Madrid-Neighborhood-Clusters

June 24, 2020

[1]:
```
# install geocoder package

import sys
!{sys.executable} -m pip install geocoder
print('package installed')
```

```
Collecting geocoder
  Downloading https://files.pythonhosted.org/packages/4f/6b/13166c909ad2f2
d76b929a4227c952630ebaf0d729f6317eb09cbceccbab/geocoder-1.38.1-py2.py3-none-
any.whl (98kB)
     |                 | 102kB 4.8MB/s ta 0:00:011
Collecting click (from geocoder)
  Downloading https://files.pythonhosted.org/packages/d2/3d/fa76db83bf75c4
f8d338c2fd15c8d33fdd7ad23a9b5e57eb6c5de26b430e/click-7.1.2-py2.py3-none-any.whl
(82kB)
     |                 | 92kB 5.1MB/s eta 0:00:011
Requirement already satisfied: requests in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geocoder)
(2.23.0)
Requirement already satisfied: six in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geocoder)
(1.15.0)
Collecting ratelim (from geocoder)
  Downloading https://files.pythonhosted.org/packages/f2/98/7e6d147fd16a10a5f821
db6e25f192265d6ecca3d82957a4fdd592cad49c/ratelim-0.1.6-py2.py3-none-any.whl
Collecting future (from geocoder)
  Downloading https://files.pythonhosted.org/packages/45/0b/38b06fd9b92dc2
b68d58b75f900e97884c45bedd2ff83203d933cf5851c9/future-0.18.2.tar.gz (829kB)
     |                 | 829kB 10.6MB/s eta 0:00:01
|                            | 204kB 10.6MB/s eta 0:00:01
Requirement already satisfied: certifi>=2017.4.17 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->geocoder) (2020.4.5.2)
Requirement already satisfied: chardet<4,>=3.0.2 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->geocoder) (3.0.4)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
```

```
requests->geocoder) (1.25.9)
Requirement already satisfied: idna<3,>=2.5 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->geocoder) (2.9)
Requirement already satisfied: decorator in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
ratelim->geocoder) (4.4.2)
Building wheels for collected packages: future
  Building wheel for future (setup.py) … done
  Stored in directory: /home/jupyterlab/.cache/pip/wheels/8b/99/a0/81daf51
dcd359a9377b110a8a886b3895921802d2fc1b2397e
Successfully built future
Installing collected packages: click, ratelim, future, geocoder
Successfully installed click-7.1.2 future-0.18.2 geocoder-1.38.1 ratelim-0.1.6
package installed
```

[2]:

```
# install geopy package

!conda install -c conda-forge geopy --yes
print('package installed')
```

```
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /home/jupyterlab/conda/envs/python

  added / updated specs:
    - geopy


The following packages will be downloaded:

    package                    |            build
    ---------------------------|-----------------
    ca-certificates-2020.6.20  |       hecda079_0         145 KB  conda-forge
    certifi-2020.6.20          |   py36h9f0ad1d_0         151 KB  conda-forge
    geographiclib-1.50         |             py_0          34 KB  conda-forge
    geopy-1.22.0               |     pyh9f0ad1d_0          63 KB  conda-forge
    ------------------------------------------------------------
                                           Total:         393 KB

The following NEW packages will be INSTALLED:

  geographiclib      conda-forge/noarch::geographiclib-1.50-py_0
  geopy              conda-forge/noarch::geopy-1.22.0-pyh9f0ad1d_0
```

```
The following packages will be UPDATED:

  ca-certificates                       2020.4.5.2-hecda079_0 -->
2020.6.20-hecda079_0
  certifi                           2020.4.5.2-py36h9f0ad1d_0 -->
2020.6.20-py36h9f0ad1d_0



Downloading and Extracting Packages
certifi-2020.6.20    | 151 KB    | ##################################### | 100%
geopy-1.22.0         | 63 KB     | ##################################### | 100%
ca-certificates-2020 | 145 KB    | ##################################### | 100%
geographiclib-1.50   | 34 KB     | ##################################### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
package installed
```

