mcpp_taller7_camila_valencia

October 10, 2016

1 Taller 7

Métodos Computacionales para Políticas Públicas - URosario

Entrega: viernes 14-oct-2016 11:59 PM
[Camila Valencia] [camila.valencia@urosario.edu.co]

1.1 Instrucciones:

- Guarde una copia de este *Jupyter Notebook* en su computador, idealmente en una carpeta destinada al material del curso.
- Modifique el nombre del archivo del *notebook*, agregando al final un guión inferior y su nombre y apellido, separados estos últimos por otro guión inferior. Por ejemplo, mi *notebook* se llamaría: mcpp_taller7_santiago_matallana
- Marque el *notebook* con su nombre y e-mail en el bloque verde arriba. Reemplace el texto "[Su nombre acá]" con su nombre y apellido. Similar para su e-mail.
- Desarrolle la totalidad del taller sobre este *notebook*, insertando las celdas que sea necesario debajo de cada pregunta. Haga buen uso de las celdas para código y de las celdas tipo *markdown* según el caso.
- Recuerde salvar periódicamente sus avances.
- Cuando termine el taller:

1. Descárguelo en PDF. Si tiene algún problema con la conversión, descárguelo en HTML.
2. Suba todos los archivos a su repositorio en GitHub, en una carpeta destinada exclusivamente para este taller, antes de la fecha y hora límites.

(Todos los ejercicios tienen el mismo valor.)

Este taller tiene dos partes. Una obligatoria, relativamente fácil, y otra voluntaria y más retadora. Los invito a intentar desarrollar el taller en su totalidad.

En este taller exploraremos los datos de crimen de Chicago.

Descargue los datos de crimen del Chicago Data Portal solo para el año 2015 (https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2).

1.1.1 Parte obligatoria

1.1.2 1.

Calcule el número de crímenes en cada Community Area en 2015. Haga un gráfico de barras que lo ilustre.

In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
plt.rcParams["figure.figsize"] = [18.0, 10.0]
plt.style.use('ggplot')
```

In [3]:
```
crimes = pd.read_csv('Crimes_-_2015.csv', parse_dates=['Date'])
```

In [4]:
```
crimes.head()
```

Out[4]:
```
         ID Case Number                Date                  Block  IUCR  \
0  10514462    HZ256372 2015-01-01 00:00:00  073XX S EXCHANGE AVE  0281
1  10515175    HZ257172 2015-11-24 17:30:00      033XX W ADAMS ST  0820
2  10077106    HY266148 2015-05-19 01:12:00   009XX W BELMONT AVE  0560
3  10111002    HY299741 2015-06-13 16:45:00    015XX S HAMLIN AVE  143A
4  10301916    HY469211 2015-01-01 00:00:00     062XX W BARRY AVE  0266

          Primary Type              Description  \
0  CRIM SEXUAL ASSAULT           NON-AGGRAVATED
1                THEFT           $500 AND UNDER
2              ASSAULT                   SIMPLE
3    WEAPONS VIOLATION  UNLAWFUL POSS OF HANDGUN
4  CRIM SEXUAL ASSAULT                PREDATORY

            Location Description Arrest Domestic  \
0  NURSING HOME/RETIREMENT HOME  False    False
1                     RESIDENCE  False    False
2                     APARTMENT   True    False
3                         ALLEY  False    False
4                     RESIDENCE   True     True

              ...              Ward  Community Area  FBI Code  \
0             ...               7.0              43        02
1             ...              28.0              27        06
2             ...              44.0               6       08A
3             ...              24.0              29        15
4             ...              36.0              19        02

  X Coordinate Y Coordinate  Year             Updated On  Latitude  \
0          NaN          NaN  2015  05/10/2016 03:56:50 PM       NaN
1          NaN          NaN  2015  05/10/2016 03:56:50 PM       NaN
```

```
2     1169640.0    1921442.0  2015  05/11/2016 03:48:18 PM  41.939943
3     1151295.0    1892216.0  2015  05/11/2016 03:48:18 PM  41.860125
4     1134262.0    1919947.0  2015  05/11/2016 03:48:18 PM  41.936539

   Longitude                       Location
0        NaN                            NaN
1        NaN                            NaN
2 -87.651925  (41.939943264, -87.651924995)
3 -87.720117  (41.860124593, -87.720116627)
4 -87.781987  (41.936538876, -87.781987083)

