

행정구역별 의료기관 현황 분석하기

승실대학교 베어드교양대학
서유희 교수
yhsuh@ssu.ac.kr

행정구역별 의료기관 현황 분석하기

2

▶ 목표

- ▶ 행정구역별로 공공보건의료기관 수를 파악하고 인구 수 대비 공공의료기관 비율을 비교 분석한다.

▶ 데이터 수집

1. 국가통계포털 : 행정구역_시군구_별_성별_인구수.xlsx
2. 공공데이터포털 : 공공보건의료기관현황.csv

▶ 데이터 정제와 탐색

1. 데이터 읽어와 행정구역 분리하기
2. 행정구역 이름으로 주소 수정
3. 행정구역별 공공보건의료기관 수 집계
4. 행정구역별 인구수 데이터 정리
5. 테이블에 필요한 컬럼 추출 후 테이블 병합

▶ 분석모델 구축 및 시각화

주요사용 데이터분석 라이브러리

3

- ▶ **pandas** : 데이터 분석과 처리를 쉽게 할 수 있게 도와주는 라이브러리
 - ▶ <http://pandas.pydata.org>
- ▶ **numpy** : Numerical Python 의 약어로 파이썬으로 고성능 수치 계산을 쉽고 빠르게 할 수 있도록 만든 라이브러리
 - ▶ <http://www.numpy.org>
- ▶ **matplotlib** : MATLAB(과학 및 공학 연산을 위한 소프트웨어)의 시각화 기능을 모델링해서 만들어진 파이썬 라이브러리
 - ▶ <http://matplotlib.org/>

파이썬 기본 자료구조

4

▶ 리스트

- ▶ $a = [10, 20, 30]$
- ▶ $b = [[1,2], [3,4], [5,6]]$
- ▶ 데이터 개별 접근 : $a[0] + a[2]$ / $a[0] = 1$

▶ 튜플

- ▶ $a = (10,20,30)$
- ▶ $b = ((1,2), (3,4), (5,6))$
- ▶ 데이터 개별 접근 : $a[0] + a[2]$ / $a[0] = 1$ (Error)

▶ 딕셔너리

- ▶ key : value 형식으로 데이터 저장
- ▶ $dict = \{ 'a' : 100, 'b' : 200, 'c' : 300 \}$
- ▶ 데이터 개별 접근 : $dict['a'] + dict ['b']$

Pandas 자료구조

5

Series

▶ 인덱싱 기능을 추가한 1차원 배열

Index	Data
1	'A'
2	'B'
3	'C'
4	'D'
5	'E'

DataFrame

– 인덱스와 컬럼으로 구성된 2차원 데이터 구조

Index	Columns		
	A	B	C
0	'Hello'	'Column B'	NaN
1	'NO INFO'	'NO INFO'	'NO INFO'
2	'A'	'Column B'	NaN
3	'A'	'Column B'	NaN
4	'A'	'Column B'	NaN

Data

데이터 정제 및 탐색

행정구역별 공공보건의료기관 수 집계

1. 데이터 읽기와 행정구역 정보 분리하기

8

- ▶ '보건복지부 공공보건 의료기관 현황.csv' 파일 가져오기
- ▶ 주소에서 시도, 군구 정보를 분리

면

병원명	설립 형태	근거 법령	관계 행정 기관	관개 공공 단체	심평원 요양 기관번호	종별 구분	병상 수	소재지 우 편번호	주소	홈페이지	대표전화	FAX	비고	
1	강원도 재활병원	시도립	강원도재활병원 설치및운영에관한조례	강원도	해당 없음	32200641	병원	165	24227	강원도 춘천시 충열로 142번길 24-16	www.grh.or.kr	033-248-7700	033-248-7723	NaN
2	강원도 삼척의료원	특수법인	지방의료원의 설립 및 운영에 관한 법률	보건복지부(강원도)	지방의료원	32100060	종합 병원	152	25920	강원도 삼척시 오십천로 418	http://ksmc.or.kr	033-572-1141	033-573-8424	NaN
3	강원도 영월의료원	특수법인	지방의료원의 설립 및 운영에 관한 법률	보건복지부(강원도)	지방의료원	32100078	종합 병원	214	26234	강원도 영월군 영월읍 중앙1로 59	http://www.youngwol.org	033-370-9117	033-370-9137	NaN
4	강원도 원주의료원	특수법인	지방의료원의 설립 및 운영에 관한 법률	보건복지부(강원도)	지방의료원	32100086	종합 병원	237	26448	강원도 원주시 서원대로 387(개운동)	www.kwmc.or.kr	033-760-4500	033-761-5121	NaN
5	강원도 강릉의료원	특수법인	지방의료원의 설립 및 운영에 관한 법률	보건복지부(강원도)	지방의료원	32100159	종합 병원	137	25535	강원도 강릉시 강강로 2007(남문동 164-1)	http://www.gnmc.or.kr	033-648-6910	033-610-1415	NaN

