

Taller 7

Métodos Computacionales para Políticas Públicas - URosario

Entrega: viernes 5-mar-2019 11:59 PM

****Francisco Monsalve Guaracao****

`francisco.monsalve@urosario.edu.co`

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/data>).

Parte obligatoria

```
In [1]: 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")
```

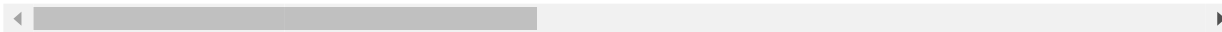
```
In [2]: crimes = pd.read_csv("Crimes_-_2001_to_present.csv", parse_dates=["Date"])
crimes.head()
```

Out[2]:

	ID	Case Number	Date	Block	IUCR	Primary Type	Description	Location Description
0	3305587	HK301958	2004-04-14 19:45:00	009XX N LAWLER AVE	2024	NARCOTICS	POSS: HEROIN(WHITE)	STREET

	ID	Case Number	Date	Block	IUCR	Primary Type	Description	Location Description
1	3305589	HK338529	2004-05-02 10:45:00	081XX S KEDZIE AVE	0560	ASSAULT	SIMPLE	STREET
2	3305592	HK336096	2003-11-24 00:01:00	021XX S HOMAN AVE	1562	SEX OFFENSE	AGG CRIMINAL SEXUAL ABUSE	RESIDENCE
3	3305593	HK322023	2004-04-24 12:35:00	005XX W 66TH ST	1821	NARCOTICS	MANU/DEL:CANNABIS 10GM OR LESS	STREET
4	3305594	HK341518	2004-05-03 14:30:00	046XX S WHIPPLE ST	1320	CRIMINAL DAMAGE	TO VEHICLE	STREET

5 rows × 22 columns



```
In [3]: #quedar sólo con el año 2015
is_2015 = crimes["Year"]==2015
```

```
In [4]: print(is_2015.head())

0    False
1    False
2    False
3    False
4    False
Name: Year, dtype: bool
```

```
In [5]: crimes_2015 = crimes[is_2015]
print(crimes_2015.shape)

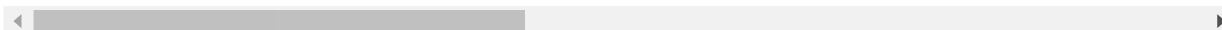
(264071, 22)
```

```
In [6]: crimes_2015.head()
```

Out[6]:

	ID	Case Number	Date	Block	IUCR	Primary Type	Description	L
46	10072163	HY260736	2015-05-14 20:00:00	038XX S ASHLAND AVE	0810	THEFT	OVER \$500	S
123	10072550	HY252513	2015-05-08 17:44:00	015XX W 69TH ST	2017	NARCOTICS	MANU/DELIVER:CRACK	SID
1458	10075231	HY264086	2015-05-17 15:00:00	012XX W PRATT BLVD	0460	BATTERY	SIMPLE	
1596	10075512	HY264097	2015-05-17 15:15:00	067XX N WESTERN AVE	0430	BATTERY	AGGRAVATED: OTHER DANG WEAPON	SID
1764	10076372	HY261267	2015-05-14 14:10:00	065XX W BRYN MAWR AVE	0460	BATTERY	SIMPLE	SC F BU

5 rows × 22 columns



1.

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

```
In [7]: crimes_2015[["Community Area", "Year"]].describe().round(2)
```

Out[7]:

	Community Area	Year
count	264071.00	264071.0
mean	37.57	2015.0
std	21.42	0.0

	Community Area	Year
min	0.00	2015.0
25%	23.00	2015.0
50%	32.00	2015.0
75%	57.00	2015.0
max	77.00	2015.0

```
In [8]: crimes_by_community_2015 = crimes_2015.groupby("Community Area")
crimes_by_community_2015.groups
```

```
Out[8]: {0.0: Int64Index([2276916, 5916362], dtype='int64'),
1.0: Int64Index([ 1458, 12653, 26518, 26534, 34677, 57636,
63623,
74726, 81935, 87812,
...
6455774, 6491012, 6570375, 6607177, 6626703, 6652579, 6661
974,
6663791, 6677217, 6760097],
dtype='int64', length=3591),
2.0: Int64Index([ 1596, 70438, 119915, 122368, 122501, 122996,
123774,
124721, 130343, 137956,
...
6435856, 6443444, 6455300, 6502416, 6520419, 6609910, 6617
589,
6718267, 6780016, 6807250],
dtype='int64', length=3111),
3.0: Int64Index([ 16866, 22449, 49172, 50121, 58109, 63109,
67835,
68271, 72373, 112281,
...
6399539, 6408844, 6408850, 6580647, 6641745, 6673235, 6718
912,
6804664, 6807556, 6812548],
dtype='int64', length=3663),
4.0: Int64Index([ 1822, 23464, 63544, 119798, 119799, 119949,
```

```

120517,
        122582, 123522, 129706,
        ...
        6281501, 6287829, 6295752, 6369257, 6443445, 6517171, 6593
403,
        6645240, 6693286, 6779544],
        dtype='int64', length=1770),
5.0: Int64Index([ 16923,  24130,  81951, 121415, 121586, 121923,
124900,
        125615, 128865, 129212,
        ...
        6326036, 6333934, 6405085, 6418493, 6449576, 6507389, 6552
727,
        6642592, 6809798, 6816507],
        dtype='int64', length=1395),
6.0: Int64Index([ 45168,  46997,  81952, 102475, 105426, 119330,
119801,
        120579, 121392, 121745,
        ...
        6569598, 6618593, 6648454, 6650974, 6654028, 6714654, 6729
396,
        6779524, 6789347, 6812114],
        dtype='int64', length=5590),
7.0: Int64Index([ 112889, 121718, 122212, 123523, 124842, 125571,
125764,
        126891, 128387, 130691,
        ...
        6508888, 6514890, 6544804, 6547724, 6560436, 6641768, 6672
370,
        6676444, 6805043, 6809617],
        dtype='int64', length=3764),
8.0: Int64Index([  3306,  34615,  34671,  49051,  54320,  60650,
81927,
        81954, 119529, 119826,
        ...
        6564610, 6626694, 6636394, 6641746, 6653318, 6683864, 6701
488,
        6702261, 6709165, 6713018],
        dtype='int64', length=9075),

```

```

9.0: Int64Index([ 125052, 2135825, 2136008, 2136179, 2136220, 2136334,
2136781,
                2137062, 2137153, 2137618,
                ...
                5942638, 5942751, 5944151, 5946558, 5946960, 5946985, 5946
987,
                6346180, 6348752, 6702302],
                dtype='int64', length=258),
10.0: Int64Index([ 1764, 120372, 120800, 123435, 124892, 12554
1, 133879,
                134684, 138504, 139286,
                ...
                6417910, 6443466, 6443992, 6453033, 6500216, 6517877, 6541
675,
                6584893, 6608796, 6804378],
                dtype='int64', length=1278),
11.0: Int64Index([ 34202, 95332, 125753, 144331, 153864, 15694
6, 2069159,
                2135099, 2135135, 2135281,
                ...
                5947250, 6283349, 6287831, 6303115, 6318882, 6323986, 6446
668,
                6499561, 6522723, 6693362],
                dtype='int64', length=1157),
12.0: Int64Index([ 120999, 125288, 126346, 1962604, 2098018, 213566
9, 2136655,
                2137178, 2137184, 2137757,
                ...
                5943701, 5946184, 5946559, 5947155, 6340154, 6376265, 6503
866,
                6517864, 6520464, 6730066],
                dtype='int64', length=450),
13.0: Int64Index([ 28058, 58255, 121913, 127312, 137479, 14948
6, 164623,
                2135198, 2135373, 2135403,
                ...
                5945966, 5945976, 5947000, 5947134, 6272800, 6313544, 6329
574,
                6340918, 6512599, 6589303],

```

```

dtype='int64', length=843),
14.0: Int64Index([ 23504,  43253, 125406, 127007, 135481, 13791
1, 138116,
               139571, 139617, 150342,
               ...
               6442845, 6496010, 6499560, 6558923, 6618511, 6626689, 6672
418,
               6730065, 6758988, 6770322],
dtype='int64', length=2536),
15.0: Int64Index([ 30545,  68900, 120254, 120823, 123085, 12559
3, 126038,
               127772, 129278, 130168,
               ...
               6518566, 6521818, 6527110, 6589302, 6599396, 6624757, 6625
672,
               6716088, 6769707, 6805022],
dtype='int64', length=3453),
16.0: Int64Index([ 27148,  55948,  56542, 120932, 122111, 12553
2, 126957,
               128059, 128470, 128584,
               ...
               6467640, 6484958, 6517876, 6552744, 6553615, 6587527, 6608
828,
               6617606, 6629393, 6753601],
dtype='int64', length=3019),
17.0: Int64Index([ 10720,  31107,  34055, 119619, 120474, 12080
1, 122200,
               122770, 124339, 125098,
               ...
               6443441, 6446219, 6455777, 6581439, 6618514, 6654011, 6656
286,
               6691045, 6695845, 6718254],
dtype='int64', length=1660),
18.0: Int64Index([ 125293, 138669, 152991, 2135230, 2135358, 213541
9, 2135989,
               2137001, 2137070, 2137329,
               ...
               5944807, 5944986, 5945614, 5947099, 6265121, 6271103, 6277
214,