[3]:
```
# install folium package
!conda install -c conda-forge folium=0.5.0 --yes
print('package installed')
```

```
Collecting package metadata (current_repodata.json): done
Solving environment: failed with initial frozen solve. Retrying with flexible
solve.
Collecting package metadata (repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /home/jupyterlab/conda/envs/python

  added / updated specs:
    - folium=0.5.0


The following packages will be downloaded:

    package                    |            build
    ---------------------------|-----------------
    altair-4.1.0               |             py_1         614 KB  conda-forge
    branca-0.4.1               |             py_0          26 KB  conda-forge
    brotlipy-0.7.0             |py36h8c4c3a4_1000         346 KB  conda-forge
    chardet-3.0.4              |py36h9f0ad1d_1006         188 KB  conda-forge
    cryptography-2.9.2         |   py36h45558ae_0         613 KB  conda-forge
    folium-0.5.0               |             py_0          45 KB  conda-forge
    pandas-1.0.5               |   py36h830a2c2_0        10.1 MB  conda-forge
```

```
    pysocks-1.7.1              |   py36h9f0ad1d_1          27 KB  conda-forge
    requests-2.24.0            |     pyh9f0ad1d_0          47 KB  conda-forge
    toolz-0.10.0               |             py_0          46 KB  conda-forge
    vincent-0.4.4              |             py_1          28 KB  conda-forge
    ------------------------------------------------------------
                                           Total:        12.0 MB

The following NEW packages will be INSTALLED:

  altair             conda-forge/noarch::altair-4.1.0-py_1
  attrs              conda-forge/noarch::attrs-19.3.0-py_0
  branca             conda-forge/noarch::branca-0.4.1-py_0
  brotlipy           conda-forge/linux-64::brotlipy-0.7.0-py36h8c4c3a4_1000
  chardet            conda-forge/linux-64::chardet-3.0.4-py36h9f0ad1d_1006
  cryptography       conda-forge/linux-64::cryptography-2.9.2-py36h45558ae_0
  entrypoints        conda-forge/linux-64::entrypoints-0.3-py36h9f0ad1d_1001
  folium             conda-forge/noarch::folium-0.5.0-py_0
  idna               conda-forge/noarch::idna-2.9-py_1
  importlib_metadata conda-forge/noarch::importlib_metadata-1.6.1-0
  jinja2             conda-forge/noarch::jinja2-2.11.2-pyh9f0ad1d_0
  jsonschema         conda-forge/linux-64::jsonschema-3.2.0-py36h9f0ad1d_1
  markupsafe         conda-forge/linux-64::markupsafe-1.1.1-py36h8c4c3a4_1
  pandas             conda-forge/linux-64::pandas-1.0.5-py36h830a2c2_0
  pyopenssl          conda-forge/noarch::pyopenssl-19.1.0-py_1
  pyrsistent         conda-forge/linux-64::pyrsistent-0.16.0-py36h8c4c3a4_0
  pysocks            conda-forge/linux-64::pysocks-1.7.1-py36h9f0ad1d_1
  pytz               conda-forge/noarch::pytz-2020.1-pyh9f0ad1d_0
  requests           conda-forge/noarch::requests-2.24.0-pyh9f0ad1d_0
  toolz              conda-forge/noarch::toolz-0.10.0-py_0
  urllib3            conda-forge/noarch::urllib3-1.25.9-py_0
  vincent            conda-forge/noarch::vincent-0.4.4-py_1



Downloading and Extracting Packages
pysocks-1.7.1        | 27 KB     | ##################################### | 100%
toolz-0.10.0         | 46 KB     | ##################################### | 100%
chardet-3.0.4        | 188 KB    | ##################################### | 100%
folium-0.5.0         | 45 KB     | ##################################### | 100%
branca-0.4.1         | 26 KB     | ##################################### | 100%
cryptography-2.9.2   | 613 KB    | ##################################### | 100%
brotlipy-0.7.0       | 346 KB    | ##################################### | 100%
altair-4.1.0         | 614 KB    | ##################################### | 100%
requests-2.24.0      | 47 KB     | ##################################### | 100%
pandas-1.0.5         | 10.1 MB   | ##################################### | 100%
vincent-0.4.4        | 28 KB     | ##################################### | 100%
Preparing transaction: done
Verifying transaction: done
```

```
Executing transaction: done
package installed
```

