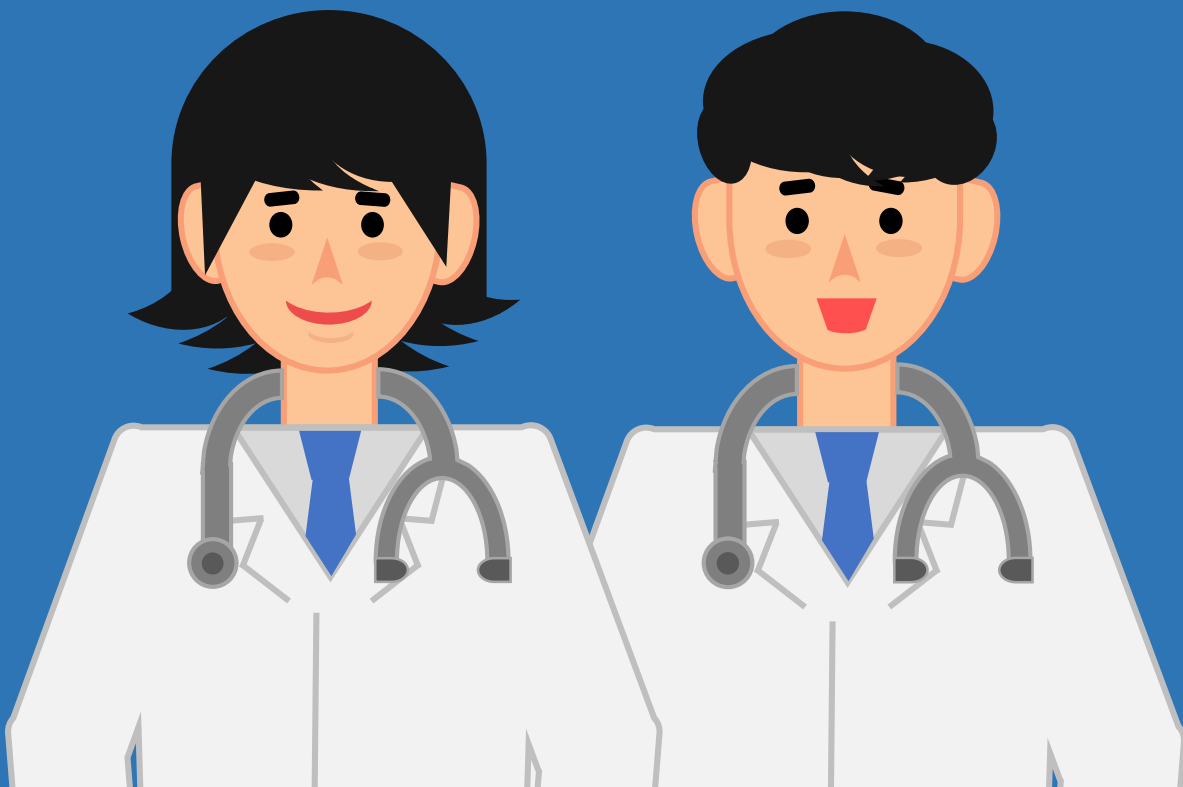




공공빅데이터청년인턴십
국민건강증진 실현을
위한 보건지수 개발

2조

김의정 김지은 김초현 마지우 이연인 이종안 전다빈

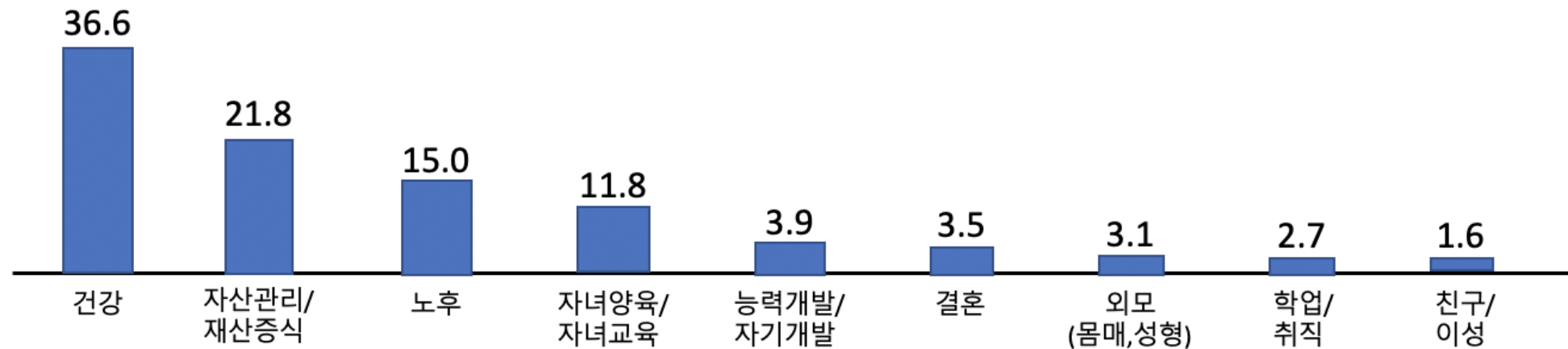


INDEX

- ① 보건지수의 필요성
- ② 보건지표 선정
- ③ 분석 프로세스
- ④ 분석 결과
- ⑤ 활용 방안
- ⑥ 결론



1. 보건지수의 필요성



2018 소비 트렌드 조사에 따르면 소비자가
가장 관심있는 분야는 **건강**

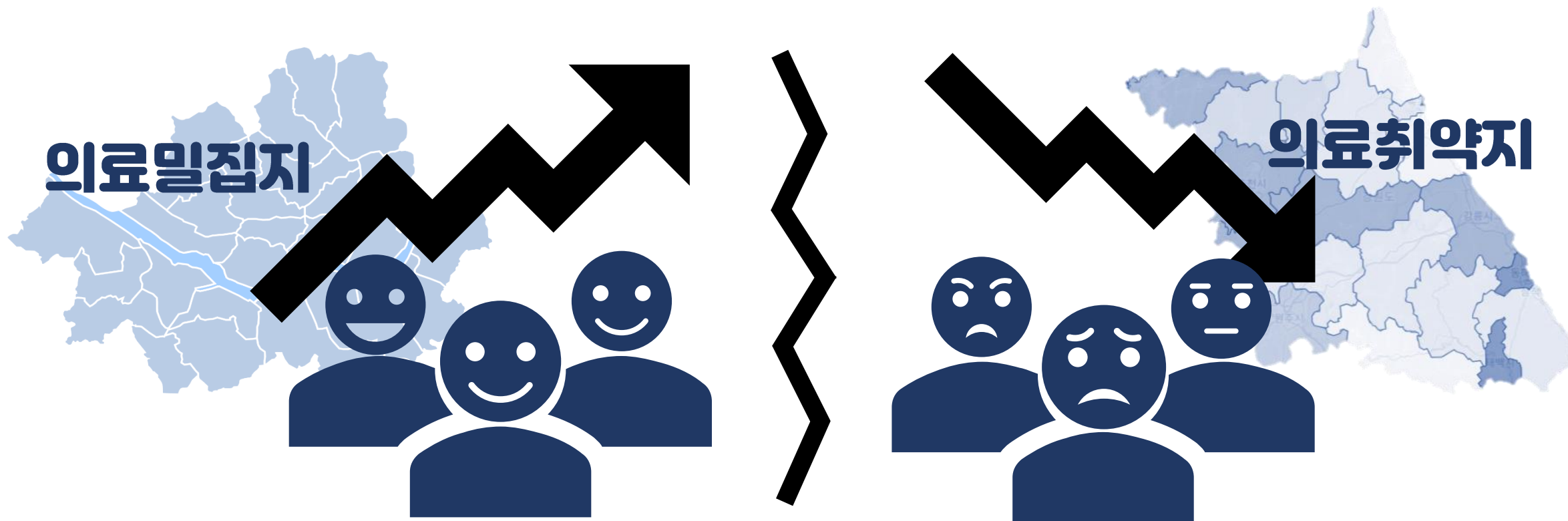


경제발전과 시민의식에 성장에 따라 건강에 대한 기대수준 향상



건강증진법 제정 뒤 건강증진사업 활성화

보건지수 개발의 필요성



국민 전체의 건강 수준은 향상되었지만
의료서비스 및 사회구조적 문제로 건강 불평등이라는 한계 발생

<의료 불균형 관련 기사>

지역 간 의료인력 격차 심각...필수 의료서비스 부족 초래

연합뉴스 2019.04.24

우리나라 보건의료 체계의 가장 큰 문제는 '지역 간 의료 불균형'

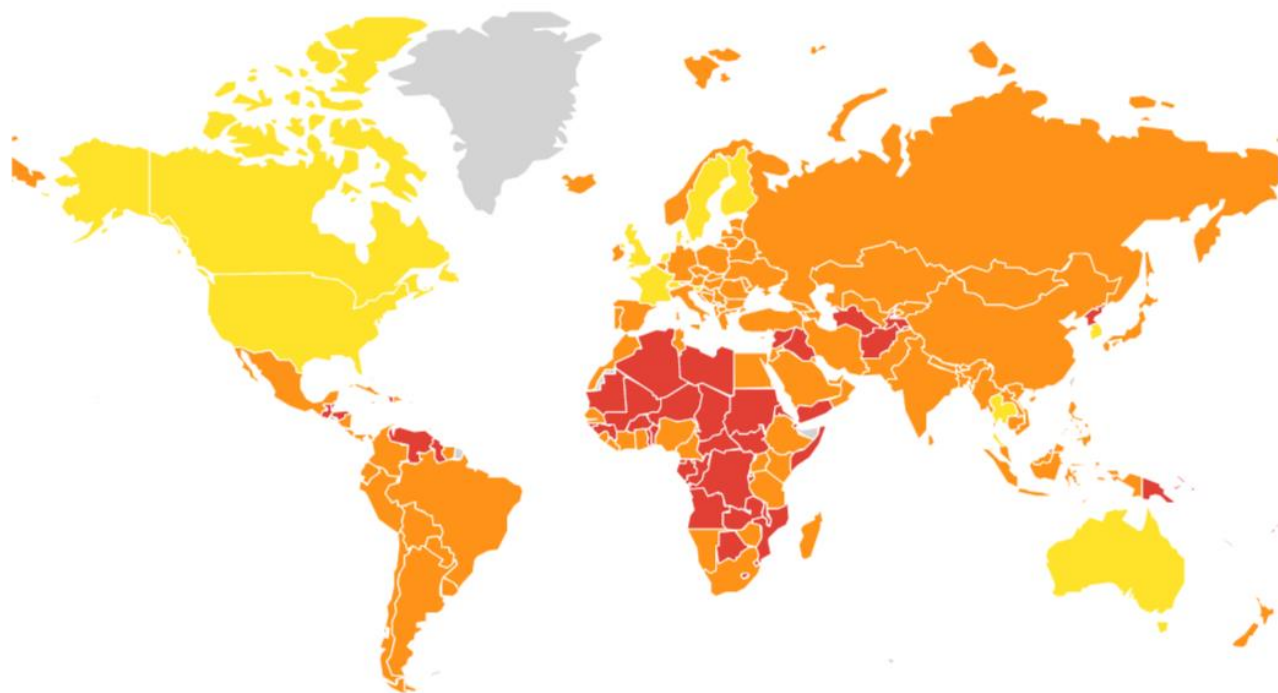
한의신문 2020.09.01

1만명 당 병상수, 지역간 4배 차이 나기도...김광수 의원 "수도권-지방 간 차별 없애야"

