pop2

July 30, 2021

[3]:
```
import csv
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
```

[26]:
```
data: [] = list()
home: [] = list()
away: object = None
result_name: str = ''
```

[27]:
```
#df = pd.read_csv('./data/202106_202106_    _ .csv', encoding='UTF-8',␣
↪thousands=',', index_col=0)
#df.to_csv('./data/202106_202106_    _ _wo_comma.csv', sep=',', na_rep='NaN')
data = csv.reader(open('./data/202106_202106_    _ _wo_comma.csv', 'rt',␣
↪encoding='UTF-8'))
next(data)
data = list(data)
```

[28]:
```
arr = []
[arr.append(int(j)) for i in data if ' ' in i[0] for j in i[3:]]
print([i for i in arr])
```

```
[16, 8, 11, 21, 13, 28, 19, 25, 24, 19, 22, 12, 10, 19, 22, 12, 14, 11, 30, 31,
45, 85, 73, 70, 81, 70, 65, 77, 89, 86, 79, 79, 71, 66, 68, 53, 65, 53, 67, 51,
60, 46, 50, 45, 46, 40, 44, 49, 49, 62, 51, 49, 70, 69, 60, 59, 62, 50, 64, 54,
70, 80, 74, 53, 57, 72, 68, 56, 44, 41, 37, 38, 43, 40, 34, 27, 29, 30, 41, 35,
32, 28, 22, 18, 17, 14, 23, 18, 11, 6, 6, 10, 9, 11, 4, 0, 2, 6, 2, 1, 9]
```

[29]:
```
plt.style.use('ggplot')
plt.plot(arr)
```

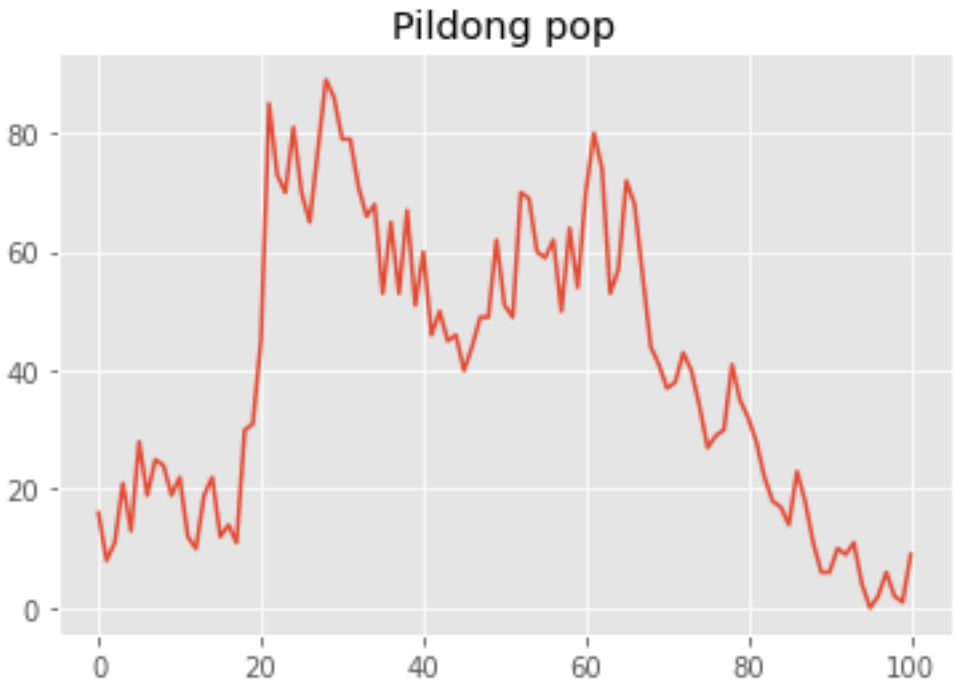
[29]:
```
[<matplotlib.lines.Line2D at 0x7ff44abe28e0>]
```