```



```

        6369126, 6716057, 6725775],
        dtype='int64', length=587),
19.0: Int64Index([ 13279,  20932,  48422,  60654,  63641,  6837
9,  81939,
        81941,  84113, 119213,
        ...
        6512574, 6516361, 6516381, 6518564, 6545487, 6595951, 6617
618,
        6621065, 6639917, 6704705],
        dtype='int64', length=4838),
20.0: Int64Index([  6219,  32117,  45824,  56521,  67856, 12673
6, 128064,
        131241, 137348, 140823,
        ...
        6301621, 6311701, 6317224, 6336509, 6382921, 6503935, 6574
726,
        6601998, 6644353, 6750553],
        dtype='int64', length=1732),
21.0: Int64Index([ 29944,  63570, 119337, 119487, 121966, 12208
8, 122223,
        122518, 124223, 124234,
        ...
        6306733, 6393910, 6395697, 6451349, 6454099, 6515649, 6528
698,
        6600370, 6685276, 6688391],
        dtype='int64', length=2263),
22.0: Int64Index([ 23453,  23623,  60788,  65570,  66289,  6838
1, 102971,
        103701, 120663, 121386,
        ...
        6550412, 6585170, 6592705, 6595879, 6620239, 6646944, 6754
336,
        6763319, 6805343, 6807698],
        dtype='int64', length=4805),
23.0: Int64Index([  5544,  5557,  5558,  6230,  8574,  858
3,  9775,
        12091, 12205, 12635,
        ...
        6641751, 6678921, 6691044, 6695025, 6697342, 6721562, 6723

```

```

292,
        6780892, 6807254, 6826331],
        dtype='int64', length=8199),
24.0: Int64Index([ 13298,  15829,  35289,  63078,  67852,  7522
1,  75251,
        104315,  119551,  119582,
        ...,
        6560444, 6582525, 6592724, 6611626, 6666393, 6707798, 6779
986,
        6805307, 6806783, 6815778],
        dtype='int64', length=7082),
25.0: Int64Index([  6240,   6745,   8598,  10321,  12209,  1276
9,  15652,
        15690,   15712,   15817,
        ...,
        6707058, 6716106, 6720116, 6723269, 6729419, 6757291, 6760
090,
        6769029, 6770327, 6778963],
        dtype='int64', length=17382),
26.0: Int64Index([  3328,   6231,   6814,   7173,   7293,   873
6,  10811,
        12083,   12092,   12093,
        ...,
        6441772, 6451350, 6456375, 6482203, 6486736, 6524078, 6615
764,
        6689134, 6691784, 6826818],
        dtype='int64', length=6029),
27.0: Int64Index([  5563,   5576,   5580,   7227,   7261,   927
6,  12085,
        12124,   12128,   12148,
        ...,
        6509494, 6524907, 6542501, 6551778, 6617604, 6654003, 6663
789,
        6758967, 6780012, 6809628],
        dtype='int64', length=5448),
28.0: Int64Index([  9753,  12663,  20410,  28067,  30564,  3837
9,  46453,
        48956,   49073,   53300,
        ...

```

```

        6594238, 6601974, 6602818, 6615307, 6626692, 6657993, 6709
174,
        6799965, 6806806, 6827866],
        dtype='int64', length=7948),
29.0: Int64Index([ 5596,  8579,  8586,  8589,  9766, 1264
1, 15809,
        15812, 16323, 16791,
        ...
        6601985, 6602820, 6606339, 6606357, 6727128, 6757282, 6761
177,
        6778972, 6778987, 6826933],
        dtype='int64', length=8410),
30.0: Int64Index([ 8601, 20483, 26617, 67150, 77568, 9668
6, 112084,
        121741, 122087, 123937,
        ...
        6487928, 6511138, 6536711, 6563739, 6566256, 6566949, 6583
216,
        6590099, 6604611, 6728686],
        dtype='int64', length=4452),
31.0: Int64Index([ 8593, 28062, 34230, 34239, 68376, 11939
6, 119418,
        121018, 122454, 124911,
        ...
        6369100, 6504511, 6514097, 6552723, 6569605, 6570476, 6587
512,
        6617605, 6754339, 6755063],
        dtype='int64', length=2528),
32.0: Int64Index([ 3458, 40925, 45422, 69378, 70317, 10662
8, 108576,
        111286, 120931, 121320,
        ...
        6594255, 6616781, 6642588, 6652604, 6670000, 6673168, 6721
548,
        6729508, 6755071, 6763312],
        dtype='int64', length=7595),
33.0: Int64Index([ 31677, 50159, 119435, 124596, 125238, 12581
9, 129640,
        131843, 131844, 133456,

```

```

        ...
        6324919, 6342649, 6420072, 6435927, 6456832, 6500224, 6507
180,
        6593384, 6598539, 6648452],
        dtype='int64', length=1884),
    34.0: Int64Index([ 30115, 122909, 131418, 132399, 138331, 211254
9, 2135022,
        2135394, 2135999, 2136250,
        ...
        6354777, 6378714, 6409578, 6489578, 6490344, 6495443, 6558
884,
        6563668, 6678164, 6678923],
        dtype='int64', length=1084),
    35.0: Int64Index([ 4856, 9264, 22907, 34220, 34591, 4474
0, 47190,
        50112, 64190, 67108,
        ...
        6348753, 6376269, 6411088, 6439717, 6506026, 6550434, 6551
844,
        6653660, 6658709, 6678138],
        dtype='int64', length=2523),
    36.0: Int64Index([ 122899, 124833, 1931658, 2112843, 2135759, 213582
4, 2136606,
        2137946, 2138212, 2138566,
        ...
        5946780, 5946873, 5947110, 5947224, 6000698, 6295735, 6353
006,
        6549788, 6758441, 6807472],
        dtype='int64', length=632),
    37.0: Int64Index([ 7305, 71812, 71813, 94629, 122905, 13415
8, 2135446,
        2135623, 2136078, 2136338,
        ...
        5946759, 5946853, 5947120, 6271144, 6283292, 6289598, 6385
386,
        6420757, 6451325, 6512654],
        dtype='int64', length=842),
    38.0: Int64Index([ 9150, 12753, 13768, 14299, 31032, 4267
7, 64123,

```

```

        64184, 64188, 64707,
        ...
        6457362, 6483497, 6508884, 6537479, 6615937, 6620236, 6622
945,
        6688441, 6714656, 6769644],
        dtype='int64', length=3153),
39.0: Int64Index([ 14284, 14292, 22915, 28037, 52284, 6473
5, 122750,
        122882, 123258, 123370,
        ...
        6290351, 6317138, 6372405, 6405995, 6406721, 6432300, 6441
750,
        6446133, 6451351, 6551974],
        dtype='int64', length=1370),
40.0: Int64Index([ 3291, 12637, 16366, 19953, 37318, 6893
2, 81986,
        110132, 121582, 121719,
        ...
        6456264, 6535084, 6546329, 6566199, 6583189, 6590086, 6614
054,
        6645999, 6696571, 6697344],
        dtype='int64', length=2767),
41.0: Int64Index([ 37897, 74675, 114211, 124926, 125076, 12638
9, 132042,
        135194, 135332, 143174,
        ...
        6270255, 6289579, 6317161, 6334798, 6340122, 6457386, 6494
147,
        6505349, 6788733, 6804767],
        dtype='int64', length=1516),
42.0: Int64Index([ 6682, 8725, 12143, 12152, 12630, 1264
0, 19961,
        20369, 20476, 27176,
        ...
        6538213, 6579872, 6584081, 6601989, 6618636, 6638173, 6759
694,
        6804872, 6804874, 6807286],
        dtype='int64', length=3752),
43.0: Int64Index([ 6609, 8553, 10322, 10731, 12157, 1223

