

<2021 금융 데이터 경진대회>

KB 국민카드 지역별 음식점 세부업종집계내역 활용

음식점 업종 및 개업 위치 선정



팀 명 : 배고파

팀장 및 팀원 : 조영준, 길경주, 한유진

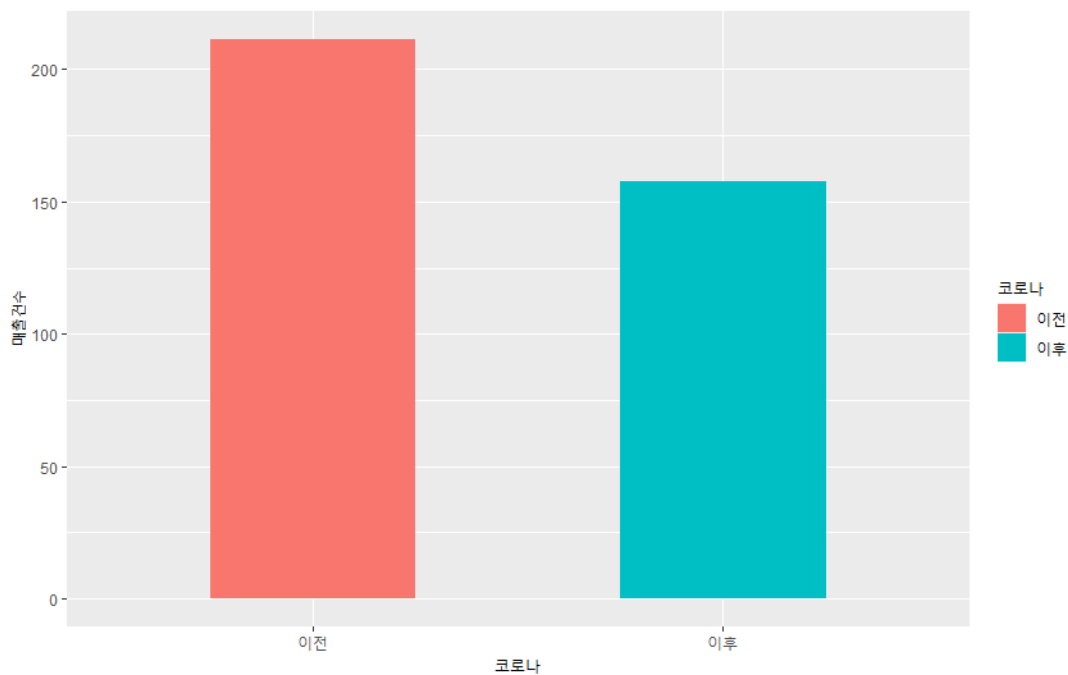
2021 금융 데이터 경진대회 결과 제출 양식

1. 주제

식당 매출 건수에 영향을 미치는 외부 환경 요소를 분석하여, 식당의 종류에 따라 개업하기 좋은 법정동에 대한 인사이트를 제공합니다.

2. 배경 및 필요성

코로나 이후 식당 매출의 기대값이 낮아진 시기에, 식당 개업을 준비하시는 분들에게 식당의 업종 및 위치 선정에 데이터 분석 결과로 도움을 줄 수 있습니다.



<실제로 매출이 줄어든 모습>

3. 아이디어 제안 및 분석 결과

- 1) 성별, 연령대, 유동인구 등의 변수가 매출 건수에 중요한 요인임을 확인하였습니다.
- 2) 식당에 자주 방문하는 성별, 연령대를 분석하였습니다.
- 3) 식당에 자주 방문하는 성별, 연령대가 주로 가는 식당을 분석하였습니다.

- 4) 해당 성별, 연령대의 유동인구가 높은 지역임에도 불구하고, 선호도가 높은 업종의 매출건수가 낮은 법정동을 파악하였습니다.
- 5) 유동인구 대비 매출건수가 낮은 법정동에 해당 업종의 식당을 개업하는 것을 추천합니다

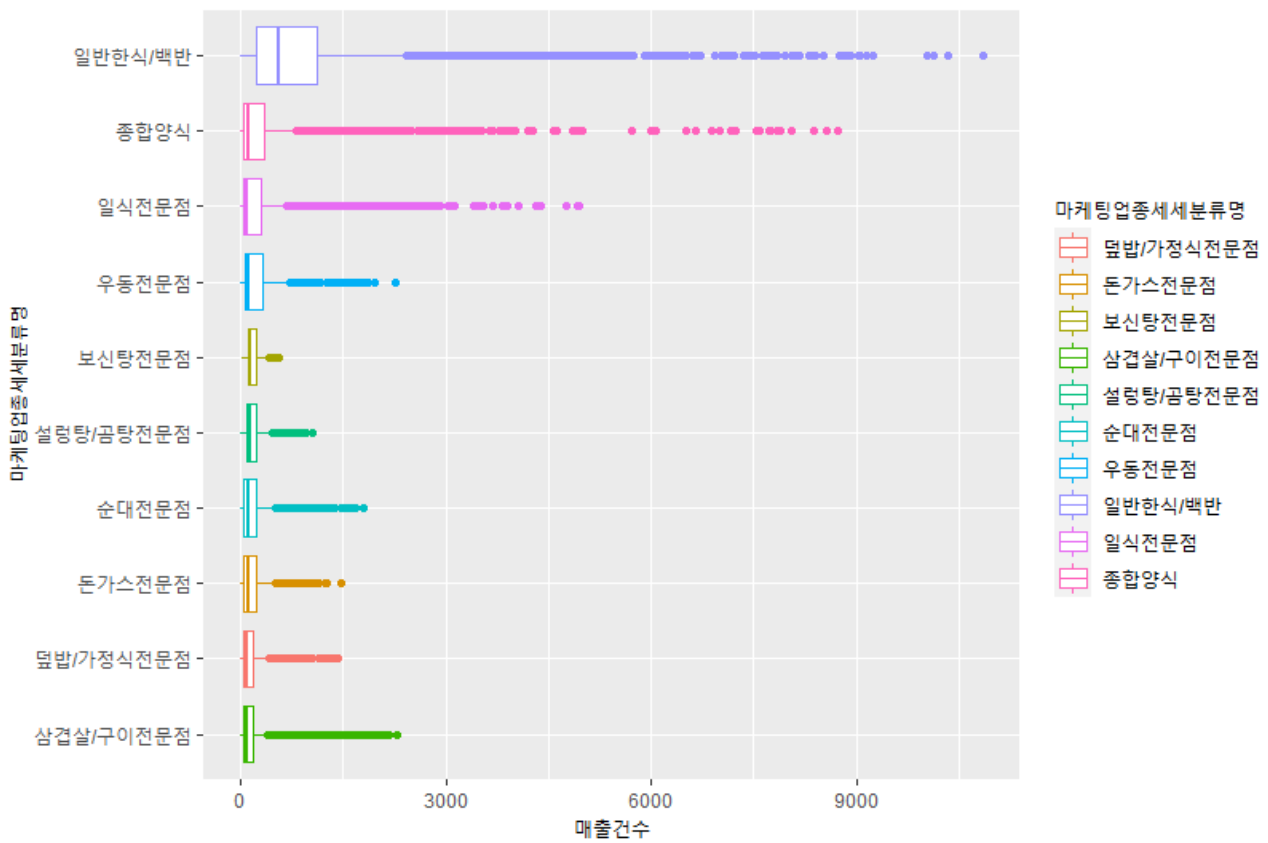
분석결과 : 서울

- 1) 매출건수에 영향을 많이 미치는 요인을 xgboost 를 사용하여 분석하였습니다.

```
> xgb.importance(model = xgbc)
      Feature      Gain      Cover      Frequency
1: 마케팅업종세분류명 4.378862e-01 0.3834397799 0.4112224449
2:      유동20대 1.850552e-01 0.0851386959 0.0705410822
3:      연령대 1.465415e-01 0.1468861105 0.1615230461
4:      유동30대 7.372330e-02 0.0308129722 0.0328657315
5:      역 2.842877e-02 0.0041396780 0.0144288577
6:      성별구분 2.602927e-02 0.0472157452 0.0737474950
7:      price 2.323104e-02 0.0049946917 0.0224448898
8:      주거30대 1.569083e-02 0.0154473352 0.0176352705
9:      주거20대 1.281532e-02 0.0224011259 0.0200400802
10:      주거10대 1.181925e-02 0.0129716154 0.0196392786
11:      학교 9.844133e-03 0.0385825847 0.0244488978
12:      유동60대 9.335102e-03 0.0287937182 0.0204408818
13:      세대수 5.248298e-03 0.0142906349 0.0144288577
14:      유동40대 4.350811e-03 0.0160073737 0.0184368737
15:      주거50대 3.839105e-03 0.0165175316 0.0136272545
16:      소비자물가지수 2.997057e-03 0.0912014297 0.0276553106
17:      주거40대 2.322443e-03 0.0094652225 0.0100200401
18:      유동50대 3.513405e-04 0.0179067320 0.0136272545
19:      미세먼지 3.328987e-04 0.0104721152 0.0096192385
20:      대학교 1.351310e-04 0.0027175829 0.0032064128
21:      초미세먼지 2.305414e-05 0.0005973247 0.0004008016
```

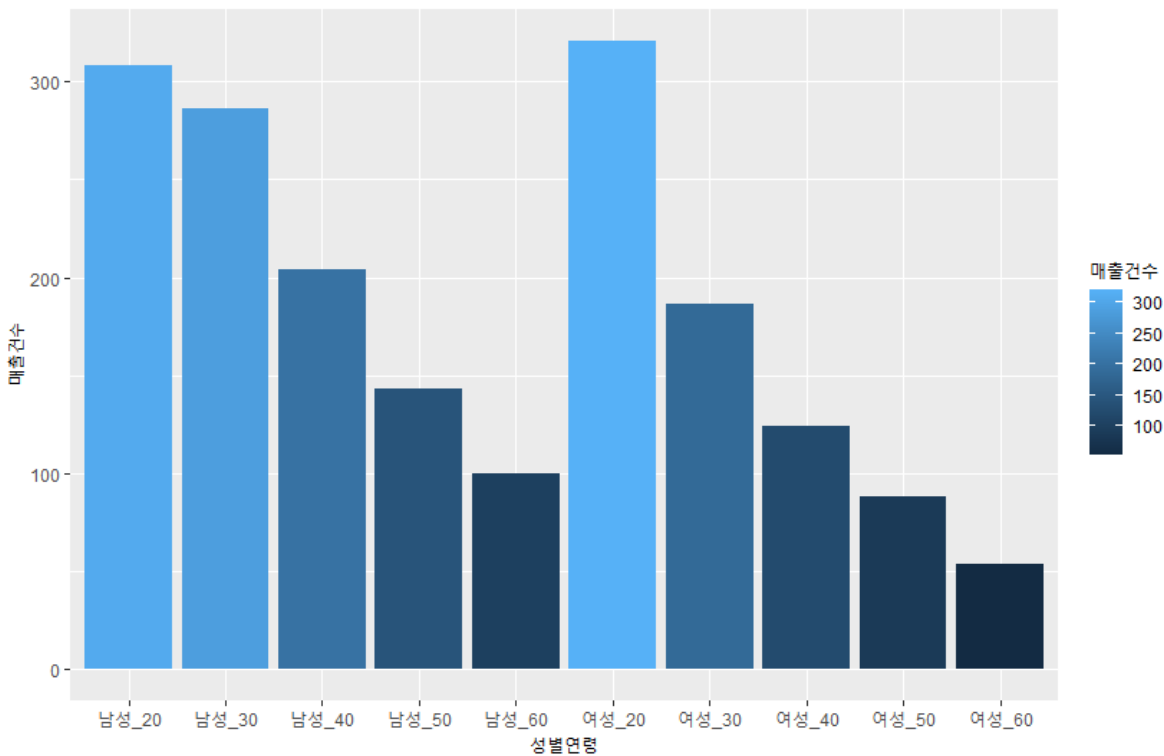
변수 중요도를 고려하였을 때, 식당의 업종과 연령대 특히 20 대 30 대가 중요한 것을 알 수 있었습니다.

- 2) 마포구 강남구 전체 지역에서 가장 매출건수가 높은 업종은 다음과 같습니다.



