

한남대학교

아파트와 빌라의 도심활성화 기여도

데이터 시각화 프로젝트 과제



수업: 데이터 시각화

교수님: 박영호 교수님

학번: 20182890

이름: 임성민

[목차]

1. 발표 주제 및 선정 이유 -----	2
2. 데이터 내용 및 출처 -----	2
3. 데이터 변수 속성 및 전처리 과정 -----	3
4. 자료 시각화 도구와 시각화 -----	5
5. 시각화 결과 해석 -----	6
6. 두 가지 결론 -----	8

1. 발표 주제 및 선정 이유




① 발표 주제

: 아파트와 빌라의 도심 활성화 기여도

② 주제 선정 이유

: 세종시 행정 거점 이주 계획 발표 이후, 집값에 대한 사람들의 관심이 뜨거워지고 있다. 보통 집값이 높은 지역은 아파트로 가득하고, 이주해 오는 사람들에게 의해 주변 상권도 같이 발전한다. 그런데 초기 유성구와 도안 신도시 혹은 내포신도시 등 집값이 높거나 아파트가 잘 팔린다고 하여 주변 지역 활성화까지는 힘든 경우를 많이 볼 수 있다. 대전의 구도심 활성화에 관심이 많아지면서 집값과 아파트의 유무, 지역민의 소비정도, 지역 활성화 정도에 어떠한 연계점을 찾고자 위와 같은 주제를 선정하였다.

2. 데이터 내용 및 출처

데이터 이름	데이터 출처	데이터 내용
아파트 실거래, 연립 다세대 실거래	공공데이터 포털 대전광역시_주택 https://data.go.kr/tcs/dss/selectFileDataDetailView.do?publicDataPk=15067893	
대전광역시 자치구별 신용카드(KB 국민카드) 매출액	공공데이터 포털 대전광역시_자치구별 신용카드(KB국민카드) 매출액 https://data.go.kr/tcs/dss/selectFileDataDetailView.do?publicDataPk=15064213	
대전광역시 도시재생활성화지역	공공데이터 포털 대전광역시_도시재생활성화지역 https://data.go.kr/tcs/dss/selectFileDataDetailView.do?publicDataPk=15068454	
시도, 읍면동 대한민국 최신 행정구역(SHP)	시도, 읍면동 데이터 2020년 5월 http://www.gisdeveloper.co.kr/?p=2332	

3. 데이터 변수 속성 및 전처리 과정

① 변수의 속성

■ 아파트 실거래

열: 기준년월, 기준일, 법정동, 지번, 아파트명, 층, 전용면적, 거래금액.일부보기

행: 64037개

```
> apt %>% head()
  기준년월 기준일 법정동 지번 아파트명 층 전용면적 거래금액.일부보기 x
1 201501 1~10      인동 179-1  현대 2 70.96      12600 NA
2 201501 1~10      인동 179-1  현대 9 70.96      15000 NA
3 201501 1~10      인동 179-1  현대 2 70.96      12600 NA
4 201501 1~10      인동 179-1  현대 9 70.96      15000 NA
5 201501 1~10      인동 179-1  현대 2 70.96      12600 NA
6 201501 1~10      인동 179-1  현대 9 70.96      15000 NA

> apt %>% summary()
  기준년월      기준일      법정동      지번      아파트명
Min.   :201401   Length:64037   Length:64037   Length:64037   Length:64037
1st Qu.:201410   Class :character   Class :character   Class :character   Class :character
Median :201507   Mode  :character   Mode  :character   Mode  :character   Mode  :character
Mean   :201576
3rd Qu.:201801
Max.   :202009

  층      전용면적      거래금액.일부보기      x
Min.   : 1.000   Min.   : 12.01   Min.   : 2000   Mode:logical
1st Qu.: 4.000   1st Qu.: 59.65   1st Qu.: 13000   NA's:64037
Median : 8.000   Median : 84.36   Median : 19500
Mean   : 8.817   Mean   : 76.78   Mean   : 21046
3rd Qu.:12.000   3rd Qu.: 84.97   3rd Qu.: 26500
Max.   :50.000   Max.   :285.00   Max.   :250000
NA's   : 4       NA's   :4
```

■ 연립 다세대 실거래

열: 기준년월, 기준일, 법정동, 지번, 연립주택명, 층, 건축년도, 전용면적, 거래금액.일부보기

행: 8159개

```
> v111 %>% head()
  기준년월 기준일 법정동 지번 연립주택명 층 건축년도 전용면적 거래금액.일부보기 x
1 201501 21~31      NA      NA      NA      NA      NA      NA      NA
2 201501 21~31      NA      NA      NA      NA      NA      NA      NA
3 201501 1~10      NA      NA      NA      NA      NA      NA      NA
4 201501 11~20      NA      NA      NA      NA      NA      NA      NA
5 201501 1~10      중문동 462-3   삼보빌라 3   1996   74.29      7800 NA
6 201501 21~31      중문동 274-1   복음타운 2   1995   84.96      12000 NA

> v111 %>% summary()
  기준년월      기준일      법정동      지번      연립주택명      층      건축년도
Min.   :201401   Length:8159   Length:8159   Length:8159   Length:8159   Min.   :~-1.000   Min.   :1971
1st Qu.:201410   Class :character   Class :character   Class :character   Class :character   1st Qu.: 2.000   1st Qu.:1990
Median :201507   Mode  :character   Mode  :character   Mode  :character   Mode  :character   Median : 2.000   Median :1993
Mean   :201581
3rd Qu.:201802
Max.   :202009

  전용면적      거래금액.일부보기      x
Min.   : 12.99   Min.   : 500   Mode:logical
1st Qu.: 45.29   1st Qu.: 5100   NA's:8159
Median : 52.50   Median : 6700
Mean   : 54.95   Mean   : 8412
3rd Qu.: 64.69   3rd Qu.: 9400
Max.   :228.54   Max.   :100000
NA's   : 4       NA's   :4

> dim(v111)
[1] 8159 10
```

■ 대전광역시 자치구별 신용카드(KB국민카드) 매출액

열: 기준년도, 구, 업종, 3월.억원, 4월.억원, 5월.억원, 6월.억원, 7월.억원

행: 307개

```
> kboard %>% head()
  기준년도 구      업종      3월.억원      4월.억원      5월.억원      6월.억원      7월.억원
1 2019년 중구      건축관련업      5.6      5.6      5.6      5.6      6.4
2 2019년 중구      관광여행      0.1      0.1      0.1      0.1      0.1
3 2019년 중구      기타유통      0.3      0.4      0.5      0.2      0.2
4 2019년 중구      기타제조/도매      1.0      0.4      0.8      0.9      0.7
5 2019년 중구      레저/스포츠      15.7      15.8      17.0      13.6      11.5
6 2019년 중구      문화/취미      24.5      11.3      13.4      11.9      9.9

> kboard %>% summary()
  기준년도 구      업종      3월.억원      4월.억원      5월.억원      6월.억원      7월.억원
Length:307   Length:307   Length:307   Min.   : 0.00   Min.   : 0.00   Min.   : 0.00   Min.   : 0.10   Min.   : 0.00
Class :character   Class :character   Class :character   1st Qu.: 1.95   1st Qu.: 2.10   1st Qu.: 2.30   1st Qu.: 2.15   1st Qu.: 1.90
Median : 9.35   Median : 9.45   Median : 10.85   Median : 9.40   Median : 9.00
Mean   : 43.89   Mean   : 45.28   Mean   : 50.34   Mean   : 45.03   Mean   : 43.74
3rd Qu.: 32.85   3rd Qu.: 32.00   3rd Qu.: 39.15   3rd Qu.: 33.25   3rd Qu.: 30.70
Max.   :854.80   Max.   :798.00   Max.   :866.80   Max.   :823.60   Max.   :825.20
NA's   :25       NA's   :21       NA's   :15       NA's   :20       NA's   :24

> dim(kboard)
[1] 307 8
```



```

# shp파일을 sf
sido_shp <- st_read("Desktop/R/min/CTPRVN_202005_/CTPRVN.shp")
# 인코딩
sido_shp$SIG_KOR_NM <- iconv(sido_shp$CTP_KOR_NM, localeToCharset(), 'UTF-8')

dj_shp <- subset(sido_shp, CTP_KOR_NM == '대전광역시')

dj_shp %>%
  plot()
### 대전 파일과 읍면동 파일을 붙이자.
st_crs(sido_shp)
emd_shp <- st_read("Desktop/R/min/EMD_202005_읍면동/EMD.shp")
## 인코딩 변경
emd_shp$EMD_KOR_NM <- iconv(emd_shp$EMD_KOR_NM, localeToCharset(), 'UTF-8')
unique(emd_shp$EMD_KOR_NM)
## 대전시 행정동 코드: 3000000000
Dj_emd_shp <- emd_shp %>%
  mutate(SIGUNGU = case_when(str_detect(EMD_CD, "A30140") ~ "중구",
                               str_detect(EMD_CD, "A30110") ~ "동구",
                               str_detect(EMD_CD, "A30230") ~ "대덕구",
                               str_detect(EMD_CD, "A30200") ~ "유성구",
                               str_detect(EMD_CD, "A30170") ~ "서구")) %>%

  filter(!is.na(SIGUNGU))
st_crs(dj_shp)
st_crs(emd_shp)
class(Dj_emd_shp)
class(emd_shp)

Dj_emd_shp <- Dj_emd_shp %>%
  mutate(SIGUNGU = case_when(str_detect(EMD_CD, "A30140") ~ "중구",
                               str_detect(EMD_CD, "A30110") ~ "동구",
                               str_detect(EMD_CD, "A30230") ~ "대덕구",
                               str_detect(EMD_CD, "A30200") ~ "유성구",
                               str_detect(EMD_CD, "A30170") ~ "서구")) %>%

  filter(!is.na(SIGUNGU))
## sf 데이터프레임 --> shapefile 변환
Dj_emd_shp_2 <- as(Dj_emd_shp, 'Spatial')
class(Dj_emd_shp_2)

## shapefile --> 데이터프레임 변환
Dj_emd_shp_1 <- fortify(Dj_emd_shp_2)

Dj_emd_shp_2$EMD_KOR_NM
## ggplot 시각화가 가능 한지 확인
Dj_emd_shp_1 %>%
  ggplot(aes(x = long, y = lat, group = group)) +
  geom_polygon(color = "white", fill="skyblue") +
  coord_fixed(1.3) +
  guides(fill=FALSE) +
  theme_void()
##### 법정동으로 붙여보기
Dj_emd_shp_2@data
### 지도 확인
plot(Dj_emd_shp_2)
shp.df <- fortify(Dj_emd_shp_2)
### join
Dj_emd_shp_2@data$id <- rownames(Dj_emd_shp_2@data)
shp.df.use <- left_join(shp.df, Dj_emd_shp_2@data, by="id")
shp.df.use <- shp.df.use[order(shp.df.use$id, shp.df.use$order),]
shp.df.use %>% head(30)

tail(apt1920)
apt20 <- rename(apt1920, "EMD_KOR_NM" = "법정동")
tail(apt20)
### Join apt20 자료 추가
shp.apt <- inner_join(shp.df.use, apt20, by="EMD_KOR_NM")

kb19_1 <- rename(kb19, "SIGUNGU" = "구")

### Join kb19 자료 추가
shp.kb19 <- left_join(shp.df.use, kb19_1, by="SIGUNGU")
shp.kb20 <- left_join(shp.df.use, kb20_1, by="SIGUNGU")

### library
library(ggmap)
library(dplyr)

# 영어 행을 하나 만들자
dfek <- shp.df.use[,c(9,10,11)]
# 빈도 행도 만들자
EMD_KOR_NM <- table(apt20$SIGUNGU)
EMD_KOR_NM.df <- data.frame(EMD_KOR_NM)

### library
library(sf)
### shape file
shp.file <- file.choose()
shp.re <- sf::st_read(shp.file, quiet=TRUE, options="ENCODING=CP949")

apt123 <- rename(apt, "EMD_KOR_NM" = "법정동")
EMD_KOR_NM <- table(apt123$시군구)
EMD_KOR_NM.df <- data.frame(EMD_KOR_NM)

### Join ??? 자료 추가
shp.use_eng <- left_join(shp.re, dfek, by="EMD_ENG_NM")
shp.use_eng_1 <- shp.use_eng[,c(3,4)]
shp.use_eng_1 <- rename(shp.use_eng_1, "EMD_KOR_NM" = "EMD_KOR_NM.y")
shp.use_eng_1 <- rename(shp.use_eng_1, "SIGUNGU" = "SIGUNGU.y")
apt20 <- rename(apt, "SIGUNGU" = "시군구")
EMD_KOR_NM.df <- rename(EMD_KOR_NM.df, "SIGUNGU" = "EMD_KOR_NM")
EMD_KOR_NM.df <- rename(EMD_KOR_NM.df, "빈도" = "Freq")
shp.use_apt <- left_join(shp.use_eng_1, EMD_KOR_NM.df, by="SIGUNGU")
shp.use_apt[is.na(shp.use_apt)] <- 0
shp.use_apt <- unique(shp.use_apt)

