



서울시 1인 가구 사업 제안

강우석, 강지철, 강태민,
김세영, 이가연, 이정은, 이주희

목 차



연구 목적 및 개요

- 연구 목적
- 데이터 선정
- 데이터 분석 방법



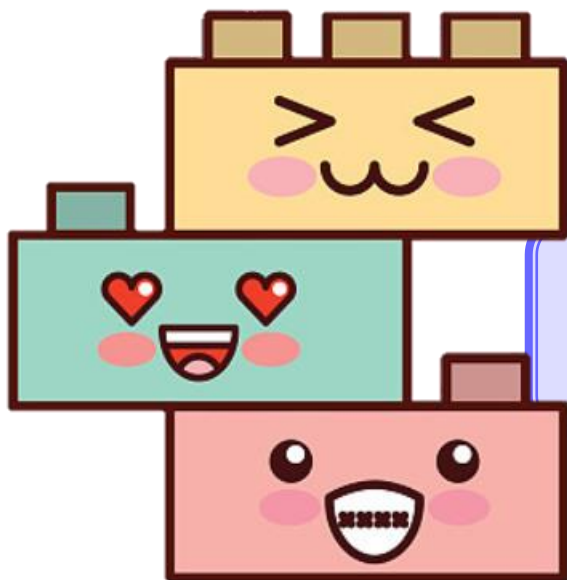
데이터 분석

- 1인 가구 밀집도 등급
- 행운동 상권 분석
- 워드 클라우드
- 기타 요인



결론

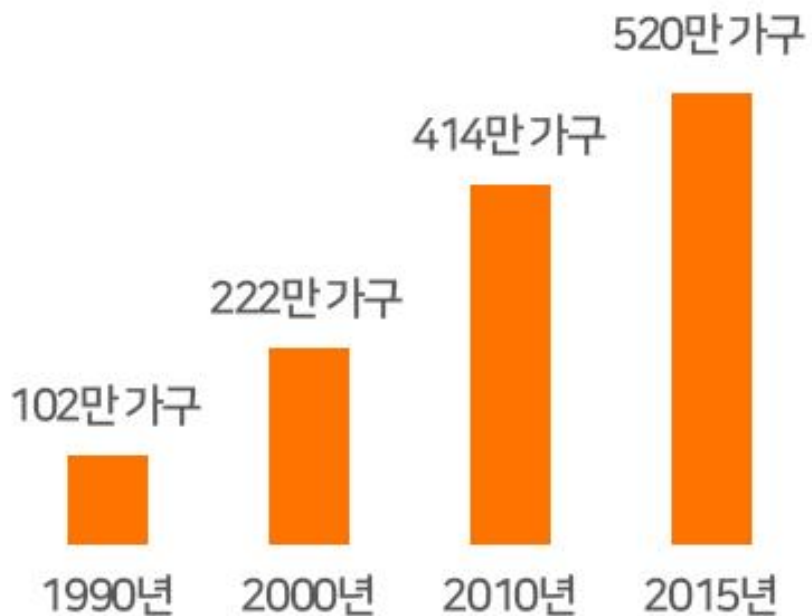
- 요약 및 결과
- 추가 및 보완점



연구 목적 및 개요

왜 1인 가구에 주목하는가?

우리나라 1인가구수



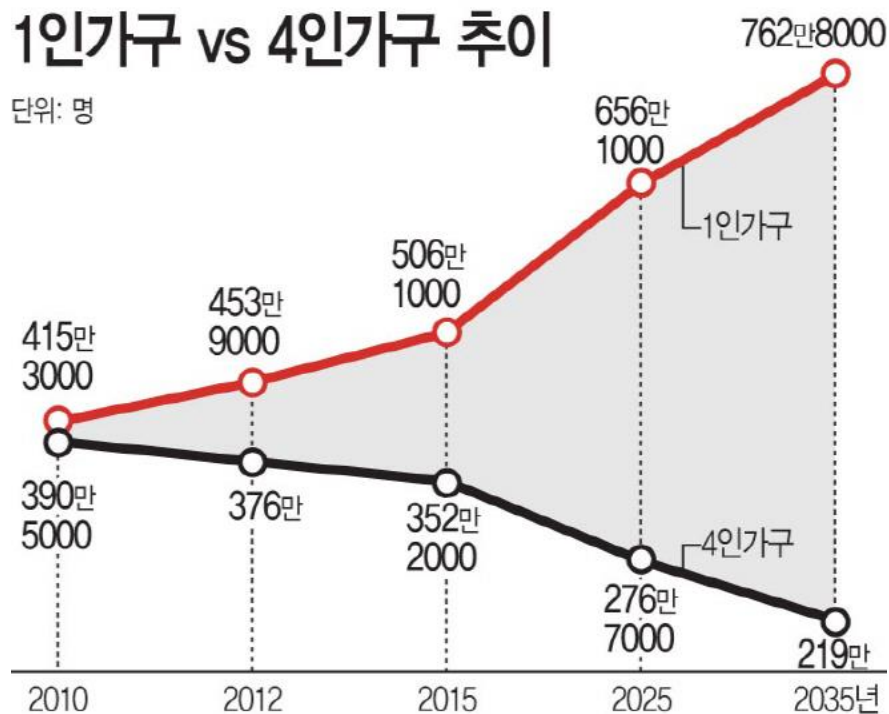
통계청의 장래인구 전망에 의하면 2035년에는
1인가구가 총 가구의 34.3%가 될 것으로 예측된다.

덜 버는 1인 가구, 2~3인 가구보다 돈 더 쓴다...소비시장 '큰 손'

f t [icon] 최종수정 2015.08.26 09:52 기사입력 2015.08.26 08:36 댓글 쓰기

1인가구 vs 4인가구 추이

단위: 명



자료: 통계청

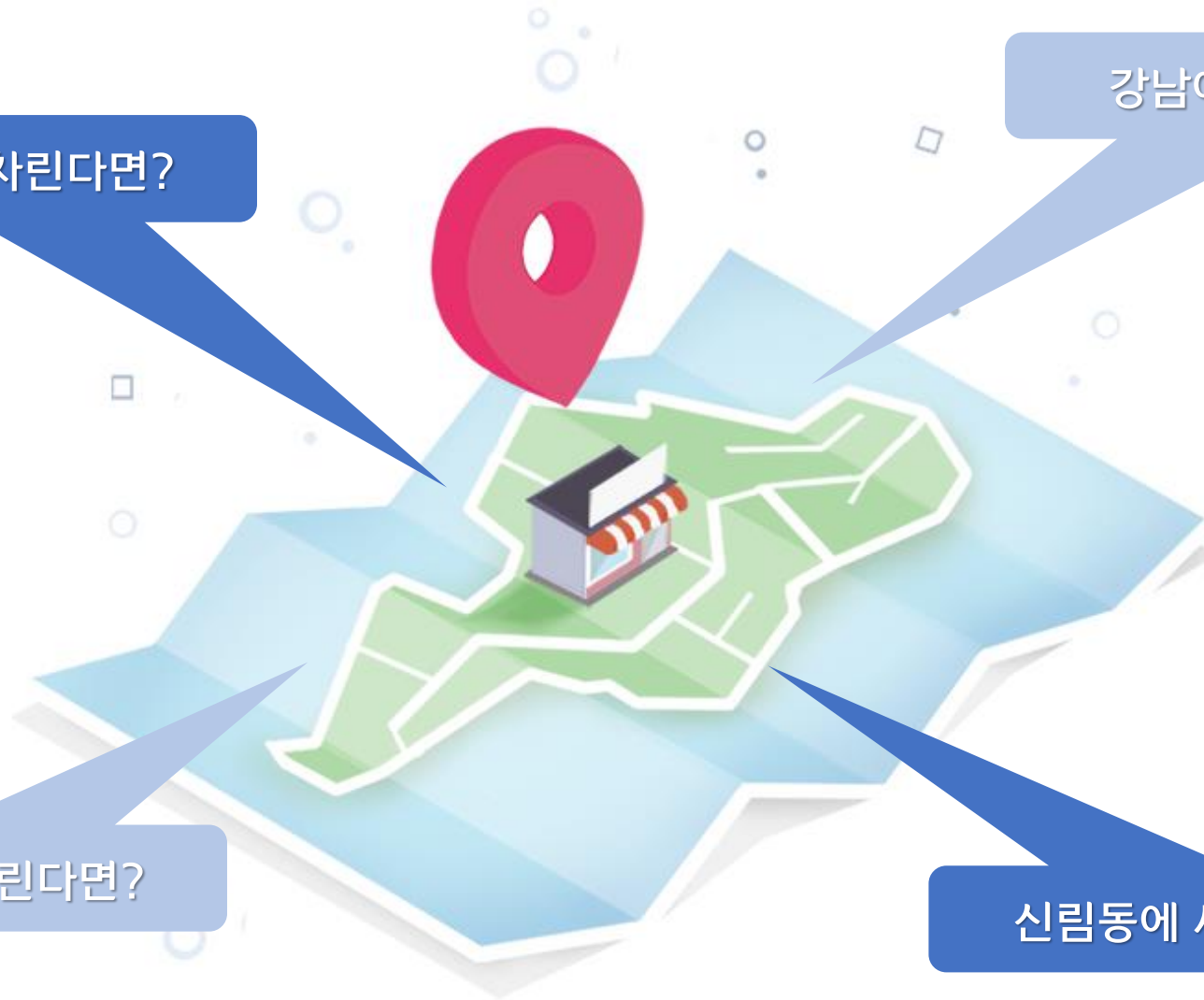
만약에 1인 가구를 타겟으로 창업을 한다면?

