

9강 지리적 데이터의 시각화 (2)

숭실대학교 정보통계보험수리학과
이정진 교수

1. 지리적 데이터의 시각화란?
2. 지도 위에 데이터 표시
3. 지리적 데이터를 색으로 표시
4. 시간에 따른 지리적 데이터 시각화

1. 지리적 데이터의 시각화란?

- 지리적 데이터의 시각화 목적
- 지리적 데이터의 시각화 방법

1 지리적 데이터의 시각화란?

▶ 지리적 데이터의 시각화

- 시각화 : 지역과 관련된 데이터를 지도와 함께 색으로 표시
- 목적
 - 지역별 특징적인 사실 부각
 - 지역 전체 상황 파악
 - 시간에 따른 변화 파악

1 지리적 데이터의 시각화란?

▶ 지리적 데이터 시각화의 방법

■ 방법

- 한 지역의 특성 데이터를 지도 위에 표시
- 지리적 데이터를 색으로 구별하여 지도 위에 표시
- 시간에 따른 지리적 데이터를 같은 색으로 구별하여 여러 개의 지도를 같이 그림

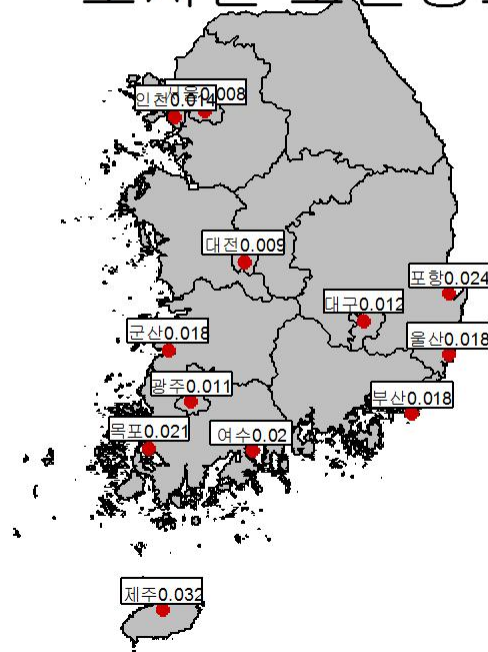
2. 지도 위에 데이터 표시

- 지도 위에 데이터 표시 방법
- 예제 5.5 오존데이터

▶ 지도 위에 데이터 표시

- 지역과 관련된 데이터를 지도 위에 표시
- 목적 : 지역별 특징적인 사실 부각

도시별 오존농도



▶ **[예제 5.5] 오존데이터** (c:/Rwork/city_pollution.txt)

X (경도)	Y (위도)	시도	이산화황	오존	이산화질소	미세먼지 농도
126.98	37.53	서울	0.008	0.008	0.044	64
129.03	35.06	부산	0.006	0.018	0.024	50
128.55	35.82	대구	0.009	0.012	0.028	56
126.68	37.48	인천	0.009	0.014	0.031	71
126.84	35.15	광주	0.007	0.011	0.033	65
127.36	36.30	대전	0.006	0.009	0.025	54
129.39	35.54	울산	0.006	0.018	0.024	40
126.61	35.58	군산	0.007	0.018	0.018	50
127.44	34.75	여수	0.010	0.020	0.016	44
126.42	34.78	목포	0.006	0.021	0.009	53
129.40	36.03	포항	0.009	0.024	0.025	53
126.55	33.46	제주	0.006	0.032	0.016	49

▶ [예제 5.5] 오존데이터 (c:/Rwork/city_pollution.txt)

```
library(sp)
```

```
# 시도별 행정지도 데이터 불러와 지도를 그림, gadm 데이터프레임이  
자동으로 생성됨
```

```
load("C:/Rwork/KOR_adm1.RData")
```

```
plot(gadm, col="grey75")
```

```
# 시도 오존 데이터 불러옴
```

```
pollution <- read.table("C:/Rwork/pollution.txt",header=T)
```


▶ [예제 5.5] 오존데이터 (c:/Rwork/city_pollution.txt)