[5 rows x 22 columns]
```

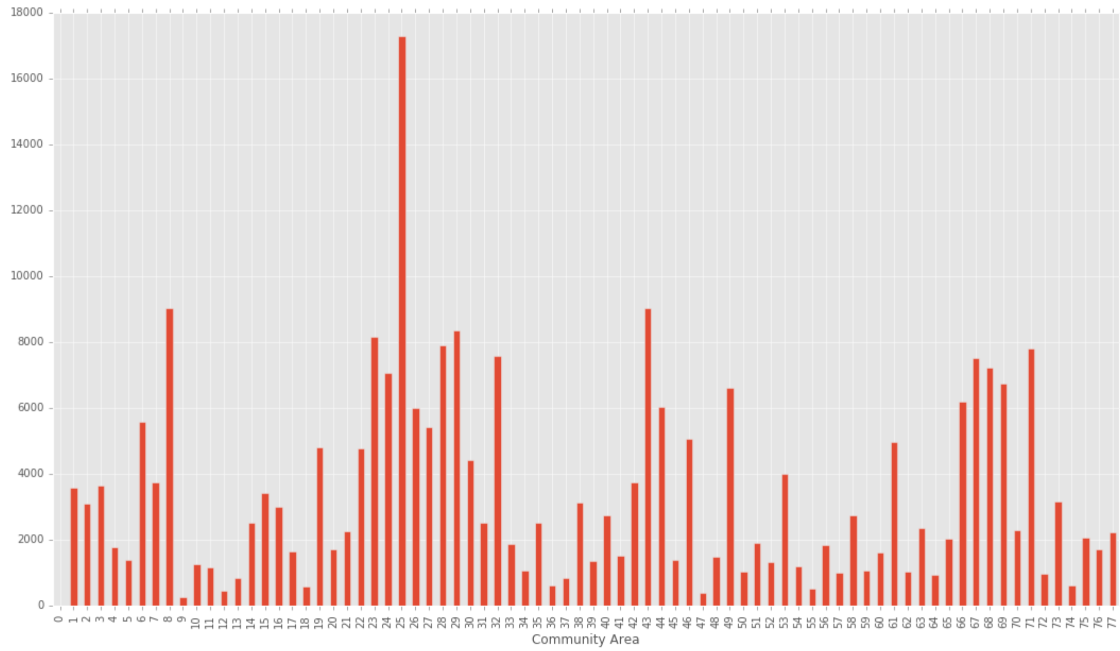
In [5]:
```
crimes_by_community = crimes.groupby('Community Area')
community_crime_count = crimes_by_community['ID'].agg('count')
community_crime_count.to_frame()
```

Out[5]:
```
                   ID
Community Area
0                   2
1                3573
2                3091
3                3645
4                1765
5                1390
6                5572
7                3750
8                9028
9                 256
10               1267
11               1154
12                445
13                840
14               2520
15               3434
16               3005
17               1651
18                586
19               4813
20               1725
21               2254
22               4777
23               8152
24               7053
25              17284
26               6012
27               5425
28               7906
```

```
29               8350
...               ...
48               1474
49               6605
50               1043
51               1918
52               1315
53               3992
54               1185
55                512
56               1843
57                993
58               2742
59               1080
60               1607
61               4970
62               1034
63               2358
64                932
65               2046
66               6192
67               7505
68               7218
69               6743
70               2303
71               7821
72                985
73               3147
74                614
75               2071
76               1719
77               2243

[78 rows x 1 columns]
```

```
In [6]: community_crime_count.plot(kind='bar');
```



1.1.3 2.

Ordene las Community Areas de acuerdo con el número de crímenes. ¿Qué Community Area (por nombre, idealmente) presenta el mayor número de crímenes? ¿El menor?

```
In [7]: datacrime = community_crime_count.to_frame()
        census = pd.read_csv('Census_Data.csv')
        census.head()
```

```
Out[7]:    Community Area Number COMMUNITY AREA NAME  PERCENT OF HOUSING CROWDED  \
        0                    1.0        Rogers Park                         7.7
        1                    2.0         West Ridge                         7.8
        2                    3.0             Uptown                         3.8
        3                    4.0      Lincoln Square                        3.4
        4                    5.0       North Center                         0.3

           PERCENT HOUSEHOLDS BELOW POVERTY  PERCENT AGED 16+ UNEMPLOYED  \
        0                              23.6                          8.7
        1                              17.2                          8.8
        2                              24.0                          8.9
        3                              10.9                          8.2
        4                               7.5                          5.2

           PERCENT AGED 25+ WITHOUT HIGH SCHOOL DIPLOMA  \
        0                                          18.2
        1                                          20.8
```

```
2                                              11.8
3                                              13.4
4                                               4.5

   PERCENT AGED UNDER 18 OR OVER 64  PER CAPITA INCOME   HARDSHIP INDEX
0                              27.5              23939             39.0
1                              38.5              23040             46.0
2                              22.2              35787             20.0
3                              25.5              37524             17.0
4                              26.2              57123              6.0
```

```
In [8]: datacrime['Nombre']=census['COMMUNITY AREA NAME']
        datacrime
```

```
Out[8]:                     ID               Nombre
        Community Area
        0                    2          Rogers Park
        1                 3573          West Ridge
        2                 3091              Uptown
        3                 3645      Lincoln Square
        4                 1765        North Center
        5                 1390           Lake View
        6                 5572        Lincoln Park
        7                 3750     Near North Side
        8                 9028         Edison Park
        9                  256        Norwood Park
        10                1267      Jefferson Park
        11                1154         Forest Glen
        12                 445          North Park
        13                 840         Albany Park
        14                2520        Portage Park
        15                3434         Irving Park
        16                3005             Dunning
        17                1651         Montclaire
        18                 586      Belmont Cragin
        19                4813             Hermosa
        20                1725            Avondale
        21                2254        Logan Square
        22                4777       Humboldt park
        23                8152           West Town
        24                7053              Austin
        25               17284  West Garfield Park
        26                6012  East Garfield Park
        27                5425      Near West Side
        28                7906      North Lawndale
        29                8350      South Lawndale
        ...                ...                 ...
        48                1474            Roseland
```

