

서울시 상권별 창업 성공률과 인구밀도, 매출, 공실률의 관계 분석

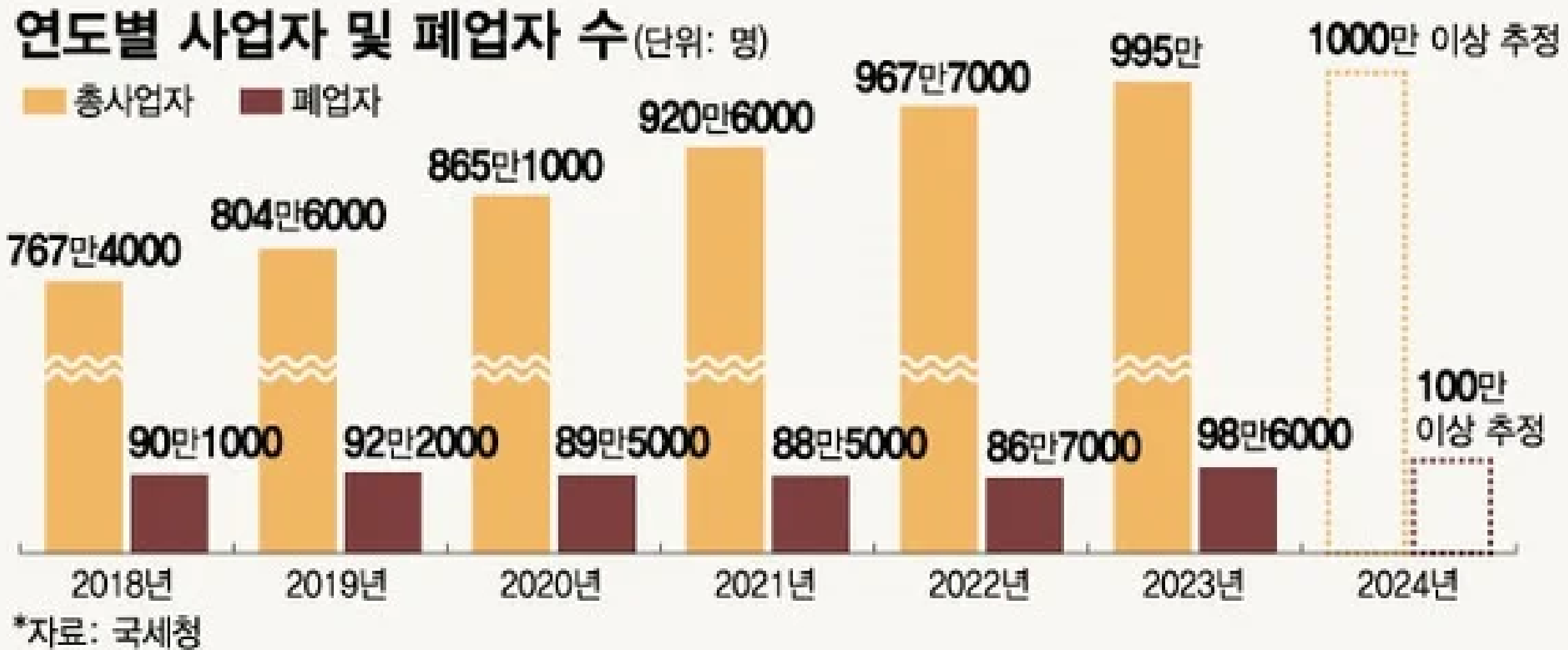


목차

- 01 주제 소개
- 02 프로젝트 개요
- 03 가설
- 04 데이터 수집 방법
- 05 데이터 탐색
- 06 데이터 모델링 및 시각화
- 07 결론



창업 대비 폐업 비율은 **79.4%**, 가게 10곳이 문을 여는 동안 8곳이 닫는다.

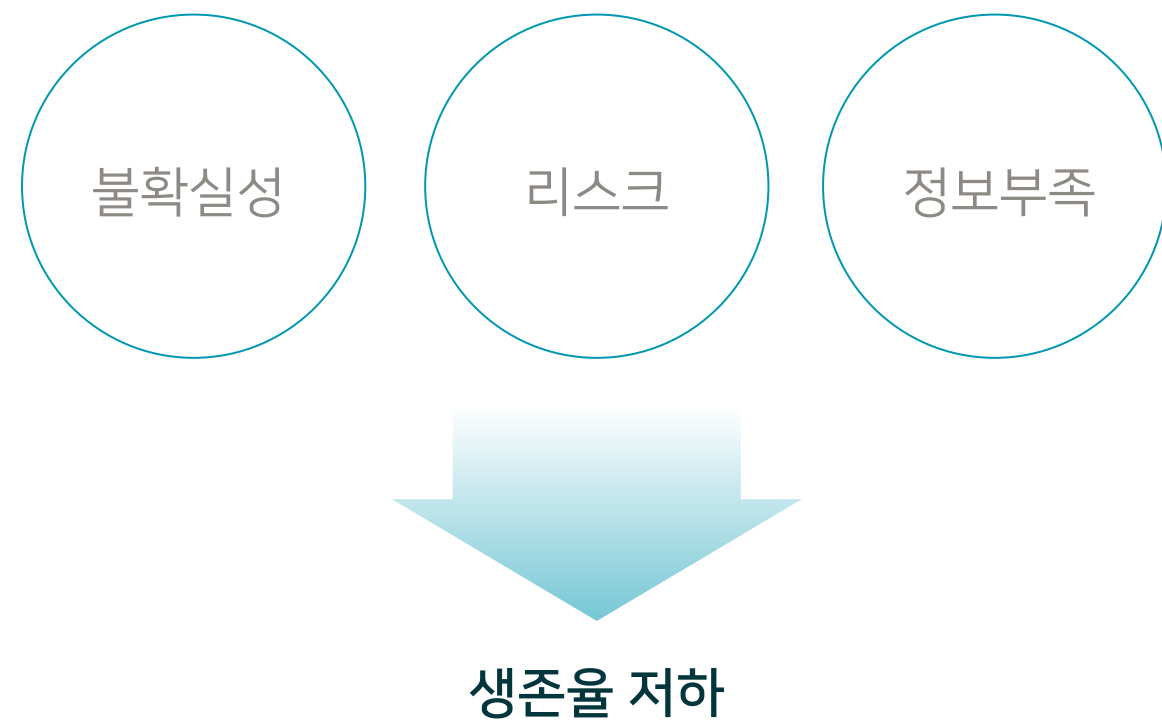


카페·편의점 등 대책없이 진출

창업을 선호하는 직종(복수 응답)으로는 카페, 베이커리, 주점 등이 포함된 외식·음료업을 선택한 응답자(68.4%)가 가장 많았다. 편의점, 쇼핑몰, 꽃집 등 유통·판매업(24.2%)이 뒤를 이었다.

전문가들은 충분한 사전 준비 없이 경쟁이 치열한 업종에 뛰어들면 성공 가능성이 크지 않다고 지적했다

창업사회 속 Pain Point



서울 시내 카페 생존율

“올해 2분기 기준 서울 시내 카페의 ‘1년 생존율’은 77.7%”라며
“네 곳 중 한 곳이 1년 이내 망할 정도로 경쟁이 치열하기 때문에
창업할 땐 철저하게 고민하고 준비해야 한다”

자료출처-한경일보

창업자의 절반 이상이 3년 내 폐업하는 현실에서,
창업 전 단계부터 데이터 기반으로 생존 가능성과 리스크를 정량적으로 예측할 수 있는 도구가 있다면 어떨까?

서울시 상권별 대표 업종의 창업 성공률 분석

주제: 서울시 상권별 대표 업종의 창업 성공률 분석

1. 창업의 범위

창업에도 많은 종류가 있습니다. 자영업, 프랜차이즈, 스타트업 등등
그중에서 자영업자에만 한정하여, 창업의 개념을 소규모의 형태에 집중하였습니다.

2. 성공의 정의

월 매출 기준, 지속 영업 기간
ex. 월매출 1,000만원 이상 3년 이상 영업중

3. 업종 샘플링

서울 전체 상권과 모든 업종을 다 분석하면 데이터가 수십만 건에 달해 해석과 활용이 어렵기 때문에,
실제 창업이 활발한 대표 상권과 주요 업종만을 샘플링하여 각 업종별로 집중적으로 비교·분석하였습니다.

주제: 서울시 상권별 대표 업종의 창업 성공률 분석

상권 유형별 대표 3곳 샘플링
대학가, 상업지구, 베드타운

각 상권별로 점포수 기준
주요 업종 3개씩 선정

성공률에 영향을 미치는 요소

평균유동인구

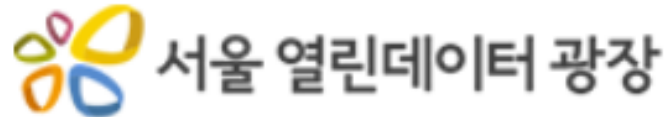
평균매출

공실률

가설1 유동인구가 많을수록 상권·업종의 생존율(성공률)이 높을 것이다.

가설2 동일 업종(유사업종) 밀집도가 높을수록 생존율(성공률)이 낮을 것이다.

가설3 공실률이 높을수록 상권·업종의 생존율(성공률)이 낮을 것이다.



서울시 상권분석서비스(점포-상권)

상권-업종 데이터 (점포-상권)

- 수집항목: 상권코드, 상권코드명, 서비스업종코드, 서비스_업종코드명, 점포수, 유사업종점포수, 개업/폐업_점포_수, 개업률/폐업률, 프랜차이즈 점포수
- 데이터 갱신일: 2025.06.11
- 제공: 서울신용보증재단 빅데이터팀
- 출처: 서울 열린데이터 광장

서울시 상권분석서비스(영역-상권)

상권-영역/위치 데이터 (상권 좌표/행정동)

- 수집항목: 상권 식별, 명칭, 상권 위치(좌표), 매핑용 행정동 코드/명
- 데이터 갱신일: 2025.06.11
- 제공: 서울신용보증재단 빅데이터팀
- 출처: 서울 열린데이터 광장

행정동 단위 서울 생활인구(내국인)

생활인구(인구밀도) 데이터

- 수집항목: 매핑용 행정동 코드, 해당 행정동의 생활 인구
- 데이터 갱신일: 2025.06.19
- 제공: 서울특별시 디지털도시국 데이터전략과
- 출처: 서울 열린데이터 광장

서울시 매장용빌딩 임대료·공실률 및 수익률 통계

매장용빌딩 공실률/임대료 데이터

- 수집항목: 대분류, 소분류 지역명, 임대료 정보, 공실률, 수익률 정보
- 데이터 갱신일: 2025.03.04
- 제공: 서울특별시 디지털도시국 데이터전략과
- 출처: 서울 열린데이터 광장

1. 서울시 상권/점포/매출/인구/공실률 4개 데이터 병합

-전처리 전 초기 데이터 형태

데이터셋	컬럼 형태
점포-상권	상권코드/명
상권-영역	상권코드/명 → 행정동코드/명, 좌표로 변환 정보 제공
생활인구	'행정동' 또는 '지역명' 단위
공실률/임대료	'행정동' 또는 '지역명' 단위