```

```

8, 12766,
    13293, 13773, 16894,
    ...
    6666391, 6678964, 6684603, 6698412, 6720769, 6750564, 6759
609,
    6793379, 6804735, 6805389],
    dtype='int64', length=9099),
44.0: Int64Index([ 13311, 16845, 19957, 20375, 34094, 4216
7, 49580,
    54329, 68868, 71409,
    ...
    6636398, 6655612, 6661119, 6676457, 6699122, 6699123, 6727
159,
    6750567, 6761304, 6769009],
    dtype='int64', length=6043),
45.0: Int64Index([ 119421, 122544, 122828, 133591, 141048, 14514
7, 155366,
    162129, 162378, 163464,
    ...
    5946913, 5947055, 5947059, 6264137, 6312826, 6377930, 6404
301,
    6460860, 6547658, 6583291],
    dtype='int64', length=1390),
46.0: Int64Index([ 13213, 16483, 23073, 33596, 42128, 5012
7, 57053,
    63122, 63669, 67139,
    ...
    6655597, 6663792, 6684597, 6690224, 6704756, 6729400, 6760
702,
    6770303, 6801243, 6833122],
    dtype='int64', length=5086),
47.0: Int64Index([ 74722, 138681, 2135032, 2137475, 2137650, 213781
1, 2138047,
    2138142, 2138561, 2138994,
    ...
    5940020, 5940171, 5942135, 5942200, 5942906, 5944770, 5946
932,
    6520379, 6584889, 6807526],
    dtype='int64', length=389),

```

```

48.0: Int64Index([ 37884,  56384,  63533, 120212, 120256, 12536
8, 126191,
               128197, 160214, 2022079,
               ...
               6508004, 6543226, 6543889, 6550433, 6582395, 6658718, 6667
193,
               6678996, 6760722, 6804633],
               dtype='int64', length=1484),
49.0: Int64Index([ 12162,  16823,  24006,  54350, 119374, 11982
1, 120802,
               121063, 121196, 121970,
               ...
               6629397, 6641781, 6649348, 6671462, 6672137, 6672357, 6682
901,
               6726511, 6727932, 6788810],
               dtype='int64', length=6644),
50.0: Int64Index([ 119917, 125244, 125510, 128257, 128276, 13245
6, 132645,
               1931633, 2136062, 2136516,
               ...
               5947091, 5947232, 6279858, 6347383, 6393199, 6423345, 6439
750,
               6595057, 6704001, 6730058],
               dtype='int64', length=1048),
51.0: Int64Index([ 20365,  68359, 111910, 119443, 120234, 12176
6, 122340,
               126774, 127073, 127346,
               ...
               6277194, 6288778, 6294897, 6300701, 6351410, 6377112, 6487
886,
               6511926, 6595025, 6806807],
               dtype='int64', length=1922),
52.0: Int64Index([ 109504, 121614, 130315, 132785, 135948, 13916
1, 141146,
               149962, 150980, 153360,
               ...
               5945736, 5946038, 5946310, 5947058, 6264184, 6353918, 6368
095,
               6499541, 6642589, 6642590],

```

```

dtype='int64', length=1318),
53.0: Int64Index([ 6640, 10211, 119398, 119490, 119779, 12045
7, 121678,
121937, 122819, 123222,
...
6426708, 6431573, 6434318, 6447284, 6538193, 6584080, 6613
216,
6660333, 6713028, 6819922],
dtype='int64', length=4013),
54.0: Int64Index([ 10233, 120054, 121256, 122155, 123595, 12685
7, 153361,
156078, 163194, 171528,
...
5946908, 5946944, 5947133, 5947142, 5947143, 6280772, 6289
662,
6349589, 6421470, 6582407],
dtype='int64', length=1189),
55.0: Int64Index([ 120085, 122665, 2046628, 2135095, 2135740, 213595
4, 2137995,
2138977, 2141309, 2141605,
...
5946109, 5946237, 5947057, 6282452, 6290376, 6359023, 6376
205,
6446216, 6457905, 6538197],
dtype='int64', length=516),
56.0: Int64Index([ 8704, 8708, 12659, 13759, 16867, 2035
2, 21055,
21475, 24135, 28059,
...
6451862, 6455819, 6498865, 6517188, 6535211, 6577253, 6620
244,
6629344, 6645992, 6729422],
dtype='int64', length=1853),
57.0: Int64Index([ 13217, 66638, 121327, 122878, 123049, 12501
3, 162500,
2135068, 2135229, 2135887,
...
5946666, 6272824, 6278014, 6293085, 6361003, 6384573, 6422
834,

```



```

        6442888, 6536709, 6555230],
        dtype='int64', length=996),
58.0: Int64Index([ 22913,  37848,  42727,  53315, 119905, 12524
0, 125456,
        126514, 126802, 127813,
        ...
        6525564, 6529478, 6579010, 6592706, 6592732, 6601960, 6608
069,
        6617613, 6648455, 6804576],
        dtype='int64', length=2757),
59.0: Int64Index([    46,  20936,  23066,  60078,  72840,  8827
3, 125822,
        130255, 135860, 141803,
        ...
        6280769, 6293195, 6340884, 6363542, 6446712, 6600391, 6661
179,
        6724743, 6756199, 6834639],
        dtype='int64', length=1083),
60.0: Int64Index([ 20467,  24122, 123878, 125768, 134290, 14037
3, 141135,
        143410, 150355, 155706,
        ...
        6287088, 6289597, 6401191, 6429344, 6493476, 6508164, 6517
865,
        6540708, 6561231, 6595855],
        dtype='int64', length=1615),
61.0: Int64Index([  4952,  16881,  22660,  23094,  26642,  2664
3,  28092,
        30533,  34206,  39452,
        ...
        6583253, 6587508, 6589253, 6594267, 6600327, 6620237, 6623
749,
        6648415, 6808339, 6815599],
        dtype='int64', length=4996),
62.0: Int64Index([  9947, 2016150, 2039361, 2123550, 2135085, 213612
1, 2136129,
        2136145, 2136390, 2136433,
        ...
        5945415, 5945568, 5946384, 5946719, 5946891, 6267711, 6296

```

```

644,
        6301824, 6335748, 6505919],
        dtype='int64', length=1037),
63.0: Int64Index([ 30550,  48421, 118242, 119910, 119968, 11996
9, 122419,
        122948, 124695, 128053,
        ...
        6383678, 6444007, 6454079, 6555294, 6557060, 6562993, 6572
028,
        6631264, 6736944, 6807558],
        dtype='int64', length=2365),
64.0: Int64Index([ 63554, 126275, 127266, 129972, 138129, 15137
0, 153042,
        175842, 2093225, 2094876,
        ...
        5945758, 5945923, 5946120, 5946762, 6377128, 6510362, 6529
736,
        6557033, 6635617, 6758385],
        dtype='int64', length=938),
65.0: Int64Index([ 41636,  67811, 120809, 121329, 121747, 12262
0, 123460,
        129869, 131636, 135359,
        ...
        6294871, 6299005, 6340896, 6439203, 6510390, 6585924, 6601
988,
        6674769, 6737474, 6788776],
        dtype='int64', length=2055),
66.0: Int64Index([  8599,  13280, 15683, 16171, 20430, 2087
8, 24124,
        30571,  31172,  31738,
        ...
        6594352, 6603840, 6617627, 6632951, 6663051, 6664515, 6666
353,
        6704755, 6705646, 6730073],
        dtype='int64', length=6220),
67.0: Int64Index([  123,  2793,  3338,  5555,  5579,  565
4,  6101,
        6232,  6615,  6683,
        ...

