

용인시 전기차 충전소 입지 선정 모델 개발 방안

-JazzIsHorse-



1 Data preprocessing

2 EDA

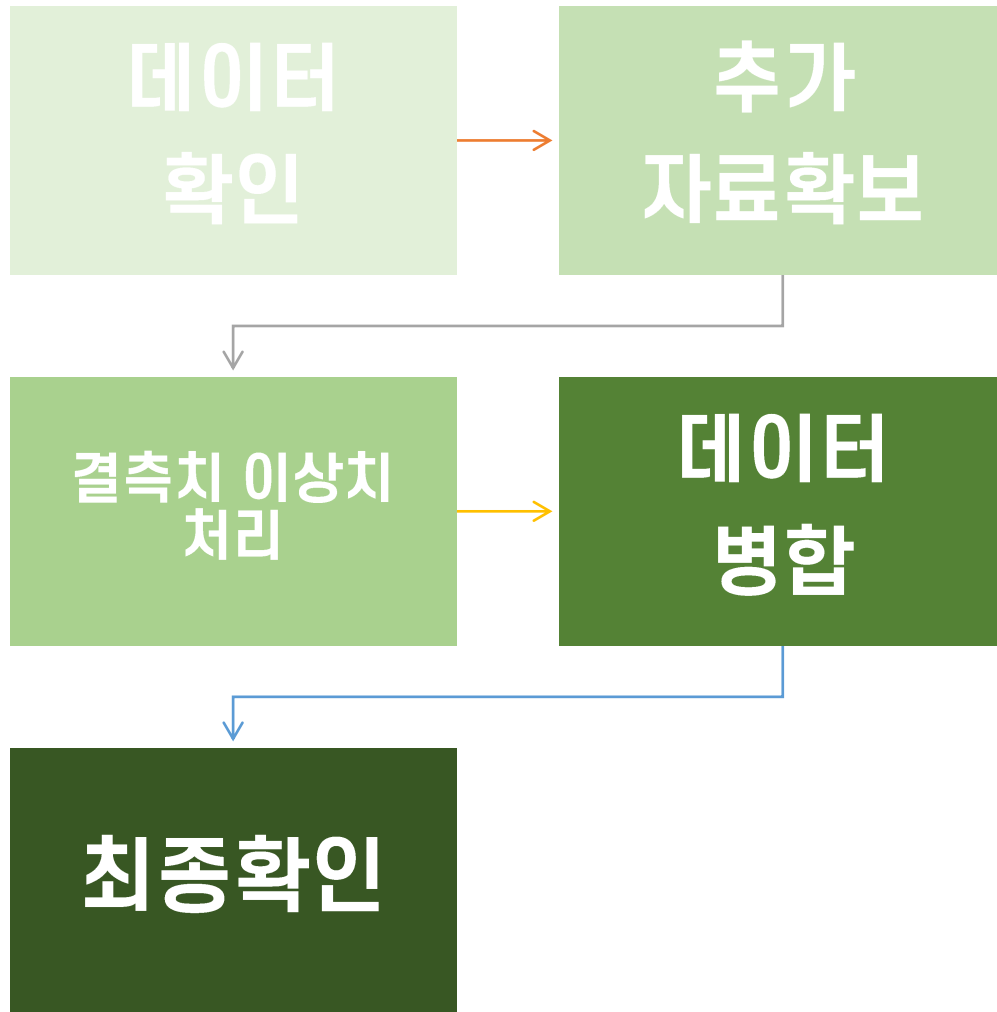
3 Conclusion

4 Reference

Part 1.

Data preprocessing





전처리 과정

1. 정의서(명세서) 기반 데이터 정보의 전반적인 확인
2. 공공데이터 포털에서 필요한 추가데이터 확보
3. 각 데이터의 이상치 및 결측치 확인 및 처리
4. 전처리가 완료된 데이터를 병합
5. 최종 확인

Part 1 Data preprocessing

자료형태확인

1. 용인시 거주자의 전기차 충전 앱 실행 292043 rows × 14 columns
2. 용인시 유동인구의 전기차 충전 앱 실행 264713 rows × 14 columns
3. 용인시 거주자의 타지역 충전 앱 실행 293171 rows × 16 columns

결측치 확인

```
pd.DataFrame([df1.isna().sum(), df2.isna().sum(), df3.isna().sum()])
```

| | 기준일자 | 요일 | 시도명 | 시군구명 | 행정동명 | 경도 | 위도 | 성별 | 연령 | 업종류 | 앱실행시간대 | 실행자수 |
|---|------|----|-----|------|------|----|----|----|----|-----|--------|------|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

용인시 전기차 충전소 현황 데이터 확인

- 한국환경공단_전기자동차 충전소 정보 // Open API에서 데이터 수집

```
{'statNm': '용인시청',  
 'statId': 'ME174147',  
 'chgerId': '01',  
 'chgerType': '06',  
 'addr': '경기도 용인시 처인구 삼가동 556',  
 'location': 'null',  
 'useTime': '24시간 이용가능',  
 'lat': '37.2406659',  
 'lng': '127.1773752',  
 'busiId': 'ME',  
 'bnm': '환경부',  
 'busiNm': '환경부',
```

갯수확인

```
len(res['items']['item'])
```

4457

결측치 확인

```
yong_elc.isna().sum()
```

```
Unnamed: 0      0  
충전소명      0  
충전소ID      0  
충전기ID      0  
충전기타입    0  
주소          4  
상세위치     2210  
이용가능시간  113  
위도          0  
경도          0  
기관 아이디   0  
기관명        0  
운영기관명    0  
운영기관연락처 116  
충전기상태    0  
상태갱신일시  193  
마지막 충전시작일시 1489  
마지막 충전종료일시 1476  
충전중 시작일시 1607
```

필요한 컬럼만 선별 후 저장

필요한 것만 가져오기

충전소명, 충전기타입, 위도, 경도, 충전용량, 충전방식, 주차료무료

```
yong_elc = yong_elc[['충전소명', '충전기타입', '위도', '경도', '충전용량', '충전방식',  
yong_elc.head(3)
```

| | 충전소명 | 충전기타입 | 위도 | 경도 | 충전용량 | 충전방식 | 주차료무료 |
|---|-------------|-------|-----------|------------|------|------|-------|
| 0 | 용인시청 | 6 | 37.240666 | 127.177375 | 50.0 | 단독 | N |
| 1 | 포곡읍사무소 | 6 | 37.278100 | 127.231259 | 50.0 | 단독 | N |
| 2 | 죽전휴게소(서울방향) | 6 | 37.332073 | 127.104848 | 50.0 | 단독 | Y |

- 충전기타입 (01:DC차데모, 02: AC완속, 03: DC차데모+AC3상, 04: DC콤보, 05: DC차데모+DC콤보)
- 충전방식 (단독/동시)
- 충전용량 kW (3, 7, 50, 100, 200)

데이터 병합

총 개수 확인

```
print(len(df1), len(df2), len(df3), ' = ', len(df1) + len(df2) + len(df3))
```

292043 264713 293171 = 849927

병합

```
df_sum = pd.concat([df1, df2, df3], join = 'outer', axis = 0)
```

```
len(df_sum)
```

849927

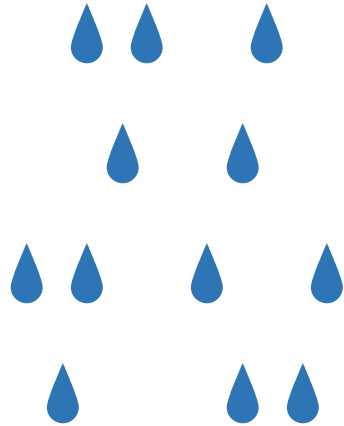
병합확인

```
df_sum.tail(2)
```

| | 기준일자 | 요일 | 시도명 | 시군구명 | 행정동명 | 경도 | 위도 | 성별 |
|--------|----------|----|-------|------|------|------------|-----------|--------|
| 293169 | 20220619 | 일 | 서울특별시 | 강남구 | 역삼1동 | 127.033035 | 37.499966 | MALE |
| 293170 | 20220605 | 일 | 경기도 | 용인시 | 영덕1동 | 127.072790 | 37.272970 | FEMALE |

```
df_sum.shape
```

(849927, 13)