```

시도 shp 파일에서 대전광역시 데이터만 추출

추출한 대전시 파일에 대전 읍면동 shp파일을 병합 후 데이터 인코딩을 utf-8 형태로 수정

대전광역시 행정동 코드(300000단위)에 부합되는 대전광역시 '구'정보를 삽입

Sf 파일로 만들어진 데이터를 shp 파일로 변환하고 데이터프레임 형태로 변형

법정동 데이터로 시각화가 가능한지 ggplot을 이용해 확인

대전시 시군 파일에 구읍면동 좌표 데이터를 join

아파트2019년, 2020년 데이터와 kb국민카드 데이터에 만들어진 좌표 데이터 병합

각 데이터의 법정동명 이름을 '법정동'으로 맞춰줌

준비한 자료에 대전광역시 '시군동' 데이터를 동 기준으로 모두 join하고 모든 행의 내용이 전부 중복인 데이터를 unique로 삭제

처음에 준비한 동 데이터 자료에 각 '동'에 맞는 좌표값만 대조하여 남긴다.

4. 자료 시각화 도구와 시각화

① Plotly

: 오픈소스 자바스크립트 라이브러리로 R내에서 ggplot2와 연계가능한 시각화 패키지이다. 대화형 시각화 자료를 만들 수 있고 지도 데이터를 다루기 수월하여 본 프로젝트에서 지도 시각화의 용도로 사용하였다.

② 시각화 핵심 코드

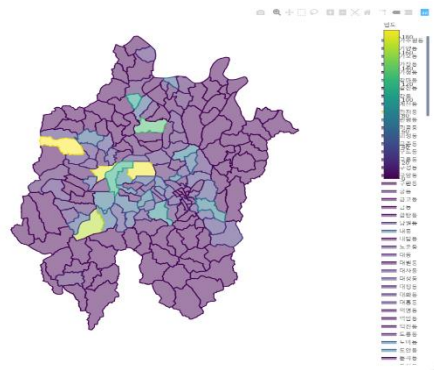
```
EMD_KOR_NM <- table(apt2019$EMD_KOR_NM)
EMD_KOR_NM.df <- data.frame(EMD_KOR_NM)
EMD_KOR_NM.df <- rename(EMD_KOR_NM.df, "EMD_KOR_NM" = "Var1")
EMD_KOR_NM.df <- rename(EMD_KOR_NM.df, "빈도" = "Freq")

shp.use_apt <- left_join(shp.use_eng_1, EMD_KOR_NM.df, by="EMD_KOR_NM")
shp.use_apt[is.na(shp.use_apt)] <- 0

table(shp.use_apt$EMD_KOR_NM)
library(plotly)
### 다각형 색상 - 동네
# 색이 겹쳐져 나온것은 동네 이름이 겹치는 것

fig <- plot_ly(shp.use_apt, split=~EMD_KOR_NM, color=~빈도)
```

[결과물]



