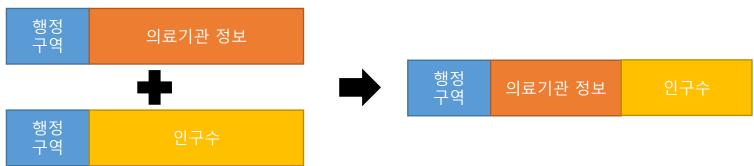
공공보건의료기관 데이터 분석

공공데이터 분석

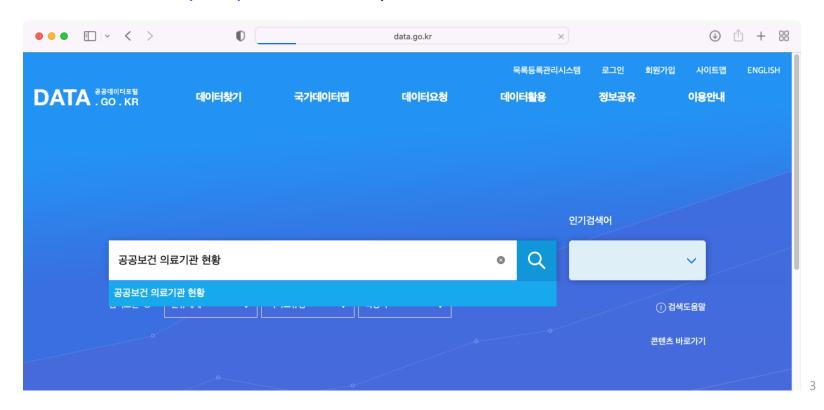
행정 구역별 공공보건의료기관 분석

- ■두 개의 자료를 합침
 - 행정 구역별 공공보건의료기관 자료
 - **행정구역**별 의료기관의 주소
 - 인구수 정보가 없음
 - 행정구역 시군구별 성별 인구수
 - **행정구역**별 인구수 정보 포함
 - 의료기관의 정보가 없음
- ■최종 목표
 - 행정구역별 인구수 대비 공공의료기관의 수 분석

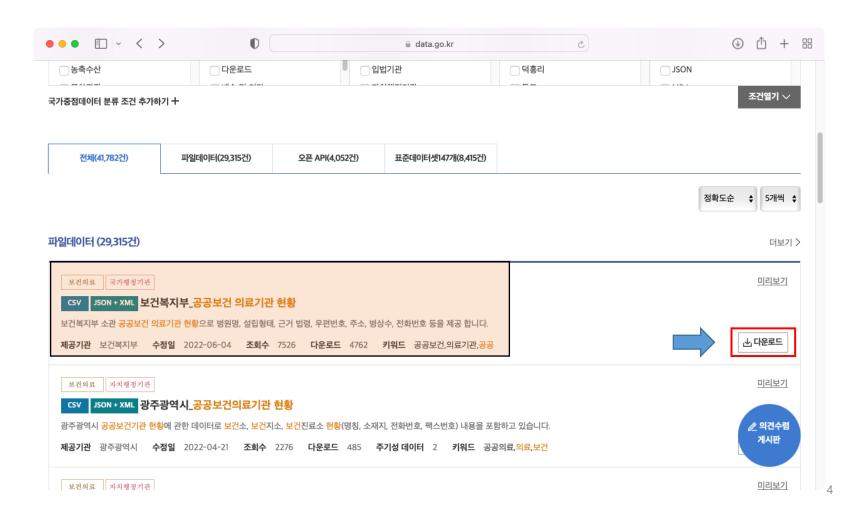


행정 구역별 공공보건의료기관 분석

- 행정구역별 공공보건의료 기관 수 파악
 - 인구수 대비 공공보건의료기간 비율을 비교
- 전국 공공보건 의료 기관 현황 파악
 - www.data.go.kr
 - '공공보건 의료기관 현황' 검색