데이터 확인

명세

데이터 병합



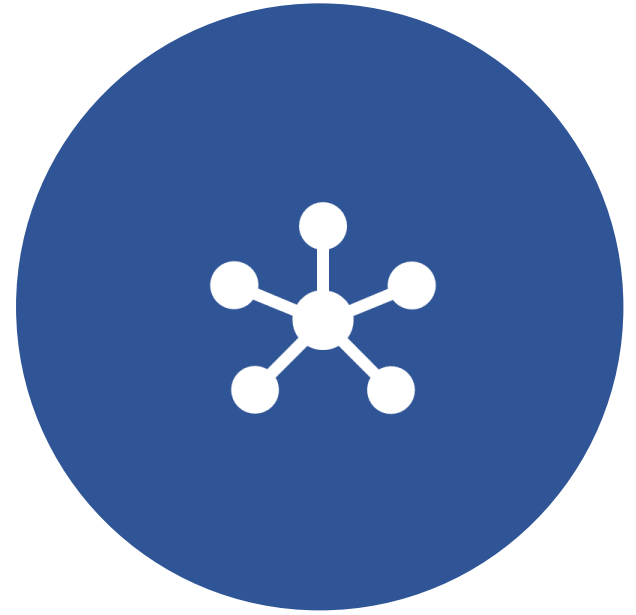
Part 2, Exploratory Data Analysis



충전소 현황 파악



전기차 현황 파악

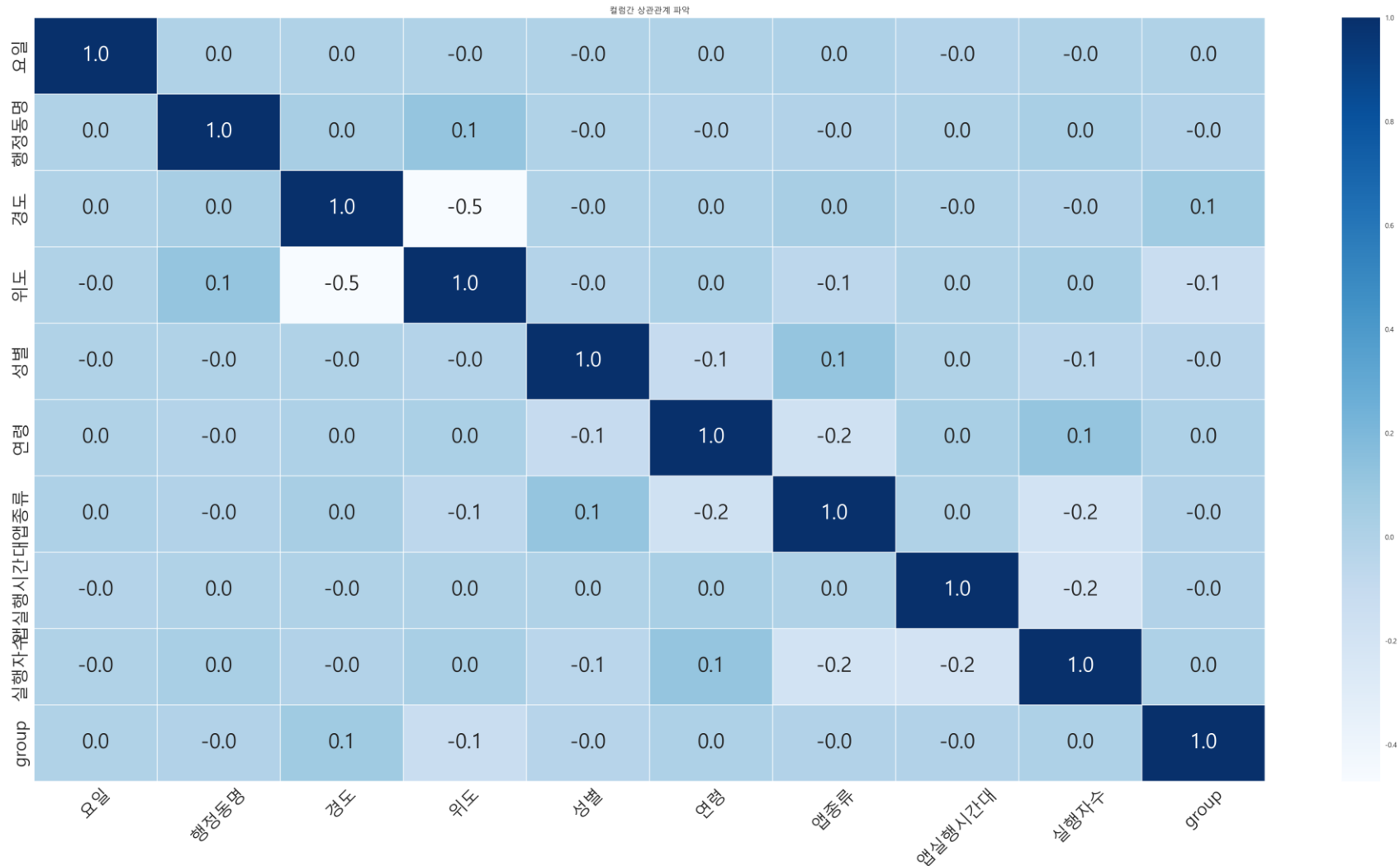


앱실행 군집화

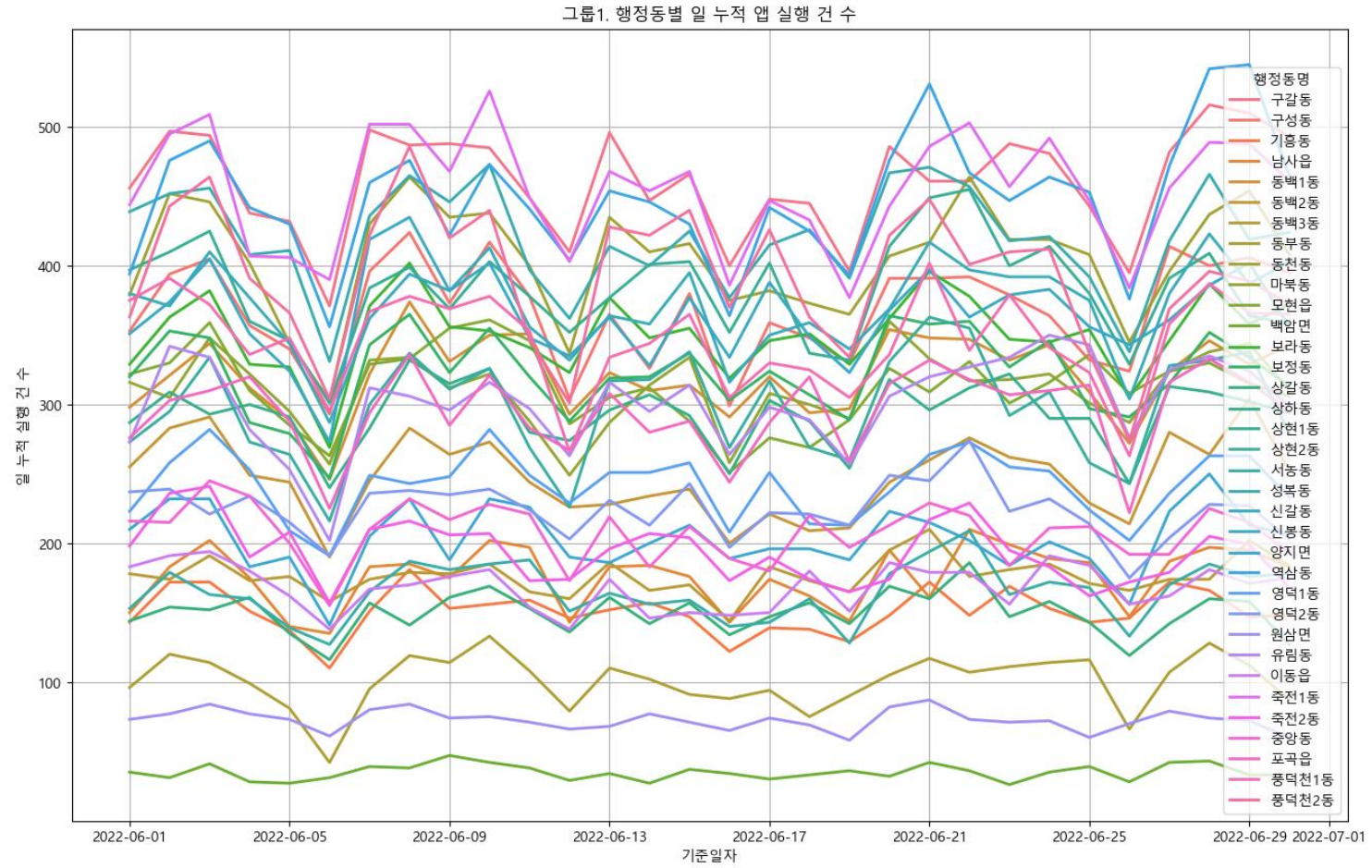

```
# 각 그룹의 일별 행정동 실행자수의 최소최대값, 평균값, 사분위수 산출
df_sqrt = df_sum.groupby(['group', '기준일자', '행정동명'])['실행자수'].describe()
df_sqrt
```

| | | | count | mean | std | min | 25% | 50% | 75% | max |
|-------|----------|------|-------|-----------|----------|-----|-----|-----|------|------|
| group | 기준일자 | 행정동명 | | | | | | | | |
| 1 | 20220601 | 구갈동 | 456.0 | 5.109649 | 3.233247 | 3.0 | 3.0 | 3.0 | 6.0 | 21.0 |
| | | 구성동 | 353.0 | 5.770538 | 3.849592 | 3.0 | 3.0 | 4.0 | 7.0 | 26.0 |
| | | 기흥동 | 143.0 | 5.972028 | 4.090429 | 3.0 | 3.0 | 4.0 | 8.0 | 27.0 |
| | | 남사읍 | 150.0 | 5.660000 | 3.301068 | 3.0 | 3.0 | 4.0 | 7.0 | 19.0 |
| | | 동백1동 | 298.0 | 5.399329 | 3.750412 | 3.0 | 3.0 | 3.0 | 7.0 | 22.0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 3 | 20220630 | 화서2동 | 1.0 | 3.000000 | NaN | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 |
| | | 화수2동 | 2.0 | 5.000000 | 2.828427 | 3.0 | 4.0 | 5.0 | 6.0 | 7.0 |
| | | 화양동 | 7.0 | 11.000000 | 6.633250 | 4.0 | 6.5 | 8.0 | 15.0 | 22.0 |
| | | 화정1동 | 3.0 | 3.000000 | 0.000000 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 |
| | | 회현동 | 26.0 | 5.000000 | 3.346640 | 3.0 | 4.0 | 4.0 | 4.0 | 20.0 |

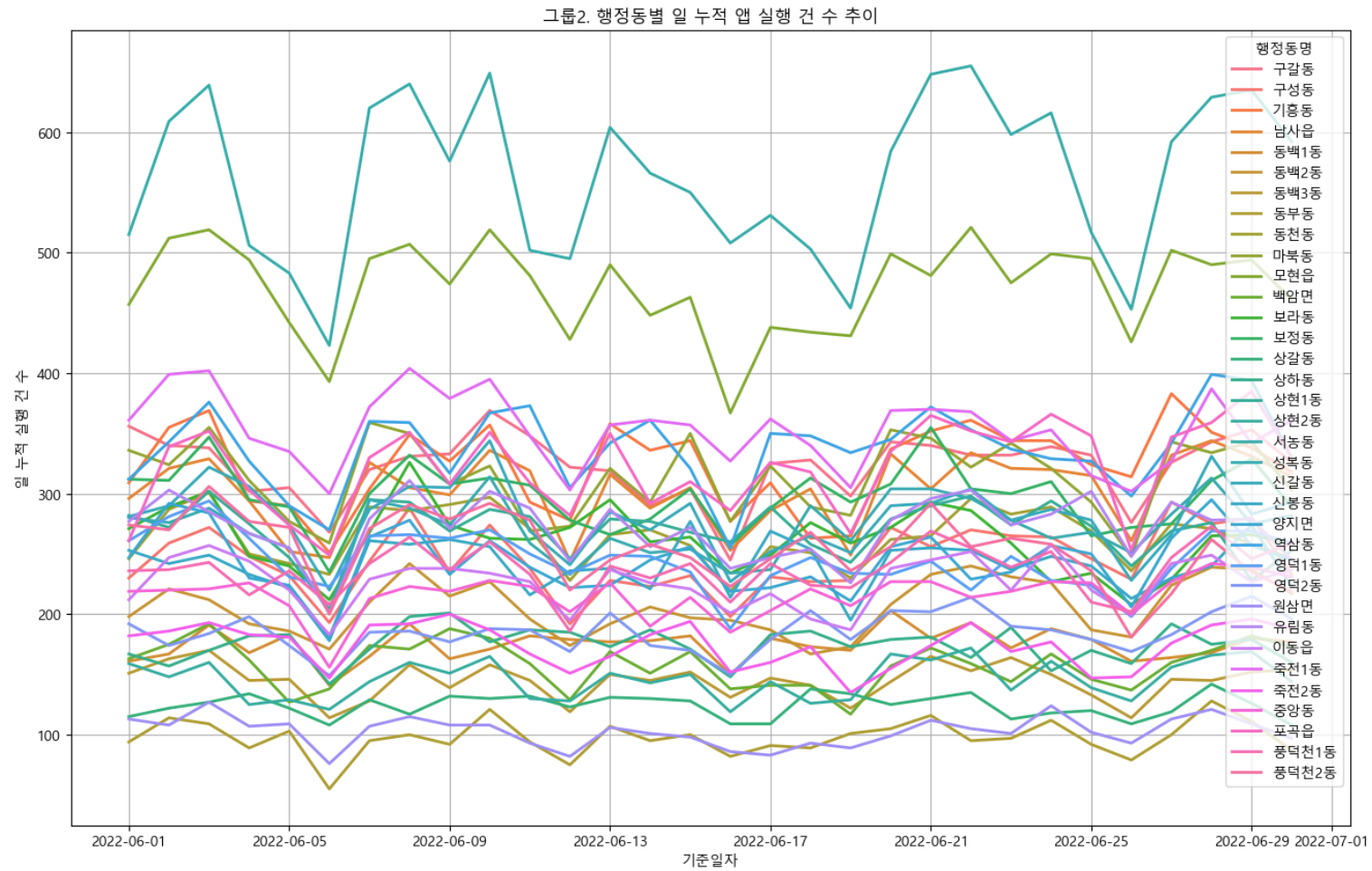
병합된 데이터를 바탕으로 그룹, 기준일자, 행정동명으로 그룹화하여 일자별 각 행정 구역의 산술 통계량을 산출.



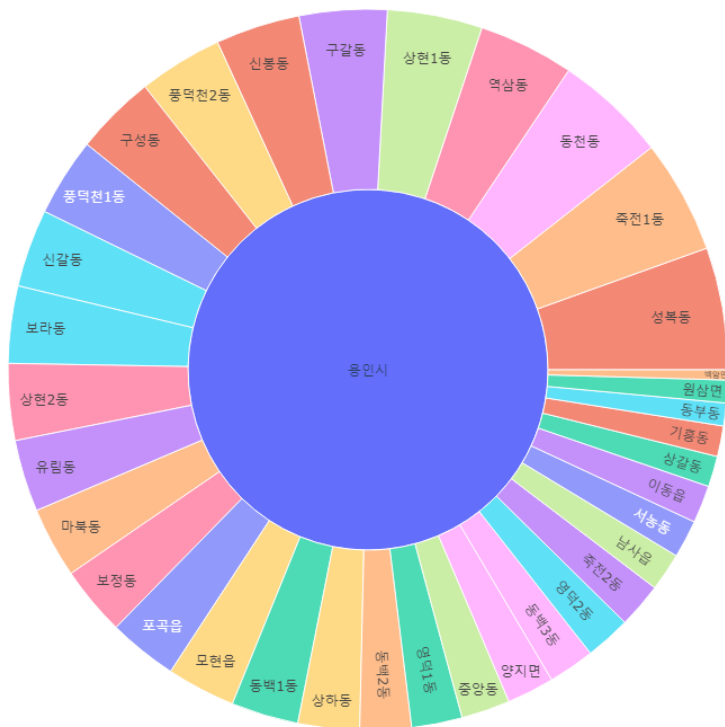
전처리된 컬럼들을 라벨 인코딩하여 각 컬럼간 상관계수를 파악.



