

카드 소비 데이터 B조

부산 지역화폐 군구별 분석 및 활성화 방안

10기 2019113421 최은진

11기 2020110478 이다현

12기 2019110467 문성원

목차

1. 분석 목적
2. 데이터 소개
3. 데이터 정제
4. 구 별 EDA
5. Clustering
6. 결론

1. 분석 목적

1. 분석목적

부산광역시 지역화폐 '동백전' 데이터를 통한 활성화 방안

2. 테이터 소개

2. 데이터 소개

6,7,8월 행정동 단위 업종별 지역화폐 거래 데이터를 합하여 하나의 Dataframe 생성

strd_date	strd_tizn_val	ctpr_nm	ccg_nm	adn_dgnm	adng_no	cardc_nm	tpbuz_cd	tpbuz_nm	bth_yr	sex_cd	trns_cnt	ttn_trns_amt
20220601	0	부산광역시	중구	대청동	21010530	코나	4010	편의점	1973	M	1	7350
20220601	0	부산광역시	중구	중앙동	21010510	부산	4010	편의점	1981	F	1	3700
...
20220801	0	부산광역시	영도구	동삼1동	21040650	하나	4010	편의점	1971	F	2	7600
20220801	0	부산광역시	영도구	동삼1동	21040650	하나	4010	편의점	2001	M	1	2200

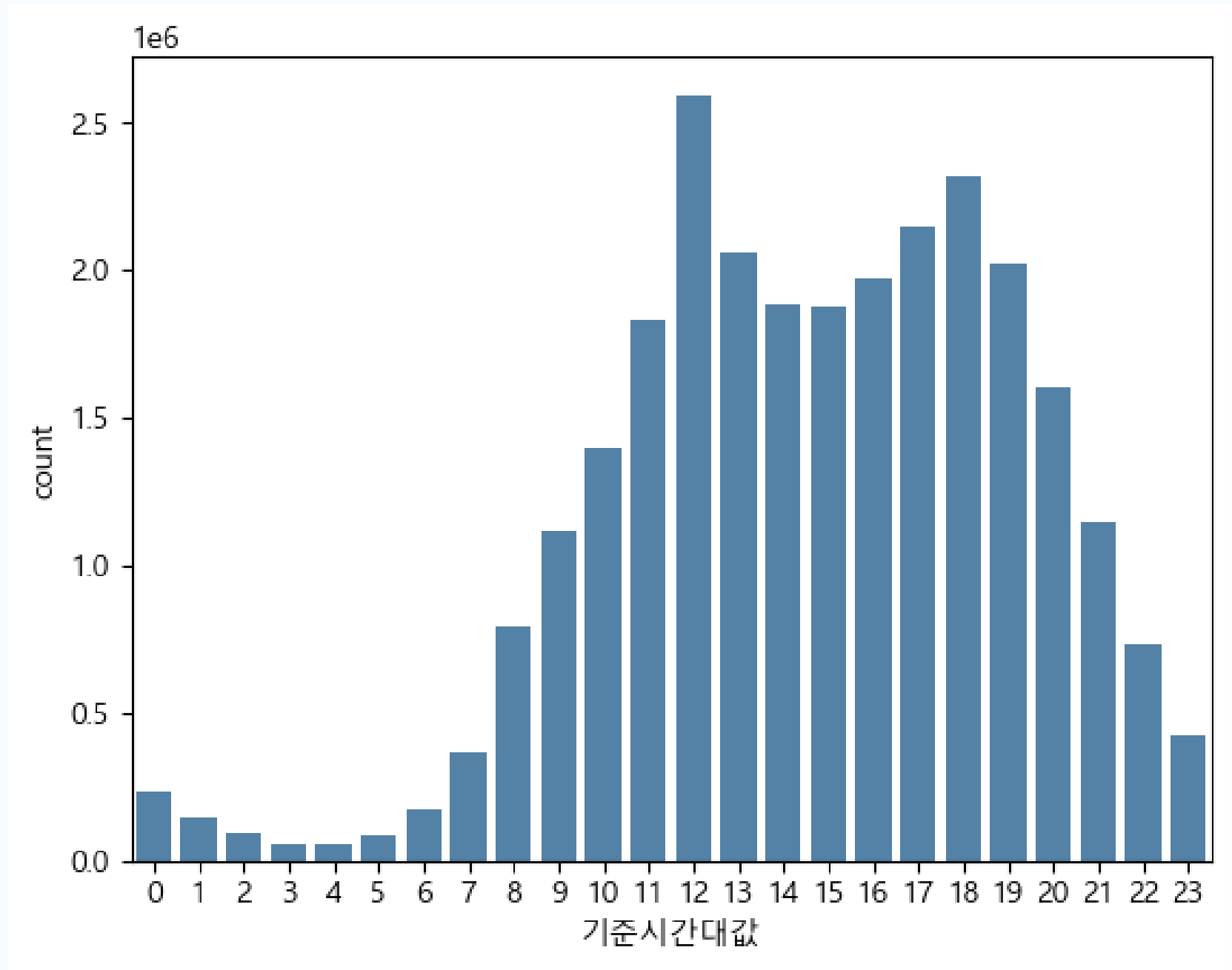
2. 데이터 소개

변수명

- strd_date : 기준일자
- strd_tizn_val : 기준시간대값
- ctp_r_nm : 시도명
- ccg_nm : 시군구명
- adn_dgnm : 행정동명
- adng_no : 행정동번호
- cardc_nm : 카드사명
- tpbuz_cd : 업종코드
- tpbuz_nm : 업종명
- bth_yr : 출생년도
- sex_cd : 성별코드
- trns_cnt : 거래건수
- ttl_trns_amt : 총거래금액

2. 데이터 소개

기준시간대값

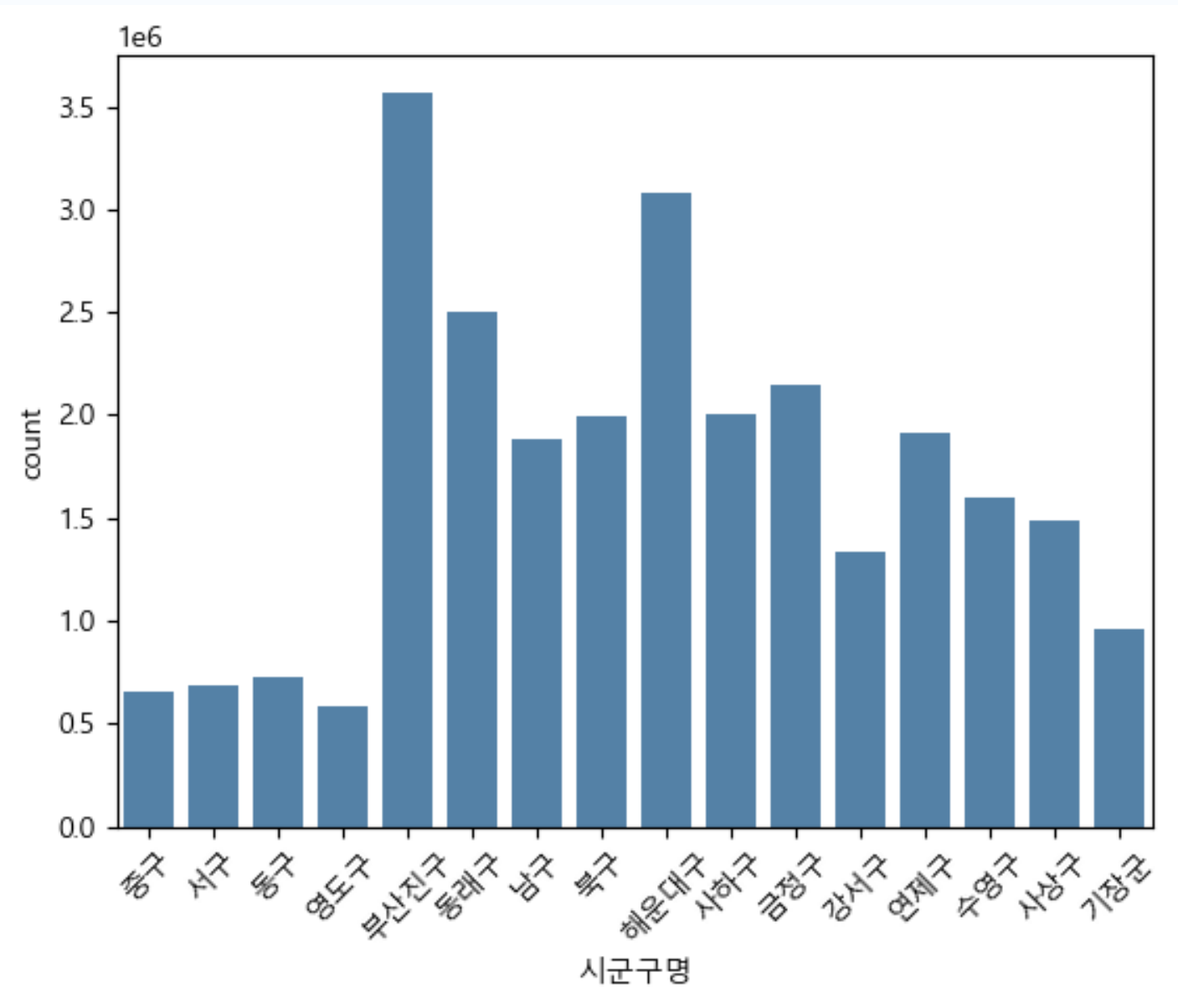


12	2594249
18	2317544
17	2149852
13	2064385
19	2022080
16	1973444
14	1884563
15	1879940
...	
7	365718
0	236182
6	176698
1	141841
2	89931
5	84899
3	59274
4	52672

→ 12시가 가장 높다

2. 데이터 소개

시군구명



강서구, 금정구, 기장군, 남구, 동구, 동래구,
부산진구, 북구, 사상구, 사하구, 서구,
수영구, 연제구, 영도구, 중구, 해운대구

총 16 개 구

2. 데이터 소개

행정동명
(205개)

대청동, 중앙동, 동광동, 서대신3동, 서대신4동, '부민동, 아미동, 초장동,

...

명장1동, 명장2동, 서1동, 삼락동, 신선동, 가락동, 일광읍, 금성동

행정동번호
(205개)