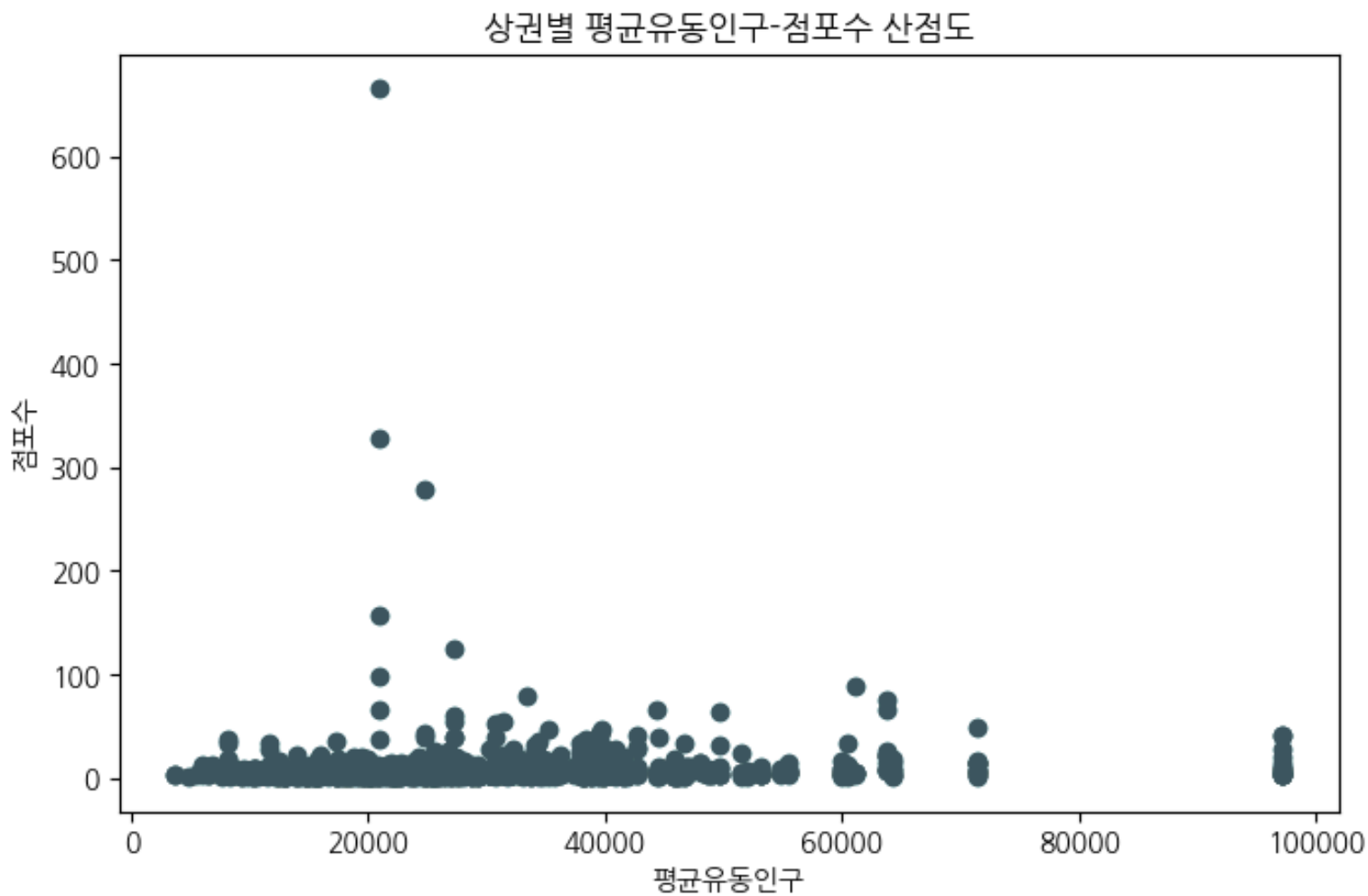
행정동코드 매핑(병합) 필요

-데이터 병합 후

	기준_년분 기_코드	상권_구 분_코드	상권_구분 _코드_명	상권_코 드	상권_코 드_명	서비스_업종_ 코드	서비스_업 종_코드_명	점포_ 수	유사_업종_ 점포_수	개업_ 율	개업_점 포_수	폐업_ 률	폐업_점 포_수	프랜차이즈 _점포_수	행정동_코 드	행정동_ 코드_명	당월_매출_금 액	평균유동인 구	생존율 (%)
0	20241	U	관광특구	3001491	이태원 관광특구	CS100001	한식음식점	122	134	8	10	3	4	12	11170650	이태원1동	1.185751e+10	11660.410198	97
1	20241	U	관광특구	3001491	이태원 관광특구	CS100001	한식음식점	122	134	8	10	3	4	12	11170650	이태원1동	1.312908e+10	11660.410198	97
2	20241	U	관광특구	3001491	이태원 관광특구	CS100001	한식음식점	122	134	8	10	3	4	12	11170650	이태원1동	1.276885e+10	11660.410198	97
3	20241	U	관광특구	3001491	이태원 관광특구	CS100001	한식음식점	122	134	8	10	3	4	12	11170650	이태원1동	1.432346e+10	11660.410198	97
4	20241	U	관광특구	3001491	이태원 관광특구	CS100001	한식음식점	122	134	8	10	3	4	12	11170650	이태원1동	1.216901e+10	11660.410198	97

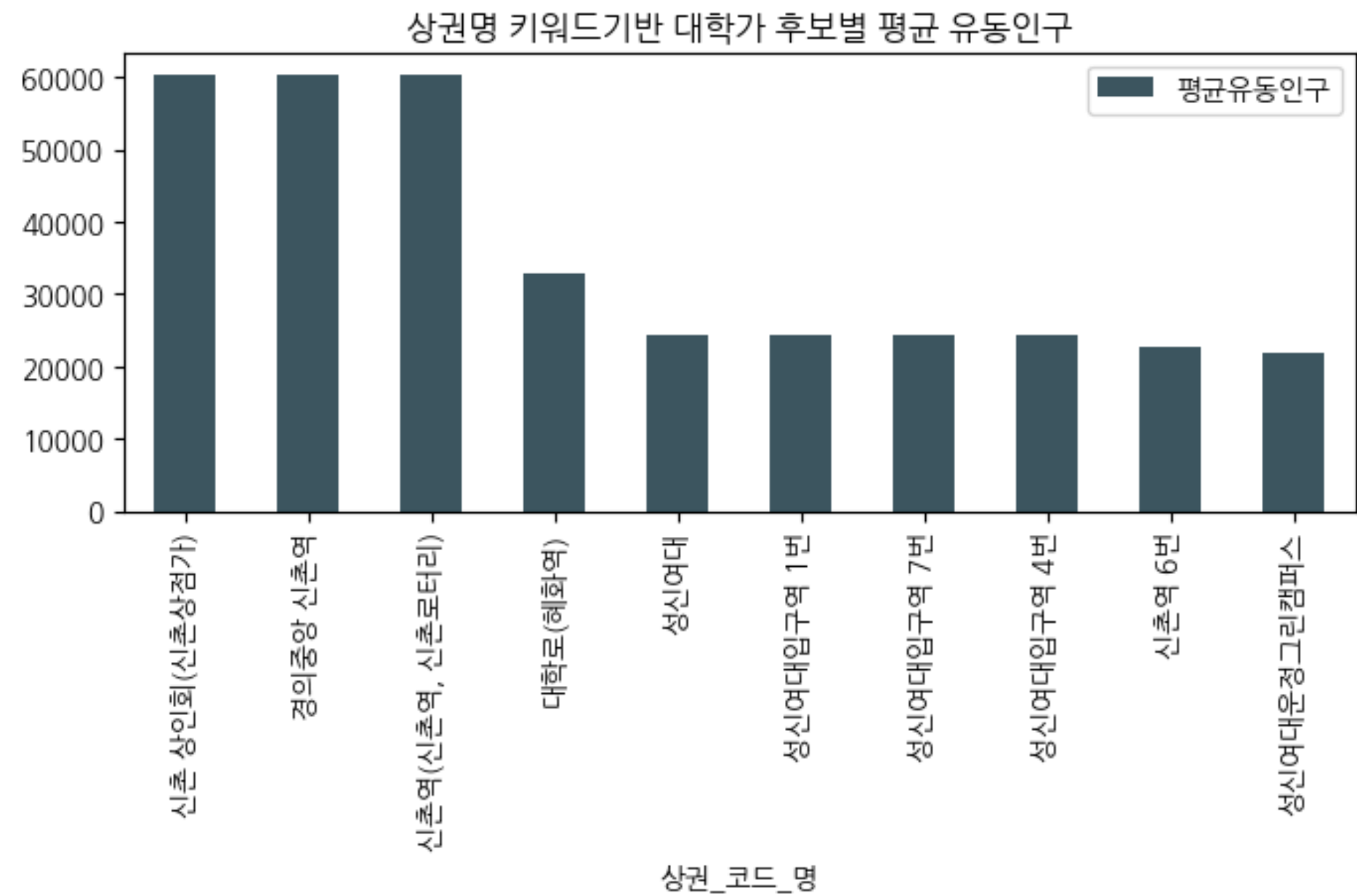
2. 상권별 인구/매출/점포수 집계)

	상권_코드_명	점포_수	당월_매출_금액	평균유동인구
1	63빌딩	5.290514	2.368735e+09	97163.050067
1145	여의도역(여의도)	40.398350	4.297847e+09	97163.050067
222	국회의사당역(국회의사당)	19.551444	2.208734e+09	97163.050067
106	경복아파트교차로	6.790584	9.747475e+08	97100.315119
1142	언주역 8번	2.504587	5.365235e+07	97100.315119
654	뱅뱅사거리	28.537008	1.708163e+09	97100.315119
207	구역삼세무서	14.605217	7.118093e+08	97100.315119
1147	역삼역 4번	3.003774	7.250497e+07	97100.315119
213	국기원	2.875912	1.171632e+09	97100.315119
495	르네상스호텔사거리	10.656085	6.235942e+08	97100.315119



3. 상권 카테고리 분류

① 대학가 신촌역, 성신여대, 대학로

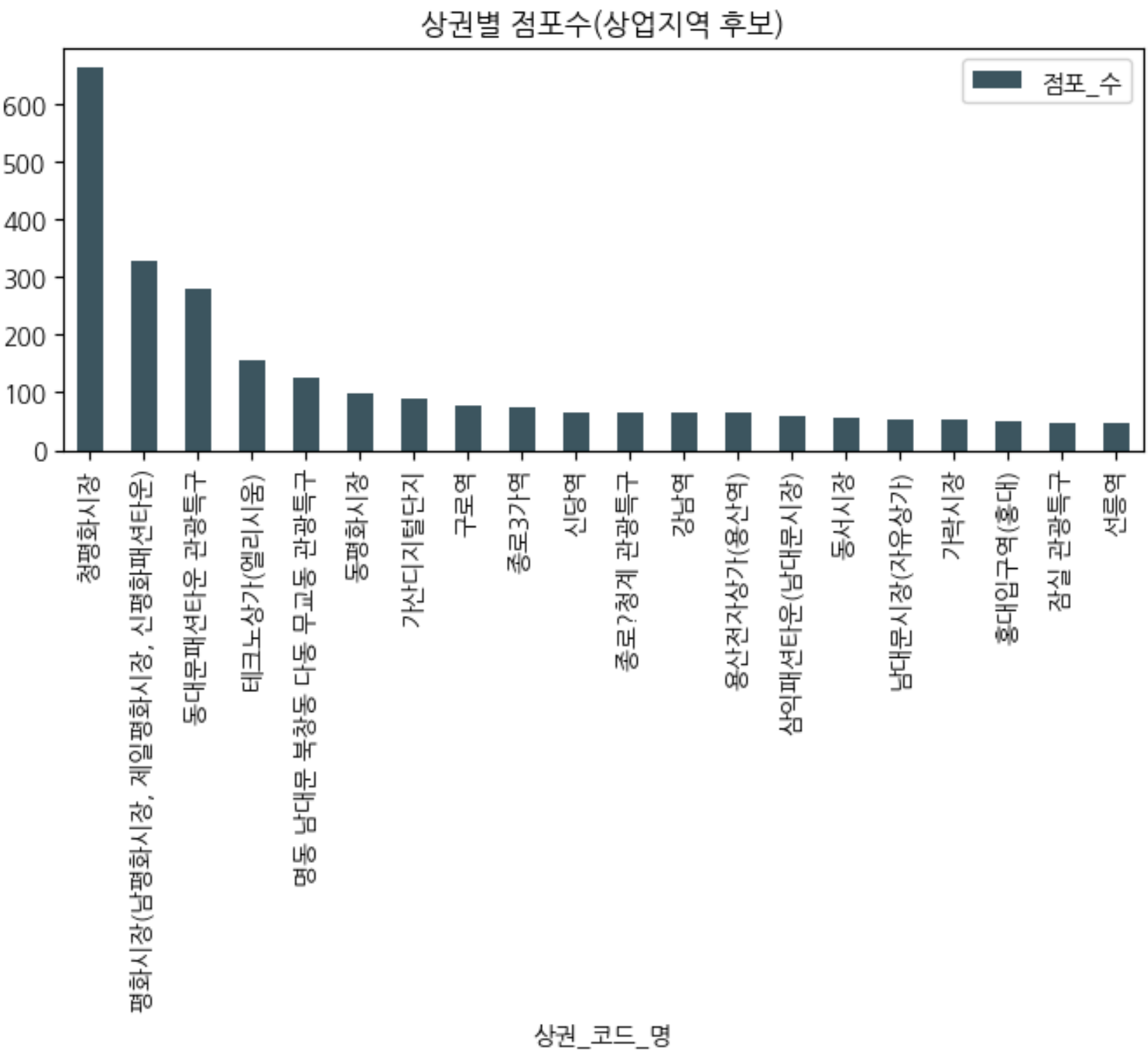


keyword = ['성신여대', '대학로', '신촌']

- 주어진 키워드(대학가 연관) 포함된 상권만 골라, 상권별로 평균 유동인구/점포 수 집계.