2. 행정구역 이름 수정하기

9

1) 시도이름에서 잘못된 시도이름 수정

▶ 창원시 -> 경상남도 창원시

시도	군구
27 창원시	의창구
31 창원시	마산합포구3.15대로

▶ 경산시 -> 경상북도 경산시

시도	군구
47 경산시	경안로

▶ 천안시 -> 충청남도 천안시

시도	군구
209 천안시	동남구
210 천안시	동남구

2) 시도이름을 표준이름으로 수정

- 1) 경기 -> 경기도, 경남 -> 경상남도
- ▶ 경북 -> 경상북도, 충북 -> 충청북도
- ▶ 서울시 -> 서울특별시
- ▶ 부산특별시 -> 부산광역시
- ▶ 대전시 -> 대전광역시
- ▶ 충남 -> 충청남도, 전남 -> 전라남도
- ▶ 전북 -> 전라북도

3) 잘못된 군구이름 수정

▶ 아란13길 -> 제주특별자치도, 제주시

시도	군구
75 제주특별자치도	아란13길

3. 행정구역별 공공보건의료기관 수 집계

10

1) 시도, 군구 컬럼 결합한 컬럼 생성

	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

2) 시도군구를 기준으로 그룹을 만들어 그룹별 의료기관수 합계 구하기

	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

3) 데이터 병합에 사용할 인덱스 설정하기

	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

행정구역별 인구 수 데이터 추가

4. 행정구역 컬럼 추가하기

12

1) 행정구역_시군구_별_성별_인구수_2.xlsx 데이터 가져오기

	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

2) 컬럼명 시도, 군구로 변경

	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

4. 행정구역 컬럼 추가하기

13

3) 병합에 사용할 공통 컬럼 '시도군구' 추가, 필요 없는 행 ('소계') 삭제

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

4) 병합의 기준이 될 인덱스를 '시도군구' 로 설정

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

테이블에 필요한 컬럼 추출하여 테이블 병합

5) 데이터 프레임 병합 후 데이터 정리

15

1) 의료기관 데이터프레임과 '시도군구별 인구수' 데이터 프레임 병합

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

2) 필요컬럼만 추출 후 이름 변경

시도	군구	count	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

5) 데이터 프레임 병합 후 데이터 정리

16

