





- ☑ 특정지역 인구구조 연령별 성별 막대 그래프
- ☑ 임의지역 인구구조 연령별 성별 막대 그래프
- 赵 임의지역 인구구조 성별 원 그래프



### 학습목표

- ☑ 특정 지역의 인구구조 연령별 성별 막대 그래프를 그릴 수 있다.
- ☑ 임의 지역의 인구구조 연령별 성별 막대 그래프를 그릴 수 있다.
- ☑ 임의 지역의 인구구조 성별 원 그래프를 그릴 수 있다.

#### LESSON 01

## 특정지역 인구구조 연령별 성별 막대 그래프

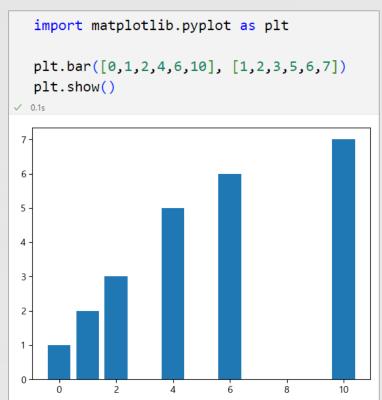


#### [1] 특정지역 인구구조 연령별 성별 막대 그래프



#### → 막대그래프 개요







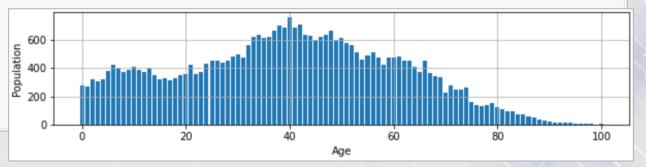
#### → 막대그래프 그리기 (1/2)



```
import matplotlib.pyplot as plt
import csv
f = open('age.csv', encoding='cp949')
data = csv.reader(f)
header = next(data)
result = []
for row in data:
    if '신도림' in row[0]:
        for i in row[3:1:
            result.append(int(i))
f.close()
plt.figure(figsize=(10, 2))
plt.bar(range(len(result)), result)
plt.xlabel('Age')
plt.ylabel('Population')
plt.grid(True)
plt.show()
```

- 목록 result에 다음을 저장
- 행에서 row[3:]

막대그래프를 수직이 아닌 수평 방향으로 그릴 수는 없을까요?



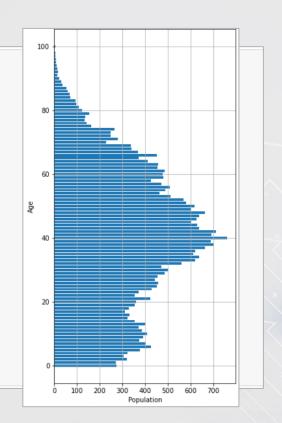


#### → 막대그래프 그리기 (2/2)

#### ◆ 수평 막대그래프를 활용하여 우리 동네의 연령별 인구수 데이터 시각화하기

**a** barh() 함수

```
import matplotlib.pyplot as plt
import csv
f = open('age.csv', encoding='cp949')
data = csv.reader(f)
header = next(data)
result = []
for row in data:
    if '신도림' in row[0]:
        for i in row[3:]:
            result.append(int(i))
f.close()
plt.figure(figsize=(5, 10))
plt.barh(range(len(result)), result)
plt.xlabel('Population')
plt.ylabel('Age')
plt.grid(True)
plt.show()
```

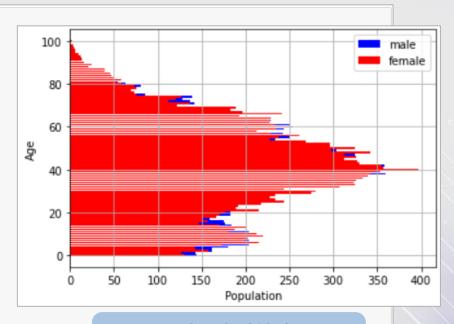




#### ⊸ 항아리 모양 그래프 그리기 (1/5)

#### ❤️ 수평 막대그래프를 활용 barh(): 남녀를 구분한 데이터 gender.csv 사용

```
import matplotlib.pyplot as plt
import csv
f = open('gender.csv', encoding='cp949')
data = csv.reader(f)
header = next(data)
male = []
female = []
for row in data:
    if '신도림' in row[0]:
        for i in row[3:104]:
            male.append(int(i))
        for i in row[106:207]:
                                  # row[106:]이라고 적어도 됩니다.
            female.append(int(i))
f.close()
plt.barh(range(len(male)), male, color='b', label='male')
plt.barh(range(len(female)), female, color='r', label='female')
plt.xlabel('Population')
plt.ylabel('Age')
plt.grid(True)
plt.legend()
plt.show()
```