```

```

        6642604, 6650922, 6663809, 6674091, 6718911, 6738049, 6753
042,
        6769008, 6769031, 6805847],
        dtype='int64', length=7537),
68.0: Int64Index([ 6748, 7283, 8643, 8745, 9255, 926
8, 9278,
        10230, 12125, 12137,
        ...
        6710161, 6760700, 6760701, 6760723, 6769030, 6769032, 6779
513,
        6788809, 6827865, 6831450],
        dtype='int64', length=7243),
69.0: Int64Index([ 8710, 10220, 15805, 16327, 16743, 1682
4, 19937,
        21575, 23054, 23603,
        ...
        6658741, 6663810, 6698301, 6718268, 6720030, 6769028, 6788
787,
        6788807, 6788808, 6805324],
        dtype='int64', length=6774),
70.0: Int64Index([ 6778, 72347, 121587, 122953, 122997, 12344
1, 127944,
        128206, 135145, 137874,
        ...
        6451280, 6456316, 6481383, 6499535, 6524924, 6566317, 6637
374,
        6705556, 6722556, 6807459],
        dtype='int64', length=2315),
71.0: Int64Index([ 5704, 8714, 9698, 12642, 16746, 1687
4, 19943,
        28057, 45814, 48509,
        ...
        6633692, 6635615, 6636395, 6642544, 6655596, 6760724, 6769
004,
        6769007, 6790091, 6804517],
        dtype='int64', length=7844),
72.0: Int64Index([ 123308, 124483, 125345, 127484, 128913, 13201
2, 140809,
        146064, 1997645, 2006285,

```

```

        ...
        5946427, 5946760, 5946885, 5947189, 6316330, 6317122, 6618
275,
        6622921, 6636397, 6716806],
        dtype='int64', length=991),
73.0: Int64Index([ 63537, 119448, 119902, 120343, 122215, 12435
0, 125609,
        126366, 126808, 127747,
        ...
        6546698, 6590098, 6600331, 6611505, 6615003, 6616798, 6653
295,
        6705511, 6755081, 6809385],
        dtype='int64', length=3165),
74.0: Int64Index([ 119860, 121675, 125026, 131603, 162630, 16643
7, 2136739,
        2136742, 2137115, 2137680,
        ...
        5945266, 5945623, 5945737, 5945858, 6265917, 6269262, 6295
672,
        6349588, 6414834, 6754353],
        dtype='int64', length=618),
75.0: Int64Index([ 122060, 122423, 124334, 128410, 133029, 13303
0, 135385,
        140441, 151260, 156583,
        ...
        6380636, 6412579, 6417127, 6454691, 6487235, 6512592, 6551
823,
        6598543, 6622902, 6741830],
        dtype='int64', length=2087),
76.0: Int64Index([ 6188, 12606, 12638, 12750, 12752, 1276
7, 12778,
        13288, 13300, 13304,
        ...
        6389756, 6389761, 6394832, 6431592, 6435077, 6446685, 6455
238,
        6520439, 6632961, 6723289],
        dtype='int64', length=1746),
77.0: Int64Index([ 3521, 44795, 81949, 90368, 121022, 12203
4, 122113,

```

```
125713, 129303, 129905,  
...  
6377974, 6388889, 6389696, 6402688, 6405086, 6417145, 6440  
786,  
6527093, 6654001, 6805796],  
dtype='int64', length=2259}}
```

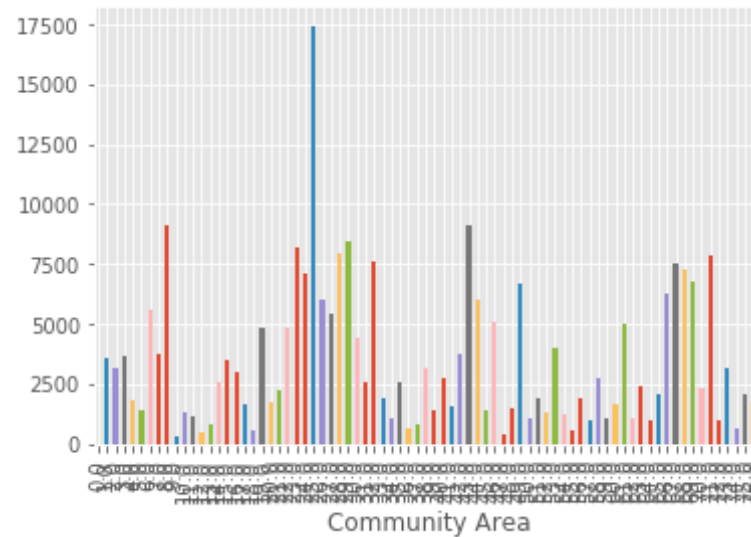
```
In [9]: community_crime_count_2015 = crimes_by_community_2015["ID"].agg("count"  
)
```

```
In [10]: crimes_by_community_2015["ID"].agg("count")
```

```
Out[10]: Community Area  
0.0      2  
1.0    3591  
2.0    3111  
3.0    3663  
4.0    1770  
5.0    1395  
6.0    5590  
7.0    3764  
8.0    9075  
9.0     258  
10.0   1278  
11.0   1157  
12.0    450  
13.0    843  
14.0   2536  
15.0   3453  
16.0   3019  
17.0   1660  
18.0    587  
19.0   4838  
20.0   1732  
21.0   2263  
22.0   4805  
23.0   8199  
24.0   7082  
25.0  17382
```

```
26.0    6029
27.0    5448
28.0    7948
29.0    8410
...
48.0    1484
49.0    6644
50.0    1048
51.0    1922
52.0    1318
53.0    4013
54.0    1189
55.0     516
56.0    1853
57.0     996
58.0    2757
59.0    1083
60.0    1615
61.0    4996
62.0    1037
63.0    2365
64.0     938
65.0    2055
66.0    6220
67.0    7537
68.0    7243
69.0    6774
70.0    2315
71.0    7844
72.0     991
73.0    3165
74.0     618
75.0    2087
76.0    1746
77.0    2259
Name: ID, Length: 78, dtype: int64
```

```
In [11]: community_crime_count_2015.plot(kind="bar");
```



En la gráfica se puede observar que el Community Area con mayor cantidad de casos de crímenes es el #25, Austin

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 [12]: freq_are = pd.crosstab(index = crimes_2015['Community Area'], columns =
      "count")
```

```
In [13]: freq_are.sort_values(["count"], ascending = False)
```

```
Out[13]:
```

```
col_0  count
```

Community Area	Weight	count
25.0	17382	
43.0	9099	
8.0	9075	
29.0	8410	
23.0	8199	
28.0	7948	
71.0	7844	
32.0	7595	
67.0	7537	
68.0	7243	
24.0	7082	
69.0	6774	
49.0	6644	
66.0	6220	
44.0	6043	
26.0	6029	
6.0	5590	
27.0	5448	
46.0	5086	
61.0	4996	
19.0	4838	
22.0	4805	
30.0	4452	

col_0	count
Community Area	
53.0	4013
7.0	3764
42.0	3752
3.0	3663
1.0	3591
15.0	3453
73.0	3165
...	...
76.0	1746
20.0	1732
17.0	1660
60.0	1615
41.0	1516
48.0	1484
5.0	1395
45.0	1390
39.0	1370
52.0	1318
10.0	1278
54.0	1189
11.0	1157
34.0	1084
59.0	1083

	col_0	count
Community Area		
	50.0	1048
	62.0	1037
	57.0	996
	72.0	991
	64.0	938
	13.0	843
	37.0	842
	36.0	632
	74.0	618
	18.0	587
	55.0	516
	12.0	450
	47.0	389
	9.0	258
	0.0	2

78 rows × 1 columns

```
In [14]: n_crimes = crimes_by_community_2015["ID"].agg("count")
n_crimes
```

```
Out[14]: Community Area
0.0      2
1.0    3591
2.0    3111
3.0    3663
4.0    1770
5.0    1395
```

6.0	5590
7.0	3764
8.0	9075
9.0	258
10.0	1278
11.0	1157
12.0	450
13.0	843
14.0	2536
15.0	3453
16.0	3019
17.0	1660
18.0	587
19.0	4838
20.0	1732
21.0	2263
22.0	4805
23.0	8199
24.0	7082
25.0	17382
26.0	6029
27.0	5448
28.0	7948
29.0	8410
	...
48.0	1484
49.0	6644
50.0	1048
51.0	1922
52.0	1318
53.0	4013
54.0	1189
55.0	516
56.0	1853
57.0	996
58.0	2757
59.0	1083
60.0	1615
61.0	4996

```
62.0    1037
63.0    2365
64.0     938
65.0    2055
66.0    6220
67.0    7537
68.0    7243
69.0    6774
70.0    2315
71.0    7844
72.0     991
73.0    3165
74.0     618
75.0    2087
76.0    1746
77.0    2259
Name: ID, Length: 78, dtype: int64
```

```
In [15]: n_crimes[n_crimes == n_crimes.max()]
```

```
Out[15]: Community Area
25.0    17382
Name: ID, dtype: int64
```

```
In [16]: n_crimes[n_crimes == n_crimes.min()]
```

```
Out[16]: Community Area
0.0      2
Name: ID, dtype: int64
```

Si no se cuenta el Community Area 0.0, entonces el de menos crímenes es el número 9, Edison Park, con 258 casos

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.

Date | CA1 CA2 CA3 CA77 01-01-2001| 10 20 02-01-2001| 9 56 03-01-2001| 10 20 .. 31-12-2015| 1 34

```
In [19]: # Create function to strip time from date field, and use it to create a
         # nother column
         def to_day(timestamp):
             return timestamp.replace(minute=0, hour=0, second=0)

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

```
In [20]: crimes["Day"]
```

```
Out[20]: 0      2004-04-14
         1      2004-05-02
         2      2003-11-24
         3      2004-04-24
         4      2004-05-03
         5      2004-05-01
         6      2004-05-01
         7      2004-05-03
         8      2004-04-27
         9      2004-05-04
        10      2004-05-02
        11      2004-05-03
        12      2004-04-28
        13      2004-05-04
        14      2004-04-14
        15      2004-05-03
        16      2004-05-03
        17      2004-04-15
        18      2004-05-01
```

19	2004-04-29
20	2004-04-15
21	2004-05-03
22	2004-05-03
23	2004-04-16
24	2004-03-27
25	2004-05-03
26	2004-04-17
27	2004-04-21
28	2004-05-03
29	2004-05-01

...