종로구에 카페를 차린다면?

강남에 식당을 차린다면?

노원구에 학원을 차린다면?

신림동에 세탁소를 차린다면?



특정 사업^{ex)카페} 을 창업하고자 한다면?

어느 지역에 **1인 가구**가 많을까?

어느 지역에 **유동인구**가 많을까?

경쟁업체 수가 얼마나 될까?

소비자의 **구매력**은?



특정 사업^{ex)카페} 을 창업하고자 한다면?

어느 지역에 **1인 가구**가 많을까?

어느 지역에 **유동인구**가 많을까?

통계 데이터를 이용하여
1인 가구 타겟 비즈니스의
맞춤형 상권 추천 시스템을 제안해보자!

경쟁업체 수가 얼마나 될까?

소비자의 **구매력**은?

데이터 선정



서울시 인구통계



서울시 상권



네이버 블로그

데이터 분석 방법

Python



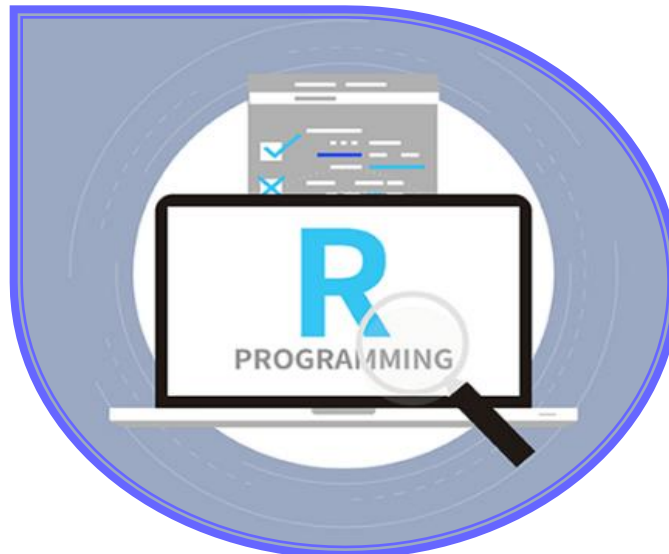
**Data
Modeling**

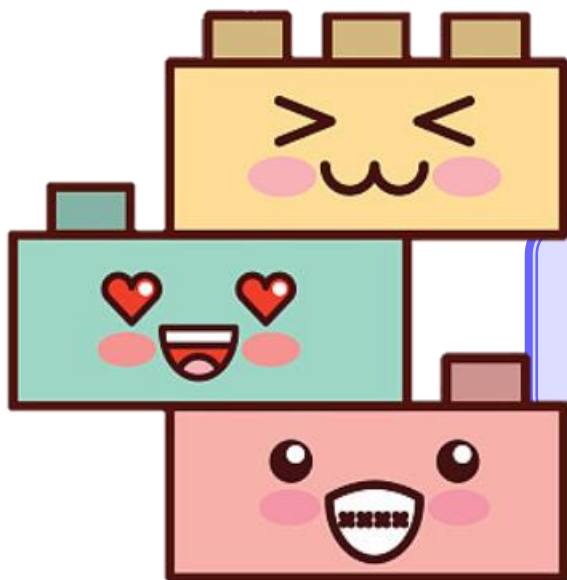


Word Cloud



R





데이터 분석

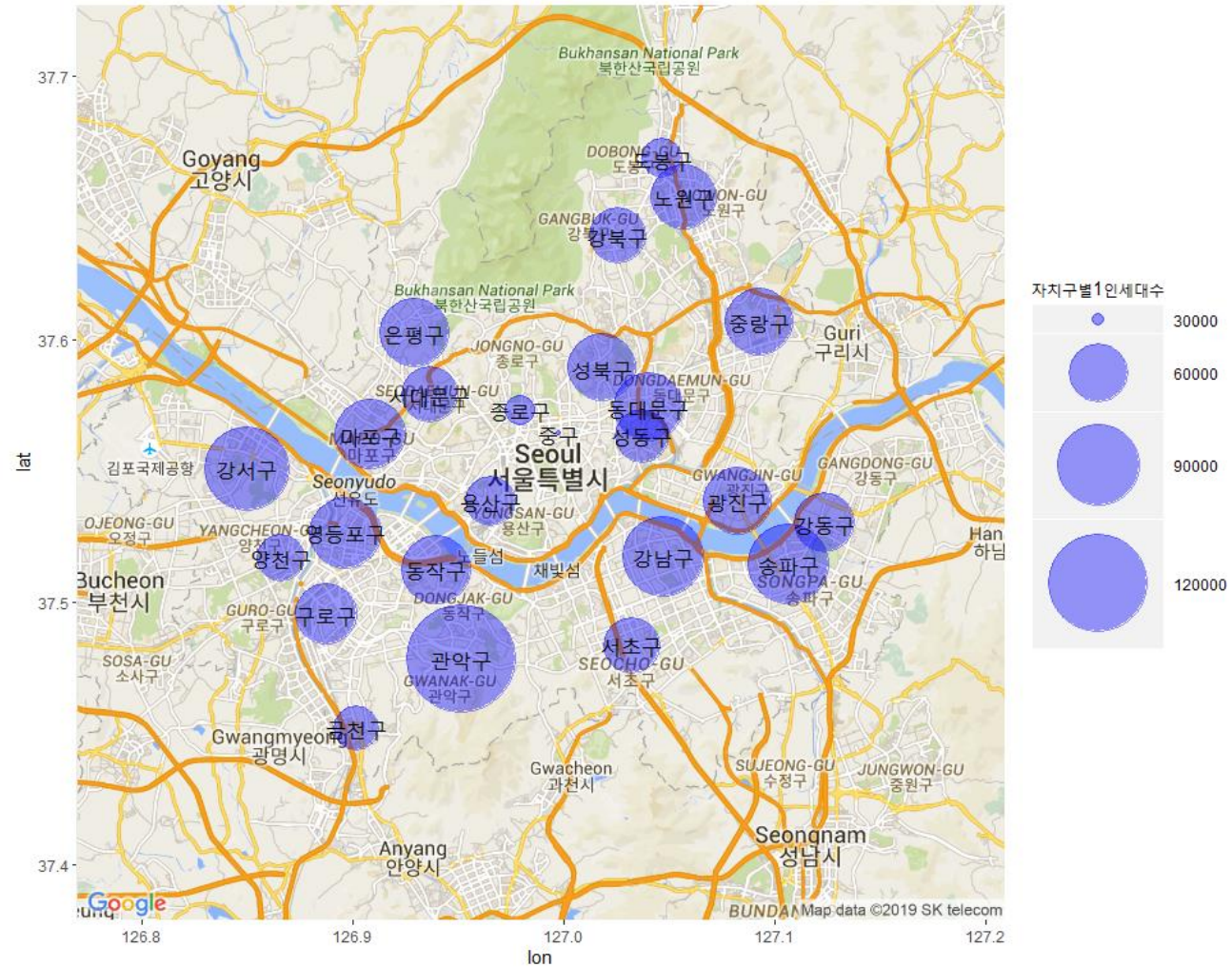
서울시 1인 세대 분포 조사

기간	자치구	행정동	전체세대수	1인세대	2인세대
	종로구	합계	4,315,921	1,740,372	911,696
		소계	73,942	36,591	14,055
		사직동	4,557	2,236	894
		삼청동	1,397	743	257
		부암동	4,268	1,626	940
		평창동	7,518	2,420	1,689
		무악동	3,177	740	690
		교남동	4,581	1,636	1,081
		가회동	2,056	930	433
		종로1,2,3,4가동	5,275	4,176	523
		종로5·6가동	3,231	2,172	500
		미화동	4,473	2,689	797
		창신1동	2,937	1,721	563
		창신2동	4,119	1,905	886
		창신3동	2,898	903	668
		송인1동	2,978	1,354	645
		송인2동	5,743	3,759	907

