EDA 분석

1. **EDA 분석을 위한 창원 분석 데이터 (총 13개)**

- 창원시 시·군·구, 읍·면·동 지명(‘address’)

- 지역 면적(‘area’)

- 교통량(‘traffic’, ‘traffic/area’)

- 기온(‘평균 기온’, ‘최대 기온’)(‘avg\_temp’, ‘max\_temp’)

- 미세먼지 데이터(‘dust\_naver’)

- 유동 인구 데이터(‘f\_pop’, ‘f\_pop/area’)

- 총 인구 수(‘p\_total’)

- 노약자 인구 데이터(‘young\_old’, ‘yo\_ratio’, ‘young\_old/area’)

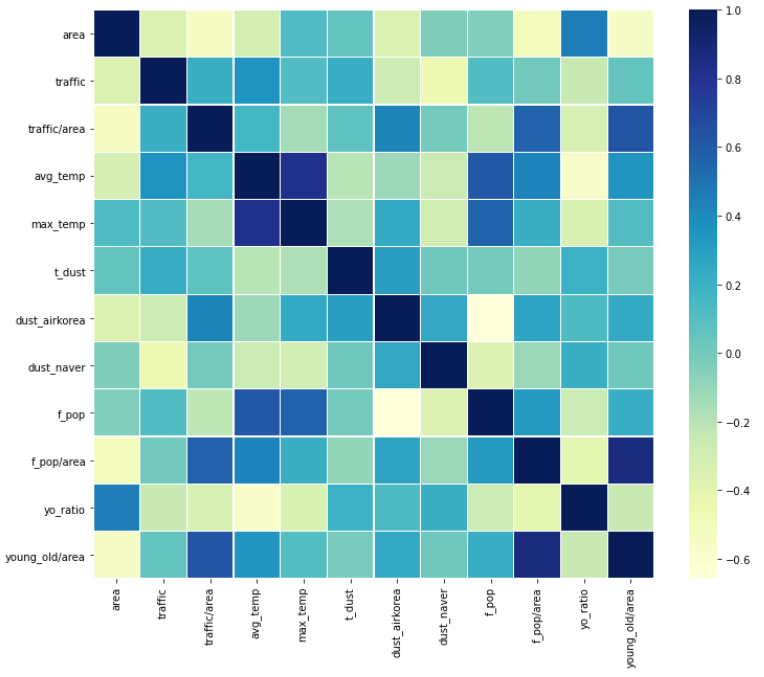
1. **상관계수 분석**

**-전체 상관계수 분석**

텍스트, 창문이(가) 표시된 사진

자동 생성된 설명

**(시각화 자료)**



**-상관계수 분석 결과**

* 상관계수가 높은 데이터 종류

①교통량 데이터(‘traffic’, ‘traffic/area’): 기온, 미세먼지 데이터

②기온 데이터(‘max\_temp’, ‘avg\_temp’): 교통량, 인구 수 및 노약자 데이터

③미세먼지 데이터(‘dust\_naver’): 교통량, 기온, 유동인구 데이터

④인구 데이터(‘f\_pop’, ‘f\_pop/area’, ‘p\_total’): 기온, 노약자 데이터

⑤노약자 데이터(‘young\_old’, ‘yo\_ratio’, ‘young\_old/area’): 기온, 인구 데이터와 상관계수가 높게 나옴

* ‘area’로부터 파생된 변수들이 서로 상관관계 높게 나옴

Ex) 지역 면적당 노약자 비율(‘young\_old/area’), 지역 면적당 교통량(‘traffic/area’)

* 교통량과 미세먼지는 양의 상관계수를 가지고 있을 것이라 예상했지만, 실제 값은 음의 상관계수를 가지고 있는 것으로 나타남

이유: 산업단지에서 나오는 미세먼지가 상관계수에 영향을 주는 것으로 판단

1. **데이터 검증 및 통합**

지명(‘address’) → 지명(‘address’)

* 행정 구역 지명 데이터 그대로 사용

교통량(‘traffic’, ‘traffic/area’) → 교통량(‘‘traffic/area’)

인구 데이터(‘f\_pop’, ‘f\_pop/area’, ‘p\_total’) → 인구 데이터(‘f\_pop/area’)

노약자 인구 데이터(‘yo\_ratio’, ‘young\_old/area’) → 노약자 인구 데이터(‘young\_old/area’)

* 각 시,군,구 지역의 전체 값을 나타내는 [‘traffic’, ‘f\_pop’, ‘p\_total’, ‘young\_old’], 4개의 변수는 지역 면적에 따라 데이터 값이 편향적으로 분포되어 있어 쿨링로드 입지선정을 위한 변수에 부적절하다고 판단
* 각 시,군,구의 데이터 값을 지역 먼적(‘area’)로 나눈 [‘‘traffic/area’, ‘f\_pop/area’, ‘young\_old/area’]을 변수로 선정
* 노약자 인구 데이터 중 노약자 인구 비율을 나타내는 ‘yo\_ratio’ 보단 지역 면적당 노약자 수를 나타내는 ‘young\_old/area’가 입지 분석을 위한 변수로 더 적합하다고 판단

기온(‘평균 기온’, ‘최대 기온’)(‘avg\_temp’, ‘max\_temp’) → 기온(‘최대 기온’) (‘max\_temp’)