청년 의사 2019.10.21

- 지역 보건의에 관한 뉴스 및 논문 자료를 크롤링
- 의료 민원 관련 키워드 추출 후 워드클라우드로 시각화

세계적으로 사용하는 보건지수



Key

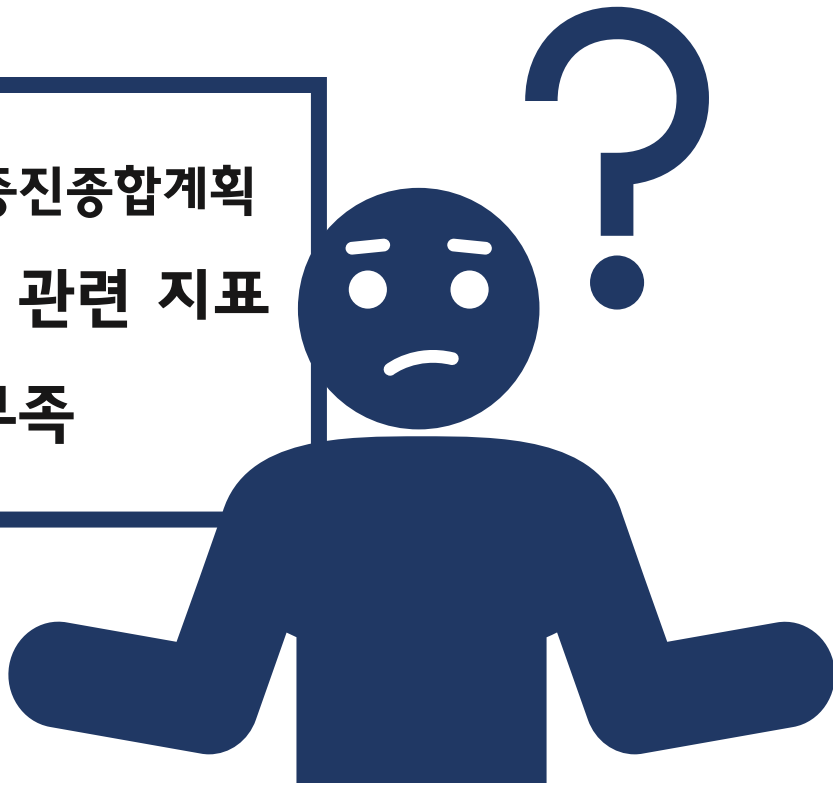
- Most Prepared
- More Prepared
- Least Prepared

Select a country to see Overall Score/Rank and access a full country page.

세계보건안전지수(GHS)

보건지수 개발의 필요성

국민건강증진종합계획
의료자원 관련 지표
부족



현재 우리나라는 의료와 사회구조를
모두 파악할 수 있는 지표가 부재

대한민국만의 보건지수 개발



보건지수 개발의 필요성



〈보건지수 개발의 의의〉

- 국민건강에 영향을 주는 다양한 요인을 고려한 보건지수를 개발
- 보건의료 관련 정책 실무자가 지역사회 건강수준과 보건의료 현황을 정확하게 진단
- 지역의 보건수준에 영향을 미치는 구체적인 요인을 파악

2. 보건지표 선정

분석 데이터 목록



의료 데이터

의료 데이터 목록

구분	활용데이터	설명	컬럼
인구	인구	인구수에 대한 데이터	지표연도, 시도코드, 시도명, 이벤트 건수 (표본집단), 대상자 건수 (모집단)
의료기관	의료기관	의료기관수에 대한 데이터	
주요시설	전문병상	전문/허가병상 수에 대한 데이터 *전문 ; 전문의료 기관 개설허가 신고상의 병상 수 *허가 : 의료기관 개설허가 신고상의 병상 수	
	허가병상		
의료인력	의사	의사/전문의/간호사/약사 수에 대한 데이터	
	전문의		
	간호사		
	약사		

분석 데이터 목록



사회적 취약계층
데이터

사회적 취약계층 데이터 목록

구분	활용데이터	설명	컬럼
수급권자 정보	기초생활보장 수급권자	기초생활보장/기초연금/장애인연금대상자/ 차상위장애인/차상위본인 부담경감 대상자 수급권자수에 대한 데이터	기초연금 수급권자수, 장애인연금 수급권자수, 차상위장애인 수급권자수, 차상위본인 부담경감대상자 수급권자수, 시도코드
	기초연금 수급권자		
	장애인연금 대상자수급권자		
	차상위 장애인수급권자		
	차상위 본인부담경감 대상자수급권자		
장애인 인구정보	장애인수	장애인 인구수에 대한 데이터	장애인수_합계, 시도코드

데이터 정제

① 데이터 추출

- 지역별 인구 대비 의료 자원의 수를 비교하기 위해 이벤트건수, 대상자건수 추출

② 데이터 전처리

- 행정구역코드를 기준으로 데이터 변환
- 전국의 지수를 비교하기 위해 실제 사용 중인 행정구역코드로 지역코드 통일
- 데이터 클리닝 : 결측치 및 이상치 확인결과 해당사항 없음을 확인

③ 데이터 표준화

- 표준화 = (이벤트건수 ÷ 대상자 건수)