[4]:
```
#import libraries
import pandas as pd  # library for data analsysis
import requests  # library to handle requests
import geocoder
import json  # library to handle JSON files
from pandas.io.json import json_normalize  # tranform JSON file into a pandas␣
↪dataframe
import numpy as np  # library to handle data in a vectorized manner
import matplotlib.cm as cm  # Matplotlib and associated plotting modules
import matplotlib.colors as colors
from geopy.geocoders import Nominatim  # convert an address into latitude and␣
↪longitude values
from sklearn.cluster import KMeans  # import k-means from clustering stage
import folium  # map rendering library
```

[5]:
```
# Get neighborhood data: 'cod_distrito', 'distrito', 'cod_barrio', 'barrio'

neighbors = pd.read_csv('madrid-distritos-barrios.csv')

#clean dataframe and drop duplicates
neighbors = neighbors.drop(neighbors.columns[[4,5]], axis=1)
neighbors = neighbors.drop_duplicates(subset = ["distrito"])
neighbors = neighbors.astype(str)
neighbors['Code'] = neighbors['cod_distrito'] + neighbors['cod_barrio']
neighbors = neighbors.drop(['cod_distrito', 'cod_barrio'], axis=1)
neighbors.rename(columns={'distrito':'Borough', 'barrio':'Neighborhood'},␣
↪inplace=True)
neighbors = neighbors[['Code', 'Borough', 'Neighborhood']]
neighbors['Latitude'] = 0.0
neighbors['Longitude'] = 0.0

neighbors
```

[5]:
```
     Code             Borough                  Neighborhood  Latitude  \
0     111              Centro                       Palacio       0.0
6     221          Arganzuela                      Imperial       0.0
13    331              Retiro                      Pacífico       0.0
19    441           Salamanca                     Recoletos       0.0
25    551           Chamartín                       El Viso       0.0
31    661              Tetuán                 Bellas Vistas       0.0
37    771            Chamberí                    Gaztambide       0.0
43    881  Fuencarral-El Pardo                      El Pardo       0.0
51    991     Moncloa-Aravaca                 Casa de Campo       0.0
58  10101              Latina                   Los Cármenes       0.0
```

```
65   11111          Carabanchel                                Comillas       0.0
72   12121                Usera                               Orcasitas       0.0
79   13131    Puente de Vallecas                              Entrevías       0.0
85   14141            Moratalaz                                 Pavones       0.0
91   15151        Ciudad Lineal                                  Ventas       0.0
100  16161            Hortaleza                                 Palomas       0.0
106  17171           Villaverde  Villaverde Alto, C.H. Villaverde       0.0
111  18181     Villa de Vallecas       Casco Histórico de Vallecas       0.0
118  20201   San Blas-Canillejas                               Simancas       0.0
126  21211              Barajas                       Alameda de Osuna       0.0

     Longitude
0          0.0
6          0.0
13         0.0
19         0.0
25         0.0
31         0.0
37         0.0
43         0.0
51         0.0
58         0.0
65         0.0
72         0.0
79         0.0
85         0.0
91         0.0
100        0.0
106        0.0
111        0.0
118        0.0
126        0.0
```

[6]:

```
# call for arcgis to assess the coordinates for each Neighborhood
for i in range(0, len(neighbors)):
        dir = neighbors.iloc[i,2] + ', ' + neighbors.iloc[i,1] + ', madrid,␣
 ↪españa'  # "Neighborhood, Borough, 'madrid', 'españa'"
        g = geocoder.arcgis(dir)
        neighbors.iloc[i,3] = g.latlng[0]
        neighbors.iloc[i,4] = g.latlng[1]

neighbors
```