- ■보건복지부_공공보건 의료기관 현황 csv 파일 다운로드
 - 파일명 변경: 공공보건의료기관현황.csv



■ 공공보건의료기관현황.csv 파일

연번	병원 명	설립형태	근거 법령	관계 행정기관	관계 공공단계	처심평원 요양기	종별구분	병상수	소재지 우편반	주소	홈페이지	대표전화	FAX	비고
	1 강원도 재활병원	시도립	강원도재활병	강원도	해당없음	32200641	병원	165	24227	강원도 춘천	www.grh.or.k	033-248-770	033-248-77	723
	2 강원도 삼척의료원	특수법인	지방의료원의	보건복지부(강원도	지방의료원	32100060	종합병원	152	25920	강원도 삼척	http://ksmc.d	033-572-114	033-573-84	124
	3 강원도 영월의료원	특수법인	지방의료원의	보건복지부(강원도	지방의료원	32100078	종합병원	214	26234	강원도 영월	http://www.y	033-370-911	033-370-91	137
	4 강원도 원주의료원	특수법인	지방의료원의	보건복지부(강원도	지방의료원	32100086	종합병원	237	26448	강원도 원주	www.kwmc.c	033-760-450	033-761-51	121
	5 강원도 강릉의료원	특수법인	지방의료원의	보건복지부(강원도	지방의료원	32100159	종합병원	137	25535	강원도 강릉	http://www.g	033-646-691	033-610-14	115
	6 강원도 속초의료원	특수법인	지방의료원의	보건복지부(강원도	지방의료원	32100141	종합병원	190	24821	강원도 속초	http://www.s	033-630-600	033-630-60	004
	7 춘천시 노인전문병원	시군구립	노인복지법 큰	강원도	해당없음	32280271	요양병원	166	24216	강원도 춘천	http://ccsilve	033-240-750	033-240-75	591
	8 강원도립강릉요양병원	시도립	강원도립강릉	강원도	해당없음	32280424	요양병원	124	25535	강원도 강릉	http://www.g	033-640-880	033-640-88	325
	9 정선군립병원	시군구립	군립의료기관	강원도	해당없음	32201184	병원	61	26150	강원도 정선	없음	033-592-312	033-592-31	125
	10 경기도의료원 수원병원	특수법인	지방의료원의	보건복지부(경기도	지방의료원	31100988	종합병원	143	16316	경기도 수원	http://www.r	031-888-011	031-244-02	242
	11 경기도의료원 이천병원	특수법인	지방의료원의	보건복지부(경기도	지방의료원	31101151	종합병원	115	17366	경기도 이천	http://www.r	031-639-480	031-636-75	500
	12 경기도의료원 안성병원	특수법인	지방의료원의	보건복지부(경기도	지방의료원	31100341	종합병원	151	17572	경기도 안성	http://www.r	031-8046-50	031-8046-5	5555
	13 경기도의료원 의정부병원	특수법인	지방의료원의	보건복지부(경기도	지방의료원	31100287	종합병원	227	11671	경기도 의정-	http://www.r	031-828-500	031-828-50)24
	14 경기도의료원 포천병원	특수법인	지방의료원의	보건복지부(경기도	지방의료원	31100317	종합병원	150	11142	경기도 포천	http://www.r	031-539-911	031-532-01	137
	15 경기도의료원 파주병원	특수법인	지방의료원의	보건복지부(경기도	지방의료원	31100325	종합병원	192	10922	경기도 파주	http://www.r	031-940-910	031-944-77	729
	16 경기도노인전문 용인병원	시도립	경기도노인전	경기도	해당없음	31282628	요양병원	199	17089	경기도 용인	www.silverne	031-288-040	031-288-04	105
	17 경기도노인전문 평택병원	시도립	경기도노인전	경기도	해당없음	31284311	요양병원	230	17723	경기도 평택	goodmsilver.	031-669-998	031-669-99	933
	18 경기도노인전문 시흥병원	시도립	경기도노인전	경기도	해당없음	31283535	요양병원	264	14978	경기도 시흥	www.gshosp	031-500-710	031-500-73	382