최종 데이터셋

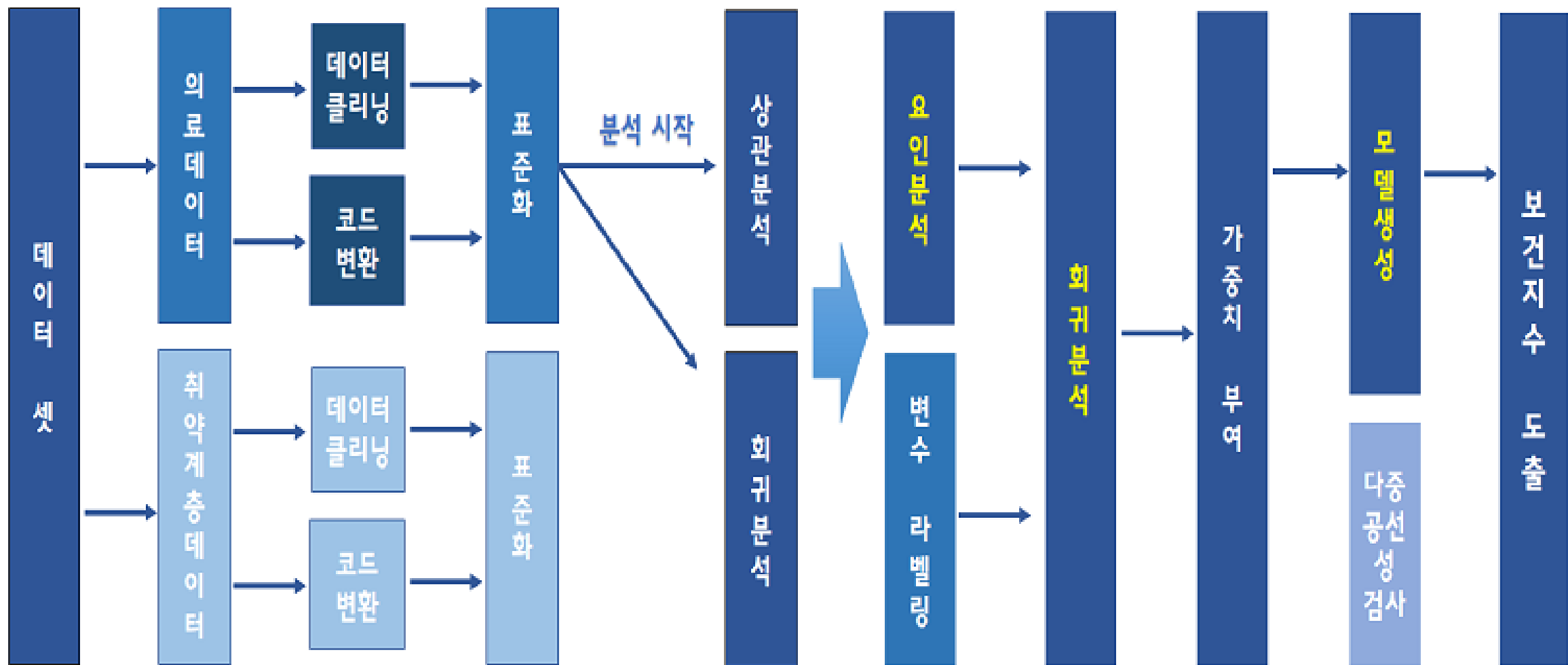
〈250개 지역(행) 32개 요인(열)〉

지역코드	지역명	의사_표준	간호사_표	약사_표준	전문의_표	전문병상_	허가병상_	의료기관_	의료장비_	권장신체활	음주율_표	흡연율_표	비만율_표	고도비만율	초고도비만	복부비만율	대사증후군	고혈압_표	당뇨병_표
11110	서울특별시종로구	0.0038	0.0047	0.0029	0.0094	0.0005	0.0023	0.0009	0.0099	0.0301	0.0234	0.0172	0.5045	0.0879	0.0119	0.3388	0.1583	0.2029	0.0831
11140	서울특별시중구	0.0024	0.0027	0.0021	0.0059	0.0004	0.0013	0.0013	0.0075	0.0259	0.0251	0.0187	0.4985	0.0896	0.0107	0.3461	0.1575	0.2138	0.0875
11170	서울특별시용산구	0.0009	0.0009	0.0008	0.0017	0.0001	0.0005	0.0005	0.0014	0.0271	0.0233	0.0178	0.4878	0.0856	0.0110	0.3328	0.1487	0.1987	0.0799
11200	서울특별시성동구	0.0006	0.0006	0.0008	0.0017	0.0001	0.0007	0.0004	0.0015	0.0281	0.0264	0.0188	0.4919	0.0856	0.0107	0.3304	0.1487	0.2024	0.0808
11215	서울특별시광진구	0.0006	0.0007	0.0008	0.0020	0.0001	0.0007	0.0004	0.0015	0.0289	0.0261	0.0186	0.4935	0.0890	0.0125	0.3238	0.1508	0.1866	0.0733
11230	서울특별시동대문구	0.0010	0.0013	0.0010	0.0028	0.0002	0.0019	0.0004	0.0030	0.0319	0.0259	0.0184	0.5137	0.0950	0.0128	0.3696	0.1699	0.2205	0.0908
11260	서울특별시중랑구	0.0005	0.0008	0.0007	0.0013	0.0001	0.0008	0.0004	0.0016	0.0309	0.0255	0.0201	0.5220	0.0960	0.0142	0.3481	0.1737	0.2284	0.0938
11290	서울특별시성북구	0.0006	0.0006	0.0007	0.0016	0.0001	0.0008	0.0004	0.0015	0.0297	0.0257	0.0186	0.5102	0.0903	0.0132	0.3456	0.1627	0.2175	0.0874
11305	서울특별시강북구	0.0004	0.0003	0.0007	0.0013	0.0001	0.0007	0.0004	0.0010	0.0313	0.0253	0.0196	0.5199	0.0977	0.0149	0.3427	0.1613	0.2489	0.0996
11320	서울특별시도봉구	0.0004	0.0005	0.0006	0.0011	0.0001	0.0010	0.0003	0.0010	0.0308	0.0239	0.0189	0.5150	0.0920	0.0126	0.3367	0.1637	0.2345	0.0932
11350	서울특별시노원구	0.0007	0.0008	0.0007	0.0018	0.0001	0.0008	0.0004	0.0023	0.0310	0.0246	0.0180	0.5073	0.0915	0.0129	0.3332	0.1463	0.2155	0.0839
11380	서울특별시은평구	0.0004	0.0004	0.0006	0.0012	0.0000	0.0008	0.0004	0.0008	0.0299	0.0250	0.0192	0.5093	0.0940	0.0135	0.3366	0.1615	0.2253	0.0905
11410	서울특별시서대문구	0.0011	0.0018	0.0012	0.0033	0.0003	0.0013	0.0004	0.0039	0.0294	0.0239	0.0178	0.5092	0.0945	0.0127	0.3356	0.1590	0.2150	0.0865
11440	서울특별시마포구	0.0006	0.0002	0.0007	0.0014	0.0000	0.0002	0.0006	0.0008	0.0282	0.0230	0.0177	0.4925	0.0898	0.0126	0.3187	0.1488	0.1868	0.0732
11470	서울특별시양천구	0.0006	0.0007	0.0007	0.0017	0.0001	0.0006	0.0004	0.0015	0.0300	0.0255	0.0182	0.5064	0.0896	0.0120	0.3529	0.1510	0.1953	0.0766
11500	서울특별시강서구	0.0004	0.0005	0.0006	0.0013	0.0001	0.0007	0.0004	0.0013	0.0290	0.0265	0.0192	0.5065	0.0940	0.0128	0.3538	0.1560	0.1971	0.0802
11530	서울특별시구로구	0.0005	0.0007	0.0008	0.0018	0.0001	0.0010	0.0003	0.0016	0.0296	0.0229	0.0184	0.5114	0.0901	0.0122	0.3390	0.1721	0.2118	0.0809
11545	서울특별시금천구	0.0005	0.0004	0.0006	0.0012	0.0000	0.0009	0.0004	0.0010	0.0306	0.0261	0.0203	0.5323	0.1002	0.0152	0.3463	0.1619	0.2175	0.0864
11560	서울특별시영등포구	0.0010	0.0013	0.0011	0.0028	0.0002	0.0015	0.0005	0.0035	0.0297	0.0243	0.0180	0.4932	0.0847	0.0115	0.3422	0.1574	0.1999	0.0750
11590	서울특별시동작구	0.0006	0.0009	0.0009	0.0020	0.0001	0.0008	0.0004	0.0016	0.0291	0.0229	0.0166	0.4839	0.0811	0.0114	0.3164	0.1446	0.1919	0.0733
11620	서울특별시관악구	0.0004	0.0003	0.0006	0.0011	0.0000	0.0005	0.0004	0.0009	0.0302	0.0239	0.0181	0.4969	0.0931	0.0136	0.3194	0.1562	0.1838	0.0718
11650	서울특별시서초구	0.0012	0.0015	0.0010	0.0035	0.0002	0.0008	0.0006	0.0027	0.0275	0.0223	0.0154	0.4540	0.0744	0.0107	0.2957	0.1247	0.1631	0.0608
11680	서울특별시강남구	0.0018	0.0021	0.0015	0.0060	0.0003	0.0015	0.0009	0.0062	0.0281	0.0226	0.0159	0.4521	0.0729	0.0093	0.2934	0.1286	0.1584	0.0587
11710	서울특별시송파구	0.0009	0.0013	0.0010	0.0029	0.0002	0.0010	0.0004	0.0025	0.0294	0.0249	0.0169	0.4752	0.0791	0.0104	0.3050	0.1371	0.1746	0.0650
11740	서울특별시강동구	0.0009	0.0012	0.0011	0.0026	0.0002	0.0014	0.0005	0.0024	0.0305	0.0250	0.0181	0.4954	0.0873	0.0123	0.3243	0.1557	0.2022	0.0785

