

# 빅데이터처리 프로젝트

## (프랜차이즈 편의점 입지 시각화)

---

컴퓨터정보과 3-A 201944004신건호



# 목차

- 프로젝트 내용
- 데이터 수집
- 데이터 변환
- 데이터 시각화
- 결론

# 프로젝트 내용

- 서울시 프랜차이즈 편의점 분포와 각 프랜차이즈 별 점포 개수 비교 및 분석
- 프랜차이즈 편의점은  
GS25, CU, 세븐일레븐, 이마트24, 미니스톱 편의점으로 한정
- 공공데이터 포털에서 상가(상권) 정보를 수집 후  
편의점으로 분류하여 편의점 데이터 수집
- Folium을 통해 서울시 지도 위에 점포들을 시각화

# 데이터 수집

## 데이터 상세

### 소상공인시장진흥공단\_상가(상권)정보

영업 중인 전국 상가업소 데이터를 제공합니다.

(상호명, 업종코드, 업종명, 지번주소, 도로명주소, 경도, 위도 등)

공공데이터 포털

(<https://www.data.go.kr/>) 에서

소상공인시장진흥공단\_상가(상권)정보

(<https://www.data.go.kr/data/15083033/fileData.do>) csv 파일 수집

 소상공인시장진흥공단\_상가(상권)정보\_서울\_202309.csv

### #공공데이터 포털에서 수집한 서울시에 있는 상가(상권) 정보 변환

```
df = pd.read_csv('/content/drive/MyDrive/소상공인시장진흥공단_상가(상권)정보_서울_202309.csv', engine='python', encoding='utf-8')
```

### #필요한 열만 지정하여 새로 dataset 만들기

```
dataset = df[['상호명', '지점명',  
             '상권업종소분류명',  
             '시도명', '시군구명', '행정동명',  
             '위도', '경도']]
```

### #상권 업종 소분류명이 '편의점' 이면서 시도명이 '서울특별시'인 경우만 추출

```
df_store = dataset[(dataset['상권업종소분류명']=='편의점')&(dataset['시도명']=='서울특별시')]
```

### #상가(상권) 정보에서 '상권 업종 소분류명' 열에 '편의점'으로 필터링하여 편의점으로 분류되어 있는 상권 수 확인

```
df_seoul_store = df_store[df_store['상권업종소분류명'].str.contains('편의점')]  
df_seoul_store.index = range(len(df_seoul_store))  
print('서울시 내 편의점 점포 수 :', len(df_seoul_store))
```

### #서울시에 구 별로 편의점의 점포 수 구분 후 출력

```
store_gu = df_seoul_store.groupby('시군구명')['상호명'].count().to_frame().sort_values(by='상호명', ascending=False)  
store_gu = store_gu.reset_index()  
store_gu = store_gu.set_index('시군구명')  
store_gu
```

#### 상호명

#### 시군구명

강남구	896
송파구	625
강서구	585
마포구	542
서초구	523
영등포구	513
관악구	464
중구	406
광진구	400
은평구	391
동대문구	382
강동구	377
구로구	362
성북구	351
서대문구	348
동작구	333
노원구	323
중랑구	320
금천구	319
종로구	313
성동구	297
양천구	294
용산구	289
강북구	281
도봉구	245

# 데이터 변환 (특정 프랜차이즈 편의점만 추출)

```
# 서울특별시와 편의점인 경우만 추출된 데이터 프레임에서 상호명에 'GS25' 또는 '지에스'가 포함되어 있는 경우만 추출
df_seoul_gs25 = df_store[df_store['상권업종소분류명'].str.contains('편의점')&df_store['상호명'].str.contains('GS25|지에스')]
df_seoul_gs25.index = range(len(df_seoul_gs25))
print('서울시 내 GS25 편의점 점포 수 :', len(df_seoul_gs25))
```