```
49                 6605                  Pullman
50                 1043            South Deering
51                 1918                East Side
52                 1315             West Pullman
53                 3992                Riverdale
54                 1185                Hegewisch
55                  512           Garfield Ridge
56                 1843          Archer Heights
57                  993           Brighton Park
58                 2742           McKinley Park
59                 1080              Bridgeport
60                 1607                 New City
61                 4970              West Elsdon
62                 1034                Gage Park
63                 2358                 Clearing
64                  932                West Lawn
65                 2046             Chicago Lawn
66                 6192          West Englewood
67                 7505                Englewood
68                 7218  Greater Grand Crossing
69                 6743                  Ashburn
70                 2303           Auburn Gresham
71                 7821                  Beverly
72                  985       Washington Height
73                 3147          Mount Greenwood
74                  614              Morgan Park
75                 2071                   O'Hare
76                 1719                Edgewater
77                 2243                  CHICAGO

[78 rows x 2 columns]
```

```
In [9]: test = datacrime.sort(['ID'], ascending=[False])
        test
```

```
C:\Users\camila.valencia\AppData\Local\Continuum\Anaconda3\lib\site-packages\ipyker
  from ipykernel import kernelapp as app
```

```
Out[9]:                    ID                   Nombre
        Community Area
        25              17284       West Garfield Park
        43               9039                  Chatham
        8                9028              Edison Park
        29               8350           South Lawndale
        23               8152                West Town
        28               7906           North Lawndale
        71               7821                  Beverly
```

32	7572	Near South Side
67	7505	Englewood
68	7218	Greater Grand Crossing
24	7053	Austin
69	6743	Ashburn
49	6605	Pullman
66	6192	West Englewood
44	6021	Avalon Park
26	6012	East Garfield Park
6	5572	Lincoln Park
27	5425	Near West Side
46	5064	Burnside
61	4970	West Elsdon
19	4813	Hermosa
22	4777	Humboldt park
30	4433	Lower West Side
53	3992	Riverdale
7	3750	Near North Side
42	3730	South Shore
3	3645	Lincoln Square
1	3573	West Ridge
15	3434	Irving Park
73	3147	Mount Greenwood
...	...	...
20	1725	Avondale
76	1719	Edgewater
17	1651	Montclaire
60	1607	New City
41	1507	Woodlawn
48	1474	Roseland
5	1390	Lake View
45	1380	South Chicago
39	1358	Washington Park
52	1315	West Pullman
10	1267	Jefferson Park
54	1185	Hegewisch
11	1154	Forest Glen
59	1080	Bridgeport
34	1074	Douglas
50	1043	South Deering
62	1034	Gage Park
57	993	Brighton Park
72	985	Washington Height
64	932	West Lawn
37	841	Grand Boulevard
13	840	Albany Park
36	630	Fuller Park
74	614	Morgan Park

```
18                  586           Belmont Cragin
55                  512           Garfield Ridge
12                  445               North Park
47                  386          Calumet Heights
9                   256             Norwood Park
0                     2              Rogers Park

[78 rows x 2 columns]
```

La community area que presenta mas crimenes es West Garfield Park numero de area 25 con 17284 crimenes respotados en el 2015y la que presenta el menor numero es Rogers Park con 2 crimenes en el 2015

1.1.4 3.

Cree una tabla cuyas filas sean días del año (yyyy-mm-dd) y las columnas las 77 Community Areas. En cada campo de la tabla deberá haber el correspondiente número de crímenes. Seleccione algunas Community Areas que le llamen la atención y haga un gráfico de serie de tiempo.

Pista: El siguiente código puede serle útil.

```
In [10]: # Create function to strip time from date field, and use it to create anot
         def to_day(timestamp):
             return timestamp.replace(minute=0,hour=0, second=0)

         crimes['Day'] = crimes['Date'].apply(to_day)
```

```
In [11]: crimes_by_day_community = crimes.groupby(['Day','Community Area'])
         crimes_by_day_community_count = crimes_by_day_community['ID'].agg('count')
         crimes_by_day_community_count.to_frame().head()
```

```
Out[11]:                            ID
         Day        Community Area
         2015-01-01 1               15
                    2                7
                    3               12
                    4                6
                    5                5
```

```
In [12]: community_crime_timeseries =crimes_by_day_community_count.unstack('Communi
         community_crime_timeseries
```

