

# 데이터분석입문

## Lecture 07. 인구 데이터와 그래프 그리기

동양미래대학교  
인공지능소프트웨어학과  
강 환수

- ❖ 01. 인구 데이터 준비하기
- ❖ 02. bar 함수로 막대그래프 그리기
- ❖ 03. pie 함수로 원그래프 그리기
- ❖ 04. scatter 함수로 산점도 그리기

# 01. 인구 데이터 준비하기

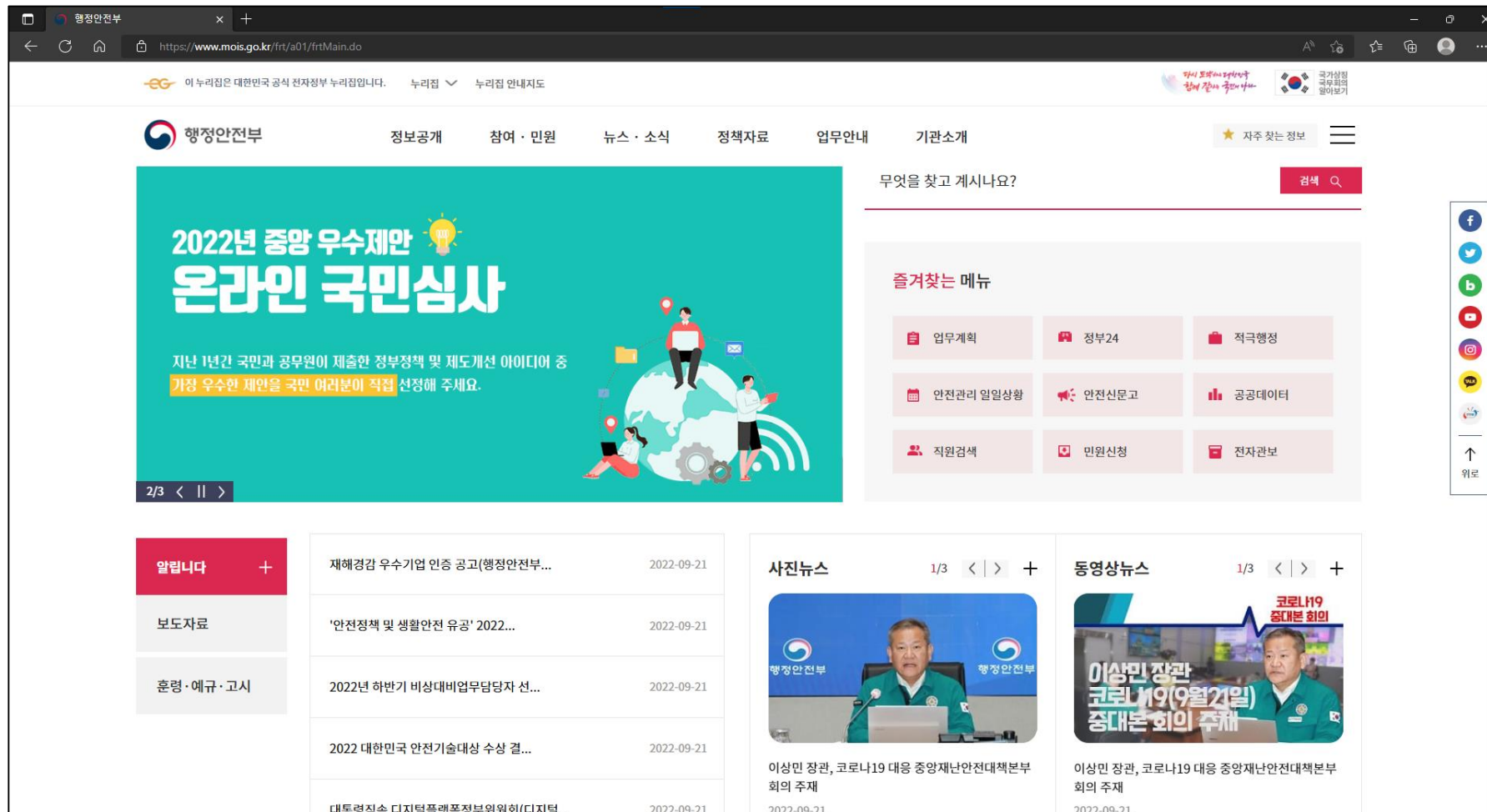
02. bar 함수로 막대그래프 그리기

03. pie 함수로 원그래프 그리기

04. scatter 함수로 산점도 그리기

## ❖ ① 인구 공공데이터 내려받기 (1/8)

- 인구 데이터 수집 → 행정안전부(<https://www.mois.go.kr/>)



## ❖ ① 인구 공공데이터 내려받기 (2/8)

- [정책자료] - [통계] - [주민등록 인구통계] 버튼 클릭



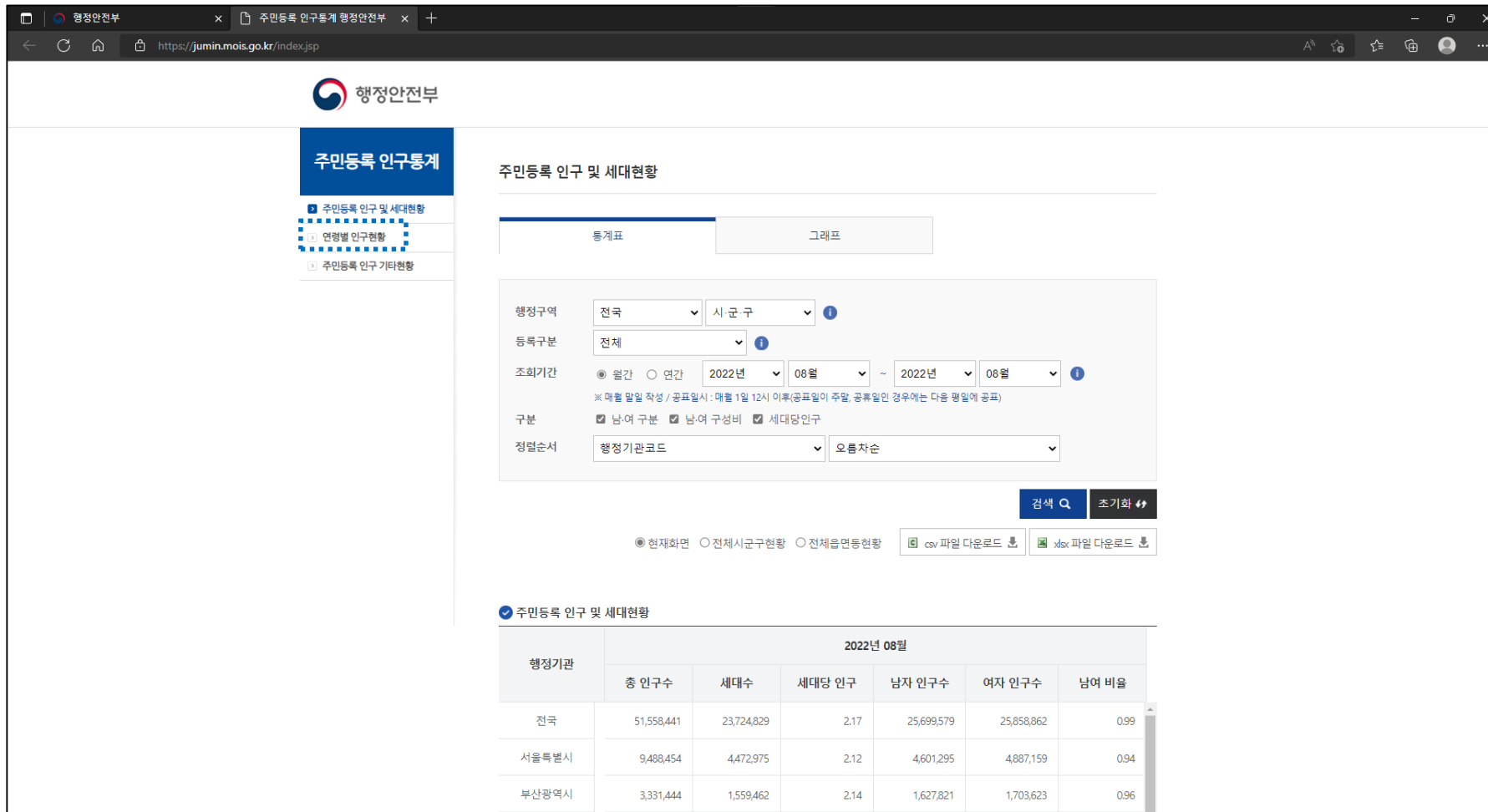
The screenshot shows the MOIS website with the following navigation path highlighted:

- 1. Policy Data (정책자료)
- 2. Statistics (통계)
- 3. Resident Registration Population Statistics (주민등록 인구통계)

The website also displays a sidebar with '알립니다' (Notice) and '보도자료' (Press Release) sections, and a main content area with '사진뉴스' (Photo News) and '동영상뉴스' (Video News) sections.

