

```
In [91]: import requests
import json
import pandas as pd
from bs4 import BeautifulSoup as bs
import csv
import time
from lxml.html import fromstring
```

```
In [2]: #us cities list with coordinates
us=pd.read_csv('uscitiesv1.4.csv')
```

```
In [3]: us.head()
```

Out[3]:

	city	city_ascii	state_id	state_name	county_fips	county_name	lat	
0	Prairie Ridge	Prairie Ridge	WA	Washington	53053	Pierce	47.1443	-122.1
1	Edison	Edison	WA	Washington	53057	Skagit	48.5602	-122.4
2	Packwood	Packwood	WA	Washington	53041	Lewis	46.6085	-121.6
3	Wautauga Beach	Wautauga Beach	WA	Washington	53035	Kitsap	47.5862	-122.5
4	Harper	Harper	WA	Washington	53035	Kitsap	47.5207	-122.5

```
In [78]: #due to the computation power, only run for 500 most populated cities in US
us_test= us.sort_values(['population'], ascending=[False])[0:100]
lat= list(us_test['lat'])
lng= list(us_test['lng'])
```

```
In [79]: us_test.head()
```

Out[79]:

	city	city_ascii	state_id	state_name	county_fips	county_name	lat	
12359	New York	New York	NY	New York	36061	New York	40.6943	-73.9
16410	Los Angeles	Los Angeles	CA	California	6037	Los Angeles	34.1140	-118
20044	Chicago	Chicago	IL	Illinois	17031	Cook	41.8373	-87.6
3640	Miami	Miami	FL	Florida	12086	Miami-Dade	25.7840	-80.2
7504	Dallas	Dallas	TX	Texas	48113	Dallas	32.7938	-96.7

```
In [92]: def get_proxies():
    url = 'https://free-proxy-list.net/'
    response = requests.get(url)
    parser = fromstring(response.text)
    proxies = set()
    for i in parser.xpath('//tbody/tr')[10]:
        if i.xpath('..//td[7][contains(text(),"yes")]'):
            #Grabbing IP and corresponding PORT
            proxy = ":".join([i.xpath('..//td[1]/text()')[0], i.xpath(
'..//td[2]/text()')[0]])
            proxies.add(proxy)
    return proxies
```

```
In [116]: #function to use latitude and longitude to search for theatre locations
         on imax website
         #input: latitude, longitude; return: name, address
def get_data(lat, lng):
    date = '2019-03-09'
    proxies = get_proxies()
    res= requests.get('https://www.imax.com/showtimes/ajax/theatres?date
=' + date + '&lat=' + str(lat) + '&lon=' + str(lng),
                      proxies=proxies)
    soup= bs(res.content, 'lxml')
    newData = json.loads(soup.select_one('p').text)
    columns = ['name','address']
    baseURL = 'https://www.imax.com'
    results = []
    for row in newData['rows']:
        soup = bs(row['row'], 'lxml')
        name = soup.select_one('.theatre-title').text.strip()
        address = soup.select_one('.theatre-address').text.strip()
        results.append([name,address])
    return results
```

```
In [118]: #get data for all the cities
    results=[]
    columns = ['name','address']
    for i in range(len(us_test)):
        results.extend(get_data(lat[i], lng[i]))
        time.sleep(0.5)
    df_100 = pd.DataFrame(results, columns = columns)
```

```
In [121]: #save results to csv
    df_100.to_csv("us100_full.csv", index=False)
```

```
In [28]: df_test2[['address','city','state','other1','other2']] = df_test2.address.str.split(", ",expand=True)
```

```
/Users/ran/anaconda2/envs/py3/lib/python3.6/site-packages/pandas/core/frame.py:2352: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy  
self[k1] = value[k2]
```

```
In [36]: df_test2
```

Out[36]:

	name	address	city	state	other1	other2
43	Edwards Camarillo Palace Stadium 12 & IMAX	680 Ventura Boulevard	Camarillo	CA 93010	None	None
44	Cinemark 22 & IMAX	2600 West Avenue I	Lancaster	CA 93534	None	None
47	Navy Pier IMAX at AMC	700 East Grand Avenue - Suite 115	Chicago	IL 60611	None	None
48	Regal City North Stadium 14 & IMAX	2600 N. Western Ave	Chicago	IL 60647	None	None
49	AMC Quarry 14 & IMAX	9201 W. 63rd Street	Hodgkins	IL 60525	None	None
50	AMC Showplace Village Crossing & IMAX	7000 Carpenter Road	Skokie	IL 60077	None	None
51	AMC Oakbrook 12 & IMAX	600 Oakbrook Center	Oak Brook	Illinois 60523	None	None
52	AMC DINE-IN Yorktown 18 & IMAX	80 Yorktown Shopping Center	Lombard	IL 60148	None	None
53	AMC Showplace Niles 12 & IMAX	301 Golf Mill Center	Niles	IL 60714	None	None
54	Cinemark at Seven Bridges & IMAX	6500 Route 53	Woodridge	IL 60517	None	None
55	AMC Streets of Woodfield & IMAX	601 North Martingale Road	Schaumburg	IL 60173	None	None
56	AMC Showplace Schererville 16 & IMAX	875 Deercreek Drive	Schererville	IN 46375	None	None
57	Regal Lincolnshire Stadium 15 & IMAX	300 Parkway Drive	Lincolnshire	IL 60069	None	None
58	AMC Showplace New Lenox 14 & IMAX	1320 W. Maple Street	New Lenox	IL 60451	None	None
59	AMC Showplace Naperville 16 & IMAX	2815 Showplace Drive	Naperville	IL 60564	None	None
60	AMC South Barrington 30 & IMAX	175 Studio Drive	South Barrington	IL 60010	None	None
61	Portage 16 & IMAX	6550 US Highway 6	Portage	IN 46368	None	None
62	Randall 15 & IMAX	550 North Randall Road	Batavia	IL 60510	None	None

	name	address	city	state	other1	other2
63	Regal South Beach Stadium 18 & IMAX	1120 Lincoln Road	Miami Beach	FL 33139	None	None
64	AMC Sunset Place 24 & IMAX	5701 Sunset Drive	South Miami	FL 33143	None	None
65	Cobb Dolphin 19 & IMAX	11471 NW 12th Street	Miami	FL 33172	None	None
66	AMC Aventura 24 & IMAX	19501 Biscayne Blvd. In Aventura	Aventura	FL 33180	None	None
67	Regal Kendall Village Stadium 16 & IMAX	8595 SW 124 Avenue	Miami	FL 33183	None	None
68	AMC Pembroke Lakes 9 & IMAX	12085 Pines Blvd.	Pembroke Pines	FL 33026	None	None
69	Autonation IMAX, Museum of Discovery & Science	401 SW 2nd Street	Fort Lauderdale	FL 33312	None	None
70	Regal Sawgrass Stadium 23 & IMAX	2600 NW 136th Avenue	Sunrise	FL 33323	None	None
71	Frank Delray MarketPlace Stadium 12 & IMAX	14775 Lyons Road	Delray Beach	FL 33446	None	None
83	AMC Cherry Hill 24 & IMAX	2121 Route 38	Cherry Hill	NJ 08002	None	None
90	Penn Cinema Riverfront & IMAX	401 South Madison Street	Wilmington	DE 19801	None	None
93	Edwards Houston Marq*E Stadium 23 & IMAX	7620 Katy Freeway	Houston	TX 77024	None	None
...
468	AMC Patton Creek 15 & IMAX	4450 Creekside Avenue	Hoover	AL 35244	None	None
469	Cobb Hollywood 16 & IMAX	4250 Old Greensboro Road	Tuscaloosa	AL 35405	None	None
475	Cinemark Tinseltown USA & IMAX	2291 Buffalo Road	Rochester	NY 14624	None	None
476	AMC Sarasota 12 & IMAX	8201 S. Tamiami Trail	Sarasota	FL 34238	None	None
480	Edwards Fresno Stadium 22 & IMAX	250 Paseo del Centro	Fresno	CA 93720	None	None