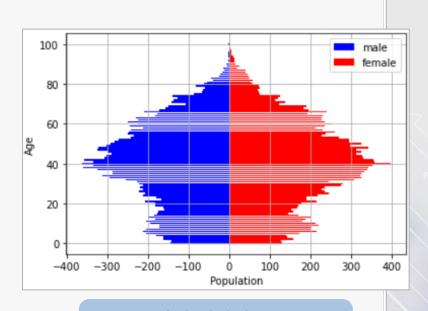
그래프가 겹쳐서 알아보기가 어렵습니다.



#### ⊸ 항아리 모양 그래프 그리기 (3/5)

#### **→ 남자는 음수, 여자는 양수, 수평 막대그래프를 활용**

```
import matplotlib.pyplot as plt
import csv
f = open('gender.csv', encoding='cp949')
data = csv.reader(f)
header = next(data)
male = []
female = []
for row in data:
    if '신도림' in row[0]:
        for i in row[3:104]:
            male.append(-int(i))
        for i in row[106:207]:
                                  # row[106:]이라고 적어도 됩니다.
            female.append(int(i))
f.close()
plt.barh(range(len(male)), male, color='b', label='male')
plt.barh(range(len(female)), female, color='r', label='female')
plt.xlabel('Population')
plt.ylabel('Age')
plt.grid(True)
plt.legend()
plt.show()
```



남성 데이터를 음수로 변경하였습니다.



#### ⊸ 항아리 모양 그래프 그리기 (4/5)

#### **◇** 찾고 싶은 지역의 이름을 입력 받아 성별 인구수 데이터 시각화하기

```
import csv
import matplotlib.pyplot as plt
f = open('gender.csv', encoding='cp949')
data = csv.reader(f)
header = next(data)
male = []
female = []
name = input('찾고 싶은 지역의 이름을 알려주세요: ')
area = ''
for row in data:
    if name in row[0] :
       area = row[0]
       for i in row[3:104] :
           male.append(-int(i))
       for i in row[106:]:
           female.append(int(i))
```

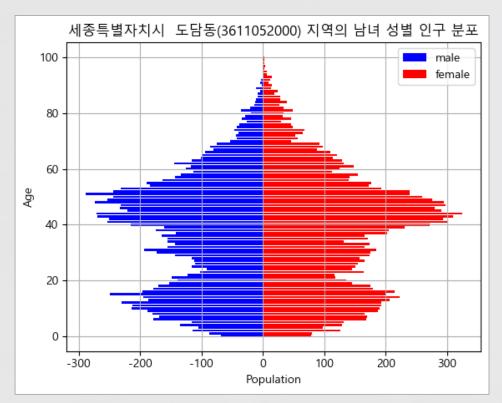
```
f.close()
plt.rc('font', family = 'Malgun Gothic')
plt.rc('axes', unicode_minus = False)

plt.barh(range(len(male)), male, color='b', label='male')
plt.barh(range(len(female)), female, color='r', label='female')
plt.title(f'{area} 지역의 남녀 성별 인구 분포')
plt.xlabel('Population')
plt.ylabel('Age')
plt.grid(True)
plt.legend()
plt.show()
```



→ 항아리 모양 그래프 그리기 (5/5)

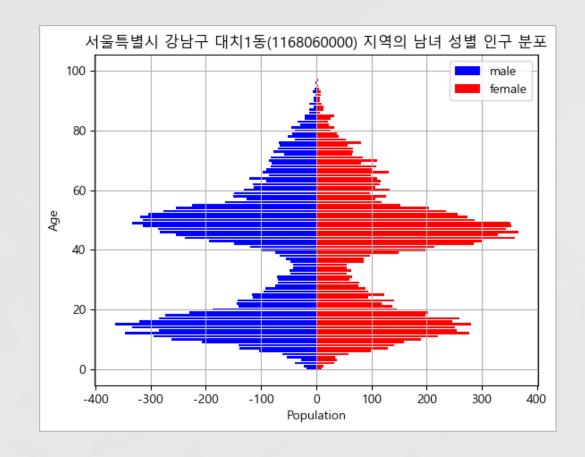
#### ❖ 찾고 싶은 지역의 이름을 입력 받아 성별 인구수 데이터 시각화하기(실행결과)