## ❖ ① 인구 공공데이터 내려받기 (3/8)

- [연령별 인구현황] 버튼 클릭



행정안전부

주민등록 인구통계

주민등록 인구 및 세대현황

통계표 | 그래프

행정구역: 전국 | 시·군·구

등록구분: 전체

조회기간: 월간 | 연간 | 2022년 | 08월 ~ 2022년 | 08월

※ 매월 말일 작성 / 공표일시: 매월 1일 12시 이후(공표일이 주말, 공휴일인 경우에는 다음 평일에 공표)

구분: 남·여 구분 | 남·여 구성비 | 세대당인구

정렬순서: 행정기관코드 | 오름차순

검색 | 초기화


현재화면 | 전체시군구현황 | 전체읍면동현황 | csv 파일 다운로드 | xlsx 파일 다운로드

주민등록 인구 및 세대현황

| 행정기관  | 2022년 08월  |            |        |            |            |       |
|-------|------------|------------|--------|------------|------------|-------|
|       | 총 인구수      | 세대수        | 세대당 인구 | 남자 인구수     | 여자 인구수     | 남여 비율 |
| 전국    | 51,558,441 | 23,724,829 | 2.17   | 25,699,579 | 25,858,862 | 0.99  |
| 서울특별시 | 9,488,454  | 4,472,975  | 2.12   | 4,601,295  | 4,887,159  | 0.94  |
| 부산광역시 | 3,331,444  | 1,559,462  | 2.14   | 1,627,821  | 1,703,623  | 0.96  |

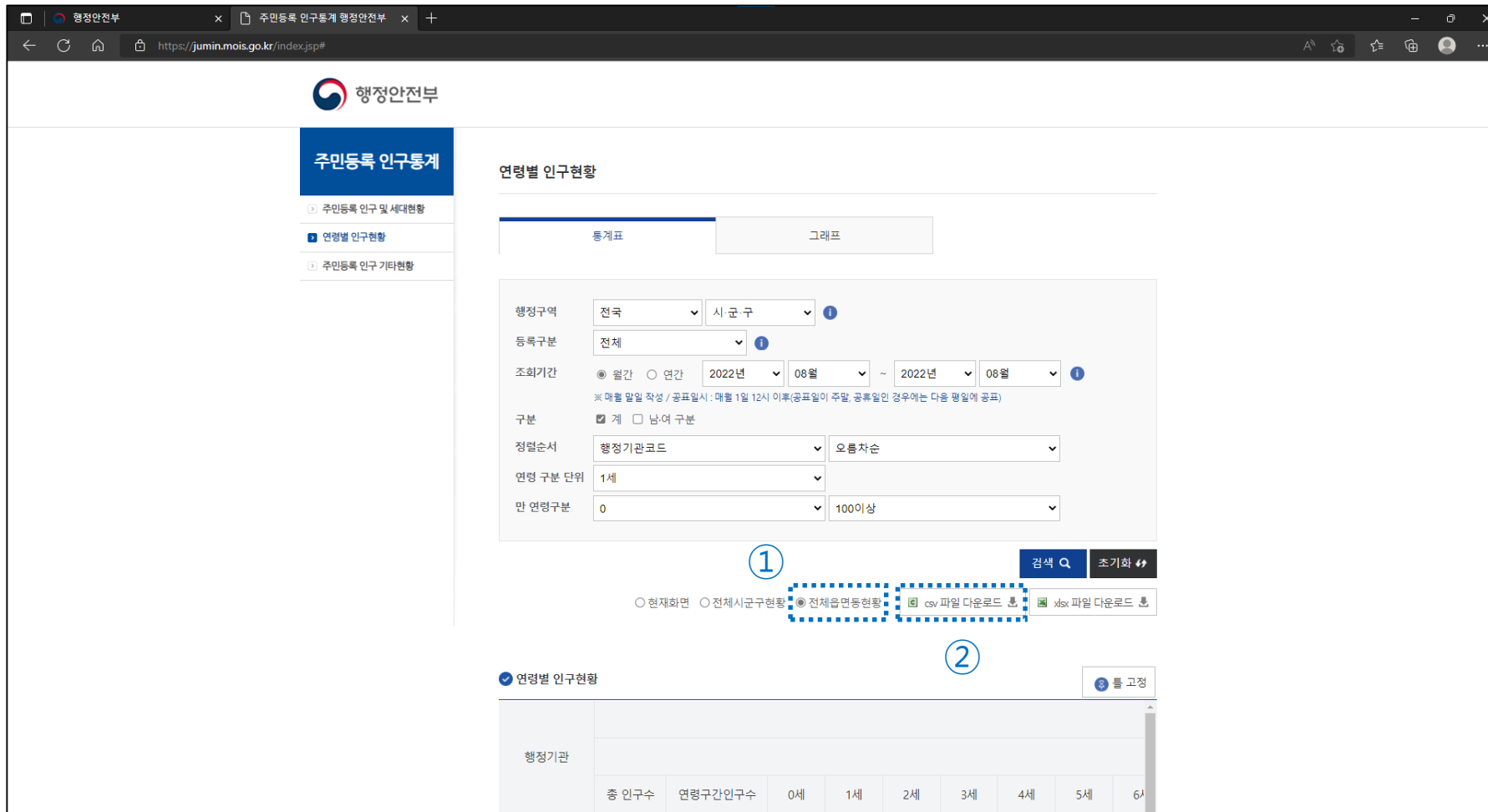
## ❖ ① 인구 공공데이터 내려받기 (4/8)

- [조회기간]: 2022년 08월 ~ 2022년 08월
- [남·여 구분] 체크 해제
- [연령 구분 단위]: 1세
- 만 연령구분: 0, 100이상
- [검색] 버튼 클릭



## ❖ ① 인구 공공데이터 내려받기 (5/8)

- [전체읍면동현황] 체크 후, [CSV 파일 다운로드] 버튼 클릭



행정안전부

주민등록 인구통계

연령별 인구현황

통계표 | 그래프

행정구역: 전국 | 시·군·구

등록구분: 전체

조회기간: 월간 | 연간 | 2022년 | 08월 ~ 2022년 | 08월

구분: ☒ 계 ☐ 남·여 구분

정렬순서: 행정기관코드 | 오름차순

연령 구분 단위: 1세

만 연령구분: 0 | 100이상

①

○ 현재화면 ○ 전체시군구현황 ● 전체읍면동현황 ○ CSV 파일 다운로드 ○ XLSX 파일 다운로드

②

● 연령별 인구현황

행정기관

| 총 인구수 | 연령구간인구수 | 0세 | 1세 | 2세 | 3세 | 4세 | 5세 | 6세 |
|-------|---------|----|----|----|----|----|----|----|
|-------|---------|----|----|----|----|----|----|----|



## ❖ ① 인구 공공데이터 내려받기 (6/8)

- 아래와 같이 경고 창이 뜨면 [확인] 버튼을 클릭하세요.

### jumin.mois.go.kr의 메시지

주민등록인구통계 시스템의 원활한 서비스 제공을 위해  
연령 구분 단위가 1세일 때는 남·여 구분없이 1개월 단위로만 다운로드  
가능합니다.

확인

### jumin.mois.go.kr의 메시지

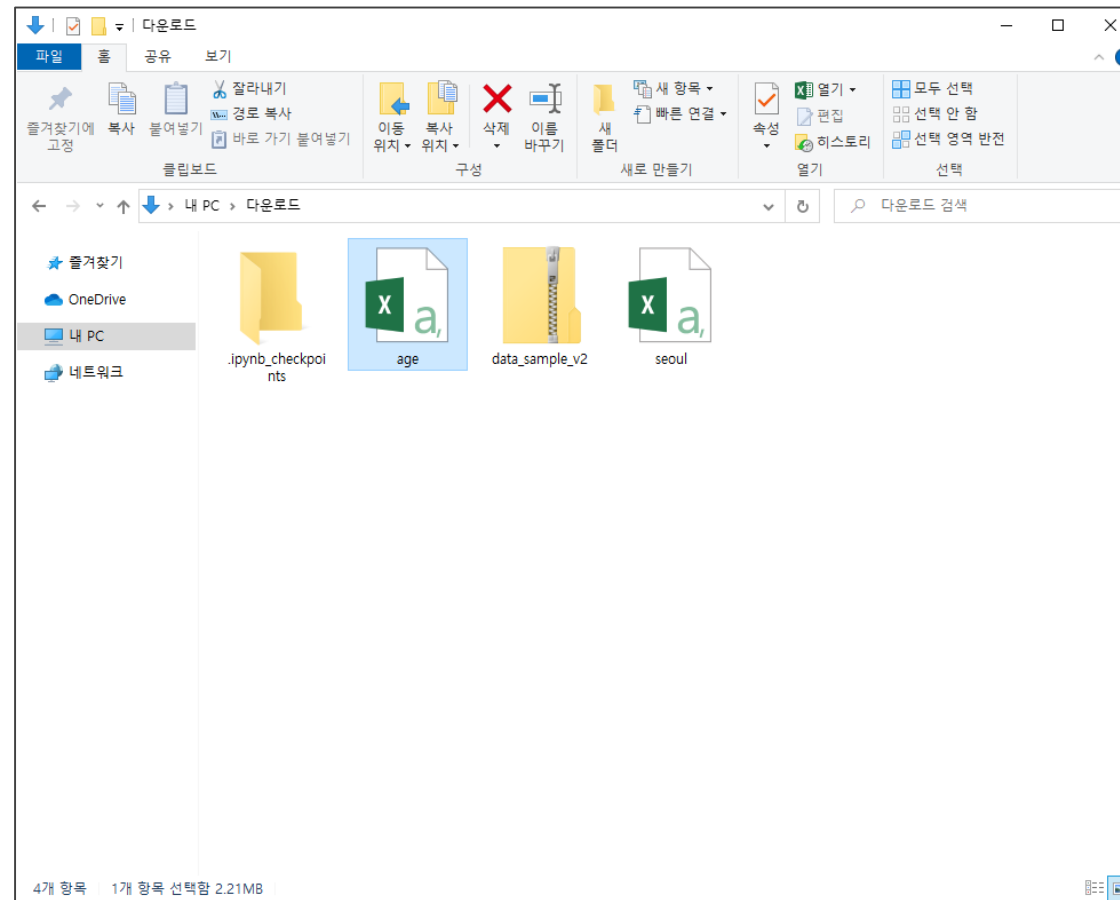
통계자료의 양에 따라 소요시간이 10~30초 소요됩니다.

엑셀 파일 [저장] 후 [파일열기]를 권장합니다.

확인

## ❖ ① 인구 공공데이터 내려받기 (7/8)

- 파일 이름을 "age.csv"로 변경합니다.
- 파일은 다운로드(Downloads) 폴더에 저장되어 있다고 가정하고 실습을 진행하겠습니다.



- age.csv 파일을 열면 전국의 읍면동, 연령별 인구를 확인할 수 있습니다.

Copyright 2023. Dongyang Mirae University all rights reserved.

## ❖ ② 인구 데이터 살펴보고 질문하기

- 어느 동네에 영유아가 가장 많을까?
- 어느 동네가 가장 고령화되었을까?
- 우리 동네에 가장 많이 살고 있는 연령은 몇 살일까?
- 내 또래 사람들이 가장 많이 사는 지역은 어디일까?

## ❖ ③ 우리 동네 인구 구조 시각화하기

- 우리 동네의 인구 구조를 시각화 하려면 어떤 단계(또는 절차)를 거치면 될까요?