3. 상권 카테고리 분류

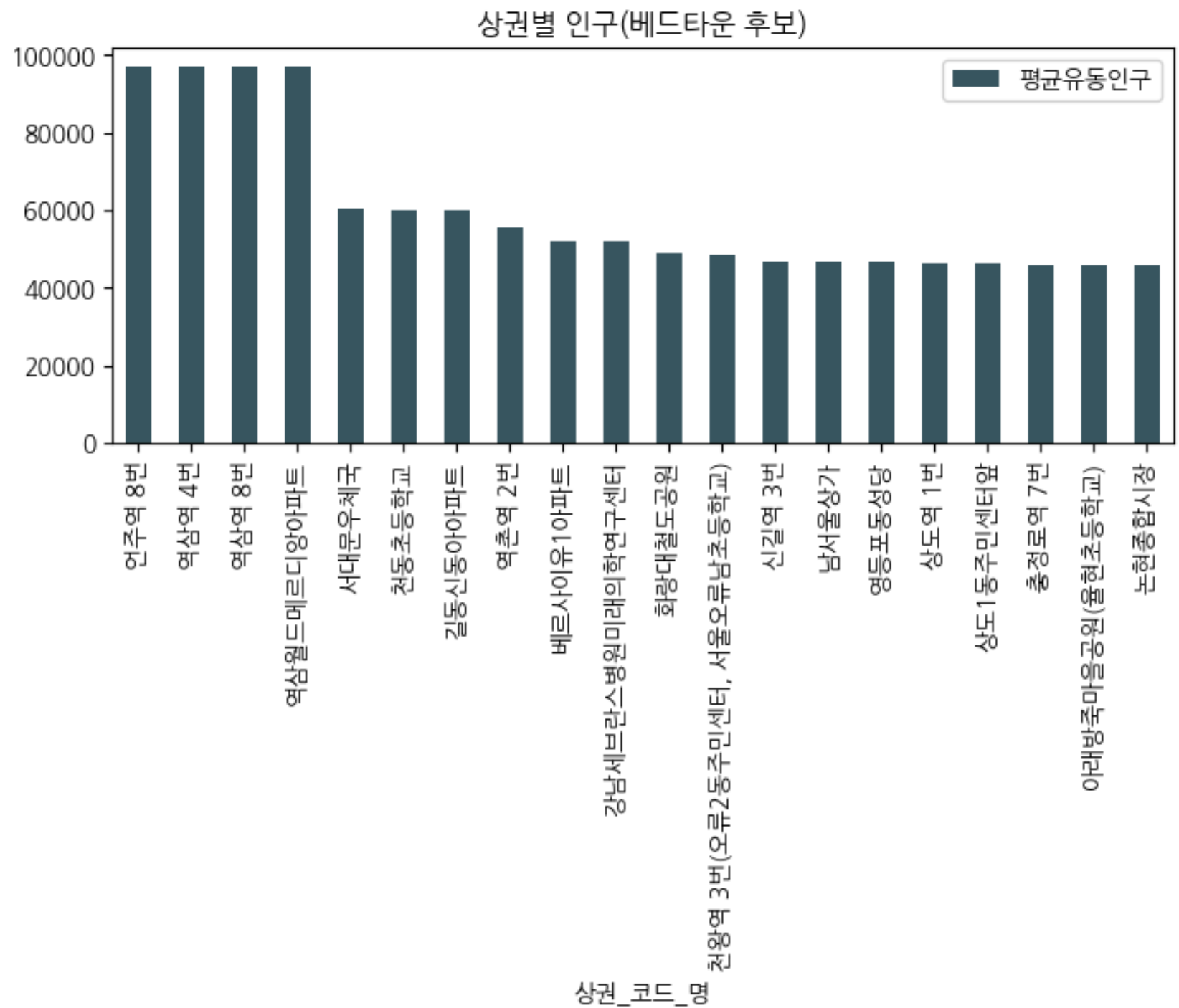
② 상업지구 청평화시장, 구로역, 용산역



- 대학가 후보(campus_areas)에 해당하지 않는 상권만 추출
- 점포수와 매출액 기준 내림차순 정렬

3. 상권 카테고리 분류

③ 베드타운 언주역 8번, 서대문우체국, 길동신동아아파트



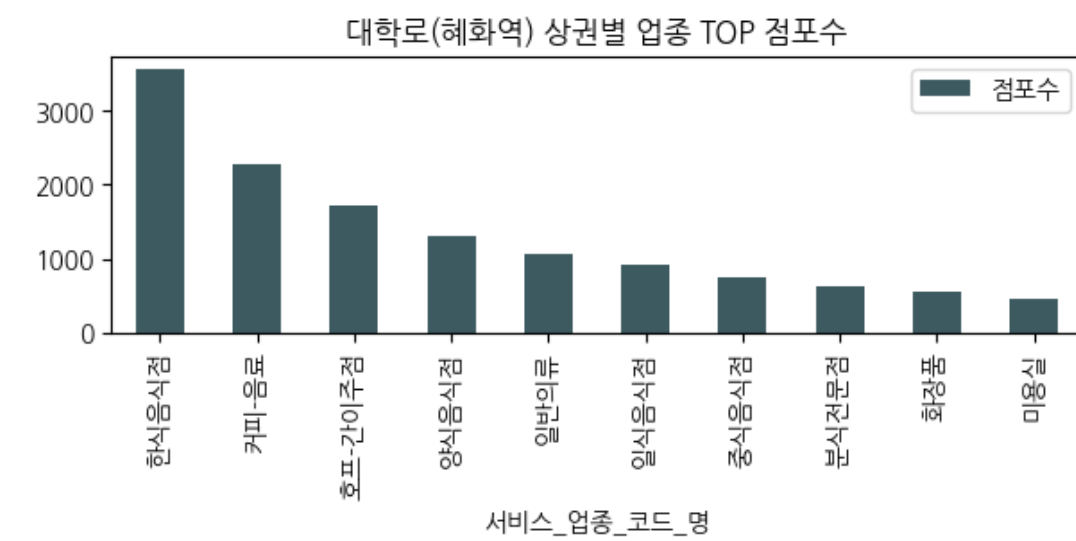
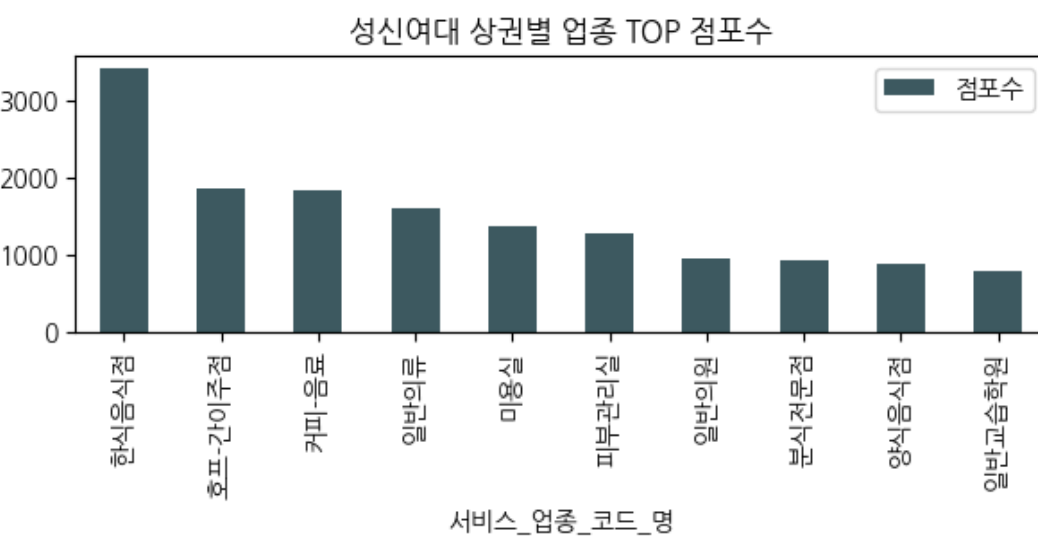
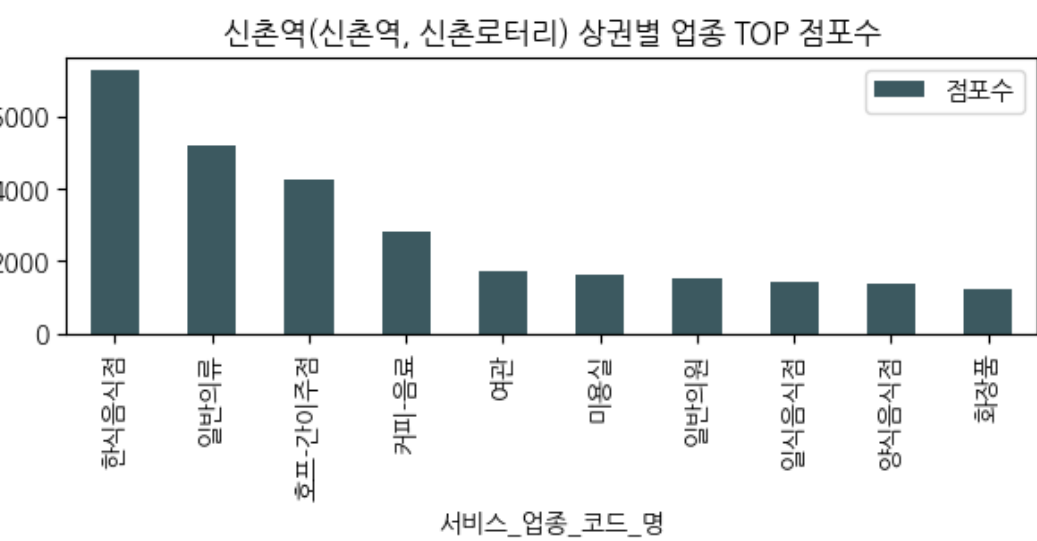
- 점포수, 매출액이 전체 하위 30%인 컷오프 값을 계산
- 점포수·매출 모두 하위 30%이면서, 이미 선정된 6개 후보 (대학가+상업지역)엔 포함되지 않은 상권만 추출
- 이 상권들을 평균 유동인구가 많은 순으로 정렬