```
Out[12]: Community Area   0     1     2     3     4    5     6     7     8    9   .
         Day                                                                       .
         2015-01-01      NaN  15.0   7.0  12.0   6.0  5.0  22.0  12.0  45.0  1.0  .
         2015-01-02      NaN   5.0   9.0   8.0   3.0  2.0  10.0   9.0  27.0  NaN  .
         2015-01-03      NaN   7.0  11.0   9.0   7.0  4.0   6.0  11.0  27.0  1.0  .
         2015-01-04      NaN  12.0   7.0   9.0  10.0  3.0  15.0   5.0  16.0  1.0  .
         2015-01-05      NaN   6.0   7.0   5.0   4.0  5.0  15.0   7.0  11.0  1.0  .
         2015-01-06      NaN   8.0   8.0   6.0   5.0  NaN  14.0   7.0  13.0  NaN  .
```

2015-01-07	NaN	6.0	2.0	5.0	5.0	1.0	8.0	6.0	17.0	1.0	.
2015-01-08	NaN	6.0	7.0	3.0	5.0	NaN	6.0	5.0	8.0	1.0	.
2015-01-09	NaN	10.0	5.0	10.0	2.0	4.0	14.0	6.0	21.0	NaN	.
2015-01-10	NaN	6.0	12.0	8.0	NaN	1.0	10.0	5.0	24.0	2.0	.
2015-01-11	NaN	8.0	6.0	11.0	5.0	4.0	20.0	4.0	26.0	1.0	.
2015-01-12	NaN	3.0	6.0	6.0	6.0	1.0	7.0	11.0	17.0	NaN	.
2015-01-13	NaN	10.0	10.0	9.0	5.0	4.0	8.0	6.0	15.0	NaN	.
2015-01-14	1.0	17.0	8.0	9.0	4.0	3.0	6.0	11.0	15.0	1.0	.
2015-01-15	NaN	9.0	8.0	8.0	6.0	7.0	9.0	11.0	18.0	1.0	.
2015-01-16	NaN	13.0	6.0	12.0	5.0	5.0	12.0	12.0	22.0	1.0	.
2015-01-17	NaN	12.0	5.0	5.0	2.0	5.0	16.0	7.0	31.0	1.0	.
2015-01-18	NaN	12.0	6.0	12.0	7.0	5.0	14.0	8.0	18.0	1.0	.
2015-01-19	NaN	3.0	12.0	7.0	3.0	5.0	10.0	11.0	25.0	NaN	.
2015-01-20	NaN	8.0	8.0	9.0	10.0	3.0	13.0	12.0	31.0	NaN	.
2015-01-21	NaN	18.0	8.0	9.0	10.0	9.0	12.0	9.0	29.0	2.0	.
2015-01-22	NaN	7.0	7.0	6.0	9.0	7.0	14.0	10.0	21.0	NaN	.
2015-01-23	NaN	12.0	11.0	6.0	4.0	6.0	8.0	6.0	27.0	1.0	.
2015-01-24	NaN	14.0	9.0	7.0	8.0	4.0	11.0	15.0	23.0	NaN	.
2015-01-25	NaN	5.0	9.0	3.0	5.0	2.0	20.0	6.0	29.0	NaN	.
2015-01-26	NaN	10.0	8.0	5.0	4.0	3.0	13.0	5.0	13.0	1.0	.
2015-01-27	NaN	7.0	8.0	6.0	3.0	3.0	6.0	3.0	13.0	NaN	.
2015-01-28	NaN	15.0	9.0	11.0	1.0	1.0	9.0	11.0	26.0	NaN	.
2015-01-29	NaN	9.0	9.0	11.0	2.0	3.0	11.0	7.0	26.0	1.0	.
2015-01-30	NaN	8.0	11.0	6.0	5.0	3.0	12.0	10.0	22.0	NaN	.
...	...	...	...	...	...	...	...	...	...	...	.
2015-12-02	NaN	12.0	6.0	11.0	4.0	7.0	12.0	11.0	23.0	1.0	.
2015-12-03	NaN	6.0	11.0	10.0	2.0	5.0	12.0	8.0	32.0	1.0	.
2015-12-04	NaN	8.0	9.0	8.0	5.0	5.0	10.0	4.0	27.0	NaN	.
2015-12-05	NaN	9.0	7.0	4.0	5.0	6.0	8.0	13.0	33.0	1.0	.
2015-12-06	NaN	9.0	11.0	11.0	2.0	NaN	19.0	8.0	24.0	NaN	.
2015-12-07	NaN	7.0	10.0	8.0	2.0	6.0	10.0	8.0	21.0	1.0	.
2015-12-08	NaN	6.0	9.0	5.0	3.0	6.0	17.0	13.0	32.0	3.0	.
2015-12-09	NaN	14.0	11.0	7.0	8.0	9.0	16.0	13.0	25.0	NaN	.
2015-12-10	NaN	3.0	9.0	9.0	8.0	6.0	25.0	13.0	30.0	1.0	.
2015-12-11	NaN	11.0	7.0	11.0	6.0	5.0	20.0	10.0	29.0	NaN	.
2015-12-12	NaN	10.0	5.0	7.0	8.0	4.0	26.0	15.0	39.0	1.0	.
2015-12-13	NaN	7.0	12.0	10.0	3.0	4.0	20.0	9.0	38.0	NaN	.
2015-12-14	NaN	15.0	13.0	11.0	9.0	6.0	10.0	10.0	27.0	1.0	.
2015-12-15	NaN	10.0	6.0	12.0	5.0	3.0	12.0	21.0	26.0	NaN	.
2015-12-16	NaN	6.0	7.0	11.0	5.0	4.0	20.0	9.0	20.0	1.0	.
2015-12-17	NaN	9.0	8.0	8.0	7.0	5.0	19.0	19.0	28.0	NaN	.
2015-12-18	NaN	13.0	7.0	6.0	6.0	2.0	17.0	20.0	28.0	NaN	.
2015-12-19	NaN	7.0	11.0	6.0	3.0	1.0	10.0	12.0	41.0	2.0	.
2015-12-20	NaN	12.0	10.0	13.0	7.0	4.0	18.0	11.0	24.0	1.0	.
2015-12-21	NaN	6.0	1.0	11.0	4.0	1.0	6.0	11.0	32.0	NaN	.
2015-12-22	NaN	14.0	11.0	16.0	5.0	4.0	13.0	9.0	26.0	NaN	.
2015-12-23	NaN	12.0	11.0	14.0	6.0	5.0	18.0	13.0	28.0	NaN	.
2015-12-24	NaN	8.0	11.0	2.0	6.0	3.0	20.0	11.0	26.0	NaN	.

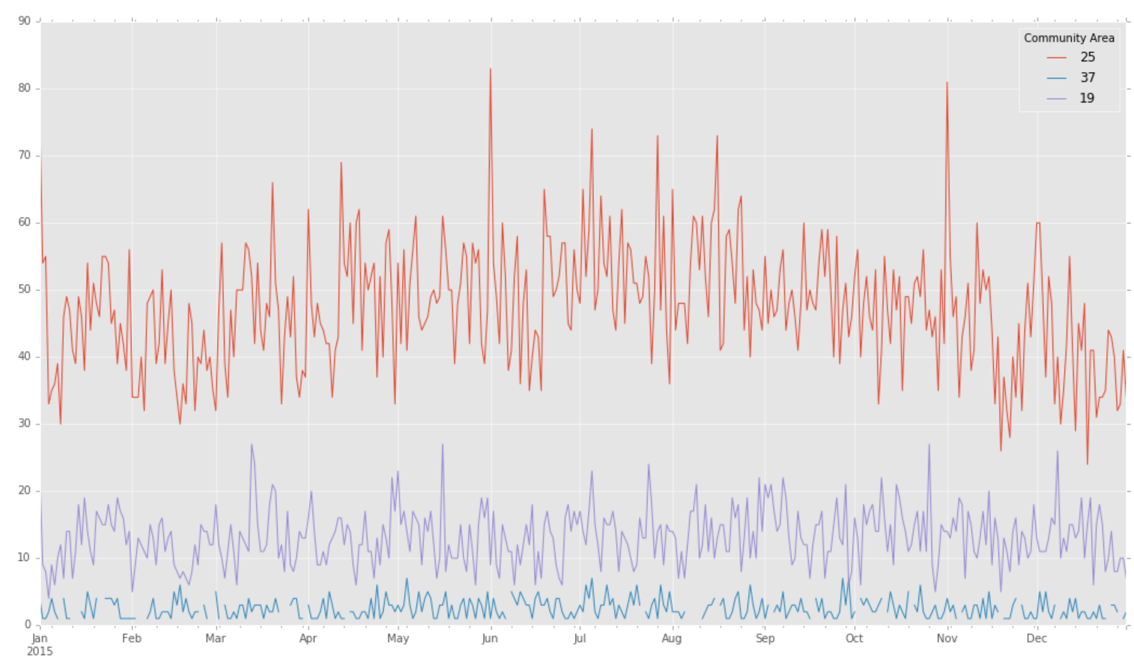
```
2015-12-25       NaN   2.0   5.0   3.0   3.0  NaN   5.0   5.0  10.0  1.0  .
2015-12-26       NaN   6.0  12.0  13.0   1.0  1.0  17.0   5.0  25.0  NaN  .
2015-12-27       NaN  13.0   8.0   6.0   3.0  1.0  16.0  10.0  35.0  NaN  .
2015-12-28       NaN   7.0   8.0   6.0   2.0  2.0  14.0   9.0  19.0  NaN  .
2015-12-29       NaN   6.0   7.0  14.0   8.0  3.0   9.0   5.0  24.0  NaN  .
2015-12-30       NaN   5.0   9.0   8.0   4.0  1.0  11.0  17.0  28.0  1.0  .
2015-12-31       NaN   8.0   4.0   9.0   4.0  3.0  19.0   5.0  26.0  NaN  .