### 알고리즘(Algorithm) 설계하기

- ✓ Step 1) 인구 데이터 파일을 읽어온다.
- ✓ Step 2) 전체 데이터에서 한 줄씩 반복해서 읽어온다.
- ✓ Step 3) 우리 동네에 대한 데이터인지 확인한다.
- ✓ Step 4) 우리 동네일 경우 0세부터 100세 이상까지의 인구수를 순서대로 저장한다.
- ✓ Step 5) 저장된 연령별 인구수 데이터를 시각화한다.

## ❖ ④ 남·여 인구 공공데이터 내려받기 (1/8)

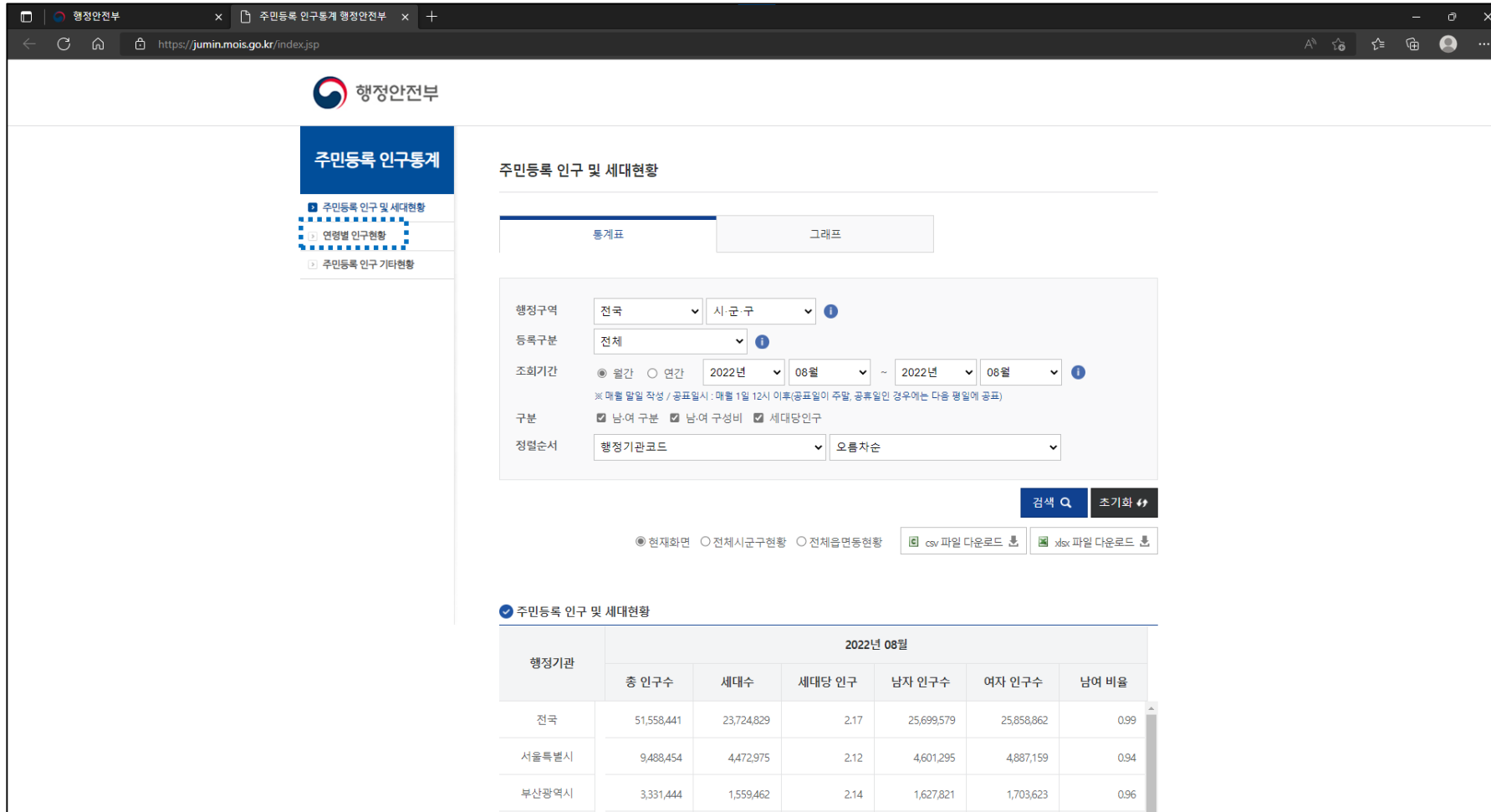
- 행정안전부(<https://www.mois.go.kr/>) - [정책자료] - [통계] - [주민등록 인구통계] 버튼 클릭



The screenshot shows the homepage of the Ministry of the Interior (www.mois.go.kr). The navigation bar at the top includes links for '정보공개' (Information Disclosure), '참여·민원' (Participation·Petition), '뉴스·소식' (News·Notice), '정책자료' (Policy Data), '업무안내' (Business Guide), and '기관소개' (Institution Introduction). The '정책자료' link is circled with a blue dashed line and labeled with a circled '1'. Below the navigation bar, the '정책자료' (Policy Data) section is highlighted. On the left side of this section, a sidebar lists various categories: '정책자료 홈', '주요업무계획', '법령정보', '전자관보', '통계' (Statistics), '간행물', '정책연구보고서', and '참고자료'. The '통계' link is circled with a blue dashed line and labeled with a circled '2'. In the main content area, under the '통계' category, there is a list of statistics. The link '주민등록 인구통계' (Residential Registration Population Statistics) is circled with a blue dashed line and labeled with a circled '3'. Below the main content area, there are three sections: '알립니다' (Notice), '사진뉴스' (Photo News), and '동영상뉴스' (Video News). The '알립니다' section contains a list of notices with dates. The '사진뉴스' and '동영상뉴스' sections each show a news item with a photo and a video thumbnail, both dated 2022-09-21.

## ❖ ④ 남·여 인구 공공데이터 내려받기 (2/8)

- [연령별 인구현황] 버튼 클릭



행정안전부

주민등록 인구통계

주민등록 인구 및 세대현황

통계표 | 그래프

행정구역: 전국 | 시·군·구

등록구분: 전체

조회기간: 월간 | 연간 | 2022년 | 08월 ~ 2022년 | 08월

※매월 말일 작성 / 공표일시: 매월 1일 12시 이후(공표일이 주말, 공휴일인 경우에는 다음 평일에 공표)

구분: 남·여 구분 | 남·여 구성비 | 세대당인구

정렬순서: 행정기관코드 | 오름차순

검색 | 초기화

현재화면 | 전체시군구현황 | 전체읍면동현황

csv 파일 다운로드 | xlsx 파일 다운로드

주민등록 인구 및 세대현황

| 행정기관  | 2022년 08월  |            |        |            |            |       |
|-------|------------|------------|--------|------------|------------|-------|
|       | 총 인구수      | 세대수        | 세대당 인구 | 남자 인구수     | 여자 인구수     | 남여 비율 |
| 전국    | 51,558,441 | 23,724,829 | 2.17   | 25,699,579 | 25,858,862 | 0.99  |
| 서울특별시 | 9,488,454  | 4,472,975  | 2.12   | 4,601,295  | 4,887,159  | 0.94  |
| 부산광역시 | 3,331,444  | 1,559,462  | 2.14   | 1,627,821  | 1,703,623  | 0.96  |

## ❖ ④ 남·여 인구 공공데이터 내려받기 (3/8)

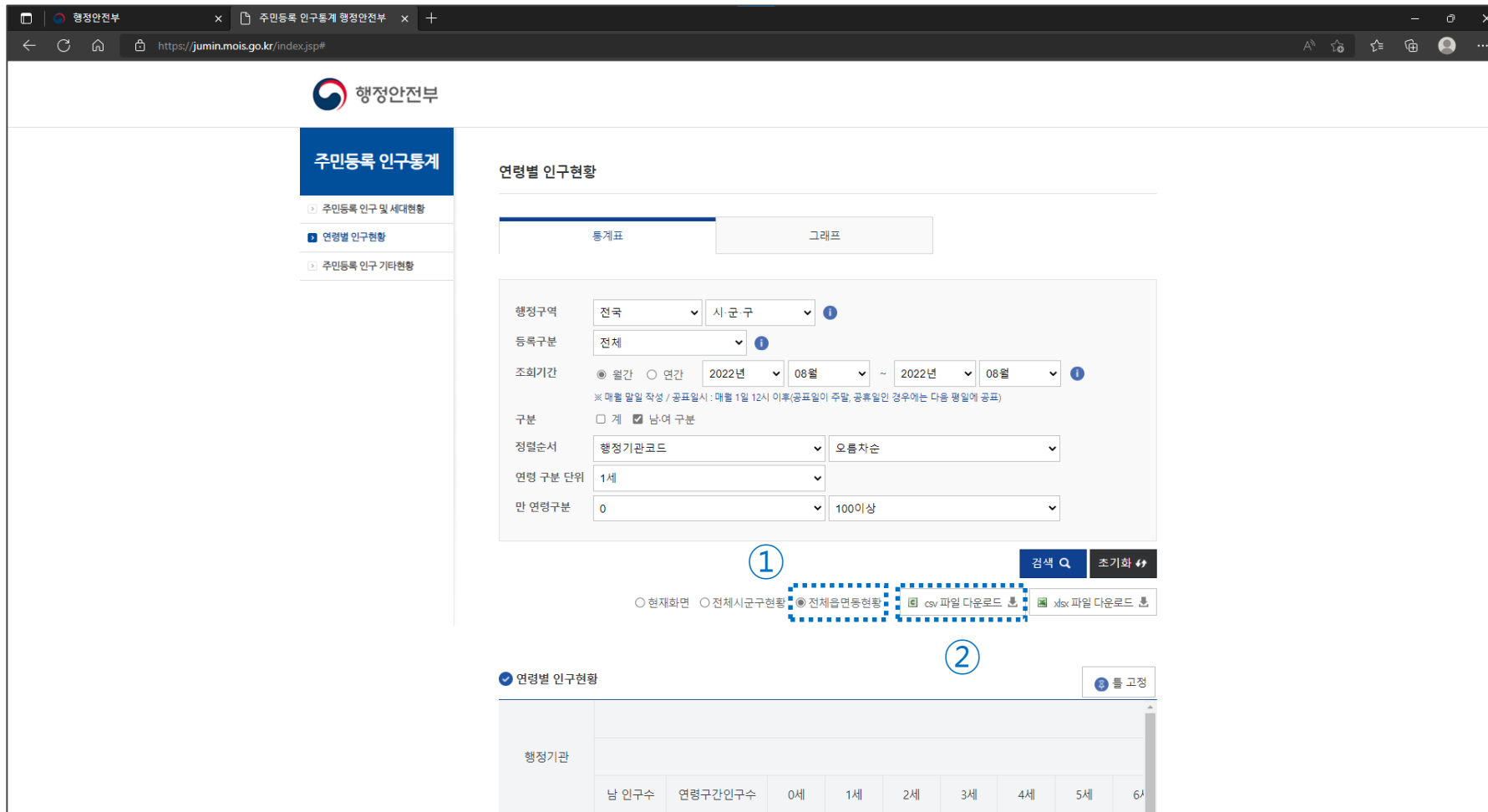
- [조회기간]: 2022년 08월 ~ 2022년 08월
- [계] 체크 해제
- [연령 구분 단위]: 1세
- 만 연령구분: 0, 100이상
- [검색] 버튼 클릭