chapter 05

데이터 탐색

4.분석을 위한 샘플링

- 대학가 상권 업종 선정하기
- : 점포 수 기준으로 내림차순으로 정렬하여 Top3 선정

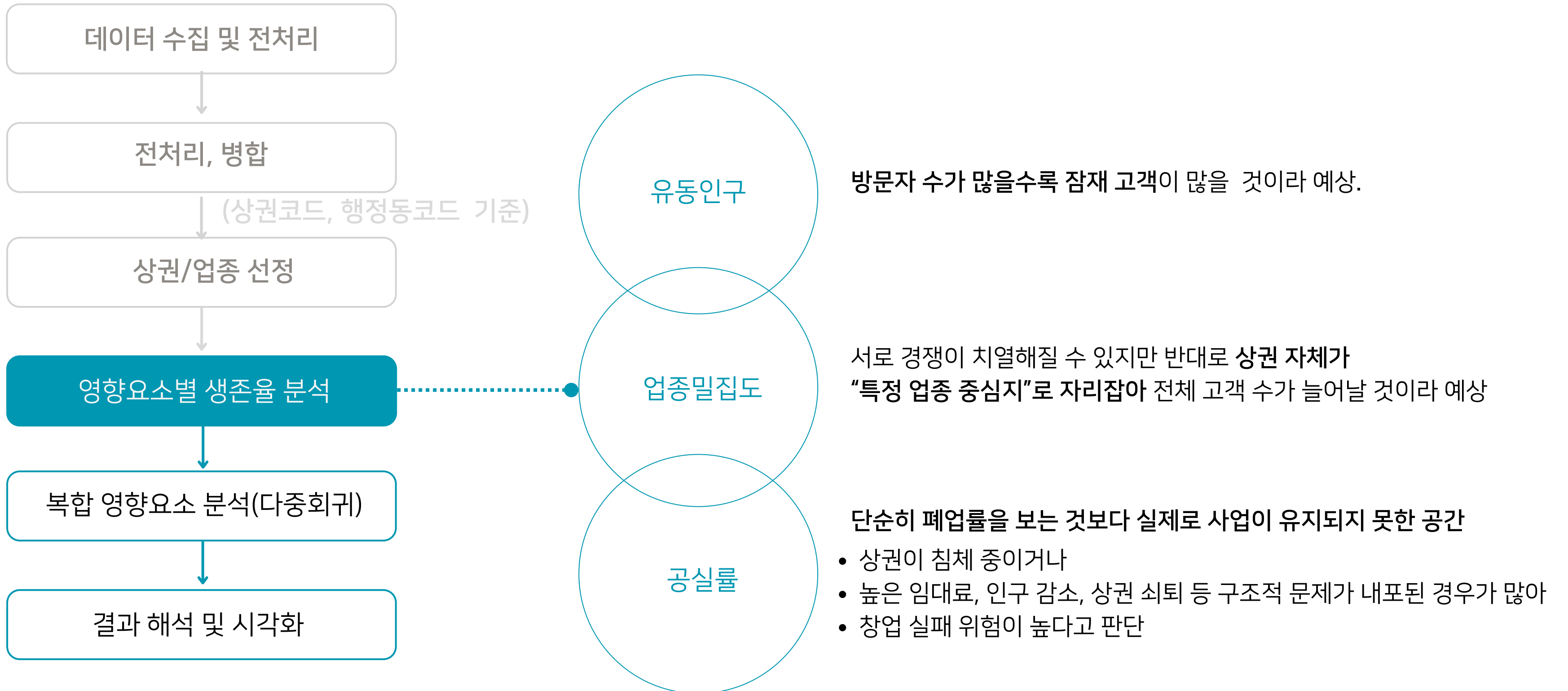


-> 각 카테고리별 대표 상권 3개(총 9개) 자동 선정

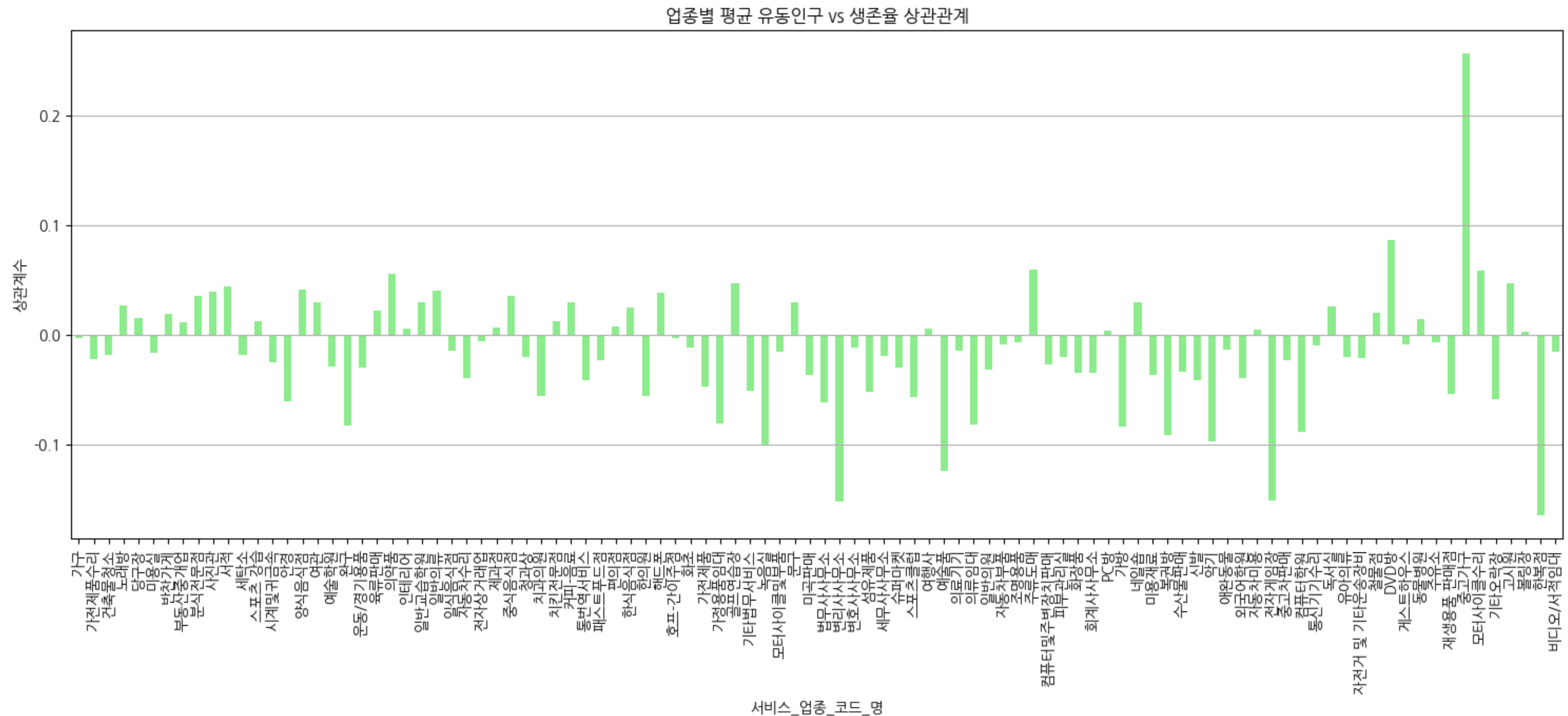
최종 분석 대상: 9개 상권 × 3개 업종 = 27개 조합

대학가			상업지구			베드타운		
신촌역	성신여대	대학로(혜화역)	청평화시장	구로역	용산역	언주역 8번	서대문우체국	길동신동아아파트
한식음식점	한식음식점	한식음식점	일반의류	조명용품	컴퓨터및주변장치판매	미용실	한식음식점	일반교습학원
일반의류	호프-간이주점	커피-음료	슈퍼마켓	일반의류	가전제품	일반의류	커피-음료	일반의류
호프-간이주점	커피-음료	호프-간이주점	부동산중개업	한식음식점	조명용품	화장품	세탁소	사진관

창업 성공(생존)률에 영향을 미치는 주요 변수: 유동인구, 업종밀집도, 공실률

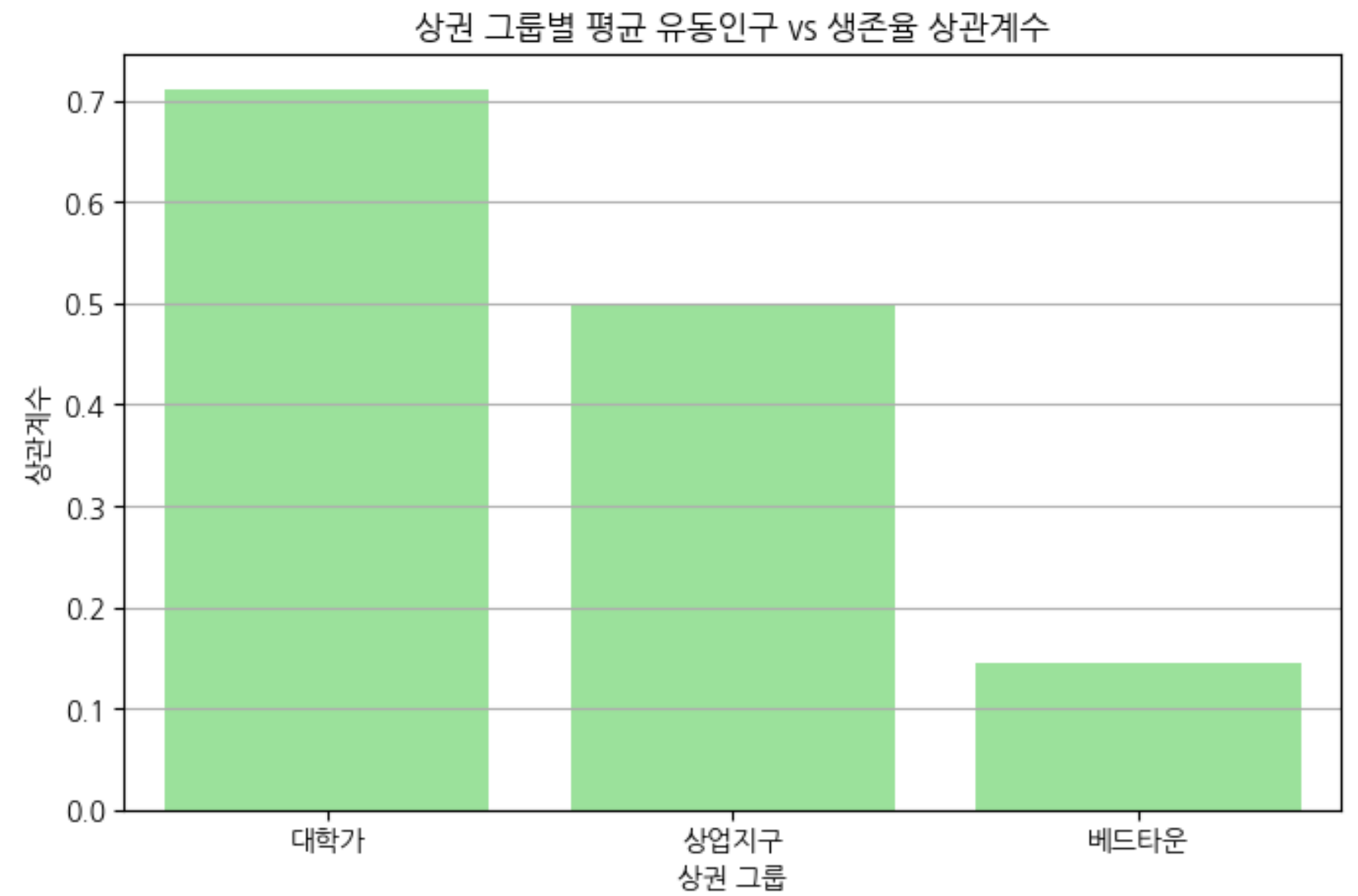
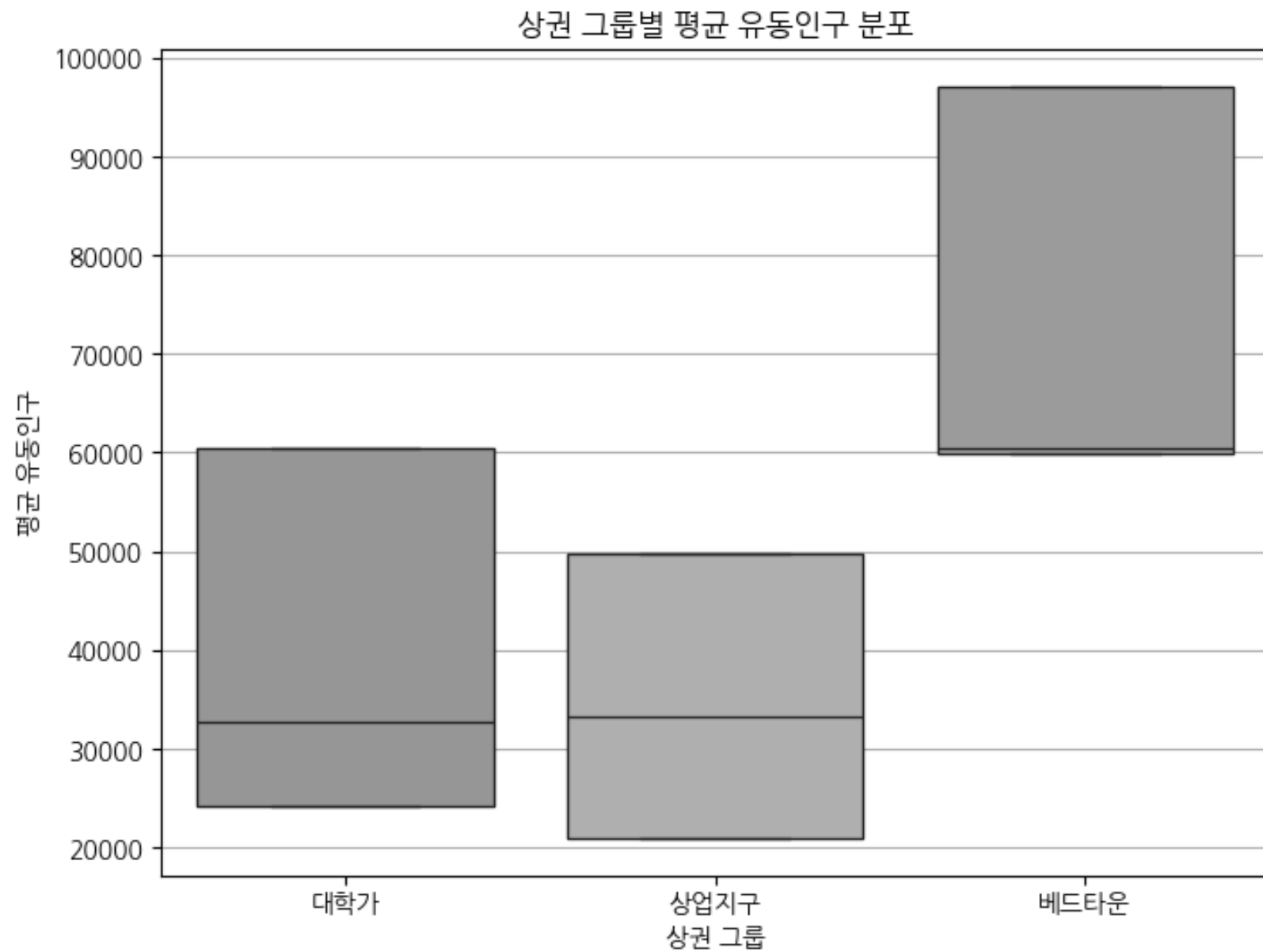