Community Area    68    69    70    71   72    73   74   75    76    77
Day
2015-01-01      29.0  25.0   9.0  45.0  2.0   8.0  2.0  5.0   6.0   8.0
2015-01-02      12.0  22.0   5.0  17.0  1.0  11.0  1.0  3.0   6.0   5.0
2015-01-03      23.0  12.0   8.0  18.0  NaN   8.0  1.0  7.0   3.0   3.0
2015-01-04      13.0  15.0   9.0  12.0  1.0   5.0  NaN  1.0   6.0   1.0
2015-01-05      16.0  12.0   8.0  17.0  NaN   5.0  2.0  2.0   7.0   5.0
2015-01-06      15.0  14.0   6.0  11.0  2.0   8.0  2.0  3.0   6.0   4.0
2015-01-07      11.0   7.0   4.0  16.0  4.0   7.0  NaN  3.0   7.0   1.0
2015-01-08       9.0   9.0   6.0  10.0  2.0   4.0  1.0  5.0   3.0   3.0
2015-01-09      18.0  14.0  10.0  20.0  1.0   9.0  2.0  8.0   5.0   2.0
2015-01-10       9.0  13.0   6.0  28.0  3.0   3.0  1.0  5.0   5.0   2.0
2015-01-11      17.0   8.0  11.0  17.0  2.0  10.0  2.0  4.0   2.0   4.0
2015-01-12      12.0  18.0   6.0  19.0  3.0   5.0  1.0  4.0   4.0   3.0
2015-01-13      19.0  12.0   9.0  11.0  2.0   6.0  NaN  4.0   6.0   6.0
2015-01-14      21.0  16.0   6.0  24.0  NaN   5.0  1.0  1.0   2.0   6.0
2015-01-15      20.0  20.0   4.0  22.0  3.0   7.0  2.0  6.0   6.0   5.0
2015-01-16      18.0  17.0   8.0  16.0  4.0   5.0  NaN  6.0   7.0   3.0
2015-01-17      29.0  13.0   7.0  15.0  2.0   6.0  2.0  5.0  10.0   3.0
2015-01-18      20.0  22.0   3.0  20.0  1.0  11.0  NaN  8.0   2.0   4.0
2015-01-19      24.0  19.0   3.0  28.0  NaN  13.0  NaN  6.0   8.0   4.0
2015-01-20      19.0  15.0   5.0  23.0  1.0  10.0  NaN  8.0   6.0   7.0
2015-01-21      11.0  12.0   7.0  29.0  5.0   3.0  2.0  6.0   3.0   9.0
2015-01-22      19.0  13.0  11.0  22.0  1.0   5.0  1.0  6.0   4.0   6.0
2015-01-23      27.0  18.0   6.0  20.0  1.0   7.0  3.0  4.0   6.0   8.0
2015-01-24      10.0  13.0   5.0  17.0  1.0   8.0  NaN  3.0   5.0  10.0
2015-01-25      15.0  13.0   7.0  15.0  3.0   4.0  1.0  6.0   3.0   5.0
2015-01-26      18.0  16.0  10.0  29.0  1.0   8.0  1.0  9.0   2.0   8.0
2015-01-27      20.0  16.0   9.0  20.0  2.0  11.0  2.0  6.0   3.0   5.0
2015-01-28      13.0  16.0   3.0  18.0  2.0   7.0  1.0  6.0   7.0   6.0
2015-01-29      19.0  20.0   4.0  24.0  2.0   6.0  1.0  8.0   5.0   5.0
2015-01-30      20.0  22.0   6.0  20.0  5.0   7.0  NaN  6.0   2.0   4.0
...              ...   ...   ...   ...  ...   ...  ...  ...   ...   ...
2015-12-02      23.0  13.0   6.0  24.0  1.0   8.0  4.0  6.0   6.0   7.0
2015-12-03      17.0  19.0   7.0  17.0  2.0   7.0  3.0  3.0   3.0   4.0
2015-12-04      18.0  15.0   3.0  23.0  2.0   7.0  NaN  6.0   5.0   3.0
2015-12-05      15.0  21.0   5.0  26.0  3.0   9.0  3.0  5.0   2.0   4.0
2015-12-06      19.0  20.0   6.0  21.0  5.0  10.0  1.0  1.0   4.0  10.0
2015-12-07      20.0  21.0   3.0  30.0  3.0   8.0  3.0  6.0   4.0   6.0
2015-12-08       9.0  16.0   9.0  32.0  2.0   7.0  2.0  1.0   2.0   9.0
```