#### [21] 특정지역 인구구조 연령별 성별 막대 그래프



#### → 대치1동 사례



#### LESSON 02

## 지역 인구구조 연령별 성별 막대 그래프



#### ☑근 지역 인구구조 연령별 성별 막대 그래프



- ⊸ 제주도에는 여성의 비율이 더 높을까? (1/8)
- ☑ 예로부터 제주도를 돌, 바람, 여성이 많다 하여 삼다도(三多島)라고 불렀다고 합니다. 정말로 그럴까요? 정말인지 확인해보기 위해서 제주도의 성별 인구 비율을 비교해보겠습니다.



#### → 제주도에는 여성의 비율이 더 높을까? (2/8)

#### **❷** 제주특별자치도의 성별 분포를 항아리 모양 그래프로 표현하기

```
import matplotlib.pyplot as plt
import csv
f = open('gender.csv', encoding='cp949')
data = csv.reader(f)
header = next(data)
male = []
female = []
                                                         - ValueError가 발생 가능
name = input('찾고 싶은 지역의 이름을 알려주세요: ')
                                                         - 문자열 '1,934'에서 콤마 (,) 때문에 문제가 발생 예상
for row in data:
   if name in row[0]:
       for i in row[3:104]:
          male.append(-int(i))
       for i in row[106:207]:
                              # row[106:]이라고 적어도 됩니다.
          female.append(int(i))
f.close()
plt.rc('font', family='Malgun Gothic')
plt.rcParams['axes.unicode_minus'] = False
plt.title(name + ' 지역의 남녀 성별 인구 분포')
plt.barh(range(len(male)), male, color='b', label='male')
plt.barh(range(len(female)), female, color='r', label='female')
plt.xlabel('Population')
plt.ylabel('Age')
plt.grid(True)
plt.legend()
plt.show()
```



#### → 제주도에는 여성의 비율이 더 높을까? (3/8)

#### 🤪 제주특별자치도의 성별 분포를 항아리 모양 그래프로 표현하기(실행결과 및 원인분석)

```
ValueError Traceback (most recent call last)
~#AppData#Local#Temp/ipykernel 15000/3042654285.py in <module>
```

14 if name in row[0]: 15 for i in row[3:104]:

---> 16 male.append(-int(i))

찾고 싶은 지역의 이름을 알려주세요: 제주특별자치도

17 for i in row[106:207]: # row[106:]이라고 적어도 됩니다.

18 female.append(int(i))

ValueError: invalid literal for int() with base 10: '1.934'

ValueError가 발생하였습니다. 문자열 '1,934'에서 콤마 (,) 때문에 문제가 발생한 것으로 예상됩니다.

	A	В	С	D	E	F	G	Н	I	J	K	L	М	N
3818	제주특별자치도 (5000000000)	339,256	339,256	1,934	2,134	2,352	2,640	2,916	3,250	3,561	3,523	3,628	4,043	3,861
3819	제주특별자치도 제주시 (5011000000)	246,637	246,637	1,482	1,597	1,769	2,039	2,181	2,459	2,661	2,672	2,739	2,994	2,827
3820	제주특별자치도 제주시 한림읍(5011025000)	10,914	10,914	39	56	42	46	65	83	107	98	99	106	95
3821	제주특별자치도 제주시 애월읍(5011025300)	19,234	19,234	87	117	118	139	193	197	215	220	248	272	239
3822	제주특별자치도 제주시 구좌읍(5011025600)	7,904	7,904	21	33	37	29	44	45	41	63	56	79	77
3823	제주특별자치도 제주시 조천읍(5011025900)	13,048	13,048	69	73	73	104	96	124	152	131	142	172	155
3824	제주특별자치도 제주시 한경면(5011031000)	4,664	4,664	19	15	19	26	19	31	36	39	43	34	36
3825	제주특별자치도 제주시 추자면(5011032000)	860	860	2	2	4	2	1	2	1	4	6	4	7
3826	제주특별자치도 제주시 우도면(5011033000)	887	887	6	3	8	6	2	5	2	1	4	8	6
3827	제주특별자치도 제주시 일도1동(5011051000)	1,284	1,284	2	1	5	2	6	8	1	6	7	6	4
3828	제주특별자치도 제주시 일도2동(5011052000)	15,980	15,980	95	85	98	113	136	148	159	172	159	176	206
3829	제주특별자치도 제주시 이도1동(5011053000)	3,755	3,755	22	15	14	21	25	18	39	31	39	32	32
3830	제주특별자치도 제주시 이도2동(5011054000)	24,252	24,252	148	139	171	190	212	257	283	267	271	307	307
3831	제주특별자치도 제주시 삼도1동(5011055000)	6,626	6,626	30	35	46	42	37	64	64	57	64	81	66
3832	제주특별자치도 제주시 삼도2동(5011056000)	4,095	4,095	22	24	11	35	19	24	20	29	29	28	34
	제조트법지치는 제조기 8단4로/6044067000)	2 5 2 5	2 5 2 5	10	15	12	1.5	40	27	10	21	20	20	20