27개 샘플 전체의 상관계수 시각화



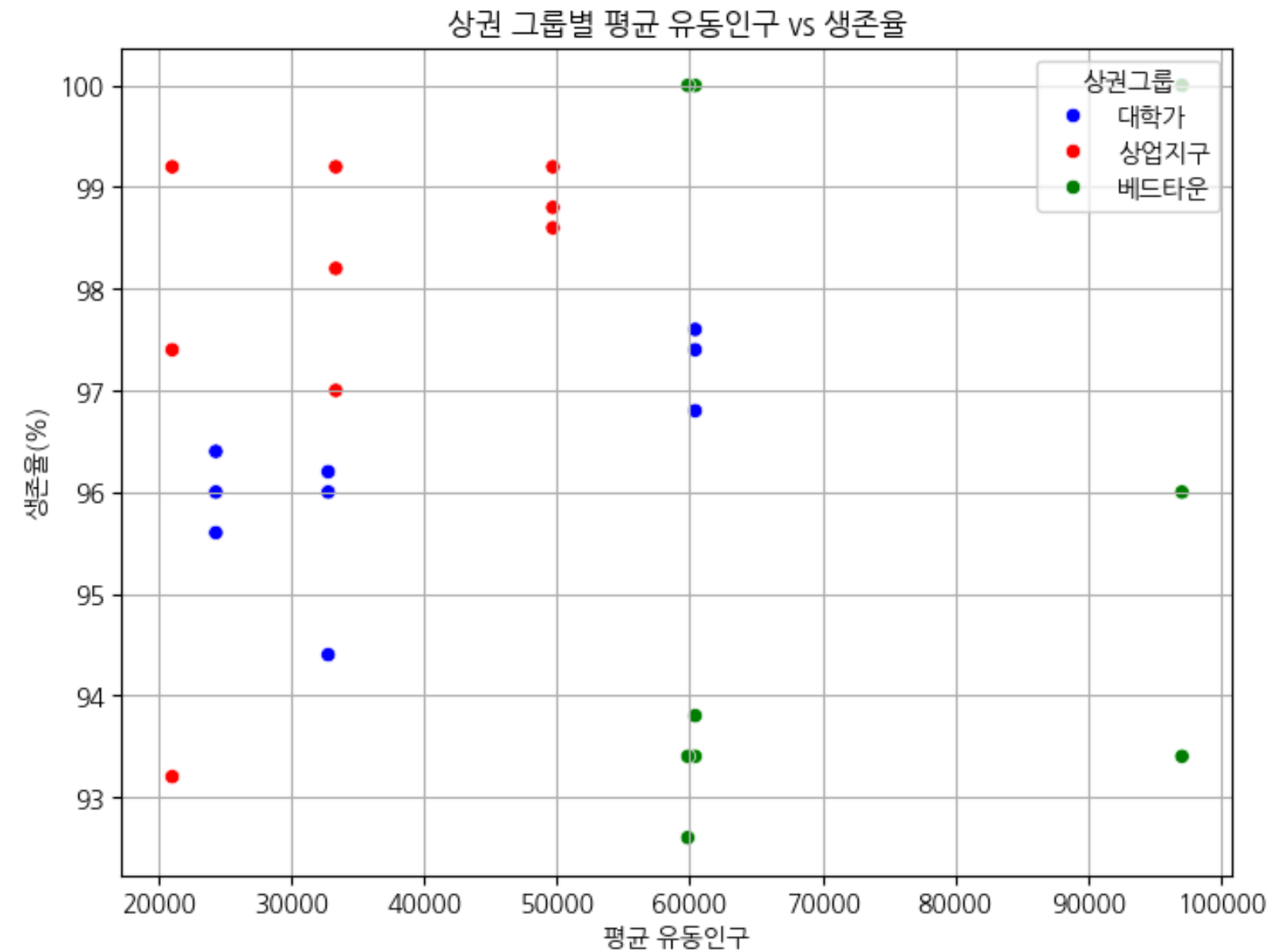
창업 성공(생존)률에 영향을 미치는 주요 변수: 유효인구, 업종밀집도, 공실률

상권별 파악:



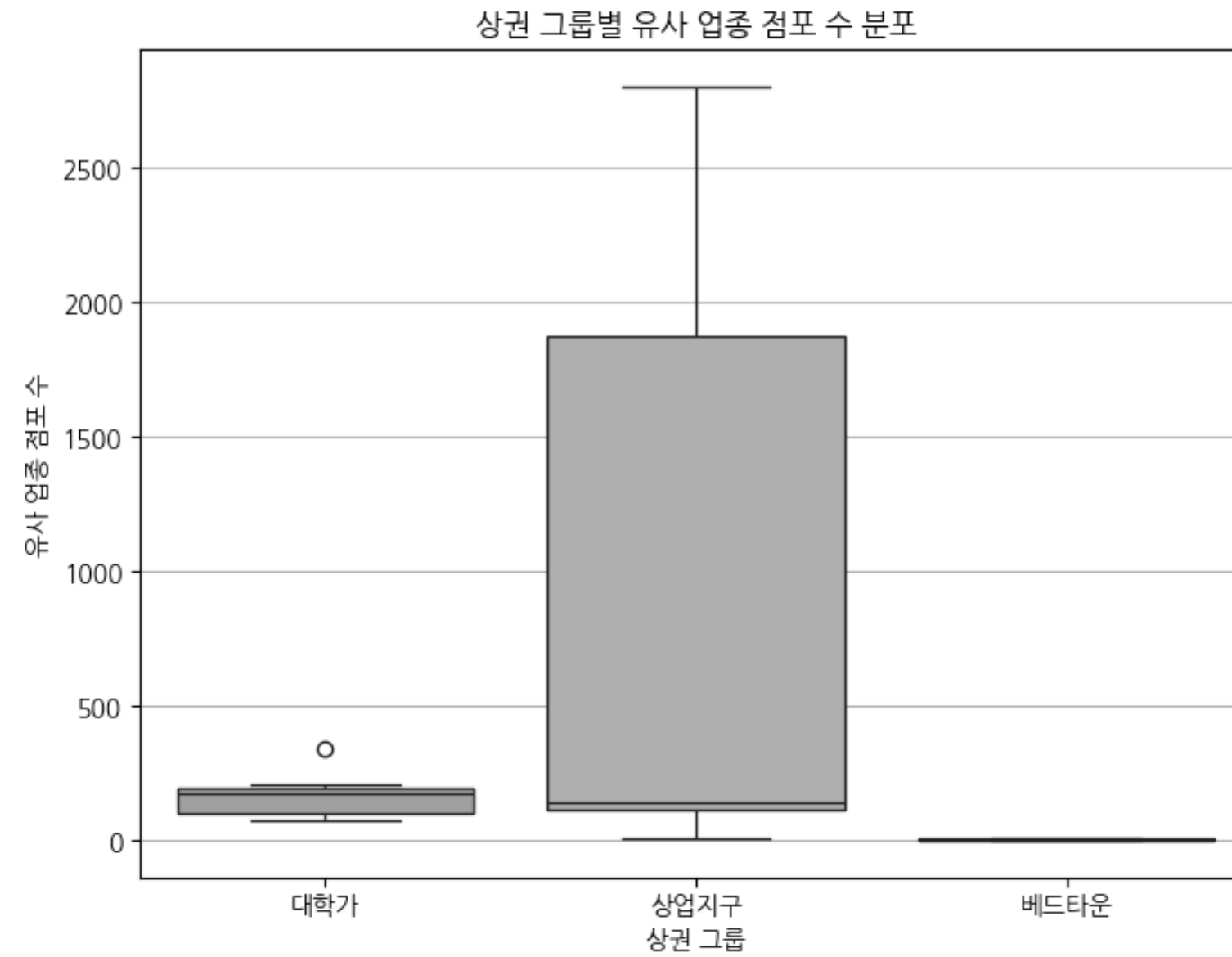
창업 성공(생존)률에 영향을 미치는 주요 변수: 유동인구, 업종밀집도, 공실률

창업 생존율(%) (상권-업종별 실제 폐업률 기반)



