

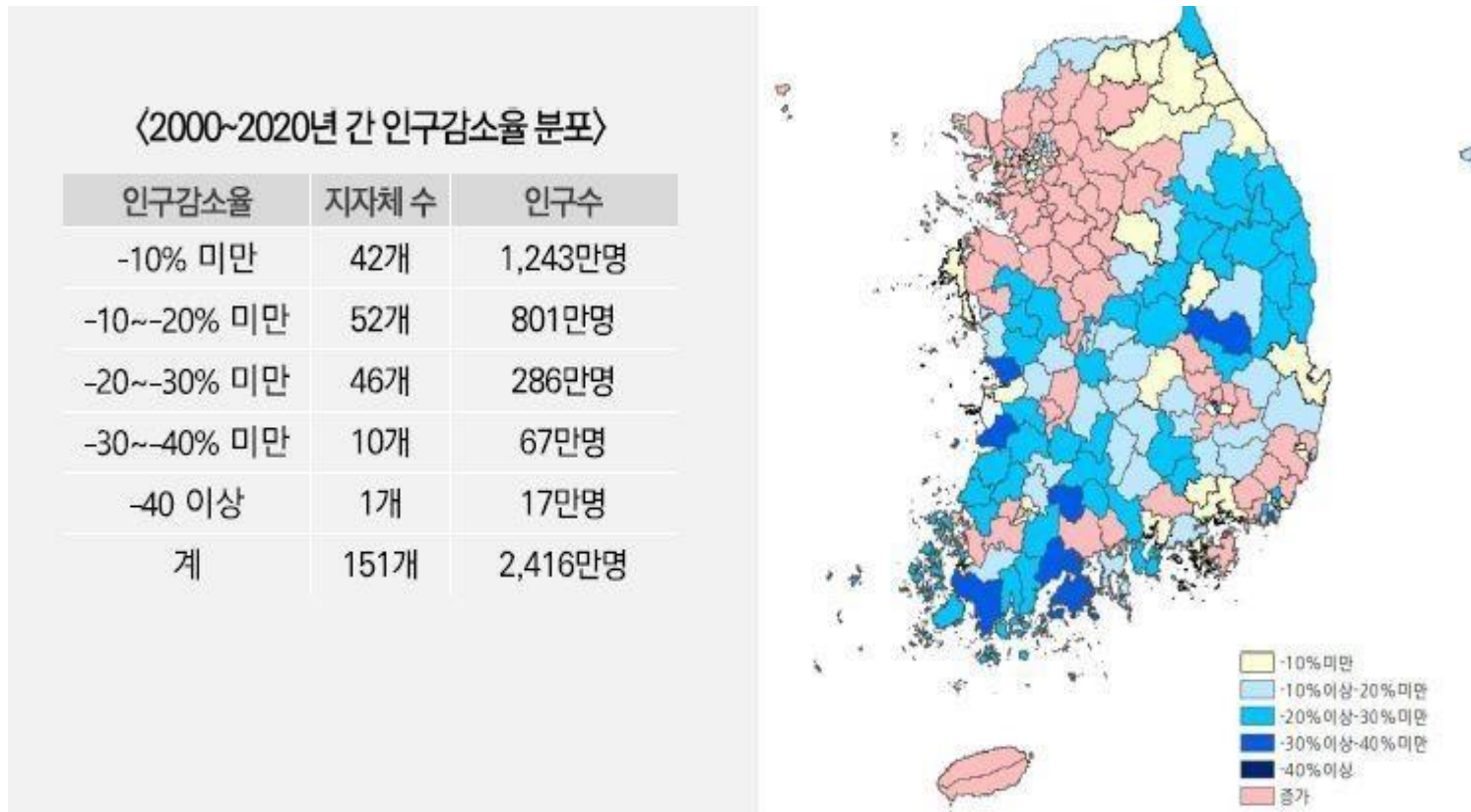
지도시각화

숙명여자대학교 경영학부 오중산

단계구분도(choropleth map) 소개

- 단계구분도의 정의

◆ 지도 상에서 통계치 결과에 따라 음영/색상/패턴별 차이를 다르게 보여주는 지도



20년 동안 인구감소와 관련된 단계구분도(출처: 국토연구원)

단계구분도 데이터: USArrests

- USArrests 데이터

- ◆ 1973년 미국 주(state)별 강력 범죄 용의자 체포 관련 데이터

- `crime <- USArrests`

- ❖ 변수: Murder, Assault, Urbanpop, Rape

- ◆ 변수명이 없는 주(state)에 대해 변수명 state 부여하고, 주 명칭을 소문자로 바꾸기

- `library(tibble) : dplyr에 내장된 패키지`

- `crime <- rownames_to_column(crime, var = "state")`

- `crime$state <- tolower(crime$state)`

USArrests 단계구분도 실습

- state 데이터 활용하기

- ◆ maps 패키지 내장 데이터(지도)이며, 지역별 위도와 경도 정보 등을 담고 있음

- `install.packages("maps") / library(maps)`

- ◆ ggplot2 패키지의 `map_data` 함수를 이용하여 state를 데이터 프레임 형태로 저장

- `library(ggplot2)`

- `states_map <- map_data("state")`

- ❖ `states_map`에는 주별로 소속된 15,000개가 넘는 여러 지역의 위치 정보가 담겨 있음

- ❖ `states_map`은 형태상 데이터 프레임이지만 내용상 지도로 볼 수 있으며, 이것을 플랫폼으로 삼아 다양한 정보(예: 범죄율, 소득, 교육수준 등...)를 반영하여 단계구분도로 표현 가능

