

# DATA\_301\_Assignment\_5A\_ISHAAN\_SATHAYE\_SRESHTA\_TALLURI

November 6, 2023

## 1 CityBikes Rest API

This section asks you to fetch JSON data from the [CityBikes REST API](#) to answer some questions about bike share programs.

Please be mindful not to send requests to the server more often than necessary. In particular, make sure you separate code that requests data from the server into a separate cell so that you do not have to repeatedly request data from the server. You will lose points if your requests are not separated into their own cell.

```
[1]: import pandas as pd
import numpy as np
import requests
import time
```

## 2 Question 1

Find all cities with bike share programs in the United States (country code “US”), along with their network ID (for example, “ford-gobike”). How many cities in the U.S. have bike share programs (at least that are in this API)?

```
[2]: response = requests.get("http://api.citybik.es/v2/networks")

bikes = response.json()
```

```
[3]: df_bikes = pd.json_normalize(bikes['networks'])
df_bikes.head()
```

```
[3]:
```

	company	href
0	[ « »]	/v2/networks/velobike-moscow \
1	[Urban Infrastructure Partner]	/v2/networks/baerum-bysykkel
2	[Comunicare S.r.l.]	/v2/networks/bicincitta-siena
3	[Cyclopolis Systems]	/v2/networks/cyclopolis-maroussi
4	[Cyclopolis Systems]	/v2/networks/cyclopolis-nafplio

  

id	name	location.city	location.country
----	------	---------------	------------------

0	velobike-moscow	Velobike	Moscow	RU	\
1	baerum-bysyssel	Bysyssel	Bærum	NO	
2	bicincitta-siena	Bicincittà	Siena	IT	
3	cyclopolis-maroussi	Cyclopolis	Maroussi	GR	
4	cyclopolis-nafplio	Cyclopolis	Nafplio	GR	

  

	location.latitude	location.longitude	
0	55.750000	37.616667	\
1	59.894550	10.546343	
2	43.318600	11.330600	
3	38.056872	23.808330	
4	37.563940	22.809340	

  

	source	gbfs_href	license.name	
0	NaN	NaN	NaN	\
1	NaN	NaN	NaN	
2	https://www.bicincitta.com/frmLeStazioni.aspx?...	NaN	NaN	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	

  

	license.url	ebikes
0	NaN	NaN
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	NaN	NaN

```
[4]: df_bikes_us = df_bikes[df_bikes["location.country"] == "US"]
len(df_bikes_us)
```

[4]: 40

There are 40 cities in the U.S. with bike share programs.

### 3 Question 2

Construct a `DataFrame` containing data about all bike stations in all networks in the United States. Save this `DataFrame` to disk using `.to_csv()`. (You will need it in Part B of this assignment.)

Use this `DataFrame` to determine the total number of bicycles in bike share programs across the United States. You may assume that the number of bikes at a station is the number of empty spaces, plus the number of available bikes.

```
[5]: df_bikes_us.head()
```

```
[5]:
28                [PBSC, Alta Bicycle Share, Inc] \
81                [BCycle, LLC]
```

```

82             [Motivate International, Inc, PBSC]
85 [Portland Bureau of Transportation (PBOT), Lyf...
87             [BCycle, LLC]

```

	href	id	name
28	/v2/networks/we-cycle	we-cycle	WE-cycle \
81	/v2/networks/austin	austin	Austin B-cycle
82	/v2/networks/bike-chattanooga	bike-chattanooga	Bike Chattanooga
85	/v2/networks/biketown	biketown	BIKETOWN
87	/v2/networks/boulder	boulder	Boulder B-cycle

	location.city	location.country	location.latitude	location.longitude
28	Aspen, CO	US	39.194951	-106.837002 \
81	Austin, TX	US	30.264080	-97.743550
82	Chattanooga, TN	US	35.045630	-85.309680
85	Portland, OR	US	45.521754	-122.681079
87	Boulder, CO	US	40.008110	-105.263850