그룹1. 용인시 거주자의 행정동별 일 누적 앱실행 건 수

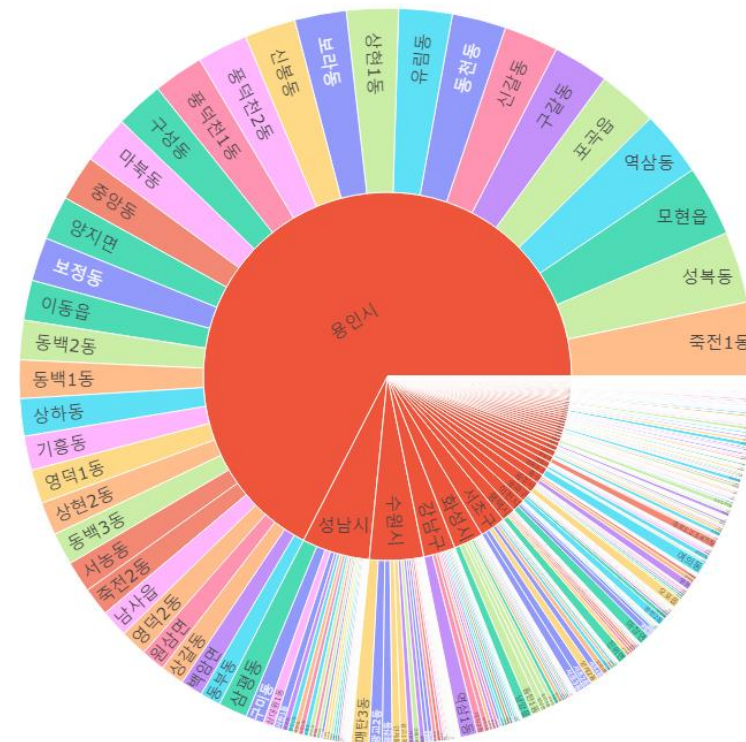
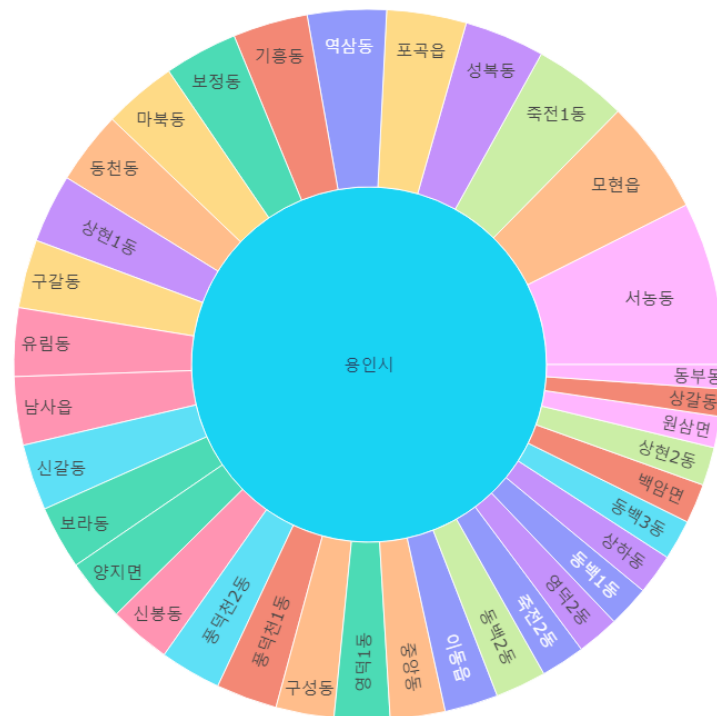