#### → 제주도에는 여성의 비율이 더 높을까? (5/8)

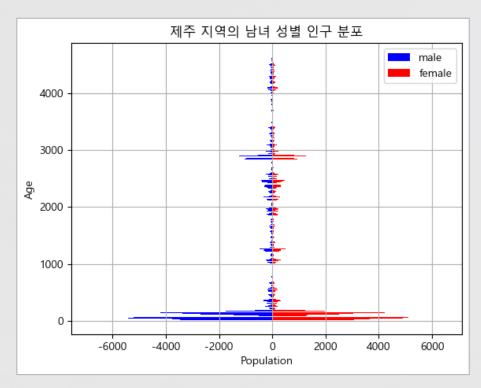
#### ❤️ replace() 함수를 활용하여 제주특별자치도의 성별 분포를 항아긴 모양 그래프로 표현하기

```
import matplotlib.pyplot as plt
import csv
f = open('gender.csv', encoding='cp949')
                                                 - replace() 함수를 활용하면 문자열에서 특정 문자를 제거 가능
data = csv.reader(f)
                                                 - replace('특정 문자', "") → 특정 문자를 제거
header = next(data)
male = []
female = []
name = input('찾고 싶은 지역의 이름을 알려주세요: ')
for row in data:
   if name in row[0]:
       for i in row[3:104]:
          i = i.replace('.', "")
          male.append(-int(i))
      for i in row[106:207]: # row[106:]이라고 적어도 됩니다.
          i = i.replace(',', "")
          female.append(int(i))
f.close()
plt.rc('font', family='Malgun Gothic')
plt.rcParams['axes.unicode_minus'] = False
plt.title(name + ' 지역의 남녀 성별 인구 분포')
plt.barh(range(len(male)), male, color='b', label='male')
plt.barh(range(len(female)), female, color='r', label='female')
plt.xlabel('Population')
plt.ylabel('Age')
plt.grid(True)
plt.legend()
plt.show()
```



#### → 제주도에는 여성의 비율이 더 높을까? (5/8)



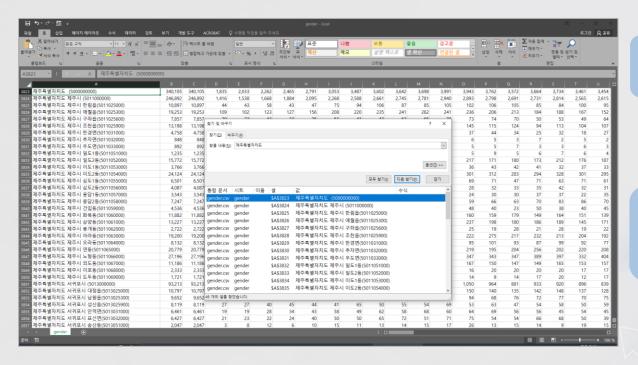


ValueError는 발생하지 않았지만, y축 Age 부분의 값이 엄청 많네요.



#### → 제주도에는 여성의 비율이 더 높을까? (6/8)

**→** replace() 함수를 활용하여 제주특별자치도의 성별 분포를 항아리 모양 그래프로 표현하기(원인분석)



"제주특별자치도"라는 문자열을 포함한 행이 무려 46개나 있네요.



즉, 46개의 데이터가 하나의 리스트에 저장되어 그래프로 표현되었던 거군요!