도시명을 넣은 사각형 크기를 위한 값 설정

```
pollution$breadth<-2/5
```

```
pollution$height<-0.1
```

```
pollution$space<-0.1
```

```
spaceDif<-0.05
```

지도위에 도시의 점 그리기

```
for (i in 1:dim(pollution)[1]) {
```

```
  coords <- SpatialPoints(data.frame(cbind(pollution$x경도[i], pollution$y위도[i])),  
    proj4string = CRS("+proj=longlat"))
```

```
  plot(coords, col ="red3", pch = 20, cex = 1.5, add = TRUE)
```

```
}
```

▶ [예제 5.5] 오존데이터 (c:/Rwork/city_pollution.txt)

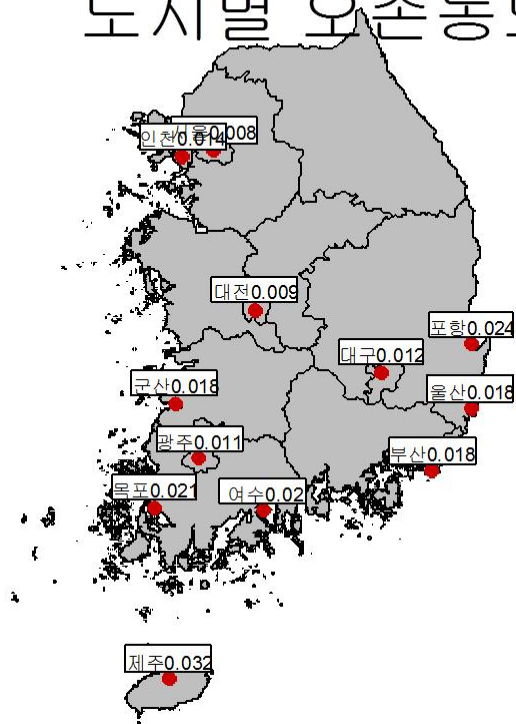
```
# 도시 옆에 이름 넣을 사각형 그리기
for (i in 1:dim(pollution)[1]) {
  a<-c(pollution$x경도[i]-pollution$breadth[i],
      pollution$x경도[i]+pollution$breadth[i],
      pollution$x경도[i]+pollution$breadth[i],
      pollution$x경도[i]-pollution$breadth[i])
  b<-c(pollution$y위도[i]+pollution$space[i]-pollution$height[i]+spaceDif,
      pollution$y위도[i]+pollution$space[i]-pollution$height[i]+spaceDif,
      pollution$y위도[i]+pollution$space[i]+pollution$height[i]+spaceDif,
      pollution$y위도[i]+pollution$space[i]+pollution$height[i]+spaceDif)
  polygon(x=a,y=b, col="white")
}
```

▶ [예제 5.5] 오존데이터 (c:/Rwork/city_pollution.txt)

```
library(stringr)
cityLabels<-str_c(pollution$시도,pollution$오존)
cityCoord<-
matrix(c(t(pollution$x경도),t(pollution$y위도+pollution$space+spaceDif)),
dim(pollution)[1])
text(cityCoord, labels = cityLabels, cex=0.6, bg="white")
text(128,38.6, labels="도시별 오존농도", cex=2)
```

▶ [예제 5.5] 오존데이터 (c:/Rwork/city_pollution.txt)

도시별 오존농도



3. 지리적 데이터를 색으로 표시

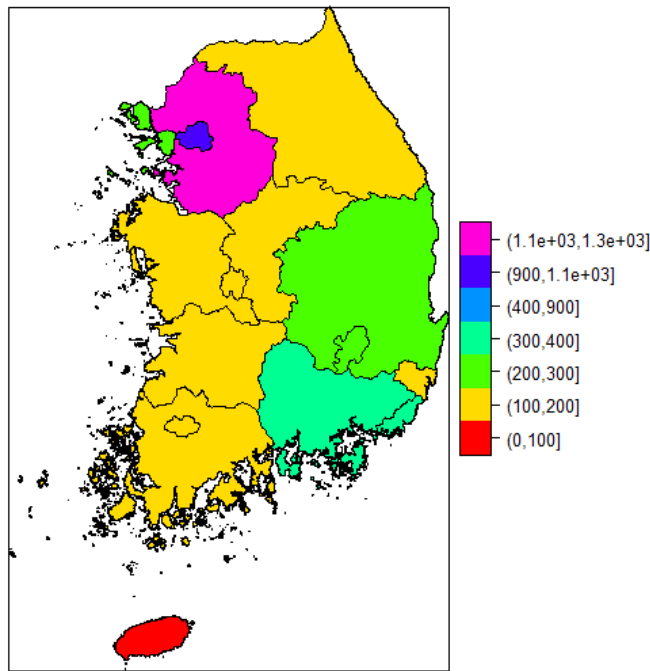
- **지리적 데이터를 색으로 표시하는 방법**
- **예제 5.6 시도별 인구데이터**

3 지리적 데이터를 색으로 표시

▶ 지리적 데이터를 색으로 표시

- 지리적 데이터를 색으로 표시
- 목적 : 지역 전체 상황 파악

2010년 시도별 인구분포



3 지리적 데이터를 색으로 표시

▶ [예제 5.6] 시도별 인구데이터 (c:/Rwork/시도인구.txt)

Code	시도	Y2010	Y2000	Y1990	Y1980
15	서울	963	985	1060	835
1	부산	339	365	379	315
4	대구	243	247	222	160
11	인천	263	246	181	108
7	광주	146	135	113	72
5	대전	149	136	104	65
16	울산	107	101	68	41
8	경기	1119	893	615	493
6	강원	146	148	157	179
2	충북	149	146	138	142
3	충남	200	184	201	295
13	전북	176	188	206	228
14	전남	172	199	250	377
9	경북	257	271	286	495
10	경남	311	297	367	332
12	제주	52	51	51	46

3 지리적 데이터를 색으로 표시

▶ [예제 5.6] 시도별 인구데이터 (c:/Rwork/city_pollution.txt)

```
# sp 패키지의 설치와 라이브러리 불러옴  
install.packages('sp')  
library(sp)
```

```
# 시도별 행정지도 데이터 불러옴, gadm 데이터프레임이 자동으로 생성됨  
load("C:/Rwork/KOR_adm1.RData")
```

```
# 시도인구 데이터 불러옴, 코드별로 정렬  
population <- read.table("C:/Rwork/시도인구.txt",header=T)  
population_sort <- population[order(population$Code),]
```


3 지리적 데이터를 색으로 표시

▶ [예제 5.6] 시도별 인구데이터 (c:/Rwork/city_pollution.txt)

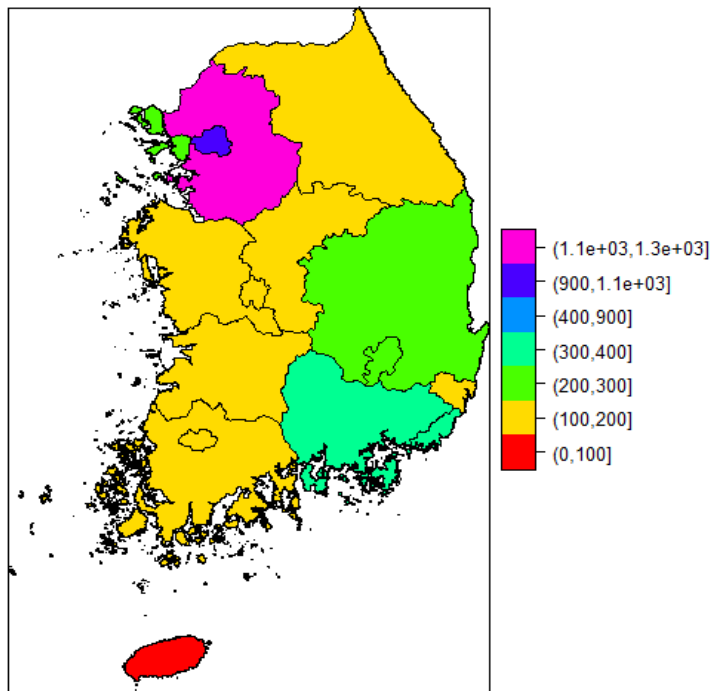
```
# 인구를 구별하기 위한 7개의 구간 설정하고 각 시도 인구를 이 구간 데이터로 변환
# (0,100], (100,200], (200,300], (300,400], (400,900], (900,1100], (1100, 1300]
interval <- c(0,100,200,300,400,900,1100,1300)
population_cut <- cut(population_sort$Y2010,breaks=interval)
gadm$population <- as.factor(population_cut)

# 각 구간의 색을 무지개 색으로 할당하고 지도를 그림
col = rainbow(length(levels(gadm$population)))
spplot(gadm, "population", col.regions=col, main="2010년 시도별 인구분포")
```

3 지리적 데이터를 색으로 표시

▶ [예제 5.6] 시도별 인구데이터 (c:/Rwork/city_pollution.txt)

2010년 시도별 인구분포

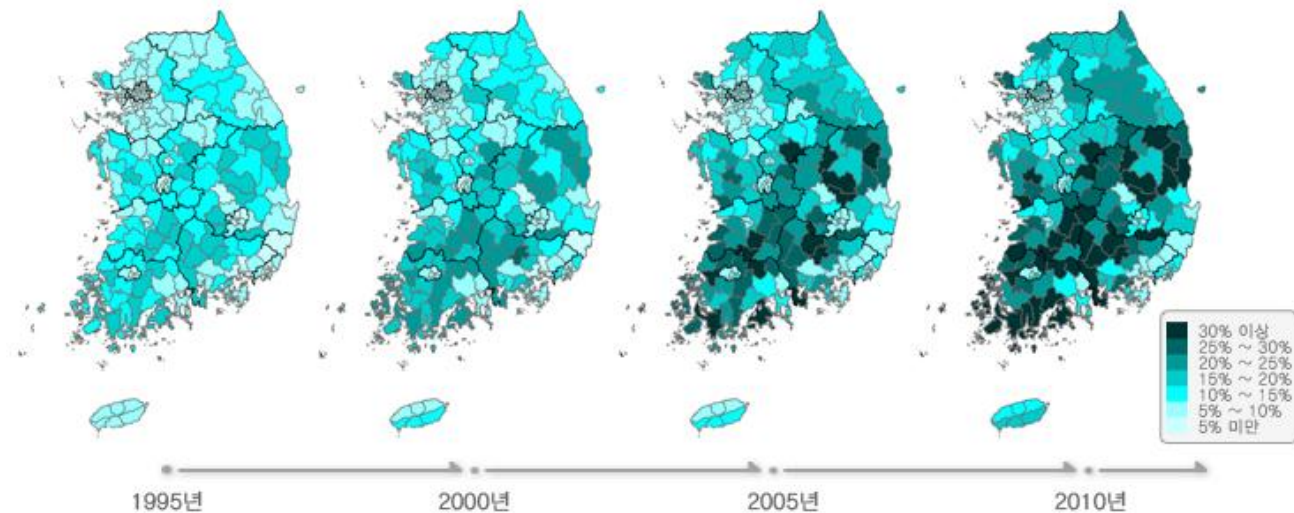


4. 시간에 따른 지리적 데이터의 시각화

- 시간에 따른 지리적 데이터의 시각화
- 예제 5.7 시도별 인구데이터

▶ 시간에 따른 지리적 데이터의 시각화

- 시간에 따른 지리적 데이터를 여러 지도에 같은 색으로 표시
- 시간에 따른 변화 파악





▶ [예제 5.7] 시도별 인구데이터 (c:/Rwork/city_pollution.txt)

2010년 시도별 인구분포의 객체를 만든다.

```
p1 <- spplot(gadm, "population", col.regions=col, main="2010년 시도별  
인구분포")
```

2000년도 시도별 인구분포의 객체를 만든다.

```
population_cut <- cut(population_sort$Y2000,breaks=interval)  
gadm$population <- as.factor(population_cut)  
col = rainbow(length(levels(gadm$population))) # 각 구간의 색 할당  
p2 <- spplot(gadm, "population", col.regions=col, main="2000년 시도별  
인구분포")
```