[30]:
```
home = []
[home.append(int(j)) for i in data if ' ' in i[0] for j in i[3:]]
print(home)
```

```
[16, 8, 11, 21, 13, 28, 19, 25, 24, 19, 22, 12, 10, 19, 22, 12, 14, 11, 30, 31,
45, 85, 73, 70, 81, 70, 65, 77, 89, 86, 79, 79, 71, 66, 68, 53, 65, 53, 67, 51,
60, 46, 50, 45, 46, 40, 44, 49, 49, 62, 51, 49, 70, 69, 60, 59, 62, 50, 64, 54,
70, 80, 74, 53, 57, 72, 68, 56, 44, 41, 37, 38, 43, 40, 34, 27, 29, 30, 41, 35,
32, 28, 22, 18, 17, 14, 23, 18, 11, 6, 6, 10, 9, 11, 4, 0, 2, 6, 2, 1, 9]
```

[31]:
```
plt.title('Pildong pop')
plt.plot(arr)
```

[31]:
```
[<matplotlib.lines.Line2D at 0x7ff4495ff1f0>]
```



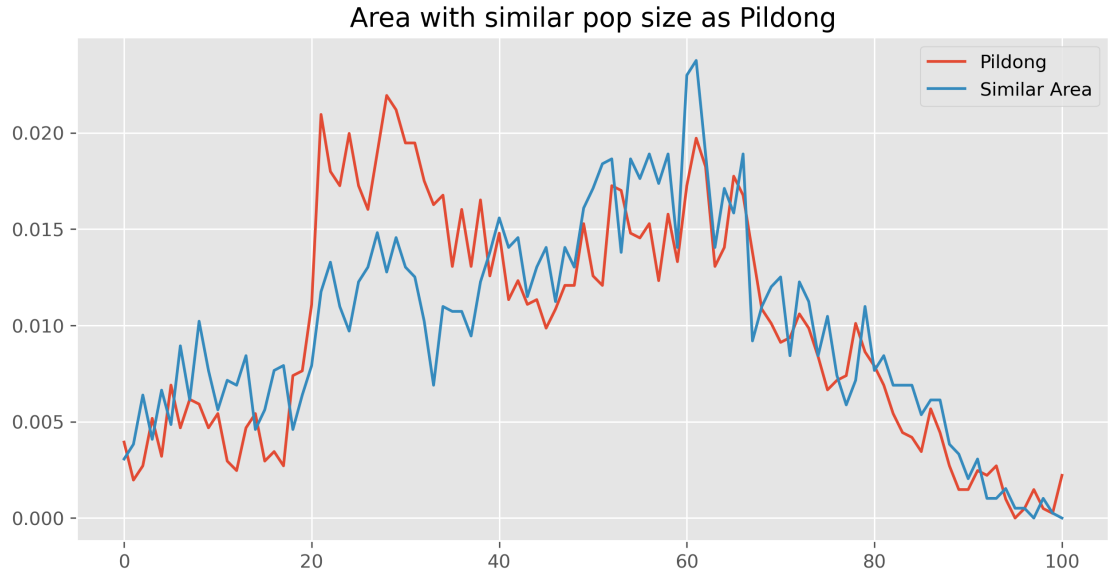
[54]:
```
home = []  #     self.home   local variable home
for i in data:
    if ' ' in i[0]:
        home = np.array(i[3:], dtype=int)/int(i[2])

away = []
result = 0
mn = 1  #
for i in data:
    away = np.array(i[3:], dtype=int)/int(i[2])
    s = np.sum((home-away)**2)
    if s < mn and ' ' not in i[0]:  # s < 1
        mn = s
        result_name = i[0]
        result = away
```

```
<ipython-input-54-861fea1535b9>:10: RuntimeWarning: invalid value encountered in
true_divide
  away = np.array(i[3:], dtype=int)/int(i[2])
```

[55]:
```
plt.style.use('ggplot')
plt.figure(figsize=(10, 5), dpi=300)
plt.title('Area with similar pop size as Pildong')
plt.plot(home, label='Pildong')
plt.plot(away, label='Similar Area')
```

```
plt.legend()
```

[55]: <matplotlib.legend.Legend at 0x7ff43b696340>



[ ]: