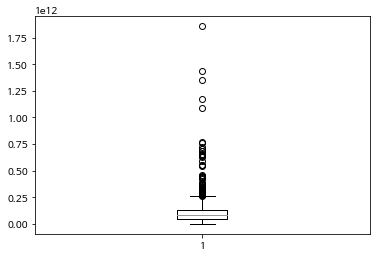
[0110] EDA 방법론

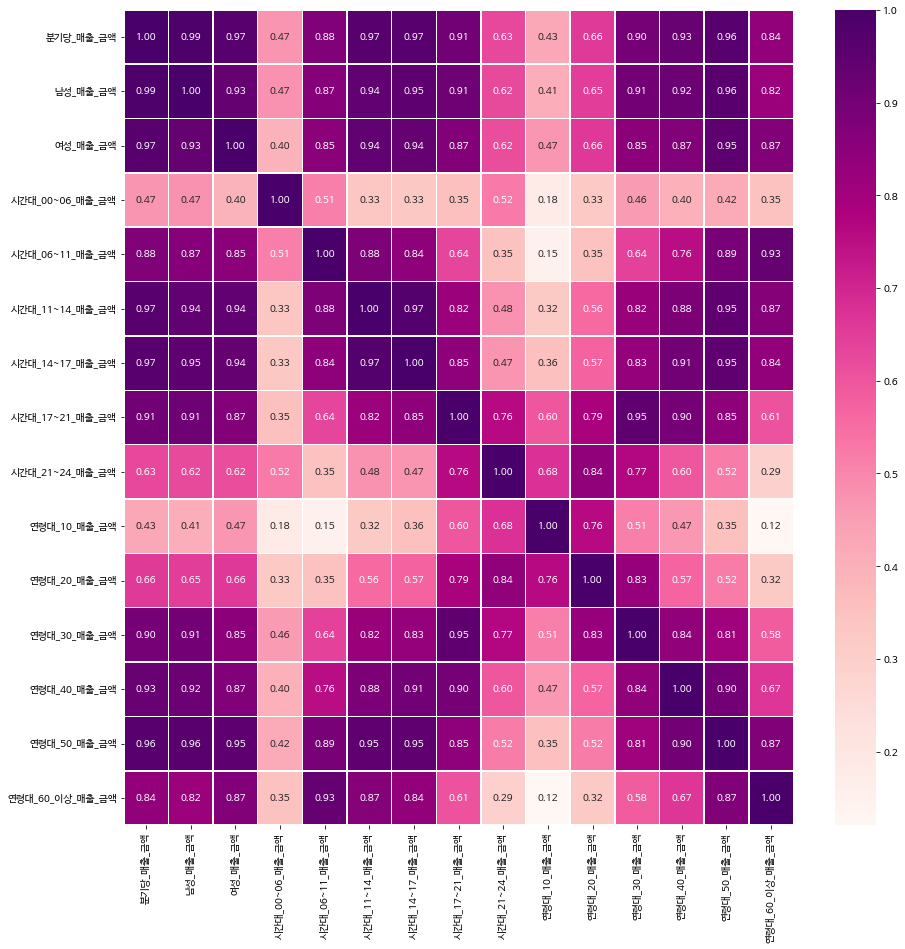
오늘 세션 내용을 바탕으로, 본인 팀 데이터에 맞는 EDA를 어떻게 진행할 지, 자유 양식의 형태로 서술하여 제출해주세요.

저희 조는 골목상권 데이터를 분석하고 있습니다.

‘상권배후지-추정매출’ 데이터에서 2022년 1,2,3분기 매출 합을 골목상권별로 계산했습니다.



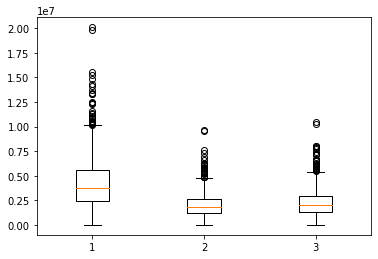
Boxplot을 통해 매출 금액을 나타냈습니다.



히트맵 분석을 통해 매출 금액끼리의 상관관계를 시각화해보았습니다.

그렇다면 매출에 영향을 주는 요인은 어떠한 것이 있는지 분석해보겠습니다.

우선 상권 근처의 생활인구가 영향을 줄 것 같아 ‘상권배후지-생활인구’ 데이터를 불러왔습니다.



순서대로 총 생활인구, 남성 생활인구, 여성 생활인구입니다. 총 생활인구에 비해 남성 및 여성 생활인구는 width가 좁은 것이 확인됩니다.

매출 금액과 merge한 이후 히트맵을 통하여 어떠한 요인이 영향을 많이 주는지 봤습니다.

텍스트, 하얀색이(가) 표시된 사진

자동 생성된 설명

유의미한 상관관계는 아니지만 그래도 연령대\_30\_생활인구\_수가 매출 금액에 가장 많은 영향을 준다는 것을 알 수 있었습니다.

앞으로 분석을 위해 ‘상권배후지-직장인구’ 데이터를 사용하여 매출 금액과 유의미한 상관 관계를 가지고 있는 요인을 찾아내어 이를 기반으로 한 소상공인 창업 컨설팅을 진행해볼 예정입니다.