# mcpp_taller7_john_caro

October 15, 2016

## 1 Taller 7

Métodos Computacionales para Políticas Públicas - URosario

**Entrega: viernes 14-oct-2016 11:59 PM**

**[John Alexander Caro Becerra]** [jhonalexbc@gmail.com]

### 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).

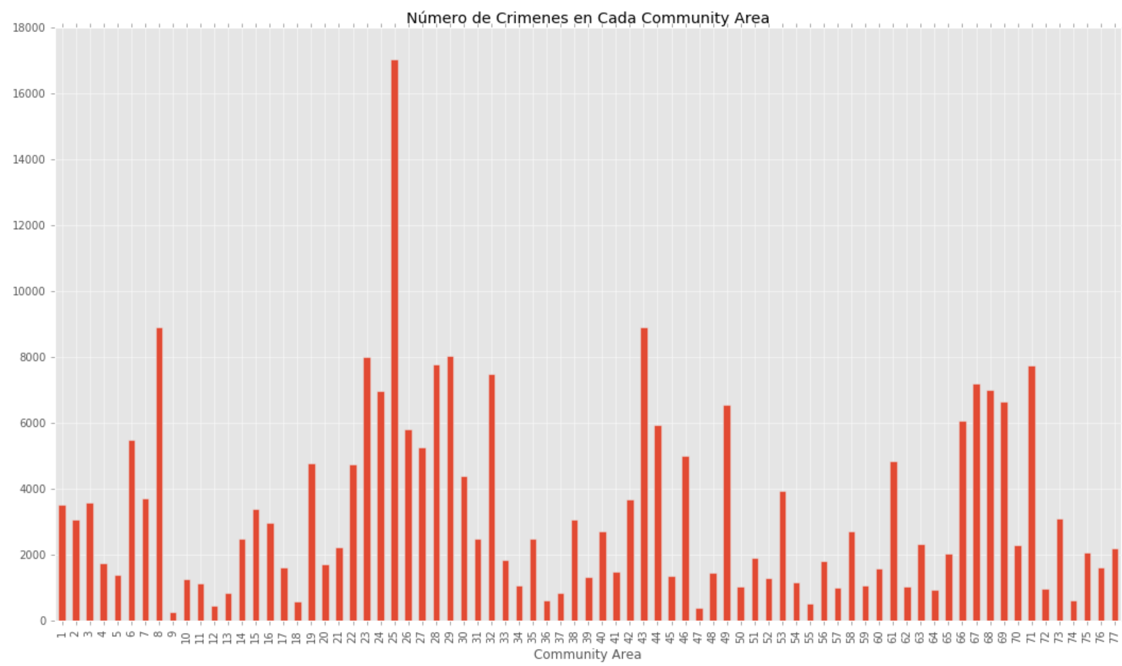
### 1.1.1 Parte obligatoria

### 1.1.2 1.

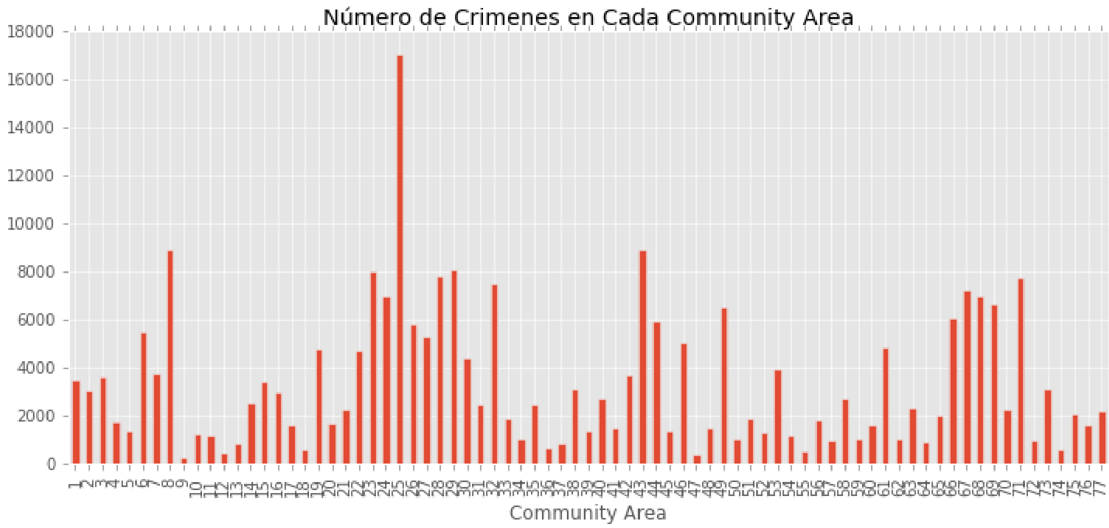
Calcule el número de crímenes en cada Community Area en 2015. Haga un gráfico de barras que lo ilustre.

```
In [1]: import pandas as pd
        import numpy as np
        from numpy.random import randn
        import matplotlib.pyplot as plt
        %matplotlib inline
        plt.rcParams["figure.figsize"] = [18.0, 10.0]
        plt.style.use('ggplot')

        crimes = pd.read_csv('Crimes_-_2001_to_present.csv', parse_dates=['Date'])
```

```
In [2]: numeros_crimenes_c_area = crimes.groupby('Community Area')['ID'].agg('count
        numeros_crimenes_c_area.plot(kind='bar',title="Número de Crimenes en Cada C
```




```
In [3]: numeros_crimenes_c_area.plot(kind='bar', figsize=(12,5),title="Número de Cr
```




### 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 [4]:
```
numeros_crimenes_c_area_ordenado = numeros_crimenes_c_area.sort_values()
C_area_menor_crimenes = numeros_crimenes_c_area_ordenado.head(1)
C_area_mayor_crimenes = numeros_crimenes_c_area_ordenado.tail(1)
numeros_crimenes_c_area_ordenado
```

Out[4]:
```
Community Area
9         254
47        380
12        444
55        506
18        572
74        608
36        622
13        828
37        834
64        927
72        976
57        986
62       1030
50       1037
34       1057
59       1063
11       1140
54       1173
```

```
10      1259
52      1303
39      1340
45      1356
5       1375
48      1458
41      1493
60      1590
76      1622
17      1626
20      1703