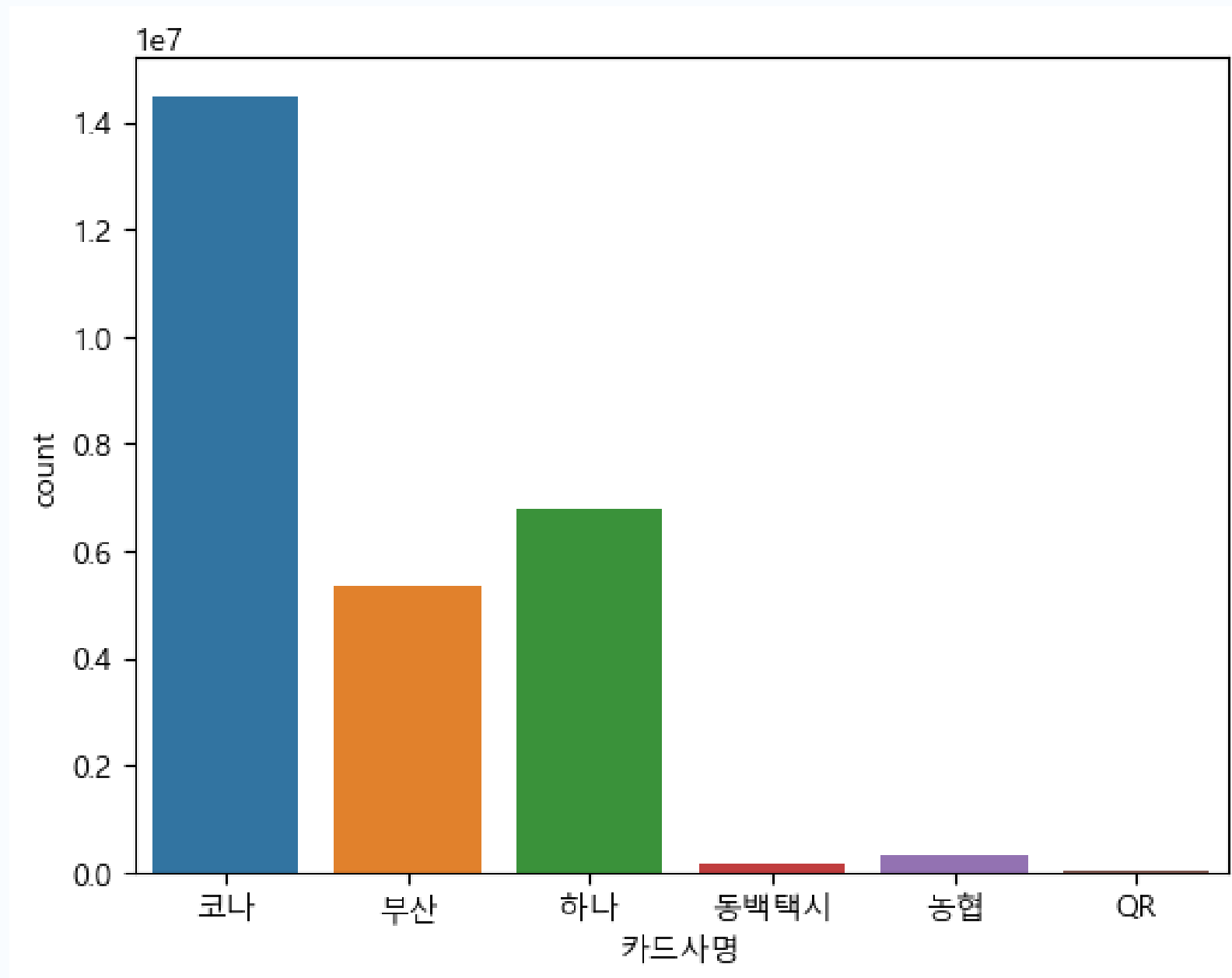
21010530, 21010510, 21010520, 21020560, 21020570, 21020590, 21020610,

...

21060640, 21110510, 21150510, 21040680, 21120550, 21510111, 21110710

2. 데이터 소개

카드사명



'코나', '부산', '하나', '동백택시', '농협', 'QR'

- 2022년 4월부터 동백전 지역화폐로 은행 부산은행 실시
- 2021년 4월 ~ 2022년 3월까지 동백전 지역화폐로 코나아이 사용

2. 데이터 소개

업종명
(210개)

편의점, 사무서비스, 골프연습장, 일반한식,
슈퍼마켓, 택시, 연쇄점, 농축수산물,

...

'정수기, 유리, 가례서비스업, 카메라, 초중고
교육기관, 기타광학제품, 렌터카

업종코드
(210개)

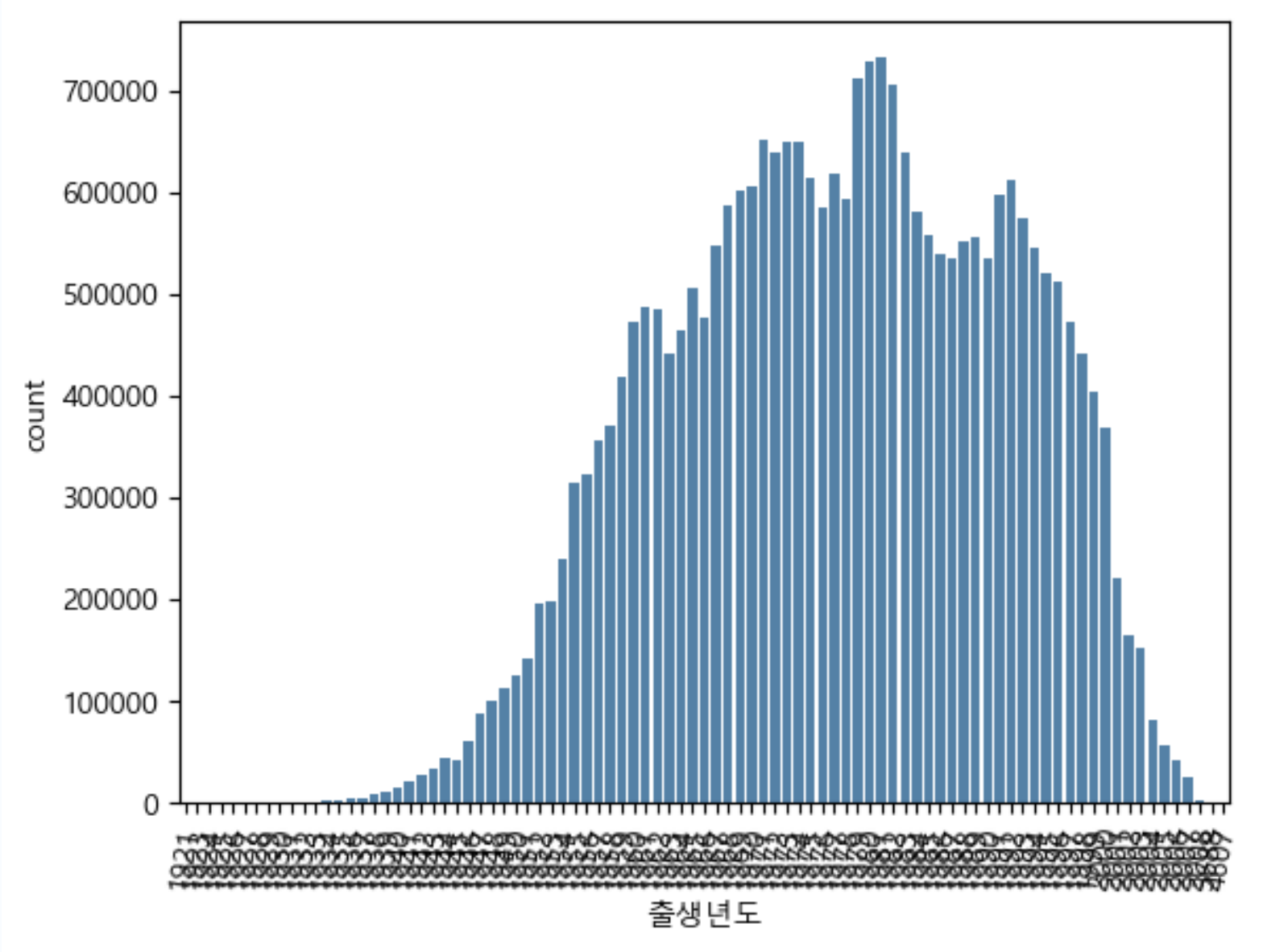
4010, 9130, 2102, 8001, 4020, 1123, 4021,
8310, 8320, 8021, 8399, 2114, 8006, 2130,

...

3210, 9005, 9135, 9110, 3401, 5121, 3499,
1130, 2201, 4199, 9911, 4481, 9397, 9396

