

2025.12.09

서울시 권역 간 소비 구조 차이 분석

: 데이터 기반 권역별 세분화 전략



팀 02

팀장 : 김서연

팀원 : 이효인, 한예원

CONTENTS

01 연구 배경

02 데이터 설명

03 데이터 전처리

04 데이터 시각화

05 데이터 분석

06 결론 및 시사점

01 연구 배경

주제 선정 및 연구 필요성

소비 품목 간 양극화

2024년 국내 소매판매는 -2.2%로
21년 만에 최대 감소를 기록

(중앙일보, 2024)

의류·신발·가방 등 준내구재 -3.7%,
승용차·가전 같은 내구재는 -2.8%, 음식료품 등 비내구재 -1.3%

(KBS WORLD, 2025)

소비 품목 양극화는
각 지역 특성에 따라 소비 구조가 다르게 재편

소비 품목 양극화 → 공간의 양극화

권역 간 소비 구조 분석의 필요성 및 연구 전략

전통적인 상권 중심 분석 또는 전국 평균 지표만으로는
지역 간 소비 격차, 권역 간 불균형을 파악할 수 없다



자치구를 서울 5대 권역으로 재분류



서울시 자치구별 추정 소득 및 추정 소비 정보와
통계기법(ANOVA/HSD)을 통해 권역별 소비 격차 실증적으로 확인

01 연구 배경

연구 목적



서울시 5대 권역의 상권에서 소비 항목별 비중의 차이가 존재하는지 검정한다.



차이가 존재할 경우 어떤 소비 항목에서, 어떤 권역이 다른 권역보다 유의하게 높은지 파악한다.



02 데이터 설명

데이터 개요

서울 열린데이터광장에서 제공하는 < 서울시 자치구별 추정 소득 및 추정 소비 정보 > 데이터 활용

데이터 이름	출처	주요 내용	활용목적
서울시 상권분석서비스 (소득 소비- 자치구)	 서울 열린데이터 광장 SEOUL OPEN DATA PLAZA	서울시 25개 자치구의 상권 분석용 공공 데이터 (월평균 소득, 총지출 소비 항목별 지출 금액 등)	권역별 소비 항목 비중 분석, 상권의 내재 수요 잠재력 분석

02 데이터 설명

데이터 불러오기

서울시 상권분석서비스(소득소비-자치구).csv을 데이터 프레임으로 불러온 결과, 자치구별 소비 항목별 _총금액 데이터들을 확인

# 데이터 불러오기															
df = pd.read_csv('/content/drive/MyDrive/서울시 상권분석서비스(소득소비-자치구).csv', encoding="cp949") df.head()															
기준_년 분기_코드	행정동 _코드	행정 동_ 코드 _명	월_평균_소 득_금액	소득 구_간_ 코드	지출_총금액	식료품_지출_총 금액	의류_신발_지출 _총금액	생활용품_지출_ 총금액	의료비_지출_총 금액	교통_지출_총금 액	교육_지출_총 금액	유흥_지출_총 금액	여가_문화_지출 _총금액	기타_지출_총 금액	음식_지출_총금 액
0	20242	11215	광진구	3192971	7 100541241000	22634699000	3375381000	1864827000	15005146000	5961023000	3393363000	4281862000	6687173000	5054612000	32283155000
1	20242	11500	강서구	3160623	7 231352477000	34067640000	6533371000	2276283000	32773565000	33678827000	6986997000	4639547000	55668619000	16818036000	37909592000
2	20242	11740	강동구	3292857	7 157864735000	23331462000	2428563000	1590537000	21974063000	3736967000	6822469000	2666873000	66151709000	3909040000	25253052000
3	20243	11260	중랑구	2598785	6 24488372000	7415487000	384896000	486318000	4031113000	1791594000	944118000	998291000	1317535000	892071000	6226949000
4	20243	11470	양천구	3796034	8 40012838000	9217440000	398730000	389703000	8155027000	2268938000	5841107000	732643000	3974345000	1369768000	7665137000

02 데이터 설명

변수 설명

주요 변수

본 연구에서는 지역 구분 변수와 경제적 지표 및 항목별 소비 지출 데이터를 주요 변수로 설정하여 권역간 소비 구조 차이를 검정 하였다.