4       1747
       ...
73      3109
15      3392
1       3519
3       3585
42      3665
7       3726
53      3949
30      4393
22      4737
19      4769
61      4842
46      5009
27      5270
6       5495
26      5794
44      5943
66      6081
49      6545
69      6653
24      6959
68      6991
67      7203
32      7489
71      7733
28      7788
23      8015
29      8039
43      8906
8       8920
25     17020
Name: ID, dtype: int64
```

```
In [5]: C_area_menor_crimenes
```

```
Out[5]: Community Area
```

```
9    254
Name: ID, dtype: int64
```

In [6]:
```
C_area_mayor_crimenes
```

Out[6]:
```
Community Area
25    17020
Name: ID, dtype: int64
```

#### 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 [7]:
```
# Create function to strip time from date field, and use it to create anoth
def to_day(timestamp):
    return timestamp.replace(minute=0,hour=0, second=0)

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

In [11]:
```
crimes_by_community_day = crimes.groupby(['Community Area', 'Day'])

crimes_by_community_day_count = crimes_by_community_day['ID'].agg('count')
crimes_by_community_day_count
```

Out[11]:
```
Community Area  Day
1               2015-01-01    13
                2015-01-02     5
                2015-01-03     7
                2015-01-04    12
                2015-01-05     6
                2015-01-06     7
                2015-01-07     6
                2015-01-08     6
                2015-01-09    10
                2015-01-10     6
                2015-01-11     8
                2015-01-12     3
                2015-01-13    10
                2015-01-14    17
                2015-01-15     9
                2015-01-16    13
                2015-01-17    12
                2015-01-18    12
                2015-01-19     3
                2015-01-20     8
                2015-01-21    18
```