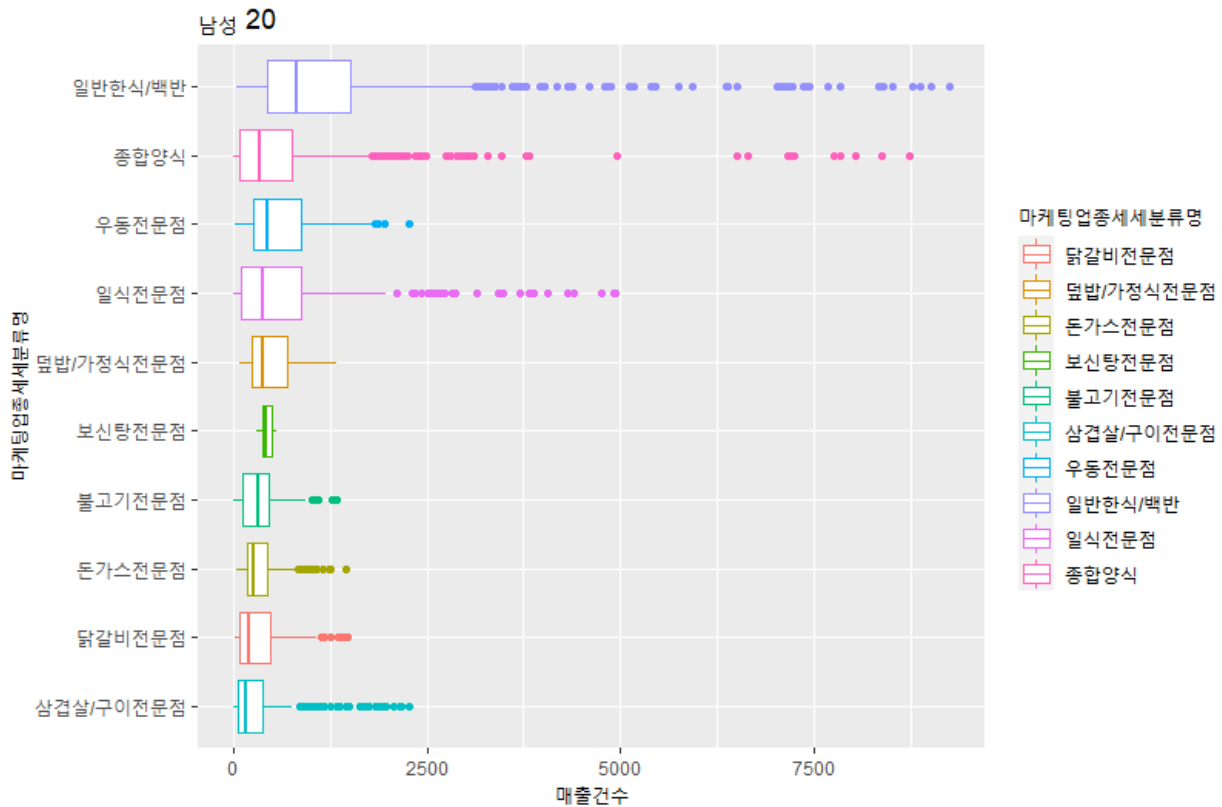
일반한식/백반, 종합양식, 일식전문점, 우동전문점 순으로 매출건수가 높음을 확인할 수 있습니다.

3) 성별과 연령대의 조합을 통해 가장 식당을 많이 방문하는 조합을 확인하였습니다.

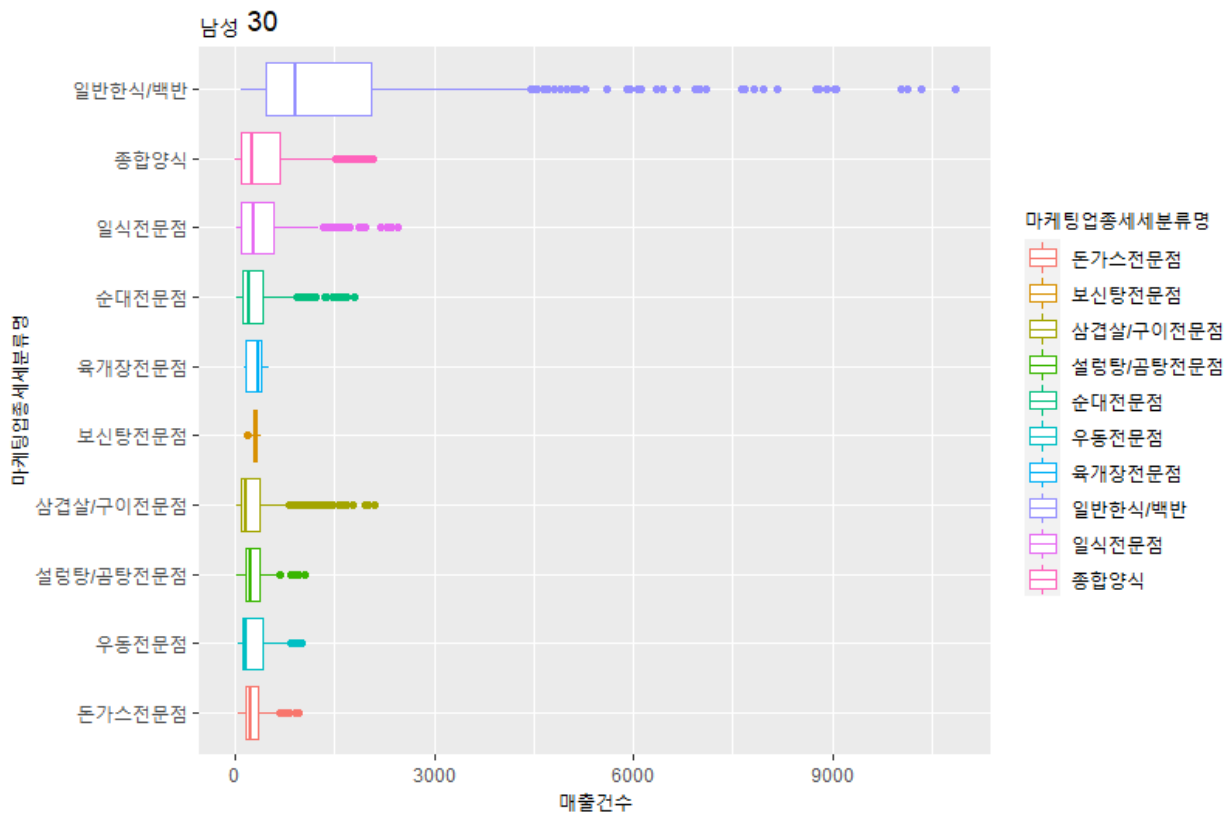


20 대남성, 30 대남성, 20 대여성의 매출 비중이 상당히 높은 것을 확인할 수 있었습니다.