그룹2. 용인시 유동인구의 행정동별 일 누적 앱실행 건 수



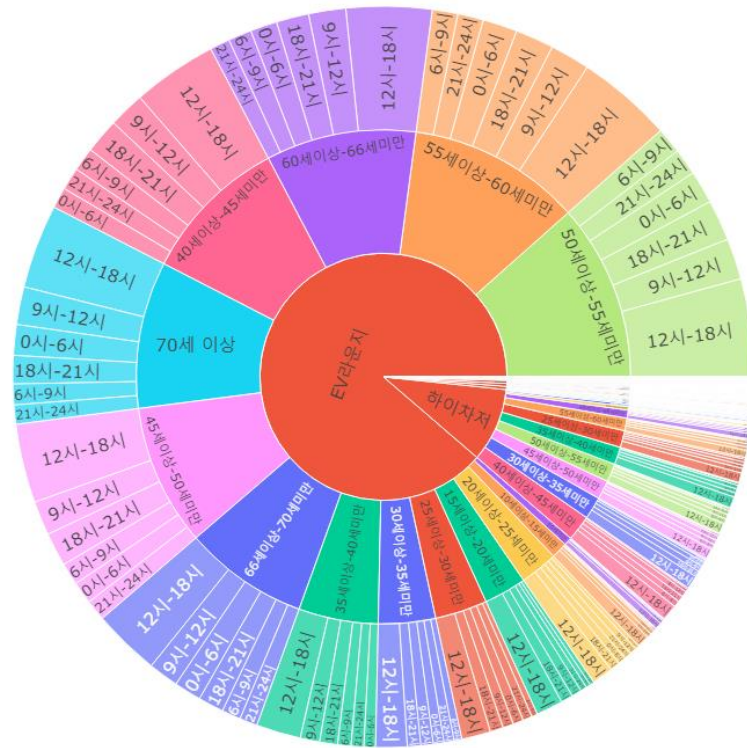
그룹1.
용인시 거주자
행정동 별 앱 실행 빈도

그룹2.
용인시 유동인구
행정동 별 앱 실행 빈도

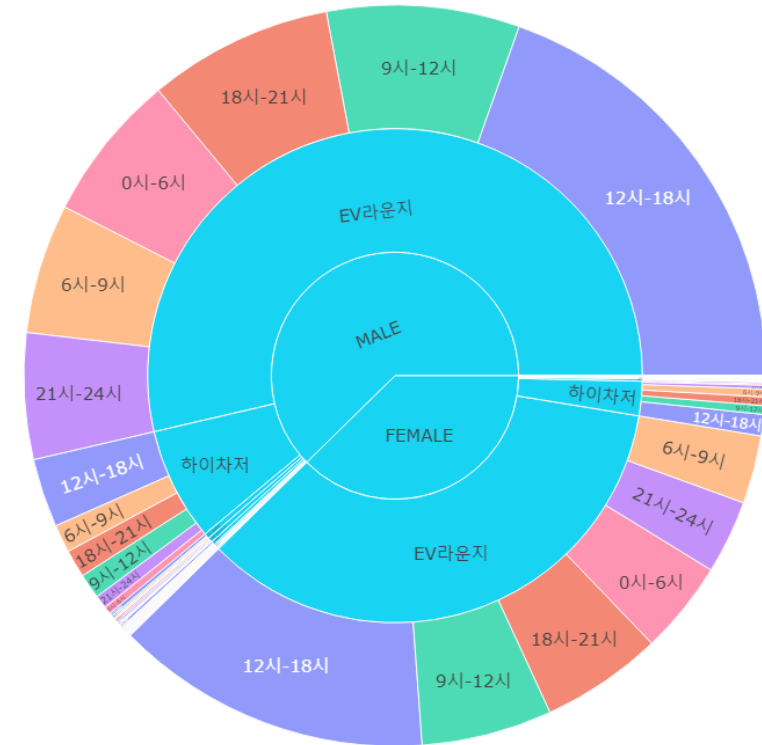


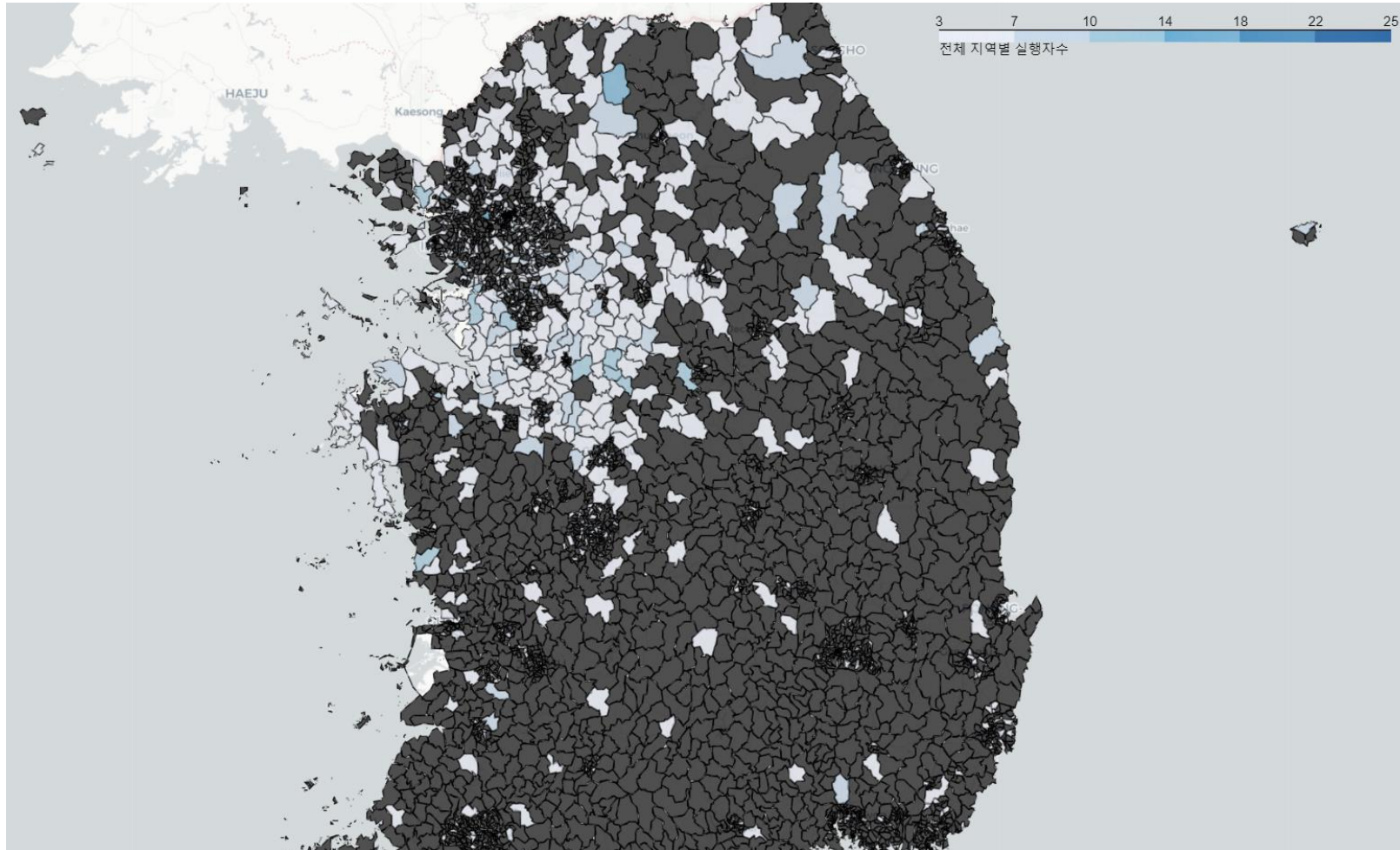
그룹3.
용인시 거주자의
시외 포함 앱 실행 빈도

앱종류, 연령별 이용시간대 빈도 분류



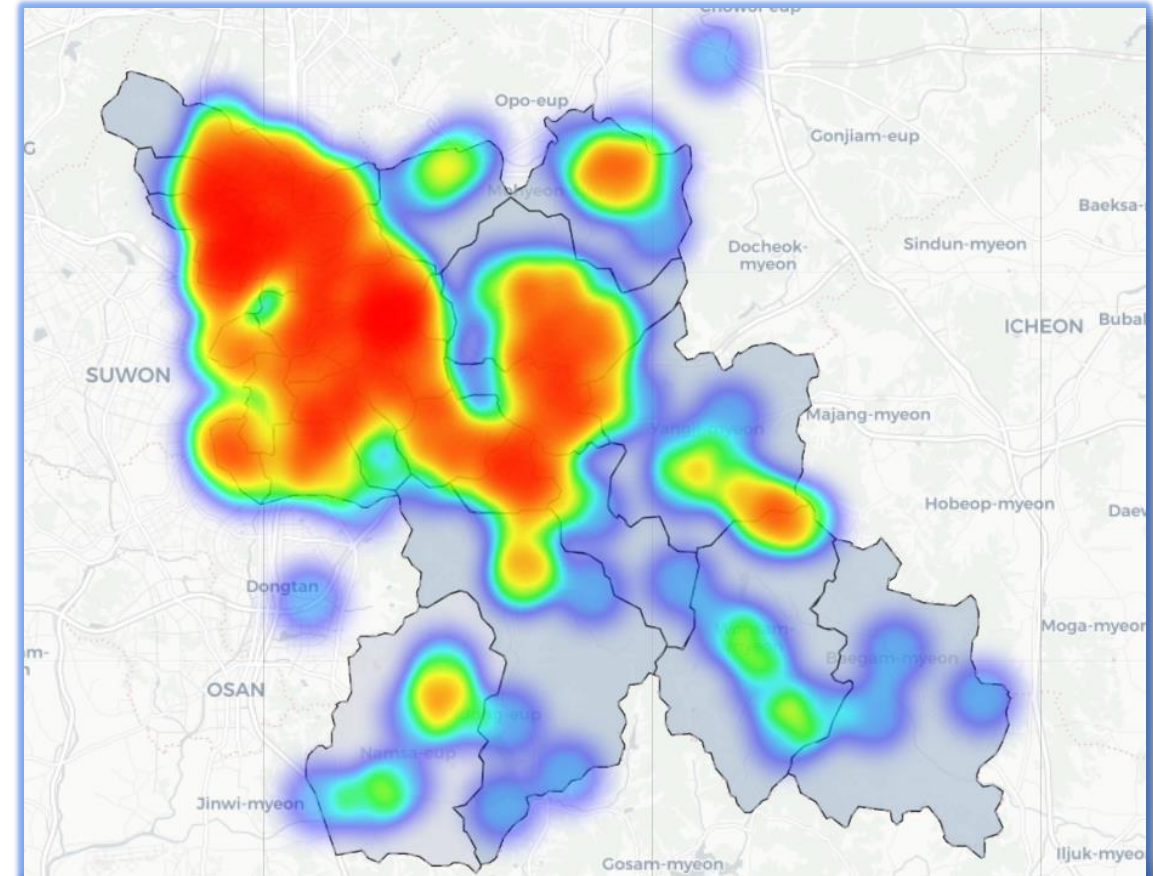
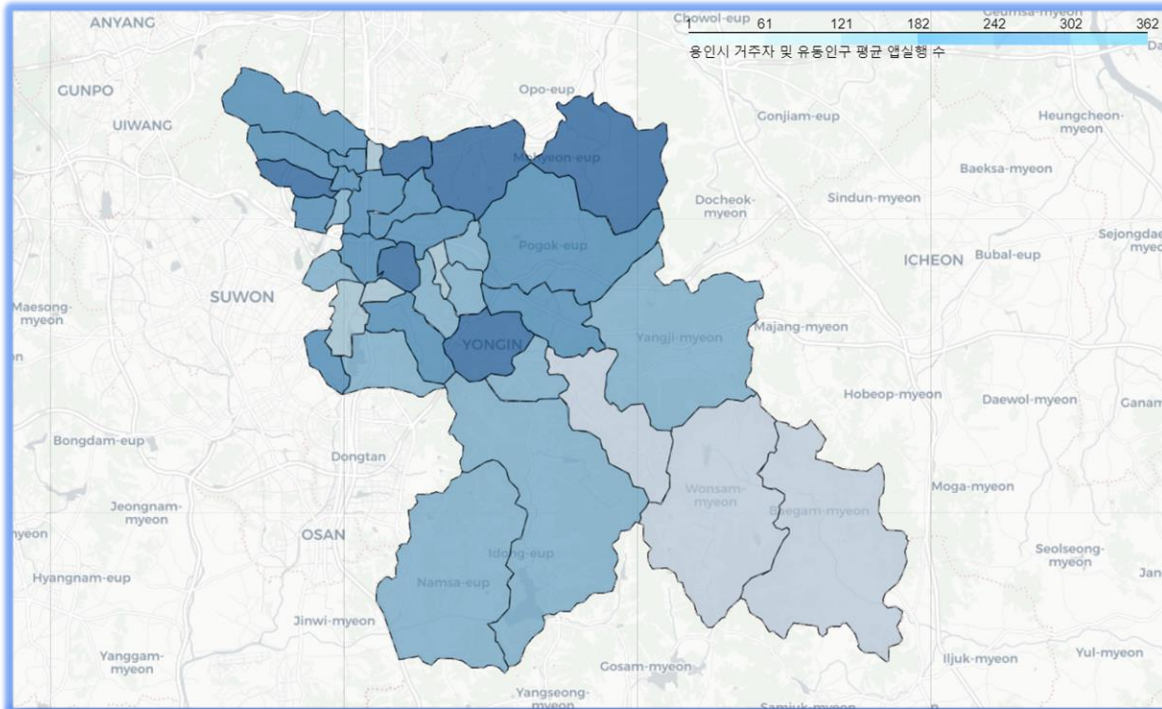
성별, 앱종류, 이용시간대 빈도 분류