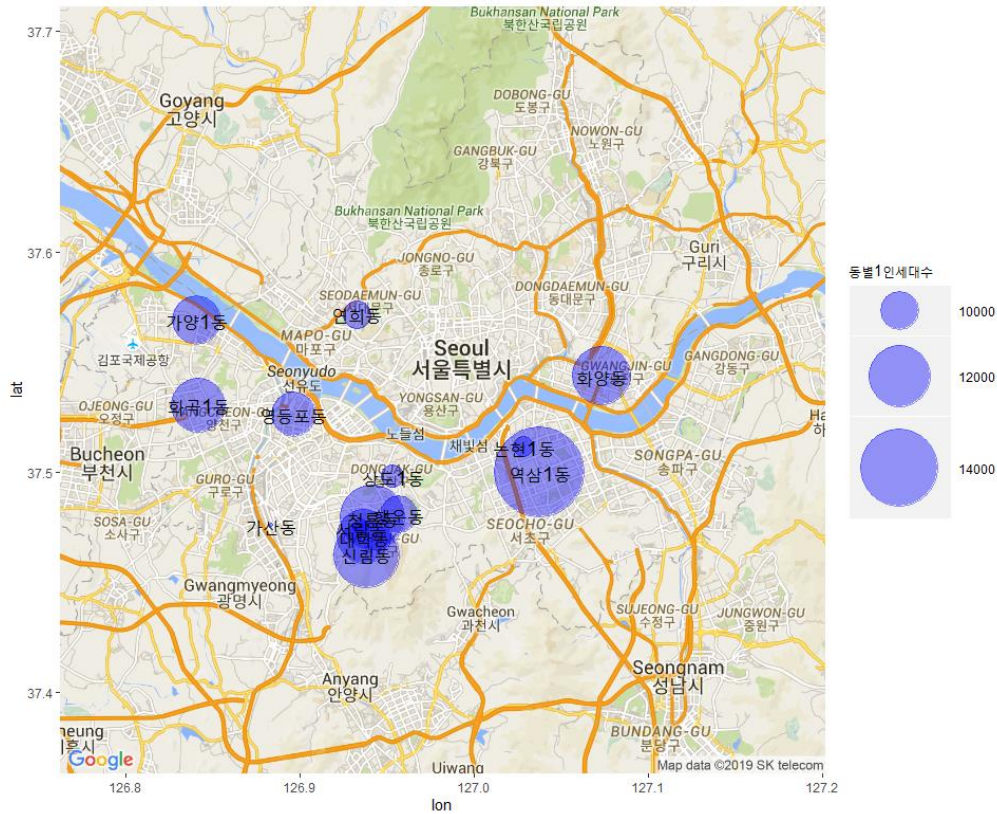
동	면적	
	면적	구성비(%)
	605,24	100
소계	23,91	4
사직동	1,23	0,2
삼청동	1,49	0,2
부암동	2,27	0,4
평창동	8,87	1,5
무악동	0,36	0,1
교남동	0,35	0,1
가회동	0,54	0,1
종로1,2,3,4가동	2,35	0,4
종로5·6가동	0,6	0,1
미화동	0,78	0,1
창신1동	0,31	0,1
창신2동	0,26	0
창신3동	0,23	0
송인1동	0,23	0
송인2동	0,35	0,1

-2018년 서울시의 구/동별 1인세대수와 각 지역의 면적 데이터-

서울시 1인 세대 분포 조사 (구별)



서울시의 1인세대수 상위 15개동

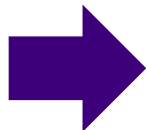


1. 역삼동
2. 신림동
3. 청룡동
4. 화양동
5. 대학동
6. 화곡1동
7. 가양1동
8. 영등포동
9. 행운동
10. 연희동

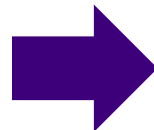
11. 용신동
12. 상도1동
13. 논현1동
14. 서림동
15. 가산동

서울시 1인 세대 분포 조사

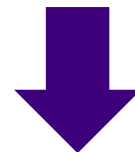
X1인세대수	면적
161.8333	1.34
2739.0833	0.96
4165.0833	1.13
3821.6667	0.40
8862.2500	2.52



면적당1인세대수
120.7711
2853.2118
3685.9145
9554.1667
3516.7659



```
#표준화  
std <- function(X){  
  return((X-min(X))/(max(X)-min(X)))  
}  
stdconc <- apply(stdconc, 2, std)
```



ST면적당1인세대수
0.000000000
0.107621861
0.140419283
0.371550459

서울시 1인 세대 분포 조사

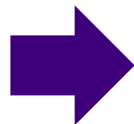
ST면적당1인세대수

0.000000000

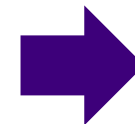
0.107621861

0.140419283

0.371550459



```
quantile(conc2[,6], probs=seq(0,1, length=6))
conrank <- c()
nrow <- nrow(conc2)
for (i in 1:nrow){
  if (conc2$ST면적당1인세대수[i] >= 0.23400820){
    conrank[i] <- "매우높음"
  } else if (conc2$ST면적당1인세대수[i] >= 0.15879761){
    conrank[i] <- "높음"
  } else if (conc2$ST면적당1인세대수[i] >= 0.11277217){
    conrank[i] <- "보통"
  } else if (conc2$ST면적당1인세대수[i] >= 0.06286156){
    conrank[i] <- "낮음"
  } else {
    conrank[i] <- "매우낮음"
  }
}
```