## ❖ ④ 남·여 인구 공공데이터 내려받기 (4/8)

- [전체읍면동현황] 체크 후, [CSV 파일 다운로드] 버튼 클릭



행정안전부

주민등록 인구통계

연령별 인구현황

통계표 | 그래프

행정구역: 전국 | 시·군·구

등록구분: 전체

조회기간: 월간 | 연간 | 2022년 | 08월 ~ 2022년 | 08월

구분: ☐ 계 ☒ 남·여 구분

정렬순서: 행정기관코드 | 오름차순

연령 구분 단위: 1세

만 연령구분: 0 | 100이상

①

○ 현재화면 ○ 전체시군구현황 ● 전체읍면동현황

②

CSV 파일 다운로드

xls 파일 다운로드

연령별 인구현황

행정기관

| 남 인구수 | 연령구간인구수 | 0세 | 1세 | 2세 | 3세 | 4세 | 5세 | 6세 |
|-------|---------|----|----|----|----|----|----|----|
|-------|---------|----|----|----|----|----|----|----|

## ❖ ④ 남·여 인구 공공데이터 내려받기 (5/8)

- 아래와 같이 경고 창이 뜨면 [확인] 버튼을 클릭하세요.

### jumin.mois.go.kr의 메시지

주민등록인구통계 시스템의 원활한 서비스 제공을 위해  
연령 구분 단위가 1세일 때는 남·여 구분없이 1개월 단위로만 다운로드  
가능합니다.

확인

### jumin.mois.go.kr의 메시지

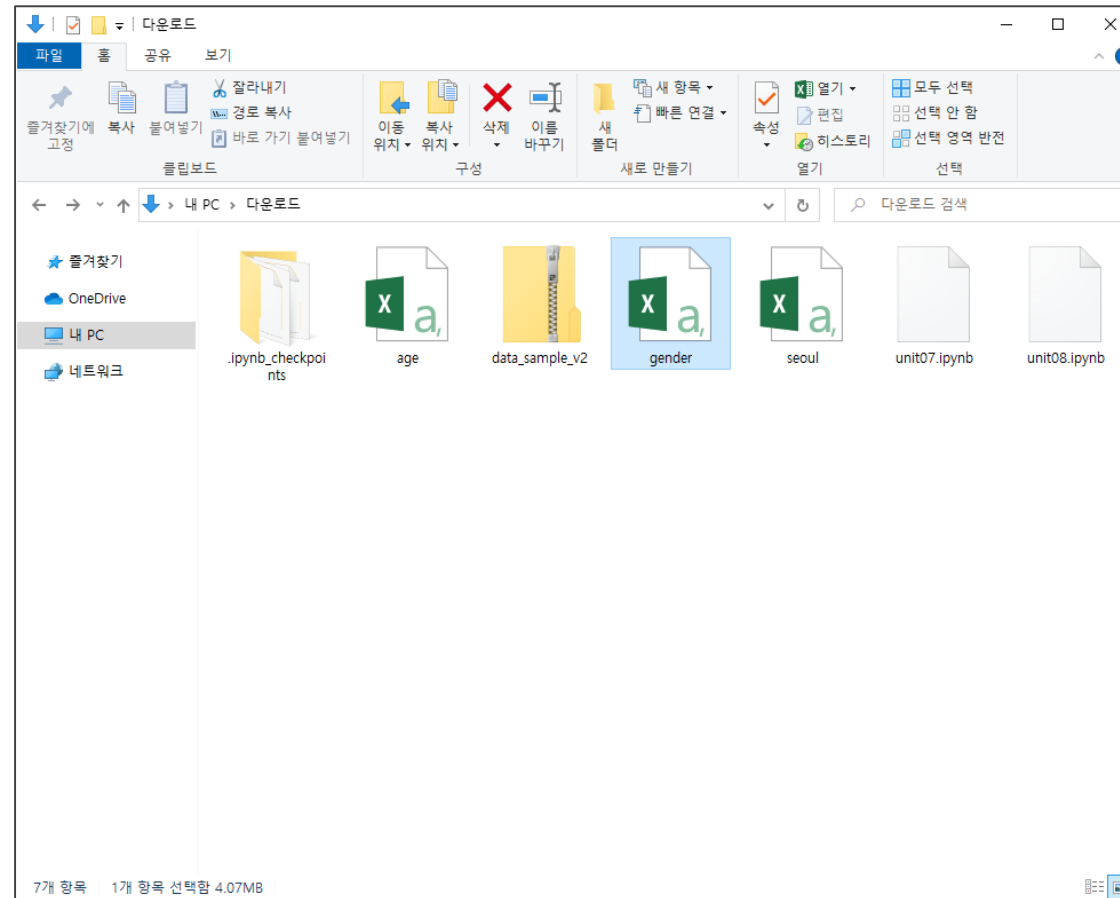
통계자료의 양에 따라 소요시간이 10~30초 소요됩니다.

엑셀 파일 [저장] 후 [파일열기]를 권장합니다.

확인

## ❖ ④ 남·여 인구 공공데이터 내려받기 (6/8)

- 파일 이름을 "gender.csv"로 변경합니다.
- 파일은 다운로드(Downloads) 폴더에 저장되어 있다고 가정하고 실습을 진행하겠습니다.



## ❖ ④ 남·여 인구 공공데이터 내려받기 (7/8)

- gender.csv 파일을 열어 봅니다

◆ B~CZ열: 남성 인구수 데이터

◆ DA~GY열: 여성 인구수 데이터

|    | CW             | CX             | CY             | CZ                 | DA              | DB                 | DC            | DD            | DE            |
|----|----------------|----------------|----------------|--------------------|-----------------|--------------------|---------------|---------------|---------------|
| 1  | 2022년08월_남_97세 | 2022년08월_남_98세 | 2022년08월_남_99세 | 2022년08월_남_100세 이상 | 2022년08월_여_총인구수 | 2022년08월_여_연령구간인구수 | 2022년08월_여_0세 | 2022년08월_여_1세 | 2022년08월_여_2세 |
| 2  | 301            | 240            | 246            | 414                | 4,887,159       | 4,887,159          | 19,816        | 21,718        | 22,053        |
| 3  | 12             | 9              | 8              | 15                 | 74,091          | 74,091             | 212           | 236           | 260           |
| 4  | 3              | 1              | 0              | 1                  | 6,374           | 6,374              | 24            | 22            | 20            |
| 5  | 2              | 2              | 1              | 1                  | 5,154           | 5,154              | 17            | 16            | 21            |
| 6  | 0              | 1              | 0              | 2                  | 1,254           | 1,254              | 2             | 4             | 3             |
| 7  | 0              | 1              | 0              | 0                  | 4,935           | 4,935              | 10            | 14            | 11            |
| 8  | 2              | 1              | 2              | 2                  | 9,570           | 9,570              | 21            | 46            | 51            |
| 9  | 0              | 0              | 0              | 1                  | 4,349           | 4,349              | 20            | 19            | 27            |
| 10 | 0              | 0              | 0              | 0                  | 5,365           | 5,365              | 27            | 32            | 42            |
| 11 | 0              | 0              | 0              | 0                  | 2,148           | 2,148              | 6             | 4             | 3             |
| 12 | 0              | 1              | 1              | 4                  | 2,937           | 2,937              | 5             | 11            | 7             |
| 13 | 0              | 0              | 1              | 1                  | 2,380           | 2,380              | 2             | 5             | 3             |
| 14 | 0              | 0              | 0              | 0                  | 3,767           | 3,767              | 14            | 8             | 9             |
| 15 | 1              | 1              | 3              | 1                  | 8,501           | 8,501              | 16            | 19            | 19            |
| 16 | 2              | 0              | 0              | 0                  | 2,257           | 2,257              | 1             | 5             | 3             |
| 17 | 1              | 0              | 0              | 0                  | 3,799           | 3,799              | 8             | 3             | 11            |
| 18 | 1              | 0              | 0              | 2                  | 3,297           | 3,297              | 18            | 13            | 16            |
| 19 | 0              | 0              | 0              | 0                  | 3,027           | 3,027              | 10            | 2             | 10            |
| 20 | 0              | 1              | 0              | 0                  | 4,977           | 4,977              | 11            | 13            | 4             |
| 21 | 7              | 7              | 3              | 14                 | 62,651          | 62,651             | 299           | 255           | 289           |
| 22 | 0              | 0              | 0              | 0                  | 1,156           | 1,156              | 5             | 5             | 3             |
| 23 | 0              | 1              | 0              | 1                  | 2,281           | 2,281              | 11            | 6             | 8             |

