

merge code

November 22, 2024

```
[1]: import pandas as pd
```

```
[2]: df1 = pd.read_csv('Children_in_each_neighborhood.csv')
df2 = pd.read_csv('Single_parents_in_each_neighborhood.csv')
df3 = pd.read_csv('income.csv')
df4 = pd.read_csv('bachelors.csv')
df5 = pd.read_csv('non_movers.csv')
df1
```

```
[2]: postal_code      data \
0      M1B  2021A0011M1B - Children - Counts - Total
1      M9V  2021A0011M9V - Children - Counts - Total
2      M2J  2021A0011M2J - Children - Counts - Total
3      M2N  2021A0011M2N - Children - Counts - Total
4      M1E  2021A0011M1E - Children - Counts - Total
..      ...
91     M5J  2021A0011M5J - Children - Counts - Total
92     M5G  2021A0011M5G - Children - Counts - Total
93     M5E  2021A0011M5E - Children - Counts - Total
94     M5C  2021A0011M5C - Children - Counts - Total
95     M5H  2021A0011M5H - Children - Counts - Total
```

	Children in each neighbor	Children Points
0	24745	96
1	20320	95
2	17845	94
3	17785	93
4	16850	92
..
91	1775	5
92	1315	4
93	1145	3
94	355	2
95	280	1

[96 rows x 4 columns]

```
[3]: df2
```

```
[3]: postal_code      data \
0      M1B  2021A0011M1B - Total one-parent families - Cou...
1      M3N  2021A0011M3N - Total one-parent families - Cou...
2      M9V  2021A0011M9V - Total one-parent families - Cou...
3      M6M  2021A0011M6M - Total one-parent families - Cou...
4      M1E  2021A0011M1E - Total one-parent families - Cou...
..      ...
91     M5G  2021A0011M5G - Total one-parent families - Cou...
92     M4T  2021A0011M4T - Total one-parent families - Cou...
93     M5E  2021A0011M5E - Total one-parent families - Cou...
94     M5C  2021A0011M5C - Total one-parent families - Cou...
95     M5H  2021A0011M5H - Total one-parent families - Cou...
```

	Single parents	Single parents Points
0	4590	96
1	4520	95
2	3925	94
3	3840	93
4	3520	92
..
91	325	5
92	315	4
93	230	3
94	95	2
95	45	1

[96 rows x 4 columns]

```
[4]: df3
```

```
[4]: Unnamed: 0 postal_code      data \
0      56      M5E  2021A0011M5E - Median after-tax income in 2020...
1      83      M8X  2021A0011M8X - Median after-tax income in 2020...
2      48      M4T  2021A0011M4T - Median after-tax income in 2020...
3      44      M4N  2021A0011M4N - Median after-tax income in 2020...
4      38      M4G  2021A0011M4G - Median after-tax income in 2020...
..      ...      ...
91     94      M9V  2021A0011M9V - Median after-tax income in 2020...
92      3      M1G  2021A0011M1G - Median after-tax income in 2020...
93     33      M3N  2021A0011M3N - Median after-tax income in 2020...
94     14      M1V  2021A0011M1V - Median after-tax income in 2020...
95     39      M4H  2021A0011M4H - Median after-tax income in 2020...
```

	After-tax Income	After-tax Income Rank Score
0	56800	96

1	56800	96
2	56400	94
3	54400	93
4	54000	92
..
91	29400	6
92	29200	4
93	28400	3
94	27400	2
95	27000	1

[96 rows x 5 columns]

```
[5]: df4
```

```
[5]: Unnamed: 0 postal_code data \
0      22      M2N  2021A0011M2N - Bachelor's degree or higher - C...
1      66      M5V  2021A0011M5V - Bachelor's degree or higher - C...
2      18      M2J  2021A0011M2J - Bachelor's degree or higher - C...
3      53      M5A  2021A0011M5A - Bachelor's degree or higher - C...
4      52      M4Y  2021A0011M4Y - Bachelor's degree or higher - C...
..      ...      ...
91     31      M3L  2021A0011M3L - Bachelor's degree or higher - C...
92     58      M5H  2021A0011M5H - Bachelor's degree or higher - C...
93     30      M3K  2021A0011M3K - Bachelor's degree or higher - C...
94     89      M9L  2021A0011M9L - Bachelor's degree or higher - C...
95     55      M5C  2021A0011M5C - Bachelor's degree or higher - C...
```

	Number of bachlors degreess	Education Score
0	39350	96
1	39245	95
2	24985	94
3	21690	93
4	20815	92
..
91	2380	5
92	2270	4
93	1965	3
94	1960	2
95	1770	1