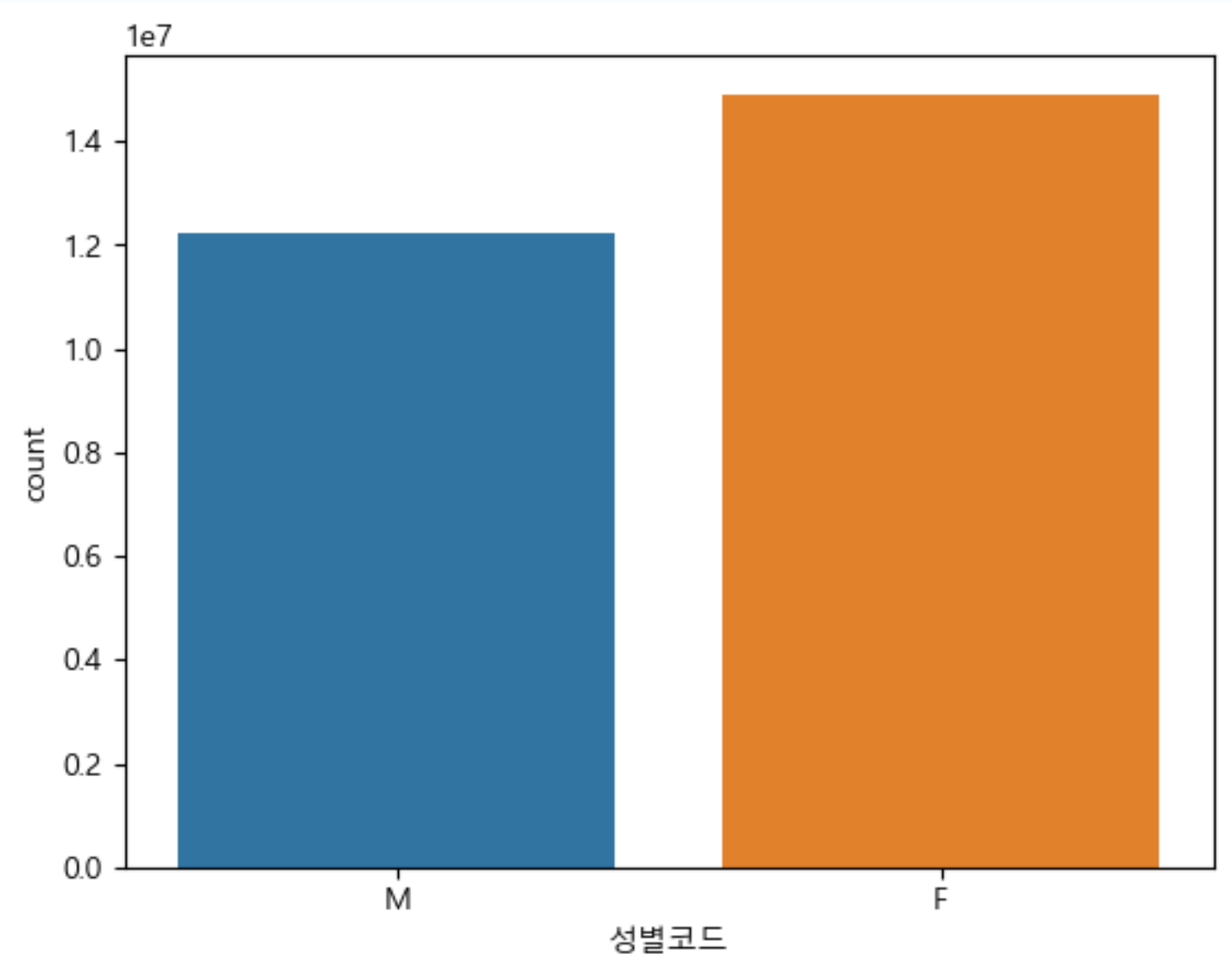
2. 데이터 소개

출생년도



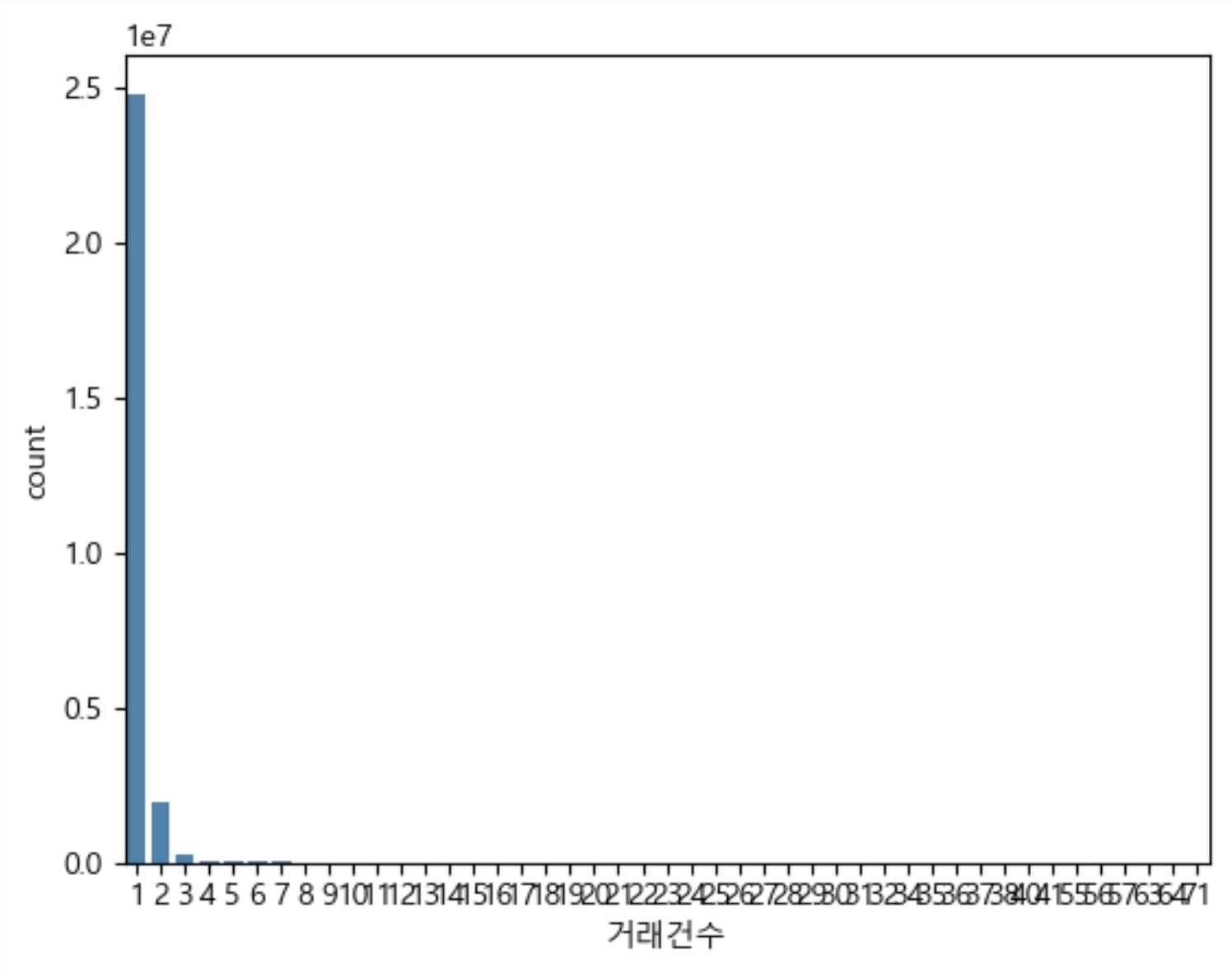
범위: 1921~4607

성별코드



2. 데이터 소개

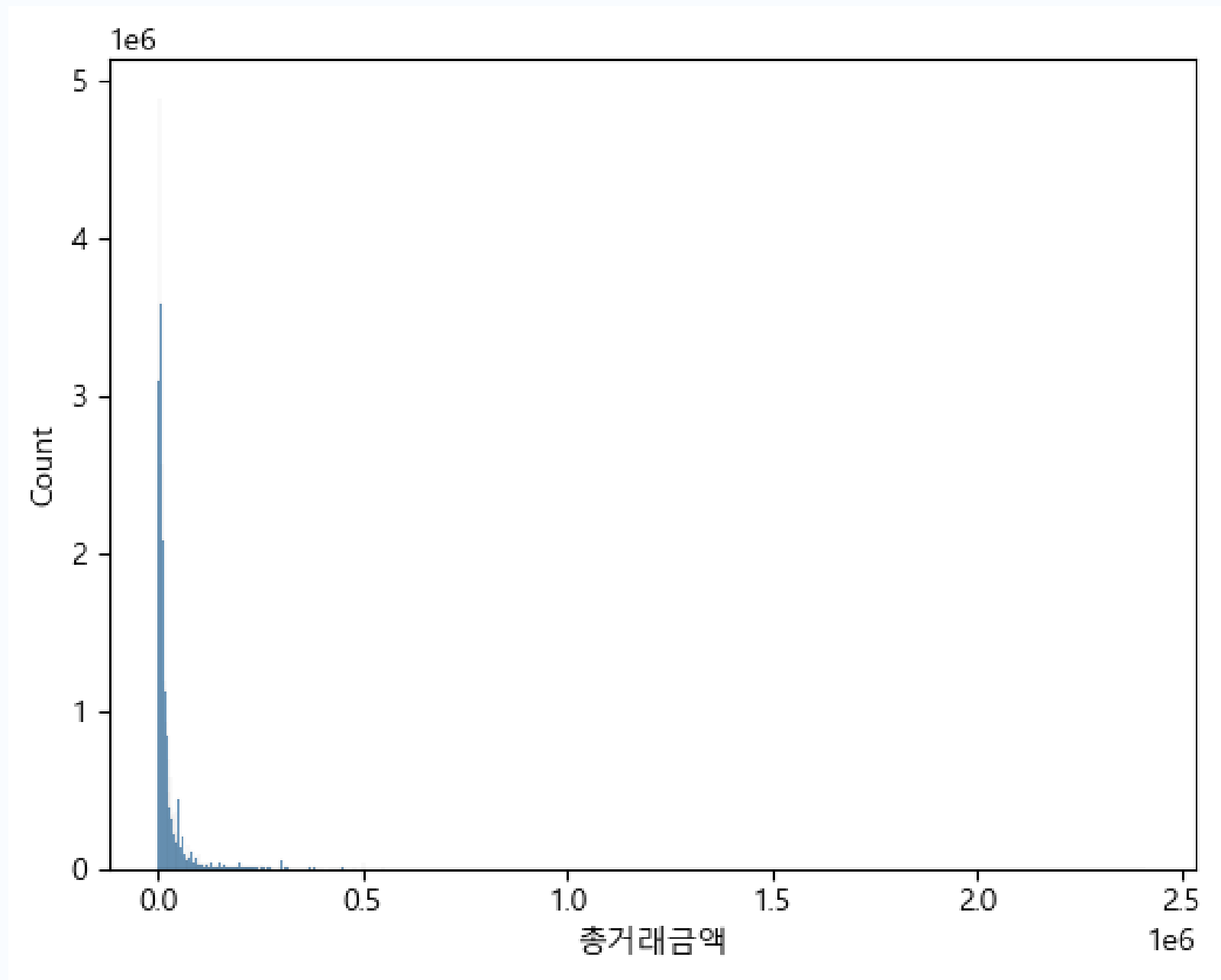
거래건수



1	24823825
2	1964625
3	264914
4	55319
5	15612
...	
71	1
56	1
37	1
55	1
63	1

2. 데이터 소개

총 거래금액



총 거래금액 사용 범위
1 ~ 2413460

3. 테이터 정제

3. 데이터 정제

- 결측치, 중복행 없음
- 변수 제거
 - 데이터 전체가 부산광역시 이므로 '시도명' 변수 제거
 - '행정동명', '행정동번호' 모두 제거
(행정동 총 205개로 너무 많아 파악 어려워 구별로 나눠 파악하기로함)
 - '출생년도' 이상값 제거 4607년, 3688년

3. 데이터 정제

- 변수 생성

- '기준일자' 변수 이용하여 '요일' 변수 생성
(Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday)
 - 월별 파악 위해 '기준일자'에서 월을 추출하여 새로운 '월' 변수 추가
 - '출생년도' 변수 이용하여 'age' 변수 생성
(17 ~ 102)
 - 'age' 변수 이용하여 'age_range' 변수 생성
(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)
- => 'age_range' 생성하여 'age', '출생년도' 변수 제거

3. 데이터 정제

- data

기준일자	기준 시간대값	시군구명	카드사명	업종코드	업종명	성별 코드	거래건수	총 거래금액	요일	월	age_range
2022-06-01	0	중구	코나	4010	편의점	M	1	7350	Wednesday	6	50
2022-06-01	0	중구	부산	4010	편의점	F	1	3700	Wednesday	6	40
...
2022-08-01	0	영도구	코나	8399	기타음료식품	F	1	7500	Monday	8	20
2022-08-01	0	영도구	하나	4010	편의점	F	2	7600	Monday	8	50
2022-08-01	0	영도구	하나	4010	편의점	M	1	2200	Monday	8	20

3. 데이터 정제

- 변수 생성

- 한 고객당 1건이 아닌 해당 일자, 시간대와 일치하는 모든 건수를 합한 결과 따라서 총 거래금액을 거래 건수로 나누어줌
'건별거래금액' 변수 생성
- 건별 거래금액이 10원 이하인 것은 있을 수 없는 거래라고 생각하여
총 111행 제거

3. 데이터 정제

- 변수 생성

- '업종코드_중분류.csv' 데이터 이용하여 '중분류' 변수 생성
- '업종명', '업종코드' 열 삭제

중분류	업종코드
백화점/할인점	4001
백화점/할인점	4002
편의점/마켓	4010
...	...
의류/패션잡화	4209
교육	5106
기타	5003

<업종코드_중분류.csv>

3. 데이터 정제

- 최종 data

기준일자	기준 시간대값	시군구명	카드사명	성별코드	거래건수	총 거래금액	요일	월	age_range	건별 거래금액	중분류
2022-06-01	0	중구	코나	M	1	7350	Wednesday	6	50	7350.00	편의점/마켓
2022-06-01	0	중구	부산	F	1	3700	Wednesday	6	40	3700.00	편의점/마켓
...
2022-08-01	0	영도구	코나	F	1	7500	Monday	8	20	7500.00	생활
2022-08-01	0	영도구	하나	F	2	7600	Monday	8	50	3800.00	편의점/마켓
2022-08-01	0	영도구	하나	M	1	2200	Monday	8	20	2200.00	편의점/마켓

3. 데이터 정제

- 추가 사용데이터 - 부산시 인구 수 데이터 이용 <busan_population.csv>

	0	10	20	30	40	50	60	70	80	90	100
중구	1081	1899	6581	5169	5374	6808	7982	5064	2281	333	10
서구	5193	7192	13333	11015	14470	17211	19388	12413	5943	738	16
동구	3960	4852	10425	11165	11970	14490	16358	10885	5226	623	23
...	
수영구	9565	11359	24002	23991	25062	27066	29027	17777	7613	949	33
사상구	10660	14326	28316	24830	28470	38222	22895	18388	6656	700	22
기장	17594	18408	17163	24168	33625	25867	22895	13370	5783	825	31

3. 데이터 정제

- 추가 사용데이터 - <부산광역시_지역화폐(동백전)가맹점 현황.csv>

순번	가맹점명	도로	지번	위도	경도	데이터기준일자
1	고봉민김밥인	부산 영도구 와치로 250 (동삼동) 1층	NaN	35.08	129.07	2022-12-22
2	미사랑스킨앤바디	부산 사하구 승학로224번길 10 ,1층(괴정동)	NaN	35.10	128.99	2022-12-22
...
126939	현대표구사	부산 중구 중앙대로 137-1 (대창동2가)	NaN	35.11	129.04	2022-12-22
126940	현혜반점	부산 사상구 덕포동 404-7,8	부산 사상구 덕포동 404-7,8	NaN	NaN	2022-12-22
126941	개인택시조합 구내식당	부산 연제구 월드컵대로 399 (거제동)	NaN	35.19	129.06	2022-12-22