## ❖ ④ 남·여 인구 공공데이터 내려받기 (8/8)

- 인구수 데이터와 인덱스(Index)의 관계

|      |     | 남성 연령별 인구수 |      |    |    |      |     |         | 여성 연령별 인구수 |      |     |     |      |     |         |
|------|-----|------------|------|----|----|------|-----|---------|------------|------|-----|-----|------|-----|---------|
| 열 이름 | 지역명 | 총인구1       | 총인구2 | 0세 | 1세 | (생략) | 99세 | 100세 이상 | 총인구1       | 총인구2 | 0세  | 1세  | (생략) | 99세 | 100세 이상 |
| 인덱스  | 0   | 1          | 2    | 3  | 4  | ...  | 102 | 103     | 104        | 105  | 106 | 107 | ...  | 205 | 206     |

|      |     | 남성 연령별 인구수 |      |    |    |      |     |         | 여성 연령별 인구수 |      |      |      |      |     |         |
|------|-----|------------|------|----|----|------|-----|---------|------------|------|------|------|------|-----|---------|
| 열 이름 | 지역명 | 총인구1       | 총인구2 | 0세 | 1세 | (생략) | 99세 | 100세 이상 | 총인구1       | 총인구2 | 0세   | 1세   | (생략) | 99세 | 100세 이상 |
| 인덱스  | 0   | 1          | 2    | 3  | 4  | ...  | 102 | 103     | -103       | -102 | -101 | -100 | ...  | -2  | -1      |

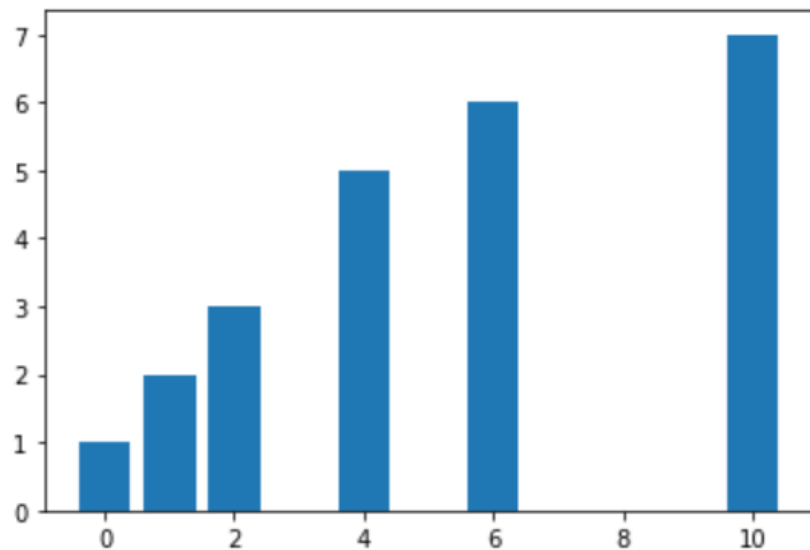
## 02. bar 함수로 막대그래프 그리기

- 01. 인구 데이터 준비하기
- 03. pie 함수로 원그래프 그리기
- 04. scatter 함수로 산점도 그리기

### ❖ 막대그래프(Bar Graph) (1/3)

- 각 데이터의 크기(값)를 막대의 길이로 표현한 그래프
- `bar([막대를 표시할 위치], [막대의 길이])` 함수

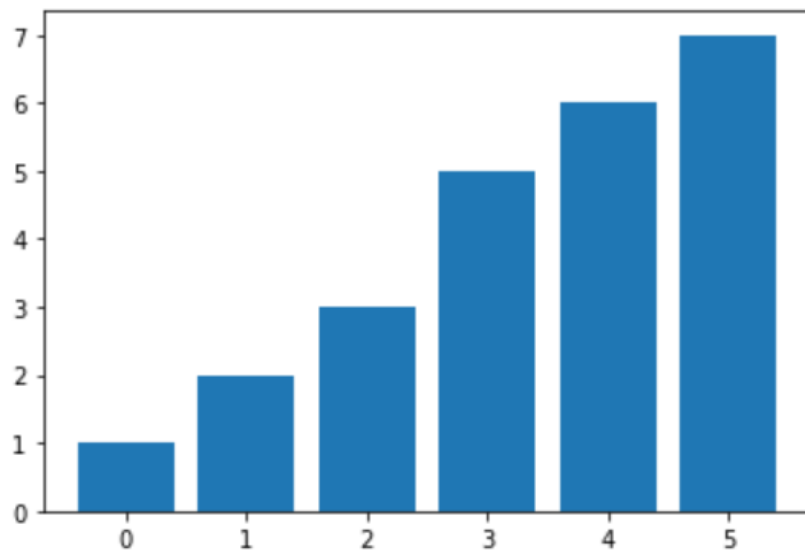
```
import matplotlib.pyplot as plt  
  
plt.bar([0, 1, 2, 4, 6, 10], [1, 2, 3, 5, 6, 7])  
plt.show()
```



### ❖ 막대그래프(Bar Graph) (2/3)

- `range()` 함수를 활용하면 막대를 표시할 위치를 쉽게 지정할 수 있습니다.

```
import matplotlib.pyplot as plt  
  
plt.bar(range(6), [1, 2, 3, 5, 6, 7])  
plt.show()
```



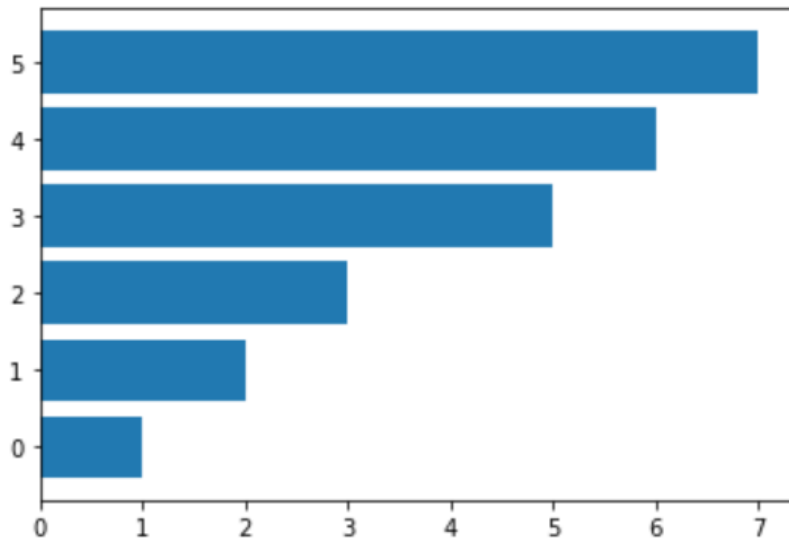
막대그래프를 수직이 아닌 수평 방향으로  
그릴 수는 없을까요?



### ❖ 막대그래프(Bar Graph) (3/3)

- `barh()` 함수를 이용하여 수평 막대그래프를 그릴 수 있습니다.

```
import matplotlib.pyplot as plt  
  
plt.barh(range(6), [1, 2, 3, 5, 6, 7])  
plt.show()
```



- `horizon`  
(명사) 수평선  
(명사) 지평선

## 03. pie 함수로 원그래프 그리기

- 01. 인구 데이터 준비하기
- 02. bar 함수로 막대그래프 그리기
- 04. scatter 함수로 산점도 그리기

## ❖ 원그래프(Pie Chart) (1/7)

- 원그래프는 전체 데이터 중 특정 데이터의 비율을 보기 쉽게 표현합니다.
- pie( ) 함수

```
import matplotlib.pyplot as plt  
  
plt.pie([10, 20])  
plt.show()
```



pie( ) 함수를 이용하면 쉽게 원그래프를 그릴 수 있습니다.

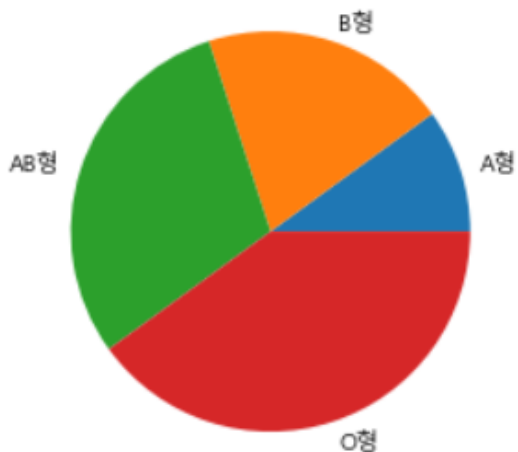
## ❖ 원그래프(Pie Chart) (2/7)

- 레이블(Label) 추가하기

```
import matplotlib.pyplot as plt

data = [10, 20, 30, 40]
name = ['A형', 'B형', 'AB형', 'O형']

plt.rc('font', family='Malgun Gothic')
plt.pie(data, labels=name)
plt.show()
```



원그래프에서는 labels에 각 항목의 이름을 입력하면 항목들의 의미를 쉽게 알아 볼 수 있습니다.

## ❖ 원그래프(Pie Chart) (3/7)

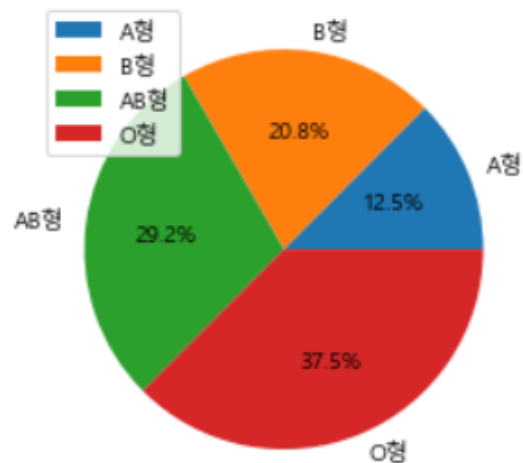
### ● 비율 및 범례 표시하기

```
import matplotlib.pyplot as plt

data = [15, 25, 35, 45]
name = ['A형', 'B형', 'AB형', 'O형']

plt.rc('font', family='Malgun Gothic')
plt.pie(data, labels=name, autopct='%.1f%%')
plt.legend()
plt.show()
```

각 항목의 비율을 표시하고 싶다면  
autopct 속성의 값을 지정해주면 됩니다.  
autopct는 **auto percent**를 의미합니다.



'%.1f%%'의 의미를 살펴보겠습니다.

- ① %는 백분율로 표현하겠다는 의미입니다.
- ② .1f는 소수점 아래 첫 번째 소수점까지 표현하겠다는 의미입니다.
- ③ %%는 표현되는 실수 뒤에 "%" 기호를 적기 위해 작성한 것 입니다.