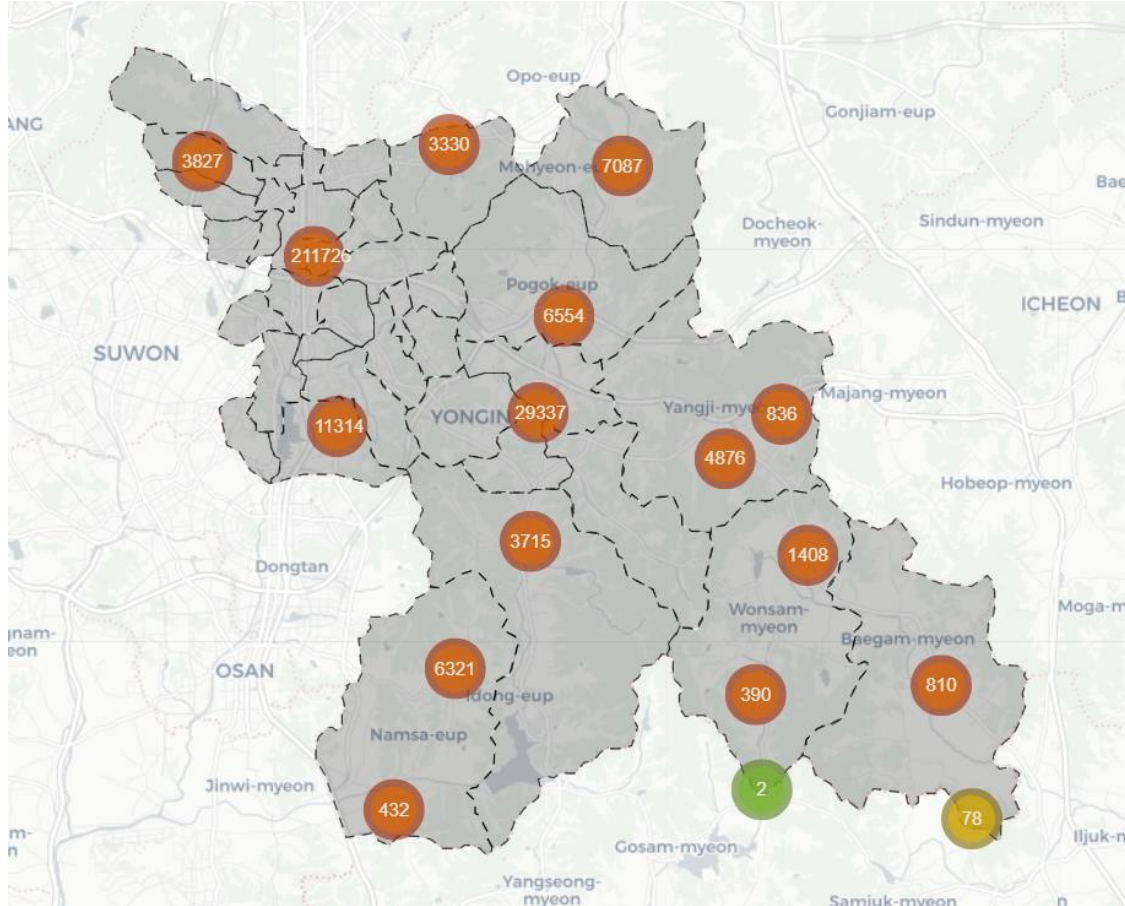




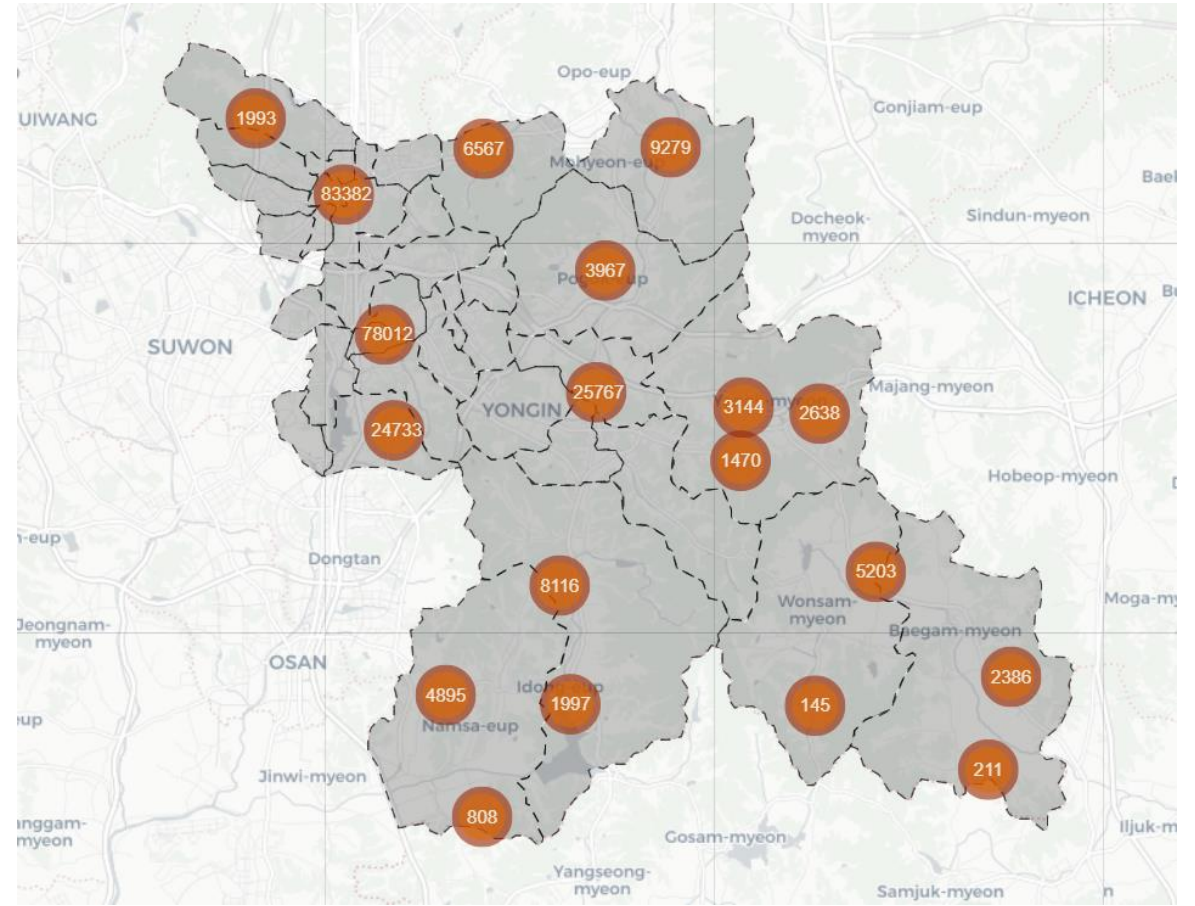
전체 그룹에서 평균 앱실행 건수를 지역별 대조.
경기도를 포함한 강원도 일대에서 상대적으로 높게 나타남.