- 의사 수, 전문병상 수, 의료기관 수 등 의료 관련 32개 항목
- 서울특별시, 강원도 등 전국 250 개의 지역

3. 분석 프로세스

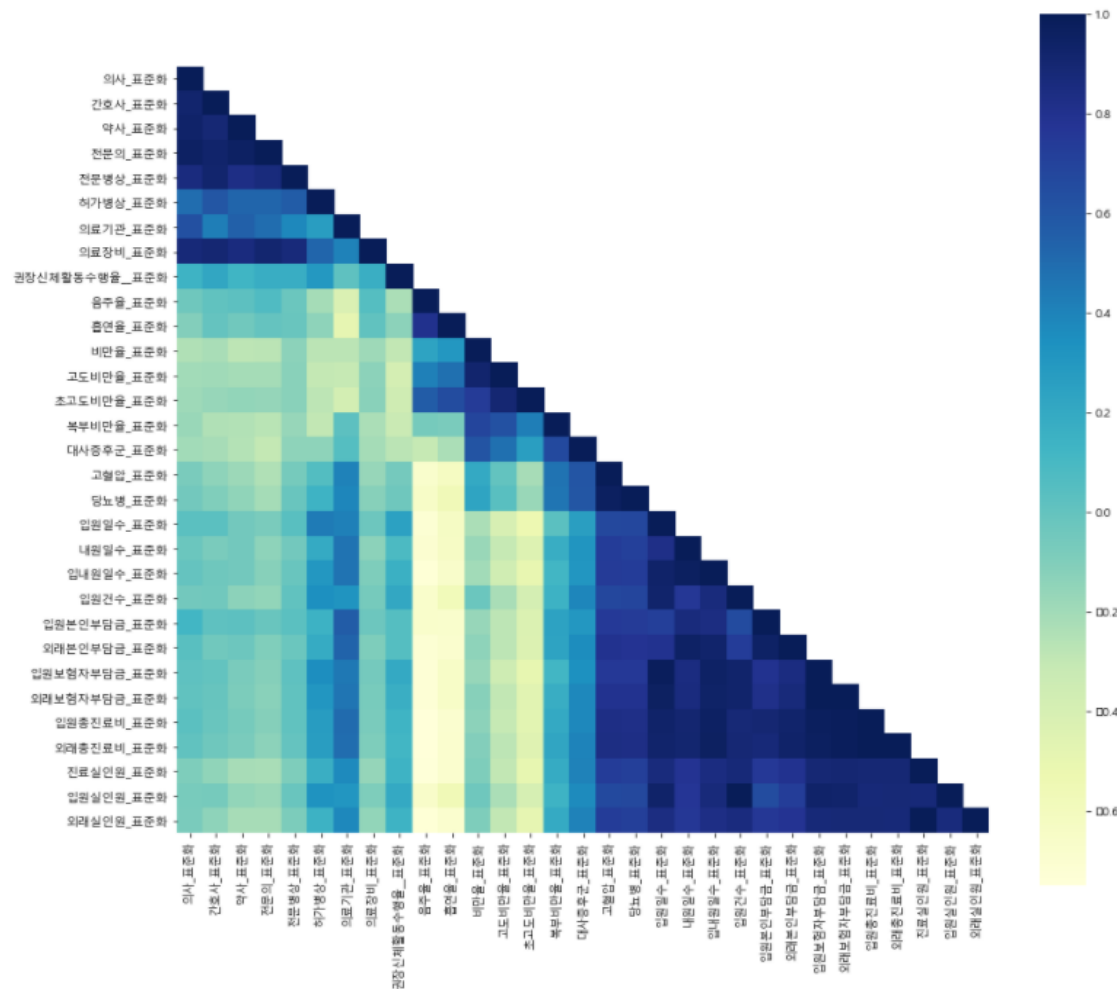
분석 프로세스 도식화



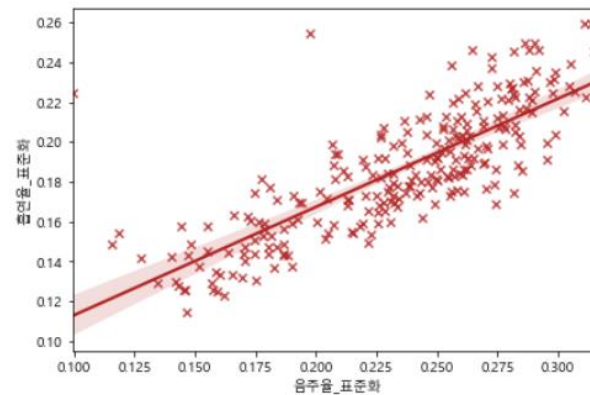
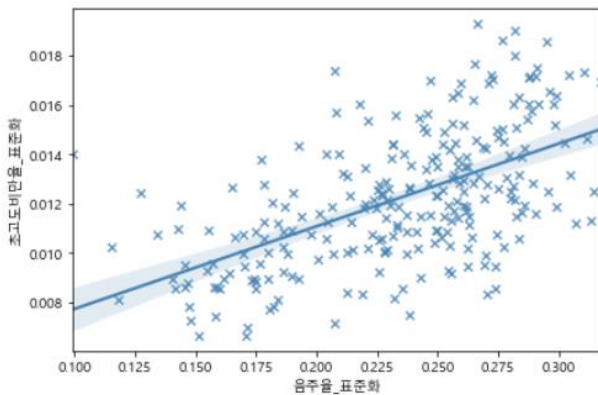
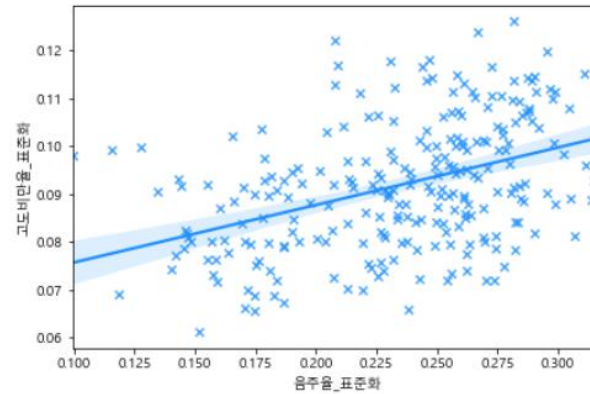
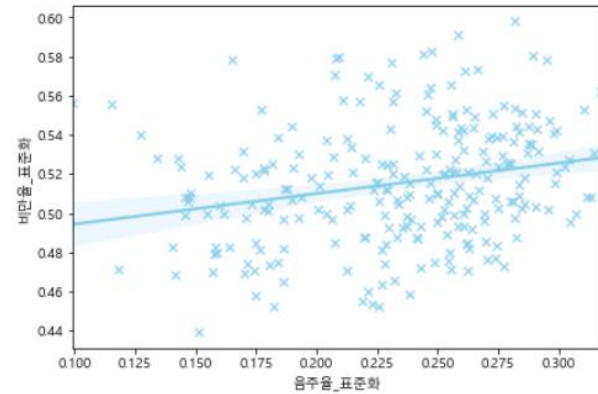
상관분석(Python)

- 32개의 변수 간 연관성 여부 확인
- 음주율, 흡연율과 고혈압 간의 관계는 0.006 정도로 현저히 낮은 것으로 나타남
- 의료장비와 의사, 간호사, 약사, 전문의 간의 관계는 0.8이상으로 매우 높게 나타남
- 특정 요소들 간의 상관관계가 매우 높게 나타난 것으로 보아 다중공선성이 발생의 가능성이 존재

<matplotlib.axes._subplots.AxesSubplot at 0x2988e8e2f70>

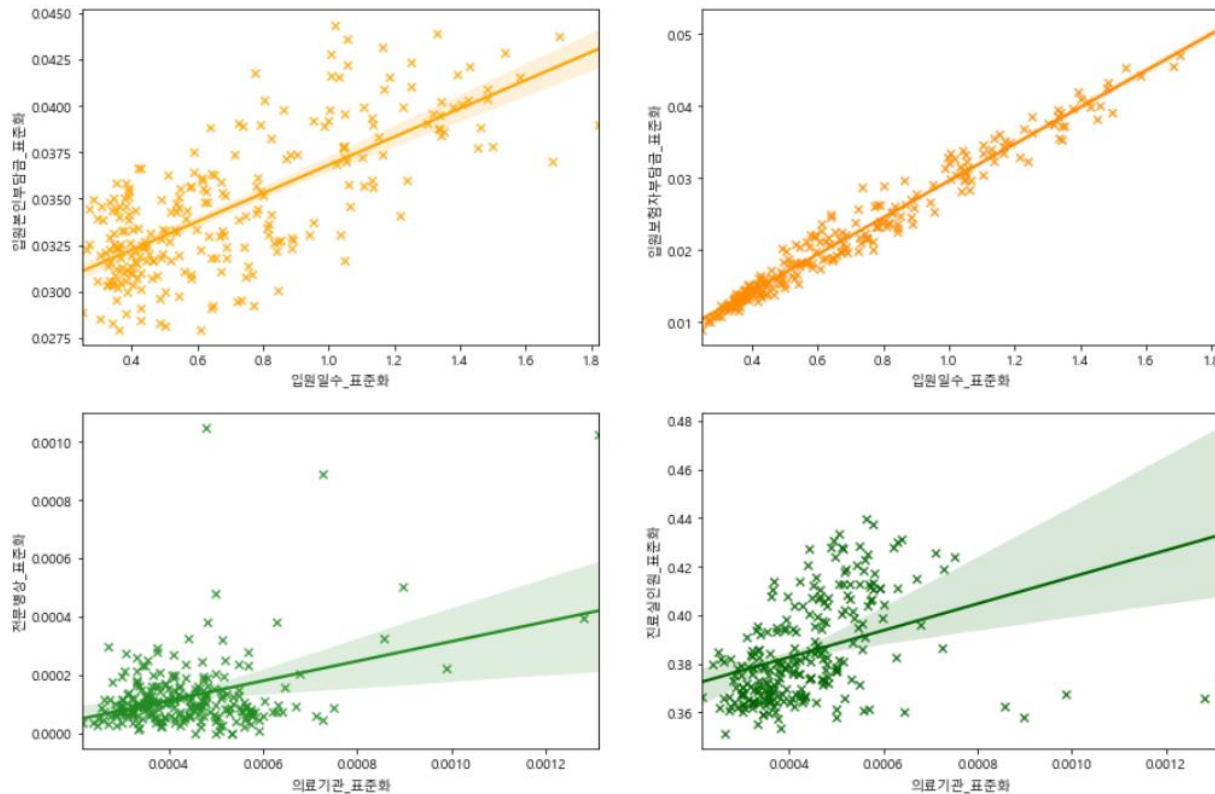


회귀분석(Python)



- 높은 상관관계를 보이는 두 변수 간 관계 예측(단순선형회귀분석)
- 음주율이 높을수록 비만율, 고도비만율, 초고도비만율이 상승
- 음주율에 비례한 비만율의 정도는 비만율, 고도비만율, 초고도비만율로 갈수록 증가
- 음주율이 높으면 흡연율의 정도도 높음을 확인

회귀분석(Python)



- 입원일수가 늘어날수록 입원본인부담금과 입원보험자부담금은 증가
- 의료기관이 많을수록 전문병상의 개수와 진료실인원이 늘어남

KMO 테스트

```
> KMO(data)
```

```
Kaiser-Meyer-Olkin factor adequacy
```

```
Call: KMO(r = data)
```

```
Overall MSA = 0.5
```

```
MSA for each item =
```

의사_표준화	간호사_표준화	약사_표준화	
0.5	0.5	0.5	
전문병상_표준화	허가병상_표준화	의료기관_표준화	의료
0.5	0.5	0.5	
권장신체활동수행율_표준화	음주율_표준화	흡연율_표준화	비
0.5	0.5	0.5	
고도비만을_표준화	초고도비만을_표준화	복부비만을_표준화	대사증후
0.5	0.5	0.5	
고혈압_표준화	당뇨병_표준화	입원일수_표준화	내
0.5	0.5	0.5	
입내원일수_표준화	입원건수_표준화	입원본인부담금_표준화	외래본인부담
0.5	0.5	0.5	
입원보험자부담금_표준화	외래보험자부담금_표준화	입원총진료비_표준화	외래총진료비
0.5	0.5	0.5	
진료실인원_표준화	입원실인원_표준화	외래실인원_표준화	
0.5	0.5	0.5	

- KMO 테스트는 0.5이상이면 요인분석으로 분석하기에 적절하다고 판단할 수 있으나, 해당 결과에는 0.5미만인 값은 존재하지 않음

요인분석(R)

- 총 4개의 요소로 그룹화, varimax 회전법 사용
- h2는 각 변수의 공통성을 의미하는 것으로 0.3이하는 다른 변수들과 공통점이 별로 없는 것으로 판단, 해당 결과에는 0.3이하는 존재하지 않음
- Proportion var는 각 요인이 설명하는 총 분산의 비율
- RC1은 총 분산의 46%, RC2는 19%, RC3는 13%, RC4는 6%로 이 네 요인이 총 변동의 84%를 설명한다고 해석 가능

```
> Varimax
Principal Components Analysis
Call: principal(r = data, nfactors = 4, rotate = "varimax")
Standardized loadings (pattern matrix) based upon correlation matrix
```