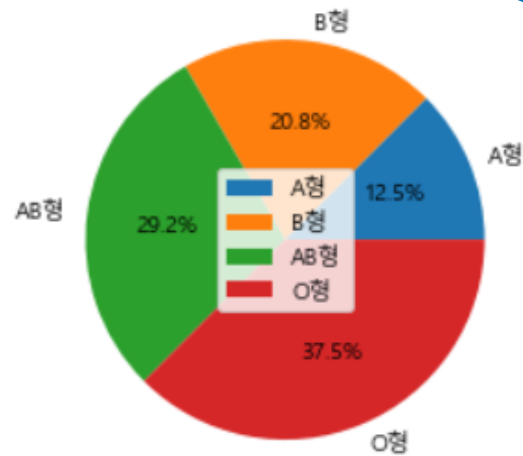
## ❖ 원그래프(Pie Chart) (4/7)

- 범례 위치 조정하기

```
import matplotlib.pyplot as plt

data = [15, 25, 35, 45]
name = ['A형', 'B형', 'AB형', 'O형']

plt.rc('font', family='Malgun Gothic')
plt.pie(data, labels=name, autopct='%.1f%%')
plt.legend(loc=10)
plt.show()
```



아래의 표를 참조하여 loc에 값을 입력하면 범례의 위치를 조절할 수 있습니다.

|   |    |      |
|---|----|------|
| 2 | 9  | 1    |
| 6 | 10 | 5, 7 |
| 3 | 8  | 4    |

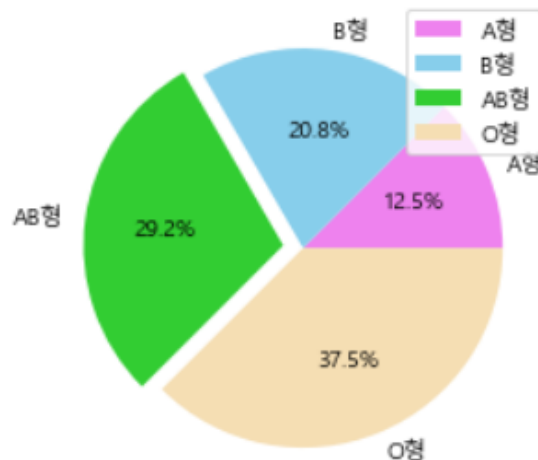
## ❖ 원그래프(Pie Chart) (5/7)

### ● 색 및 돌출 효과 정하기

```
import matplotlib.pyplot as plt

data = [15, 25, 35, 45]
name = ['A형', 'B형', 'AB형', 'O형']
color = ['violet', 'skyblue', 'limegreen', 'wheat']

plt.rc('font', family='Malgun Gothic')
plt.pie(data, labels=name, autopct='%.1f%%', colors=color, explode=(0, 0, 0.1, 0))
plt.legend(loc=1)
plt.show()
```



색은 colors 속성으로 설정할 수 있습니다.  
돌출 효과는 explode 속성으로 설정할 수 있습니다.

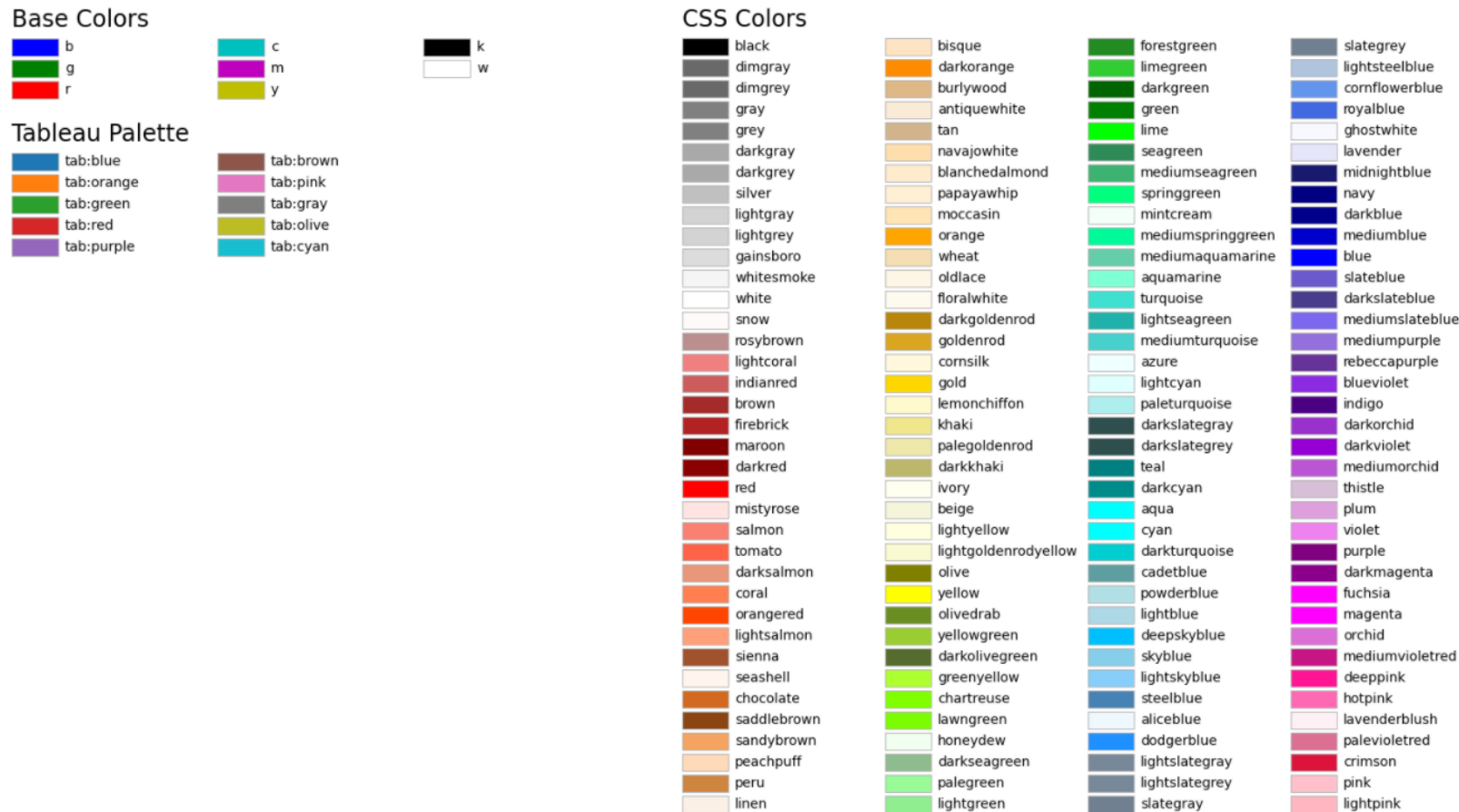
- explode  
(동사) 폭발하다  
(동사) 폭발시키다

# 03. pie 함수로 원그래프 그리기



## ❖ 원그래프(Pie Chart) (6/7)

- matplotlib 라이브러리에서 사용할 수 있는 다양한 색의 이름
- [URL] [https://matplotlib.org/stable/gallery/color/named\\_colors.html](https://matplotlib.org/stable/gallery/color/named_colors.html)





## ❖ 원그래프(Pie Chart) (7/7)

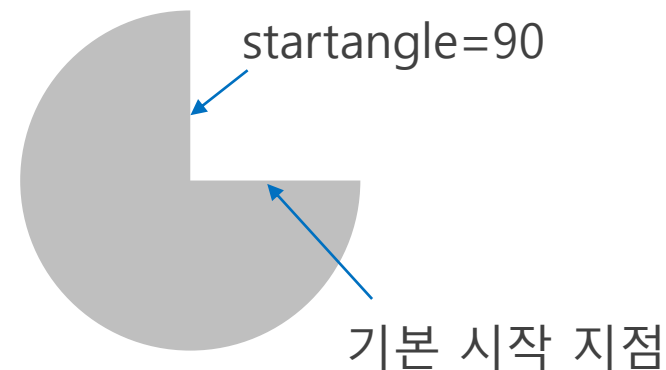
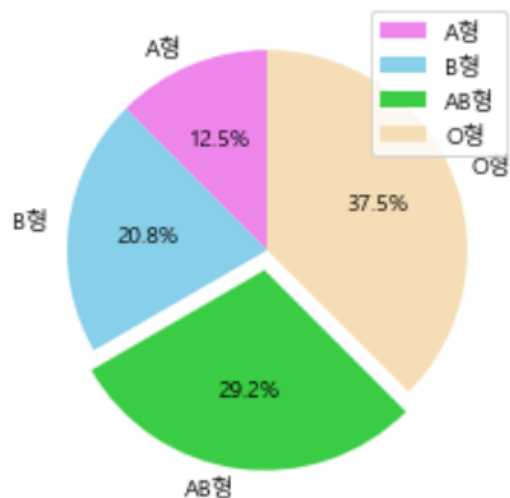
- 원그래프(Pie Chart)의 시작 각도 조정하기

```
import matplotlib.pyplot as plt

data = [15, 25, 35, 45]
name = ['A형', 'B형', 'AB형', 'O형']
color = ['violet', 'skyblue', 'limegreen', 'wheat']

plt.rc('font', family='Malgun Gothic')
plt.pie(data, labels=name, autopct='%1f%%', colors=color, explode=(0, 0, 0.1, 0), startangle=90)
plt.legend(loc=1)
plt.show()
```

startangle 속성에  
시작 각도를 지정하여 줍니다.



