Data:

For this project, Hyderabad city neighborhood names, their respective latitude and longitude coordinates are required.

Get Neighborhood names:

After a quick google search, it was found that there is a Wikipedia page that provides the information on neighbourhoods of Hyderabad city. Below is the link for the Wikipedia webpage.

Wikipedia link:
https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Hyderabad

To get data from Wikipedia, let's web-scrape this Wikipedia webpage and get the required data. For the web-scraping, beautiful soup python package is used.

On inspecting the Wikipedia webpage, our required data in in the "<li>" tag.

Python requests library is used to get the data from the URL. Now, beautiful soup object is used to web scrape the data from the webpage. The web-scraped data is stored in a pandas data frame. The web scraping steps are shown below.

Web-Scraping the Wikipedi Page to get neighbourhoods of Hyderabad

```
wikipedia_url = 'https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Hyderabad'
html = requests.get(wikipedia_url)
if html.status_code == 200:
    print('Successfully retrieved response from the url \n')

html = html.text
#print(html)
```

```
Successfully retrieved response from the url
```

Using Beautiful Soup on fetched data

```
soup = BeautifulSoup(html, 'html.parser')
#print(soup.prettify())
```

Extracting the required data from Beautiful soup object

```
scraped_data = []
data = soup.find("div", {"class":"mw-content-ltr"})
hood=data.findAll('li')
#len(hood)
filtered_data = hood[41:285]
for row in filtered_data:
    scraped_data.append(row.a.text)
```

storing the web scraped data into pandas dataframe

```
wiki_data = pd.DataFrame(scraped_data,columns=['Neighbourhood'])
wiki_data.head(10)
```

	Neighbourhood
0	Ameerpet
1	Begumpet
2	SR Nagar
3	Prakash Nagar
4	Punjagutta
5	Balkampet
6	Sanathnagar
7	Bharat Nagar
8	Erragadda
9	Borabanda

Get coordinates for neighborhoods:

We got the neighborhood names of Hyderabad city by web scraping the Wikipedia webpage. The latitude and longitude coordinates of the neighborhoods are not present in the Wikipedia webpage. To get the coordinates, we use the openstreetmap.org website's nominatim API. An API request is to be made to the following URL.

URL:
"https://nominatim.openstreetmap.org/search?q={}&limit=1&format=json"

The response we get contains the coordinates for a requested address in json format. By using this API, coordinates for all the neighborhoods of Hyderabad city are obtained.

```
url = "https://nominatim.openstreetmap.org/search?q={}&limit=1&format=json".format('hyderabad')
result = requests.get(url).text
result
```

'[{"place_id":259328421,"licence":"Data © OpenStreetMap contributors, ODbL 1.0. https://osm.org/copyright","osm_type":"relation","osm_id":7868535,"boundingbox":["17.2916377","17.5608321","78.2387067","78.6223912"],"lat":"17.360589","lon":"78.4740613","display_name":"Hyderabad, Bahadurpura mandal, Hyderabad, Telangana, India","class":"boundary","type":"administrative","importance":0.6836118022682846,"icon":"https://nominatim.openstreetmap.org/ui/mapicons//poi_boundary_administrative.p.20.png"}]'

```
# Storing the data into a DataFrame
hyd_coords = pd.DataFrame(temp, columns=['Latitude', 'Longitude'])
hyd_coords.head()
```

	Latitude	Longitude
0	17.4375012	78.4482505
1	17.4440199	78.4624821
2	17.4452312	78.4449117
3	17.2300647	80.1331686
4	17.426957	78.4523925

The challenge:

By using the nominatim API, coordinates for 14 neighborhoods are not obtained. This is mainly because,

1. For some neighbourhoods, the API doesn't provide coordinates due to unknown reason.
2. Few neighborhood names in the Wikipedia webpage are misspelt. This was found by manually searching for coordinates on google website.

This challenge was overcome by searching coordinates for remaining neighborhoods, on the below mentioned website.

Website: https://www.latlong.net/

Finally, the coordinates for all the neighborhoods were obtained and stored in a data frame.

Now combine the data from two data frames and store it into a new data frame.

The final data frame is as shown below.

```
hyd_data.head()
```

	Neighbourhood	Latitude	Longitude
0	Ameerpet	17.437501	78.448251
1	Begumpet	17.444020	78.462482
2	SR Nagar	17.445231	78.444912
3	Prakash Nagar	17.230065	80.133169
4	Punjagutta	17.426957	78.452393

Geopy library to get coordinates of Hyderabad city:

Geopy is a python that can be used to fetch coordinates of an address. This library was used to get the coordinates of Hyderabad city.

```
from geopy.geocoders import Nominatim # convert an address into latitude and longitude values

address = 'Toronto, Ontario'

geolocator = Nominatim(user_agent="Toronto_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of Toronto are {}, {}.'.format(latitude, longitude))
```

```
The geograpical coordinate of Toronto are 43.6534817, -79.3839347.
```

Neighborhood location data using Foursquare API:

To get the nearby venues of a neighborhood, Foursquare API was used. Foursquare API can provide location data of an address. It provides diverse information about venues, users, photos, check-in's, geo-tagging…etc. This API was used in this project to get near by venue details of a neighborhood. To get data from the API, a search query is to be sent to the foursquare API. The response from the API contains the requested data in json format. The venues data is stored in a data frame.

```
LIMIT = 100
radius = 500
url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    neighborhood_latitude,
    neighborhood_longitude,
    radius,
    LIMIT)
url
```

```
'https://api.foursquare.com/v2/venues/explore?&client_id=PBZRJ1RQMS3C0YMDITUULDNOSW2P2F0I2ECWPRGFDJMVOBBA&client_secret=UU3GBPL
KT25QAY5VO3S4RHPNHS5JUFDOANJHRUJPAAA4ZT3J&v=20180605&ll=43.7532586,-79.3296565&radius=500&limit=100'
```

	name	categories	lat	lng
0	Blue Fox	Indian Restaurant	17.437054	78.445912
1	Kakatiya Deluxe Mess	Diner	17.433435	78.447090
2	Minerva Coffee Shop	Indian Restaurant	17.437295	78.446074
3	Santosh Dhaba	Vegetarian / Vegan Restaurant	17.439442	78.448259
4	Sher-e-Punjab Dhaba	Indian Restaurant	17.438454	78.452262