USArrests 단계구분도 실습

• 강력범죄 단계구분도 만들기

- ◆ map_id 명령어 사용을 위해 mapproj 패키지 설치 및 로딩
- ◆ ggChoropleth 함수 사용을 위해 [ggiraphExtra](#) 패키지 설치 및 로딩
 - ggChoropleth 함수는 단계구분도를 그려주는 함수
- ◆ ggChoropleth를 이용한 단계 구분도 만들기
 - ggChoropleth(data = crime, aes(fill = Murder, map_id = state), map = states_map)
 - ❖ crime 데이터 프레임의 Murder 측정결과를 states_map에 주(state)를 기준으로 반영하여 단계구분도로 표현하라는 의미
 - ❖ 주의) aes를 활용하려면 ggplot2 패키지 불러와야 함
 - ggChoropleth(data = crime, aes(fill = Assault, map_id = state), map = states_map, interactive = T)
 - ❖ Export에서 'Save As Web Page...' 선택하여 html 형식으로 저장

한국 시도별 인구 단계구분도 실습

- kormaps2014 패키지

- ◆ 대한민국 지역통계 데이터와 지도 데이터를 담고 있음

- korpop1: 2015년 시·도별 인구 데이터
 - korpop2: 2015년 시·군·구별 인구 데이터
 - korpop3: 2015년 읍·면·동별 인구 데이터
 - kormap1: 2014년 시·도별 행정지도
 - kormap2: 2014년 시·군·구별 행정지도
 - kormap3: 2014년 읍·면·동별 행정지도

- ◆ 사용안내 사이트: <https://rpubs.com/cardiomoon/222145>

한국 시도별 인구 단계구분도 실습

- kormaps2014 패키지 설치 및 로딩

- ◆ 문자열처리를 패키지 stringi 설치 및 로딩

- `install.packages("stringi") / library(stringi)`

- ◆ github에 있는 패키지 설치를 위한 패키지 devtools 설치 및 로딩

- `install.packages("devtools") / library(devtools)`

- ◆ `devtools::install_github("cardiomoon/kormaps2014")`

- kormaps2014는 CRAN에 등록되어 있지 않아서 이러한 방식으로 설치해야 함

- ◆ `library(kormaps2014)`

- ◆ 주의사항: kormaps2014가 CRAN에 등록된 공식 패키지가 아니라 설치 및 시행과정에 오류가 빈번히 발생함!

한국 시도별 인구 단계구분도 실습

• 변수명 바꾸기

- ◆ 단계구분도에서 한글변수명 깨지는 오류를 막기 위해 변수명을 영문으로 변경
 - `korpop1 <- rename(korpop1, pop = 총인구_명, name = 행정구역별_읍면동)`

• 변수 측정결과에 대한 한글 깨짐 방지

- ◆ 아래와 같이 인코딩 방식을 변경해야만 지도에서 시도명(한글)이 깨지지 않음
 - `korpop1$name <- iconv(korpop1$name, "UTF-8","CP949")`

■ 인터랙티브 단계구분도 실행

- ◆ `ggChoropleth(data = korpop1, aes(fill = pop, map_id = code, tooltip = name), map = kormap1, interactive = T)`
 - `code`는 광역시도별 고유번호이며, `tooltip`은 지도상 표시할 내용

한국 시도별 결핵환자수 단계구분도 실습

• tbc 데이터

◆ korpmap2014에 내장된 데이터로 2001~2015년 결핵환자수(NewPts) 데이터 포함

- 데이터프레임 생성: `new_tbc <- tbc`

◆ 인터랙티브 단계구분도 실행

- 한글인코딩 방식변경: `new_tbc$name <- iconv(new_tbc$name, "UTF-8","CP949")`
- `ggChoropleth(data = new_tbc, aes(fill = NewPts, map_id = code, tooltip = name), map = kormap1, interactive = T)`
 - ❖ 지도에 2001년 데이터만 출력되므로, 특정 연도를 확인하려면 `filter` 함수를 통해 연도를 추출해야 함
- 2002년도 결핵환자수 인터랙티브 단계구분도 실행
 - ❖ `new_tbc_2002 <- new_tbc %>% filter(year == 2002)`
 - ❖ `ggChoropleth(data = new_tbc_2002, aes(fill = NewPts, map_id = code, tooltip = name), map = kormap1, interactive = T)`
 - ✓ 주의) 앞서 `new_tbc`의 `name` 변수 측정값 인코딩 방식을 CP949로 바꾸었으므로 다시 바꾸면 안됨!