[6]:

```
      Code              Borough                       Neighborhood   Latitude  \
0      111               Centro                            Palacio  40.409630
6      221           Arganzuela                           Imperial  40.400210
13     331               Retiro                           Pacífico  40.401910
```

```
19     441            Salamanca                           Recoletos  40.429720
25     551            Chamartín                             El Viso  40.450000
31     661               Tetuán                       Bellas Vistas  40.454570
37     771             Chamberí                          Gaztambide  40.434900
43     881  Fuencarral-El Pardo                            El Pardo  40.511644
51     991      Moncloa-Aravaca                       Casa de Campo  40.435470
58   10101               Latina                        Los Cármenes  40.402800
65   11111          Carabanchel                            Comillas  40.390940
72   12121                Usera                           Orcasitas  40.369850
79   13131   Puente de Vallecas                           Entrevías  40.367835
85   14141           Moratalaz                              Pavones  40.416670
91   15151        Ciudad Lineal                              Ventas  40.453490
100  16161            Hortaleza                             Palomas  40.474440
106  17171           Villaverde  Villaverde Alto, C.H. Villaverde  40.350000
111  18181     Villa de Vallecas       Casco Histórico de Vallecas  40.362510
118  20201  San Blas-Canillejas                            Simancas  40.443730
126  21211              Barajas                     Alameda de Osuna  40.458180

     Longitude
0    -3.879790
6    -3.696180
13   -3.676030
19   -3.679750
25   -3.700000
31   -3.705520
37   -3.715510
43   -3.754898
51   -3.731700
58   -3.731780
65   -3.724200
72   -3.712310
79   -3.664759
85   -3.650000
91   -3.654340
100  -3.641100
106  -3.700000
111  -3.617100
118  -3.609770
126  -3.589530
```

[7]:
```
# Foursquare credentials
CLIENT_ID = 'XZQ0BJKCXSL1CH4MYIYXHE20YQYNEM1WFS4ZEBVPSKDQH4J5'
CLIENT_SECRET = '2JM1AYFVJ5XRXWRS4FCPP0RPB1XCINKSHUQGKBOI0P0A5KCO'
VERSION = '20180604'
```

[8]:
```
# define the venues dataframe
venues_columns = ['Neighborhood',
```

```
                    'Neighborhood Latitude',
                    'Neighborhood Longitude',
                    'Venue',
                    'Venue Latitude',
                    'Venue Longitude',
                    'Venue Category']

madrid_venues = pd.DataFrame(columns=venues_columns)
```

[9]:

```
# function finding venues for a given location
def getNearbyVenues(names, latitudes, longitudes):
    radius=500
    LIMIT=100
    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?
↪&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item␣
↪in venue_list])
    nearby_venues.columns = ['Neighborhood',
                    'Neighborhood Latitude',
                    'Neighborhood Longitude',
                    'Venue',
```

```
                'Venue Latitude',
                'Venue Longitude',
                'Venue Category']

    return(nearby_venues)
```

[11]:
```
# call the function
madrid_venues = getNearbyVenues(names=neighbors['Neighborhood'],
                                   latitudes=neighbors['Latitude'],
                                   longitudes=neighbors['Longitude'])

madrid_venues
```

```
Palacio
Imperial
Pacífico
Recoletos
El Viso
Bellas Vistas
Gaztambide
El Pardo
Casa de Campo
Los Cármenes
Comillas
Orcasitas
Entrevías
Pavones
Ventas
Palomas
Villaverde Alto, C.H. Villaverde
Casco Histórico de Vallecas
Simancas
Alameda de Osuna
```

[11]:
```
         Neighborhood  Neighborhood Latitude  Neighborhood Longitude  \
0             Palacio               40.40963                -3.87979
1             Palacio               40.40963                -3.87979
2             Palacio               40.40963                -3.87979
3             Palacio               40.40963                -3.87979
4             Palacio               40.40963                -3.87979
..                ...                    ...                     ...
482  Alameda de Osuna               40.45818                -3.58953
483  Alameda de Osuna               40.45818                -3.58953
484  Alameda de Osuna               40.45818                -3.58953
485  Alameda de Osuna               40.45818                -3.58953
486  Alameda de Osuna               40.45818                -3.58953
```

```
                         Venue  Venue Latitude  Venue Longitude  \
0                   Proverbium       40.408192        -3.877232
1              The London Walk       40.408197        -3.880682
2                   Go!Sushing       40.408424        -3.880492
3                VIPS Boadilla       40.408106        -3.880623
4          El Rincón Del Bierzo      40.409606        -3.880125
..                         …             …                …
482                      Dia %       40.455036        -3.586630
483            El Kiosko de Pepe     40.454098        -3.589498
484  Hiper Bazar Padre Nuestro       40.455040        -3.585681
485  Gin Terrace Hilton Madrid       40.454964        -3.585670
486              Plaza Del Navio     40.454987        -3.585642

              Venue Category
0        Italian Restaurant
1                 Irish Pub
2       Japanese Restaurant
3              Burger Joint
4    Mediterranean Restaurant
..                        …
482           Grocery Store
483               Bookstore
484          Shop & Service
485             Cocktail Bar
486                   Plaza

[487 rows x 7 columns]
```