```
                2015-01-22     7
                2015-01-23    12
                2015-01-24    14
                2015-01-25     5
                2015-01-26    10
                2015-01-27     6
                2015-01-28    15
                2015-01-29     9
                2015-01-30     8
                              ..
77              2015-12-01     2
                2015-12-02     8
                2015-12-03     4
                2015-12-04     3
                2015-12-05     4
                2015-12-06     9
                2015-12-07     6
                2015-12-08     9
                2015-12-09     6
                2015-12-10     2
                2015-12-11     9
                2015-12-12     3
                2015-12-13     5
                2015-12-14    12
                2015-12-15     7
                2015-12-16     5
                2015-12-17     7
                2015-12-18     4
                2015-12-19     4
                2015-12-20     1
                2015-12-21     5
                2015-12-22     8
                2015-12-23     7
                2015-12-24     5
                2015-12-25     4
                2015-12-26     8
                2015-12-27     2
                2015-12-28     9
                2015-12-29     4
                2015-12-30     5
Name: ID, dtype: int64
```

In [12]:
```
community_timeseries = crimes_by_community_day_count.unstack('Community Ar
community_timeseries
```

Out[12]:
```
Community Area     1     2     3     4    5     6     7     8    9    10  .
Day                                                                      .
2015-01-01      13.0   7.0  11.0   4.0  5.0  22.0  12.0  43.0  1.0  5.0  .
```

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

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2015-12-20 | 12.0 | 8.0 | 13.0 | 7.0 | 3.0 | 18.0 | 10.0 | 21.0 | 1.0 | 2.0 | . |
| 2015-12-21 | 6.0 | 2.0 | 11.0 | 4.0 | NaN | 6.0 | 11.0 | 32.0 | NaN | 3.0 | . |
| 2015-12-22 | 13.0 | 11.0 | 15.0 | 5.0 | 4.0 | 13.0 | 9.0 | 26.0 | NaN | NaN | . |
| 2015-12-23 | 12.0 | 11.0 | 14.0 | 6.0 | 5.0 | 18.0 | 12.0 | 28.0 | NaN | 3.0 | . |
| 2015-12-24 | 8.0 | 11.0 | 2.0 | 6.0 | 3.0 | 19.0 | 11.0 | 26.0 | NaN | 7.0 | . |
| 2015-12-25 | 2.0 | 6.0 | 3.0 | 3.0 | NaN | 5.0 | 5.0 | 10.0 | 1.0 | 5.0 | . |
| 2015-12-26 | 6.0 | 11.0 | 12.0 | 1.0 | 1.0 | 17.0 | 5.0 | 26.0 | NaN | 9.0 | . |
| 2015-12-27 | 14.0 | 8.0 | 6.0 | 3.0 | 1.0 | 16.0 | 11.0 | 32.0 | NaN | 1.0 | . |
| 2015-12-28 | 7.0 | 8.0 | 5.0 | 2.0 | 2.0 | 10.0 | 8.0 | 19.0 | NaN | 3.0 | . |
| 2015-12-29 | 6.0 | 7.0 | 12.0 | 8.0 | 3.0 | 8.0 | 5.0 | 25.0 | NaN | 1.0 | . |
| 2015-12-30 | 5.0 | 8.0 | 7.0 | 4.0 | 1.0 | 11.0 | 15.0 | 27.0 | 1.0 | 6.0 | . |
| 2015-12-31 | NaN | NaN | NaN | NaN | NaN | 1.0 | NaN | NaN | NaN | NaN | . |

| Community Area | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 |
|---|---|---|---|---|---|---|---|---|---|---|
| Day | | | | | | | | | | |
| 2015-01-01 | 29.0 | 23.0 | 9.0 | 44.0 | 2.0 | 8.0 | 2.0 | 5.0 | 6.0 | 8.0 |
| 2015-01-02 | 12.0 | 21.0 | 5.0 | 17.0 | 1.0 | 11.0 | 1.0 | 2.0 | 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 | 3.0 | 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 | 19.0 | 20.0 | 4.0 | 21.0 | 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 | 5.0 | 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 | 2.0 | 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 | 26.0 | 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 | 5.0 | 6.0 |
| 2015-01-29 | 19.0 | 20.0 | 4.0 | 24.0 | 2.0 | 6.0 | 1.0 | 8.0 | 4.0 | 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 | 23.0 | 1.0 | 8.0 | 4.0 | 6.0 | 6.0 | 8.0 |
| 2015-12-03 | 16.0 | 19.0 | 7.0 | 17.0 | 2.0 | 7.0 | 3.0 | 3.0 | 2.0 | 4.0 |