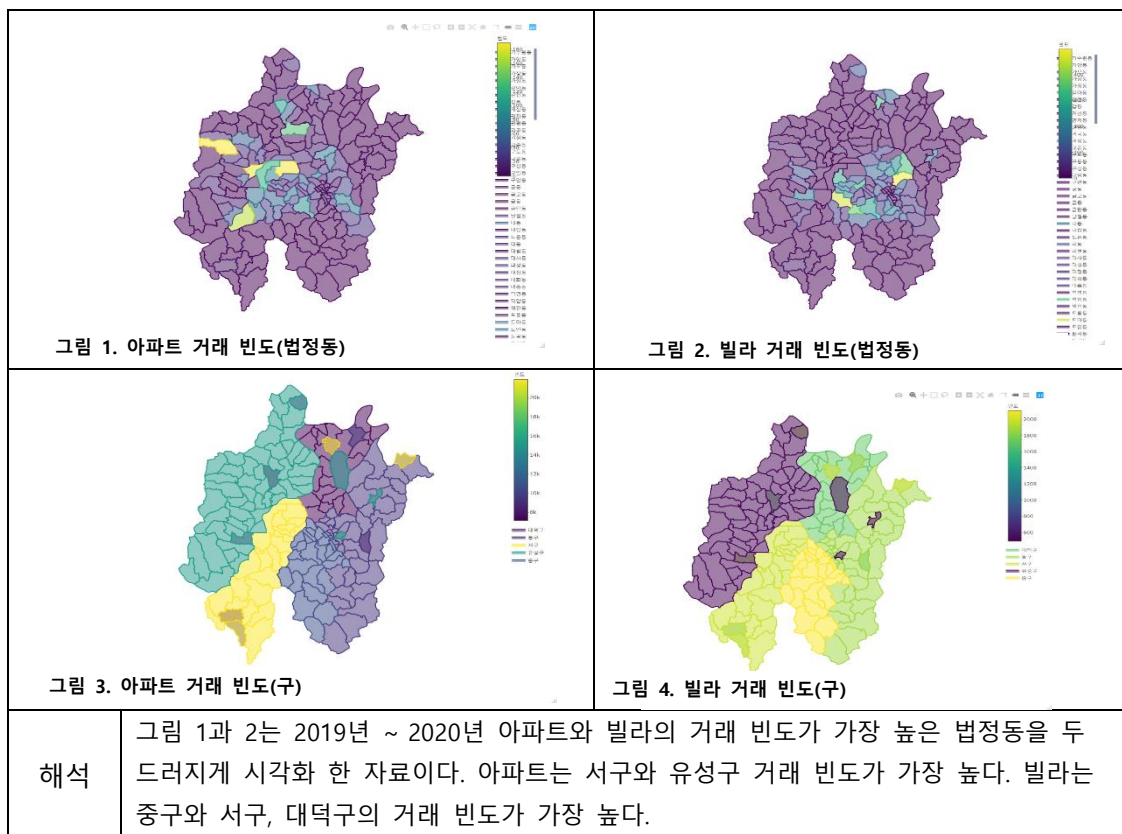
깔끔하게 정제한 데이터를 활용해 변수이름 열(동이름)과 빈도열을 생성한다.

동이름열을 기준으로 만들어 둔 시구동 데이터에 join한다.

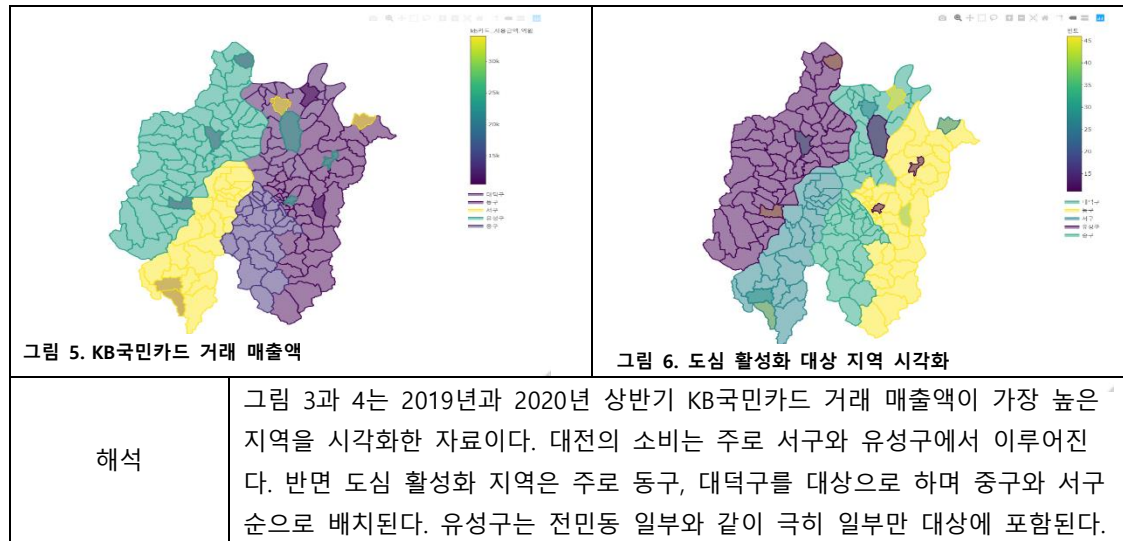
거래량이 없는 행의 NA값을 0으로 바꿔주고 동이름으로 나누고 빈도로 색깔을 입힌 plotly 형식의 지도를 생성한다.