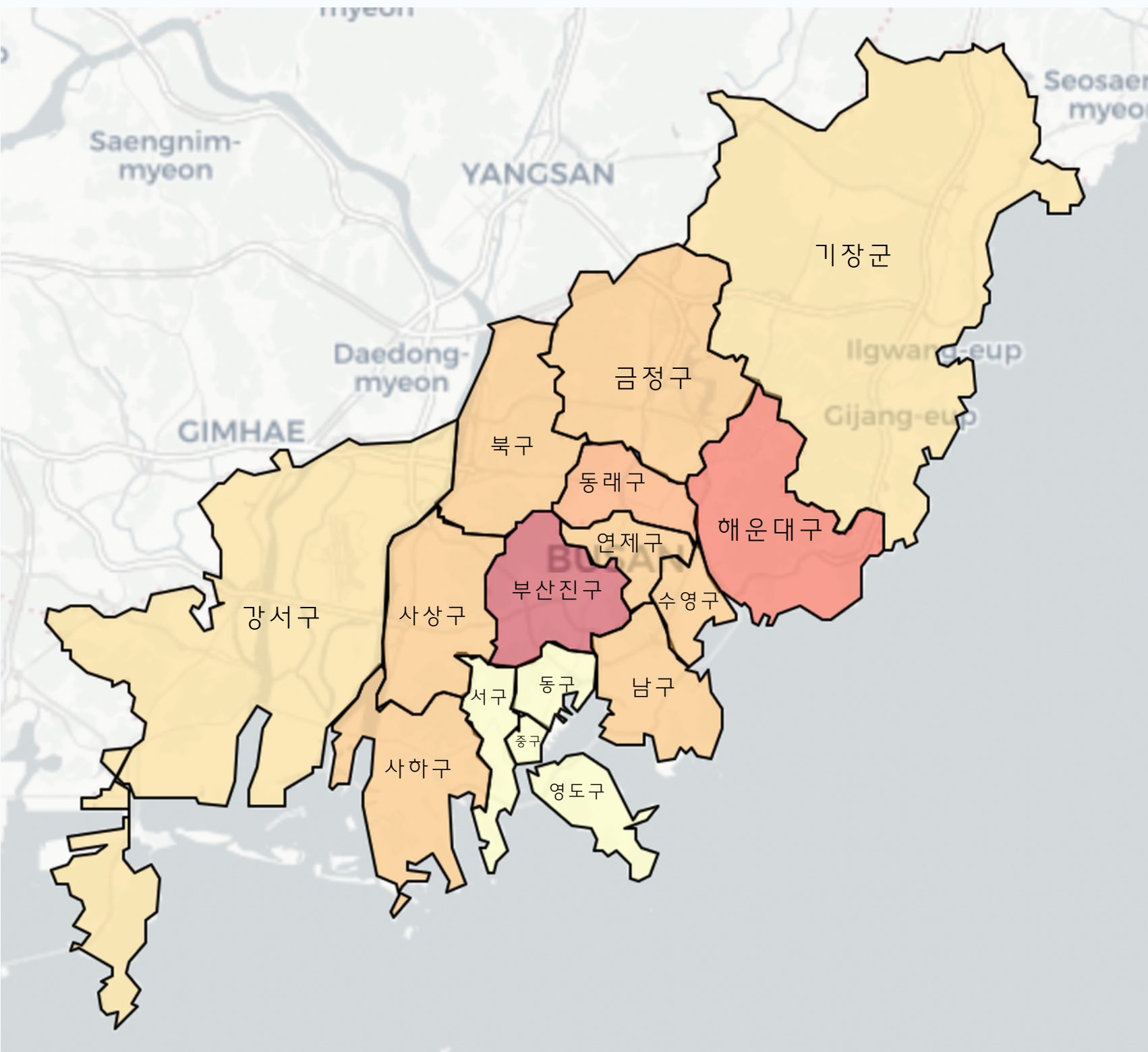
3. 데이터 정제

- 추가 사용데이터 - 부산시에 해당하지 않는 가맹점이 있어 제거 후 '시군구명' 열 생성

순번	가맹점명	도로	지번	위도	경도	데이터기준일자	시군구명
1	고봉민김밥인	부산 영도구 와치로 250 (동삼동) 1층	NaN	35.08	129.07	2022-12-22	영도구
2	미사랑스킨앤바디	부산 사하구 승학로224번길 10 ,1층(괴정동)	NaN	35.10	128.99	2022-12-22	사하구
...	
126939	현대표구사	부산 중구 중앙대로 137-1 (대창동2가)	NaN	35.11	129.04	2022-12-22	중구
126940	현혜반점	부산 사상구 덕포동 404-7,8	부산 사상구 덕포동 404-7,8	NaN	NaN	2022-12-22	사상구
126941	개인택시조합 구내식당	부산 연제구 월드컵대로 399 (거제동)	NaN	35.19	129.06	2022-12-22	연제구

3. 데이터 정제

- 추가 사용데이터 - 가맹점 수 분포



시군구명	가맹점 수
부산진구	16031
해운대구	13122
동래구	10508
사하구	9564
금정구	8710
...	...
강서구	5772
중구	5132
동구	4991
서구	3333
영도구	3218

4. 구별 EDA

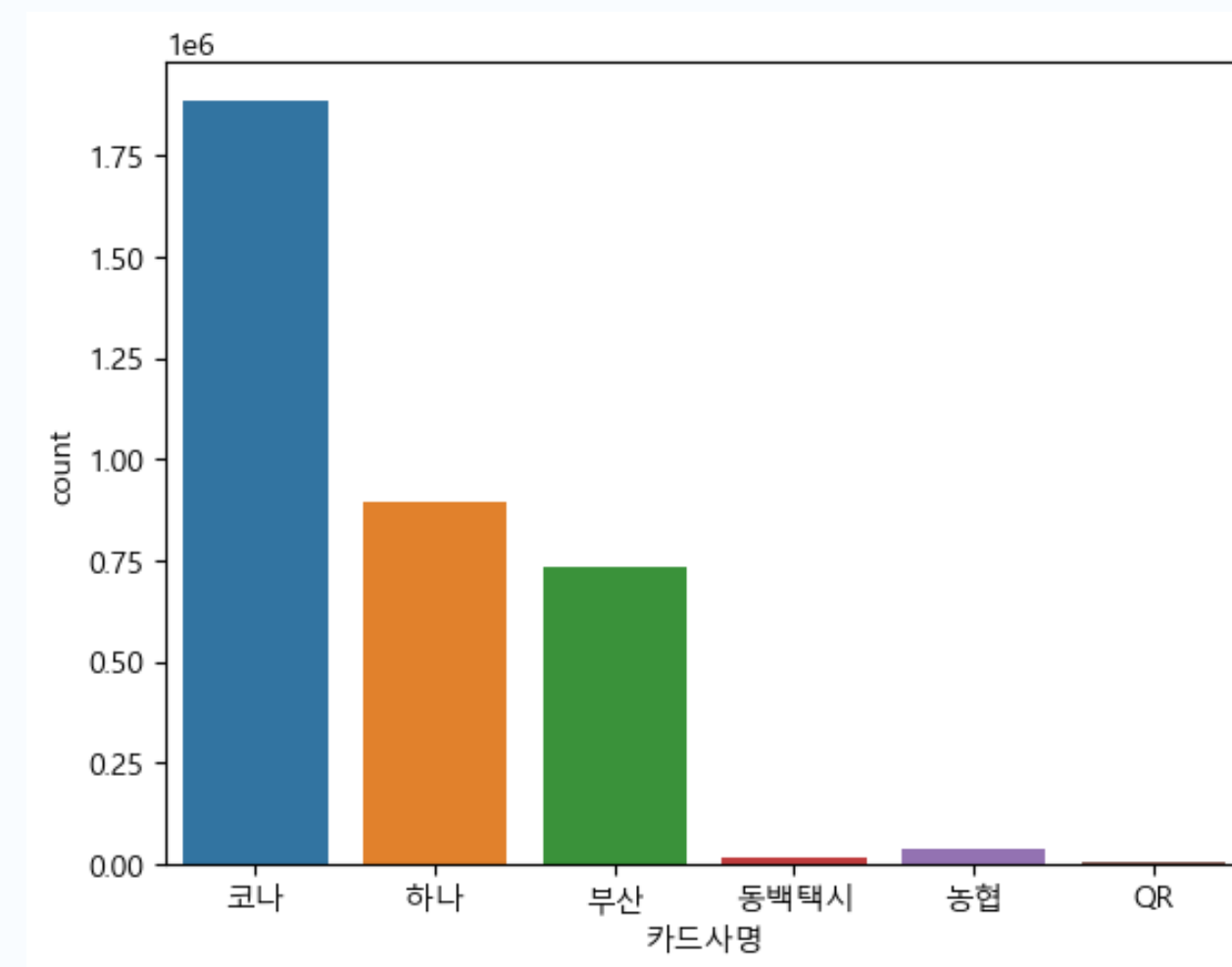
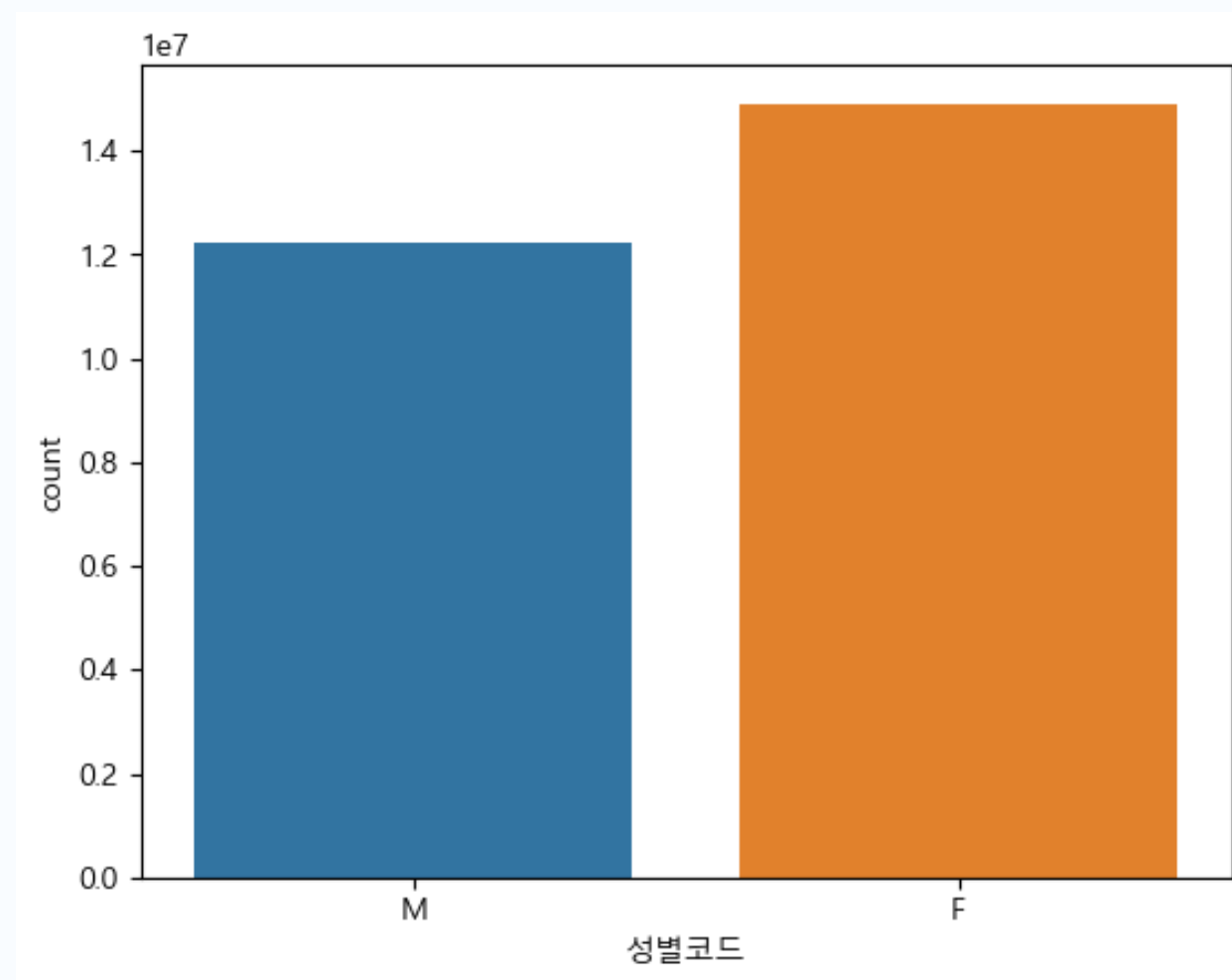
4. 구 별 EDA

