status.\_json

self.file.write(json.dumps(tweet) + ‘\\n’)

tweet\_list.append(status)

self.num\_tweets += 1

if self.num\_tweets < 100:

return True

else:

return False

self.file.close()

# Create Streaming object and authenticate

l = MyStreamListener()

stream = tweepy.Stream(auth, l)

# This line filters Twitter Streams to capture data by keywords:

stream.filter(track=[‘apples’,’oranges’])