## 04. scatter 함수로 산점도 그리기

- 01. 인구 데이터 준비하기
- 02. bar 함수로 막대그래프 그리기
- 03. pie 함수로 원그래프 그리기

## ❖ 산점도(Scatter Plot) (1/10)

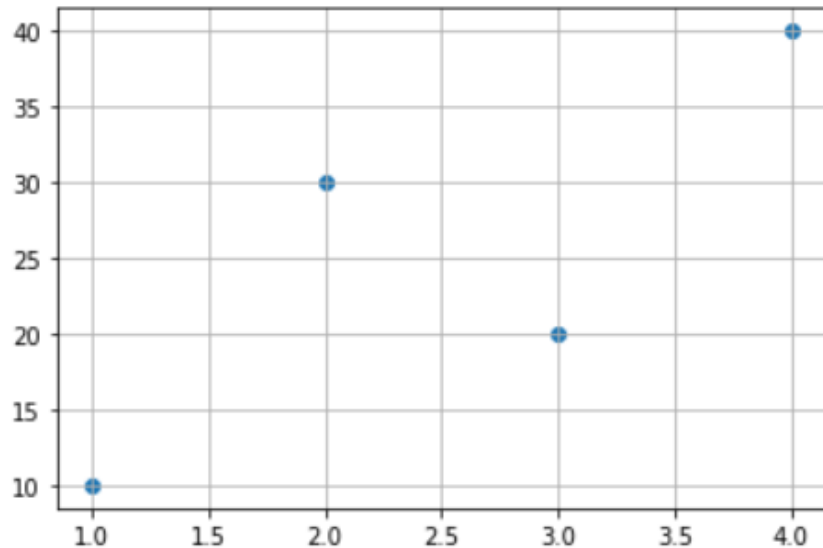
- 산점도는 가로축과 세로축을 기준으로 두 요소가 서로 어떤 관계를 맺고 있는지를 파악하기 쉽게 나타낸 그래프입니다.
- 예를 들어,
  - ◆ 가로축을 제주도 지역의 남성 인구수
  - ◆ 세로축을 제주도 지역의 여성 인구수

scatter( ) 함수와 colorbar( ) 함수를 활용하여  
실습을 진행해 보겠습니다.

## ❖ 산점도(Scatter Plot) (2/10)

- scatter( ) 함수로 표현하기

```
import matplotlib.pyplot as plt  
  
plt.scatter([1, 2, 3, 4], [10, 30, 20, 40])  
plt.grid(True)  
plt.show()
```



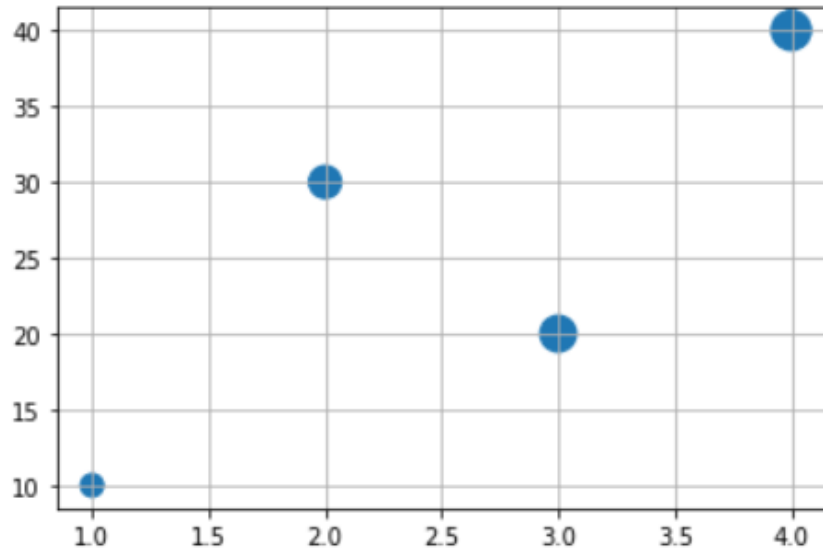
(1, 10), (2, 30), (3, 20), (4, 40) 좌표에  
점이 찍히게 됩니다.

## ❖ 산점도(Scatter Plot) (3/10)

- 점(Point)의 크기를 지정하여 버블 차트(Bubble Chart)로 표현하기

```
import matplotlib.pyplot as plt

plt.scatter([1, 2, 3, 4], [10, 30, 20, 40], s=[100, 200, 250, 300])
plt.grid(True)
plt.show()
```



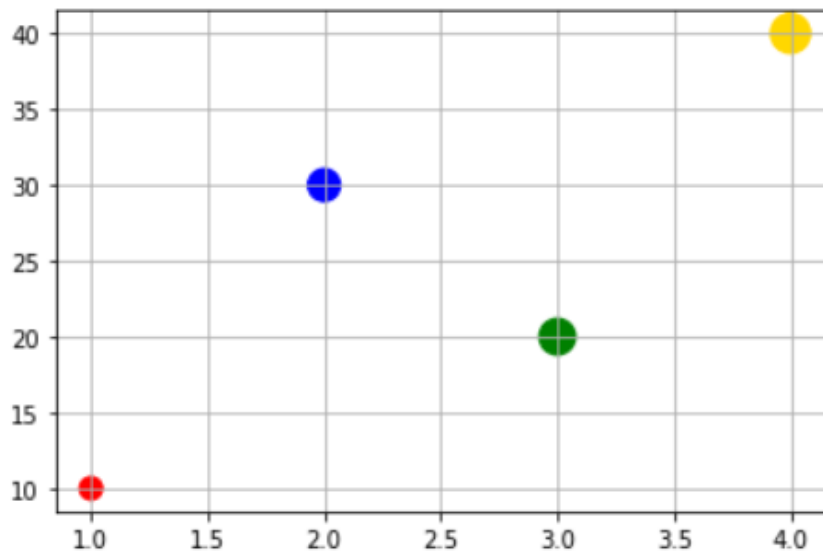
s 속성은 Size를 의미합니다.  
버블의 원하는 크기를 s 속성에 지정해줍니다.

## ❖ 산점도(Scatter Plot) (4/10)

- 버블의 색상 지정하기

```
import matplotlib.pyplot as plt

plt.scatter([1, 2, 3, 4], [10, 30, 20, 40], s=[100, 200, 250, 300], c=['red', 'blue', 'green', 'gold'])
plt.grid(True)
plt.show()
```



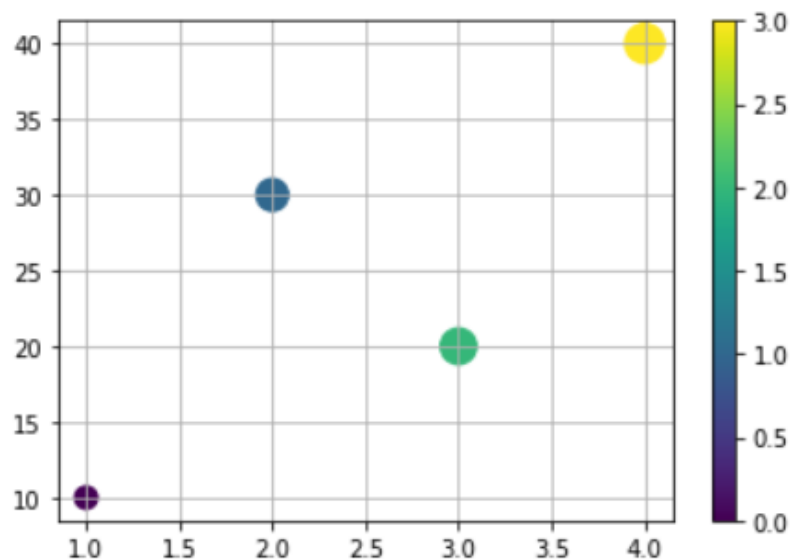
c 속성은 표현하고 싶은 색상의 수를 의미합니다.  
각 버블의 원하는 색상을 c 속성에 지정해 줄 수 있습니다.

## ❖ 산점도(Scatter Plot) (5/10)

### ● 컬러바(Color Bar) 추가하기

```
import matplotlib.pyplot as plt

plt.scatter([1, 2, 3, 4], [10, 30, 20, 40], s=[100, 200, 250, 300], c=range(4))
plt.grid(True)
plt.colorbar()
plt.show()
```



colorbar( ) 함수를 사용하면  
그래프 오른쪽에 컬러바가 추가됩니다.

4가지 색상을 사용한다는 의미로  
c=range(4)라고 수정하였습니다.

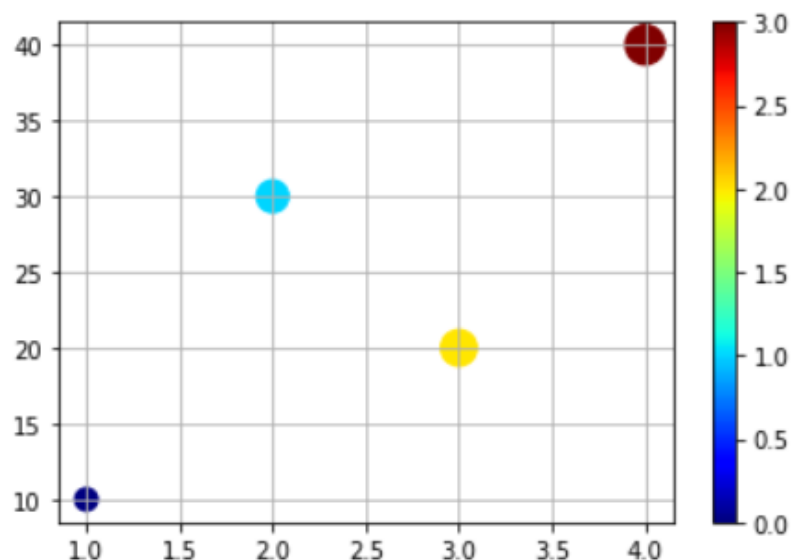
각 데이터는 정해진 색상 개수에 따라  
해당하는 컬러바의 색을 부여 받습니다.