6834651	2004-05-01
6834652	2004-05-03
6834653	2004-05-03
6834654	2004-05-02
6834655	2004-05-03
6834656	2004-05-03
6834657	2004-05-02
6834658	2004-04-01
6834659	2004-04-05
6834660	2004-04-13
6834661	2004-04-05
6834662	2004-05-03
6834663	2004-05-03
6834664	2004-05-04
6834665	2004-05-03
6834666	2004-05-04
6834667	2004-04-15
6834668	2004-03-27
6834669	2004-04-29
6834670	2004-04-07
6834671	2004-05-02
6834672	2004-04-21
6834673	2004-05-03
6834674	2004-04-07
6834675	2004-05-02
6834676	2004-05-03
6834677	2004-03-24

```
6834678    2004-05-03
6834679    2004-04-12
6834680    2004-05-03
Name: Day, Length: 6834681, dtype: datetime64[ns]
```

```
In [21]: crimes["Day"] = crimes["Date"].dt.day
crimes
```

Out[21]:

	ID	Case Number	Date	Block	IUCR	Primary Type	Desc
0	3305587	HK301958	2004-04-14 19:45:00	009XX N LAWLER AVE	2024	NARCOTICS	POSS: HEROIN(V
1	3305589	HK338529	2004-05-02 10:45:00	081XX S KEDZIE AVE	0560	ASSAULT	S
2	3305592	HK336096	2003-11-24 00:01:00	021XX S HOMAN AVE	1562	SEX OFFENSE	AGG CRIMINAL S
3	3305593	HK322023	2004-04-24 12:35:00	005XX W 66TH ST	1821	NARCOTICS	MANU/DEL:CANNABIS OF
4	3305594	HK341518	2004-05-03 14:30:00	046XX S WHIPPLE ST	1320	CRIMINAL DAMAGE	TO VE
5	3305596	HK337559	2004-05-01 23:00:00	078XX S KOLMAR AVE	0520	ASSAULT	AGGRAVATED:KNIFE/CL
6	3305599	HK336504	2004-05-01 13:00:00	014XX W 13TH ST	1350	CRIMINAL TRESPASS	TO STATE SUP
7	3305600	HK340621	2004-05-03 15:00:00	062XX S WHIPPLE ST	051A	ASSAULT	AGGRAVATED: HAM
8	3305601	HK329059	2004-04-27 09:00:00	015XX W 13TH ST	2093	NARCOTICS	FOUND SUSPECT NARC

	ID	Case Number	Date	Block	IUCR	Primary Type	Desc
9	3305602	HK341462	2004-05-04 02:00:00	045XX N SHERIDAN RD	0820	THEFT	\$500 AND U
10	3305603	HK338835	2004-05-02 18:40:00	061XX S AUSTIN AVE	0560	ASSAULT	\$
11	3305605	HK340163	2004-05-03 13:40:00	025XX N CLARK ST	0560	ASSAULT	\$
12	3305607	HK330043	2004-04-28 12:10:00	014XX N LAKE SHORE DR	2028	NARCOTICS	POSS: SYNTHETIC D
13	3305608	HK341401	2004-05-04 01:07:00	009XX W LAKESIDE PL	0460	BATTERY	\$
14	3305609	HK302287	2004-04-14 22:55:00	056XX W MADISON ST	1811	NARCOTICS	POSS: CANNABIS 30G
15	3305610	HK340413	2004-05-03 14:15:00	0000X N STATE ST	0860	THEFT	RETAIL
16	3305614	HK339414	2004-05-03 05:42:00	076XX S CAMPBELL AVE	0486	BATTERY	DOMESTIC BATTERY S
17	3305618	HK303769	2004-04-15 18:03:00	053XX W MADISON ST	2092	NARCOTICS	SOLICIT NARCOTI PUBL
18	3305619	HK337261	2004-05-01 17:40:00	026XX S PULASKI RD	0325	ROBBERY	VEHICULAR HIJA
19	3305620	HK335358	2004-04-29 22:45:00	026XX W 80TH ST	0460	BATTERY	\$

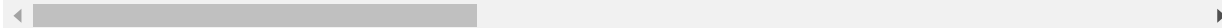
	ID	Case Number	Date	Block	IUCR	Primary Type	Desc
20	3305622	HK304038	2004-04-15 20:45:00	048XX W GLADYS AVE	2014	NARCOTICS	MANU/DELIVER: H (V
21	3305625	HK340513	2004-05-03 16:00:00	005XX W LAKE ST	0810	THEFT	OVE
22	3305626	HK340771	2004-05-03 19:10:00	043XX S LAMON AVE	0454	BATTERY	AGG PO HANDS M I
23	3305627	HK306070	2004-04-16 21:10:00	001XX S LAVERGNE AVE	2027	NARCOTICS	POSS: (
24	3305628	HK265638	2004-03-27 11:30:00	011XX N LECLAIRE AVE	2024	NARCOTICS	POSS: HEROIN(V
25	3305629	HK339278	2004-05-03 00:10:00	056XX S SAWYER AVE	0486	BATTERY	DOMESTIC BATTERY S
26	3305630	HK308269	2004-04-17 20:28:52	010XX N KEDZIE AVE	2027	NARCOTICS	POSS: (
27	3305631	HK332062	2004-04-21 13:39:00	0000X W WASHINGTON ST	5000	OTHER OFFENSE	OTHER CRIME AC PE
28	3305633	HK341513	2004-05-03 21:00:00	032XX N LINDER AVE	0910	MOTOR VEHICLE THEFT	AUTOM
29	3305636	HK340564	2004-05-01 13:00:00	013XX S BLUE ISLAND AVE	0620	BURGLARY	UNLAWFUL
...	
6834651	3305539	HK337130	2004-05-01 18:00:00	009XX N RICHMOND ST	0560	ASSAULT	§

	ID	Case Number	Date	Block	IUCR	Primary Type	Desc
6834652	3305540	HK340026	2004-05-03 12:15:00	017XX W SUPERIOR ST	1330	CRIMINAL TRESPASS	TC
6834653	3305541	HK341242	2004-05-03 22:45:46	010XX N RICHMOND ST	0560	ASSAULT	S
6834654	3305546	HK339116	2004-05-02 22:20:00	008XX W BELLE PLAINE AVE	0460	BATTERY	S
6834655	3305548	HK339841	2004-05-03 10:00:00	003XX S LEAVITT ST	1320	CRIMINAL DAMAGE	TO VE
6834656	3305549	HK339492	2004-05-03 08:01:00	020XX W WASHINGTON BLVD	1320	CRIMINAL DAMAGE	TO VE
6834657	3305551	HK339206	2004-05-02 22:35:00	035XX N BROADWAY	0460	BATTERY	S
6834658	3305552	HK276638	2004-04-01 18:00:00	037XX W DIVERSEY AVE	2091	NARCOTICS	FORFEIT PRO
6834659	3305554	HK282846	2004-04-05 02:05:00	045XX W LEXINGTON ST	2027	NARCOTICS	POSS: (
6834660	3305555	HK299545	2004-04-13 15:10:00	002XX S CANAL ST	2091	NARCOTICS	FORFEIT PRO
6834661	3305556	HK282931	2004-04-05 05:57:00	008XX S KOLMAR AVE	2027	NARCOTICS	POSS: (
6834662	3305558	HK339914	2004-05-03 10:27:00	013XX W HUBBARD ST	1200	DECEPTIVE PRACTICE	STOLEN BUY/RECEIV