#### → 제주도에는 여성의 비율이 더 높을까? (7/8)

#### ❤️ break 문을 활용하여 제주특별자치도의 성별 분포를 항아리 모양 그래프로 표현하기

```
import matplotlib.pyplot as plt
import csv
f = open('gender.csv', encoding='cp949')
data = csv.reader(f)
header = next(data)
male = []
female = []
name = input('찾고 싶은 지역의 이름을 알려주세요: ')
for row in data:
    if name in row[0]:
       for i in row[3:104]:
           i = i.replace(',', "")
           male.append(-int(i))
       for i in row[106:207]: # row[106:]이라고 적어도 됩니다.
           i = i.replace(',', "")
           female.append(int(i))
        break
f.close()
plt.rc('font', family='Malgun Gothic')
plt.rcParams['axes.unicode_minus'] = False
plt.title(name + ' 지역의 남녀 성별 인구 분포')
plt.barh(range(len(male)), male, color='b', label='male')
plt.barh(range(len(female)), female, color='r', label='female')
plt.xlabel('Population')
plt.ylabel('Age')
plt.grid(True)
plt.legend()
plt.show()
```

사용자로부터 입력 받은 내용이 포함된 데이터 중 처음 만나는 데이터만 male, female 리스트에 저장하도록 하겠습니다.

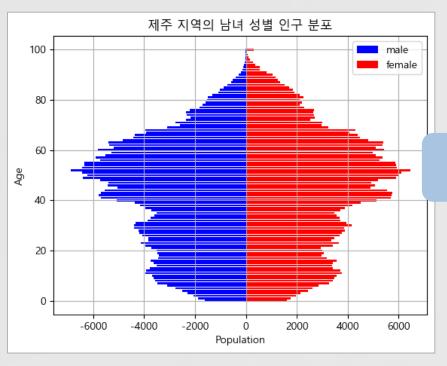


break 문을 추가하면 for 반복문을 멈출 수 있습니다.



#### → 제주도에는 여성의 비율이 더 높을까? (8/8)

#### ❤️ break 문을 활용하여 '제주'의 성별 분포를 항아리 모양 그래프로 표현하기(실행결과)



남성과 여성 중 어떤 성별의 비율이 높을까요? 항아리 모양의 그래프로는 알기 어려운 정보입니다.

#### LESSON 03

# 임민지역 인구구조 성별 원 그래프





#### → 제주도의 성별 인구 비율 표현하기 (1/5)

#### 🤪 제주특별자치도의 성별 인구 합계 구하기

```
import csv
f = open('gender.csv', encoding = 'cp949')
data = csv.reader(f)
header = next(data)
# print(header)
|name = input('궁금한 지역의 이름을 입력하세요: ')
for row in data:
   if (name in row[0]):
       male = 0
       female = 0
       for i in row[3:104]: # 3번 방에 0세, ..., 103번 방에 100세 이상
        i = i.replace(',', "") #'1,934' ---> '1934'
          male = male + int(i)
       for i in row[106:207]: # 106번 방에 0세, ..., 206번에 100세 이상
          i = i.replace(',', "") #'1,934' ---> '1934'
         female = female + int(i)
       break
f.close()
print('남성 인구 합: ', male, '명')
print('여성 인구 합: ', female, '명')
궁금한 지역의 이름을 입력하세요: 제주특별자치도
남성 인구 합: 339256 명
```