	source	gbfs_href	license.name
28	NaN	https://asp.publicbikesystem.net/ube/gbfs/v1/g...	NaN \
81	NaN	https://gbfs.bcycle.com/bcycle_austin/gbfs.json	NaN
82	NaN	https://chat.publicbikesystem.net/ube/gbfs/v1/	NaN
85	NaN	https://gbfs.biketownpdx.com/gbfs/gbfs.json	NaN
87	NaN	https://gbfs.bcycle.com/bcycle_boulder/gbfs.json	NaN

	license.url	ebikes
28	NaN	NaN
81	NaN	NaN
82	NaN	True
85	NaN	NaN
87	NaN	NaN

```
[6]: df_bikes_us.to_csv("bikes_us.csv")
```

```
[7]: df_bikes_us.reset_index(inplace=True)
df_bikes_us.head()
```

```
[7]:
   index  company
0     28  [PBSC, Alta Bicycle Share, Inc] \
1     81  [BCycle, LLC]
2     82  [Motivate International, Inc, PBSC]
3     85  [Portland Bureau of Transportation (PBOT), Lyf...
4     87  [BCycle, LLC]
```

	href	id	name
0	/v2/networks/we-cycle	we-cycle	WE-cycle \
1	/v2/networks/austin	austin	Austin B-cycle

```

2 /v2/networks/bike-chattanooga bike-chattanooga Bike Chattanooga
3 /v2/networks/biketown biketown BIKETOWN
4 /v2/networks/boulder boulder Boulder B-cycle

location.city location.country location.latitude location.longitude \
0 Aspen, CO US 39.194951 -106.837002 \
1 Austin, TX US 30.264080 -97.743550
2 Chattanooga, TN US 35.045630 -85.309680
3 Portland, OR US 45.521754 -122.681079
4 Boulder, CO US 40.008110 -105.263850

source gbfs_href license.name
0 NaN https://asp.publicbikesystem.net/ube/gbfs/v1/g... NaN \
1 NaN https://gbfs.bcycle.com/bcycle_austin/gbfs.json NaN
2 NaN https://chat.publicbikesystem.net/ube/gbfs/v1/ NaN
3 NaN https://gbfs.biketownpdx.com/gbfs/gbfs.json NaN
4 NaN https://gbfs.bcycle.com/bcycle_boulder/gbfs.json NaN

license.url ebikes
0 NaN NaN
1 NaN NaN
2 NaN True
3 NaN NaN
4 NaN NaN

```

```

[8]: bike_stations = []
for network in df_bikes_us["id"]:
    url = "http://api.citybik.es/v2/networks/" + network
    response = requests.get(url)
    bike_stations.extend(response.json().get("network").get("stations"))
    time.sleep(0.5)

```

```

[9]: df_bike_stations = pd.json_normalize(bike_stations)
df_bike_stations.head()

```

```

[9]: empty_slots free_bikes id latitude \
0 3 10 2b0452ef99fa25ce9db62a8c9bb917aa 39.364081 \
1 5 2 63cf805d373b399ae1daa391373d7a42 39.387865
2 4 3 2ccb7b9041a6901cc76e92983e61cd9d 39.386449
3 9 2 d4e64b89339be8aafa37203a16957568 39.399257
4 4 3 edd6bdd575f305389a942010c19b3fd0 39.359289

longitude name timestamp
0 -107.034000 Basalt BRT Downvalley 2023-11-07T02:02:59.882000Z \
1 -107.083800 Triangle Park 2023-11-07T02:02:59.794000Z
2 -107.083000 Willits Town Center 2023-11-07T02:02:59.885000Z
3 -107.090126 JW Dr | El Jebel Rd 2023-11-07T02:02:59.892000Z

```

4 -107.023232 Roaring Fork Club Housing 2023-11-07T02:02:59.790000Z

	extra.address	extra.altitude	extra.ebikes	...	extra.returning
0	Basalt	0.0	0.0	...	1 \
1	Basalt	0.0	1.0	...	1
2	Basalt	0.0	0.0	...	1
3	Basalt	0.0	0.0	...	1
4	Basalt	NaN	1.0	...	1

  

	extra.slots	extra.uid	extra.post_code	extra.rental_uris.android
0	13.0	1433	NaN	NaN \
1	7.0	1420	NaN	NaN
2	7.0	1439	NaN	NaN
3	11.0	1474	81623	NaN
4	7.0	1485	81621	NaN