[12]:
```python
madrid_venues.groupby('Neighborhood').count()
```

[12]:
```
                                  Neighborhood Latitude  \
Neighborhood
Alameda de Osuna                                     23
Bellas Vistas                                        48
Casa de Campo                                         5
Casco Histórico de Vallecas                           3
Comillas                                             11
El Viso                                              56
Gaztambide                                          100
Imperial                                             51
Los Cármenes                                          3
Orcasitas                                             5
Pacífico                                             52
Palacio                                              12
Palomas                                              19
Pavones                                               3
Recoletos                                            45
```

Simancas	11
Ventas	35
Villaverde Alto, C.H. Villaverde	5

	Neighborhood Longitude	Venue \
Neighborhood		
Alameda de Osuna	23	23
Bellas Vistas	48	48
Casa de Campo	5	5
Casco Histórico de Vallecas	3	3
Comillas	11	11
El Viso	56	56
Gaztambide	100	100
Imperial	51	51
Los Cármenes	3	3
Orcasitas	5	5
Pacífico	52	52
Palacio	12	12
Palomas	19	19
Pavones	3	3
Recoletos	45	45
Simancas	11	11
Ventas	35	35
Villaverde Alto, C.H. Villaverde	5	5

	Venue Latitude	Venue Longitude \
Neighborhood		
Alameda de Osuna	23	23
Bellas Vistas	48	48
Casa de Campo	5	5
Casco Histórico de Vallecas	3	3
Comillas	11	11
El Viso	56	56
Gaztambide	100	100
Imperial	51	51
Los Cármenes	3	3
Orcasitas	5	5
Pacífico	52	52
Palacio	12	12
Palomas	19	19
Pavones	3	3
Recoletos	45	45
Simancas	11	11
Ventas	35	35
Villaverde Alto, C.H. Villaverde	5	5

Venue Category

```
Neighborhood
Alameda de Osuna                           23
Bellas Vistas                              48
Casa de Campo                               5
Casco Histórico de Vallecas                 3
Comillas                                   11
El Viso                                    56
Gaztambide                                100
Imperial                                   51
Los Cármenes                                3
Orcasitas                                   5
Pacífico                                   52
Palacio                                    12
Palomas                                    19
Pavones                                     3
Recoletos                                  45
Simancas                                   11
Ventas                                     35
Villaverde Alto, C.H. Villaverde            5
```

[13]:
```
madrid_venues.head()
```

[13]:
```
  Neighborhood  Neighborhood Latitude  Neighborhood Longitude  \
0      Palacio              40.40963                -3.87979
1      Palacio              40.40963                -3.87979
2      Palacio              40.40963                -3.87979
3      Palacio              40.40963                -3.87979
4      Palacio              40.40963                -3.87979

                  Venue  Venue Latitude  Venue Longitude  \
0            Proverbium       40.408192        -3.877232
1       The London Walk       40.408197        -3.880682
2            Go!Sushing       40.408424        -3.880492
3         VIPS Boadilla       40.408106        -3.880623
4  El Rincón Del Bierzo       40.409606        -3.880125

             Venue Category
0        Italian Restaurant
1                 Irish Pub
2       Japanese Restaurant
3              Burger Joint
4  Mediterranean Restaurant
```

[14]:
```
madrid_venues.shape
```

[14]:
```
(487, 7)
```

[15]:
```
madrid_onehot = pd.get_dummies(madrid_venues[['Venue Category']], prefix="",
 ↪prefix_sep="")
madrid_onehot
```