	name	address	city	state	other1	other2
491	AMC Southroads 20 & IMAX	4923 E 41st Street	Tulsa	OK 74135	None	None
492	Cinemark Tulsa & IMAX	10802 East 71st Street	Tulsa	OK 74133	None	None
508	Cinemark Carefree Circle & IMAX	3305 Cinema Point	Colorado Springs	CO 80922	None	None
510	Regal Gulf Coast Town Center 16 & IMAX	10028 Gulf Center Drive	Fort Myers	FL 33913	None	None
511	Regal Hollywood Stadium 20 & IMAX	6006 Hollywood Blvd.	Naples	FL 34109	None	None
512	Citadel Mall IMAX Stadium 16	2072 Sam Rittenberg Boulevard	Charleston	SC 29407	None	None
531	Regal Crossgates Stadium 18 & IMAX	One Crossgates Mall Road	Albany	NY 12203	None	None
532	Dordoi Plaza & IMAX	115	Ibraimov Str.	Bishkek	Kyrgyzstan` 720021	None
533	Celebration! Cinema Grand Rapids North & IMAX	2121 Celebration Dr. NE	Grand Rapids	MI 49525	None	None
534	AMC Mall of Louisiana 15 & IMAX	9168T Picardy Ave Extension	Baton Rouge	LA 70836	None	None
535	AMC Baton Rouge 16 & IMAX	16040 Hatteras Ave.	Baton Rouge	LA 70816	None	None
536	Regal Pinnacle Stadium 18 & IMAX	11240 Parkside Drive	Knoxville	TN 37922	None	None
542	Regal Sandhill Stadium 16 & IMAX	450 Town Center Place	Columbia	SC 29229	None	None
543	Reading Cinemas Valley Plaza 16 & IMAX	2000 Wible Road	Bakersfield	CA 93304	None	None
581	AMC West Melbourne 12 & IMAX	4255 Norfolk Parkway	West Melbourne	FL 32904	None	None
584	Warren West IMAX	9150 West 21st	Wichita	KS 67205	None	None
585	Blank IMAX Dome, Science Center Iowa	401 West Martin Luther King Jr. Parkway	Des Moines	IA 50309	None	None
586	The Palms Theatre & IMAX	200 NE Westgate Drive	Waukee	Iowa 50263	None	None

	name	address	city	state	other1	other2
595	Regal Rancho Mirage Stadium 16 & IMAX	72-777 Dinah Shore Drive	Rancho Mirage	CA 92270	None	None
642	Regal Simpsonville Stadium 14 & IMAX	Simpsonville	SC 29681	None	None	None
643	AMC Chenal 9 & IMAX	17825 Chenal Parkway	Little Rock	AR 72211	None	None
645	Penn Cinema & IMAX	541 Airport Road	Lititz	PA 17543	None	None
646	New Vision Fitchburg 18 & IMAX	3301 Town Center Blvd.	Fitchburg	WI 53719	None	None
656	Galaxy Luxury+ Sparks & IMAX	1170 Scheels Drive	Sparks	NV 89434	None	None
657	AMC CityPlace 20 & IMAX	545 Hibiscus Street	West Palm Beach	FL 33401	None	None

195 rows × 6 columns

```
In [46]: max(df_test2.groupby('city').city.count())
```

```
Out[46]: 4
```

```
In [58]: #since NY has the largest population, calculated the number of theatres
         in New York City
         #get data for all the cities
lat= list(us_test['lat'])
lng= list(us_test['lng'])
results_ny=get_data(lat[0], lng[0])
df_ny = pd.DataFrame(results_ny, columns = columns)
len(df_ny)
```

```
Out[58]: 24
```