Untitled

October 15, 2024

```
[ ]: !pip install pandas
     import pandas as pd
     import sqlite3 as sq
```

```
[8]: conn = sq.connect('census.db')
     corsor = conn.cursor()
```

```
[ ]: !pip install ipython-sql
     %load_ext sql
```

```
[13]: %sql sqlite:///census.db
```

```
[14]: path = r'https://data.cityofchicago.org/resource/jcxq-k9xf.csv'
      data = pd.read_csv(path)
      data
```

```
[14]:      ca 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
     ..   …                    …                           …
     73  74.0      Mount Greenwood                         1.0
     74  75.0          Morgan Park                         0.8
     75  76.0               O'Hare                         3.6
     76  77.0            Edgewater                         4.1
     77   NaN              CHICAGO                         4.7

         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
     ..                               …                           …
     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
..                                             ...
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
..                                 ...                  ...             ...
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]
```

[15]: `data.to_sql('sensus', conn, if_exists="replace", index=False)`

[15]: 78

[17]: `%sql select * from sensus limit 5;`

```
 * sqlite:///census.db
Done.
```

[17]: [(1.0, 'Rogers Park', 7.7, 23.6, 8.7, 18.2, 27.5, 23939, 39.0),
     (2.0, 'West Ridge', 7.8, 17.2, 8.8, 20.8, 38.5, 23040, 46.0),
     (3.0, 'Uptown', 3.8, 24.0, 8.9, 11.8, 22.2, 35787, 20.0),
     (4.0, 'Lincoln Square', 3.4, 10.9, 8.2, 13.4, 25.5, 37524, 17.0),
     (5.0, 'North Center', 0.3, 7.5, 5.2, 4.5, 26.2, 57123, 6.0)]

```

```
[18]: %sql select count(*) from sensus;
```

```
 * sqlite:///census.db
Done.
```

```
[18]: [(78,)]
```

```
[21]: %%sql
      select count(*) as total_hardship_index from sensus where hardship_index > 50;
```

```
 * sqlite:///census.db
Done.
```

```
[21]: [(38,)]
```

```
[22]: %sql select max(hardship_index) from sensus;
```

```
 * sqlite:///census.db
Done.
```

```
[22]: [(98.0,)]
```

```
[29]: %sql select community_area_name from sensus order by hardship_index desc limit⏎
      ↪1;
      %sql select community_area_name from sensus where hardship_index = (select⏎
      ↪max(hardship_index) from sensus);
```

```
 * sqlite:///census.db
Done.
 * sqlite:///census.db
Done.
```

```
[29]: [('Riverdale',)]
```

```
[30]: %sql select community_area_name, per_capita_income_ as per_capita_income from⏎
      ↪sensus where per_capita_income > 60000;
```

```
 * sqlite:///census.db
Done.
```

```
[30]: [('Lake View', 60058),
       ('Lincoln Park', 71551),
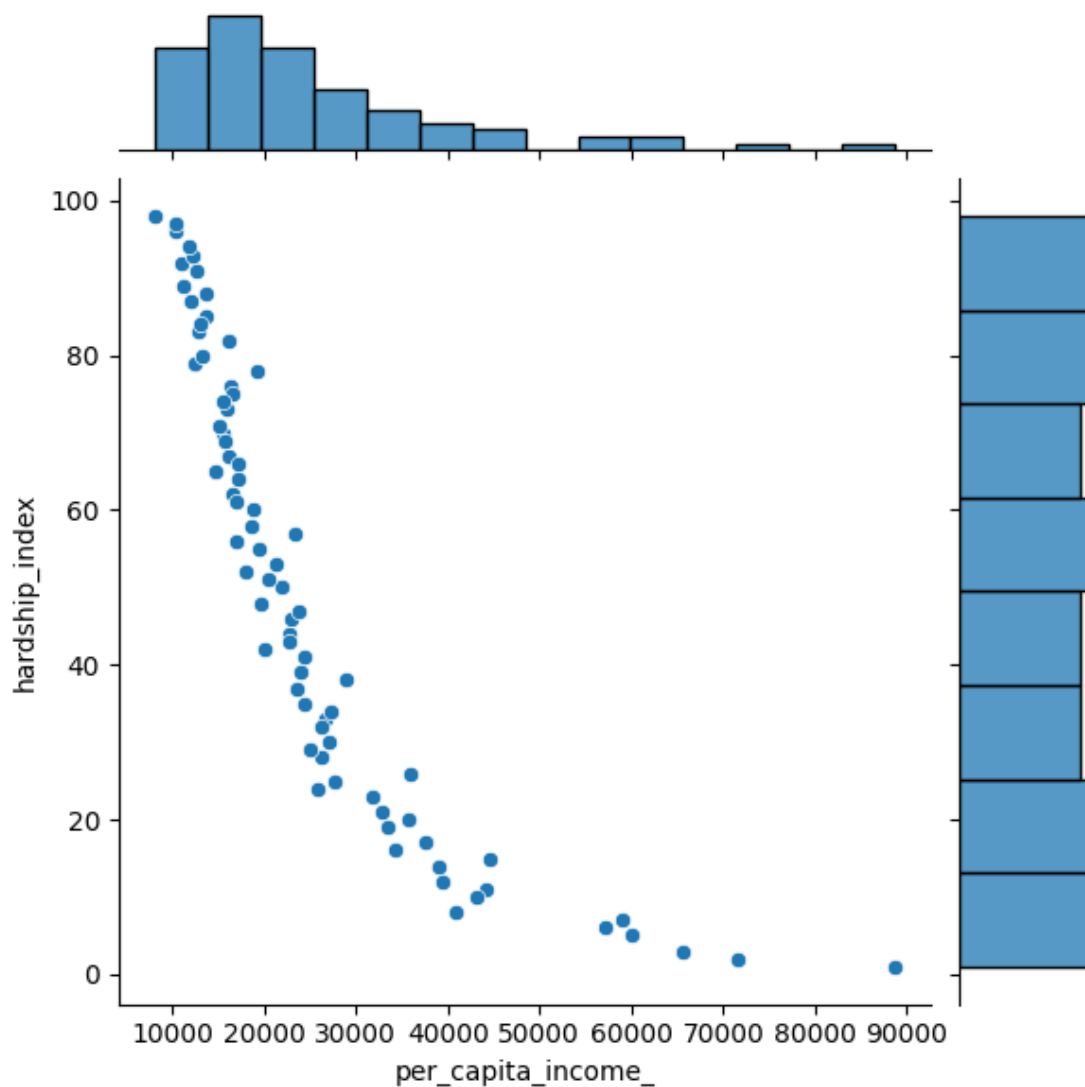       ('Near North Side', 88669),
       ('Loop', 65526)]
```

```
[ ]: !pip install matplotlib
     import matplotlib.pyplot as plt
     !pip install seaborn
```

[44]:
```
# %matplotlib inline
import seaborn as sns

table = %sql select per_capita_income_, hardship_index from sensus;
df = table.DataFrame()

g = sns.jointplot(data=df, x='per_capita_income_', y='hardship_index')
# plt.plot(g)
```

 * sqlite:///census.db
Done.

[ ]:

```
[ ]:
```

```
[ ]:
```

```
[ ]:
```

```
[ ]:
```

```
[ ]:
```

```
[ ]:
```

```
[ ]:
```

```
[ ]:
```