■ 공공보건의료기관 현황 데이터 준비

```
import pandas as pd import numpy as np data = pd.read_csv('공공보건의료기관현황.csv', index_col=0, encoding='euc_kr') print(data.columns) data.head()
```

```
Index(['병원 명', '설립형태', '근거 법령', '관계 행정기관', '관계 공공단체', '심평원 요양기관번호', '종별구분', '병상수', '소재지 우편번호', '주소', '홈페이지', '대표전화', 'FAX', '비고'], dtype='object')
```

연번	병원 명	설립형태	근거 법령	관계 행정기관	관계 공공단체	심평원 요양기관번호	종별구분	병상수	소재지 우편
1	강원도 재활병원	시도립	강원도재활병원설치및운영에관한조례	강원도	해당없음	32200641	병원	165	24227
2	강원도 삼척의료원	특수법인	지방의료원의 설립 및 운영에 관한 법률	보건복지부(강원도)	지방의료원	32100060	종합병원	152	25920
3	강원도 영월의료원	특수법인	지방의료원의 설립 및 운영에 관한 법률	보건복지부(강원도)	지방의료원	32100078	종합병원	214	26234
4	강원도 원주의료원	특수법인	지방의료원의 설립 및 운영에 관한 법률	보건복지부(강원도)	지방의료원	32100086	종합병원	237	26448
5	강원도 강릉의료원	특수법인	지방의료원의 설립 및 운영에 관한 법률	보건복지부(강원도)	지방의료원	32100159	종합병원	137	25535

■주소에서 시도, 군구 정보 분리

```
addr = pd.DataFrame(data['주소'])
addr = addr['주소'].apply(lambda v : v.split()[:2]) # 시도, 군구 분리(공백기준)
addr = pd.DataFrame(addr.tolist(), columns=('시도', '군구')) # DataFrame 생성
addr
```

	시도	군구
0	강원도	춘천시
1	강원도	삼척시
2	강원도	영월군
3	강원도	원주시
4	강원도	강릉시
5	강원도	속초시
6	강원도	춘천시

■['시도'] 컬럼에서 고유한 값 찾기

창원시, 경산시, 천안시는 광역시, 특별시가 아님

```
print(addr['시도'].unique())
```

['강원도' '경기도' '경기' '경남' '<mark>창원시</mark>' '경상남도' '경상북도' '<mark>경산시</mark>' '경북' '인천광역시' '대구광역시' '전라남도' '대전광역시' '광주광역시' '제주특별자치도' '부산광역시' '전라북도' '충북' '서울특별시' '서울시' '부산특별시' '대전시' '충남' '전남' '충청남도' '울산광역시' '전북' '<mark>천안시</mark>' '충청북도']

■ '창원시' 주소 데이터 변경

```
addr[addr['시도'] == '창원시']
```

	시도	군구
27	창원시	의창구
31	창원시	마산합포구3.15대로

• 창원시 주소 데이터 변경: 경상남도 창원시

```
addr.iloc[27] = ['경상남도', '창원시']
addr.iloc[31] = ['경상남도', '창원시']
```

• 변경된 내용 확인

addr.iloc[27]

```
    27

    시도
    경상남도

    군구
    창원시
```

■동일하게 '경산시', '천안시' 주소 데이터 변경

```
addr[addr['시도'] == '경산시']

| N도 | 군구 |
| 47 | 경산시 | 경안로 |
| addr.iloc[47] = ['경상북도', '경산시']
| addr.iloc[47] |
| 47 |
| N도 | 경상북도
```

```
addr[addr['시도'] == '천안시']
```

```
시도군구209천안시동남구210천안시동남구
```

군구

경산시

```
addr.iloc[209] = ['충청남도', '천안시']
addr.iloc[210] = ['충청남도', '천안시']
```

■시도, 구군 변경 내용 확인

print(addr['시도'].unique())

```
['강원도' '경기도' '<mark>경기' '경남</mark>' '경상남도' '경북' '인천광역시' '대구광역시'
'전라남도' '대전광역시' '광주광역시' '제주특별자치도' '부산광역시' '전라북도' '<mark>충북</mark>'
'서울특별시' '<mark>서울시' '부산특별시' '대전시</mark>' '<mark>충남</mark>' '전남' '충청남도' '울산광역시' '<mark>전북</mark>'
'충청북도']
```