	RC1	RC2	RC3	RC4	h2	u2	com
의사_표준화	-0.02	0.95	-0.10	-0.18	0.95	0.047	1.1
간호사_표준화	-0.04	0.96	-0.10	0.04	0.94	0.063	1.0
약사_표준화	-0.11	0.94	-0.15	-0.14	0.93	0.072	1.1
전공_의_표준화	-0.15	0.96	-0.16	-0.07	0.97	0.033	1.1
전공_의_의_표준화	0.00	0.93	-0.01	0.05	0.87	0.130	1.0
외과_의_표준화	0.32	0.63	-0.20	0.40	0.70	0.300	2.5
의류기_표준화	0.43	0.53	-0.11	-0.50	0.73	0.272	3.0
의류기_의_표준화	-0.09	0.93	-0.07	0.02	0.88	0.116	1.0
간호사_의_표준화	0.18	0.17	-0.43	0.35	0.37	0.632	2.7
간호사_의_의_표준화	-0.77	0.00	0.20	0.25	0.70	0.303	1.4
간호사_의_의_의_표준화	-0.69	-0.02	0.27	0.43	0.73	0.267	2.1
간호사_의_의_의_의_표준화	-0.07	-0.14	0.93	0.12	0.90	0.096	1.1
간호사_의_의_의_의_의_표준화	-0.27	-0.10	0.92	0.12	0.95	0.052	1.2
간호사_의_의_의_의_의_의_표준화	-0.46	-0.09	0.77	0.18	0.84	0.157	1.8
간호사_의_의_의_의_의_의_의_표준화	0.21	-0.15	0.79	-0.27	0.77	0.232	1.5
간호사_의_의_의_의_의_의_의_의_표준화	0.43	-0.14	0.71	-0.10	0.72	0.281	1.8
간호사_의_의_의_의_의_의_의_의_의_표준화	0.82	-0.08	0.30	-0.33	0.88	0.122	1.6
간호사_의_의_의_의_의_의_의_의_의_의_표준화	0.82	-0.02	0.33	-0.26	0.85	0.145	1.5
간호사_의_의_의_의_의_의_의_의_의_의_의_표준화	0.96	0.05	-0.17	0.13	0.96	0.039	1.1
간호사_의_의_의_의_의_의_의_의_의_의_의_의_표준화	0.89	-0.04	-0.07	-0.18	0.83	0.171	1.1
간호사_의_의_의_의_의_의_의_의_의_의_의_의_의_표준화	0.96	0.00	-0.11	-0.06	0.93	0.069	1.0
간호사_의_의_의_의_의_의_의_의_의_의_의_의_의_의_표준화	0.94	0.00	0.01	0.21	0.93	0.069	1.1
간호사_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_표준화	0.83	0.06	-0.03	-0.39	0.84	0.158	1.4
간호사_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_표준화	0.89	0.00	-0.02	-0.31	0.89	0.112	1.2
간호사_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_표준화	0.98	0.03	-0.09	0.04	0.97	0.030	1.0
간호사_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_표준화	0.98	0.02	-0.03	0.00	0.97	0.030	1.0
간호사_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_의_표준화	0.98	0.02	-0.03	-0.12	0.98	0.023	1.0
간호사_의_표준화	0.98	0.01	0.00	-0.11	0.98	0.021	1.0
간호사_의_표준화	0.93	-0.12	-0.05	-0.02	0.88	0.120	1.0
간호사_의_표준화	0.94	-0.03	-0.01	0.23	0.93	0.068	1.1
간호사_의_표준화	0.92	-0.11	-0.05	-0.05	0.86	0.143	1.0

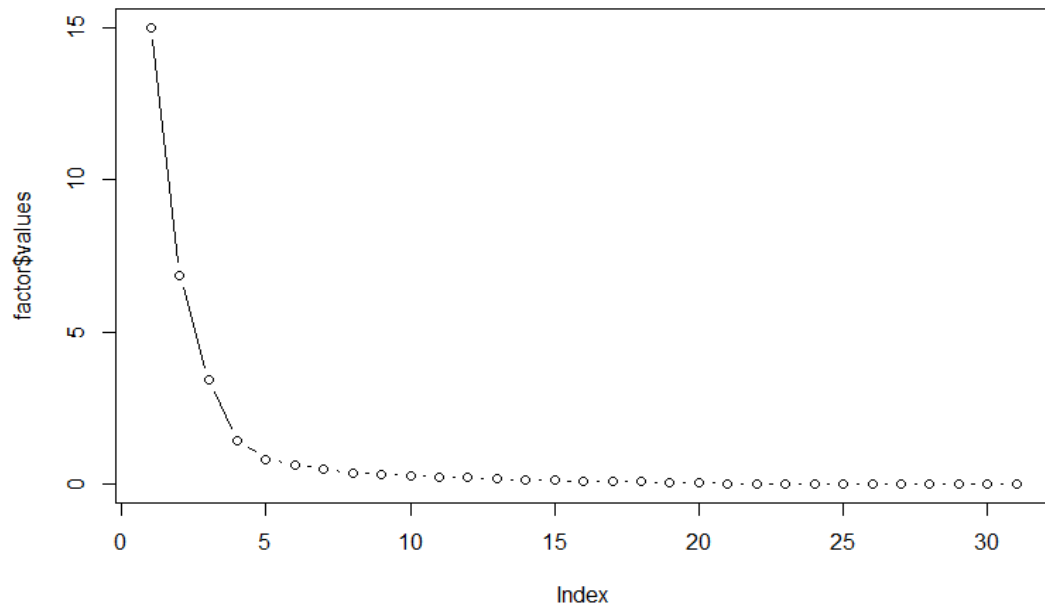
```
SS loadings          RC1  RC2  RC3  RC4
Proportion var       0.46 0.19 0.13 0.06
Cumulative var       0.46 0.66 0.79 0.85
Proportion Explained 0.55 0.23 0.15 0.07
Cumulative Proportion 0.55 0.77 0.93 1.00
```

```
Mean item complexity = 1.4
Test of the hypothesis that 4 components are sufficient.
```

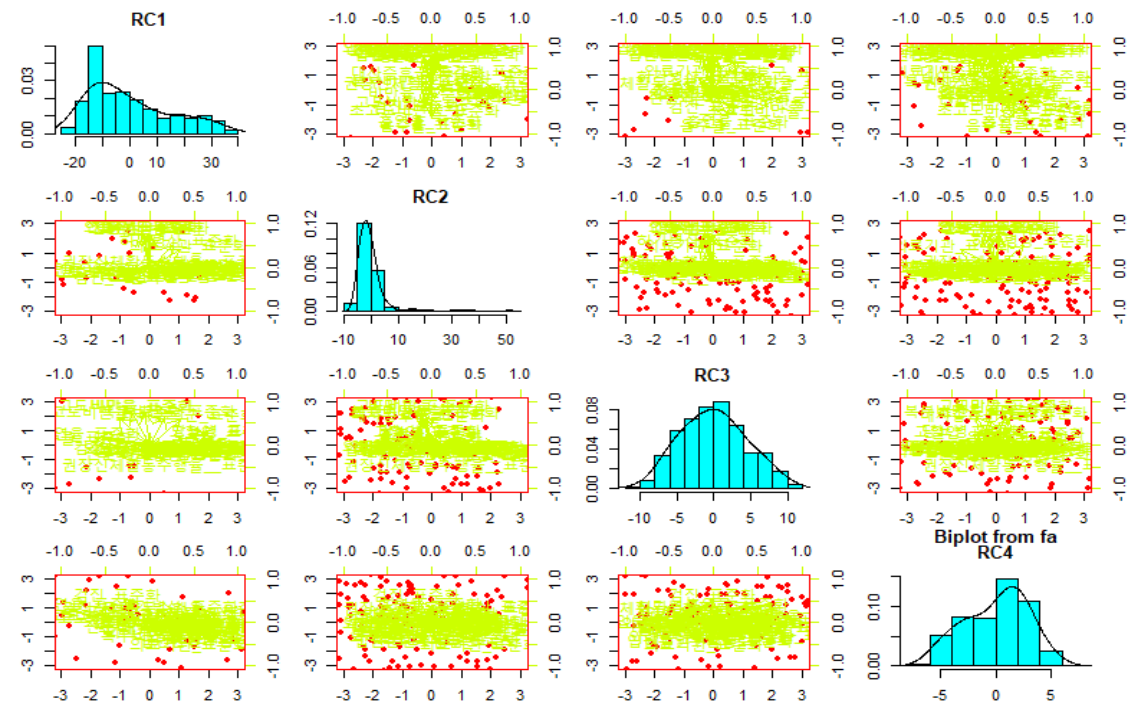
```
The root mean square of the residuals (RMSR) is 0.03
with the empirical chi square 241.78 with prob < 1
```

```
Fit based upon off diagonal values = 1
```

요인분석(R)



- 위의 'scree plot'을 사용하여 적절한 요인변수의 개수를 찾을 수 있음
- 처음 4개의 요인 값이 고유 값(Y축)에서 1 이상을 가지기 때문에 4개의 factor 생성하는 것이 적절함을 확인



- 결과 값이 평행할수록 상관관계가 높다고
- 볼 수 있는데 해당 그래프는 결과 값들이 평행하게 나왔으므로 서로 간의 연관성이 있는 것으로 보임

변수 라벨링

그룹	요소	그룹화 명
1	고혈압, 당뇨병, 입원일수, 내원일수, 입내원일수, 입원건수, 입원본인부담금, 외래본인부담금, 입원보험자부담금, 외래보험자부담금, 입원총진료비, 외래총진료비, 진료실인원, 입원실인원, 외래실인원	의료소비
2	의사, 간호사, 약사, 전문의, 전문병사, 허가병사, 의료기관, 의료장비	의료제공
3	비만율, 고도비만율, 초고도비만율, 복부비만율, 대사증후군	의료위험군
4	권장신체활동수행율, 음주율, 흡연율	개인적 습관

회귀분석(R)

- 보건지수 제안을 위해 각 변수 간의 가중치를 둘 필요가 있음
- 라벨링 한 변수의 회귀계수 값을 가중치 값으로 사용할 예정



- 설명력이 가장 높은 회귀식 선정
- 선정한 회귀식의 회귀계수로 가중치 부여
- 부여한 가중치를 기준으로 보건지수 예측 모델을 생성

김호곤. 2013. "회귀계수를 이용한 디자인 요소의 가중치 부여 방법", 『한국디자인트렌드학회』

김태일. 1999. "수리적 기법에 의한 평가모형체계의 가중치 부여방식에 관한 논의", 『한국행정학보』

World Health Organization : WHO. 2019. "GHS INDEX : GLOBAL HEALTH SECURITY INDEX", 『WHO』