등급

매우낮음

낮음

보통

매우높음

보통

매우높음

매우낮음

낮음

매우낮음

높음

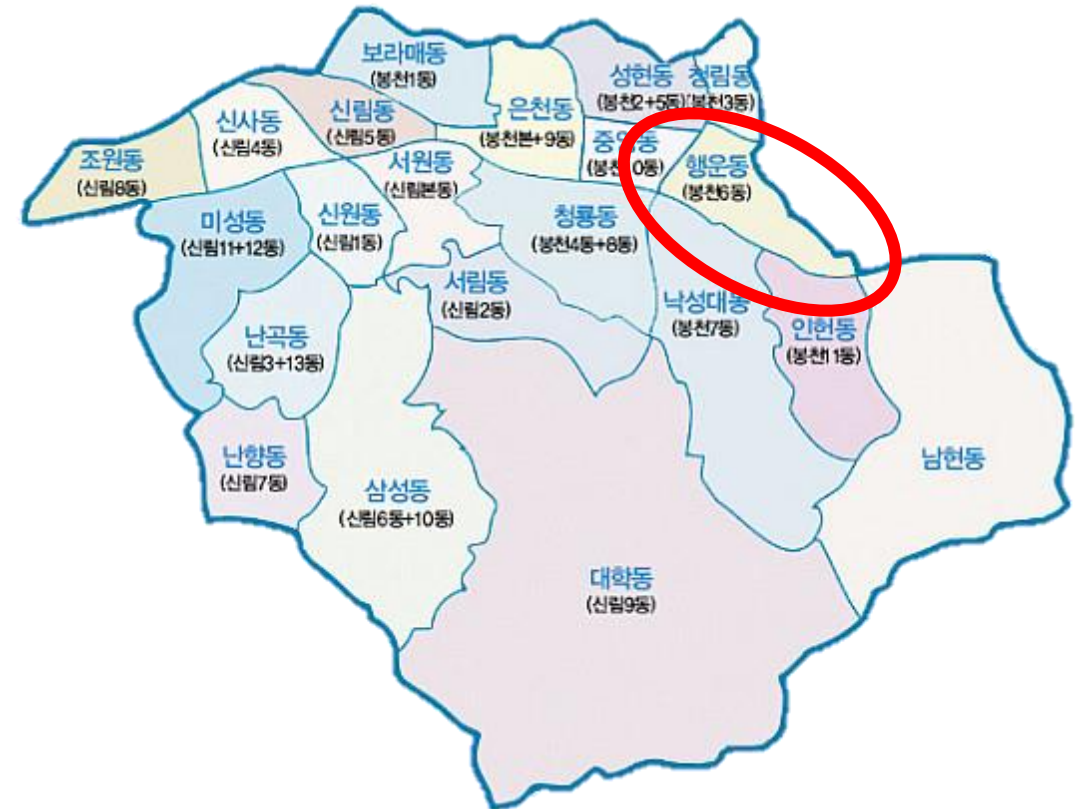
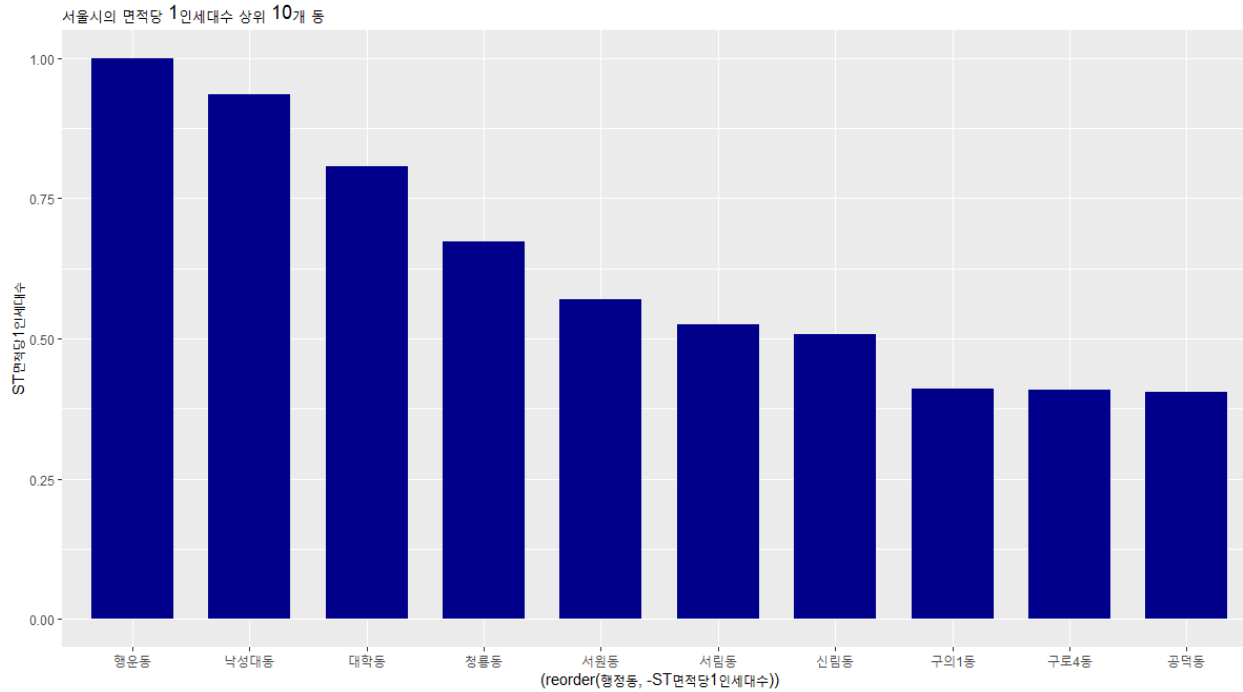
높음

매우낮음

서울시 1인 세대 분포 조사

행정동	전체세대수	X1인세대수	면적	면적당1인세대수	ST면적당1인세대수	등급
가락1동	540.1667	161.8333	1.34	120.7711	0.000000000	매우낮음
가락2동	12253.6667	2739.0833	0.96	2853.2118	0.107621861	낮음
가락본동	11645.6667	4165.0833	1.13	3685.9145	0.140419283	보통
가리봉동	5918.7500	3821.6667	0.40	9554.1667	0.371550459	매우높음
가산동	12778.0833	8862.2500	2.52	3516.7659	0.133757076	보통
가양1동	17989.2500	10840.2500	1.36	7970.7721	0.309185747	매우높음
가양2동	8109.6667	3748.6667	4.70	797.5887	0.026657618	매우낮음
가양3동	7796.0000	2932.2500	1.00	2932.2500	0.110734916	낮음
가회동	2068.4167	921.2500	0.54	1706.0185	0.062437686	매우낮음
갈현1동	11139.8333	4292.6667	0.97	4425.4296	0.169546353	높음
갈현2동	12512.5833	4001.3333	0.96	4168.0556	0.159409237	높음
강일동	11213.6667	2174.0000	2.80	776.4286	0.025824192	매우낮음

1인 가구 밀집도 상위 10개 동



행운동 상권 분석 코드

상점 데이터 불러오기

```
In [3]: 1 shop = pd.read_csv('store.csv', na_values = '?', encoding='cp949') # 엑셀 파일 읽기 #encoding='utf-8'
        2 shop.head()
```

Out [3]:

	상가업소 번호	상 호 명	지 점 명	상 권 업 종 대 분 류 코 드	상 권 업 종 대 분 류 코 드	상 권 업 종 대 분 류 코 드	상 권 업 종 소 분 류 명	표준산 업분류 코드	...	건물관리번호	건물 명	도로 명주소	구우편 번호	신우편 번호	동정 보	층 정 보	호 정 보	경도			
0	19911397	커피빈코리아 선릉로93길점	코리아선 릉로93 길점	Q	음식	Q12	커피점/ 카페	Q12A01	커피전문점/ 카페/ 다방	I56220	...	1168010100106960042022041	NaN	서울 특별시 강남구 선릉 로93 길 6	135080	6149.0	NaN	1	NaN	127.047883	3
1	19911801	프론트포즈	NaN	Q	음식	Q09	우유주점	Q09A01	호프/ 맥주	I56219	...	1154510200101620001017748	NaN	서울 특별시 금천구 가산 로 34-6	153010	8545.0	NaN	1	NaN	126.899220	3

자료 : '상가업소정보', 소상공인시장진흥공단

행운동 상권 분석 코드

동별 업종 지점수 (중분류 업종 기준)

In [46]:

```
1 shopdatadong=pd.crosstab(shop,행정동명,shop,상권업종중분류명)
2 shopdatadong.to_csv('shopdatadong.csv',encoding='EUC-KR')
3 shopdatadong.head()
```

Out [46]:

상권업 종중분 류명	PC/오락/ 당구/볼 링등	가 구 소 매	가방/신 발/액세 서리	가전 제품 소매	가정/주 방/인테리어	개인/가 정용품 수리	개인 서비스	건강/ 미용 식품	경마/경 륜/성인 오락	광 고/ 인 쇄	...	학원-예 능취미 체육	학원-음 악미술 무용	학원-자 격/국가 고시	학원-창 업취업 취미	학 원-컴 퓨터	학 원기 타	한 식	행사/ 이벤트	호텔/ 콘도	화 장 품 소 매
행정동 명																					
가락1 동	0	0	6	1	3	0	0	1	0	29	...	0	0	1	0	0	1	27	0	0	2
가락2 동	7	3	2	3	27	5	1	5	1	1	...	4	23	1	1	1	10	67	0	0	17
가락분 동	20	13	4	4	18	6	2	11	3	3	...	2	14	0	0	0	10	168	0	0	6
가리봉 동	16	3	11	5	14	1	1	6	1	5	...	2	1	0	0	1	4	61	0	2	7
가산동	33	18	95	18	89	38	7	11	2	236	...	7	5	6	1	0	27	316	4	3	53

5 rows x 94 columns

행운동 상권 분석 코드

1인가구 관련 업종 20개의 지점수

```
In [10]: 1 shopdatadong_20 = shopdatadong[['커피점/카페', '패스트푸드', '분식', '한식',
2   '일식/수산물', '음/식료품소매', '중식', '양식',
3   '병원', '화장품소매', '이/미용/건강', 'PC/오락/당구/볼링등',
4   '별식/퓨전요리', '부동산중개', '약국/한약방', '사무/문구/컴퓨터',
5   '종합소매점', '세탁/가사서비스', '제과제빵떡케익', '닭/오리요리']]
6 shopdatadong_20
```

Out [10]:

상권업종 중분류명	커피 점/카 페	패스 트푸 드	분 식	한 식	일식/ 수산 물	음/식료 품소매	중 식	양 식	병 원	화장 품소 매	이/미 용/건 강	PC/오락/당 구/볼링등	별식/퓨 전요리	부동 산중 개	약국/ 한약 방	사무/문 구/컴퓨 터	종합 소매 점	세탁/가 사서비 스	제과제 빵떡케 익	닭/오 리요 리
행정동명																				
가락1동	2	1	6	27	23	60	0	0	4	2	3	0	0	5	1	0	107	2	3	0
가락2동	29	10	19	67	20	10	4	0	34	17	72	7	3	28	8	8	33	23	8	18
가락본동	51	18	31	168	49	9	14	12	32	6	77	20	9	30	5	12	67	22	6	12