```
2015-12-04      18.0  15.0   3.0  21.0  1.0   7.0  NaN  6.0   2.0   3.0
2015-12-05      13.0  20.0   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  22.0  5.0  10.0  1.0  1.0   4.0   9.0
2015-12-07      20.0  21.0   3.0  30.0  3.0   8.0  2.0  6.0   4.0   6.0
2015-12-08       9.0  16.0   9.0  29.0  1.0   6.0  2.0  1.0   1.0   9.0
2015-12-09      11.0  16.0   3.0  20.0  6.0   7.0  1.0  4.0   4.0   6.0
2015-12-10      19.0  16.0   8.0  20.0  2.0   8.0  4.0  5.0   3.0   2.0
2015-12-11      17.0  19.0   3.0  19.0  4.0  10.0  3.0  4.0  10.0   9.0
2015-12-12      17.0  16.0   9.0  19.0  1.0   7.0  2.0  NaN   5.0   3.0
2015-12-13      17.0  17.0   7.0  14.0  NaN   2.0  3.0  3.0   5.0   5.0
2015-12-14      18.0  29.0  11.0  12.0  4.0   4.0  2.0  6.0   3.0  12.0
2015-12-15      19.0  19.0  11.0  14.0  NaN   8.0  1.0  7.0   3.0   7.0
2015-12-16      18.0  15.0  10.0  23.0  1.0   9.0  2.0  4.0   4.0   5.0
2015-12-17      15.0  19.0   4.0  15.0  NaN  10.0  NaN  6.0   7.0   7.0
2015-12-18      15.0  16.0   7.0  19.0  3.0   9.0  NaN  4.0   5.0   4.0
2015-12-19      10.0  12.0   5.0  21.0  1.0   7.0  NaN  2.0   3.0   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  18.0  2.0   9.0  1.0  1.0   2.0   5.0
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      18.0  23.0   8.0  19.0  1.0   9.0  4.0  8.0   2.0   7.0
2015-12-24      16.0  22.0   4.0  19.0  3.0   3.0  2.0  5.0   3.0   5.0
2015-12-25      13.0  15.0   3.0  18.0  1.0   5.0  NaN  6.0   3.0   4.0
2015-12-26      15.0  18.0   6.0  17.0  1.0  10.0  NaN  3.0   7.0   8.0
2015-12-27      11.0  19.0   3.0  26.0  2.0   8.0  2.0  1.0   4.0   2.0
2015-12-28      12.0  22.0   9.0  14.0  2.0   6.0  2.0  2.0   3.0   9.0
2015-12-29      18.0  16.0   7.0  18.0  NaN   8.0  3.0  2.0   2.0   4.0
2015-12-30      11.0  23.0   6.0  14.0  2.0   8.0  1.0  7.0   5.0   5.0
2015-12-31       NaN   NaN   NaN   NaN  NaN   NaN  NaN  NaN   NaN   NaN