회귀분석(R)

```
> summary(a)
```

```
Call:
lm(formula = 의료소비 ~ 의료제공 + 의료위험군 + 개인적습관, data = data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-0.23211 -0.06747 -0.01718  0.05422  0.32469
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.3407      0.1071  12.515 < 2e-16 ***
의료제공     20.0310      3.9208   5.109 6.5e-07 ***
의료위험군   0.1491      0.3630   0.411  0.682
개인적습관  -3.1976      0.2346 -13.632 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.09416 on 246 degrees of freedom
Multiple R-squared:  0.6823,    Adjusted R-squared:  0.6749
F-statistic: 72.37 on 3 and 246 DF,  p-value: < 2.2e-16
```

```
> summary(c)
```

```
Call:
lm(formula = 의료위험군 ~ 의료소비 + 의료제공 + 개인적습관, data = data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-0.04362 -0.01133 -0.00018  0.01138  0.03833
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.245006   0.018308  13.382 < 2e-16 ***
의료소비     0.004597   0.011192   0.411  0.682
의료제공    -3.525435   0.688312  -5.122 6.11e-07 ***
개인적습관  -0.059063   0.054442  -1.085  0.279
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.01653 on 246 degrees of freedom
Multiple R-squared:  0.3719,    Adjusted R-squared:  0.3642
F-statistic: 10.38 on 3 and 246 DF,  p-value: 1.852e-06
```

```
> summary(b)
```

```
Call:
lm(formula = 의료제공 ~ 의료소비 + 의료위험군 + 개인적습관, data = data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-0.0032876 -0.0008166 -0.0001833  0.0004726  0.0083344
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0023096  0.0021140   1.093  0.27567
의료소비     0.0047887  0.0009373   5.109 6.50e-07 ***
의료위험군  -0.0273338  0.0053367  -5.122 6.11e-07 ***
개인적습관   0.0134164  0.0047285   2.837  0.00493 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.001456 on 246 degrees of freedom
Multiple R-squared:  0.4171,    Adjusted R-squared:  0.4098
F-statistic: 18.91 on 3 and 246 DF,  p-value: 4.554e-11
```

```
> summary(d)
```

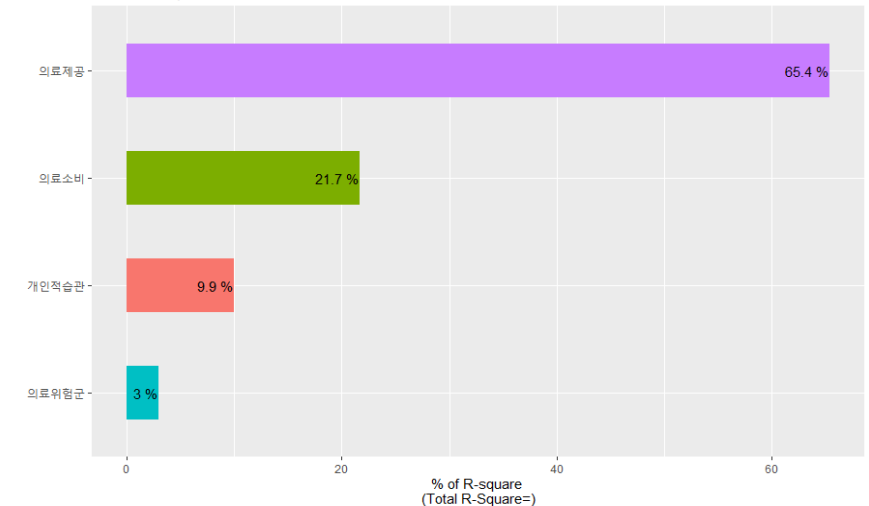
```
Call:
lm(formula = 개인적습관 ~ 의료소비 + 의료제공 + 의료위험군, data = data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-0.044317 -0.014486 -0.000452  0.011941  0.062365
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.338560   0.018018  18.791 < 2e-16 ***
의료소비    -0.134576   0.009872 -13.632 < 2e-16 ***
의료제공     2.361916   0.832444   2.837  0.00493 **
의료위험군  -0.080618   0.074311  -1.085  0.27904
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.01932 on 246 degrees of freedom
Multiple R-squared:  0.6452,    Adjusted R-squared:  0.6357
F-statistic: 63.84 on 3 and 246 DF,  p-value: < 2.2e-16
```

Relative Importance of Predictor Variables



- 총 4개의 모델 중에 Adjusted R-squared가 67.49%로 가장 높은 첫번째 회귀식을 사용하기로 결정
- 4가지 요소들의 상대적 중요도(R-square값)를 그래프로 시각화

최종 회귀식 도출

〈 최종 모델 〉

$$Y = X_1 + (20.0310)X_2 + (0.1491)X_3 + (-3.1976)X_4$$

- 가중치를 부여하여 모델을 도출한 결과, X2(의료공급)가 보건지수에 가장 큰 영향을 주는 것으로 나타남

김호곤. 2013. "회귀계수를 이용한 디자인 요소의 가중치 부여 방법", 『한국디자인트렌드학회』

김태일. 1999. "수리적 기법에 의한 평가모형체계의 가중치 부여방식에 관한 논의", 『한국행정학보』

World Health Organization : WHO. 2019. "GHS INDEX : GLOBAL HEALTH SECURITY INDEX", 『WHO』

다중공선성 검정

```
> vif(model)
의료소비      의료제공  의료위험군  개인적습관
1.882603     1.230593     1.126642     1.778570
```

- 사용 패키지 및 함수: `car`, `vif()`
- 다중공선성을 나타내는 수치인 VIF(분산팽창요인)이 클수록 다중공선성이 높음
- VIF 값이 10 이상이면 다중공선성이 존재한다고 간주
- `vif()` 함수는 $GVIF^{(1/(2*Df))}$ 값을 반환하므로 해당 함수의 결과 값이 2 이하이면 다중공선성에 의한 문제가 없다고 판단

4. 분석 결과

보건지수 도출

지역코드	지역명	보건지수
11110	서울특별시종로구	0.683
11140	서울특별시중구	0.696
11170	서울특별시용산구	0.616
11200	서울특별시성동구	0.649
11215	서울특별시광진구	0.615
11230	서울특별시동대문구	0.668
11260	서울특별시중랑구	0.683
11290	서울특별시성북구	0.643
11305	서울특별시강북구	0.692
11320	서울특별시도봉구	0.657
11350	서울특별시노원구	0.634

보건지수 도출

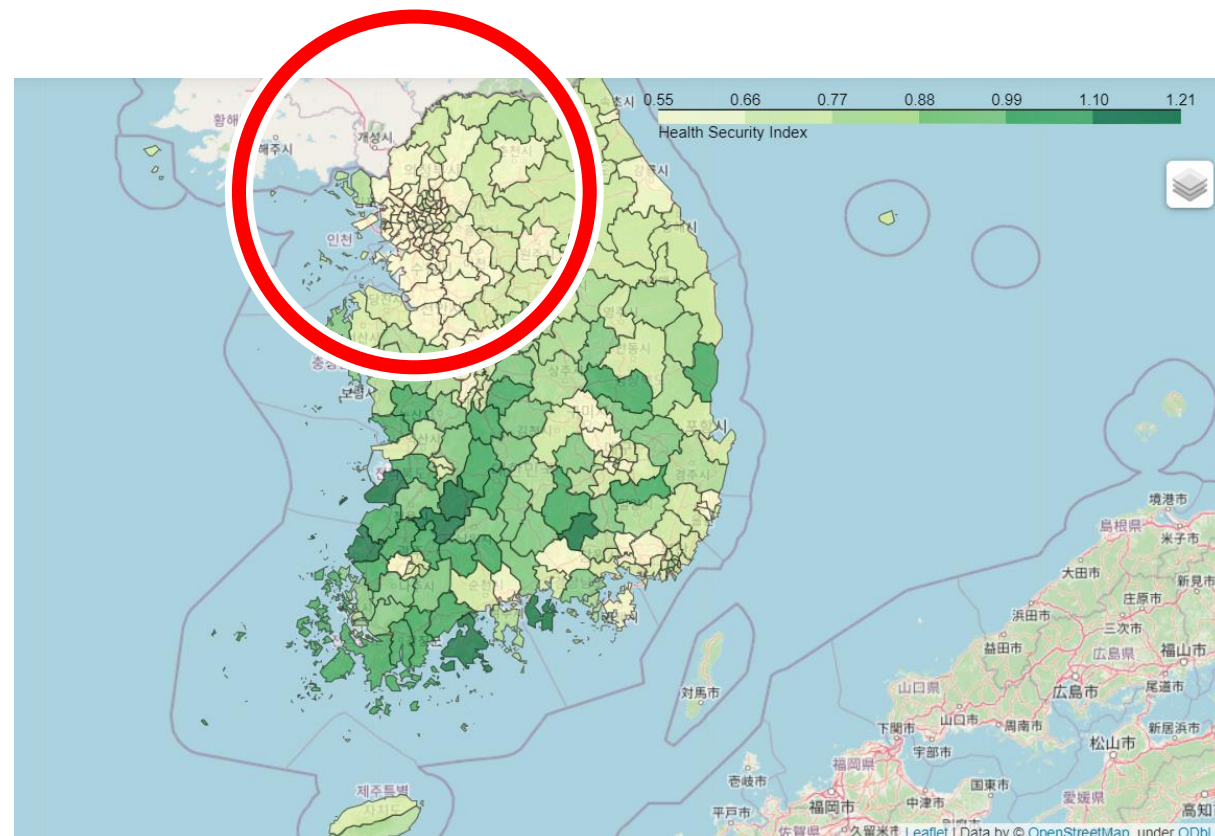
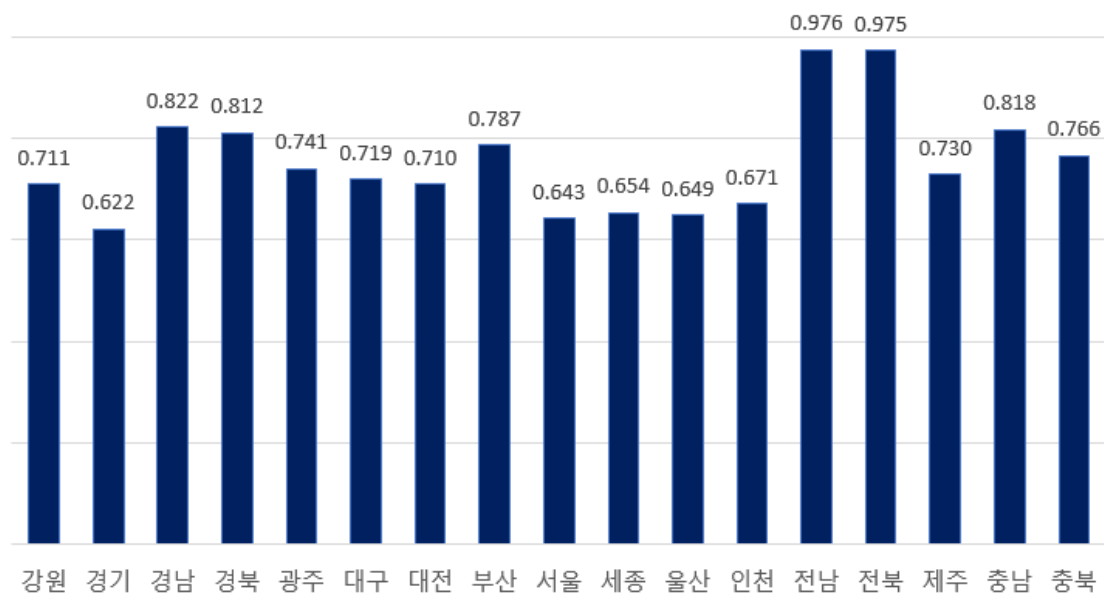
최소값	최대값	평균
0.548	1.214	0.761