[15]:
```
     American Restaurant  Arepa Restaurant  Argentinian Restaurant  \
0                      0                 0                       0
1                      0                 0                       0
2                      0                 0                       0
3                      0                 0                       0
4                      0                 0                       0
..                   …                 …                       …
482                    0                 0                       0
483                    0                 0                       0
484                    0                 0                       0
485                    0                 0                       0
486                    0                 0                       0

     Art Gallery  Asian Restaurant  Athletics & Sports  BBQ Joint  Bakery  \
0              0                 0                   0          0       0
1              0                 0                   0          0       0
2              0                 0                   0          0       0
3              0                 0                   0          0       0
4              0                 0                   0          0       0
..           …                 …                   …        …     …
482            0                 0                   0          0       0
483            0                 0                   0          0       0
484            0                 0                   0          0       0
485            0                 0                   0          0       0
486            0                 0                   0          0       0

     Bar  Beer Bar  …  Tea Room  Thai Restaurant  Theater  Toy / Game Store  \
0      0         0  …         0                0        0                 0
1      0         0  …         0                0        0                 0
2      0         0  …         0                0        0                 0
3      0         0  …         0                0        0                 0
4      0         0  …         0                0        0                 0
..   …         …  …       …              …      …               …
482    0         0  …         0                0        0                 0
483    0         0  …         0                0        0                 0
484    0         0  …         0                0        0                 0
485    0         0  …         0                0        0                 0
486    0         0  …         0                0        0                 0

     Trade School  Train Station  Travel Agency  Wine Shop  Women's Store  \
0               0              0              0          0              0
1               0              0              0          0              0
2               0              0              0          0              0
```

```
3              0            0            0          0            0
4              0            0            0          0            0
..           ...          ...          ...        ...          ...
482            0            0            0          0            0
483            0            0            0          0            0
484            0            0            0          0            0
485            0            0            0          0            0
486            0            0            0          0            0

     Yoga Studio
0              0
1              0
2              0
3              0
4              0
..           ...
482            0
483            0
484            0
485            0
486            0

[487 rows x 123 columns]
```

[16]:

```python
madrid_onehot.insert(loc=0, column='Neighborhood',␣
 ↪value=madrid_venues['Neighborhood'] )
madrid_onehot.shape
```

[16]: (487, 124)

[17]:

```python
madrid_onehot
```

[17]:

```
         Neighborhood  American Restaurant  Arepa Restaurant  \
0             Palacio                    0                 0
1             Palacio                    0                 0
2             Palacio                    0                 0
3             Palacio                    0                 0
4             Palacio                    0                 0
..                ...                  ...               ...
482  Alameda de Osuna                    0                 0
483  Alameda de Osuna                    0                 0
484  Alameda de Osuna                    0                 0
485  Alameda de Osuna                    0                 0
486  Alameda de Osuna                    0                 0

     Argentinian Restaurant  Art Gallery  Asian Restaurant  \
0                         0            0                 0
```

```
1                        0            0                 0
2                        0            0                 0
3                        0            0                 0
4                        0            0                 0
..                     …          …                …
482                      0            0                 0
483                      0            0                 0
484                      0            0                 0
485                      0            0                 0
486                      0            0                 0

     Athletics & Sports  BBQ Joint  Bakery  Bar  …  Tea Room  \
0                     0          0       0    0  …         0
1                     0          0       0    0  …         0
2                     0          0       0    0  …         0
3                     0          0       0    0  …         0
4                     0          0       0    0  …         0
..                  …        …      …   …  …       …
482                   0          0       0    0  …         0
483                   0          0       0    0  …         0
484                   0          0       0    0  …         0
485                   0          0       0    0  …         0
486                   0          0       0    0  …         0

     Thai Restaurant  Theater  Toy / Game Store  Trade School  Train Station  \
0                  0        0                 0             0              0
1                  0        0                 0             0              0
2                  0        0                 0             0              0
3                  0        0                 0             0              0
4                  0        0                 0             0              0
..               …      …               …           …            …
482                0        0                 0             0              0
483                0        0                 0             0              0
484                0        0                 0             0              0
485                0        0                 0             0              0
486                0        0                 0             0              0

     Travel Agency  Wine Shop  Women's Store  Yoga Studio
0                0          0              0            0
1                0          0              0            0
2                0          0              0            0
3                0          0              0            0
4                0          0              0            0
..             …        …            …          …
482              0          0              0            0
483              0          0              0            0
484              0          0              0            0
```

```
485                0          0               0           0
486                0          0               0           0

[487 rows x 124 columns]
```

[18]:
```python
madrid_grouped = madrid_onehot.groupby('Neighborhood').mean().reset_index()
madrid_grouped.head()
```