```
2015-12-09       11.0  18.0   4.0  20.0  6.0   7.0  1.0  4.0   4.0   6.0
2015-12-10       20.0  16.0   8.0  20.0  2.0   8.0  4.0  5.0   3.0   2.0
2015-12-11       19.0  19.0   3.0  19.0  4.0  10.0  3.0  4.0  10.0   9.0
2015-12-12       19.0  18.0   9.0  19.0  1.0   7.0  2.0  NaN   5.0   3.0
2015-12-13       18.0  16.0   7.0  14.0  NaN   2.0  3.0  3.0   6.0   5.0
2015-12-14       18.0  30.0  12.0  13.0  4.0   5.0  2.0  6.0   3.0  12.0
2015-12-15       19.0  19.0  11.0  15.0  NaN   8.0  1.0  7.0   2.0   7.0
2015-12-16       19.0  15.0  10.0  24.0  1.0   9.0  2.0  4.0   4.0   6.0
2015-12-17       16.0  20.0   4.0  16.0  NaN  10.0  NaN  6.0   7.0   8.0
2015-12-18       17.0  17.0   7.0  20.0  3.0   9.0  NaN  4.0   5.0   4.0
2015-12-19       10.0  13.0   5.0  22.0  1.0   7.0  NaN  2.0   4.0   4.0
2015-12-20       15.0  23.0   5.0  15.0  3.0   7.0  2.0  2.0   5.0   1.0
2015-12-21       12.0  15.0   3.0  19.0  2.0   9.0  1.0  1.0   2.0   4.0
2015-12-22       26.0  16.0   3.0  16.0  2.0   6.0  NaN  5.0   2.0   8.0
2015-12-23       19.0  23.0   8.0  19.0  1.0  10.0  4.0  8.0   2.0   7.0
2015-12-24       17.0  22.0   4.0  19.0  3.0   3.0  2.0  5.0   3.0   5.0
2015-12-25       13.0  16.0   3.0  18.0  1.0   5.0  NaN  6.0   3.0   4.0
2015-12-26       17.0  18.0   6.0  17.0  1.0  10.0  NaN  3.0   7.0   8.0
2015-12-27       13.0  19.0   3.0  26.0  2.0   8.0  2.0  1.0   4.0   2.0
2015-12-28       12.0  23.0   9.0  14.0  2.0   6.0  2.0  2.0   3.0   9.0
2015-12-29       19.0  16.0   7.0  18.0  NaN   8.0  3.0  3.0   2.0   5.0
2015-12-30       11.0  23.0   6.0  14.0  2.0   8.0  1.0  7.0   4.0   4.0
2015-12-31       18.0  16.0   4.0  19.0  1.0   8.0  1.0  3.0   2.0   4.0