- 부산광역시 16개구

(강서구, 금정구, 기장군, 남구, 동구, 동래구, 부산진구, 북구, 사상구, 사하구, 서구, 수영구, 연제구, 영도구, 중구, 해운대구)

- 구 공통점

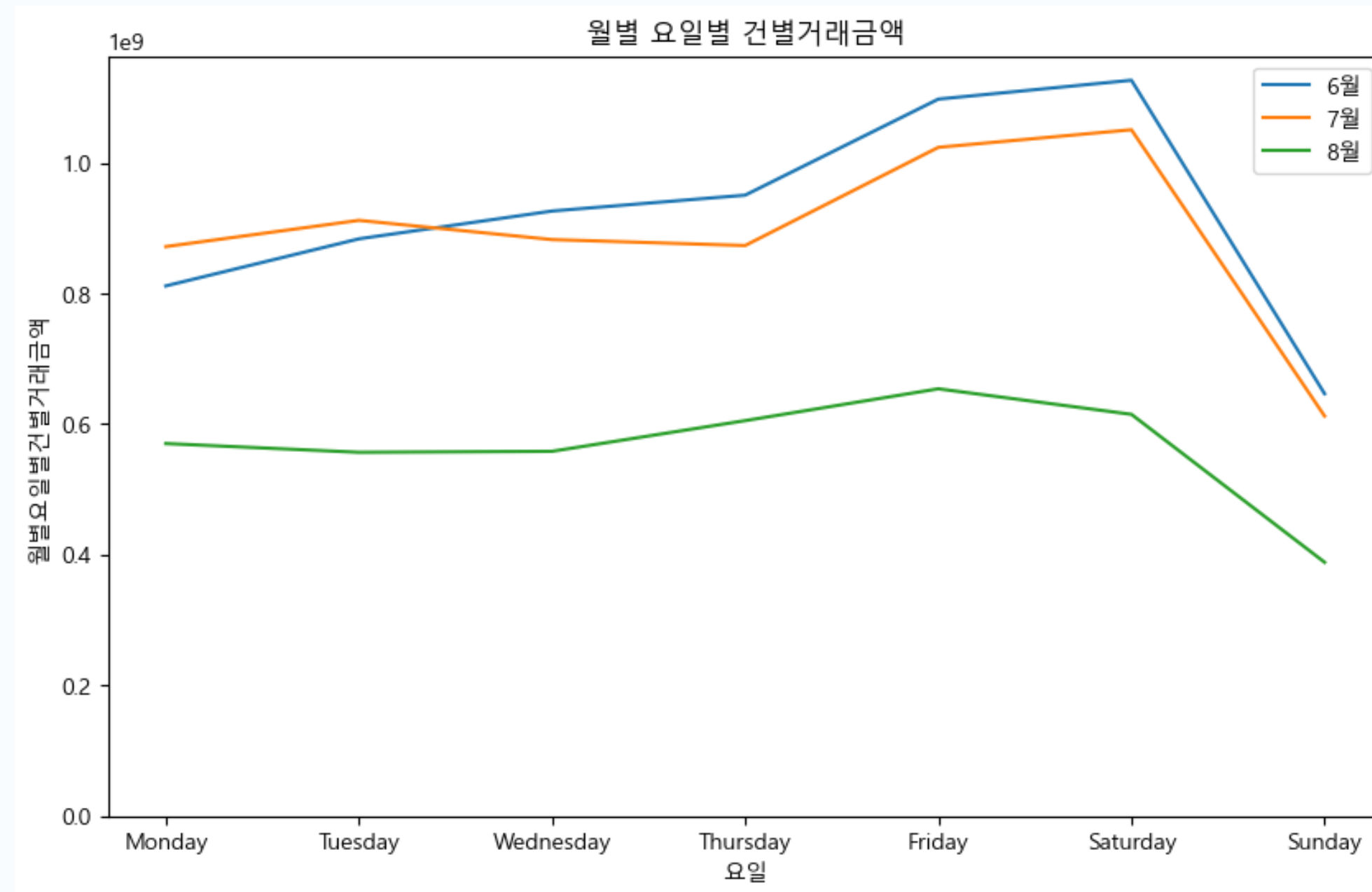
- 여성의 성별이 더 많고, 사용하는 카드사는 코나, 하나, 부산은행 순서



4. 구 별 EDA

- 구 공통점

- 6,7,8월 모두 금요일, 토요일에 가장 많은 금액 사용



4. 구 별 EDA

- 중분류

시군구명	중분류	행 개수	중분류	거래 금액 평균	중분류	거래 금액 중위수	중분류	거래 금액 최대값
강서구	편의점/마켓	437872	전자제품	243433	전자제품	250000	생활	1100000
강서구	양식	209551	교육	122102	교육	100000	전자제품	1000000
강서구	생활	209029	자동차/에너지	52273	자동차/에너지	50000	교육	960000
...
해운대구	편의점/마켓	942002	전자제품	326988	전자제품	400000	생활	1530000
해운대구	생활	610811	교육	172812	교육	150000	문화/취미	1353000
해운대구	양식	482389	자동차/에너지	48752	자동차/에너지	50000	전자제품	1340000

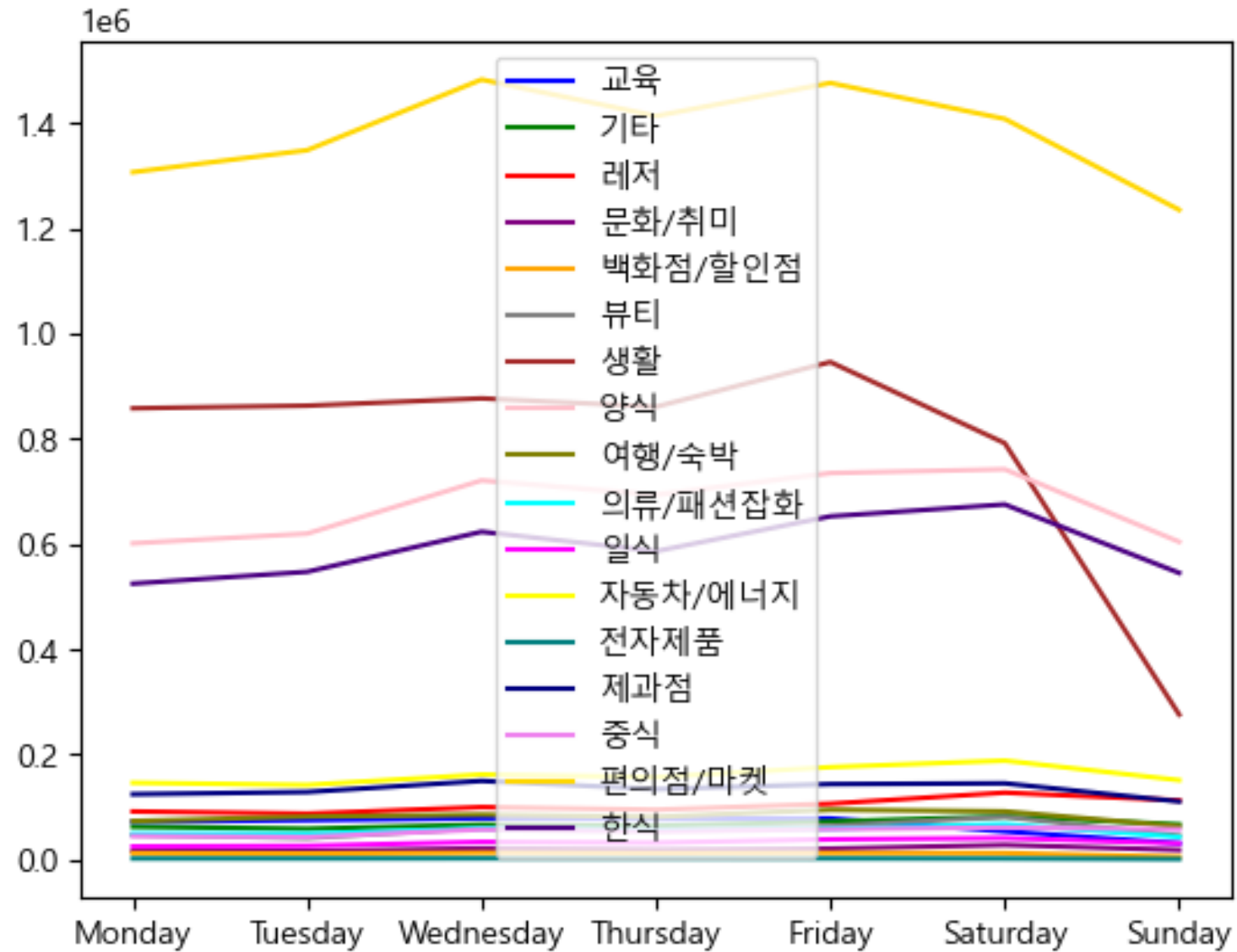
4. 구 별 EDA

• 중분류로 분류한 특징들

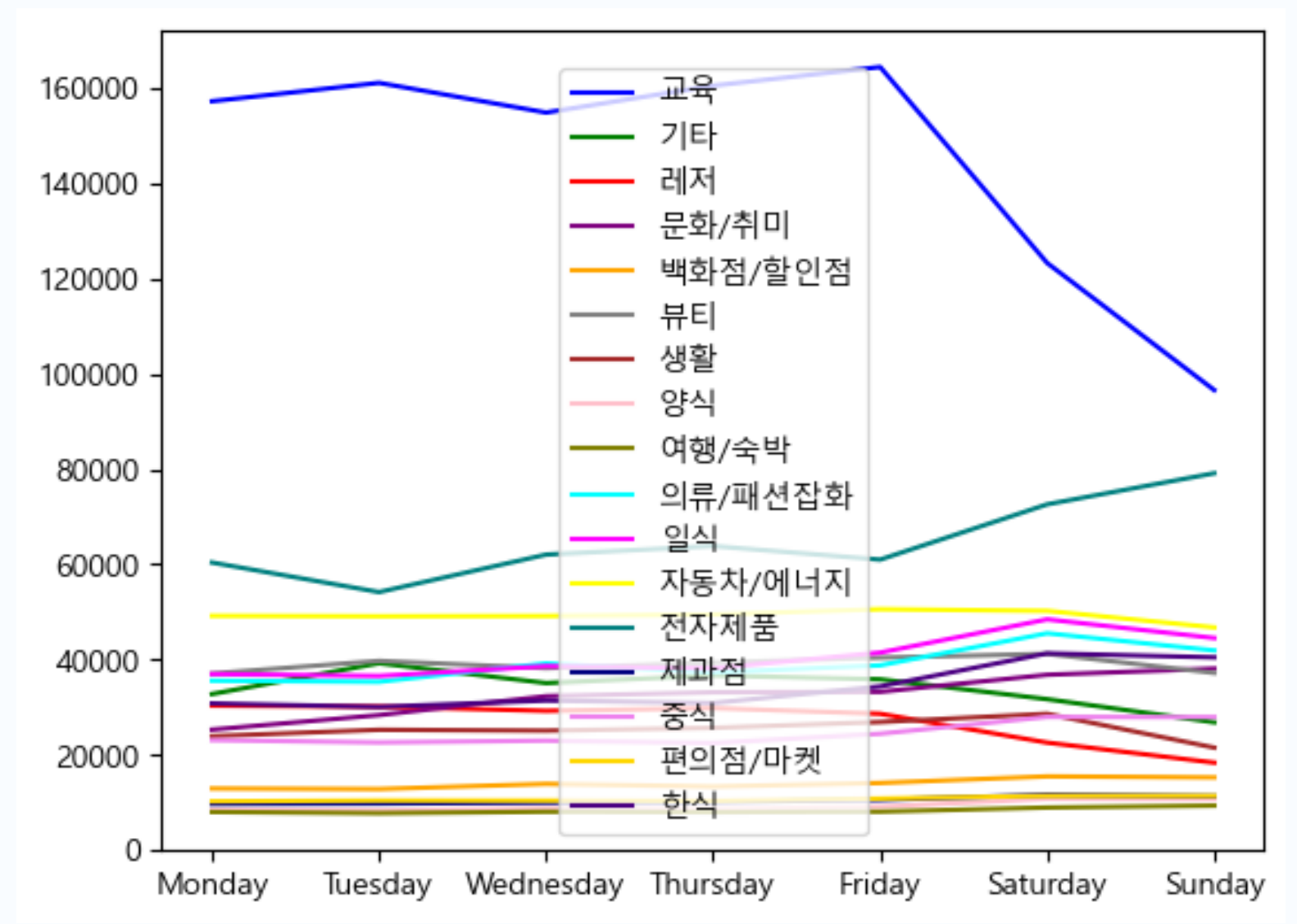
- 대부분 편의점, 생활, 양식 등 음식 가매점의 거래 건수가 많음.
- 건당 거래 금액 평균값의 상위권은 대부분 교육, 자동차/에너지, 일식이 차지
연제구, 해운대구만 문화/취미가 상위권에 등장
- 건당 거래 금액 중위값의 상위권도 대부분 교육, 자동차/에너지, 일식 등.
중구만 의류/패션잡화가 등장 -> 중구에 광복패션거리가 있기 때문.
- 거래 금액의 최대값은 생활, 교육, 뷰티, 레저가 골고루 분포
생활, 교육, 뷰티, 레저는 값비싼 활동이 존재