#서울시에 구 별로 편의점의 점포 수 구분 후 출력

```
gs25_gu = df_seoul_gs25.groupby('시군구명')['상호명'].count().to_frame().sort_values(by='상호명', ascending=False)
gs25_gu = gs25_gu.reset_index()
gs25_gu = gs25_gu.set_index('시군구명')
gs25_gu
```

# 서울특별시와 편의점인 경우만 추출된 데이터 프레임에서 상호명에 'CU' 또는 '씨유'가 포함되어 있는 경우만 추출

```
df_seoul_cu = df_store[df_store['상권업종소분류명'].str.contains('편의점')&df_store['상호명'].str.contains('CU|씨유')]
df_seoul_cu.index = range(len(df_seoul_cu))
print('서울시 내 CU 편의점 점포 수 :', len(df_seoul_cu))
```

#서울시에 구 별로 편의점의 점포 수 구분 후 출력

```
cu_gu = df_seoul_cu.groupby('시군구명')['상호명'].count().to_frame().sort_values(by='상호명', ascending=False)
cu_gu = cu_gu.reset_index()
cu_gu = cu_gu.set_index('시군구명')
cu_gu
```

서울시 내 GS25 편의점 점포 수 : 3002    서울시 내 CU 편의점 점포 수 : 2388    서울시 내 세븐일레븐 편의점 점포 수 : 2806    서울시 내 이마트24 편의점 점포 수 : 743    서울시 내 미니스톱 편의점 점포 수 : 200

상호명 	
시군구명 	
강남구	262
강서구	178
마포구	173
영등포구	167
송파구	162
관악구	152
서초구	140
은평구	134
동작구	122
구로구	118
중랑구	112
광진구	111
성북구	108
강동구	104
동대문구	103
노원구	101
서대문구	99
중구	98
양천구	88
용산구	88
금천구	84
종로구	77
강북구	77
성동구	73
도봉구	71

상호명 	
시군구명 	
강남구	173
강서구	136
송파구	133
관악구	119
서초구	114
마포구	113
영등포구	106
광진구	100
동대문구	99
은평구	97
금천구	94
강동구	91
서대문구	87
중구	86
종로구	86
중랑구	82
성북구	81
노원구	80
구로구	79
동작구	79
성동구	79
강북구	77
양천구	74
용산구	64
도봉구	59