[18]:
```
                 Neighborhood  American Restaurant  Arepa Restaurant  \
0             Alameda de Osuna                  0.0               0.0
1                Bellas Vistas                  0.0               0.0
2                Casa de Campo                  0.0               0.0
3  Casco Histórico de Vallecas                  0.0               0.0
4                     Comillas                  0.0               0.0

   Argentinian Restaurant  Art Gallery  Asian Restaurant  Athletics & Sports  \
0                     0.0          0.0               0.0                 0.0
1                     0.0          0.0               0.0                 0.0
2                     0.0          0.0               0.0                 0.0
3                     0.0          0.0               0.0                 0.0
4                     0.0          0.0               0.0                 0.0

   BBQ Joint    Bakery       Bar  …  Tea Room  Thai Restaurant  Theater  \
0        0.0  0.043478  0.043478  …       0.0              0.0      0.0
1        0.0  0.062500  0.083333  …       0.0              0.0      0.0
2        0.0  0.000000  0.000000  …       0.0              0.0      0.0
3        0.0  0.333333  0.000000  …       0.0              0.0      0.0
4        0.0  0.000000  0.000000  …       0.0              0.0      0.0

   Toy / Game Store  Trade School  Train Station  Travel Agency  Wine Shop  \
0          0.000000           0.0            0.0       0.043478        0.0
1          0.020833           0.0            0.0       0.000000        0.0
2          0.000000           0.0            0.0       0.000000        0.0
3          0.000000           0.0            0.0       0.000000        0.0
4          0.000000           0.0            0.0       0.000000        0.0

   Women's Store  Yoga Studio
0            0.0          0.0
1            0.0          0.0
2            0.0          0.0
3            0.0          0.0
4            0.0          0.0

[5 rows x 124 columns]
```

[19]:
```python
# function returning categories for each neighborhood
def return_most_common_venues(row, num_top_venues):
```

```
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_venues]
```

[20]:
```
# get top ten common venues for each Neighborhood
num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['Neighborhood']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = madrid_grouped['Neighborhood']

for ind in np.arange(madrid_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] =␣
↪return_most_common_venues(madrid_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

[20]:
```
                Neighborhood 1st Most Common Venue 2nd Most Common Venue  \
0             Alameda de Osuna            Restaurant      Tapas Restaurant
1                Bellas Vistas    Spanish Restaurant                   Bar
2                Casa de Campo                   Pub               Stadium
3  Casco Histórico de Vallecas           Pizza Place        Scenic Lookout
4                     Comillas           Coffee Shop  Fast Food Restaurant

  3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue  \
0                 Plaza             Bookstore            Hobby Shop
1         Grocery Store                Bakery           Supermarket
2    Spanish Restaurant                  Pool                   Gym
3                Bakery           Yoga Studio                  Food
4         Grocery Store           Supermarket           Flea Market

  6th Most Common Venue 7th Most Common Venue 8th Most Common Venue  \
0    Fried Chicken Joint    Italian Restaurant         Metro Station
1      Tapas Restaurant           Pizza Place           Coffee Shop
2         Grocery Store  Fast Food Restaurant     Convenience Store
3      Department Store          Dessert Shop                 Diner
```

```
4                Plaza        Farmers Market   Dumpling Restaurant

  9th Most Common Venue 10th Most Common Venue
0           Cocktail Bar            Pizza Place
1      Seafood Restaurant         Farmers Market
2        Cuban Restaurant          Deli / Bodega
3          Discount Store             Donut Shop
4      Falafel Restaurant      Electronics Store
```

[21]:
```
# Clustering for Neighborhood

# set number of clusters
kclusters = 5

madrid_grouped_clustering = madrid_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).
 ↪fit(madrid_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

[21]:
```
array([0, 0, 0, 1, 0, 0, 0, 0, 2, 4], dtype=int32)
```

[22]:
```
# add clustering labels
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

madrid_merged = neighbors
```

[23]:
```
# merge madrid_grouped with madrid_data to add latitude/longitude for each␣
 ↪neighborhood
madrid_merged = madrid_merged.join(neighborhoods_venues_sorted.
 ↪set_index('Neighborhood'), on='Neighborhood')

madrid_merged.head()
```