4. 구 별 EDA

• 중분류



- 거래건수 편의점/마켓, 생활, 양식, 한식 순



- 거래금액 평균 교육, 전자제품, 자동차/에너지 순

4. 구 별 EDA

• 기 준 시 간 대 값

시군구명	기준시간대값	행 개수	기준시간대값	거래 금액 평균	기준시간대값	거래 금액 중위수	기준시간대값	거래 금액 최대값
강서구	12	129711	20	26185	19	13410	18	1100000
강서구	18	113918	19	26176	20	12450	11	1000000
강서구	17	111384	15	25589	18	12240	16	1000000
...
해운대구	12	295642	19	25912	19	11000	10	1530000
해운대구	18	249615	20	25820	18	10000	12	1500000
해운대구	13	242192	15	25051	20	10000	17	1500000

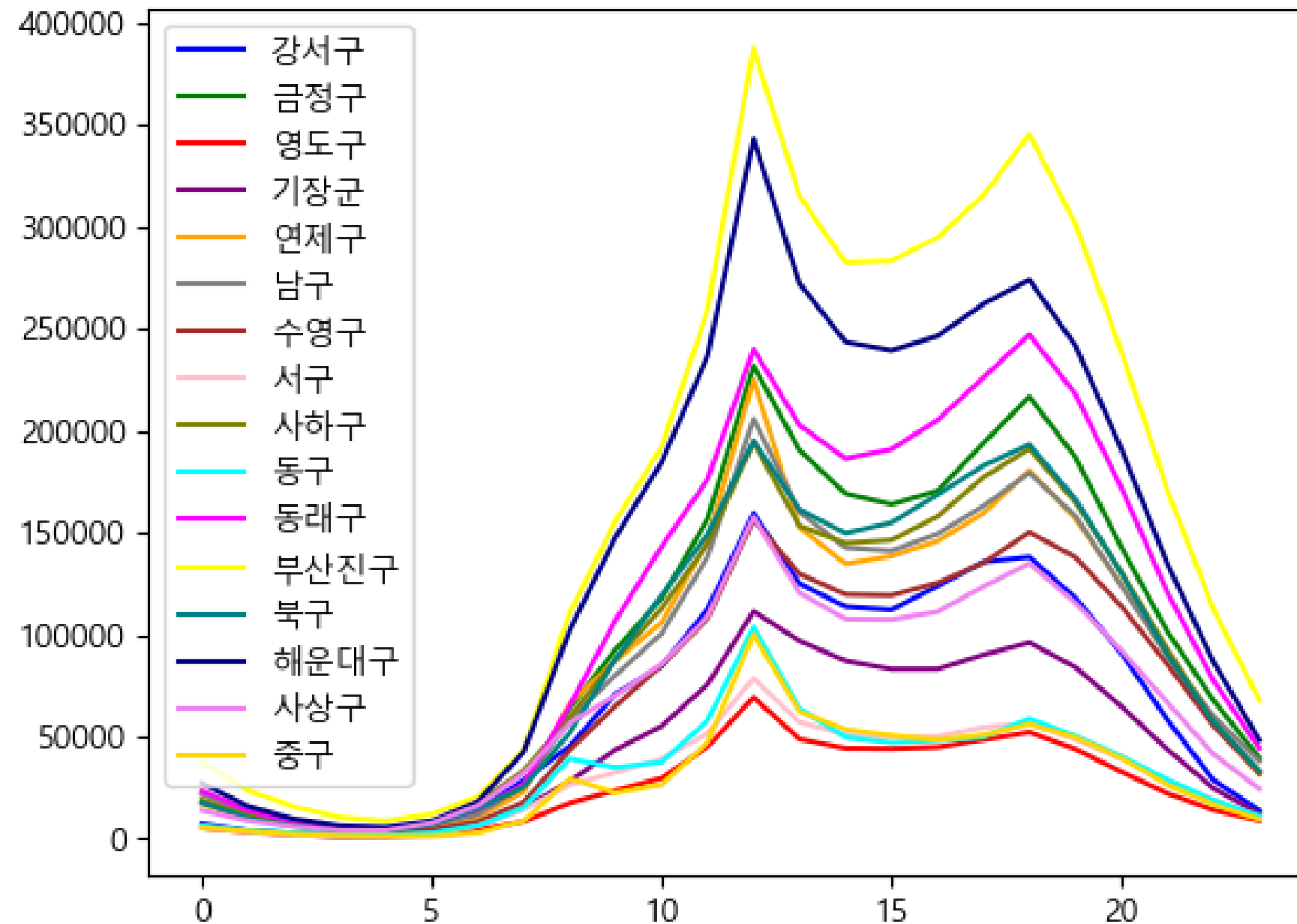
4. 구 별 EDA

- 기준시간대값

- 행 개수는 대부분 12시, 18시 근처 등 식사 시간
- 거래 금액 평균은 대부분 20시, 19시 등 퇴근 후 식사 혹은 회식 시간
(15시도 종종 보이는데 학원 혹은 취미 활동)
- 거래 금액 중위수는 평균과 유사
- 거래 금액 최대값은 골고루 분포