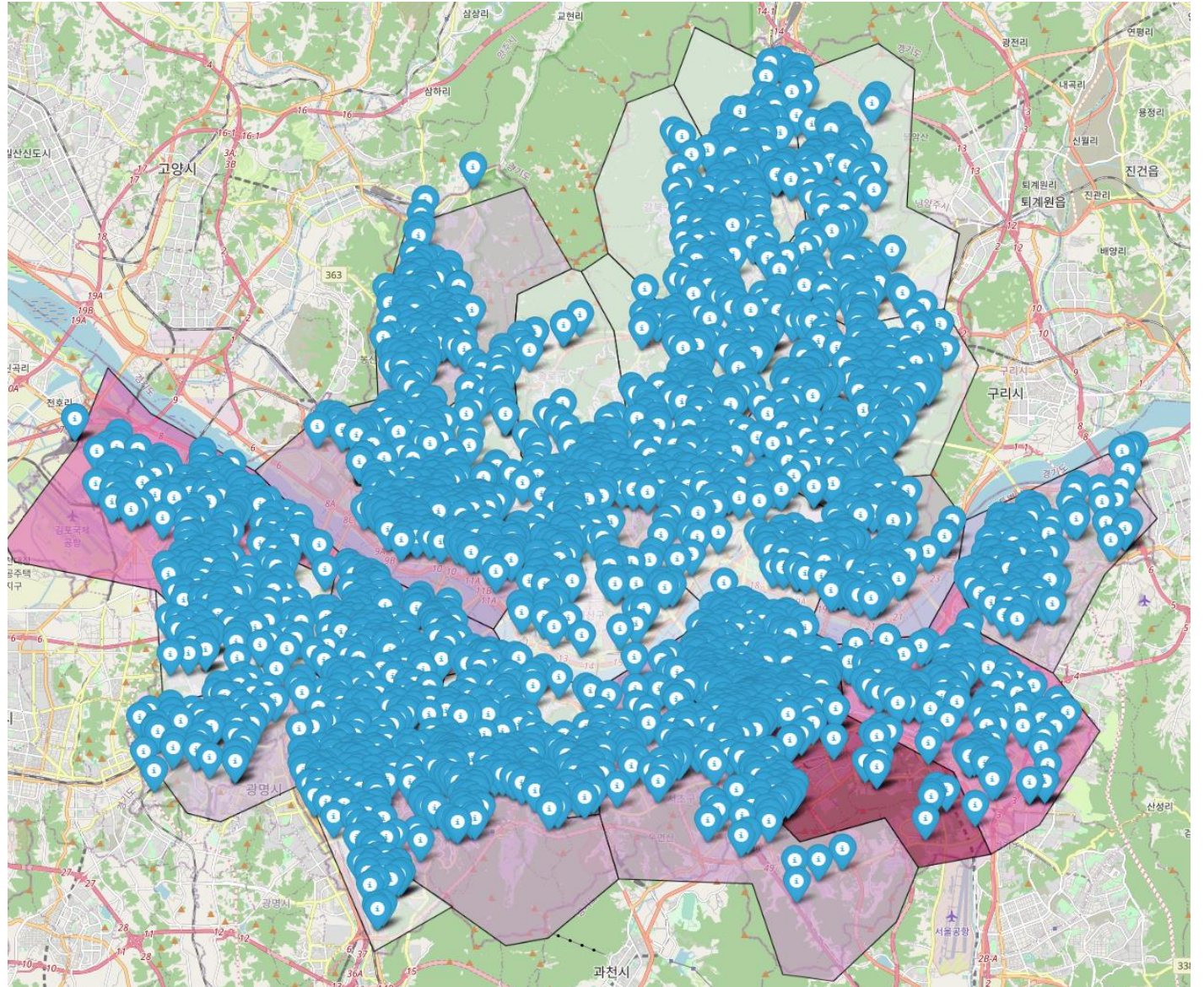
상호명 	
시군구명 	
강남구	288
송파구	213
서초구	167
마포구	162
강서구	157
영등포구	139
관악구	135
중구	130
강동구	116
동대문구	107
광진구	107
종로구	93
성북구	89
구로구	88
용산구	79
노원구	79
서대문구	78
강북구	78
금천구	77
양천구	77
은평구	76
동작구	75
성동구	68
중랑구	66
도봉구	62

상호명 	
시군구명 	
강남구	53
강서구	44
송파구	44
중구	43
영등포구	43
서초구	43
은평구	37
동대문구	37
성동구	34
광진구	33
성북구	29
서대문구	28
마포구	28
구로구	26
중랑구	26
관악구	25
금천구	24
용산구	22
양천구	21
노원구	20
종로구	20
강동구	18
동작구	16
강북구	15
도봉구	14

상호명 	
시군구명 	
강남구	27
서초구	14
강북구	12
구로구	12
강서구	11
금천구	11
성북구	10
송파구	9
마포구	9
성동구	7
영등포구	7
도봉구	6
동작구	6
강동구	6
노원구	6
광진구	6
은평구	6
동대문구	5
관악구	5
양천구	5
종로구	5
중랑구	5
용산구	4
중구	4
서대문구	2