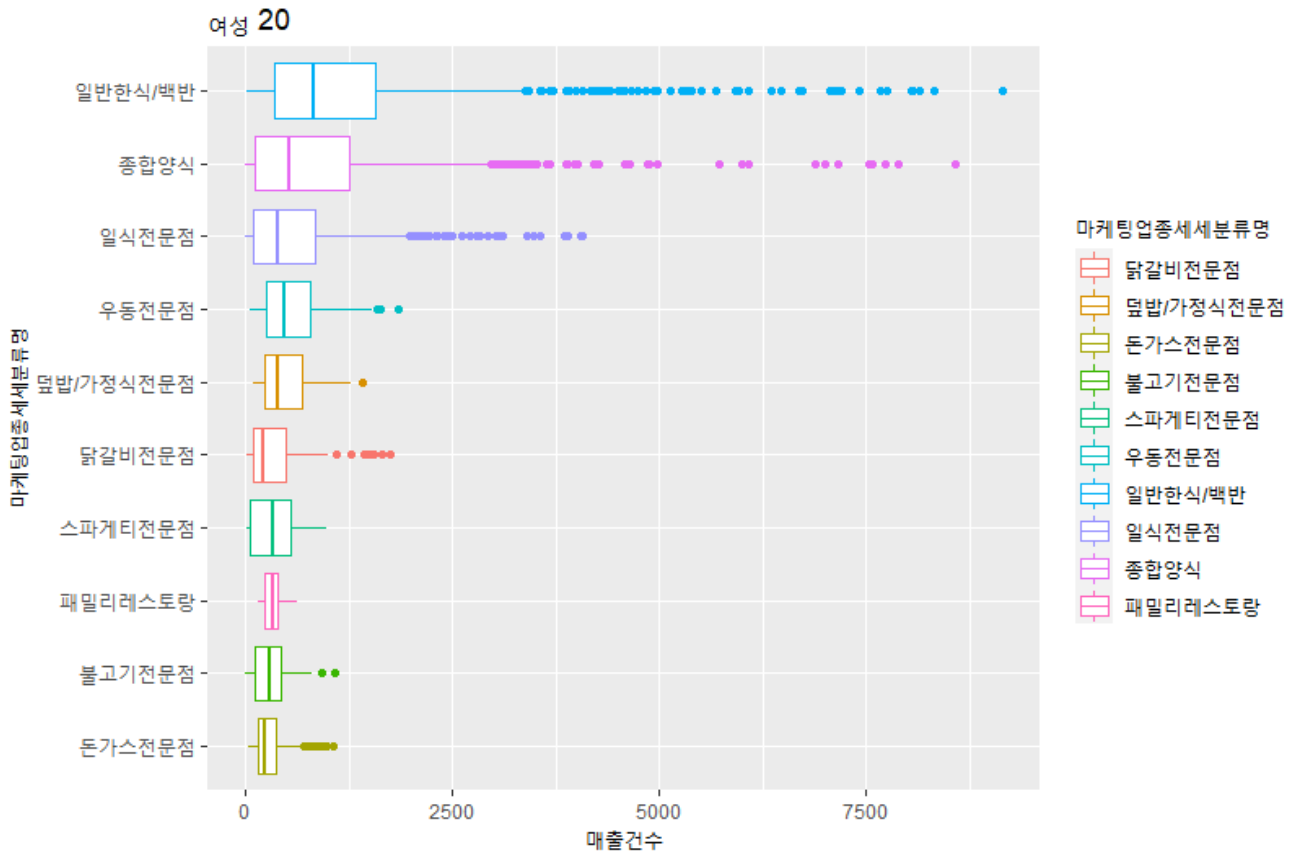
4) 20 대,30 대 남성과 20 대 여성이 자주 가는 식당 업종 분석



일반 한식/백반, 종합양식, 우동전문점 순으로 선호도가 높은 것을 확인할 수 있습니다.



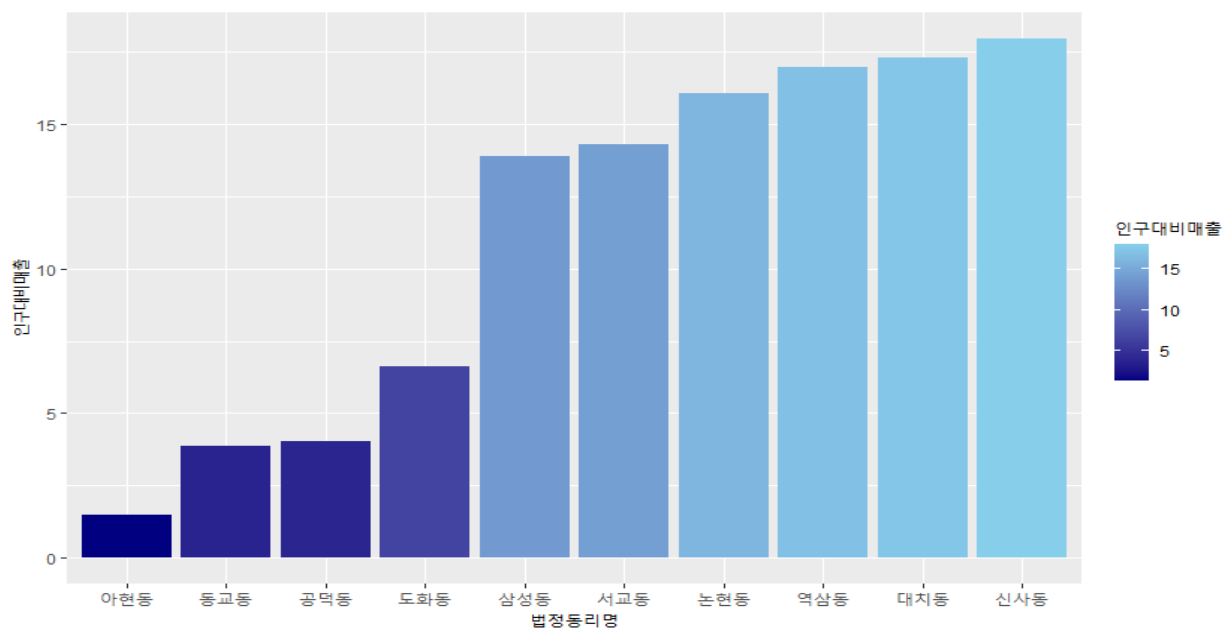
일반 한식/백반, 종합양식, 일식전문점 순으로 선호도가 높은 것을 확인할 수 있습니다.