[23]:
```
   Code      Borough Neighborhood  Latitude  Longitude  Cluster Labels  \
0   111       Centro      Palacio  40.40963   -3.87979             0.0
6   221  Arganzuela     Imperial  40.40021   -3.69618             0.0
13  331       Retiro     Pacífico  40.40191   -3.67603             0.0
19  441    Salamanca    Recoletos  40.42972   -3.67975             0.0
25  551    Chamartín      El Viso  40.45000   -3.70000             0.0

   1st Most Common Venue 2nd Most Common Venue 3rd Most Common Venue  \
0     Spanish Restaurant   Japanese Restaurant                 Plaza
6       Tapas Restaurant    Spanish Restaurant                Bakery
```

```
13    Spanish Restaurant                Bakery         Grocery Store
19    Spanish Restaurant            Restaurant              Boutique
25    Spanish Restaurant           Pizza Place                Bakery

   4th Most Common Venue     5th Most Common Venue 6th Most Common Venue  \
0   Gym / Fitness Center                      Pool            Restaurant
6             Restaurant             Grocery Store    Chinese Restaurant
13                   Bar               Pizza Place            Restaurant
19      Tapas Restaurant  Mediterranean Restaurant           Coffee Shop
25   Japanese Restaurant                Restaurant      Tapas Restaurant

   7th Most Common Venue 8th Most Common Venue     9th Most Common Venue  \
0          Shopping Mall          Burger Joint  Mediterranean Restaurant
6     Falafel Restaurant      Arepa Restaurant      Gym / Fitness Center
13      Tapas Restaurant      Food & Drink Shop     Athletics & Sports
19    Mexican Restaurant                Bakery                    Lounge
25       Thai Restaurant                 Plaza                      Park

   10th Most Common Venue
0        Department Store
6             Beer Garden
13       Asian Restaurant
19                    Spa
25            Coffee Shop
```

[24]:
```
neighborhoods_venues_sorted.head()
```

[24]:
```
   Cluster Labels                 Neighborhood 1st Most Common Venue  \
0               0              Alameda de Osuna            Restaurant
1               0                 Bellas Vistas    Spanish Restaurant
2               0                 Casa de Campo                   Pub
3               1  Casco Histórico de Vallecas           Pizza Place
4               0                      Comillas           Coffee Shop

  2nd Most Common Venue 3rd Most Common Venue 4th Most Common Venue  \
0      Tapas Restaurant                 Plaza             Bookstore
1                   Bar         Grocery Store                Bakery
2               Stadium    Spanish Restaurant                  Pool
3        Scenic Lookout                Bakery          Yoga Studio
4  Fast Food Restaurant         Grocery Store           Supermarket

  5th Most Common Venue 6th Most Common Venue 7th Most Common Venue  \
0            Hobby Shop    Fried Chicken Joint    Italian Restaurant
1           Supermarket      Tapas Restaurant           Pizza Place
2                   Gym         Grocery Store  Fast Food Restaurant
3                  Food      Department Store          Dessert Shop
4           Flea Market                 Plaza         Farmers Market
```

```
  8th Most Common Venue 9th Most Common Venue 10th Most Common Venue
0         Metro Station           Cocktail Bar            Pizza Place
1           Coffee Shop     Seafood Restaurant         Farmers Market
2     Convenience Store       Cuban Restaurant          Deli / Bodega
3                 Diner         Discount Store             Donut Shop
4   Dumpling Restaurant     Falafel Restaurant      Electronics Store
```

[25]:
```
#searching location for 'madrid, españa'

address = 'madrid, españa'

geolocator = Nominatim(user_agent="mad_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geographical coordinate of Madrid are {}, {}.'.format(latitude,␣
 ↪longitude))
```

```
The geograpical coordinate of Madrid are 40.4167047, -3.7035825.
```

[26]:
```
# Map with clusters for 'Madrid, España'
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)
```

[27]:
```
# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]
```

[28]:
```
# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(madrid_merged['Latitude'],␣
 ↪madrid_merged['Longitude'], madrid_merged['Neighborhood'],␣
 ↪madrid_merged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        #color=rainbow[cluster-1],
        fill=True,
        #fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)
```

[29]:
```
map_clusters
```

```
[29]: <folium.folium.Map at 0x7f3af3215b38>
```