[96 rows x 5 columns]

```
[14]: df5['mobility_score'] =df5['Score']
df5
```

```
[14]:      Unnamed: 0 postal_code      data \
0          22          M2N  2021A0011M2N - Non-movers - Counts - Total
1           0          M1B  2021A0011M1B - Non-movers - Counts - Total
2          18          M2J  2021A0011M2J - Non-movers - Counts - Total
3         101          M9V  2021A0011M9V - Non-movers - Counts - Total
4          14          M1V  2021A0011M1V - Non-movers - Counts - Total
..      ...      ...
98         55          M5C  2021A0011M5C - Non-movers - Counts - Total
99         58          M5H  2021A0011M5H - Non-movers - Counts - Total
100        86          M7R  2021A0011M7A - Non-movers - Counts - Total
101        87          M7Y  2021A0011M7A - Non-movers - Counts - Total
102        85          M7A  2021A0011M7A - Non-movers - Counts - Total

      value  Score  mobility_score
0    59225.0    95             95
1    59145.0    94             94
2    51585.0    93             93
3    48560.0    92             92
4    45655.0    91             91
..      ...      ...
98    2310.0     3              3
99    1775.0     2              2
100     0.0     1              1
101     0.0     1              1
102     0.0     1              1

[103 rows x 6 columns]
```

```
[15]: result = pd.merge(df1, df2, on='postal_code', how='outer', suffixes=('_df1',
↳ '_df2'))
result = pd.merge(result, df3, on='postal_code', how='outer', suffixes=('',
↳ '_df3'))
result = pd.merge(result, df4, on='postal_code', how='outer', suffixes=('',
↳ '_df4'))
result = pd.merge(result, df5, on='postal_code', how='outer', suffixes=('',
↳ '_df5'))
```

```
[16]: result
df= result
df
```

```
[16]:      postal_code      data_df1 \
0          M1B  2021A0011M1B - Children - Counts - Total
1          M9V  2021A0011M9V - Children - Counts - Total
2          M2J  2021A0011M2J - Children - Counts - Total
3          M2N  2021A0011M2N - Children - Counts - Total
4          M1E  2021A0011M1E - Children - Counts - Total
```

..
98	M5L	NaN
99	M5K	NaN
100	M7R	NaN
101	M7Y	NaN
102	M7A	NaN

	Children in each neighbor	Children Points \
0	24745.0	96.0
1	20320.0	95.0
2	17845.0	94.0
3	17785.0	93.0
4	16850.0	92.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	data_df2	Single parents \
0	2021A0011M1B - Total one-parent families - Cou...	4590.0
1	2021A0011M9V - Total one-parent families - Cou...	3925.0
2	2021A0011M2J - Total one-parent families - Cou...	3295.0
3	2021A0011M2N - Total one-parent families - Cou...	3355.0
4	2021A0011M1E - Total one-parent families - Cou...	3520.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	Single parents Points	Unnamed: 0 \
0	96.0	0.0
1	94.0	94.0
2	89.0	18.0
3	90.0	22.0
4	92.0	2.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

data After-tax Income \

0	2021A0011M1B - Median after-tax income in 2020...	31600.0
1	2021A0011M9V - Median after-tax income in 2020...	29400.0
2	2021A0011M2J - Median after-tax income in 2020...	33200.0
3	2021A0011M2N - Median after-tax income in 2020...	36400.0
4	2021A0011M1E - Median after-tax income in 2020...	32800.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	After-tax Income Rank Score	Unnamed: 0_df4 \
0	19.0	0.0
1	6.0	94.0
2	35.0	18.0
3	47.0	22.0
4	33.0	2.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	data_df4 \
0	2021A0011M1B - Bachelor's degree or higher - C...
1	2021A0011M9V - Bachelor's degree or higher - C...
2	2021A0011M2J - Bachelor's degree or higher - C...
3	2021A0011M2N - Bachelor's degree or higher - C...
4	2021A0011M1E - Bachelor's degree or higher - C...
..	...
98	NaN
99	NaN
100	NaN
101	NaN
102	NaN

	Number of bachlors degreess	Education Score	Unnamed: 0_df5 \
0	14160.0	82.0	0
1	9715.0	52.0	101
2	24985.0	94.0	18
3	39350.0	96.0	22
4	11120.0	62.0	2
..
98	NaN	NaN	61
99	NaN	NaN	60