■시도 이름을 표준이름으로 변경

```
## 표준 행정구역 이름으로 수정 : 경기 -> 경기도, 경남 -> 경상남도, ...
addr_aliases = {'경기':'경기도', '경남':'경상남도', '경북':'경상북도',
'충북':'충청북도', '서울시':'서울특별시', '부산특별시':'부산광역시',
'대전시':'대전광역시', '충남':'충청남도', '전남':'전라남도', '전북':'전라북도'}
# dict.get(key[,default])
# - key에 해당하는 값이 없으면 default값을 리턴
addr['시도']= addr['시도'].apply(lambda v: addr_aliases.get(v, v))
```

```
print(addr['시도'].unique())
```

```
['강원도' '경기도' '경상남도' '경상북도' '인천광역시' '대구광역시' '전라남도' '대전광역시' '광주광역시' '제주특별자치도' '부산광역시' '전라북도' '충청북도' '서울특별시' '충청남도' '울산광역시']
```

■ 군구 이름 확인

```
print(addr['군구'].unique())
```

```
['춘천시', '삼척시', '영월군', '원주시', '강릉시', '속초시', '정선군', '수원시',
                      '파주시', '용인시', '평택시', '시흥시',
'동두천시', '안산시', '부천시', '통영시', '사천시', '창원시', '김해시', '양산시', '거창군',
                     김천시',
                            '안동시',
                                   '울진군', '경주시',
                     '의성군', '청도군', '고령군', '칠곡군',
                           '태백시', '동해시', '화성시',
                            '청주시', '종로구',
                     '진주시',
              '해운대구', '유성구', '가평군', '양주시', '고양시', '홍천군', '양구군',
                            '양평군',
                                   '수성구',
                            '나주시',
                                   '창녕군',
             '계양구',
                            '보성군',
             '여수시', '신안군', '장성군', '완주군', '부안군', '정읍시', '남원시',
'군산시', '고창군', '진안군', '제주시', '서귀포시', '천안시', '보령시', '서산시', '서천군',
'홍성군', '제천시', '충주시', '영동군', '단양군']
```

■ '아란13길' 데이터 변경

addr[addr['군구'] == '아란13길']

	시도	군구
75	제주특별자치도	아란13길

■ '아란13길'을 '제주특별자치도', '제주시'로 수정

```
addr.iloc[75] = ['제주특별자치도', '제주시']
addr.iloc[75]

75
시도 제주특별자치도
```

print(addr['군구'].unique())

제주시

군구

```
['춘천시' '삼척시' '영월군' '원주시' '강릉시' '속초시' '정선군' '수원시' '이천시' '안성시' '의정부시' '포천시' '파주시' '용인시' '평택시' '시흥시' '여주시' '남양주시' '동두천시' '안산시' '부천시' '통영시' '사천시' '창원시' '김해시' '양산시' '거창군' '남해군' '의령군' '포항시' '김천시' '안동시' '울진군' '경주시' '구미시' '영주시' '상주시' '문경시' '경산시' '의성군' '청도군' '고령군' '칠곡군' '봉화군' '울릉군' '부평구' '북구' '순천시' '대덕구' '태백시' '동해시' '화성시' '광산구' '남구' '중구' '제주시' '서구' '전주시' '진주시' '청주시' '종로구' '성남시' '동구' '화순군' '강동구' '사상구' '달서구' '해운대구' '유성구' '가평군' '양주시' '고양시' '홍천군' '양구군' '청원군' '계룡시' '논산시' '함평군' '양평군' '수성구' '달성군' '연수구' '노원구' '기장군' '공주시' '강북구' '광진구' '나주시' '창녕군' '목포시' '고흥군' '연제구' '동매로' '서초구' '은평구' '중랑구' '강남구' '동작구' '동대문구' '양천구' '성동구' '송파구' '울주군' '계양구' '옹진군' '보성군' '광양시' '영광군' '무안군' '진도군' '강진군' '곡성군' '여수시' '신안군' '장성군' '완주군' '부안군' '정읍시' '남원시' '군산시' '고창군' '건안군' '서귀포시' '천안시' '보령시' '서산시' '서천군' '홍성군' '제천시' '충주시' '영동군' '단양군']
```

- 행정구역별 공공보건의료기관의 수 구하기
 - '시도'와 '군구' 컬럼을 합친 새로운 컬럼 추가

```
addr['시도군구'] = addr.apply(lambda r: r['시도'] + ' ' + r['군구'], axis=1) addr.head()
```

	시도	군구	시도군구
0	강원도	춘천시	강원도 춘천시
1	강원도	삼척시	강원도 삼척시
2	강원도	영월군	강원도 영월군
3	강원도	원주시	강원도 원주시
4	강원도	강릉시	강원도 강릉시

• 'count' 컬럼 추가

```
addr['count'] = 0
addr.head()
```

	시도	군구	시도군구	count
0	강원도	춘천시	강원도 춘천시	0
1	강원도	삼척시	강원도 삼척시	0
2	강원도	영월군	강원도 영월군	0
3	강원도	원주시	강원도 원주시	0
4	강원도	강릉시	강원도 강릉시	0

- 행정구역별 공공보건의료기관의 수 구하기
 - ['시도'], ['군구'], ['시도군구']를 기준으로 그룹 생성

	시도	군구	시도군구	count
0	강원도	강릉시	강원도 강릉시	4
1	강원도	동해시	강원도 동해시	1
2	강원도	삼척시	강원도 삼척시	1
3	강원도	속초시	강원도 속초시	1
4	강원도	양구군	강원도 양구군	1

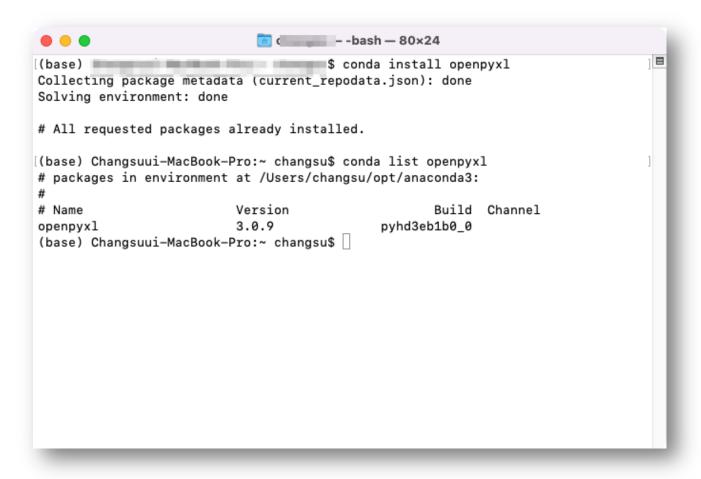
- 행정구역별 공공보건의료기관의 수 구하기
 - ['시도군구'] 컬럼을 데이터프레임 병합에 사용할 인덱스로 설정

```
addr_group = addr_group.set_index('시도군구')
addr_group.head()
```

	시도	군구	count	
시도군구				
강원도 강릉시	강원도	강릉시	4	
강원도 동해시	강원도	동해시	1	
강원도 삼척시	강원도	삼척시	1	
강원도 속초시	강원도	속초시	1	
강원도 양구군	강원도	양구군	1	

openpyxl 라이브러리 설치

- openpyxl 라이브러리 설치
 - conda install openpyxl



행정구역별 인구 데이터 준비

- 행정구역별 인구수 데이터 준비
 - 행정구역 이름 데이터 불러오기
 - 행정구역시군구별_성별인구수.xlsx
 - http://kosis.kr 에서 '행정구역'으로 데이터 검색

population = pd.read_excel('행정구역시군구별_성별인구수.xlsx') population.head()