행운동 상권 분석 코드

동별 1인가구수 데이터 가져오기

```
In [11]: 1 hum1 = pd.read_csv('grp_total_round_revised.csv', na_values = '?', encoding='cp949') # 엑셀 파일 읽기 #encoding='utf-8'
2 hum1.head()
```

Out [11]:

	Unnamed: 0	X	자치구	행정동	전체세대수	X1인세대
0	1	1	송파구	가락1동	540	162
1	2	2	송파구	가락2동	12254	2739
2	3	3	송파구	가락본동	11646	4165
3	4	4	구로구	가리봉동	5919	3822
4	5	5	금천구	가산동	12778	8862

```
In [12]: 1 pivot_hum1 = pd.pivot_table(hum1, index='행정동') #지역별 평균값 #, aggfunc=np.average
2 pivot_hum1 = pivot_hum1['X1인세대']
3 pivot_hum1 = pd.DataFrame(pivot_hum1)
4 pivot_hum1.columns = ['1인가구수']
5 pivot_hum1.head()
```

Out [12]:

1인가구수	
행정동	
가락1동	162.0
가락2동	2739.0
가락본동	4165.0
가리봉동	3822.0
가산동	8862.0

행운동 상권 분석 코드

1인가구수 대비 지점수 비율 및 정규화

In [34]:

```
1 shop_30_hum = pd.concat([pivot_hum1, shopdatadong_20 ], axis=1, sort=False) #sort 제외 가능
2 shop_30_hum.head()
```

Out[34]:

	1인가구수	커피점/카페	패스트푸드	분식	한식	일식/수산물	음/식료품소매	중식	양식	병원	...	이/미용/건강	PC/오락/당구/볼링등	별식/퓨전요리	부동산중개	약국/한약방	사무/문구/컴퓨터	종합소매점	세탁/가사서비스	제과제빵떡케익	닭/오리요리
가락1동	162.0	2.0	1.0	6.0	27.0	23.0	60.0	0.0	0.0	4.0	...	3.0	0.0	0.0	5.0	1.0	0.0	107.0	2.0	3.0	0.0
가락2동	2739.0	29.0	10.0	19.0	67.0	20.0	10.0	4.0	0.0	34.0	...	72.0	7.0	3.0	28.0	8.0	8.0	33.0	23.0	8.0	18.0
가락본동	4165.0	51.0	18.0	31.0	168.0	49.0	9.0	14.0	12.0	32.0	...	77.0	20.0	9.0	30.0	5.0	12.0	67.0	22.0	6.0	12.0
가리본동	3822.0	13.0	6.0	11.0	61.0	5.0	29.0	46.0	2.0	3.0	...	47.0	16.0	12.0	15.0	4.0	12.0	37.0	17.0	3.0	7.0
가산동	8862.0	175.0	30.0	73.0	316.0	79.0	28.0	31.0	35.0	27.0	...	87.0	33.0	38.0	96.0	17.0	75.0	143.0	50.0	17.0	32.0

5 rows × 21 columns

행운동 상권 분석 코드

```

21 shop_30_hum['종합소매점 비율'] = shop_30_hum['종합소매점'] / shop_30_hum['1인가구수']
22 shop_30_hum['세탁/가사서비스 비율'] = shop_30_hum['세탁/가사서비스'] / shop_30_hum['1인가구수']
23 shop_30_hum['제과제빵떡케익 비율'] = shop_30_hum['제과제빵떡케익'] / shop_30_hum['1인가구수']
24 shop_30_hum['닭/오리요리 비율'] = shop_30_hum['닭/오리요리'] / shop_30_hum['1인가구수']
25
26 shop_30_hum.head()

```

Out[14]:

	1인가구수	커피점/카페	패스트푸드	분식	한식	일식/수산물	음식료/음소매	중식	양식	병원	...	이/미용/건강비율	PC/오락/당구/볼링등비율	별식/퓨전요리비율	부동산중개비율	약국/한약방비율	사무/문구/컴퓨터비율	종합소매점비율	세탁/가사서비스비율	제과제빵/떡/케익비율
가락1동	162.0	2.0	1.0	6.0	27.0	23.0	60.0	0.0	0.0	4.0	...	0.018519	0.000000	0.000000	0.030864	0.006173	0.000000	0.660494	0.012346	0.018519
가락2동	2739.0	29.0	10.0	19.0	67.0	20.0	10.0	4.0	0.0	34.0	...	0.026287	0.002556	0.001095	0.010223	0.002921	0.002921	0.012048	0.008397	0.002556
가락본동	4165.0	51.0	18.0	31.0	168.0	49.0	9.0	14.0	12.0	32.0	...	0.018487	0.004802	0.002161	0.007203	0.001200	0.002881	0.016086	0.005282	0.001200

행운동 상권 분석 코드

```
In [37]: 1 shop_30_hum_norm = shop_30_hum[['커피점/카페', '패스트푸드', '분식', '한식',
2   '일식/수산물', '음/식료품소매', '중식', '양식',
3   '병원', '화장품소매', '이/미용/건강', 'PC/오락/당구/볼링등',
4   '별식/퓨전요리', '부동산중개', '약국/한약방', '사무/문구/컴퓨터',
5   '종합소매점', '세탁/가사서비스', '제과제빵떡케익', '닭/오리요리']] / max_value
6 shop_30_hum_norm.head()
```

Out [37]:

	커피점/ 카페	패스트푸 드	분식	한식	일식/수 산물	음/식료 품소매	중식	양식	병원	화장품소 매	이/미용/ 건강	PC/오 락/당구/ 볼링등	별식/퓨 전요리	부동산중 개	약국/한 약방
가 락 1 동	0.002907	0.005814	0.021201	0.023622	0.070336	0.130435	0.000000	0.000000	0.008214	0.008403	0.004926	0.000000	0.000000	0.010823	0.003322
가 락 2 동	0.042151	0.058140	0.067138	0.058618	0.061162	0.021739	0.041237	0.000000	0.069815	0.071429	0.118227	0.04375	0.032258	0.060606	0.026578
가 락 본 동	0.074128	0.104651	0.109541	0.146982	0.149847	0.019565	0.144330	0.028436	0.065708	0.025210	0.126437	0.12500	0.096774	0.064935	0.016611
가 리 봉 동	0.018895	0.034884	0.038869	0.053368	0.015291	0.063043	0.474227	0.004739	0.006160	0.029412	0.077176	0.10000	0.129032	0.032468	0.013289
가 산 동	0.254360	0.174419	0.257951	0.276465	0.241590	0.060870	0.319588	0.082938	0.055441	0.222689	0.142857	0.20625	0.408602	0.207792	0.056478

행운동 상권 분석 코드

1인가구 비중이 높은 행운동의 업종별 밀집도 비교

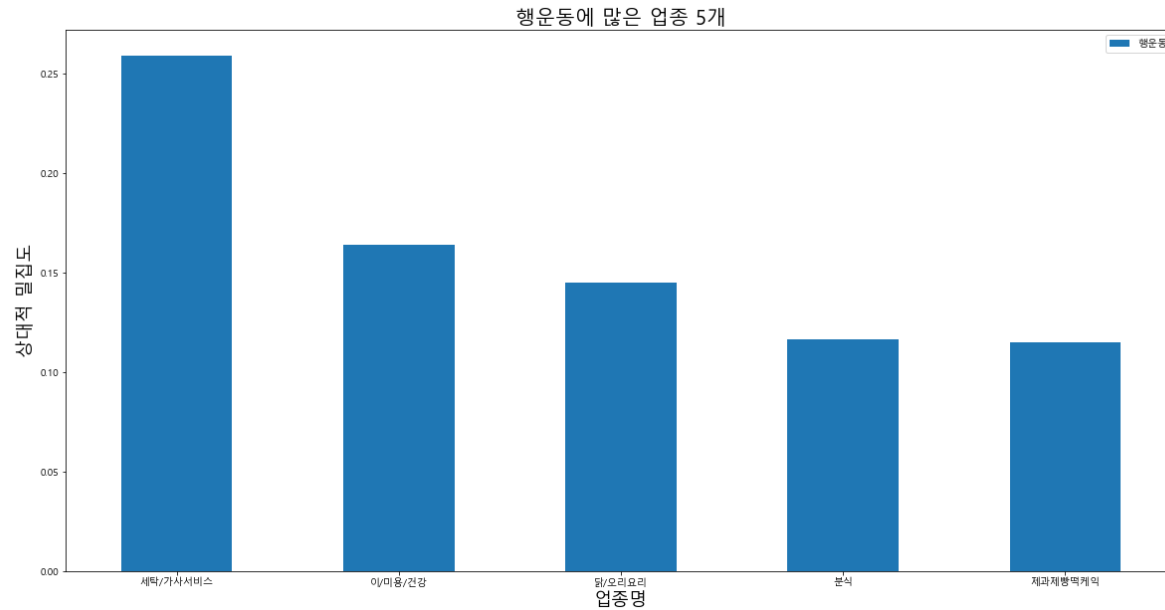
```
In [70]: 1 pd.DataFrame(shop_30_hum_norm.iloc[400,:]).head()
```