## ❖ 산점도(Scatter Plot) (6/10)

- 컬러바에 사용될 색상 지정하기

```
import matplotlib.pyplot as plt

plt.scatter([1, 2, 3, 4], [10, 30, 20, 40], s=[100, 200, 250, 300], c=range(4), cmap='jet')
plt.grid(True)
plt.colorbar()
plt.show()
```

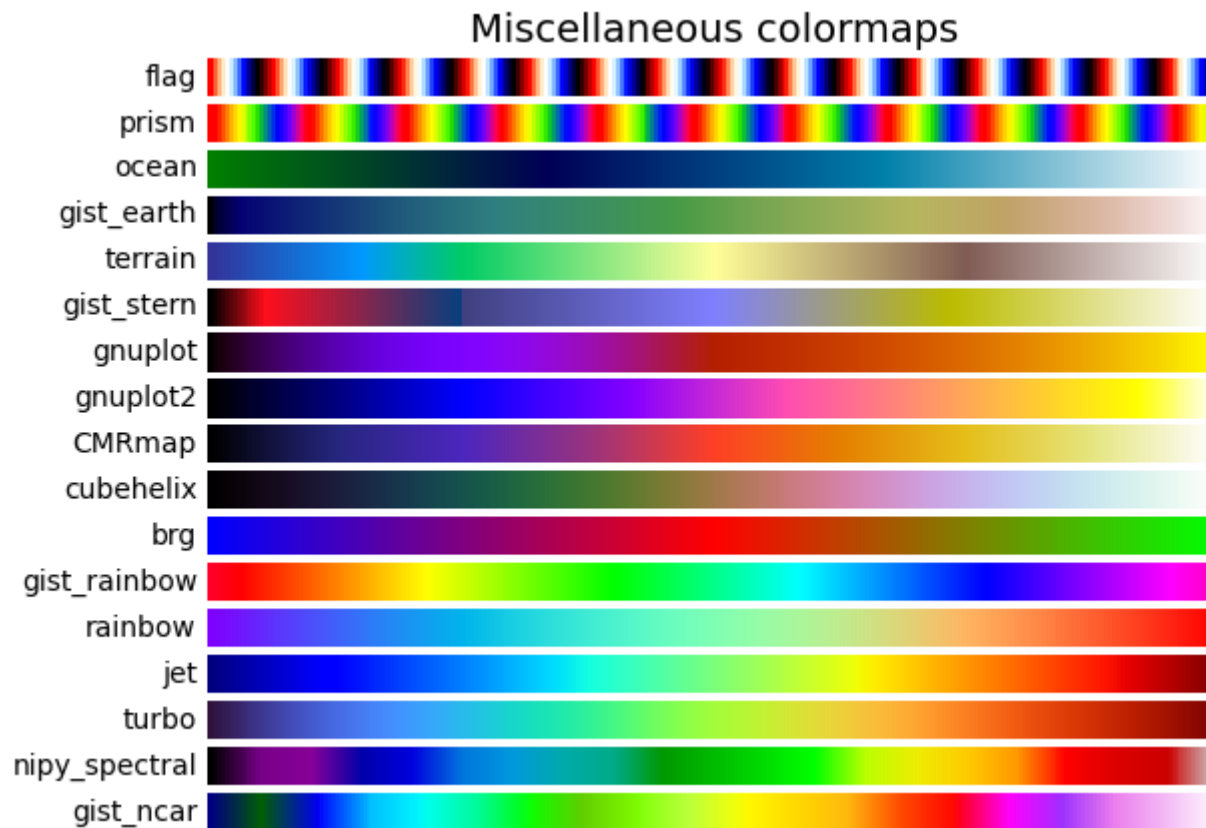


cmap 속성은 Color Map을 의미합니다.  
컬러바에 사용될 색상의 종류를  
지정할 수 있습니다.



## ❖ 산점도(Scatter Plot) (7/10)

- matplotlib 라이브러리에서 사용할 수 있는 다양한 컬러맵의 종류
- [URL] <https://matplotlib.org/stable/tutorials/colors/colormaps.html>



이외에도 다양한 컬러맵이 있습니다.

## ❖ 산점도(Scatter Plot) (8/10)

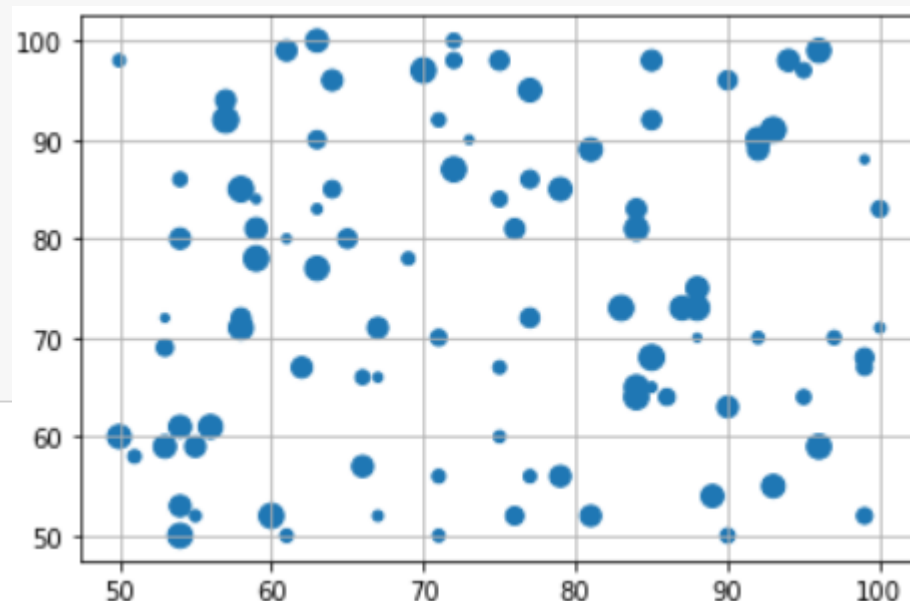
- 위치, 크기가 서로 다른 100개의 점을 산점도로 시각화하기

```
import matplotlib.pyplot as plt
import random

x = []
y = []
size = []

for i in range(100):
    x.append(random.randint(50, 100))
    y.append(random.randint(50, 100))
    size.append(random.randint(10, 100))

plt.scatter(x, y, s=size)
plt.grid(True)
plt.show()
```



## ❖ 산점도(Scatter Plot) (9/10)

- 컬러맵, 컬러바 추가하기

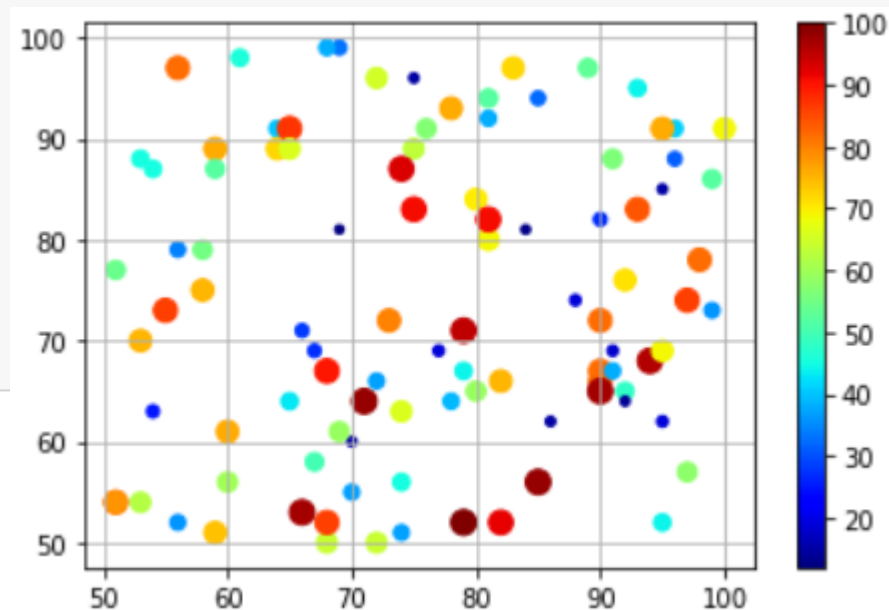
```
import matplotlib.pyplot as plt
import random

x = []
y = []
size = []

for i in range(100):
    x.append(random.randint(50, 100))
    y.append(random.randint(50, 100))
    size.append(random.randint(10, 100))

plt.scatter(x, y, s=size, c=size, cmap='jet')
plt.grid(True)
plt.colorbar()
plt.show()
```

작은 점들이 큰 점에 가려서 잘 안 보이네요.  
어떻게 해야 할까요?



## ❖ 산점도(Scatter Plot) (10/10)

### ● 버블의 투명도 설정하기

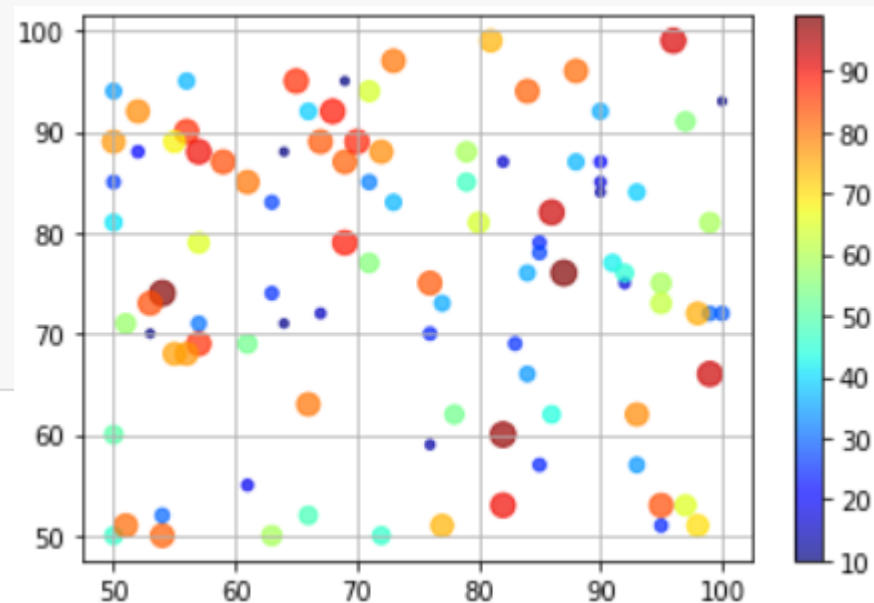
```
import matplotlib.pyplot as plt
import random

x = []
y = []
size = []

for i in range(100):
    x.append(random.randint(50, 100))
    y.append(random.randint(50, 100))
    size.append(random.randint(10, 100))

plt.scatter(x, y, s=size, c=size, cmap='jet', alpha=0.7)
plt.grid(True)
plt.colorbar()
plt.show()
```

alpha 속성값의 범위는 0부터 1까지 입니다.  
0에 가까울 수록 투명해지고, 1에 가까울 수록 불투명합니다.



- ❖ 01. 인구 데이터 준비하기
- ❖ 02. bar 함수로 막대그래프 그리기
- ❖ 03. pie 함수로 원그래프 그리기
- ❖ 04. scatter 함수로 산점도 그리기

# THANK YOU!

## Q & A

- Name: 강환수
- Office: 동양미래대학교 2호관 706호 (02-2610-1941)
- E-mail: [hsknag@dongyang.ac.kr](mailto:hsknag@dongyang.ac.kr)
- Homepage: <https://github.com/ai7dnn/2023-DA>