20 대 여성의 경우에도, 일반 한식/백반 식당의 매출건수가 가장 높음을 확인할 수 있습니다.

일반 한식/백반, 종합 양식, 일식전문점의 선호도가 높은 것을 확인할 수 있습니다.

5) 20 대 30 대의 유동인구가 평균보다 높은 지역 중에서, 유동인구 대비 매출이 적은 지역들을 확인하였습니다.



분석결과 : 분당구

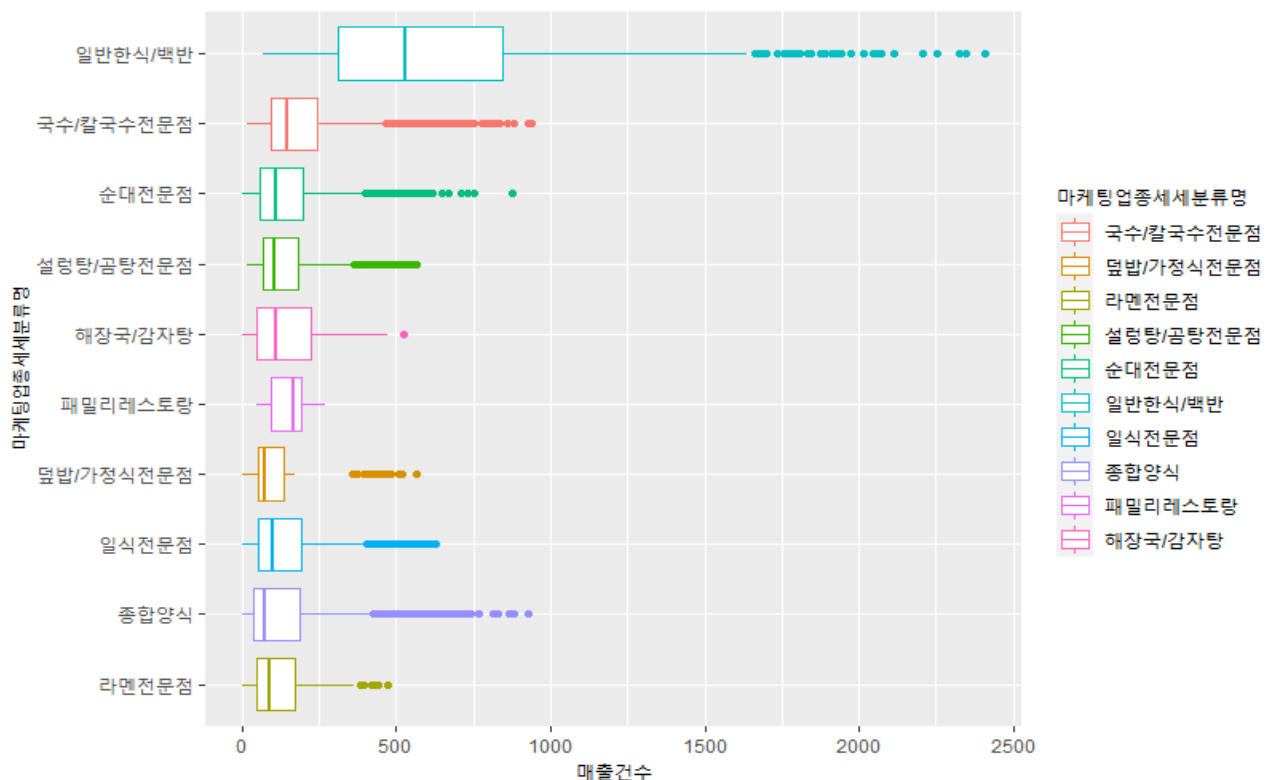
1) 매출건수에 영향을 많이 미치는 요인을 xgboost 를 사용하여 분석하였습니다.

> xgb.importance(model = xgbc)

	Feature	Gain	Cover	Frequency
1:	마케팅업종세분류명	0.6280929975	0.481911342	0.372860636
2:	학교	0.0702455590	0.014508853	0.031377343
3:	성별구분	0.0678715362	0.066708328	0.059494703
4:	연령대	0.0611830464	0.104899870	0.179299104
5:	주거총인구	0.0584114147	0.045853713	0.041564792
6:	주거50대	0.0318877774	0.060405930	0.055419723
7:	주거60대이상	0.0245223506	0.031812260	0.014669927
8:	주거10대	0.0208436447	0.015169401	0.062347188
9:	소비자물가지수	0.0123737508	0.067491000	0.040749796
10:	주거20대	0.0113876890	0.030983645	0.028932355
11:	price	0.0073908649	0.023860601	0.039119804
12:	지하철역하차	0.0023504612	0.018712639	0.021189894
13:	주거30대	0.0015156516	0.002359971	0.013854931
14:	미세먼지	0.0010850097	0.020239541	0.021189894
15:	역	0.0003718628	0.002651100	0.002852486
16:	주거40대	0.0002655880	0.001264371	0.005297474
17:	초미세먼지	0.0002007956	0.011167436	0.009779951

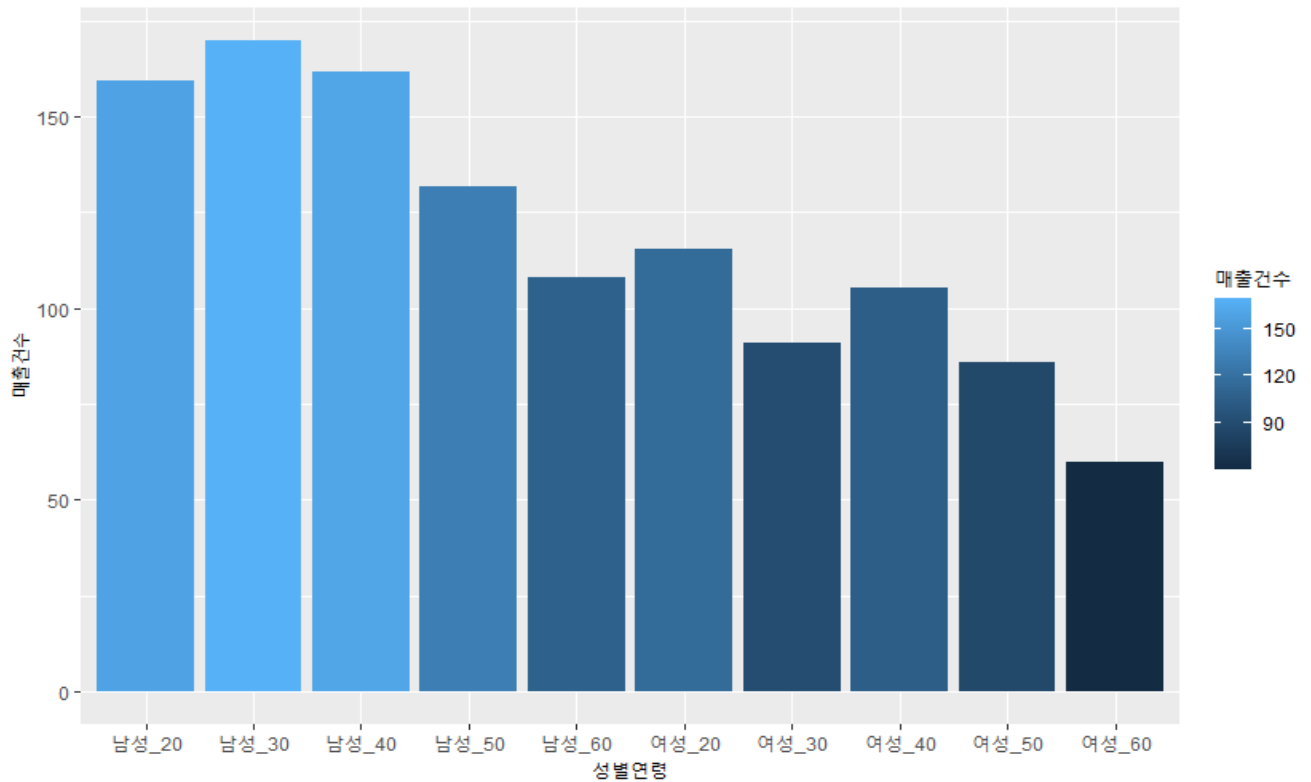
분당구의 유동인구를 반영하지 못하여, 식당 업종만이 가장 중요한 변수임을 확인하였습니다.

2) 분당구 전체 지역에서 가장 매출건수가 높은 업종은 다음과 같습니다.



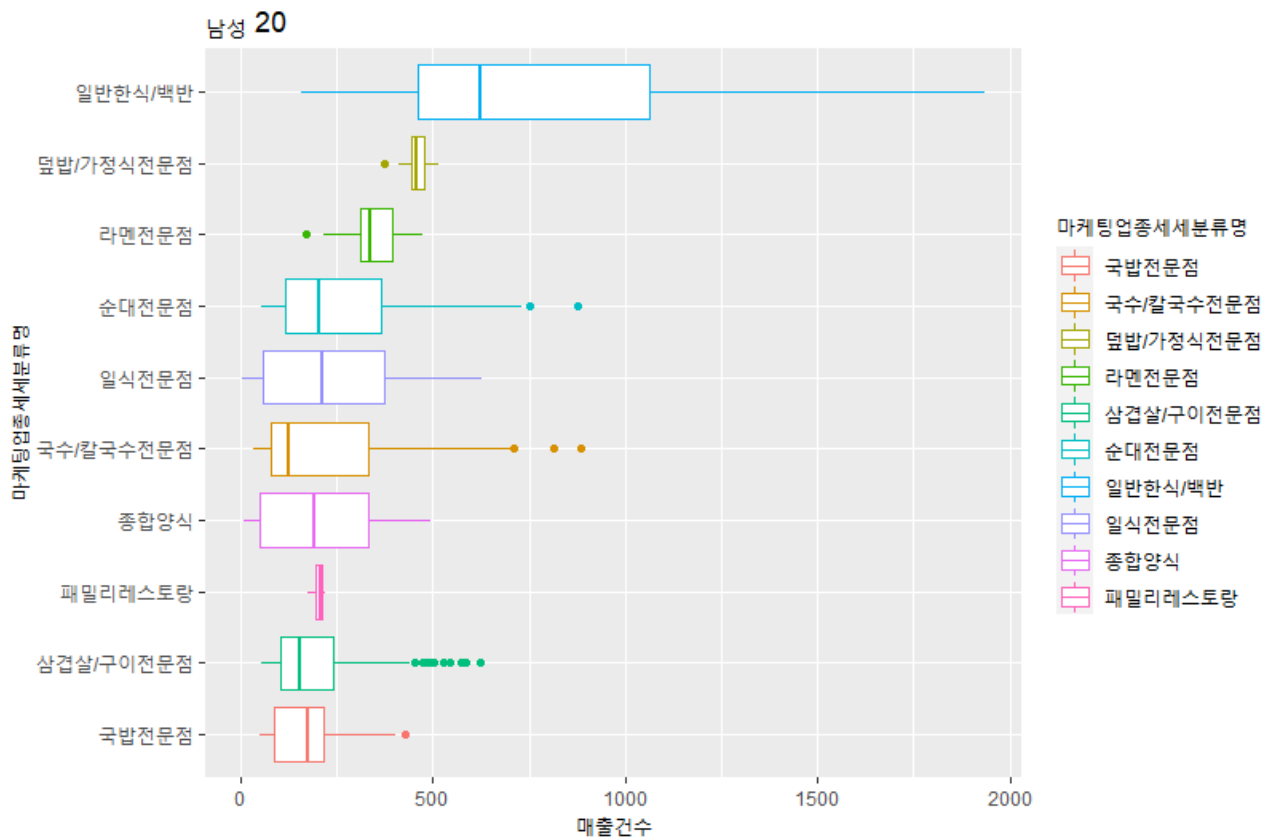
일반한식/백반, 국수/칼국수전문점, 순대전문점 순으로 매출건수가 높음을 확인할 수 있습니다.