Out [70]:

행운동	
커피점/카페	0.047965
패스트푸드	0.110465
분식	0.116608
한식	0.052493
일식/수산물	0.073394

행운동 상권 분석

-상위 5개 업종 분석-

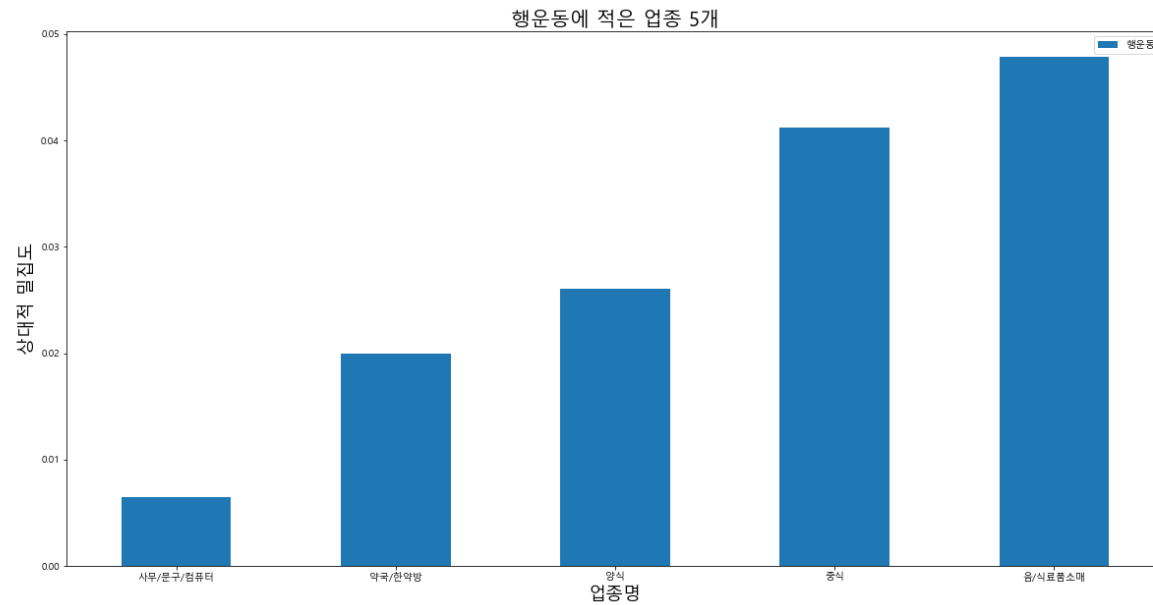


1. 세탁/가사 서비스
2. 미용건강
3. 닭/오리 요리
4. 분식
5. 제과제빵케익

Why? -> 과잉경쟁

행운동 상권 분석

-하위 5개 업종 분석-

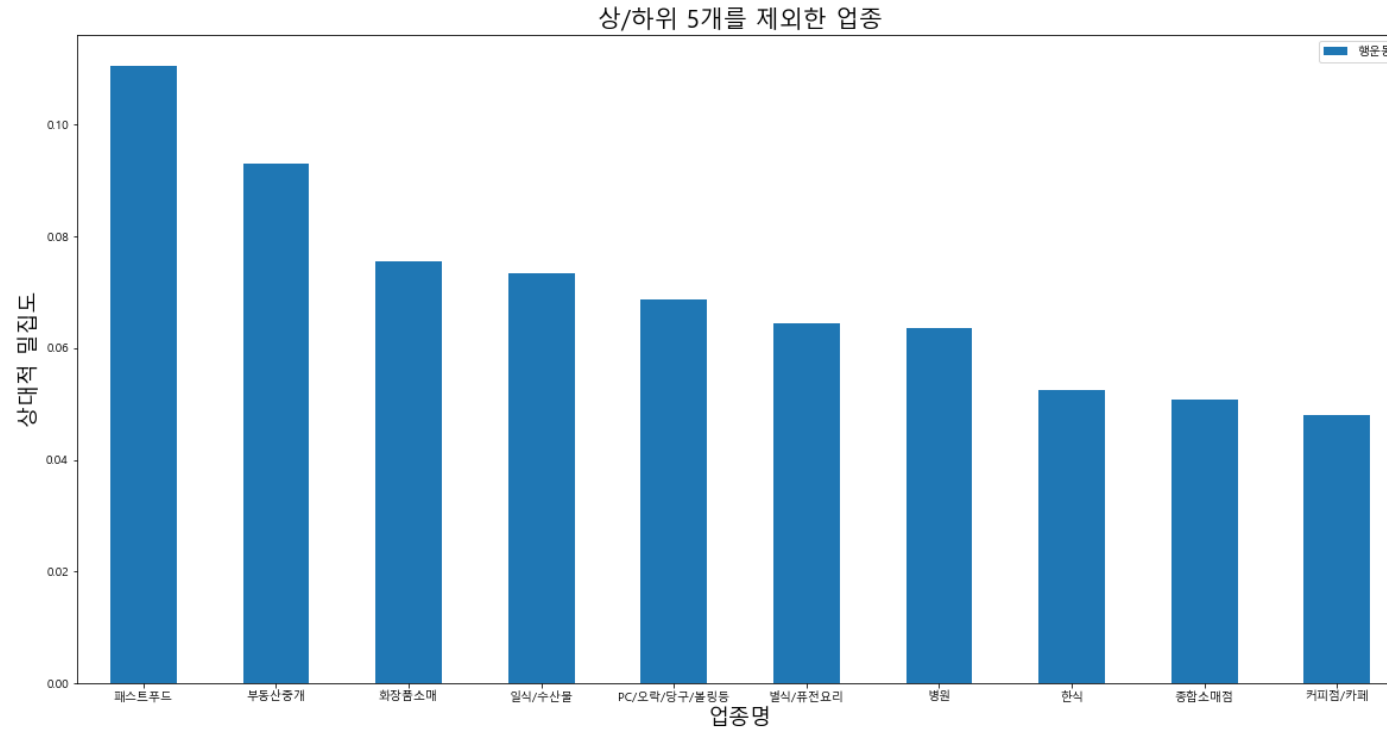


1. 사무/문구/컴퓨터
2. 약국/한약방
3. 양식
4. 중식
5. 음/식료품소매

Why ? -> 수요부족

행운동 상권 분석

-중위 10개 업종 분석-



“패스트푸드” “부동산중개” “화장품소매” “일식/수산물” “PC/오락/당구/볼링장”
“별식/퓨전요리” “병원” “한식” “종합소매점” “커피점/카페”

네이버 뉴스 및 블로그 크롤링



검색



혼밥

검색

상세검색 ▾

뉴스 >

사는 곳 >

kma >

1조 최고 >

더보기



정렬

정확도

최신순

기간

전체

1일

1주

1개월

1년

직접입력

연관검색어 ?

[상권 분석](#)

[솔로 이코노미](#)

[다들 고생하셨어요](#)



- 1인가구 밀집지역인 **관악구**에서 혼밥을 많이 한다!
- 편의점, 배달, 도시락 등 **간편식**을 많이 찾는 편



1인 가구

검색

상세검색 ▾

뉴스 >

연관검색어 ?

[상권 분석](#)

[솔로 이코노미](#)

[다들 고생하셨어요](#)

블로그 >

kma >

1조 최고 >

더보기



정렬

정확도

최신순

기간

전체

1일

1주

1개월

1년

직접입력



• 가격민감도가 높은 편

• 생필품, 소매 관련 키워드 (홈플러스, 상품

• 병원 -> 건강



행운동

검색

상세검색 ▾

뉴스 >

연관검색어 ?