[365 rows x 77 columns]
```

In [14]: `community_timeseries.describe()`

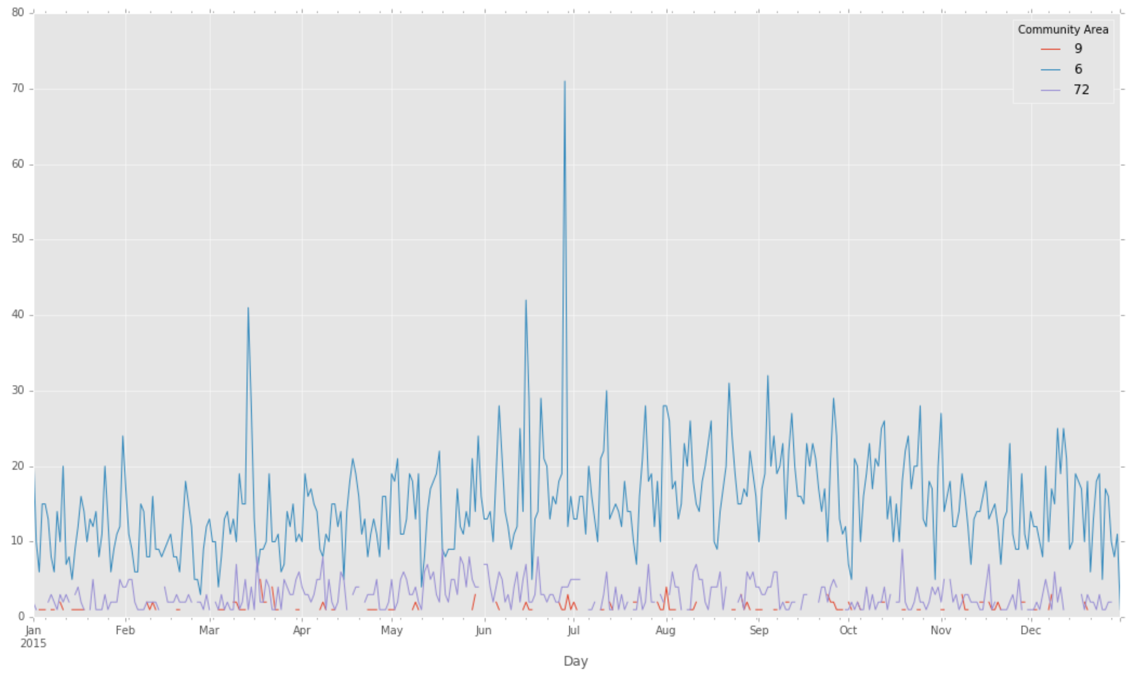
```
C:\Users\Usuario\Anaconda3\lib\site-packages\numpy\lib\function_base.py:3834: Runti
  RuntimeWarning)
```

Out[14]:

```
Community Area          1           2           3           4           5
count          364.000000  364.000000  364.000000  357.000000  347.00000
mean             9.667582    8.403846    9.848901    4.893557    3.96253
std              3.396959    3.024070    3.656267    2.318680    2.09769
min              2.000000    2.000000    2.000000    1.000000    1.00000
25%                   NaN         NaN         NaN         NaN         Na
50%                   NaN         NaN         NaN         NaN         Na
75%                   NaN         NaN         NaN         NaN         Na
max             20.000000   21.000000   29.000000   13.000000   12.00000

Community Area          6           7           8           9          1
```

```
count           365.000000  364.000000  364.000000  180.000000  352.000000
mean             15.054795   10.236264   24.505495    1.411111    3.576705
std               6.664426    3.826715    7.202117    0.707414    1.924448
min               1.000000    2.000000    5.000000    1.000000    1.000000
25%              10.000000         NaN         NaN         NaN         NaN
50%              14.000000         NaN         NaN         NaN         NaN
75%              19.000000         NaN         NaN         NaN         NaN
max              71.000000   21.000000   53.000000    5.000000   10.000000

Community Area    ...              68          69          70          71
count             ...      364.000000  364.000000  364.000000  364.000000
mean              ...       19.206044   18.277473    6.260989   21.244505
std               ...        5.626469    5.052609    2.679483    5.210510
min               ...        7.000000    5.000000    1.000000   10.000000
25%               ...             NaN         NaN         NaN         NaN
50%               ...             NaN         NaN         NaN         NaN
75%               ...             NaN         NaN         NaN         NaN
max               ...       41.000000   33.000000   16.000000   44.000000

Community Area          72          73          74          75          76
count           332.000000  364.000000  289.000000  362.000000  360.000000
mean              2.939759    8.541209    2.103806    5.668508    4.505556
std               1.820324    3.082154    1.275961    2.567764    2.141244
min               1.000000    1.000000    1.000000    1.000000    1.000000
25%                    NaN         NaN         NaN         NaN         NaN
50%                    NaN         NaN         NaN         NaN         NaN
75%                    NaN         NaN         NaN         NaN         NaN
max               9.000000   20.000000    9.000000   15.000000   11.000000