100		NaN	NaN	86
101		NaN	NaN	87
102		NaN	NaN	85

		data_df5	value	Score \
0	2021A0011M1B - Non-movers - Counts - Total	59145.0	94	
1	2021A0011M9V - Non-movers - Counts - Total	48560.0	92	
2	2021A0011M2J - Non-movers - Counts - Total	51585.0	93	
3	2021A0011M2N - Non-movers - Counts - Total	59225.0	95	
4	2021A0011M1E - Non-movers - Counts - Total	42740.0	89	
..	
98	2021A0011M5J - Non-movers - Counts - Total	11065.0	14	
99	2021A0011M5J - Non-movers - Counts - Total	11065.0	14	
100	2021A0011M7A - Non-movers - Counts - Total	0.0	1	
101	2021A0011M7A - Non-movers - Counts - Total	0.0	1	
102	2021A0011M7A - Non-movers - Counts - Total	0.0	1	

	mobility_score
0	94
1	92
2	93
3	95
4	89
..	...
98	14
99	14
100	1
101	1
102	1

[103 rows x 20 columns]

```
[9]: df['final score'] = df['Children Points']*0.15+df['Single parents Points']*0.
    ↪05+df['After-tax Income Rank Score']*0.35+df['Education Score']*0.4+0.
    ↪1*df['mobility_score']
```

```
[10]: df
dfSort = df.sort_values(by = 'final score', ascending = False)
dfSort
```

	postal_code		data_df1 \
67	M5V	2021A0011M5V - Children - Counts - Total	
32	M6S	2021A0011M6S - Children - Counts - Total	
35	M6P	2021A0011M6P - Children - Counts - Total	
3	M2N	2021A0011M2N - Children - Counts - Total	
34	M5M	2021A0011M5M - Children - Counts - Total	
..	

98	M5L	NaN
99	M5K	NaN
100	M7R	NaN
101	M7Y	NaN
102	M7A	NaN

	Children in each neighbor	Children Points \
67	5705.0	29.0
32	10095.0	64.0
35	9180.0	61.0
3	17785.0	93.0
34	9265.0	62.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	data_df2	Single parents \
67	2021A0011M5V - Total one-parent families - Cou...	1340.0
32	2021A0011M6S - Total one-parent families - Cou...	1500.0
35	2021A0011M6P - Total one-parent families - Cou...	1660.0
3	2021A0011M2N - Total one-parent families - Cou...	3355.0
34	2021A0011M5M - Total one-parent families - Cou...	855.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	Single parents Points	Unnamed: 0 \
67	45.0	66.0
32	50.0	80.0
35	56.0	78.0
3	90.0	22.0
34	29.0	60.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	data	After-tax Income \
67	2021A0011M5V - Median after-tax income in 2020...	54000.0

32	2021A0011M6S - Median after-tax income in 2020...	49200.0
35	2021A0011M6P - Median after-tax income in 2020...	42800.0
3	2021A0011M2N - Median after-tax income in 2020...	36400.0
34	2021A0011M5M - Median after-tax income in 2020...	52400.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	After-tax Income Rank Score	Unnamed: 0_df4 \
67	92.0	66.0
32	83.0	80.0
35	71.0	78.0
3	47.0	22.0
34	88.0	60.0
..
98	NaN	NaN
99	NaN	NaN
100	NaN	NaN
101	NaN	NaN
102	NaN	NaN

	data_df4 \
67	2021A0011M5V - Bachelor's degree or higher - C...
32	2021A0011M6S - Bachelor's degree or higher - C...
35	2021A0011M6P - Bachelor's degree or higher - C...
3	2021A0011M2N - Bachelor's degree or higher - C...
34	2021A0011M5M - Bachelor's degree or higher - C...
..	...
98	NaN
99	NaN
100	NaN
101	NaN
102	NaN

	Number of bachlors degreess	Education Score	Unnamed: 0_df5 \
67	39245.0	95.0	68
32	15230.0	85.0	84
35	18145.0	91.0	82
3	39350.0	96.0	22
34	12470.0	72.0	62
..
98	NaN	NaN	61
99	NaN	NaN	60
100	NaN	NaN	86

101	NaN	NaN	87
102	NaN	NaN	85

		data_df5	value	Score	final score
67	2021A0011M5V - Non-movers - Counts - Total	37820.0	84	81.40	
32	2021A0011M6S - Non-movers - Counts - Total	30465.0	62	79.30	
35	2021A0011M6P - Non-movers - Counts - Total	33125.0	74	76.75	
3	2021A0011M2N - Non-movers - Counts - Total	59225.0	95	75.65	
34	2021A0011M5M - Non-movers - Counts - Total	23500.0	50	74.75	
..	
98	2021A0011M5J - Non-movers - Counts - Total	11065.0	14	NaN	
99	2021A0011M5J - Non-movers - Counts - Total	11065.0	14	NaN	
100	2021A0011M7A - Non-movers - Counts - Total	0.0	1	NaN	
101	2021A0011M7A - Non-movers - Counts - Total	0.0	1	NaN	
102	2021A0011M7A - Non-movers - Counts - Total	0.0	1	NaN	

[103 rows x 20 columns]

```
[11]: dfSort.to_csv('sort.csv')
```

```
[ ]:
```

```
[ ]:
```