[상권 분석](#)

[솔로 이코노미](#)

[다들 고생하셨어요](#)

블로그 >

kma >

1조 최고 >

더보기



정렬

정확도

최신순

기간

전체

1일

1주

1개월

1년

직접입력



- 가전제품
- 병원관련 키워드
- 부동산 관련 키워드

워드클라우드와 행운동 상관분석을 바탕으로

3개 업종 선정

업종 추천



종합소매점



부동산중개



병원



워드클라우드와 행운동 상관분석을 바탕으로

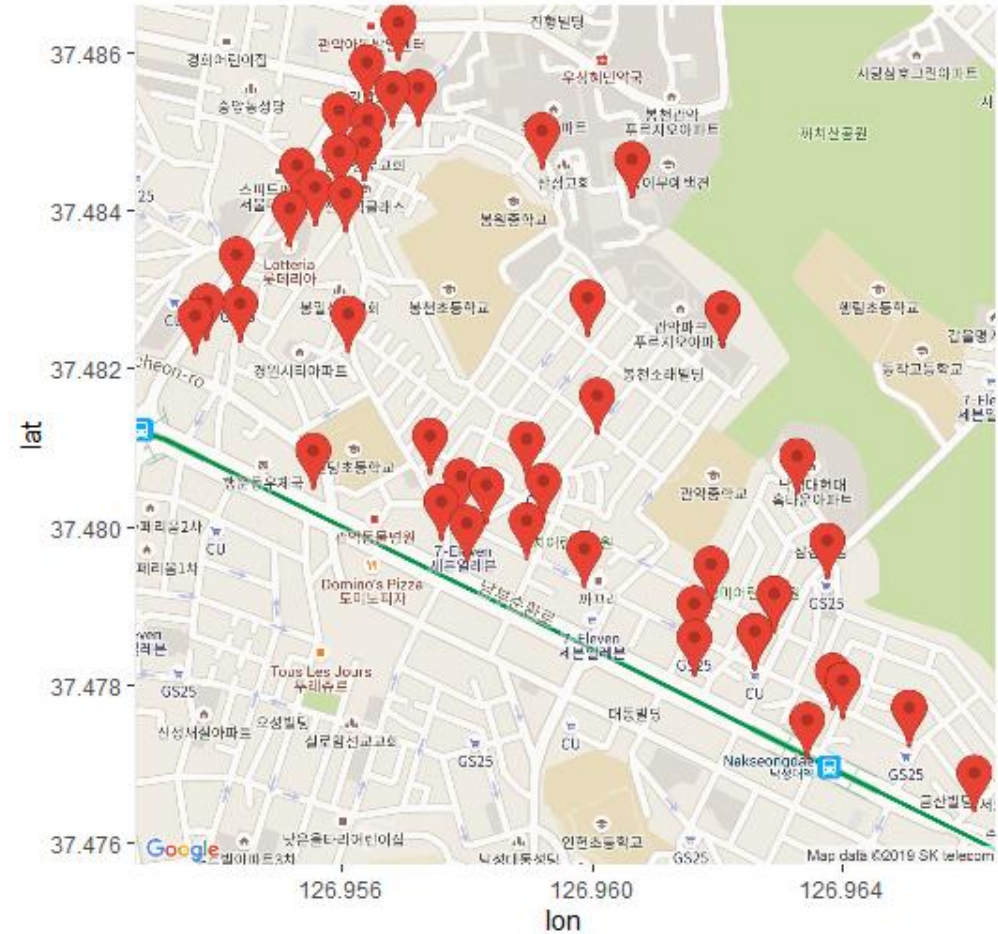
3개 업종 선정

52

업종 추천



종합소매점

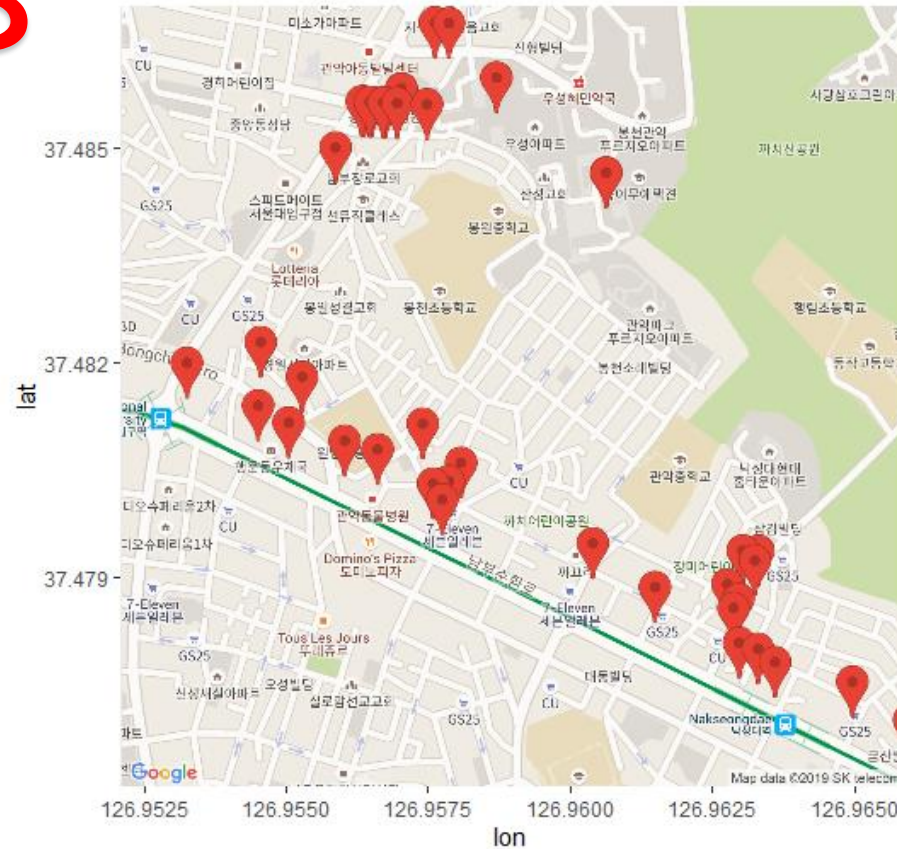


워드클라우드와 행운동 상관분석을 바탕으로

3개 업종 선정

43

업종 추천

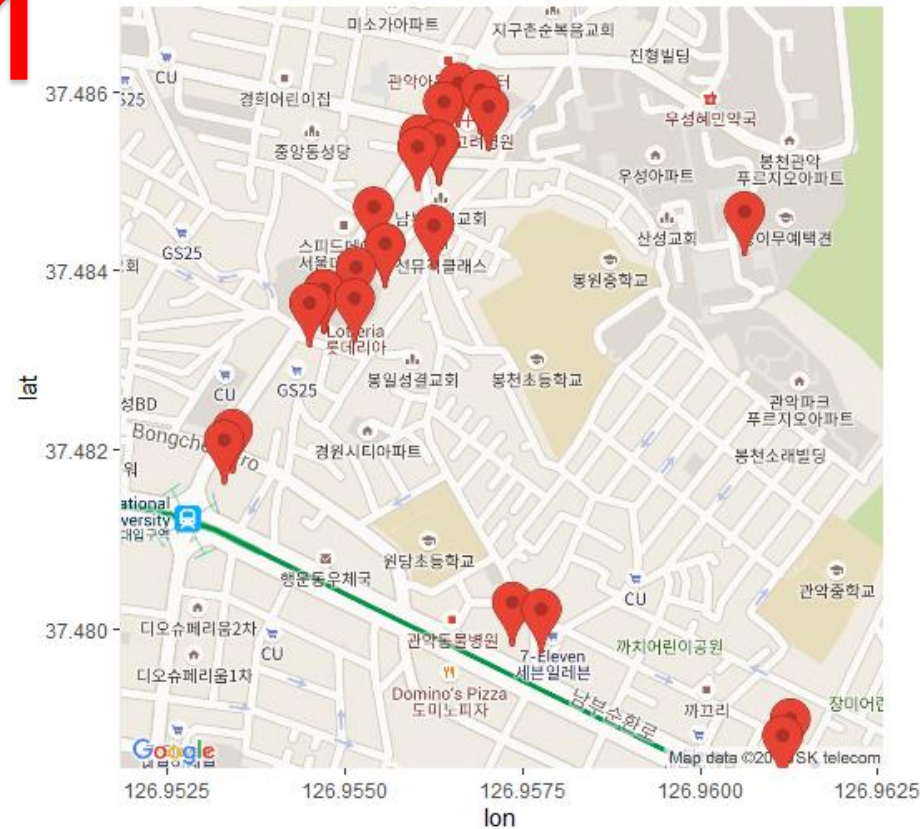


부동산중개

워드클라우드와 행운동 상관분석을 바탕으로

3개 업종 선정

31



병원

업종 추천



기타 요인



행정동	전체 평균 연령	agerank
가락1동	42.275	보통
가락2동	39.450	매우낮음
가락본동	41.200	낮음
가리봉동	48.350	매우높음
가산동	40.825	낮음
가양1동	38.000	매우낮음
가양2동	48.025	매우높음
가양3동	45.375	매우높음
가회동	45.625	매우높음
갈현1동	43.750	높음
갈현2동	42.225	보통