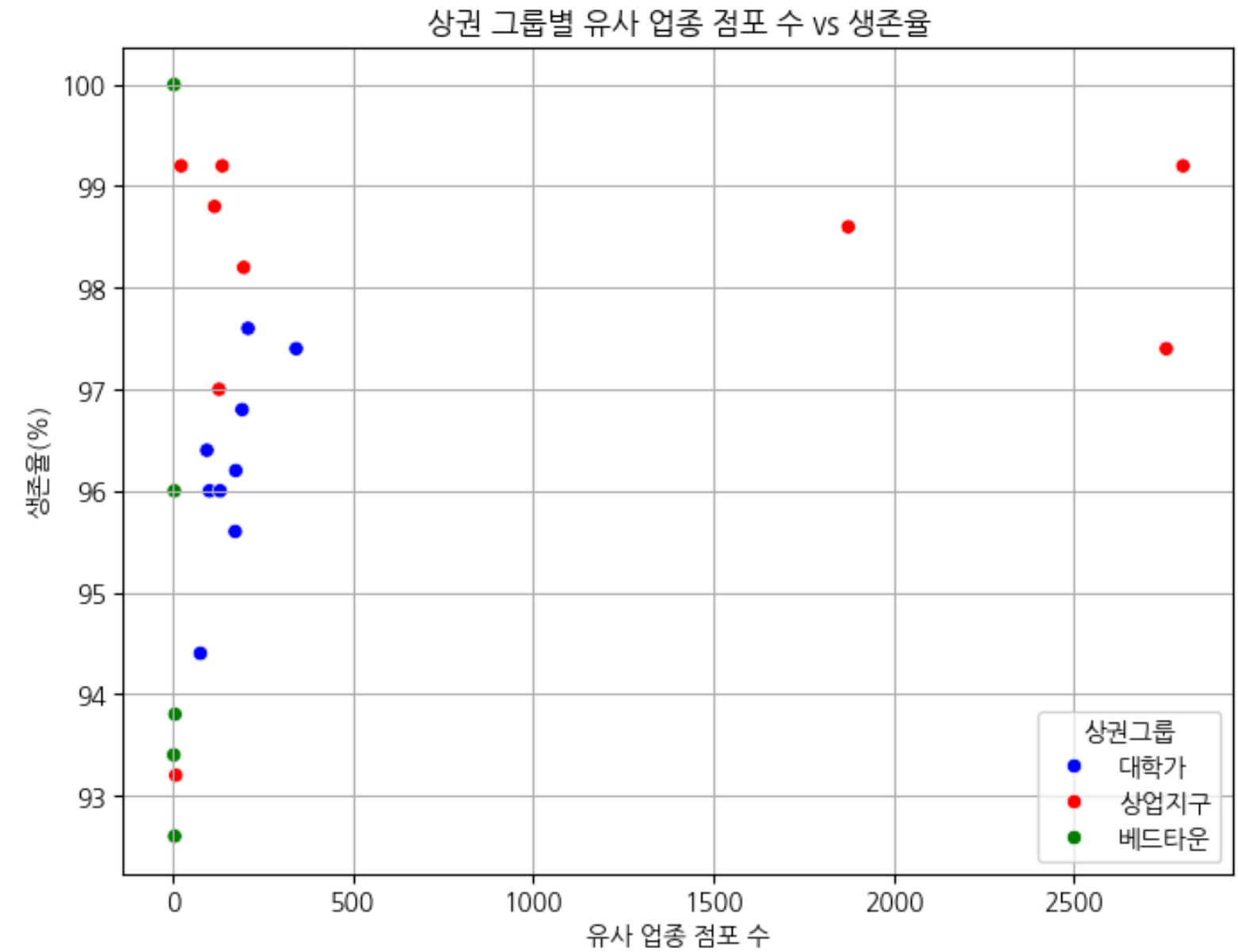
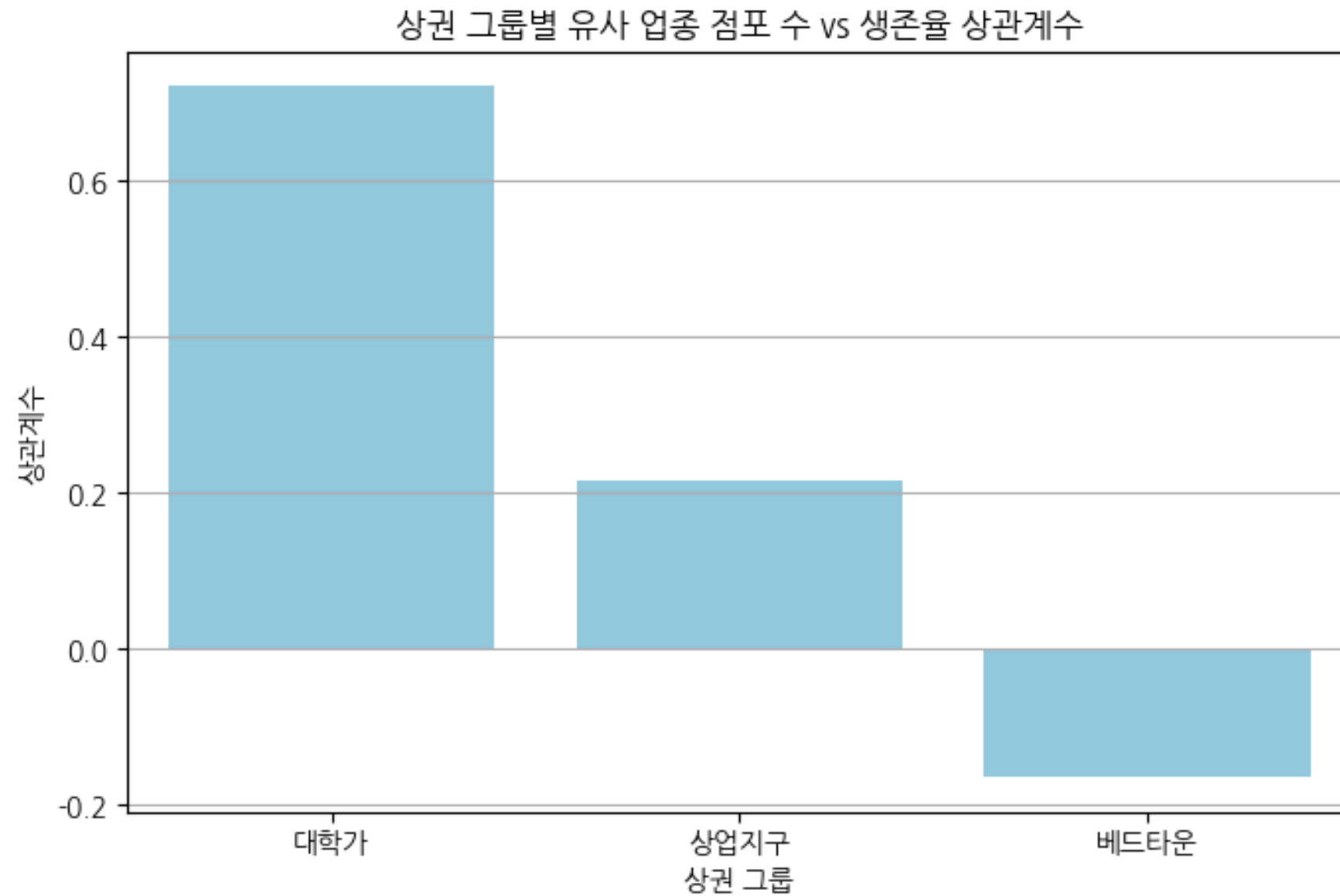
창업 성공(생존)률에 영향을 미치는 주요 변수: 유동인구, 업종밀집도, 공실률

상권별 파악:

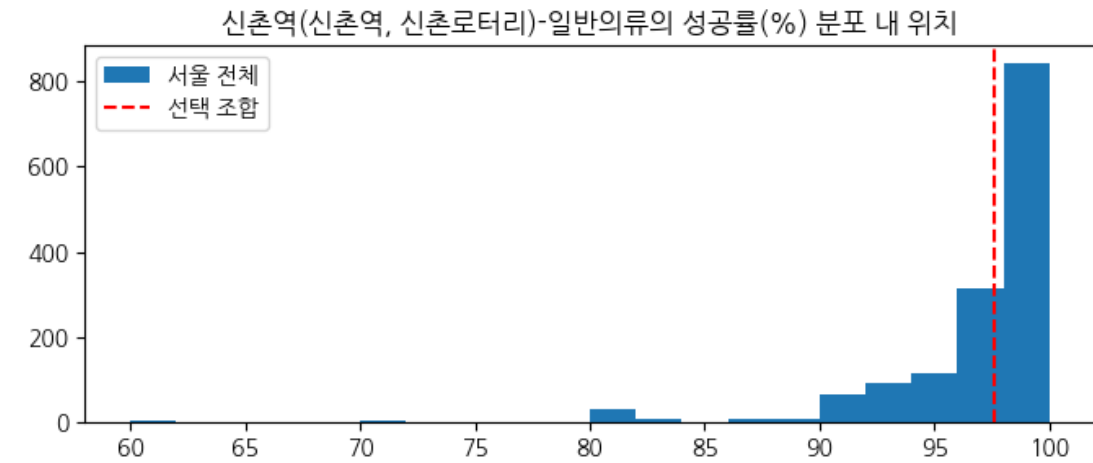
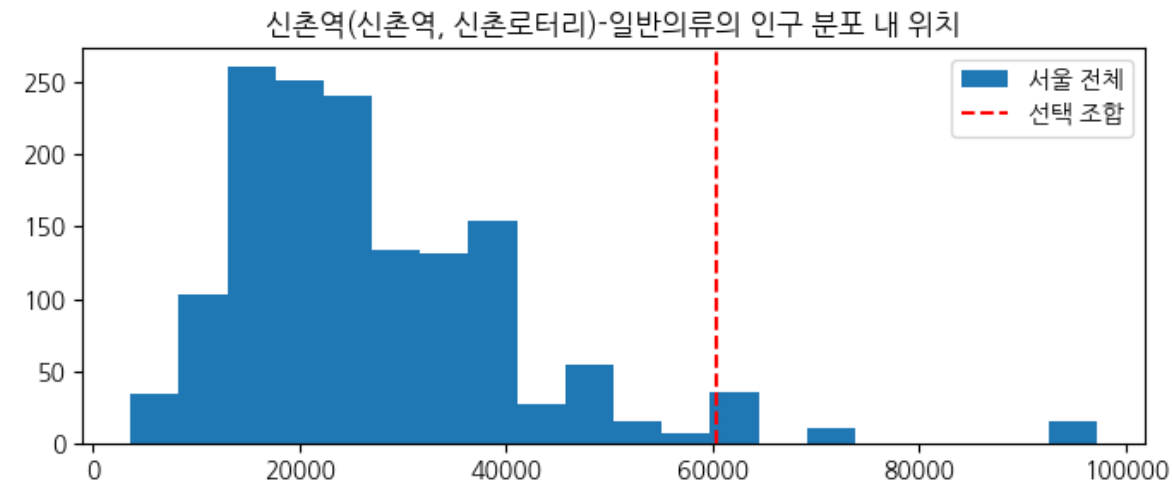


창업 성공(생존)률에 영향을 미치는 주요 변수: 유동인구, 업종밀집도, 공실률

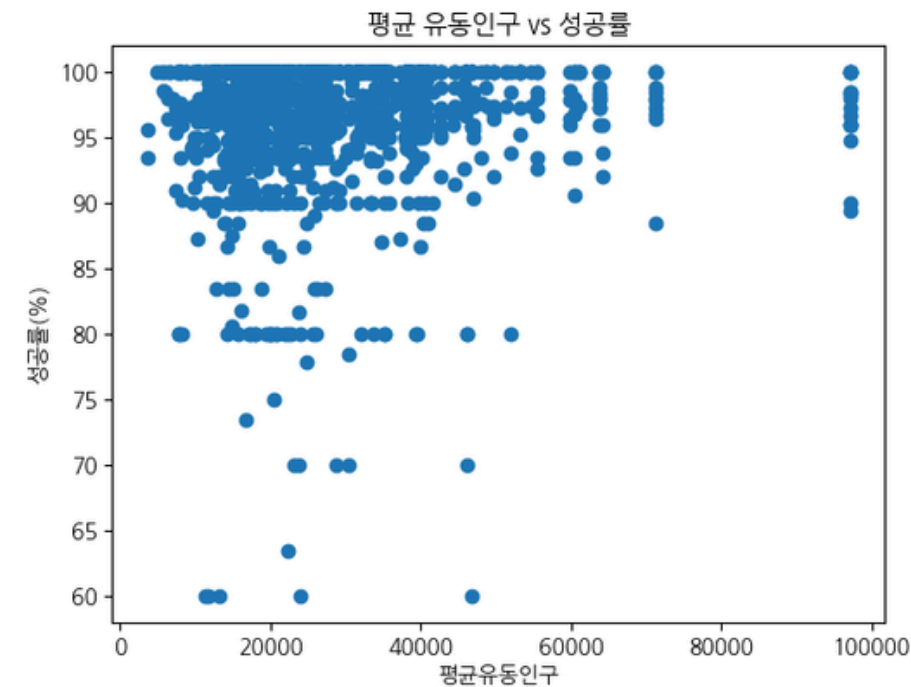
상권별 파악:



창업 성공(생존)률에 영향을 미치는 주요 변수: 유동인구, 업종밀집도, 공실률

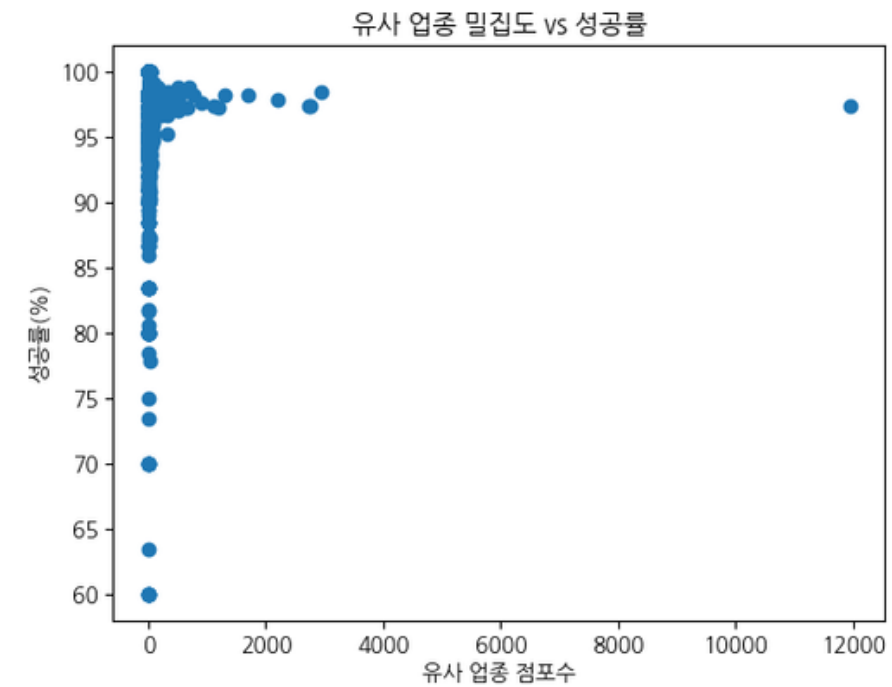


상권 **신촌역**
상권명 **일반의류**



상관계수: 0.012236202482510427

유동인구

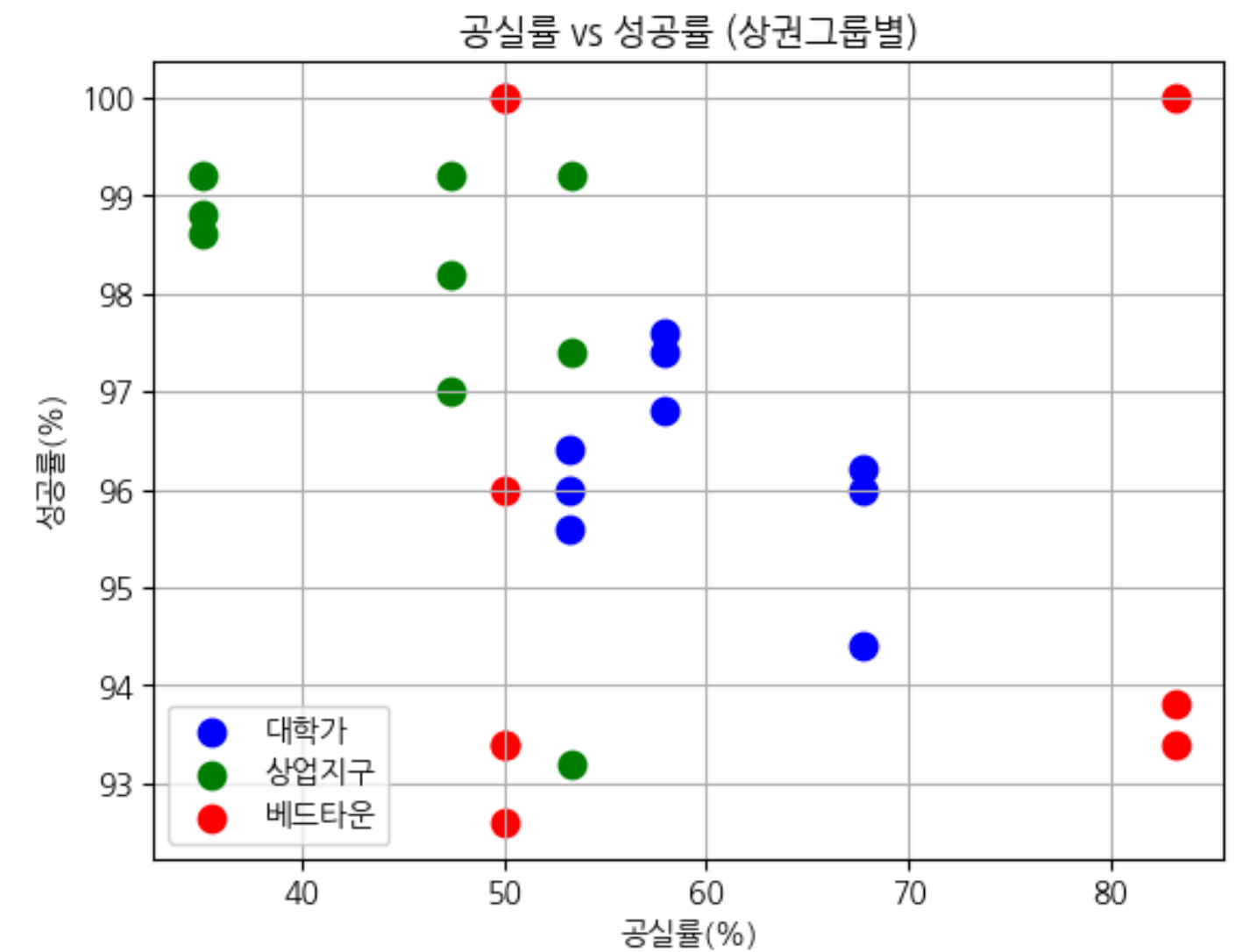
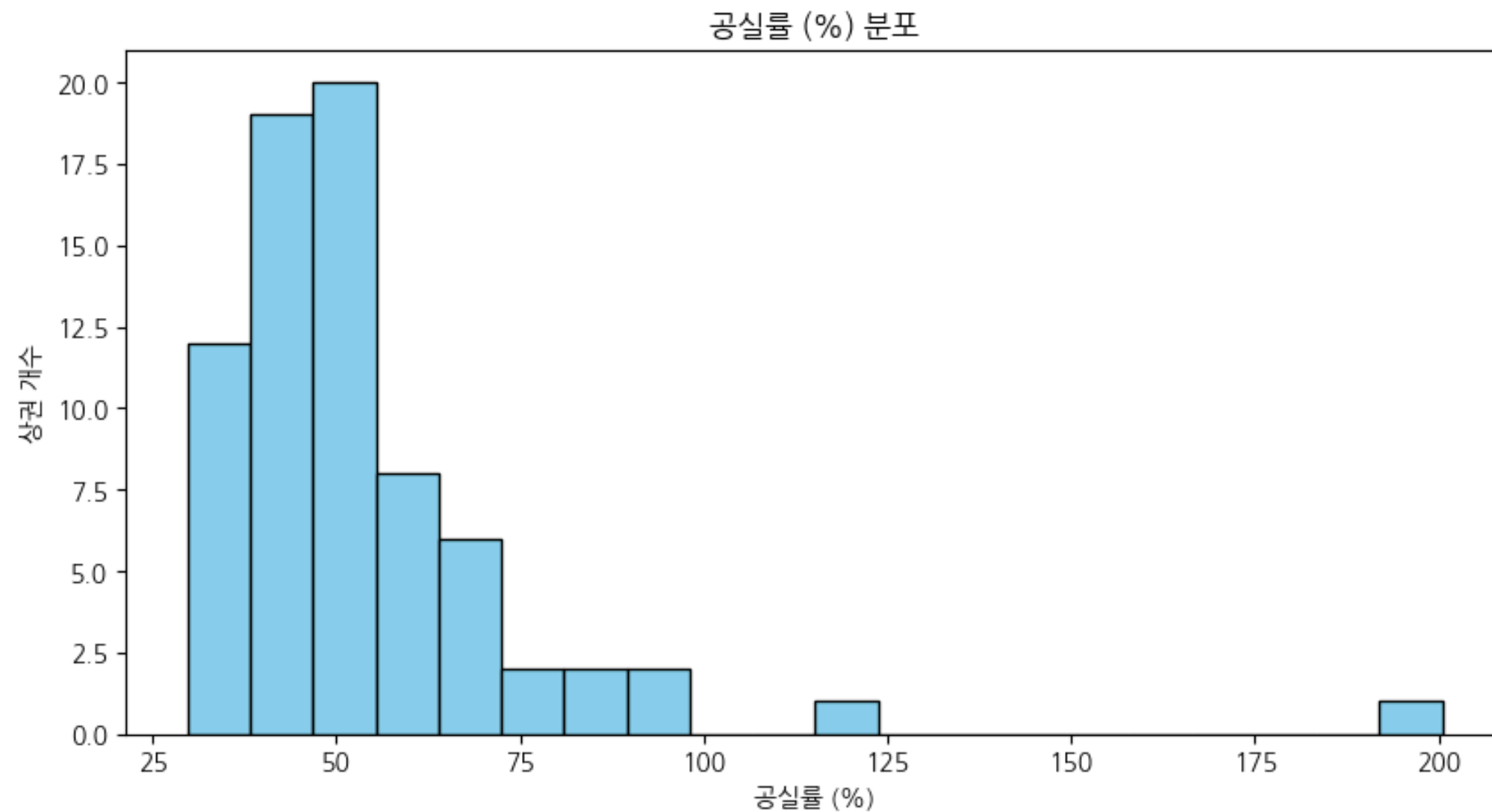


상관계수: 0.04057937912976923

업종밀집도

창업 성공(생존)률에 영향을 미치는 주요 변수: 유동인구, 업종밀집도, **공실률**

	공실률(%)	생존율(%)
대학가	59.6	96.266667
베드타운	61.066667	95.844444
상업지구	45.266667	97.866667