시도	평균 보건지수
강원도	0.711
경기도	0.622
경상남도	0.822
경상북도	0.812
광주광역시	0.741
대구광역시	0.719
대전광역시	0.710
부산광역시	0.787
서울특별시	0.643

시도	평균 보건지수
세종특별시	0.654
울산광역시	0.649
인천광역시	0.671
전라남도	0.976
전라북도	0.975
제주특별시	0.730
충청남도	0.818
충청북도	0.766

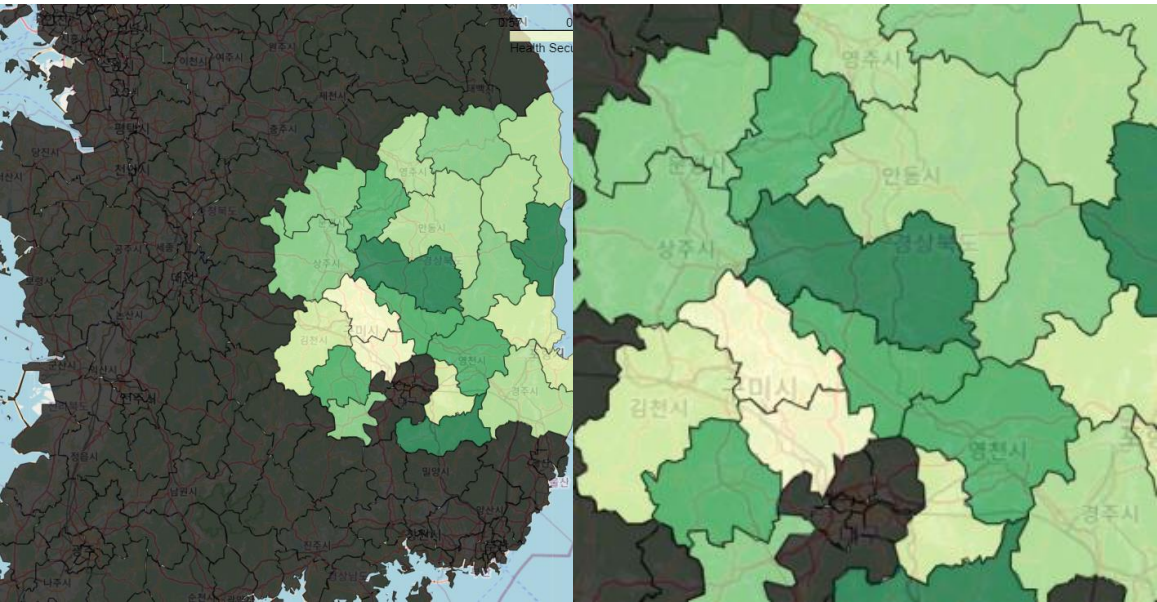
전국 보건지수 시각화

〈전국 보건지수 평균〉



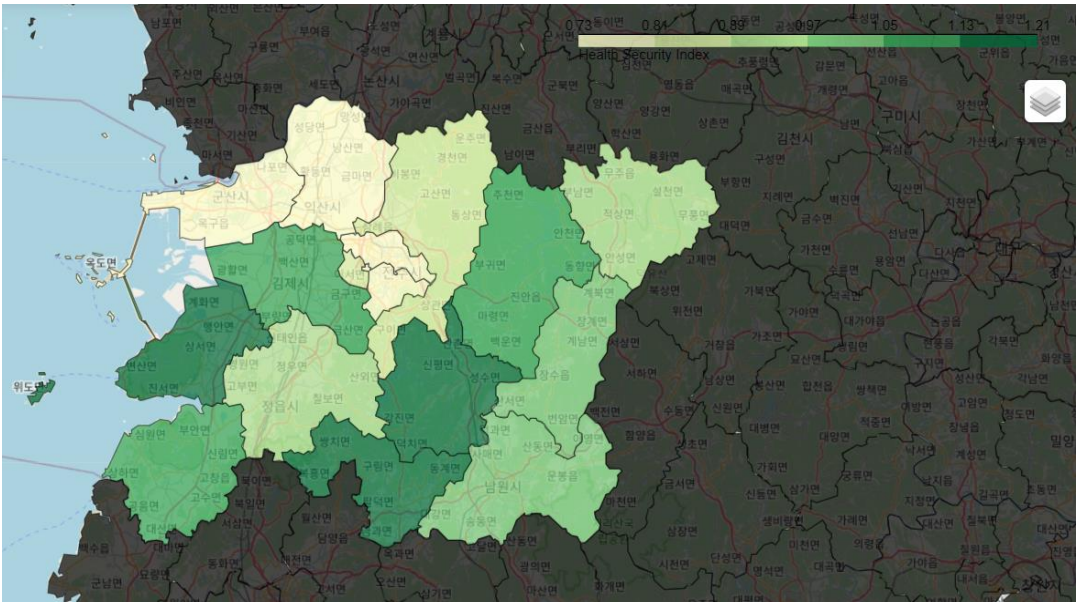
각 도별 보건지수 시각화

〈경상북도 구미시, 청도군〉



구분	지역명	보건지수
경상북도	청도군	1.041
	구미시	0.570

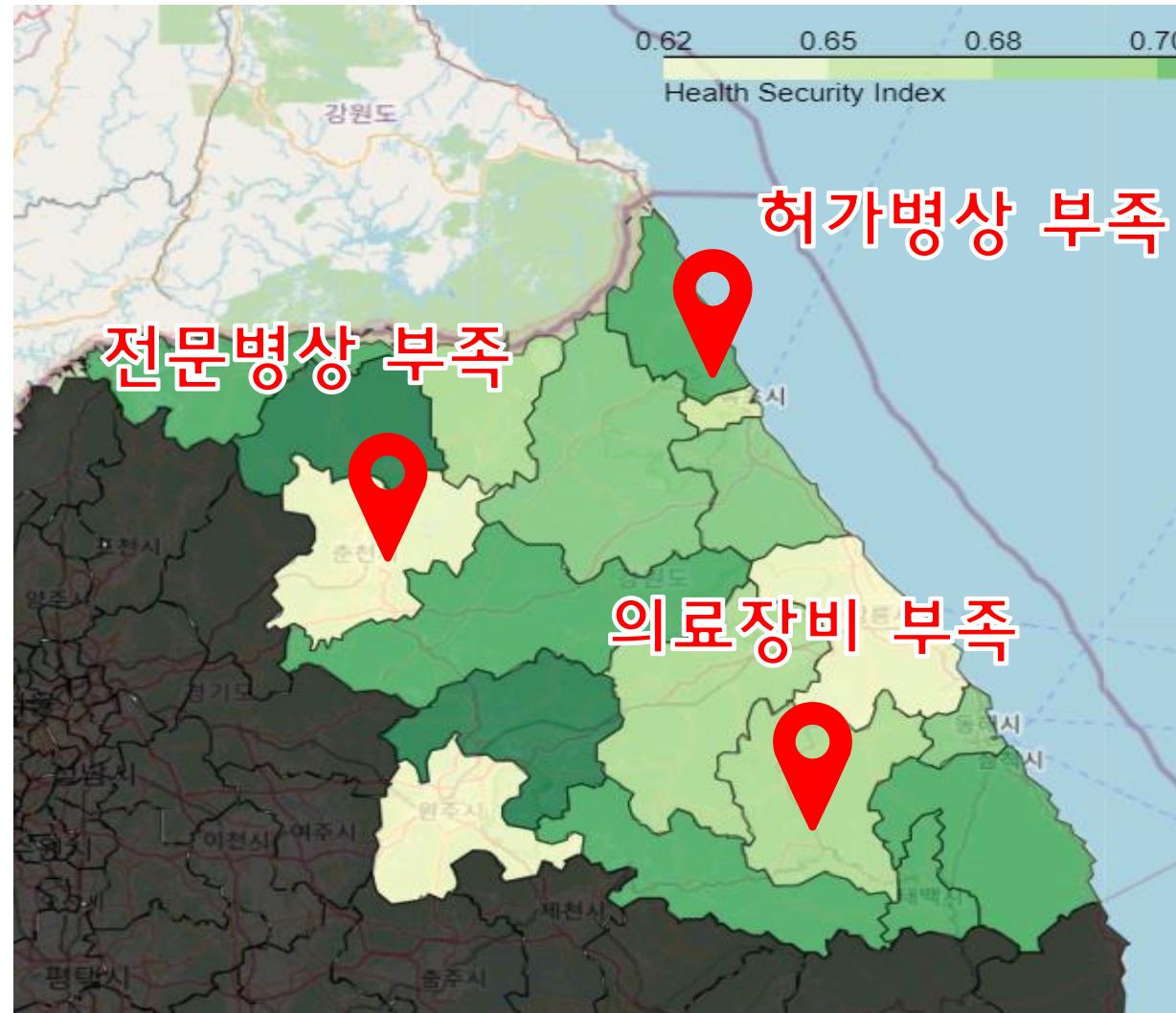
〈전라북도〉