[365 rows x 78 columns]
```

```
In [13]: community_crime_timeseries[[25,37,19]].plot();
```



1.1.5 Parte voluntaria

Descargue la base de datos de información socioeconómica (https://data.cityofchicago.org/Health-Human-Services/Census-Data-Selected-socioeconomic-indicators-in-C/kn9c-c2s2).

1.1.6 4.

Cree una tabla que agregue el número de crímenes por Community Area. Una esa tabla con la de datos socioeconómicos y cree un "scatter plot" de número de crímenes vs ingreso per cápita. Explique la relación en palabras.

```
In [14]: crimes = pd.read_csv('Crimes_-_2015.csv')
         crimes_by_community = crimes.groupby('Community Area')
         community_crime_count = crimes_by_community['ID'].agg('count')
         datacrime = community_crime_count.to_frame()
```

```
In [15]: census = pd.read_csv('Census_Data.csv')
```

```
In [16]: census['CRIME']=datacrime['ID']
         census
```

```
Out[16]:     Community Area Number     COMMUNITY AREA NAME  PERCENT OF HOUSING CROW
         0                     1.0            Rogers Park
         1                     2.0            West Ridge
         2                     3.0                Uptown
         3                     4.0         Lincoln Square
         4                     5.0           North Center
         5                     6.0              Lake View
         6                     7.0           Lincoln Park
         7                     8.0        Near North Side
         8                     9.0            Edison Park
         9                    10.0           Norwood Park
         10                   11.0         Jefferson Park
         11                   12.0            Forest Glen
         12                   13.0             North Park
         13                   14.0            Albany Park                        1
         14                   15.0           Portage Park
         15                   16.0            Irving Park
         16                   17.0                Dunning
         17                   18.0             Montclaire
         18                   19.0         Belmont Cragin                        1
         19                   20.0                Hermosa
         20                   21.0               Avondale
         21                   22.0           Logan Square
```

```
22                    23.0           Humboldt park                        1
23                    24.0               West Town
24                    25.0                  Austin
25                    26.0       West Garfield Park
26                    27.0       East Garfield Park
27                    28.0           Near West Side
28                    29.0          North Lawndale
29                    30.0          South Lawndale                        1
..                     ...                     ...
48                    49.0                Roseland
49                    50.0                 Pullman
50                    51.0           South Deering
51                    52.0               East Side
52                    53.0            West Pullman
53                    54.0               Riverdale
54                    55.0               Hegewisch
55                    56.0          Garfield Ridge
56                    57.0          Archer Heights
57                    58.0           Brighton Park                        1
58                    59.0           McKinley Park
59                    60.0              Bridgeport
60                    61.0                New City                        1
61                    62.0             West Elsdon                        1
62                    63.0               Gage Park                        1
63                    64.0                Clearing
64                    65.0               West Lawn
65                    66.0            Chicago Lawn
66                    67.0          West Englewood
67                    68.0               Englewood
68                    69.0  Greater Grand Crossing
69                    70.0                 Ashburn
70                    71.0          Auburn Gresham
71                    72.0                 Beverly
72                    73.0      Washington Height
73                    74.0        Mount Greenwood
74                    75.0             Morgan Park
75                    76.0                 O'Hare
76                    77.0               Edgewater
77                     NaN                 CHICAGO

    PERCENT HOUSEHOLDS BELOW POVERTY  PERCENT AGED 16+ UNEMPLOYED  \
0                               23.6                          8.7
1                               17.2                          8.8
2                               24.0                          8.9
3                               10.9                          8.2
4                                7.5                          5.2
5                               11.4                          4.7
6                               12.3                          5.1
```

7	12.9	7.0
8	3.3	6.5
9	5.4	9.0
10	8.6	12.4
11	7.5	6.8
12	13.2	9.9
13	19.2	10.0
14	11.6	12.6
15	13.1	10.0
16	10.6	10.0
17	15.3	13.8
18	18.7	14.6
19	20.5	13.1
20	15.3	9.2
21	16.8	8.2
22	33.9	17.3
23	14.7	6.6
24	28.6	22.6
25	41.7	25.8
26	42.4	19.6
27	20.6	10.7
28	43.1	21.2
29	30.7	15.8
..	...	...
48	19.8	20.3
49	21.6	22.8
50	29.2	16.3
51	19.2	12.1
52	25.9	19.4
53	56.5	34.6
54	17.1	9.6
55	8.8	11.3
56	14.1	16.5
57	23.6	13.9
58	18.7	13.4
59	18.9	13.7
60	29.0	23.0
61	15.6	16.7
62	23.4	18.2
63	8.9	9.5
64	14.9	9.6
65	27.9	17.1
66	34.4	35.9
67	46.6	28.0
68	29.6	23.0
69	10.4	11.7
70	27.6	28.3
71	5.1	8.0