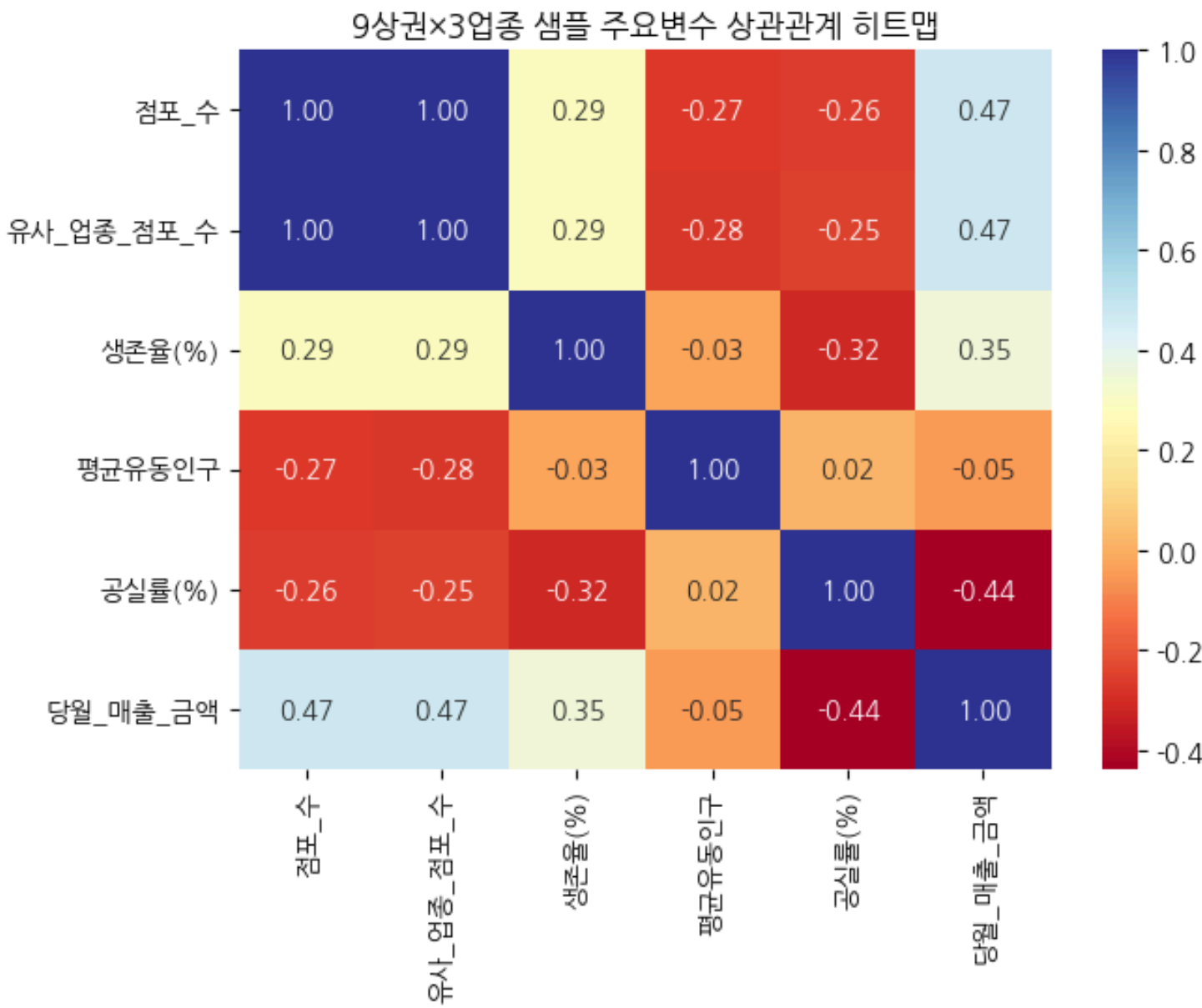


전체적으로 공실률과 생존율은 상권별로 상반된 패턴

변수별 상관관계 분석

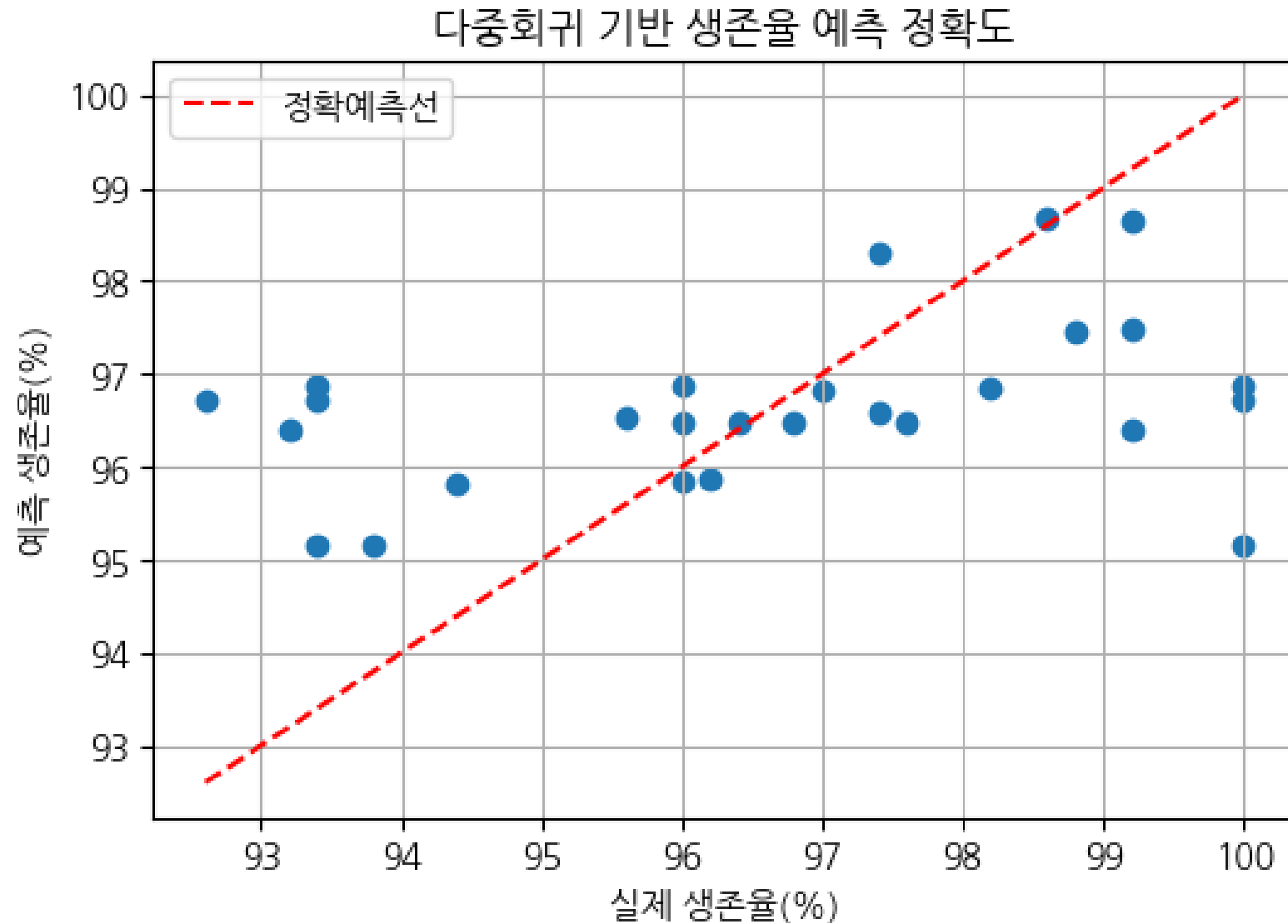
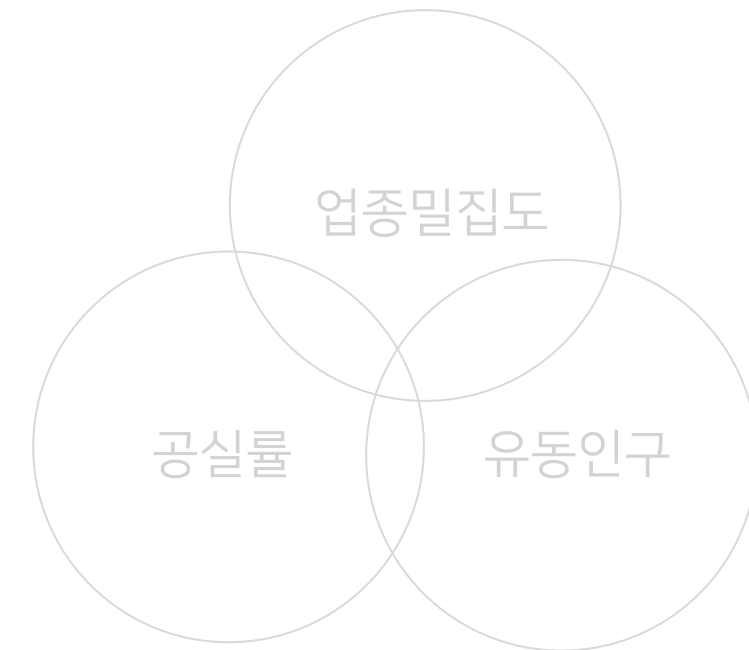
생존율과의 상관계수	유사_업종_점포_수	평균유동인구	공실률
대학가	0.721217433072597	0.710959459851494	-0.360118
상업지구	0.216766296992153	0.500181016108926	-0.479366
베드타운	-0.163819394916481	0.144300103236919	-0.025647
	0.285746	-0.027102	-0.321439

- 대학가는 유사업종이 많고, 유동인구가 많을수록 생존율이 높다.
 - 상업지구는 공실률이 높을수록 생존율이 크게 낮아지는 경향이 뚜렷하다.
 - 베드타운은 변수(인구, 경쟁, 공실률)와 생존율 간의 뚜렷한 상관관계가 거의 없다.
- > 업종 밀집도가 생존율과 가장 높은 양의 상관관계를 보임



다중회귀분석

단일 변수 영향뿐 아니라, 여러 변수의 복합적 영향력을 확인하기 위해 다중회귀분석 모델을 도입



$$\text{생존율} = a \cdot \text{유동인구} + b \cdot \text{업종밀집도} + c \cdot \text{공실률} + d$$

결과

- 평균유동인구: 0.0000004
- 유사업종수: 0.000687
- 공실률: -0.047298

평균유동인구와 유사업종수는 생존율에 아주 작은 영향
공실률은 1% 오를 때 생존율이 약 0.05% 감소함.

전체적으로 공실률이 생존율에 상대적으로 더 큰 영향.

종합적인 요소로 생존율 파악

사건 코드	서비스_업종 코드명	생존율(%)	예측_생존율(%)	예측_생존율
0	신촌역(신촌역, 신촌로터리)	한식음식점	97.4	96.577776
1	신촌역(신촌역, 신촌로터리)	일반의류	97.6	96.485938
2	신촌역(신촌역, 신촌로터리)	호프-간이주점	96.8	96.474664
3	성신여대	한식음식점	95.6	96.530744
4	성신여대	호프-간이주점	96.4	96.476714
5	성신여대	커피-음료	96	96.481663
6	대학로(혜민교차로, 혜민교차로)	한식음식점	96.2	95.881903
7	대학로(혜민교차로, 혜민교차로)	커피-음료	96	95.852069
8	대학로(혜민교차로, 혜민교차로)	호프-간이주점	94.4	95.814399
9	청평화시장	일반의류	97.4	98.287801



결과 요약

인구밀도

- 대학가: 0.71 (강한 양의 상관관계)
- 상업지구: 0.50 (중간 정도)
- 베드타운: 0.14 (관계 약함)

“유동인구는 성공률 예측에 특히 대학가·상업지구에서 유의미한 영향”

유사업종

- 대학가: 0.72
- 상업지구: 0.22
- 베드타운: -0.16

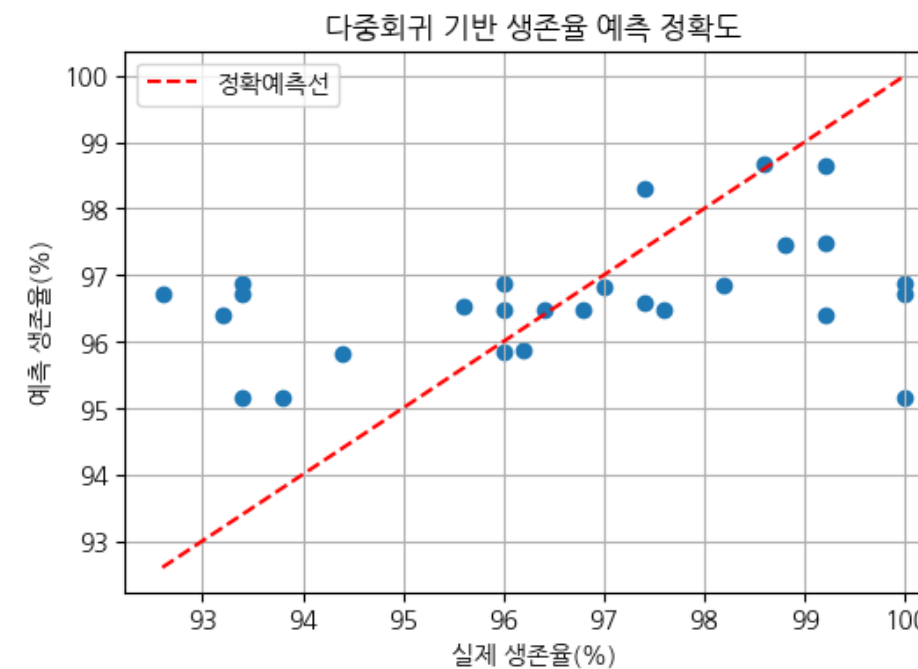
“경쟁이 심한 곳이 오히려 트렌드가 형성되어 살아남는
경우 존재”

공신풀

- 대학가: -0.36
- 상업지구: -0.48
- 베드타운: -0.03

“공실률이 상권의 건강도를 직접적으로 나타내며, 특히
상업지구에서 민감”

인구밀도 + 유사업종 + 공실률



- 평균유동인구: 0.000004
- 유사업종수: 0.000687
- 공실률: -0.047298

상권별(대학가, 상업지구, 베드타운)

상권별로 유동인구·업종 밀집도·공실률의
영향력이 다르게 나타남

대학가는 **유동인구·업종 밀집도**가 성공률에 가장 큰 영향, 상업지구는 **공실률 영향**이 두드러짐

베드타운은 변수 영향력 약함, 외부 환경요소가 더 중요할 수 있음

가설검증

가설1 유동인구가 많을수록 상권·업종의 생존율이 높을 것이다.

~~유사업종 밀집도와 생존율의 상관관계는 확인되지 않음~~

가설2 동일 업종(유사업종) 밀집도가 높을수록 생존율이 낮을 것이다.

유동인구가 많은 지역일수록 창업 생존율이 높아지는 경향

가설3 공실률이 높을수록 상권·업종의 생존율이 낮을 것이다.

공실률이 높은 지역일수록 실제로 생존율이 낮아지는 경향

결론 및 시사점

1) 분석 결과, 창업 성공에 가장 영향력이 큰 변수
는 유동인구와 공실률임이 확인됨

2) 특정 상권(대학가 등)에서는 업종 밀집
도도 긍정적 효과

-> 유사업종 밀집도와 공실률 = 생존율 파악

입력값 기반 상권/업종별 분석으로,
예비창업자는 내가 원하는 곳의 생존율과 위험도를
즉시 파악할 수 있습니다.
반복 공실 구역에는 정책적 재생 또는 창업 리스크 경보 등
새로운 대안도 제시할 수 있습니다

> 공실(빈 점포) 데이터 활용

단순히 '폐업률'이 아닌,
'실제로 비어 있는 공간'을 효율적으로 탐색
반복 공실 구역을 리스크 알림 또는
"공방, 공유주방, 리페어숍 등" 재생 정책 추천 가능

> 맞춤 샘플링 기반

27개(9상권×3업종) 샘플로
지역·업종 맞춤형 분석이 가능
→ 나만의 조건, 나만의 상권에서
창업 성공 가능성을 객관적으로 진단할 수 있음

[선택 가능한 상권]
1. 구로역
2. 길동신동아파트
3. 대학로(해화역)
4. 서대문우체국
5. 성신여대
6. 신촌역(신촌역, 신촌로터리)
7. 연주역 8번
8. 용산전자상가(용산역)
9. 청평화시장
분석할 상권명을 입력하세요 (n:종류): 신촌역(신촌역, 신촌로터리)
[해당 상권 내 업종]: '일반의류' '한식음식점' '호프-간이주점'
분석할 업종명을 입력: 일반의류

> 입력값 기반 분석

사용자가 원하는 상권·업종을 직접 입력하여
"해당 조합의 생존율, 경쟁도, 유동인구, 공실률"
등 구체적 수치를 즉시 파악할 수 있음