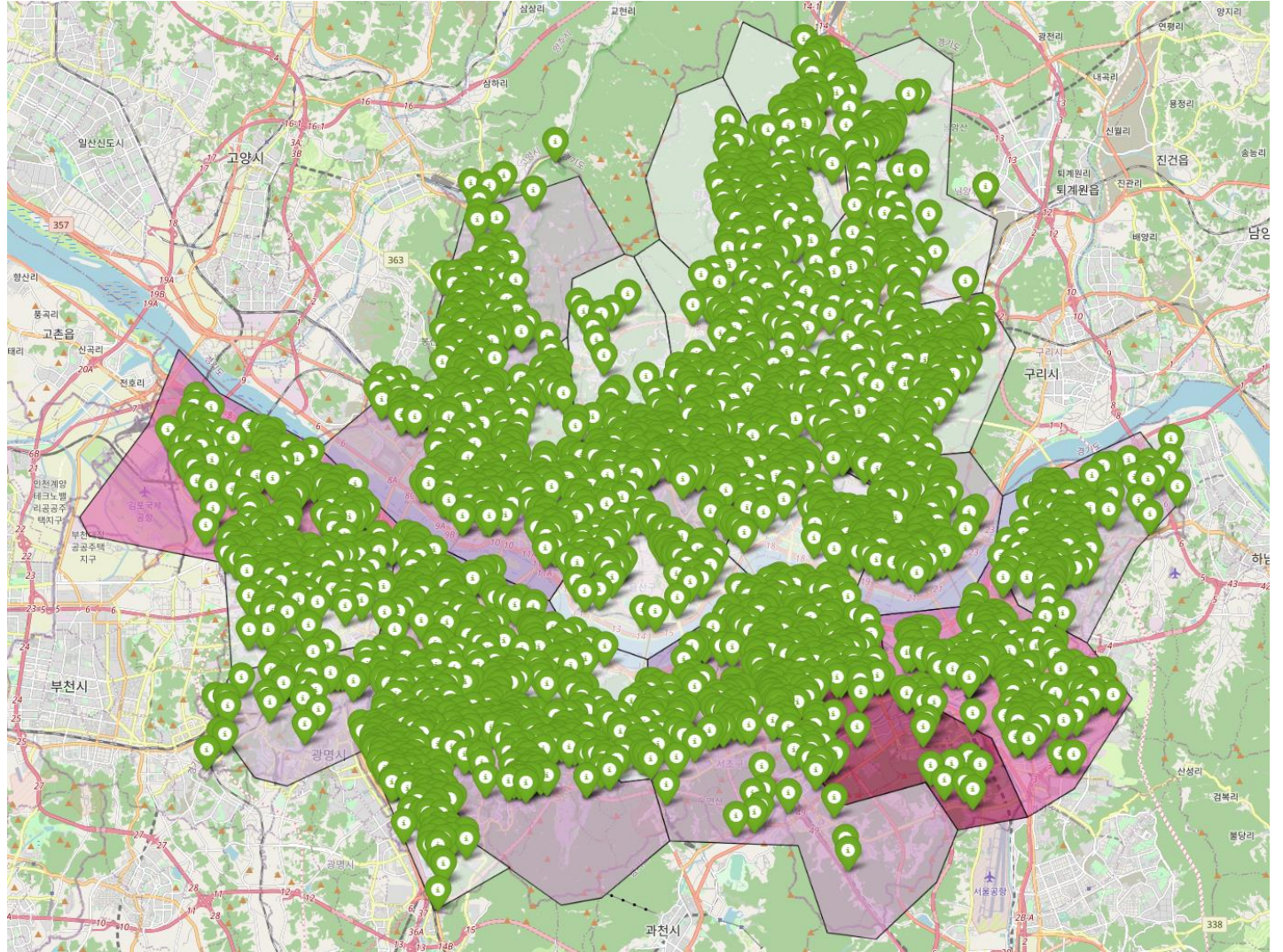


# 데이터 시각화 (서울시 'GS25' 편의점 시각화)



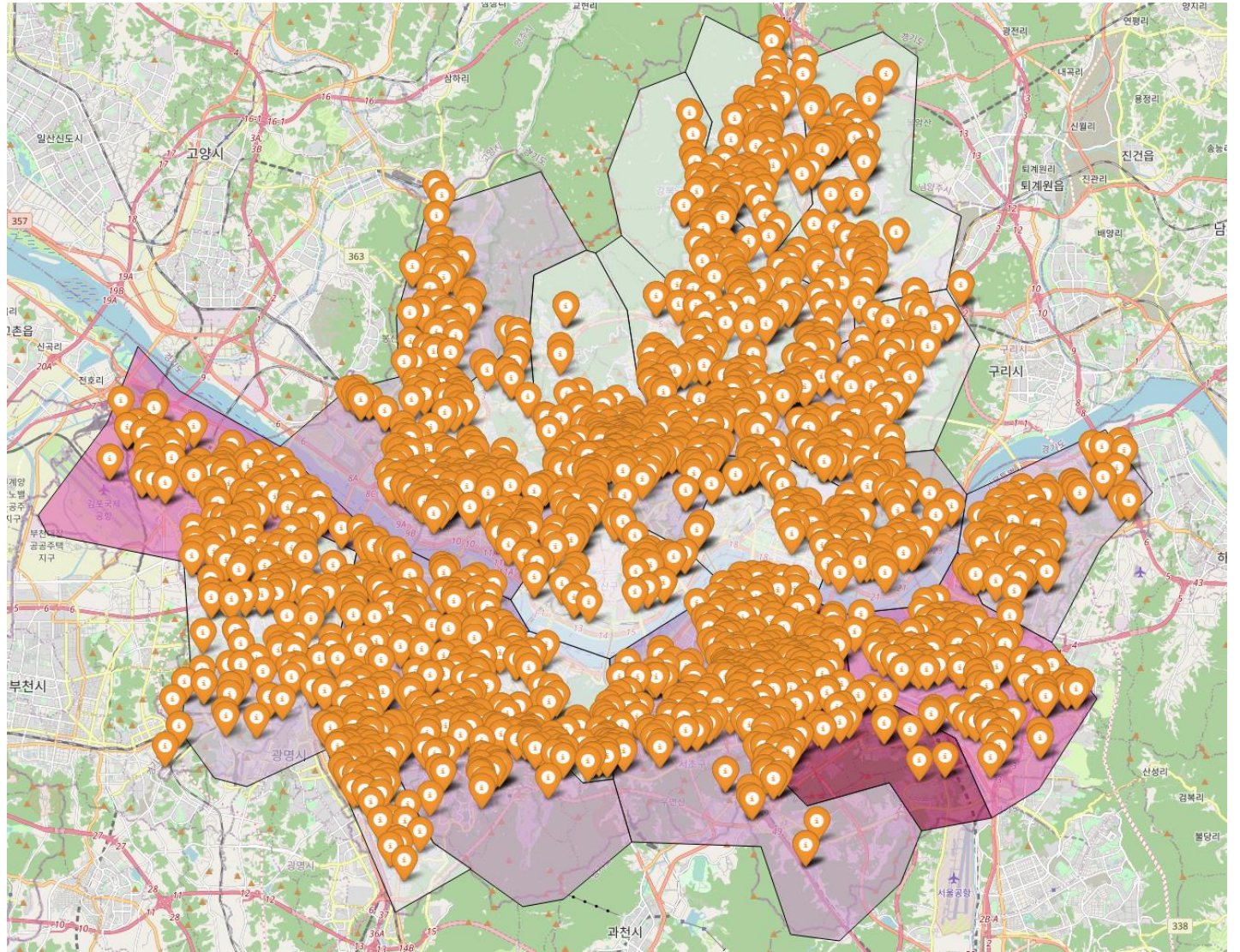


# 데이터 시각화 (서울시 'CU' 편의점 