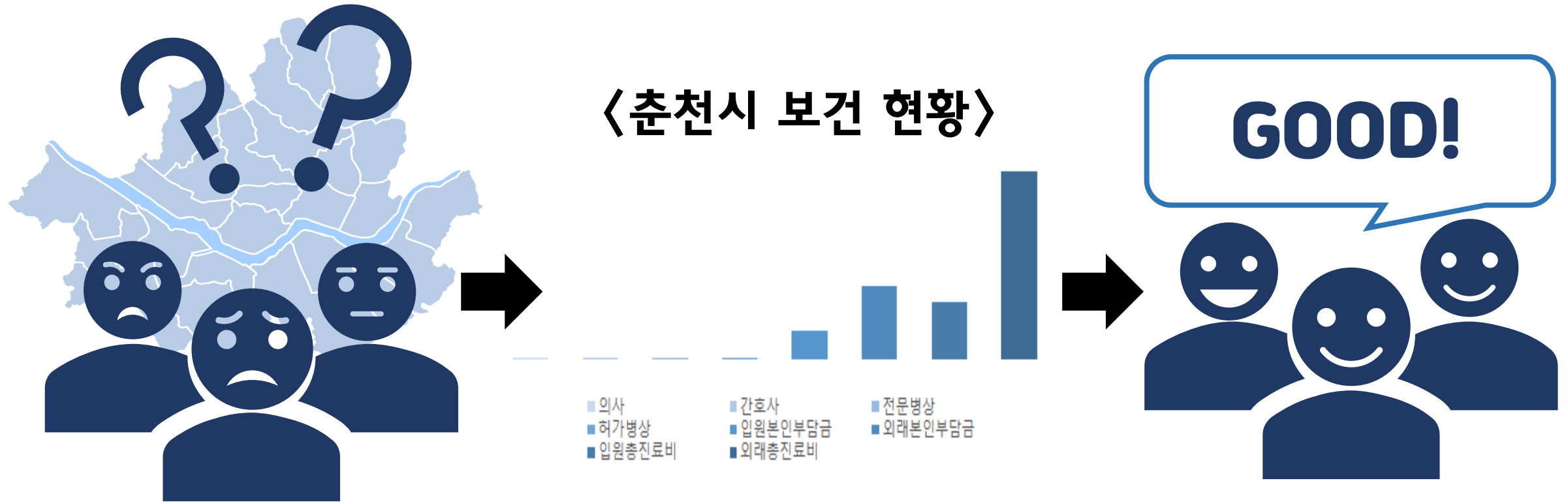
구분	지역명	보건지수
보건 최상위 지역	전라북도 순창군	1.214
보건 최하위 지역	전라북도 완산군	0.727

5. 활용 방안

활용 방안 1



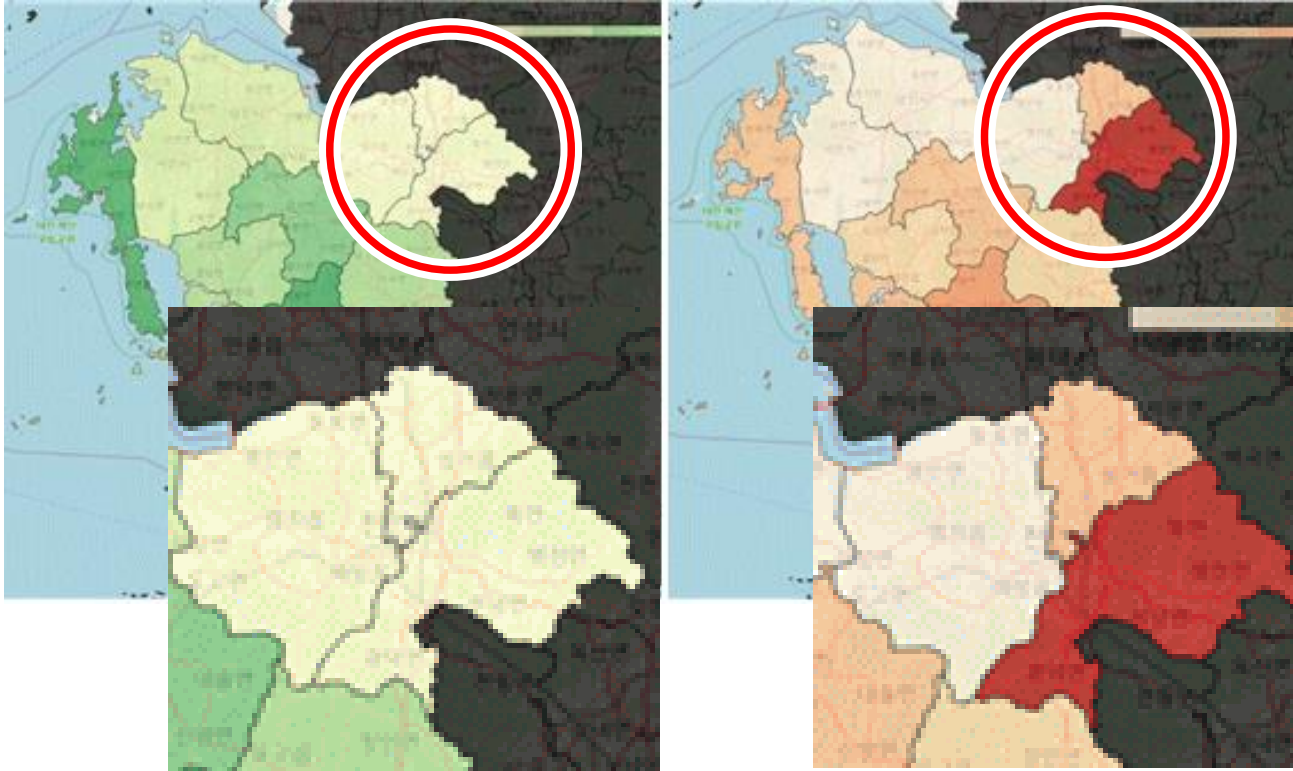
활용 방안 2



일반국민 누구라도 직관적으로 상세한 보건현황을 이해할 수 있음

활용 방안 3

〈충청남도 천안시 동남구〉

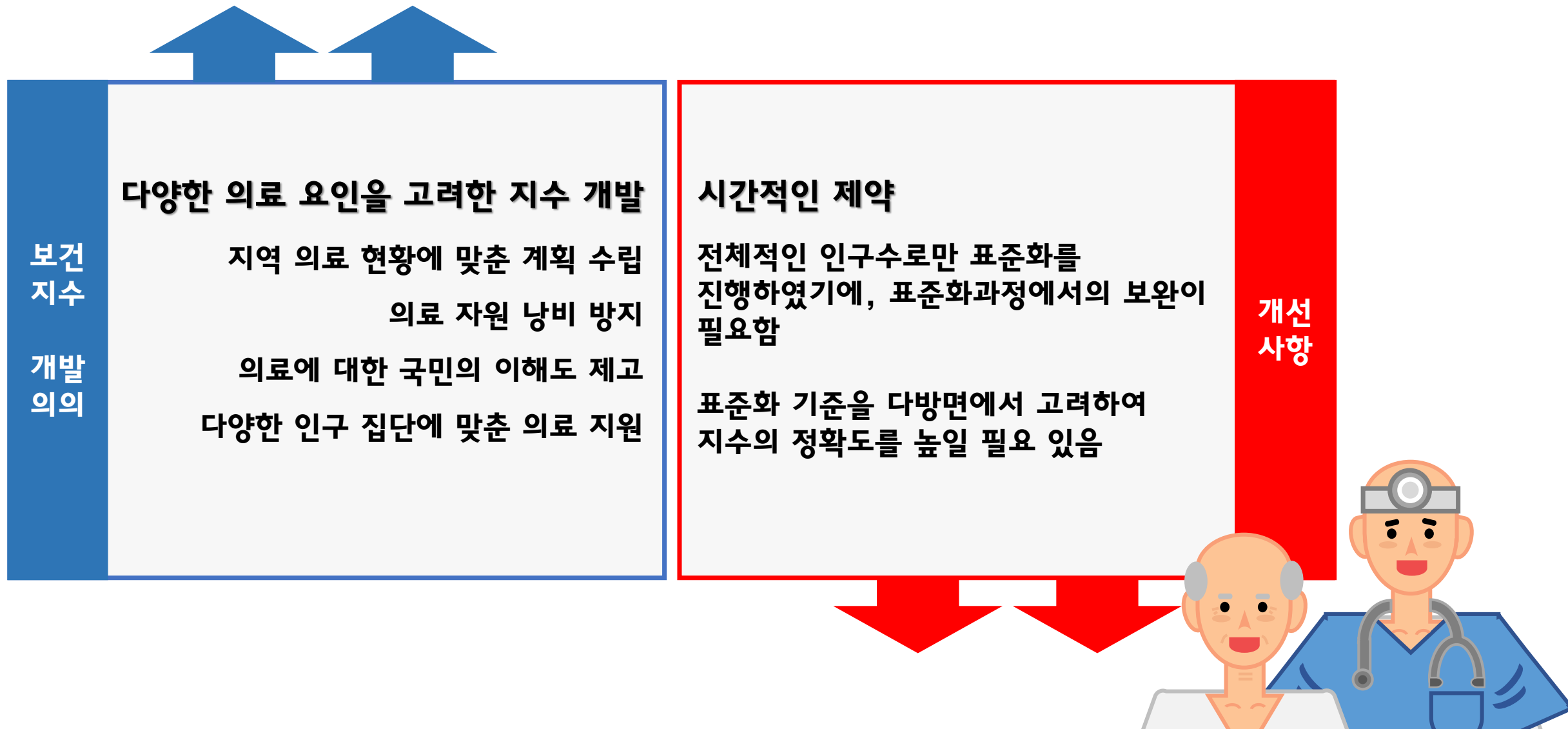


• 보건지도

• 취약계층 분포도

- 충청남도 천안시 동남구는 취약계층 분포도가 높으나, 보건지수가 낮은 지역으로 분석됨
- 특히 취약계층 중 65세 이상 노인인구 수가 전국 평균 노인인구 수보다 높은 것으로 집계됨
- 해당 지역의 경우 인구대비 전문병상의 수가 낮아 보건지수가 낮은 것으로 나타남

6. 결론



감사합니다
THANK YOU