4. 구 별 EDA

• 기준시간대 값



시간대별 거래건수

- 전체적인 분포는 비슷하나 부산진구, 해운대구가 많음

-> 인구수, 가맹점 수 많기 때문인 것으로 보임

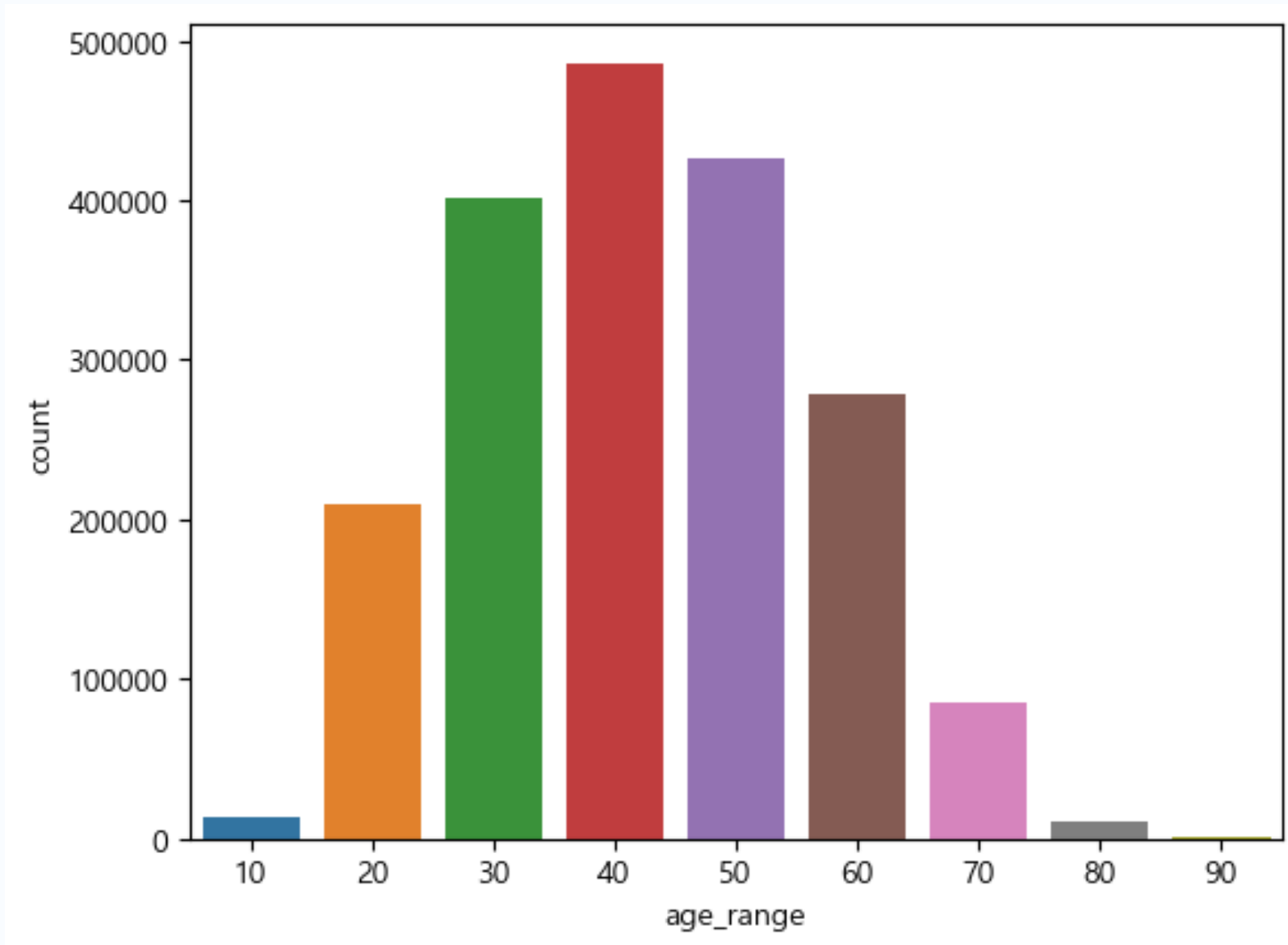
4. 구 별 EDA

• 나 이 대

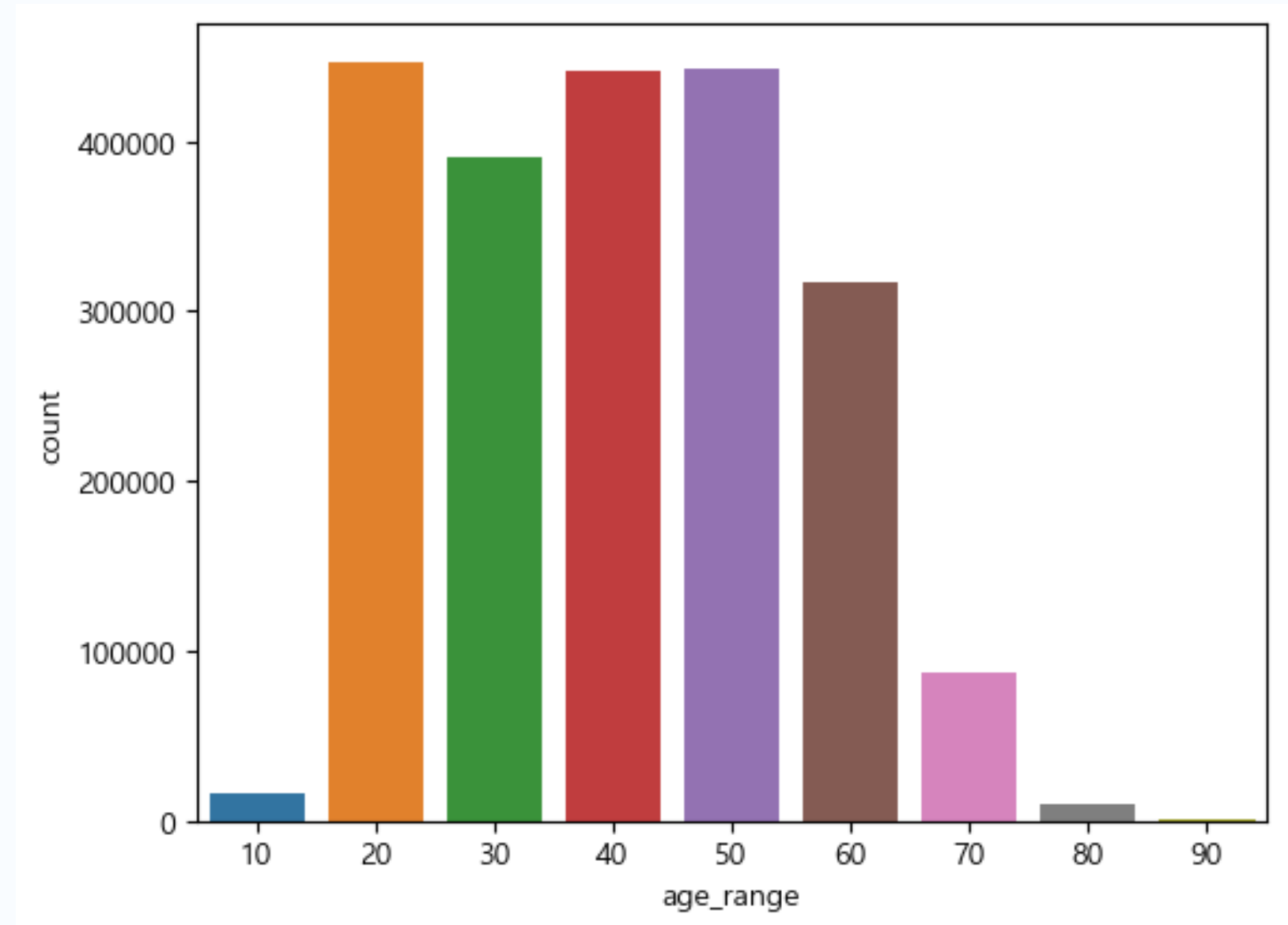
시군구명	나이대	행 개수	나이대	거래 금액 평균	나이대	거래 금액 중위수	나이대	거래 금액 최대값
강서구	40	473223	10	27465	70	12300	40	1100000
강서구	30	381647	70	24983	60	12000	30	1000000
강서구	50	214758	40	24734	80	12000	50	960000
...
해운대구	40	778916	10	30145	80	11000	50	1530000
해운대구	50	698286	40	25491	60	10000	40	1500000
해운대구	30	619444	80	23990	70	10000	60	1500000

4. 구 별 EDA

- 나이대



다른 구 평균 나이대



금정구

- 금정구는 20대가 다른 구들에 비해 거래건수 가장 많다.

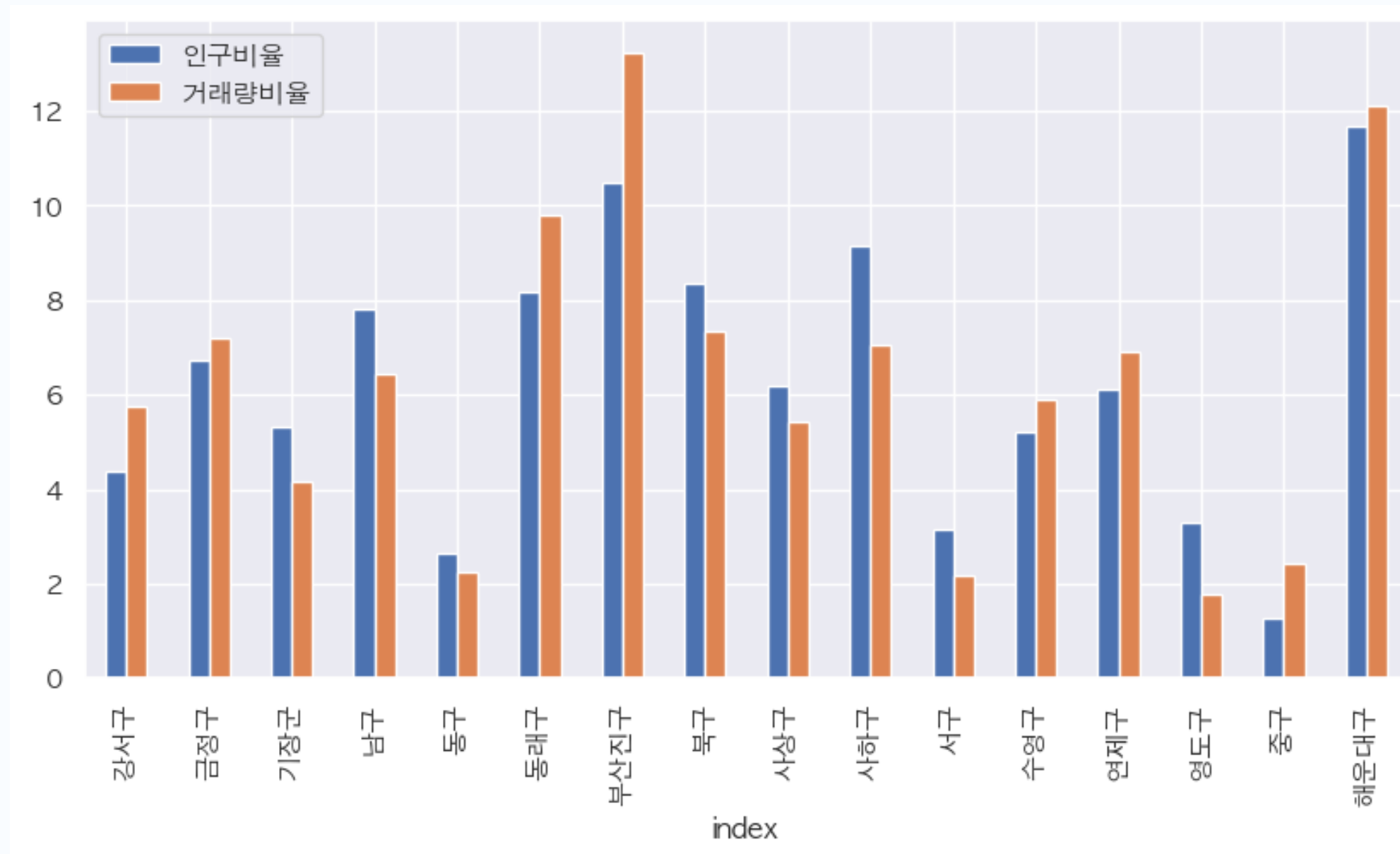
4. 구 별 EDA

• 나 이 대

- 행 개수는 30대, 40대, 50 대가 가장 많이 사용
(금정구에서는 20대가 많이 사용 -> 부산대학교 때문)
- 평균은 10대, 40대, 90대가 가장 많이 사용
(10대는 교육, 40대는 어린 자식, 90대는 의료로 추정)
- 중위수는 60대 이상이 많이 사용
(의료로 추정)
- 최대값은 30~60대 사이
(돈을 가장 많이 가지고 있는 시기, 간혹 20대도 존재)

4. 구 별 EDA

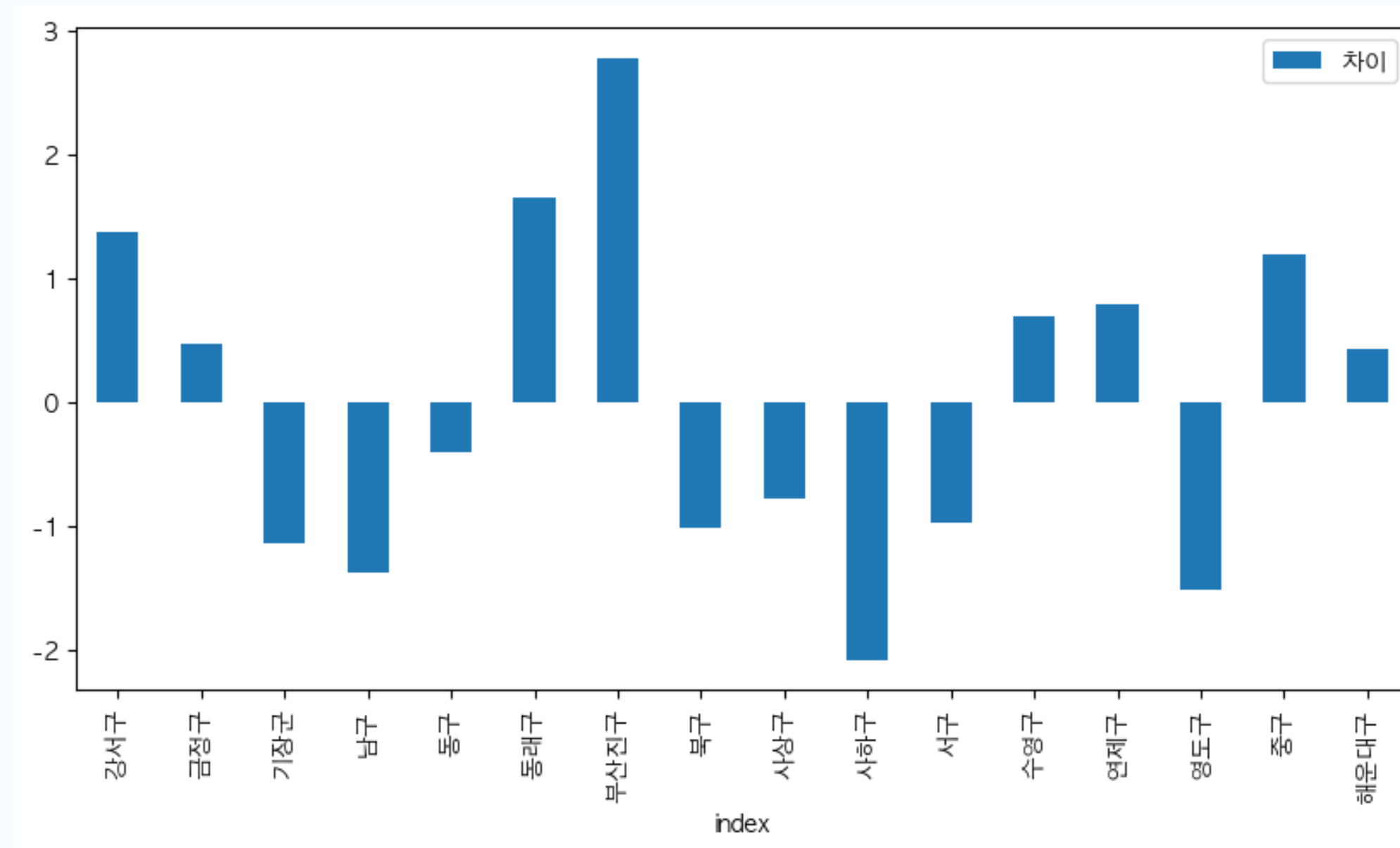
- 인구 비율 대비 거래 금액



각 구의 인구 수와 거래 금액을 부산 전체 수치 중 차지하는 비율로 환산

4. 구 별 EDA

- 인구비율 대비 구매 금액 비율 차이 (거래량비율 - 인구비율)



부산진구에서 인구 대비 소비 금액이 가장 크다. -> 타지역 인구도 많이 소비할 것
사하구, 영도구는 인구 대비 구매 비율 낮음. -> 가맹점 부족