* 지역별 평균 기온과 최대 기온을 비교해 봤을 때 평균 기온은 동일한 시,군,구 별로 차이가 거의 없어 입지선정을 위한 변수에 부적절하다고 판단
* 쿨링로드 설치 목적 및 기대효과 중 하나인 여름철 폭염 대비와 도시 열섬 완화를 위해서 평균 기온보단 최대 기온이 입지선정을 위한 변수에 더 적합하다고 판단

1. **데이터 분석 및 시각화**
2. *최종 선정된 데이터 상위 10개 지역 추출*

* **지역 면적당 교통량** -  **최대 기온(max\_temp)**

텍스트이(가) 표시된 사진

자동 생성된 설명 텍스트, 테이블이(가) 표시된 사진

자동 생성된 설명

- 마산합포구, 마산회원구 비율이 높음 - 의창구 및 성산구가 온도가 높음

- 주거 단지 위치와 연관성 높음

* **미세먼지**  - **지역 면적당 유동인구량**

테이블이(가) 표시된 사진

자동 생성된 설명 텍스트이(가) 표시된 사진

자동 생성된 설명

- 산업단지와 주거 밀집지역 위치와 일치 - 성산구, 마산합포구 비율이 높음

- **지역 면적당 노약자비율**

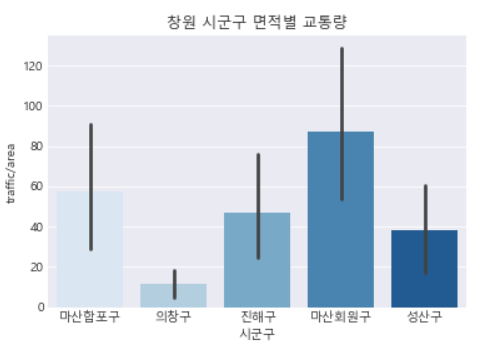
테이블이(가) 표시된 사진

자동 생성된 설명

- 구별로 비교적 균등하게 분포되어 있음

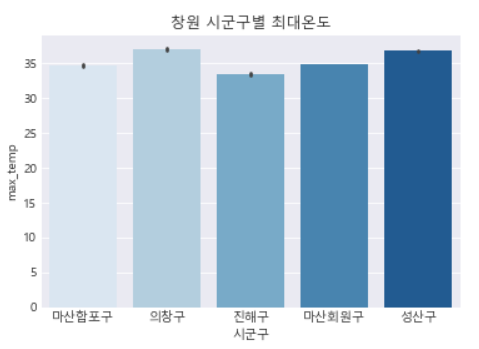
1. *시도별 데이터 시각화*

* **지역 면적당 교통량**

지도이(가) 표시된 사진

자동 생성된 설명

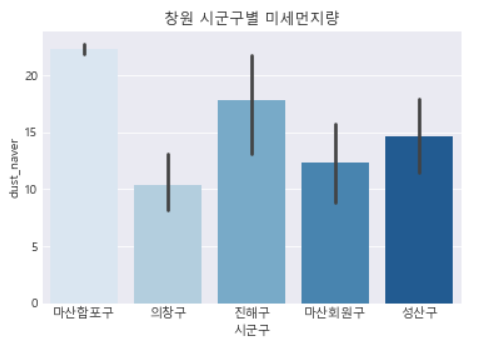
* 1. 마산회원구 ② 마산합포구 ③진해구 ④성산구 ⑤의창구
* **최대기온**

지도이(가) 표시된 사진

자동 생성된 설명

1. 의창구 ② 성산구 ③ 마산합포구 ④ 마산회원구 ⑤ 진해구

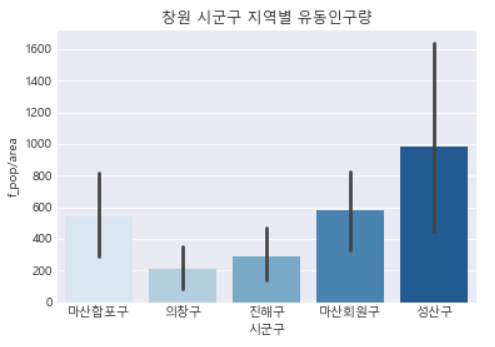
* **미세먼지**

지도이(가) 표시된 사진

자동 생성된 설명

1. 마산합포구 ② 진해구 ③ 성산구 ④ 마산회원구 ⑤ 의창구

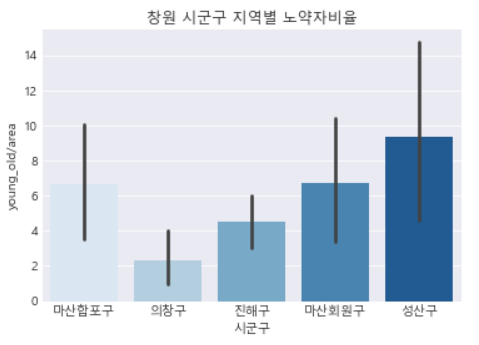
* **지역 면적당 유동인구량**

지도이(가) 표시된 사진

자동 생성된 설명

1. 성산구 ② 마산회원구 ③ 마산합포구 ④ 진해구 ⑤ 의창구

* **지역 면적당 노약자비율**

지도이(가) 표시된 사진

자동 생성된 설명

1. 성산구 ② 마산회원구 ③ 마산합포구 ④ 진해구 ⑤ 의창구