3) 성별과 연령대의 조합을 통해 가장 식당을 많이 방문하는 조합을 확인하였습니다.

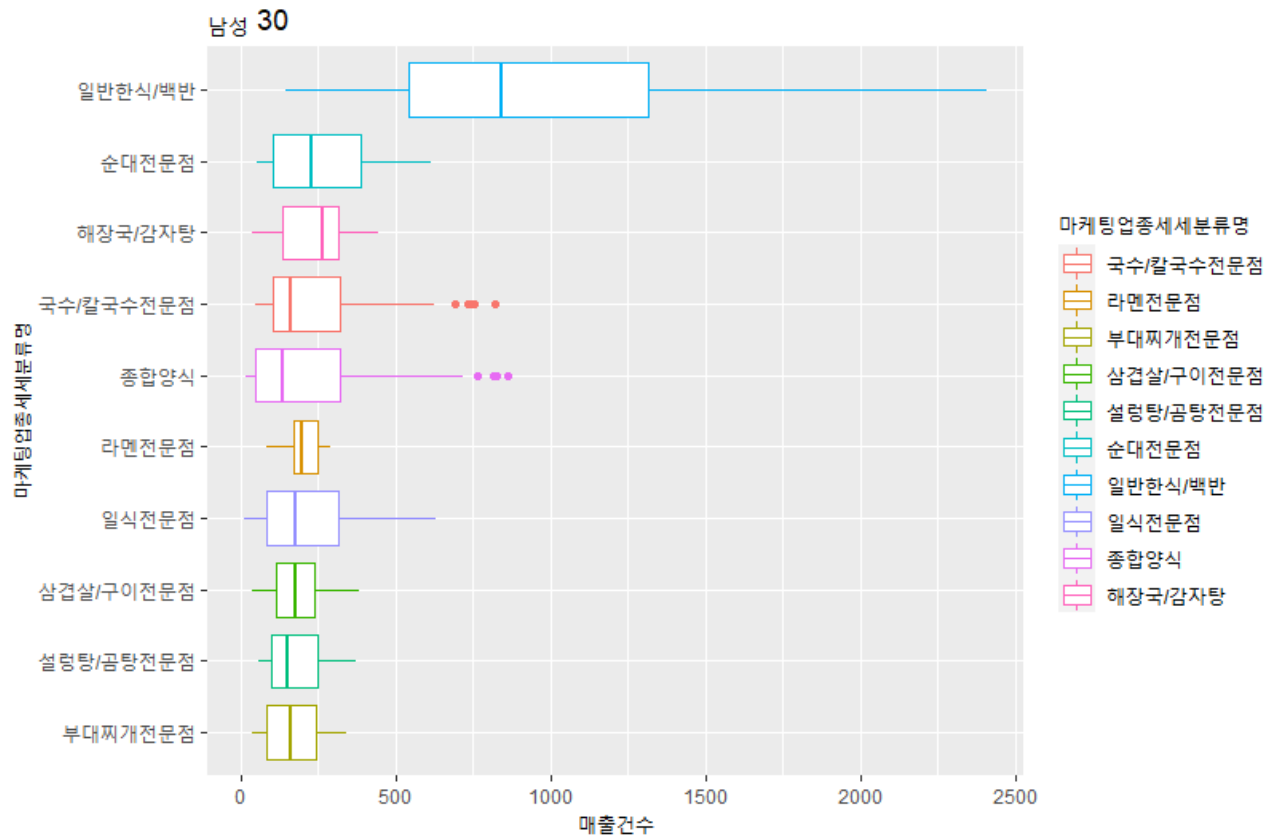


남성 손님이 여성 손님에 비해 전체적으로 많음을 확인할 수 있습니다.