	ID	Case Number	Date	Block	IUCR	Primary Type	Desc
6834663	3305561	HK339507	2004-05-03 08:15:00	045XX N HAZEL ST	0560	ASSAULT	S
6834664	3305562	HK341512	2004-05-04 00:00:00	028XX N LOWELL AVE	0915	MOTOR VEHICLE THEFT	TRUCK, BUS, MOTOR
6834665	3305563	HK340386	2004-05-03 15:37:00	0000X S OAKLEY BLVD	2825	OTHER OFFENSE	HARASSME TELEF
6834666	3305565	HK341369	2004-05-04 00:47:01	021XX S FAIRFIELD AVE	1090	ARSON	ATTEMPT A
6834667	3305566	HK303276	2004-04-15 11:35:00	057XX S CICERO AVE	2091	NARCOTICS	FORFEIT PRO
6834668	3305567	HK334802	2004-03-27 00:01:00	031XX N CLARK ST	1242	DECEPTIVE PRACTICE	COMPUTER I
6834669	3305569	HK332780	2004-04-29 17:00:00	064XX S MOZART ST	0560	ASSAULT	S
6834670	3305572	HK288817	2004-04-07 19:45:00	009XX N LAWLER AVE	2024	NARCOTICS	POSS: HEROIN(V
6834671	3305573	HK338062	2004-05-02 10:20:00	028XX W 65TH ST	0560	ASSAULT	S
6834672	3305574	HK316822	2004-04-21 21:23:00	051XX W MAYPOLE AVE	2017	NARCOTICS	MANU/DELIVER:(
6834673	3305575	HK339520	2004-05-03 08:25:00	012XX W WILSON AVE	0460	BATTERY	S

	ID	Case Number	Date	Block	IUCR	Primary Type	Desc
6834674	3305576	HK289217	2004-04-07 23:45:00	083XX S BALTIMORE AVE	1811	NARCOTICS	POSS: CANNABIS 30G
6834675	3305577	HK338421	2004-05-02 14:00:00	071XX S MAPLEWOOD AVE	0560	ASSAULT	S
6834676	3305579	HK339423	2004-05-03 05:40:00	002XX S LAFLIN ST	0486	BATTERY	DOMESTIC BATTERY S
6834677	3305580	HK259786	2004-03-24 16:30:00	041XX W MAYPOLE AVE	2024	NARCOTICS	POSS: HEROIN(\
6834678	3305581	HK339734	2004-05-03 10:15:00	046XX W 79TH ST	0330	ROBBERY	AGGRA
6834679	3305582	HK297447	2004-04-12 13:30:00	049XX W FERDINAND ST	2093	NARCOTICS	FOUND SUSPECT NARC
6834680	3305585	HK339990	2004-05-03 11:30:00	004XX N TROY ST	141C	WEAPONS VIOLATION	UNLAWFUL USE OTHER WI

6834681 rows × 23 columns



```
In [24]: crimes_2015.rename(columns={'Community Area': 'CommunityArea'}, inplace=True)
crimes_2015.head()
```

C:\Users\MARIO MONSALVE\Anaconda3\lib\site-packages\pandas\core\frame.py:3781: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
return super(DataFrame, self).rename(**kwargs)
```

Out[24]:

	ID	Case Number	Date	Block	IUCR	Primary Type	Description	Location Description
46	10072163	HY260736	2015-05-14 20:00:00	038XX S ASHLAND AVE	0810	THEFT	OVER \$500	S
123	10072550	HY252513	2015-05-08 17:44:00	015XX W 69TH ST	2017	NARCOTICS	MANU/DELIVER:CRACK	SID
1458	10075231	HY264086	2015-05-17 15:00:00	012XX W PRATT BLVD	0460	BATTERY	SIMPLE	
1596	10075512	HY264097	2015-05-17 15:15:00	067XX N WESTERN AVE	0430	BATTERY	AGGRAVATED: OTHER DANG WEAPON	SID
1764	10076372	HY261267	2015-05-14 14:10:00	065XX W BRYN MAWR AVE	0460	BATTERY	SIMPLE	SC F BU

5 rows × 22 columns