제공데이터 기반 용인시 거주자 및 유동 인구 평균 앱 실행 수를 행정구역별로 지도 시각화

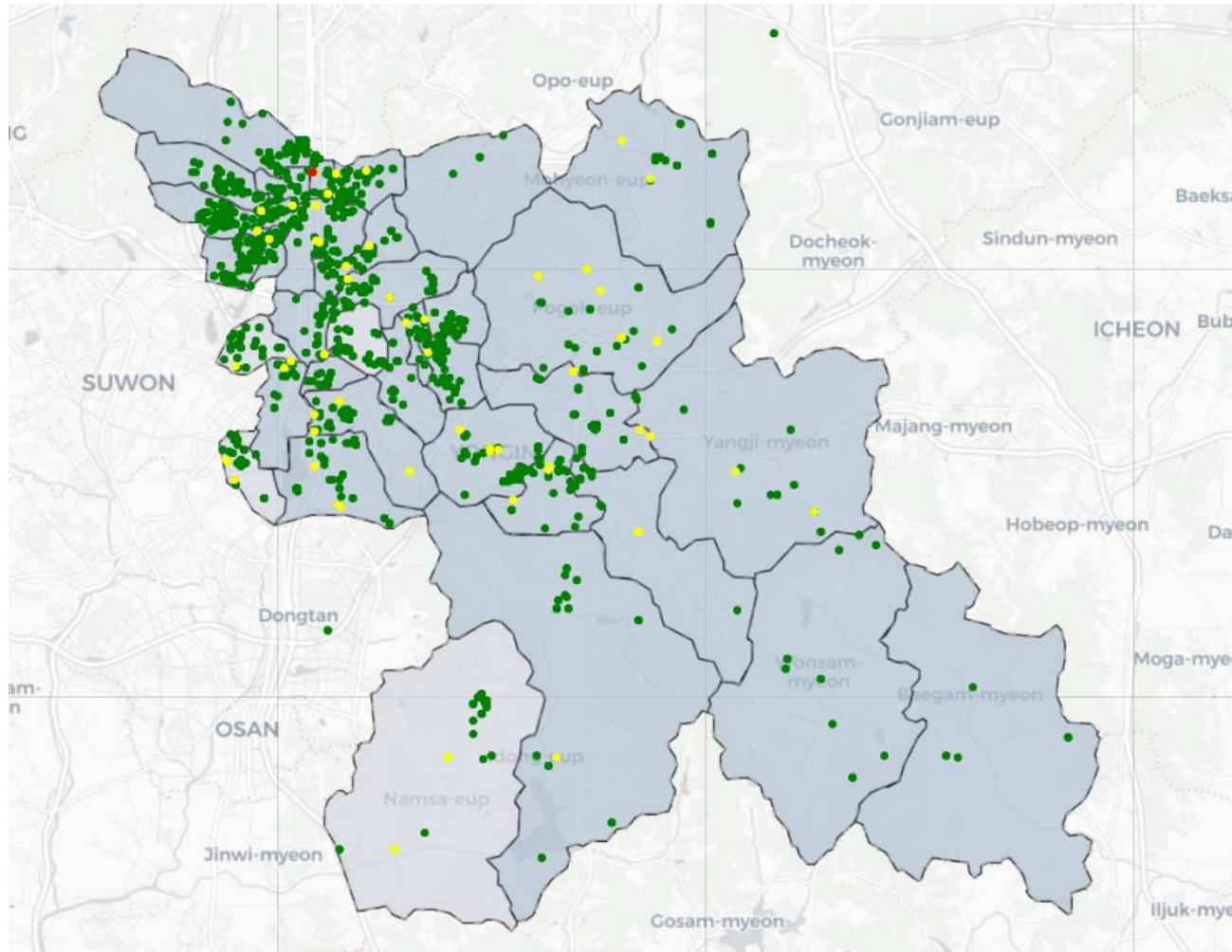




그룹1.
용인시 거주자의 행정동별 앱실행 지도 시각화



그룹2.
용인시 유동인구의 행정동별 앱실행 지도 시각화



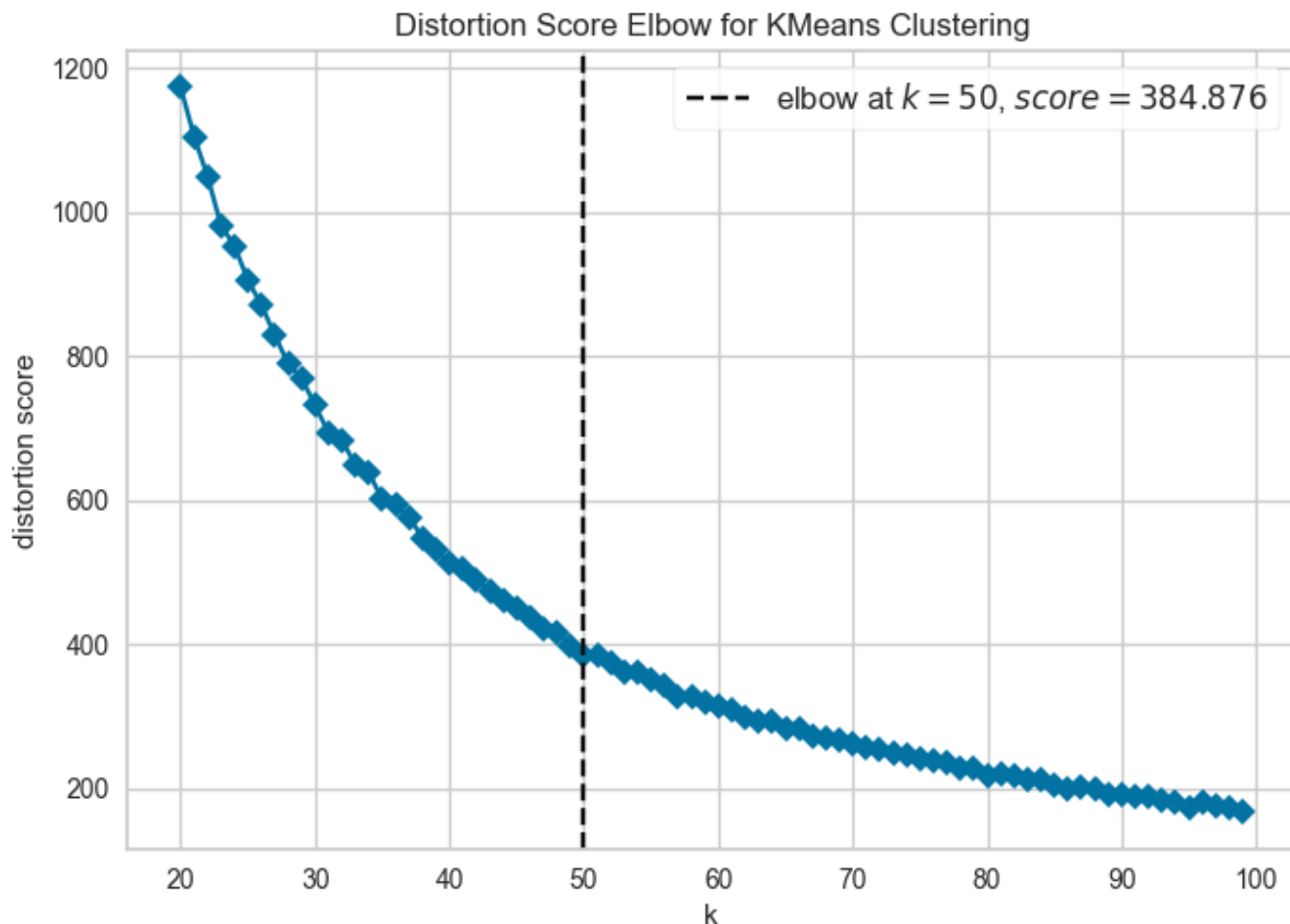
한국환경공단에서 제공하는 전기차 충전
소 OpenAPI 데이터를 수집하여
지도 시각화.

색상 구분

Red : 초급속(300 ~ 350 kW)

Yellow : 급속(50 ~ 200 kW)

Green : 완속(3, 7 kW)