Community Area         77
count           362.00000
mean              6.10221
std               2.55641
min               1.00000
25%                   NaN
50%                   NaN
75%                   NaN
max              14.00000

[8 rows x 77 columns]
```

```
In [15]: community_timeseries[[9,6,72]].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 [79]: n_c=numeros_crimenes_c_area

In [41]: crimes2 = pd.read_csv('Census_Data_-_Selected_socioeconomic_indicators_in_

In [49]: crimes2

Out[49]:     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

[78 rows x 9 columns]
```

```
In [69]: precapita=crimes2.loc[0:76,['Community Area Number',"PER CAPITA INCOME "]]

In [70]: precapita

Out[70]:     Community Area Number  PER CAPITA INCOME
         0                     1.0              23939
         1                     2.0              23040
         2                     3.0              35787
         3                     4.0              37524
         4                     5.0              57123
         5                     6.0              60058
         6                     7.0              71551
         7                     8.0              88669
         8                     9.0              40959
         9                    10.0              32875
         10                   11.0              27751
         11                   12.0              44164
         12                   13.0              26576
         13                   14.0              21323
         14                   15.0              24336
         15                   16.0              27249
         16                   17.0              26282
         17                   18.0              22014
         18                   19.0              15461
         19                   20.0              15089
         20                   21.0              20039
         21                   22.0              31908
         22                   23.0              13781
         23                   24.0              43198
         24                   25.0              15957
         25                   26.0              10934
         26                   27.0              12961
         27                   28.0              44689
         28                   29.0              12034
         29                   30.0              10402
         ..                    ...                ...
         47                   48.0              28887
         48                   49.0              17949
         49                   50.0              20588
         50                   51.0              14685
         51                   52.0              17104
         52                   53.0              16563
```

```
53                    54.0                 8201
54                    55.0                22677
55                    56.0                26353
56                    57.0                16134
57                    58.0                13089
58                    59.0                16954
59                    60.0                22694
60                    61.0                12765
61                    62.0                15754
62                    63.0                12171
63                    64.0                25113
64                    65.0                16907
65                    66.0                13231
66                    67.0                11317
67                    68.0                11888
68                    69.0                17285
69                    70.0                23482
70                    71.0                15528
71                    72.0                39523
72                    73.0                19713
73                    74.0                34381
74                    75.0                27149
75                    76.0                25828
76                    77.0                33385