시각화)



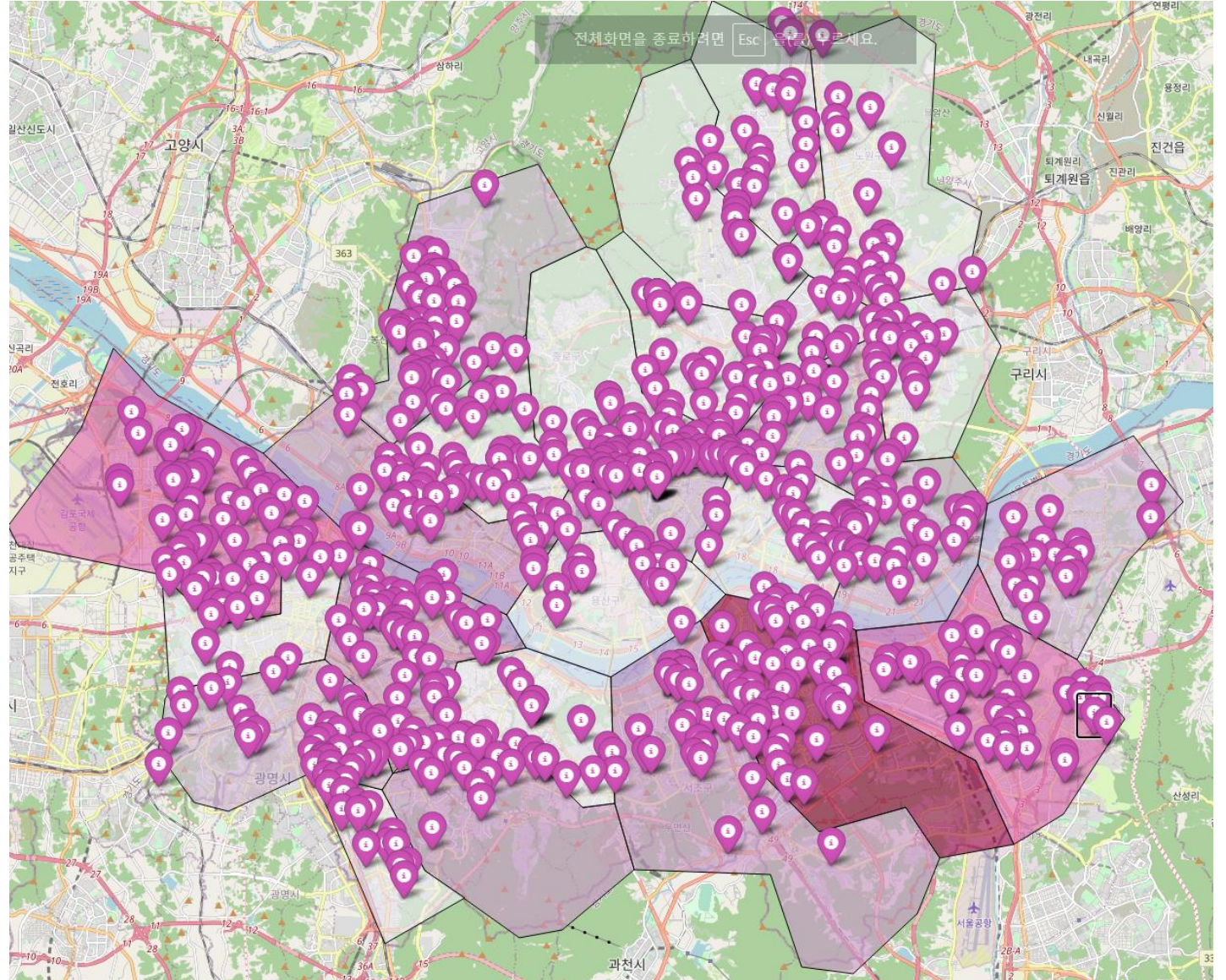


# 데이터 시각화 (서울시 '세븐일레븐' 편의점 시각화)



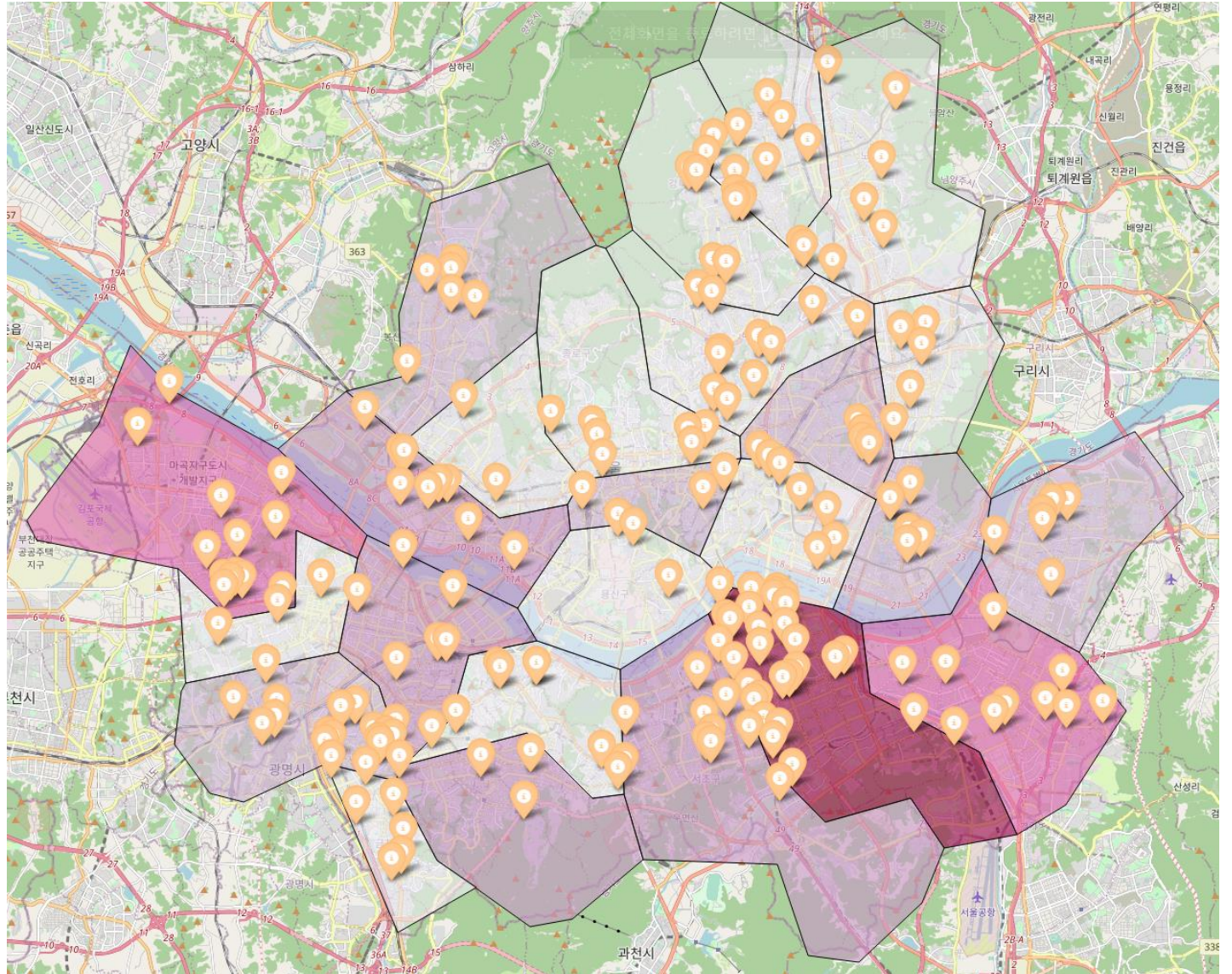


# 데이터 시각화 (서울시 '이마트24' 편의점 시각화)





# 데이터 시각화 (서울시 '미니스톱' 편의점 시각화)



# 결론

- 원래 계획은 전국을 기준으로 시각화를 하려 했으나, 너무 많은 데이터와 환경 문제로 인해 서울시로 축소하여 진행하게 된 부분이 아쉬운 부분이다.
- 편의점 점포 수 예상은 GS25, CU, 세븐일레븐, 미니스톱, 이마트24 순서였지만 예상과 다른 결과를 알 수 있어서 흥미로웠다.
- 편의점이 많은 줄은 알고 있었지만 직접 시각화를 하여 결과를 확인해 보니 한 눈에 볼 수 있어 좋은 경험이었던 것 같다.