	행정구역(시군구)별(1)	행정구역(시군구)별(2)	총인구수 (명)	남자인구수 (명)	여자인구수 (명)
0	전국	소계	51847509	25862863	25984646
1	서울특별시	소계	9733509	4745088	4988421
2	서울특별시	종로구	151215	73688	77527
3	서울특별시	중구	126201	61946	64255
4	서울특별시	용산구	229385	110701	118684

행정구역별 인구 데이터 수정

■ 행정구역별 인구수 데이터

• 컬럼의 이름 변경

```
population = population.rename(columns = {'행정구역(시군구)별(1)': '시도',
'행정구역(시군구)별(2)': '군구'})
population.head()
```

	시도	군구	총인구수 (명)	남자인구수 (명)	여자인구수 (명)
0	전국	소계	51847509	25862863	25984646
1	서울특별시	소계	9733509	4745088	4988421
2	서울특별시	종로구	151215	73688	77527
3	서울특별시	중구	126201	61946	64255
4	서울특별시	용산구	229385	110701	118684

행정구역별 인구 데이터 수정

- **행정구역별 데이터**에서 병합에 사용할 공통 컬럼 추가
 - ['군구'] 컬럼에서 공백 제거

```
row_count = population.shape[0] # row의 개수

for element in range(0, row_count):
    population['군구'][element] = population['군구'][element].strip()
```

• ['시도']와 ['군구'] 연결하여 ['시도군구'] 컬럼 추가

```
population['시도군구']= population.apply(lambda r: r['시도'] + ' '
+ r['군구'], axis=1)
population.head()
```

	시도	군구	총인구수 (명)	남자인구수 (명)	여자인구수 (명)	시도군구
0	전국	소계	51847509	25862863	25984646	전국 소계
1	서울특별시	소계	9733509	4745088	4988421	서울특별시 소계
2	서울특별시	종로구	151215	73688	77527	서울특별시 종로구
3	서울특별시	중구	126201	61946	64255	서울특별시 중구
4	서울특별시	용산구	229385	110701	118684	서울특별시 용산구

행정구역별 인구 데이터 수정

• ['군구'] 컬럼에서 '소계' 항목 제거

```
populatin = population[population.군구 != '소계']
```

• ['시도군구'] 컬럼을 병합에 사용할 인덱스로 설정

```
population = population.set_index('시도군구') population.head()
```

	시도	군구	총인구수 (명)	남자인구수 (명)	여자인구수 (명)
시도군구					
서울특별시 종로구	서울특별시	종로구	151215	73688	77527
서울특별시 중구	서울특별시	중구	126201	61946	64255
서울특별시 용산구	서울특별시	용산구	229385	110701	118684
서울특별시 성동구	서울특별시	성동구	300410	147020	153390
서울특별시 광진구	서울특별시	광진구	351263	170164	181099

- ■데이터 병합
 - 내부 조인(inner join)
 - 둘 이상의 데이터프레임에서 조건에 맞는 행을 연결
 - 외부 조인(outer join)
 - 한쪽 데이터프레임에만 존재하는 데이터를 다른 데이터프레임에 결합

■ merge 메소드

- 데이터프레임 연결: 왼쪽 데이터프레임 기준
- left_on, right_on: 왼쪽, 오른쪽 데이터프레임에서 값이 일치해야 할 항목
- left_index=True, right_index=True: 인덱스를 기준으로 병합

- 의료기관 데이터와 행정구역별 인구수 데이터 병합
 - 의료기관 데이터프레임: addr_group
 - 행정구역별 인구수 데이터프레임: population