▶ [예제 5.7] 시도별 인구데이터 (c:/Rwork/city_pollution.txt)

1990년도 시도별 인구분포의 객체를 만든다.

```
population_cut <- cut(population_sort$Y1990,breaks=interval)
```

```
gadm$population <- as.factor(population_cut)
```

```
col = rainbow(length(levels(gadm$population))) # 각 구간의 색 할당
```

```
p3 <- spplot(gadm, "population", col.regions=col, main="1990년 시도별  
인구분포")
```

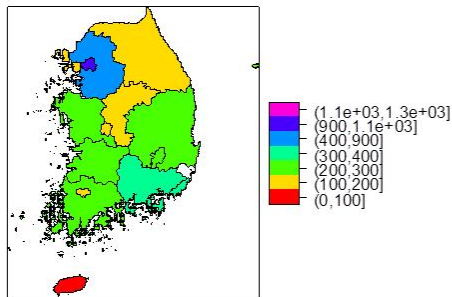


▶ [예제 5.7] 시도별 인구데이터 (c:/Rwork/city_pollution.txt)

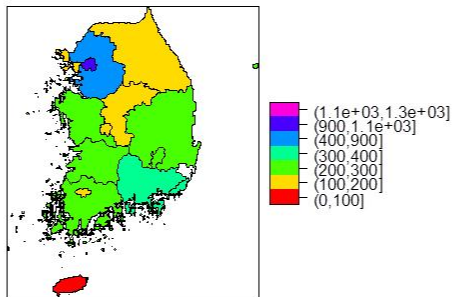
```
# 1980년도 시도별 인구분포의 객체를 만든다.  
population_cut <- cut(population_sort$Y1980,breaks=interval)  
gadm$population <- as.factor(population_cut)  
col = rainbow(length(levels(gadm$population))) # 각 구간의 색 할당  
p4 <- spplot(gadm, "population", col.regions=col, main="1980년 시도별  
인구분포")  
print(p4, pos=c(0,0.5,0.5,1),more=TRUE)  
print(p3, pos=c(0.5,0.5,1,1),more=TRUE)  
print(p2, pos=c(0,0,0.5,0.5),more=TRUE)  
print(p1, pos=c(0.5,0,1,0.5),more=TRUE)
```

▶ [예제 5.7] 시도별 인구데이터 (c:/Rwork/city_pollution.txt)

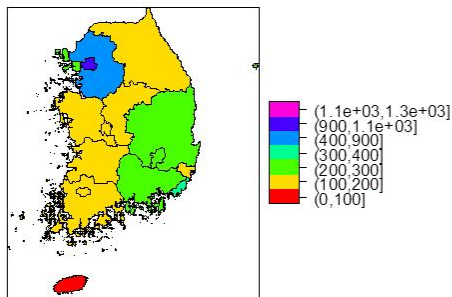
1980년 시도별 인구분포



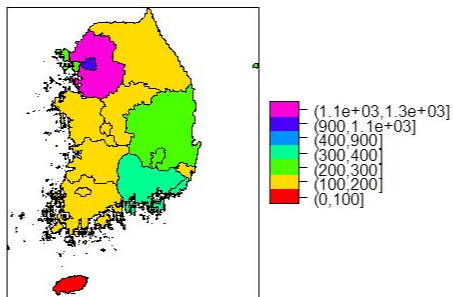
1990년 시도별 인구분포



2000년 시도별 인구분포



2010년 시도별 인구분포





■ 지리적 데이터의 시각화 : 지리적 데이터를 지도와 함께 색으로 표시

■ 목적

- 지역 전체 상황 또는 지역 사이의 연관성 파악
- 지역별 특징적인 사실 부각
- 시간에 따른 변화 흐름 파악

■ 방법

- 한 지역의 특성 데이터를 지도 위에 표시
- 지리적 데이터를 색으로 구별하여 지도 위에 표시
- 시간에 따른 지리적 데이터를 같은 색으로 구별하여 여러 개 그림



다음시간안내

텍스트 데이터의 시각화 (1)