행정동	전체 평균 연령	agerank
403 행운동	40.75	낮음

행운동 거주 인구 평균연령이 낮은 편

기타 요인



<창업 입지 선정 요건> 4

유동인구 4

접근성

역 등이 주변에 있는지 4

성장가능성

상권 및 입지의 성장가능성, 잠재력 4

가시성

가시성이 확보되어 점포를 알릴 수 있어야 함. 4

주변 점포

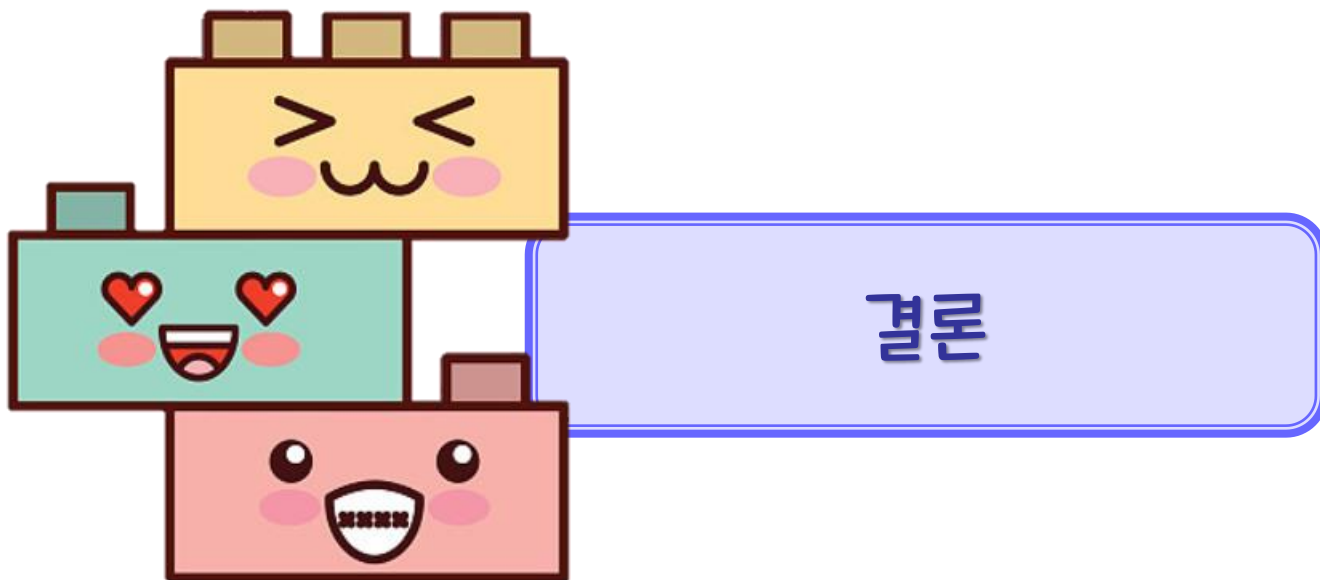
호혜적인 사업 또는 경쟁 매장의 규모 4

임대료

적정 임대료 = 하루 매출 * 3 (매출 대비 적정 점포 임대비) 4

구매력

소비자의 소득 수준. 4





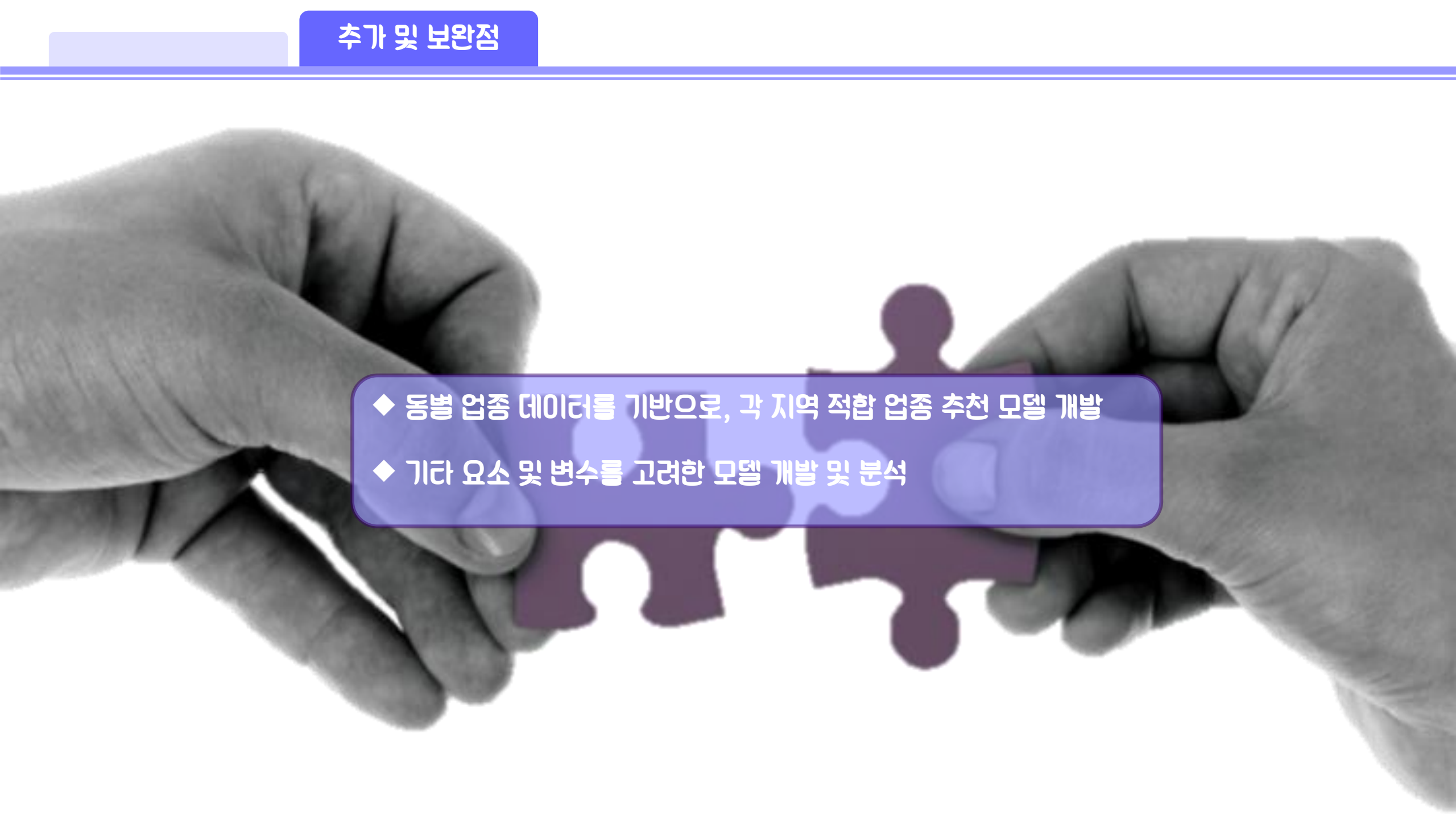
❖ 서울시 인구 밀집도 -> 행운동이 Highest

❖ 행운동에서 중위 10개 업종 추출

❖ 웹 크롤링을 통해 1인가구 비즈니스 관련 키워드 분석

❖ 행운동 3개 업종 추천

❖ 기타 창업 관련 요인 분석

- 
- A grayscale background image showing two hands holding several interlocking puzzle pieces. The puzzle pieces are arranged in a horizontal line, with the hands positioned at the ends, suggesting a process of assembly or completion. The puzzle pieces are a light gray color, and the hands are a darker gray.
- ◆ 동별 업종 데이터를 기반으로, 각 지역 적합 업종 추천 모델 개발
 - ◆ 기타 요소 및 변수를 고려한 모델 개발 및 분석