[77 rows x 2 columns]
```

In [77]:
```
n_c.iloc[0:77,]
```

Out[77]:
```
                   ID
Community Area
1                3519
2                3059
3                3585
4                1747
5                1375
6                5495
7                3726
8                8920
9                 254
10               1259
11               1140
12                444
13                828
14               2501
15               3392
16               2960
17               1626
```

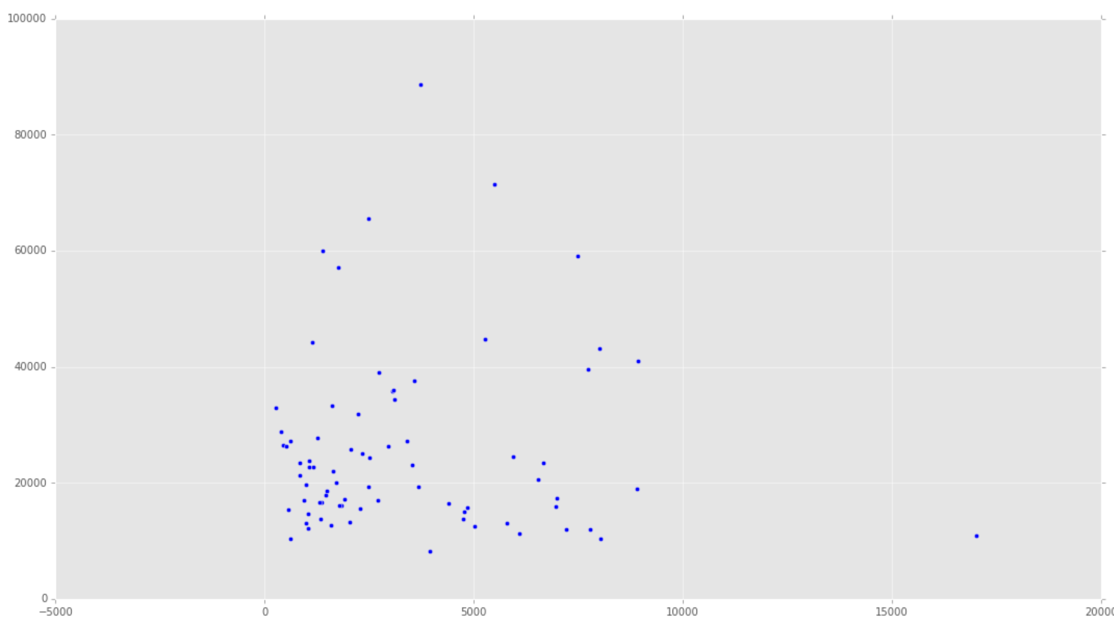
```
18                 572
19                4769
20                1703
21                2232
22                4737
23                8015
24                6959
25               17020
26                5794
27                5270
28                7788
29                8039
30                4393
...                ...
48                1458
49                6545
50                1037
51                1900
52                1303
53                3949
54                1173
55                 506
56                1796
57                 986
58                2706
59                1063
60                1590
61                4842
62                1030
63                2334
64                 927
65                2027
66                6081
67                7203
68                6991
69                6653
70                2279
71                7733
72                 976
73                3109
74                 608
75                2052
76                1622
77                2209

[77 rows x 1 columns]
```

```
In [60]: precapita
```

```
Out[60]:     Community Area Number  PER CAPITA INCOME
         0                     1.0              23939
         1                     2.0              23040
         2                     3.0              35787
         3                     4.0              37524
         4                     5.0              57123
         5                     6.0              60058
         6                     7.0              71551
         7                     8.0              88669
         8                     9.0              40959
         9                    10.0              32875
         10                   11.0              27751
         11                   12.0              44164
         12                   13.0              26576
         13                   14.0              21323
         14                   15.0              24336
         15                   16.0              27249
         16                   17.0              26282
         17                   18.0              22014
         18                   19.0              15461
         19                   20.0              15089
         20                   21.0              20039
         21                   22.0              31908
         22                   23.0              13781
         23                   24.0              43198
         24                   25.0              15957
         25                   26.0              10934
         26                   27.0              12961
         27                   28.0              44689
         28                   29.0              12034
         29                   30.0              10402
         ..                    ...                ...
         47                   48.0              28887
         48                   49.0              17949
         49                   50.0              20588
         50                   51.0              14685
         51                   52.0              17104
         52                   53.0              16563
         53                   54.0               8201
         54                   55.0              22677
         55                   56.0              26353
         56                   57.0              16134
         57                   58.0              13089
         58                   59.0              16954
         59                   60.0              22694
         60                   61.0              12765
         61                   62.0              15754
         62                   63.0              12171
```

```
63                    64.0              25113
64                    65.0              16907
65                    66.0              13231
66                    67.0              11317
67                    68.0              11888
68                    69.0              17285
69                    70.0              23482
70                    71.0              15528
71                    72.0              39523
72                    73.0              19713
73                    74.0              34381
74                    75.0              27149
75                    76.0              25828
76                    77.0              33385

[77 rows x 2 columns]
```

```
In [94]: precapita["Crimes"]=n_c.iloc[0:77,]

         #df_tmp["tmp_factorial"] = df_tmp["tmp"].apply(np.math.factorial)

In [97]: plt.scatter(precapita["Crimes"],precapita["PER CAPITA INCOME "])
         plt.show()
```




La relación que se puede observar es que a medida que el ingreso per capita aumenta, el número de crimenes disminuye, en cambio en donde el ingreso per capita es muy bajo, el número de crimenes es mayor.