```
In [26]: df = pd.DataFrame(crimes_2015, columns = ['ID', 'Case Number', 'Date', 'Block', 'IUCR', 'Primary Type', 'Description', 'Location Description', 'Arrest', 'Domestic', 'CommunityArea', 'FBI Code', 'X Coordinate', 'Y Coordinate', 'Year', 'Updated On', 'Latitude', 'Longitude'])
df
```

Out[26]:

	ID	Case Number	Date	Block	IUCR	Primary Type	Description
46	10072163	HY260736	2015-05-14 20:00:00	038XX S ASHLAND AVE	0810	THEFT	OVER \$500

	ID	Case Number	Date	Block	IUCR	Primary Type	Description
123	10072550	HY252513	2015-05-08 17:44:00	015XX W 69TH ST	2017	NARCOTICS	MANU/DELIVER:CRAC
1458	10075231	HY264086	2015-05-17 15:00:00	012XX W PRATT BLVD	0460	BATTERY	SIMPL
1596	10075512	HY264097	2015-05-17 15:15:00	067XX N WESTERN AVE	0430	BATTERY	AGGRAVATED: OTHE DANG WEAPO
1764	10076372	HY261267	2015-05-14 14:10:00	065XX W BRYN MAWR AVE	0460	BATTERY	SIMPL
1822	10076562	HY258821	2015-05-13 15:00:00	026XX W SUMMERDALE AVE	1310	CRIMINAL DAMAGE	TO PROPERT
2793	10079467	HY210091	2015-04-04 11:14:35	055XX S HERMITAGE AVE	1811	NARCOTICS	POSS: CANNABI 30GMS OR LES
3291	10081005	HY216968	2015-04-10 14:20:00	058XX S INDIANA AVE	2014	NARCOTICS	MANU/DELIVER HEROIN (WHITE
3306	10081049	HY207873	2015-04-02 14:54:00	013XX N HUDSON AVE	2027	NARCOTICS	POSS: CRAC
3328	10081102	HY162782	2015-02-24 16:47:43	042XX W ADAMS ST	2024	NARCOTICS	POSS: HEROIN(WHITE
3338	10081113	HY163595	2015-02-25 13:46:33	062XX S HONORE ST	2014	NARCOTICS	MANU/DELIVER HEROIN (WHITE
3458	10081341	HY269702	2015-05-21 18:00:00	0000X E RANDOLPH ST	0810	THEFT	OVER \$50

	ID	Case Number	Date	Block	IUCR	Primary Type	Description
3521	10081459	HY269899	2015-05-10 23:00:00	012XX W BERWYN AVE	0820	THEFT	\$500 AND UNDE
4856	10084532	HY270896	2015-05-22 16:05:00	005XX E 35TH ST	0460	BATTERY	SIMPL
4952	10084715	HY273644	2015-05-24 18:44:00	046XX S DAMEN AVE	502R	OTHER OFFENSE	VEHICLE TITLE/RE/ OFFENS
5544	10086159	HY215057	2015-04-08 19:58:00	009XX N CENTRAL PARK AVE	143A	WEAPONS VIOLATION	UNLAWFUL POSS O HANDGU
5555	10086231	HY208946	2015-04-03 12:09:00	012XX W 72ND ST	1821	NARCOTICS	MANU/DEL:CANNABI 10GM OR LES
5557	10086263	HY209362	2015-04-03 17:10:00	007XX N ST LOUIS AVE	2014	NARCOTICS	MANU/DELIVER HEROIN (WHITE
5558	10086278	HY209341	2015-04-03 16:45:00	005XX N AVERS AVE	1821	NARCOTICS	MANU/DEL:CANNABI 10GM OR LES
5563	10086292	HY209287	2015-04-03 16:40:00	033XX W LEXINGTON ST	1821	NARCOTICS	MANU/DEL:CANNABI 10GM OR LES
5576	10086327	HY208975	2015-04-03 13:06:00	034XX W LAKE ST	2014	NARCOTICS	MANU/DELIVER HEROIN (WHITE
5579	10086351	HY208895	2015-04-03 11:15:00	059XX S ASHLAND AVE	2017	NARCOTICS	MANU/DELIVER:CRAC
5580	10086362	HY208819	2015-04-03 11:19:00	032XX W MAYPOLE AVE	2014	NARCOTICS	MANU/DELIVER HEROIN (WHITE

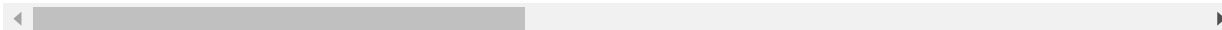
	ID	Case Number	Date	Block	IUCR	Primary Type	Description
5596	10086397	HY208805	2015-04-03 10:21:00	011XX S MOZART ST	2014	NARCOTICS	MANU/DELIVER HEROIN (WHITE
5654	10086564	HY269917	2015-05-21 21:15:00	062XX S JUSTINE ST	0486	BATTERY	DOMESTIC BATTER SIMPL
5704	10086765	HY230531	2015-04-21 12:39:00	078XX S HERMITAGE AVE	2017	NARCOTICS	MANU/DELIVER:CRAC
6101	10087770	HY175717	2015-03-07 19:51:42	062XX S LAFLIN ST	2024	NARCOTICS	POSS: HEROIN(WHITE
6188	10088067	HY186498	2015-03-16 10:30:00	100XX W OHARE ST	2091	NARCOTICS	FORFEIT PROPERT
6219	10088204	HY163990	2015-02-25 18:54:01	021XX N KILDARE AVE	2022	NARCOTICS	POSS: COCAIN
6230	10088266	HY233111	2015-04-23 15:55:00	008XX N CENTRAL PARK AVE	2017	NARCOTICS	MANU/DELIVER:CRAC
...
6806806	10119473	HY307894	2015-06-18 04:15:00	002XX N CANAL ST	0281	CRIM SEXUAL ASSAULT	NON-AGGRAVATE
6806807	10354586	HY545796	2015-12-21 20:00:00	109XX S DOTY AVE E	1150	DECEPTIVE PRACTICE	CREDIT CARD FRAU
6807250	11222762	JB140894	2015-10-01 00:01:00	022XX W MORSE AVE	1565	SEX OFFENSE	INDECEN SOLICITATION/CHIL
6807254	11224083	JB142898	2015-06-22 00:00:00	011XX N HAMLIN AVE	4651	OTHER OFFENSE	SEX OFFENDER: FAI REG NEW AD

	ID	Case Number	Date	Block	IUCR	Primary Type	Description
6807286	11224387	JB143451	2015-01-01 00:00:00	061XX S EBERHART AVE	1752	OFFENSE INVOLVING CHILDREN	AGG CRIM SEX ABUS FAM MEMBE
6807459	10000032	HY190005	2015-03-18 22:00:00	078XX S KEDZIE AVE	2024	NARCOTICS	POSS: HEROIN(WHITE
6807472	22094	HY422267	2015-09-14 02:11:00	008XX E BOWEN AVE	0110	HOMICIDE	FIRST DEGRE MURDE
6807526	11225228	JB144399	2015-12-12 00:00:00	091XX S WOODLAWN AVE	1153	DECEPTIVE PRACTICE	FINANCIAL IDENTIT THEFT OVER \$ 30
6807556	10000000	HY189952	2015-03-18 20:47:00	042XX N MARINE DR	1330	CRIMINAL TRESPASS	TO LAN
6807558	10000001	HY189865	2015-03-10 17:35:00	032XX W 55TH ST	1130	DECEPTIVE PRACTICE	FRAUD O CONFIDENCE GAM
6807698	10000410	HY190250	2015-03-18 21:00:00	023XX N ELSTON AVE	1320	CRIMINAL DAMAGE	TO VEHICL
6808339	10002687	HY191937	2015-03-20 12:00:00	046XX S DAMEN AVE	1150	DECEPTIVE PRACTICE	CREDIT CARD FRAU
6809385	10004656	HY194254	2015-03-01 00:00:00	092XX S WALLACE ST	1153	DECEPTIVE PRACTICE	FINANCIAL IDENTIT THEFT OVER \$ 30
6809617	10005477	HY190993	2015-03-19 18:30:00	003XX W NORTH AVE	0486	BATTERY	DOMESTIC BATTER SIMPL
6809628	10005513	HY174537	2015-03-06 20:36:00	028XX W WASHINGTON BLVD	2024	NARCOTICS	POSS: HEROIN(WHITE

	ID	Case Number	Date	Block	IUCR	Primary Type	Description
6809798	10005872	HY192562	2015-03-20 23:05:00	041XX N BELL AVE	0583	STALKING	CYBERSTALKING
6812114	10012492	HY194773	2015-03-22 22:30:00	032XX N SHEFFIELD AVE	0486	BATTERY	DOMESTIC BATTER SIMPL
6812548	10013348	HY202725	2015-03-29 14:12:00	009XX W MONTROSE AVE	0486	BATTERY	DOMESTIC BATTER SIMPL
6815599	10021537	HY211185	2015-04-05 11:55:00	046XX S DAMEN AVE	0860	THEFT	RETAIL THEF
6815778	10021894	HY208003	2015-04-02 08:00:00	015XX N HOYNE AVE	0620	BURGLARY	UNLAWFUL ENTR
6816507	10024019	HY213559	2015-04-07 14:15:00	031XX N WESTERN AVE	2825	OTHER OFFENSE	HARASSMENT B TELEPHON
6819922	10032520	HY211911	2015-04-06 01:17:00	125XX S EMERALD AVE	1310	CRIMINAL DAMAGE	TO PROPERT
6826331	10049611	HY180562	2015-03-11 12:25:00	009XX N HAMLIN AVE	1821	NARCOTICS	MANU/DEL:CANNABI 10GM OR LES
6826818	10050889	HY224865	2015-04-16 15:11:00	045XX W VAN BUREN ST	2031	NARCOTICS	POSS METHAMPHETAMINE
6826933	10051339	HY186540	2015-03-16 06:00:00	033XX W FILLMORE ST	1812	NARCOTICS	POSS: CANNABI MORE THAN 30GM
6827865	10054165	HY216204	2015-04-09 19:49:15	058XX S UNION AVE	1811	NARCOTICS	POSS: CANNABI 30GMS OR LES

	ID	Case Number	Date	Block	IUCR	Primary Type	Description
6827866	10054197	HY179008	2015-03-10 13:50:00	002XX S CANAL ST	2091	NARCOTICS	FORFEIT PROPERT
6831450	10063898	HY226113	2015-04-17 18:14:16	056XX S CARPENTER ST	2027	NARCOTICS	POSS: CRAC
6833122	10067511	HY253179	2015-05-09 09:30:00	081XX S COLFAX AVE	0495	BATTERY	AGGRAVATED OF SENIOR CITIZE
6834639	10072001	HY254030	2015-05-09 11:00:00	038XX S HONORE ST	0560	ASSAULT	SIMPL

264071 rows × 18 columns



In [27]: `pd.crosstab(df.CommunityArea, df.Date, margins=True)`

Out[27]:

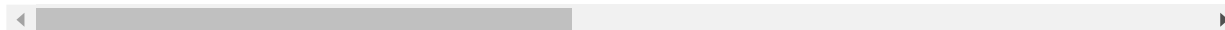
	2015-01-01 00:00:00	2015-01-01 00:01:00	2015-01-01 00:02:00	2015-01-01 00:03:00	2015-01-01 00:04:00	2015-01-01 00:05:00	2015-01-01 00:06:00	2015-01-01 00:07:00	2015-01-01 00:08:00
CommunityArea									
0.0	0	0	0	0	0	0	0	0	0
1.0	0	4	0	0	0	0	0	0	0
2.0	2	2	0	0	0	0	0	0	0
3.0	3	2	0	0	0	0	0	0	0
4.0	3	1	0	0	0	0	0	0	0
5.0	0	0	0	0	0	0	0	0	0
6.0	3	1	0	0	0	0	0	0	0
7.0	3	3	0	0	0	0	0	0	0

Date	2015-01-01 00:00:00	2015-01-01 00:01:00	2015-01-01 00:02:00	2015-01-01 00:03:00	2015-01-01 00:04:00	2015-01-01 00:05:00	2015-01-01 00:06:00	2015-01-01 00:07:00	2015-01-01 00:08:00
CommunityArea									
8.0	1	3	0	0	0	0	0	0	0
9.0	0	0	0	0	0	0	0	0	0
10.0	2	0	0	0	0	0	0	0	0
11.0	1	2	0	0	0	0	0	0	0
12.0	1	0	0	0	0	0	0	0	0
13.0	0	0	0	0	0	0	0	0	0
14.0	1	1	0	0	0	0	0	0	0
15.0	2	3	0	0	0	0	0	0	0
16.0	1	1	0	0	0	0	0	0	0
17.0	0	0	0	0	0	0	0	0	0
18.0	0	0	0	0	0	0	0	0	0
19.0	5	3	0	0	0	0	0	0	0
20.0	1	0	0	0	0	0	0	0	0
21.0	2	1	0	1	0	0	0	0	0
22.0	5	3	0	0	0	0	0	0	0
23.0	4	4	0	0	0	0	0	0	0
24.0	2	4	0	0	0	0	0	0	0
25.0	8	16	0	1	0	0	0	0	1
26.0	0	2	0	0	0	0	0	0	0
27.0	3	3	0	0	0	0	0	0	0
28.0	3	2	0	0	0	1	0	0	0
29.0	3	10	0	0	0	0	0	0	0
...

Date	2015-01-01 00:00:00	2015-01-01 00:01:00	2015-01-01 00:02:00	2015-01-01 00:03:00	2015-01-01 00:04:00	2015-01-01 00:05:00	2015-01-01 00:06:00	2015-01-01 00:07:00	2015-01-01 00:08:00
CommunityArea									
49.0	3	6	0	0	0	0	0	0	0
50.0	0	0	0	0	0	0	0	0	0
51.0	0	2	0	0	0	0	0	0	0
52.0	1	4	0	0	0	0	0	0	0
53.0	1	4	0	0	0	0	0	0	0
54.0	0	1	0	0	0	0	1	0	0
55.0	0	1	0	0	0	0	0	0	0
56.0	2	1	0	0	0	0	0	0	0
57.0	0	0	0	0	0	0	0	0	0
58.0	1	5	0	0	0	0	0	0	0
59.0	0	0	0	0	0	0	0	0	0
60.0	5	1	0	0	0	0	0	0	0
61.0	2	3	0	0	0	0	0	0	0
62.0	1	1	0	0	0	0	0	0	0
63.0	4	5	0	0	0	0	0	0	0
64.0	0	4	0	0	0	0	0	0	0
65.0	2	4	0	0	0	0	0	0	0
66.0	4	5	1	0	0	0	0	0	0
67.0	4	3	0	0	0	1	0	0	0
68.0	0	1	0	0	0	0	0	0	0
69.0	6	4	0	0	1	0	0	0	0
70.0	0	0	0	0	0	0	0	0	0
71.0	7	8	0	0	0	1	0	0	0

Date	2015-01-01 00:00:00	2015-01-01 00:01:00	2015-01-01 00:02:00	2015-01-01 00:03:00	2015-01-01 00:04:00	2015-01-01 00:05:00	2015-01-01 00:06:00	2015-01-01 00:07:00	2015-01-01 00:08:00
CommunityArea									
72.0	0	0	0	0	0	0	0	0	0
73.0	1	1	0	0	0	0	0	0	0
74.0	0	0	0	0	0	0	0	0	0
75.0	1	2	0	0	0	0	0	0	0
76.0	1	1	0	0	0	0	0	0	0
77.0	0	1	0	0	0	0	0	0	0
All	139	180	1	2	1	3	1	2	0

79 rows × 112240 columns