5. CLUSTERING

5. CLUSTERING

k-modes 범주형 클러스터링
kmeans -> 일반 클러스터링
Random Forest -> k means 결과를 y로 두고 결과 해석

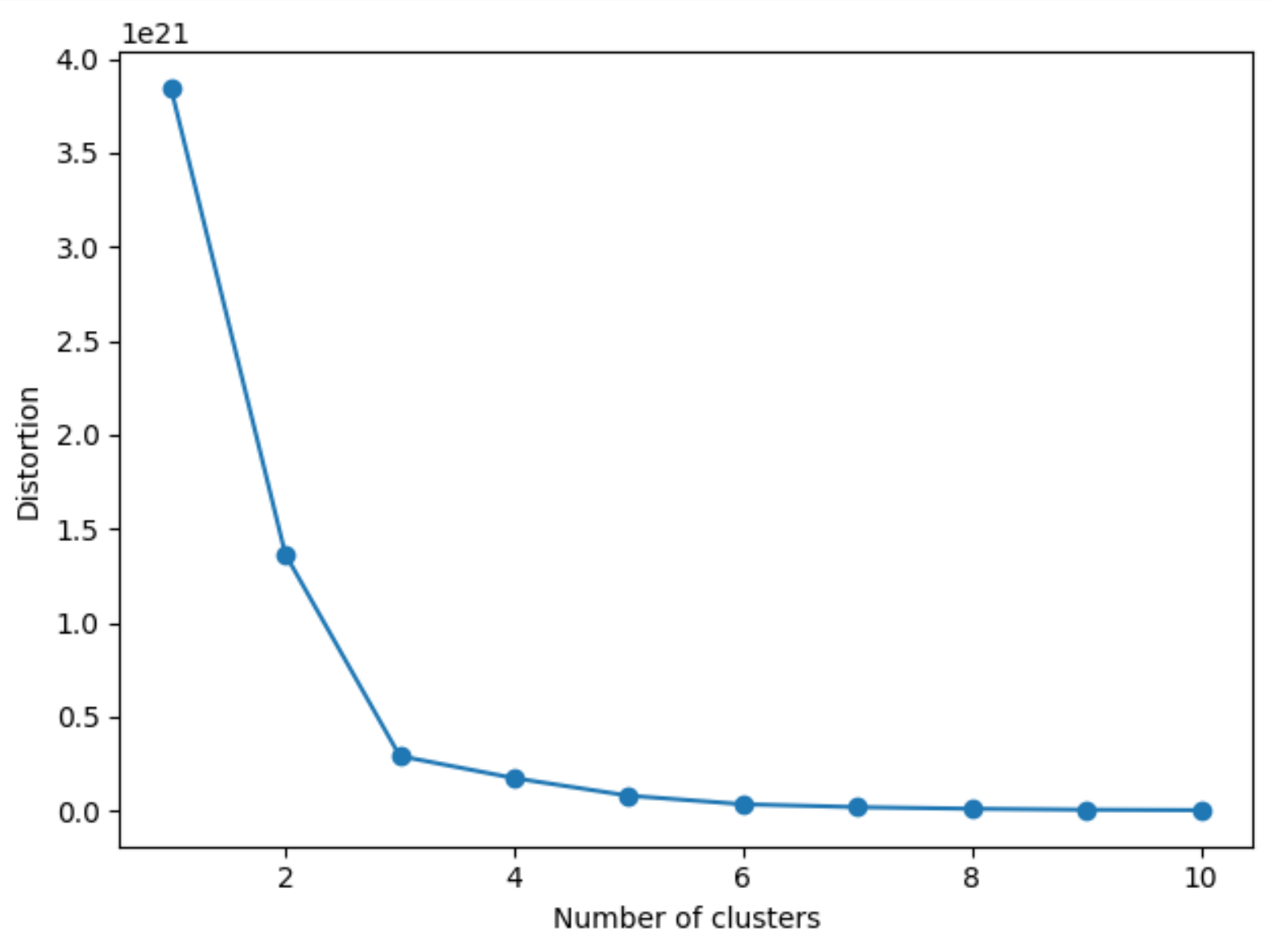
• 구별 특징 변수들 생성

- 거래수, 연령평균, 0~100 나이 연령대, 인구수, 동백전 가맹점 개수들, 인구와 거래량 차이, 평일 거래액, 주말거래액 등

시군구명	거래수	연령대평균	0대	...	거래량비율	인구-거래량	평일거래총액	주말거래총액
강서구	1336118	1978.97	19269		5.759530	1.366549	25366662075	9597086362
금정구	2151442	1978.35	11406		7.202067	0.463848	31485862279	12234932153
기장군	963690	1977.32	17594		4.165540	-1.150514	17274598923	8012689688
...								
연제구	1912041	1976.39	13518		6.903581	0.78764	31132946346	10775864398
영도구	582429	1975.77	4709		1.767899	-1.522023	7646486199	3085706806
중구	655696	1977.46	1081		2.441727	1.182256	10462440555	4360285617
해운대구	3075269	1976.85	25965		12.114723	0.417851	53311395861	20232123125

5. CLUSTERING

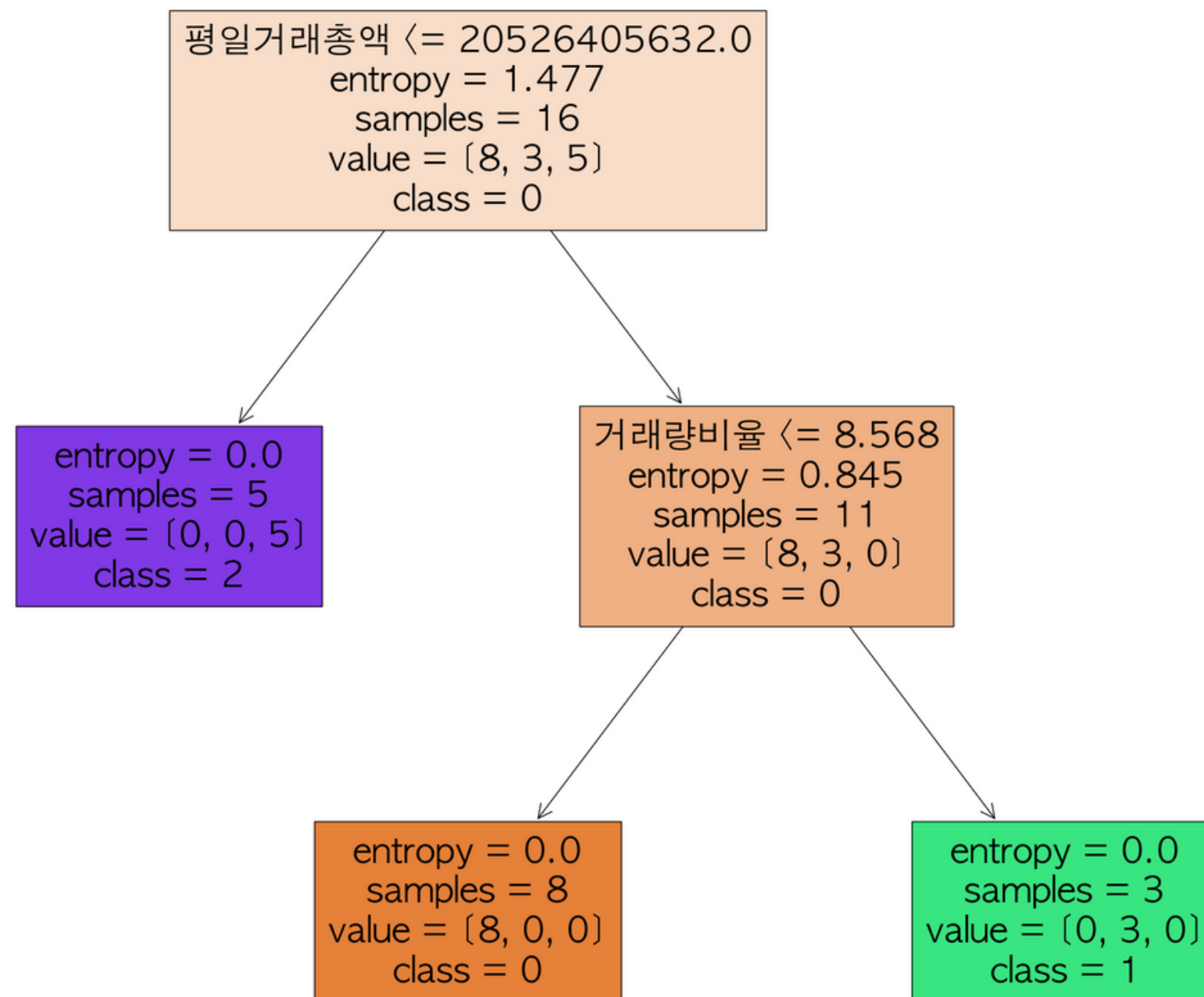
Elbow 기법으로 찾은 적절한 군집 수 : 3



시군구명	Class	시군구명	Class	시군구명	Class
강서구	0	동래구	1	기장군	2
금정구	0	부산진구	1	동구	2
남구	0	해운대구	1	서구	2
북구	0			영도구	2
사상구	0			중구	2
사하구	0				
수영구	0				
연제구	0				

5. CLUSTERING

Random Forest로 분석한 클래스 분류 기준



2가지 거래 변수로만 클래스 구분

-첫번째 분기점 변수: 평일거래총액
기장군, 동구, 서구, 영도구, 중구

인구수가 적은 편이고, 타지역으로 평일에
출퇴근을 많이 하는 주거 지역들

-두번째 분기점 변수 : 거래량비율
동래구, 부산진구, 해운대구

세 구는 부산시 안에서 가장 많은 소비가
이루어진다. 상업이 발달한 지역들.

6. 결론

6. 결론

• 고령층 동백전 사용 활성화 방안

- 동백전은 30~50대가 가장 많이 사용
- 60대 이상의 고령층도 타겟으로 하여 고령화 시대의 흐름에 맞추어 고령층의 동백전 사용을 활성화하기 위한 방안이 필요
- 동백전은 카드로만 사용이 가능하며 2021년 4월 개편 이후 신규 가입자는 동백전 앱 또는 동백전 홈페이지에서 신청가능, 웹으로 신청하려면 동백전 앱을 깔고 회원가입 필요
따라서 노인들을 위한 앱 이용 교육 등 취약계층인 고령층 대상 배려 정책 마련하거나 고령층은 화폐 사용이 더 익숙하므로 동백전을 화폐화하여 사용 확대
- 상대적으로 60대 이상 비율이 많은 수영구, 영도구 등의 고령층 비율이 높은 지역의 경우 교육 분야에서의 사용 금액이 가장 많음
교육 분야의 가맹점 수 확대하여 동백전 활성화
- 고령층을 위한 의료 분야 가맹점 확대

6. 결론

• 20대 동백전 사용 활성화 방안

- 20대의 비율이 높은 금정구에서는 생활, 레저 분야에서 가장 많은 금액 사용
- 다른 구에서도 생활, 레저 분야의 가맹점 수를 늘리고 캐시백 적립 등 동백전 사용 혜택 홍보

• 구별 특성을 고려한 가맹점 다양화

- 다른 구들과 다르게 특정 분야의 거래가 돋보이는 지역은 특성화 거리로 홍보
해운대구, 연제구 : 건당 거래금액 중 유일하게 문화/취미 거래 금액이 높은 지역.
문화/취미 업종 중 화원, 액세서리, 민예공예품 가맹점이 많은 골목은 전문 거리로 활성화
- 중구 : 광복패션거리의 존재성을 지역내외에 홍보
광복패션거리에서 귀금속, 신발 등의 구매 금액이 높음.
- 주거 지역 가맹점의 영업시간 조정 홍보
평일에 거래가 적은 구역은 주민들이 출근 전, 출근 후에 방문할 수 있도록 영업 시간을 조정해보도록 제안한다.
ex) 아침에 일찍 여는 베이커리, 야간 영업을 하는 병원