	시도_x	군구_x	count	시도_y	군구_y	총인구수 (명)	남자인구수 (명)	여자인구수 (명)
시도군구								
강원도 강릉시	강원도	강릉시	4	강원도	강릉시	213328	105775	107553
강원도 동해시	강원도	동해시	1	강원도	동해시	90417	45782	44635
강원도 삼척시	강원도	삼척시	1	강원도	삼척시	66806	33811	32995
강원도 속초시	강원도	속초시	1	강원도	속초시	81840	40470	41370
강원도 양구군	강원도	양구군	1	강원도	양구군	22526	11937	10589

• 병합 후 이름이 같은 컬럼에는 _x, _y가 붙음

- ■필요한 컬럼만 추출해서 새로운 데이터프레임 생성
 - ['시도_x'], ['군구_x'], ['count'], ['총인구수 (명)']

```
local medicalcener population = addr population merge[['시도 x', '군구 x',
                                            'count', '총인구수 (명)']]
local_medicalcenter_population = local_medicalcenter_population.rename(columns=
           {'시도_x':'시도', '군구_x':'군구',
            'count': '공공의료기관수', '총인구수 (명)':'인구수'})
local medical institution.head()
```

	시도	군구	공공의료기관수	인구수
시도군구				
강원도 강릉시	강원도	강릉시	4	213328
강원도 동해시	강원도	동해시	1	90417
강원도 삼척시	강원도	삼척시	1	66806
강원도 속초시	강원도	속초시	1	81840
강원도 양구군	강원도	양구군	1	22526

■ 인구대비 공공의료기관 비율 계산 및 컬럼 추가

local_medicalcenter_population.head()

	시도	군구	공공의료기관수	인구수	인구대비의료기관비율
시도군구					
강원도 강릉시	강원도	강릉시	4	213328	1.875047
강원도 동해시	강원도	동해시	1	90417	1.105987
강원도 삼척시	강원도	삼척시	1	66806	1.496872
강원도 속초시	강원도	속초시	1	81840	1.221896
강원도 양구군	강원도	양구군	1	22526	4.439315

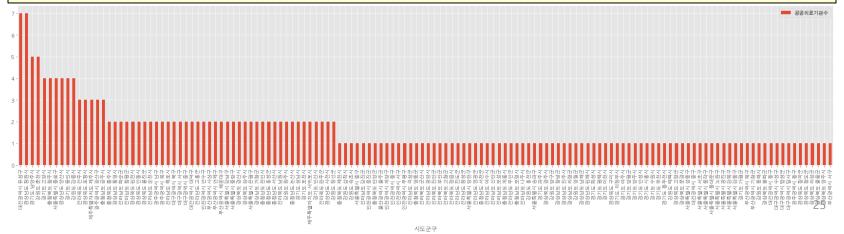
공공보건의료기관 현황 시각화 #1

- ■지역별 공공의료기관수 현황
 - ['공공의료기관수'] 컬럼 내림차순 정렬

```
import matplotlib.pyplot as plt
import platform

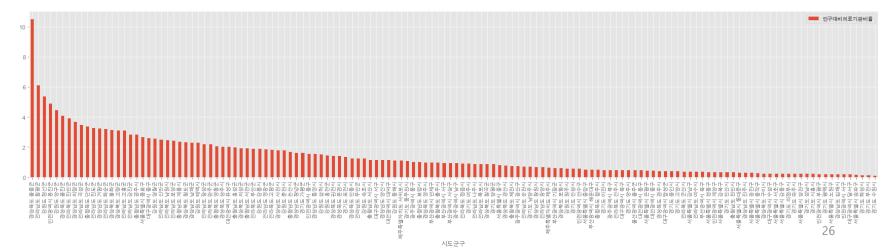
plt.style.use('ggplot')
if platform.system() == 'Windows':
    plt.rc('font', family='Malgun Gothic')
else:
    plt.rc('font', family='AppleGothic')

mc_ratio = local_medicalcenter_population[['공공의료기관수']]
mc_ratio = mc_ratio.sort_values('공공의료기관수', ascending=False)
mc_ratio.plot(kind='bar', rot=90, figsize=(30, 6))
```



공공보건의료기관 현황 시각화 #2

■ 인구대비 공공의료기관 비율





Questions?