```
In [28]: crimes_by_community = crimes.groupby("Community Area")
```

```
In [29]: community_crime_count = crimes_by_community["ID"].agg("count")
```

```
In [30]: community_crime_count.to_frame()
```

Out[30]:

	ID
Community Area	
0.0	91
1.0	95307
2.0	77651
3.0	91025
4.0	43556
5.0	37324

Community Area		ID
6.0	123246	
7.0	97504	
8.0	213163	
9.0	6006	
10.0	26467	
11.0	24613	
12.0	11171	
13.0	20788	
14.0	55559	
15.0	79270	
16.0	70847	
17.0	38211	
18.0	14721	
19.0	114937	
20.0	37781	
21.0	58644	
22.0	131335	
23.0	196705	
24.0	183776	
25.0	395579	
26.0	115924	
27.0	117955	
28.0	183464	
29.0	179757	

Community Area		ID
...	...	
48.0		33621
49.0		166000
50.0		24953
51.0		40322
52.0		31037
53.0		102182
54.0		27375
55.0		13502
56.0		51894
57.0		22477
58.0		60881
59.0		25823
60.0		39943
61.0		128643
62.0		24174
63.0		57487
64.0		24574
65.0		46605
66.0		154121
67.0		182713
68.0		165992
69.0		153145
70.0		56674

	ID
Community Area	
71.0	176146
72.0	22442
73.0	73607
74.0	13975
75.0	49827
76.0	37079
77.0	61211

78 rows × 1 columns

```
In [31]: N_crimes = crimes_by_community["ID"].agg("count")
```

```
In [32]: freq = pd.crosstab(index = crimes['Community Area'], columns = "count")
```

```
In [33]: freq.T
```

Out[33]:

Community Area	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	...	68.0
col_0												
count	91	95307	77651	91025	43556	37324	123246	97504	213163	6006	...	165992

1 rows × 78 columns



```
In [34]: crimes["Day"][crimes["Community Area"]].to_frame()
```

Out[34]:

	Day
Community Area	

Day	
Community Area	
25.0	3.0
70.0	3.0
30.0	30.0
68.0	20.0
58.0	26.0
70.0	3.0
28.0	3.0
66.0	3.0
28.0	3.0
3.0	24.0
64.0	2.0
7.0	3.0
8.0	27.0
3.0	24.0
25.0	3.0
32.0	1.0
70.0	3.0
25.0	3.0
30.0	30.0
70.0	3.0
25.0	3.0
28.0	3.0
56.0	3.0
25.0	3.0

Community Area	Day	
25.0	3.0	
63.0	1.0	
23.0	16.0	
32.0	1.0	
15.0	3.0	
28.0	3.0	
...	...	
24.0	27.0	
24.0	27.0	
24.0	27.0	
3.0	24.0	
28.0	3.0	
28.0	3.0	
6.0	1.0	
22.0	3.0	
26.0	17.0	
28.0	3.0	
26.0	17.0	
24.0	27.0	
3.0	24.0	
20.0	15.0	
28.0	3.0	
30.0	30.0	
56.0	3.0	

Day	
Community Area	
6.0	1.0
66.0	3.0
25.0	3.0
66.0	3.0
25.0	3.0
3.0	24.0
46.0	14.0
66.0	3.0
28.0	3.0
26.0	17.0
70.0	3.0
25.0	3.0
23.0	16.0

6834681 rows × 1 columns

```
In [35]: crimes_by_day = crimes.groupby("Day")
```

```
In [36]: crimes_by_day["ID"].agg("count")
```

```
Out[36]: Day
1      268349
2      221680
3      221055
4      220151
5      222934
6      222644
7      222658
8      222706
```

```
9      223164
10     226313
11     222791
12     224468
13     222559
14     225563
15     232586
16     224963
17     225702
18     224135
19     223295
20     228481
21     223773
22     221278
23     222267
24     219143
25     216790
26     218935
27     220746
28     221943
29     208605
30     205468
31     129536
Name: ID, dtype: int64
```

```
In [37]: crimes_by_community["Day"].agg("count")
```

```
Out[37]: Community Area
0.0      91
1.0    95307
2.0    77651
3.0    91025
4.0    43556
5.0    37324
6.0   123246
7.0    97504
8.0   213163
9.0     6006
10.0   26467
```

11.0	24613
12.0	11171
13.0	20788
14.0	55559
15.0	79270
16.0	70847
17.0	38211
18.0	14721
19.0	114937
20.0	37781
21.0	58644
22.0	131335
23.0	196705
24.0	183776
25.0	395579
26.0	115924
27.0	117955
28.0	183464
29.0	179757
	...
48.0	33621
49.0	166000
50.0	24953
51.0	40322
52.0	31037
53.0	102182
54.0	27375
55.0	13502
56.0	51894
57.0	22477
58.0	60881
59.0	25823
60.0	39943
61.0	128643
62.0	24174
63.0	57487
64.0	24574
65.0	46605
66.0	154121

```
67.0    182713
68.0    165992
69.0    153145
70.0     56674
71.0    176146
72.0     22442
73.0     73607
74.0     13975
75.0     49827
76.0     37079
77.0      61211
Name: Day, Length: 78, dtype: int64
```

```
In [38]: pd.crosstab(index = crimes['Community Area'], columns = "Day")
```

```
Out[38]:
```

col_0	Day
Community Area	
0.0	91
1.0	95307
2.0	77651
3.0	91025
4.0	43556
5.0	37324
6.0	123246
7.0	97504
8.0	213163
9.0	6006
10.0	26467
11.0	24613
12.0	11171

col_0	Day
Community Area	
13.0	20788
14.0	55559
15.0	79270
16.0	70847
17.0	38211
18.0	14721
19.0	114937
20.0	37781
21.0	58644
22.0	131335
23.0	196705
24.0	183776
25.0	395579
26.0	115924
27.0	117955
28.0	183464
29.0	179757
...	...
48.0	33621
49.0	166000
50.0	24953
51.0	40322
52.0	31037
53.0	102182

col_0	Day
Community Area	
54.0	27375
55.0	13502
56.0	51894
57.0	22477
58.0	60881
59.0	25823
60.0	39943
61.0	128643
62.0	24174
63.0	57487
64.0	24574
65.0	46605
66.0	154121
67.0	182713
68.0	165992
69.0	153145
70.0	56674
71.0	176146
72.0	22442
73.0	73607
74.0	13975
75.0	49827
76.0	37079
77.0	61211

78 rows × 1 columns

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>).

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 [ ]: socio_eco=pd.read_csv("Census_Data_-_Selected_socioeconomic_indicators_
in
_Chicago__2008__2012.csv")
socioec.head(30)
```