  

	extra.rental_uris.ios	extra.description	extra.last_update	extra.online
0	NaN	NaN	NaN	NaN \
1	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN

  

	extra.photo
0	NaN
1	NaN
2	NaN
3	NaN
4	NaN

[5 rows x 26 columns]

```
[18]: df_bike_stations[df_bike_stations["name"] == 0].head()
```

```
[18]: Index(['empty_slots', 'free_bikes', 'id', 'latitude', 'longitude', 'name',
        'timestamp', 'extra.address', 'extra.altitude', 'extra.ebikes',
        'extra.has_ebikes', 'extra.last_updated', 'extra.normal_bikes',
        'extra.payment', 'extra.payment-terminal', 'extra.renting',
        'extra.returning', 'extra.slots', 'extra.uid', 'extra.post_code',
        'extra.rental_uris.android', 'extra.rental_uris.ios',
        'extra.description', 'extra.last_update', 'extra.online',
        'extra.photo'],
        dtype='object')
```

```
[10]: (df_bike_stations["empty_slots"] + df_bike_stations["free_bikes"]).sum()
```

```
[10]: 141113
```

Currently, there are 141113 bicycles in the share program in the US.

## 4 Question 3

You have just finished touring Coit Tower in San Francisco, which is located at latitude 37.802747 and longitude -122.405861. Using your `DataFrame` from Question 2, find the nearest bike station with an available bike, based on taxicab distance (a.k.a. Manhattan distance).

*Hint:* You can check your answer using Google Maps!

```
[11]: df_bike_stations_available = df_bike_stations[df_bike_stations["free_bikes"] > 0]
      coit_tower = [(37.802139, -122.405853)]
      df_bike_stations_available.shape
```

```
[11]: (6739, 26)
```

```
[12]: from sklearn.metrics.pairwise import manhattan_distances

      df_bike_stations_available["distance_to_coit"] = manhattan_distances(
          df_bike_stations_available[["latitude", "longitude"]], coit_tower
      )
```

```
/var/folders/q8/mqm68gfx7pjfpqftf7y_v6140000gn/T/ipykernel_78394/3544440540.py:3
: 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: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
      df_bike_stations_available["distance_to_coit"] = manhattan_distances(
```

```
[13]: df_bike_stations_available.sort_values("distance_to_coit", inplace=True)
      df_bike_stations_available.head()
```

```
/var/folders/q8/mqm68gfx7pjfpqftf7y_v6140000gn/T/ipykernel_78394/849673195.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
      df_bike_stations_available.sort_values("distance_to_coit", inplace=True)
```

```
[13]:
```

	empty_slots	free_bikes	id	latitude
7869	23	4	c91d4150a09e136d5cca66643c9f936d	37.802055
7752	11	6	a6b2ae23c98c47b86a9e84062ea5c3ee	37.798014
7511	17	5	28754859653e0a656ba53e1fa0ea21db	37.804770
7738	10	13	0c7d51b231f5c05163176b80b5c825bf	37.804648

```
7718          10          19  8744f3ffea155dd0b5e223031ace4513  37.800496
```

	longitude	name	timestamp
7869	-122.401715	Battery St at Filbert St	2023-11-07T02:06:25.123000Z \
7752	-122.405950	Broadway at Kearny St	2023-11-07T02:06:24.778000Z
7511	-122.403234	The Embarcadero at Sansome St	2023-11-07T02:06:25.069000Z
7738	-122.402087	Cruise Terminal at Pier 27	2023-11-07T02:06:25.099000Z
7718	-122.410887	Powell St at Columbus Ave	2023-11-07T02:06:25.078000Z

	extra.address	extra.altitude	extra.ebikes	...	extra.slots
7869	NaN	NaN	2.0	...	31.0 \
7752	NaN	NaN	6.0	...	19.0
7511	NaN	NaN	0.0	...	23.0
7738	NaN	NaN	9.0	...	23.0
7718	NaN	NaN	19.0	...	30.0