여성 인구 합: 337512 명



#### ⊸ 제주도의 성별 인구 비율 표현하기 (2/5)

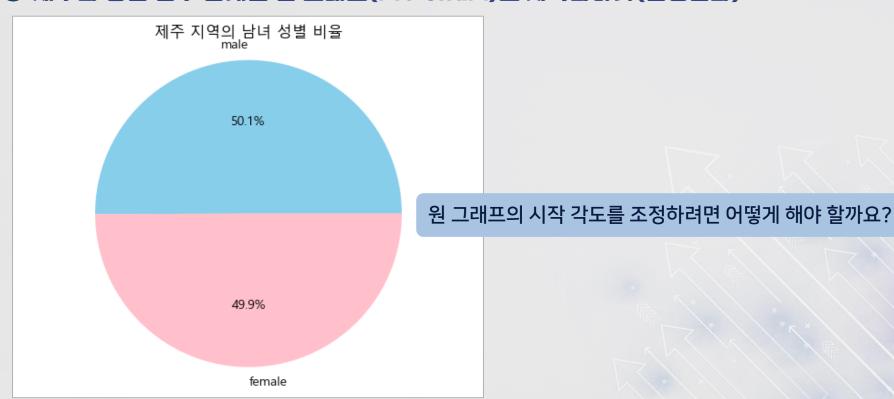
#### **● 제주특별자치도의 성별 인구 합계를 원 그래프(Pie Chart)로 시각화하기**

```
import csv
import matplotlib.pyplot as plt
f = open('gender.csv', encoding = 'cp949')
data = csv.reader(f)
header = next(data)
# print(header)
name = input('궁금한 지역의 이름을 입력하세요: ')
for row in data:
    if (name in row[0]):
       male = 0
       female = 0
       for i in row[3:104]: # 3번 방에 0세, ..., 103번 방에 100세 이상
          i = i.replace(',', "") #'1,934' ---> '1934'
           male = male + int(i)
       for i in row[106:207]: # 106번 방에 이세, ..., 206번에 100세 이상
           i = i.replace(',', "") #'1,934' ---> '1934'
           female = female + int(i)
       break
f.close()
labels = ['male', 'female']
colors = ['skyblue', 'pink']
plt.rc('font', family='Malgun Gothic')
plt.title(name + ' 지역의 성별 비율')
plt.pie([male, female], labels=labels, colors=colors, autopct='%,1f%%')
plt.show()
```



#### ⊸ 제주도의 성별 인구 비율 표현하기 (3/5)







#### → 제주도의 성별 인구 비율 표현하기 (4/5)



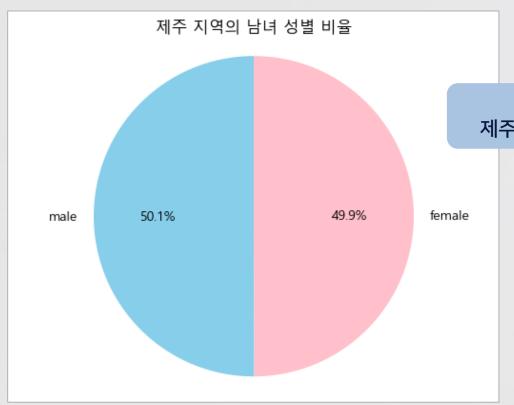
#### ❤️ 원그래프(Pie Chart)의 시작 각도 조정하기

```
import csv
import matplotlib.pyplot as plt
f = open('gender.csv', encoding = 'cp949')
data = csv.reader(f)
header = next(data)
                                                                                 startangle 속성에
# print(header)
                                                                          시작 각도를 지정하여 줍니다.
name = input('궁금한 지역의 이름을 입력하세요: ')
for row in data:
   if (name in row[0]):
      male = 0
      female = 0
      for i in row[3:104]: # 3번 방에 이세, ..., 103번 방에 100세 이상
                                                                                                                         startangle 90
        i = i.replace(',', "") #'1,934' ---> '1934'
         male = male + int(i)
      for i in row[106:207]: # 106번 방에 0세, ..., 206번에 100세 이상
          i = i.replace(',', "") #'1,934' ---> '1934'
          female = female + int(i)
      break
f.close()
labels = ['male', 'female']
colors = ['skyblue', 'pink']
                                                                                                                                    기본 시작 지점
plt.rc('font', family='Malgun Gothic')
plt.title(name + ' 지역의 성별 비율')
plt.pie([male, female], labels=labels, colors=colors, autopct='%.1f%%', startangle=90)
plt.show()
```



#### → 제주도의 성별 인구 비율 표현하기 (5/5)





확인 결과 2021년 기준으로 제주도의 남녀 인구 비율은 거의 동일하네요!

### SUMMARY

## 학습정긴





•••

- ❖ 행정안전부에서 인구 데이터를 내려 받기
  - >> age.xls, gender.xls
- 🧔 특정 지역의 연령별 인구 수 그래프
  - '지역명' in row[0]
  - >> 0세부터 100세 이상: row[3:]
- 🧔 표준입력으로 받은 지역의 연령별 인구 수 그래프

```
result = []
name = input('인구 구조가 알고 싶은 지역의 이름(읍면동 단위)을 입력해주세요 : ')
for row in data :
    if name in row[0] :
        for i in row[3:] :
            result.append(int(i))

f.close()
```