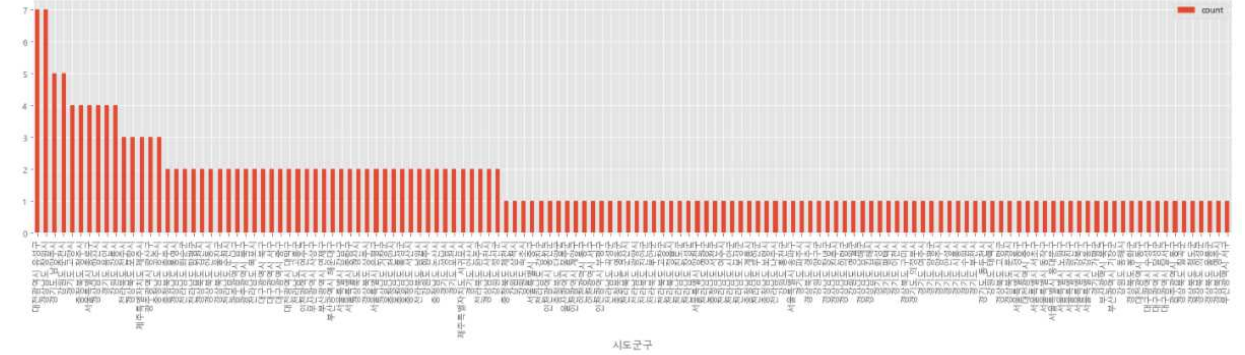
3) 시도군구의 인구대비 의료기관 수 비율 구하기

시도	군구	count	인구수	MC_ratio
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

분석 모델 구축 및 시각화

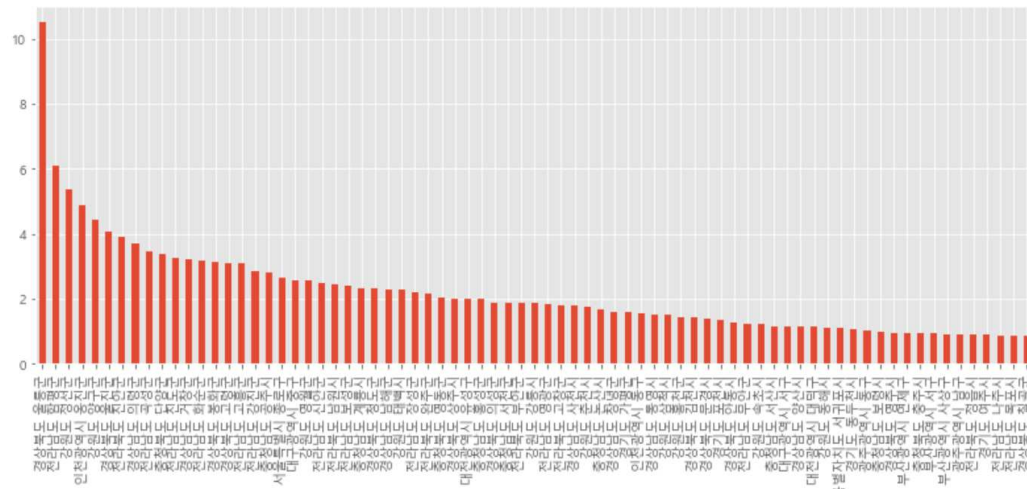
행정구역별 공공보건의료기관 수

18



행정구역별 인구수 대비 공공보건의료기관 비율

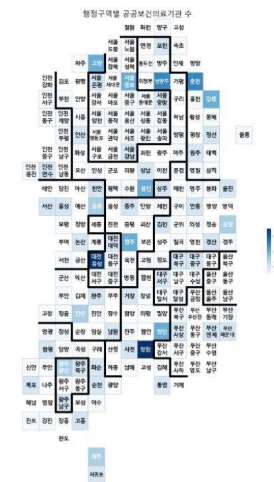
19



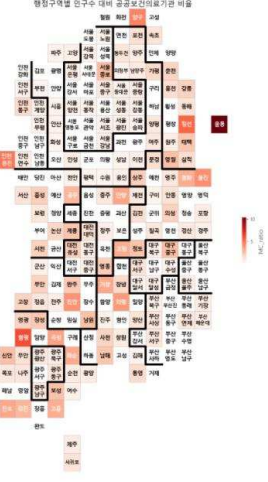
블록맵 시각화

20

▶ 행정구역별 공공보건의료기관 수



▶ 행정구역별 공공보건의료기관 비율



다음시간에 배울 내용

주	주제	온라인	오프라인
1	인공지능의 과거 현재와 미래	1. 강의 및 교과목 소개(공통, 핵심만) 2. 인공지능의 과거와 현재 3. 인공지능의 미래와 다양한 시선 4. 인공지능 개발환경 구축과 사용법(Anaconda/Colab)	1. 강의 및 교과목 소개(분반별, 자세히) 2. 다양한 인공지능 기술 경험하기 (자연어처리, 시각, 음성) 3. 인공지능 챗봇만들기(IBM 왓슨 어시스턴트)
2	공공데이터를 이용한 사회문제 발견과 해결책 모색	1. 빅데이터의 정의와 가치 2. 공공데이터 수집하기 3. 공공데이터로부터 새로운 인사이트 발견하기 - 행정구역별 인구 데이터와 공공의료기관 현황 데이터 분석	1. 서울시 CCTV설치 현황 분석하기 2. 서울시 범죄발생 현황 분석하기
3	인공지능의 개요 및 머신러닝을 이용한 예측	1. 인공지능의 정의와 분류 2. 인공지능 학습방법 이해하기 3. 인공지능 알고리즘 소개	1. 머신러닝을 이용한 이미지 식별(구글 티처블 머신) 2. 머신러닝을 이용한 보스톤 집값 예측
4	인공지능과 데이터 윤리	1. 데이터의 불완전성과 결함에 따른 예측 오류와 차별 2. 데이터 왜곡에 따른 분석과 예측 결과 비교	1. 데이터 편향성이 예측에 미치는 영향 (구글 티처블 머신) 2. 데이터 왜곡에 따른 예측 결과 비교 - 타이타닉호 생존자 예측
5	인공지능과 알고리즘 윤리	1. 알고리즘 기반 의사결정 시스템의 한계 2. 윤리가 적용된 인공지능 알고리즘	1. 알고리즘에 따른 예측 결과 비교 - 보스톤 집값 예측 - 폐암환자 생존 여부 예측
6	인공지능에 대한 다양한 이슈와 우리의 자세 고찰	1. 인공지능의 윤리적/법적 쟁점 (자율주행자동차, AI로봇, 트랜스 휴먼 등) 2. 인공지능시대 사회, 경제적 불평등 문제 3. 인공지능과 프라이버시 4. 인공지능의 윤리적 대응과 규제	1. 자율주행 자동차의 행동학습 시나리오 경험하기 2. 비윤리적 데이터 생성과 수집(웹 크롤링을 이용한 데이터 수집)
7		기말고사	