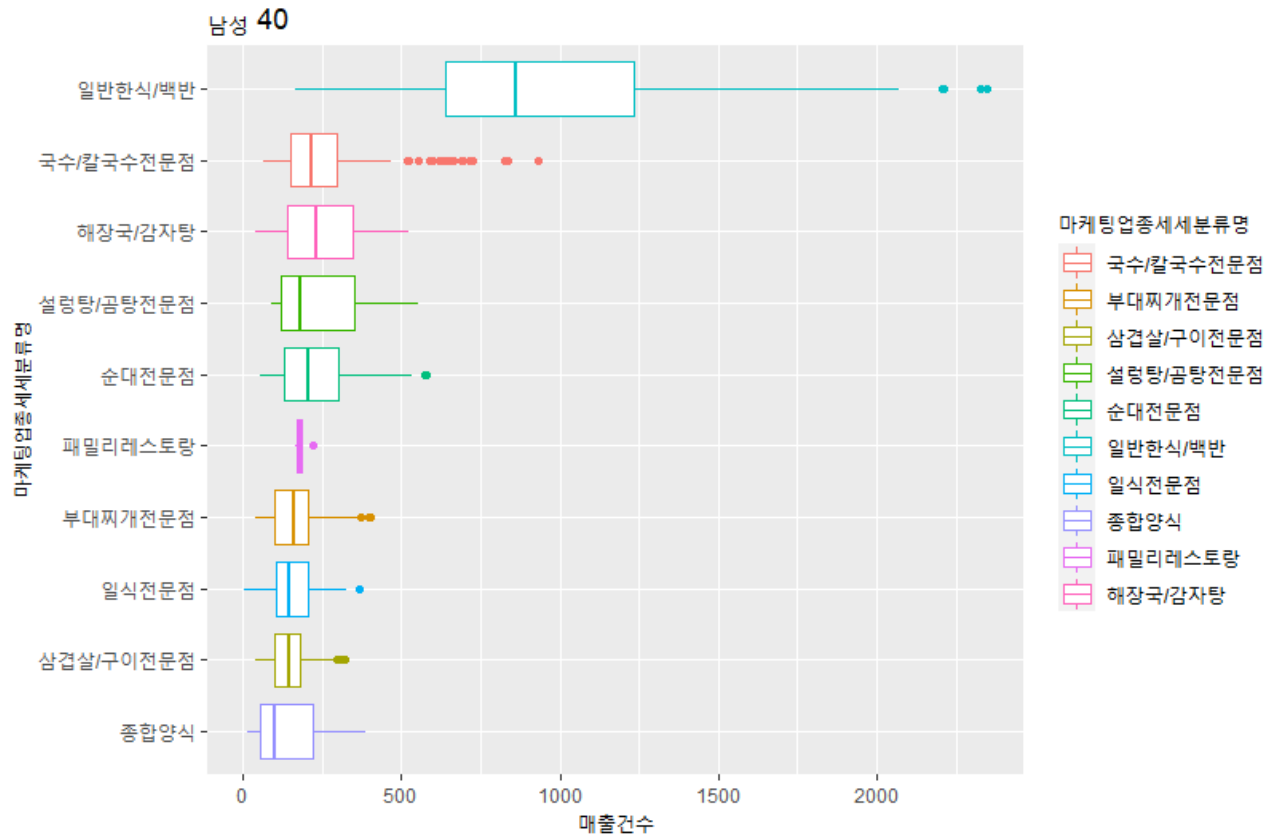
4) 20 대,30 대,40 대 남성이 자주 가는 식당 업종 분석



일반 한식/백반, 덮밥/가정식전문점, 라면전문점 순으로 선호도가 높은 것을 확인할 수 있습니다.



일반 한식/백반, 순대전문점, 해장국/감자탕 순으로 선호도가 높은 것을 확인할 수 있습니다.



일반 한식/백반, 국수/칼국수전문점, 해장국/감자탕 순으로 선호도가 높은 것을 확인할 수 있습니다.

분석결과 : 한계점

경기도의 유동인구 데이터를 반영하지 못하여, 추가적인 분석이 이루어지지 못했습니다.

행정동을 기준으로 한 유동인구/주거인구를 법정동으로 바꾸는 과정에서 지역 간 약간의 오차가 발생하였습니다.

4. 기대효과

식당 개업을 앞둔 자영업자는 해당 카드사의 데이터 분석 결과를 바탕으로 식당 업종 및 위치 선정에 대한 자문을 받고, 식당 개업 시 해당 카드사에 카드로 결제할 경우 할인 혜택을 제공하는 방식으로 상호 간 이윤추구 효과를 기대할 수 있습니다. 또한 KB pay 와 같은 간편결제 앱에서 제공하는 스마트 오더를 이용해 식당을 예약하면, 가격을 할인해 주는 방식으로 식당 홍보를 통한 매출증대를 도모하고, 동시에 카드사는 카드 수수료 수입 증가를 기대할 수 있습니다.

5. 활용 데이터

변수명	내용	데이터 타입	사용 여부	데이터 출처
기준년도	날짜	int		금융데이터거래소
기준년월	날짜	int		금융데이터거래소
한글시도명	지역	chr		금융데이터거래소
한글시군구명	지역	chr		금융데이터거래소
법정동리명	지역	chr		금융데이터거래소
마케팅업종중분류명	업종분류	chr	삭제	금융데이터거래소
마케팅업종세분류명	업종분류	chr	삭제	금융데이터거래소
마케팅업종세세분류명	업종분류	chr		금융데이터거래소
성별구분	성별	chr		금융데이터거래소
연령대	나이	chr		금융데이터거래소
신용카드보유여부	신용카드보유여부	int	합침	금융데이터거래소
체크카드보유여부	체크카드보유여부	int	합침	금융데이터거래소
회원수	카드보유여부에 따른 회원수	int	삭제	금융데이터거래소
매출건수	매출건수	int		금융데이터거래소

순위	매출순위	int	삭제	금융데이터거래소
소비자물가지수	소비자물가지수	num		http://data.seoul.go.kr/dataList/datasetList.do
미세먼지	미세먼지 수치	num		경기 데이터 드림
초미세먼지	초미세먼지 수치	num		경기 데이터 드림
price	면적당 집값	num		http://buking.kr/r/md
지하철역하차	지하철역 하차인구	num		서울 열린데이터 광장
학교	초,중,고등학교 수	int		서울 열린데이터 광장, 경기데이터드림
역	지하철역 수	int		서울 열린데이터 광장
대학교	대학교 수	int		서울 열린데이터 광장
세대수	세대수	int		서울 열린데이터 광장, 경기데이터드림
유동10대	10~19세 유동인구	num		서울 열린데이터 광장
유동20대	20~29세 유동인구	num		서울 열린데이터 광장
유동30대	30~39세 유동인구	num		서울 열린데이터 광장
유동40대	40~49세 유동인구	num		서울 열린데이터 광장
유동50대	50~59세 유동인구	num		서울 열린데이터 광장
유동60대	60세 이상 유동인구	num		서울 열린데이터 광장
유동총인구수	유동 총 인구수	num		서울 열린데이터 광장

주거10대	10~19세 주거인구	int		서울 열린데이터 광장
주거20대	20~29세 주거인구	int		서울 열린데이터 광장
주거30대	30~39세 주거인구	int		서울 열린데이터 광장
주거40대	40~49세 주거인구	int		서울 열린데이터 광장
주거50대	50~59세 주거인구	int		서울 열린데이터 광장
주거60대이상	60세 이상 주거인구	int		서울 열린데이터 광장
주거총인구	주거 총 인구수	int		서울 열린데이터 광장

※ 작성 시 유의사항

- 분량 제한 없음 /글자 폰트 크기 11 포인트(한글 및 워드로 작성)
- 도표, 이미지 등 활용 가능
- 설명을 위한 추가자료 첨부 가능
- 제출 시 표지를 함께 제출하되, 식별이 가능하도록 참가자 명(팀의 경우 팀장 및 팀원명)을 작성하여 제출

2021 년 9 월 6 일

참가자(대표자) (인 또는 서명)

금융보안원 귀중