5. 시각화 결과 해석

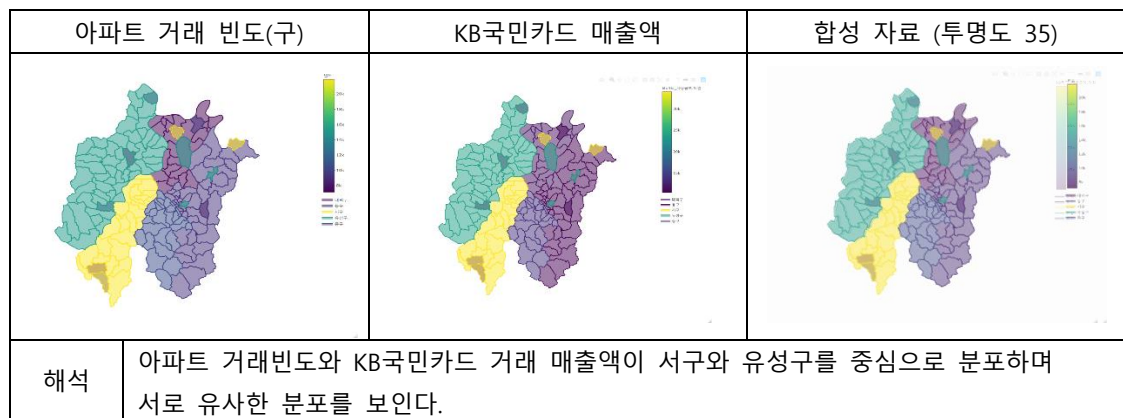
① 아파트와 빌라가 가장 많이 거래되는 지역 시각화



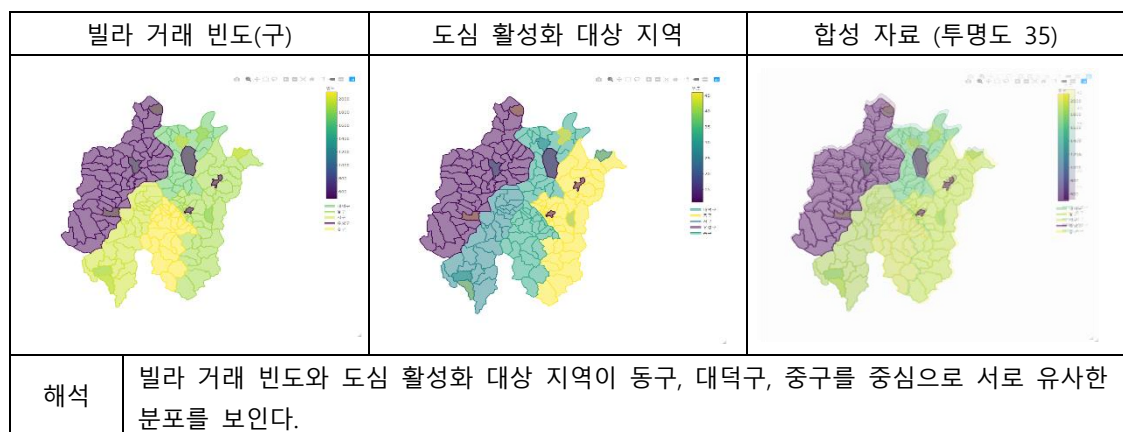
② KB국민카드 거래 매출액 & 도심 활성화 지역



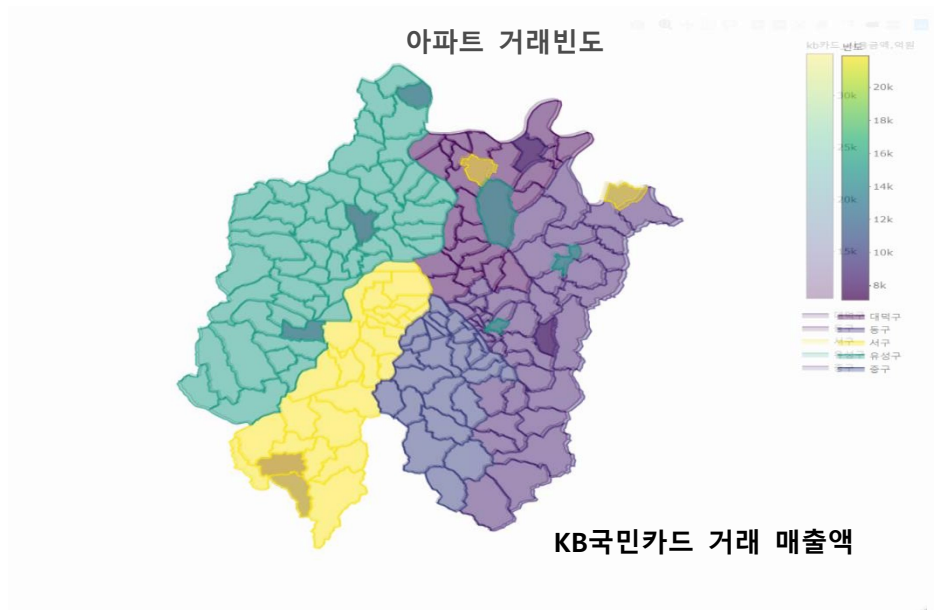
③ 아파트 거래 빈도 지도와 국민카드 거래 매출액 지도 합성



④ 빌라 거래 빈도 지도와 도심활성화 사업 대상 지도를 겹친 자료

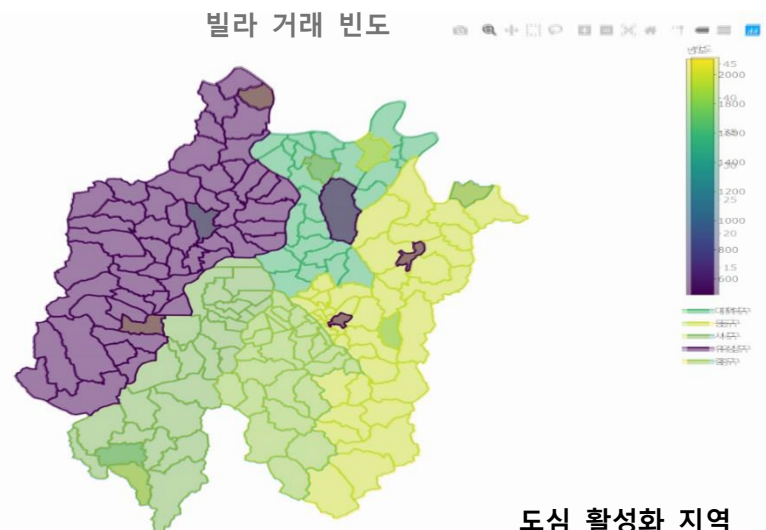


6. 두 가지 결론



▷ 부동산 거래방향과 소비 액수는 서로 밀접한 관련이 있다.

: 시각화 자료에서 유성구와 서구 신도심 부근은 부동산 투기와 소비 활동이 모두 활발하게 일어나고 있음을 확인할 수 있었다. 반면 빌라 거래가 활발한 지역은 주로 구도심 지역이며 사람들의 평균 소비 액수가 낮다. 또한 '구도심 활성화 사업의 성과가 부진하다'는 것과 '소비 액수 분포를 통해 현재 도심 활성화 사업의 성과가 미미하다'는 것을 알 수 있다.



도심 활성화 지역

▷ 부동산 분포와 도심 활성화 정도는 밀접한 관련이 있다.

: 시각화 자료를 보면 아파트 거래의 유무가 도심 활성화에 얼마나 영향을 미치는지 알 수 있다. 구도심과 일부 퇴화지역에 아파트를 늘리는 사업이 장기적인 소비와 도심 활성화에 많은 도움이 된다는 사실도 예측할 수 있었다.