```
72                              16.9                           20.8
73                               3.4                            8.7
74                              13.2                           15.0
75                              15.4                            7.1
76                              18.2                            9.2
77                              19.7                           12.9

    PERCENT AGED 25+ WITHOUT HIGH SCHOOL DIPLOMA  \
0                                           18.2
1                                           20.8
2                                           11.8
3                                           13.4
4                                            4.5
5                                            2.6
6                                            3.6
7                                            2.5
8                                            7.4
9                                           11.5
10                                          13.4
11                                           4.9
12                                          14.4
13                                          32.9
14                                          19.3
15                                          22.4
16                                          16.2
17                                          23.5
18                                          37.3
19                                          41.6
20                                          24.7
21                                          14.8
22                                          35.4
23                                          12.9
24                                          24.4
25                                          24.5
26                                          21.3
27                                           9.6
28                                          27.6
29                                          54.8
..                                           ...
48                                          16.9
49                                          13.1
50                                          21.0
51                                          31.9
52                                          20.5
53                                          27.5
54                                          19.2
55                                          19.3
56                                          35.9
```

57	45.1
58	32.9
59	22.2
60	41.5
61	37.0
62	51.5
63	18.8
64	33.6
65	31.2
66	26.3
67	28.5
68	16.5
69	17.7
70	18.5
71	3.7
72	13.7
73	4.3
74	10.8
75	10.9
76	9.7
77	19.5

	PERCENT AGED UNDER 18 OR OVER 64	PER CAPITA INCOME	HARDSHIP INDEX
0	27.5	23939	39.0
1	38.5	23040	46.0
2	22.2	35787	20.0
3	25.5	37524	17.0
4	26.2	57123	6.0
5	17.0	60058	5.0
6	21.5	71551	2.0
7	22.6	88669	1.0
8	35.3	40959	8.0
9	39.5	32875	21.0
10	35.5	27751	25.0
11	40.5	44164	11.0
12	39.0	26576	33.0
13	32.0	21323	53.0
14	34.0	24336	35.0
15	31.6	27249	34.0
16	33.6	26282	28.0
17	38.6	22014	50.0
18	37.3	15461	70.0
19	36.4	15089	71.0
20	31.0	20039	42.0
21	26.2	31908	23.0
22	38.0	13781	85.0
23	21.7	43198	10.0
24	37.9	15957	73.0

```
25                               43.6           10934           92.0
26                               43.2           12961           83.0
27                               22.2           44689           15.0
28                               42.7           12034           87.0
29                               33.8           10402           96.0
..                                ...             ...            ...
48                               41.2           17949           52.0
49                               38.6           20588           51.0
50                               39.5           14685           65.0
51                               42.8           17104           64.0
52                               42.1           16563           62.0
53                               51.5            8201           98.0
54                               42.9           22677           44.0
55                               38.1           26353           32.0
56                               39.2           16134           67.0
57                               39.3           13089           84.0
58                               35.6           16954           61.0
59                               31.3           22694           43.0
60                               38.9           12765           91.0
61                               37.7           15754           69.0
62                               38.8           12171           93.0
63                               37.6           25113           29.0
64                               39.6           16907           56.0
65                               40.6           13231           80.0
66                               40.7           11317           89.0
67                               42.5           11888           94.0
68                               41.0           17285           66.0
69                               36.9           23482           37.0
70                               41.9           15528           74.0
71                               40.5           39523           12.0
72                               42.6           19713           48.0
73                               36.8           34381           16.0
74                               40.3           27149           30.0
75                               30.3           25828           24.0
76                               23.8           33385           19.0
77                               33.5           28202            NaN

    CRIME
0       2
1    3573
2    3091
3    3645
4    1765
5    1390
6    5572
7    3750
8    9028
9     256
```

10	1267
11	1154
12	445
13	840
14	2520
15	3434
16	3005
17	1651
18	586
19	4813
20	1725
21	2254
22	4777
23	8152
24	7053
25	17284
26	6012
27	5425
28	7906
29	8350
..	...
48	1474
49	6605
50	1043
51	1918
52	1315
53	3992
54	1185
55	512
56	1843
57	993
58	2742
59	1080
60	1607
61	4970
62	1034
63	2358
64	932
65	2046
66	6192
67	7505
68	7218
69	6743
70	2303
71	7821
72	985
73	3147
74	614

```
75    2071
76    1719
77    2243

[78 rows x 10 columns]
```

```
In [17]: census.plot.scatter( x='PER CAPITA INCOME ',y='CRIME',);
```