

```

1  # coding: utf-8
2
3  from datetime import datetime, timedelta
4  from urllib.request import urlopen
5  import os
6
7
8  def scrape_station(station):
9      '''
10         This function scrapes the weather data web pages from wunderground.com
11         for the station you provide it.
12
13         You can look up your city's weather station by performing a search for
14         it on wunderground.com then clicking on the "History" section.
15         The 4-letter name of the station will appear on that page.
16         '''
17
18         # Scrape between July 1, 2014 and July 1, 2015
19         # You can change the dates here if you prefer to scrape a different range
20         current_date = datetime(year=2014, month=7, day=1)
21         end_date = datetime(year=2015, month=7, day=1)
22
23         # Make sure a directory exists for the station web pages
24         os.mkdir(station)
25
26         # Use .format(station, YYYY, M, D)
27         lookup_URL =
28         'http://www.wunderground.com/history/airport/{}/{}/{}/DailyHistory.html'
29
30         while current_date != end_date:
31
32             if current_date.day == 1:
33                 print(current_date)
34
35                 formatted_lookup_URL = lookup_URL.format(station,
36                                                         current_date.year,
37                                                         current_date.month,
38                                                         current_date.day)
39                 html = urlopen(formatted_lookup_URL).read().decode('utf-8')
40
41                 out_file_name = '{}/{}-{}-{}.html'.format(station, current_date.year,
42                                                         current_date.month,
43                                                         current_date.day)
44
45                 with open(out_file_name, 'w') as out_file:
46                     out_file.write(html)
47
48                 current_date += timedelta(days=1)
49
50 # Scrape the stations used in this article
51 for station in ['KCLT', 'KCQT', 'KHOU', 'KIND', 'KJAX',

```