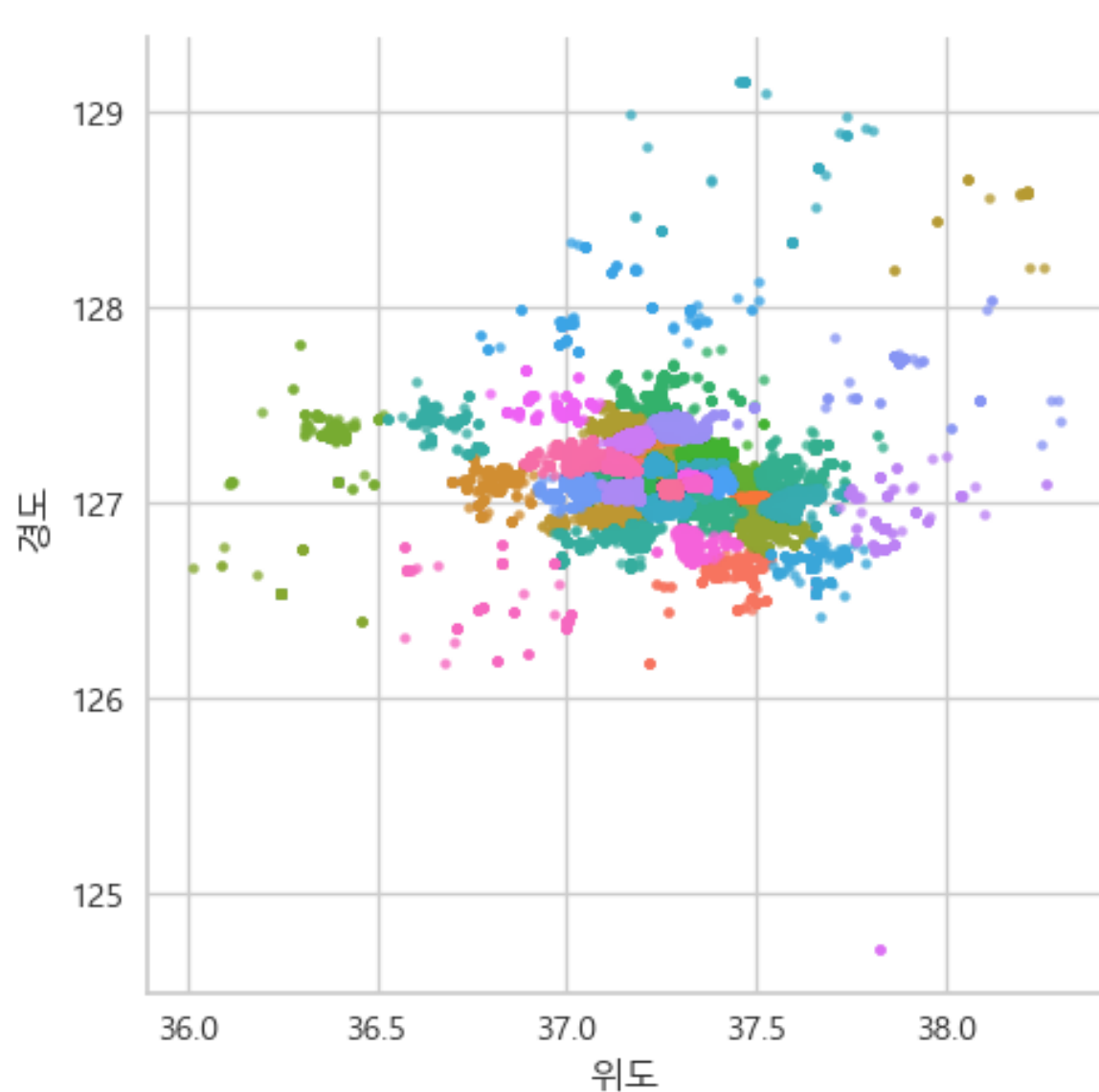


Unsupervised Learning Elbow for KMeans Clustering

Elbow Method

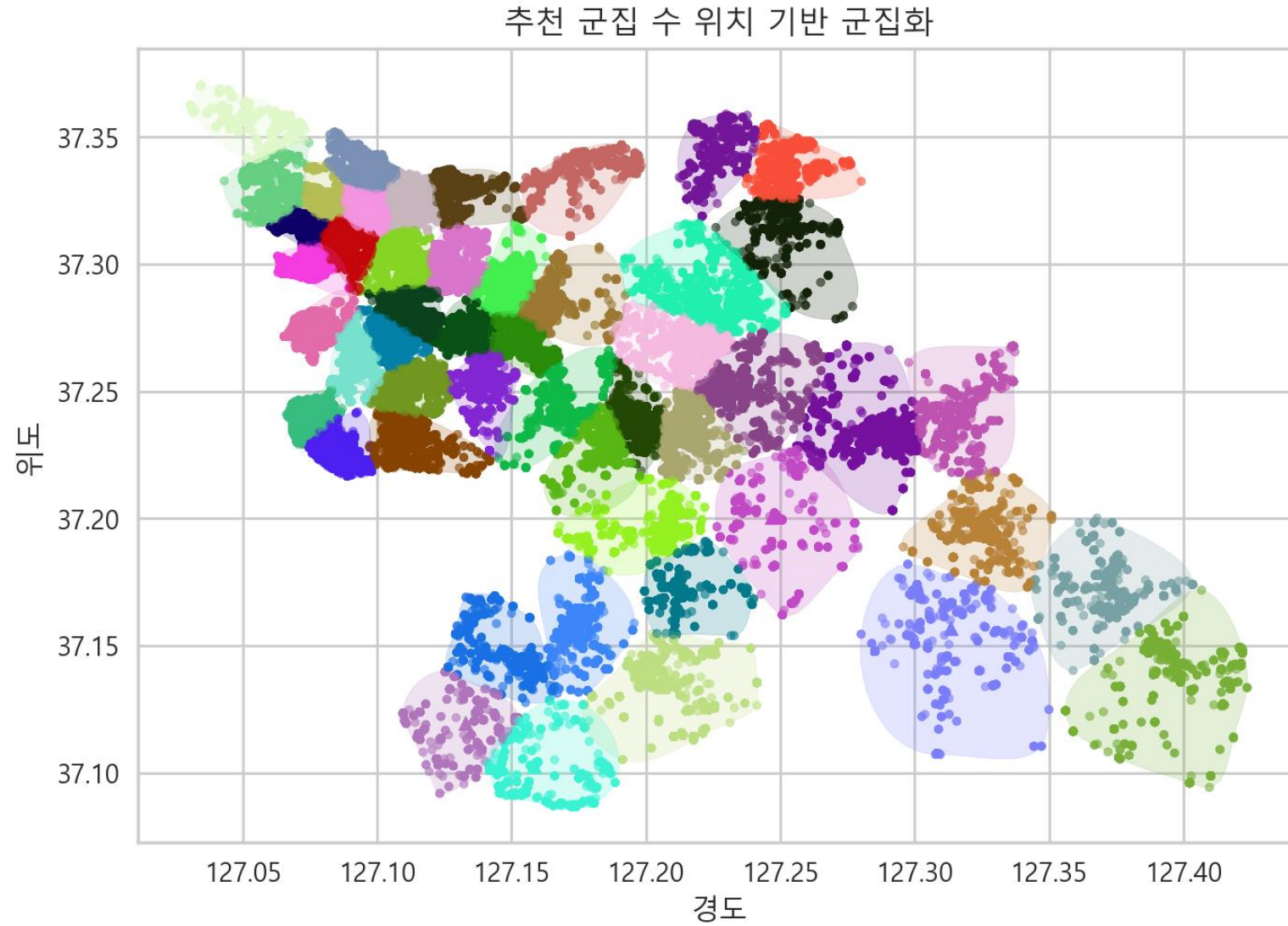
k 개수에 따라 cost function J 를 그려보았을 때, 특정 K 이후 cost가 거의 변하지 않는 elbow point가 있다면 해당 k 를 선택하는 방법.

그룹 1,2,3을 병합한 88만개의 앱실행 데이터의 [위도, 경도]를 통해 도출한 최적 군집의 개수는 50개로 나타남.



| 군집 번호 | | | |
|-------|----|----|----|
| 0 | 13 | 26 | 38 |
| 1 | 14 | 27 | 39 |
| 2 | 15 | 28 | 40 |
| 3 | 16 | 29 | 41 |
| 4 | 17 | 30 | 42 |
| 5 | 18 | 31 | 43 |
| 6 | 19 | 32 | 44 |
| 7 | 20 | 33 | 45 |
| 8 | 21 | 34 | 46 |
| 9 | 22 | 35 | 47 |
| 10 | 23 | 36 | 48 |
| 11 | 24 | 37 | 49 |
| 12 | 25 | | |

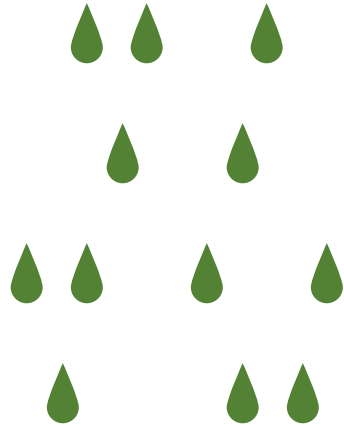
**앞서 산출한 결과인 $k = 50$ 을 바탕으로
전체 위경도 데이터에 대한 군집화를 시행.**



Part 3,

Conclusion





등록전기차수

총군집수

군집내앱실행수

군집별산출

- 용인시 내 군집별 앱활성화 비율 산출
 - 용인시 분류 군집별 앱실행수 / 용인시 전체 앱실행수 * 100
- 차후 수요 파악을 위해 현행 전기차 수의 시계열 추이 조사를 시행한 후 구역 내 점진적 증가 비율을 판단해 보급대수 지수에 반영
- 등록전기차를 K-elbow로 산정된 군집수로 나누어 필요 개수를 산출함.

```
# 현재 용인시 등록 전기차
ev_car = 5957
print('총 충전소 필요개수 :', ev_car / 50, '\n앱실행 : ', cluster_cnt[19])
```

총 충전소 필요개수 : 119.14
앱실행 : 28325

```
# 예제 군집 19번에 대해서 산출
print(f' 용인시에는 총 {(ev_car / 50)} 전기차 충전소가 필요하며,')
print(f' {19}번 군집은 전체 앱실행 비율의 {round((cluster_cnt[19] / sum(cluster_cnt)) * 100)}% 를 차지합니다.')
print(f' 따라서 {19}번 군집에는 {round((ev_car / 50) * (cluster_cnt[19] / sum(cluster_cnt)))}개의 전기차 충전소가 필요합니다.')
```

용인시에는 총 119.14 전기차 충전소가 필요하며,
19번 군집은 전체 앱실행 비율의 5% 를 차지합니다.
따라서 19번 군집에는 6개의 전기차 충전소가 필요합니다.

산출식





산출식

최적 개수 산출

군집별 최적 개수 = ((등록전기차 수 * 예상 증가율) /
범위 내 총 군집 수) * 개별 군집의 비율

최적 위치 산출

군집별 최적 개수 * (개별군집기준점 - 각 좌표간 거리)
이중군집화

1. 용인시 전기차 완속 충전소 입지 선정 모델 개발 데이터셋(빅콘테스트 제공)
2. 한국환경공단_전기자동차 충전소 정보_OpenAPI
3. 경기도 용인시_전기차 등록현황_20221001 (공공데이터 포털)
3. 공공시설 입지선정을 위한 입지모델 구축 및 적용에 관한 연구 -금산군 문화시설을 중심으로-(윤정미 외)
4. 경기도_전기차_충전시설_적정_설치기준_연구-(강철구 외)
5. 공공빅데이터를 활용한 산업단지내 공장입지 선정을 위한 매력도 지수 생성 및 시각화 연구-(이민희 외)





Thank you