	extra.uid	extra.post_code
7869	6732f1f5-616c-46bc-81e6-67c5319b3c73	NaN \
7752	07303af0-cf6e-46f2-bf9c-9ebb61f21adc	NaN
7511	23a2bad4-3ea1-4545-87eb-acbee3efccfc	NaN
7738	d82d9e94-bb62-4001-b6ad-50603199554d	NaN
7718	fb39d594-8fdb-440e-80c5-8a544b339007	NaN

	extra.rental_uris.android
7869	https://sfo.lft.to/lastmile_qr_scan \
7752	https://sfo.lft.to/lastmile_qr_scan
7511	https://sfo.lft.to/lastmile_qr_scan
7738	https://sfo.lft.to/lastmile_qr_scan
7718	https://sfo.lft.to/lastmile_qr_scan

	extra.rental_uris.ios	extra.description	extra.last_update
7869	https://sfo.lft.to/lastmile_qr_scan	NaN	NaN \
7752	https://sfo.lft.to/lastmile_qr_scan	NaN	NaN
7511	https://sfo.lft.to/lastmile_qr_scan	NaN	NaN
7738	https://sfo.lft.to/lastmile_qr_scan	NaN	NaN
7718	https://sfo.lft.to/lastmile_qr_scan	NaN	NaN

	extra.online	extra.photo	distance_to_coit
7869	NaN	NaN	0.004221
7752	NaN	NaN	0.004223
7511	NaN	NaN	0.005250
7738	NaN	NaN	0.006275
7718	NaN	NaN	0.006677

```
[5 rows x 27 columns]
```

```
[14]: df_bike_stations_available.iloc[0]
```

```
[14]: empty_slots                23
      free_bikes                  4
      id                        c91d4150a09e136d5cca66643c9f936d
      latitude                   37.802055
      longitude                  -122.401715
      name                      Battery St at Filbert St
      timestamp                  2023-11-07T02:06:25.123000Z
      extra.address              NaN
      extra.altitude             NaN
      extra.ebikes               2.0
      extra.has_ebikes           True
      extra.last_updated         1699322696.0
      extra.normal_bikes         NaN
      extra.payment              [key, creditcard]
      extra.payment-terminal     True
      extra.renting              1
      extra.returning            1
      extra.slots                31.0
      extra.uid                  6732f1f5-616c-46bc-81e6-67c5319b3c73
      extra.post_code            NaN
      extra.rental_uris.android  https://sfo.lft.to/lastmile_qr_scan
      extra.rental_uris.ios      https://sfo.lft.to/lastmile_qr_scan
      extra.description          NaN
      extra.last_update          NaN
      extra.online               NaN
      extra.photo                NaN
      distance_to_coit           0.004221
      Name: 7869, dtype: object
```

The nearest bike station to Coit Tower is at Battery St at Filbert St with 4 available bikes, which on Apple Maps is 1160 Battery St San Francisco, CA 94111 United States and located at 37.80208° N, 122.40172° W.

## 4.1 Submission Instructions

- After you have completed the notebook, select **Runtime > Run all**
- After the notebook finishes rerunning check to make sure that you have no errors and everything runs properly. Fix any problems and redo this step until it works.
- Rename this notebook by clicking on “DATA 301 Assignment 05 - YOUR NAMES HERE” at the very top of this page. Replace “YOUR NAMES HERE” with the first and last names of you and your partner (if you worked with one).
- Expand all cells with View > Expand Sections
- Save a PDF version: File > Print > Save as PDF
  - Under “More Settings” make sure “Background graphics” is checked
  - Printing Colab to PDF doesn’t always work so well and some of your output might get cutoff. That’s ok.
  - It’s not necessary, but if you want a more nicely formatted PDF you can uncomment and run the code in the following cell. (Here’s a [video](#) with other options.)



- Download the notebook: File > Download .ipynb
- Submit the notebook and PDF in Canvas. If you worked in a pair, only one person should submit in Canvas.

```
[15]: # !wget -nc https://raw.githubusercontent.com/brpy/colab-pdf/master/colab_pdf.py  
# from colab_pdf import colab_pdf  
# colab_pdf('DATA 301 Lab4A - YOUR NAMES HERE.ipynb')
```