범주형 변수

변수명	설명	유형
분기	분기 정보	범주형
행정동	행정동 이름	범주형
자치구코드	자치구 구분 코드	범주형
구	자치구 이름	범주형
권역	서울시 5개 권역	범주형

수치형 변수

변수명	설명	유형
월평균소득	분기 정보	수치형
총지출	자치구별 총 지출	수치형
지출_총금액	각 소비 항목별 총금액	수치형

특히, 권역 변수 같은 경우 서울시 5개 권역 (동북, 동남, 서남, 서북, 도심) 으로 그룹화

02 데이터 설명

변수 설명

파생 변수

전권 대비 15% P 상승

본 연구에서는 전체 소비항목 중 분석 목적과 연관성이 높은 4개 항목(여가·교육·의료·교통)만 선별하여 파생 변수를 생성하였다.

변수명	설명	의미
교육비중	교육_지출_총금액 / 총지출	총지출 대비 교육지출 비율
여가비중	여가_문화_지출_총금액 / 총지출	총지출 대비 여가지출 비율
의료비중	의료비_지출_총금액 / 총지출	총지출 대비 의료지출 비율
교통비중	교통_지출_총금액 / 총지출	총지출 대비 교통지출 비율

02 데이터 설명

탐색적 자료 분석

전권 대비 15% P 상승

▶ df.groupby('권역')[['교육비중','여가비중','의료비중','교통비중']].agg(['mean','std','min','max'])

... 권역	교육비중				여가비중				의료비중				교통비중			
	mean	std	min	max												
도심권	0.011518	0.008673	0.004869	0.046077	0.391142	0.298068	0.068975	0.799012	0.034731	0.034191	0.004682	0.112479	0.066014	0.073588	0.000555	0.231633
동남권	0.029851	0.013521	0.012074	0.056549	0.397979	0.146920	0.114823	0.590554	0.114767	0.047772	0.048649	0.216122	0.023700	0.015738	0.009268	0.063503
동북권	0.044230	0.014721	0.023063	0.086920	0.072181	0.022318	0.038383	0.152957	0.158190	0.036979	0.103692	0.275727	0.069668	0.021917	0.027497	0.142177
서남권	0.042354	0.046132	0.001911	0.151578	0.277014	0.329519	0.023019	0.954283	0.107377	0.069125	0.012519	0.322766	0.155347	0.194438	0.001963	0.575496
서북권	0.032417	0.018911	0.006949	0.067132	0.295395	0.367353	0.022907	0.905172	0.272602	0.237915	0.013532	0.682504	0.026333	0.020705	0.002711	0.064068

03 데이터 전처리

변수명 변경

#데이터 전처리

변수명	예시	새로운 변수명
기준_년분기_코드	20242	분기
행정동_코드_명	11215	행정동
월_평균_소득_금액	3192971	월평균소득
지출_총금액	100541241000	총지출



```
▶ df = df.rename(columns={  
    "기준_년분기_코드": "분기",  
    "행정동_코드_명": "행정동",  
    "월_평균_소득_금액": "월평균소득",  
    "지출_총금액": "총지출",  
})
```

분석 편의를 위해 변수명을 표준화

데이터 타입 확인 및 변환



데이터 타입 확인 & 변환
df["분기"] = df["분기"].astype(str)

<분기 타입 변환 >

- ① 숫자로 두면 연속형 변수로 오해되어 오류 발생 방지
- ② 분기를 범주형 변수로 명확하게 처리 목적

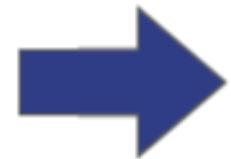
03 데이터 전처리

결측치 확인 및 제거

전체 대비 15% P 상승



결측치 확인
df.isnull().sum()



...	0
분기	0
행정동_코드	0
행정동	0
월평균소득	0
소득_구간_코드	0
총지출	0
식료품_지출_총금액	0
의류_신발_지출_총금액	0
생활용품_지출_총금액	0
의료비_지출_총금액	0
교통_지출_총금액	0
교육_지출_총금액	0
유홍_지출_총금액	0
여가_문화_지출_총금액	0
기타_지출_총금액	0
음식_지출_총금액	0

결측치 없음을 확인

03 데이터 전처리

지역 단위 재구성 및 파생 변수 생성

① 행정동 코드의 앞 5자리를 잘라 자치구 코드를 추출

```
▶ # 자치구(구) 변수 다시 만들기 (행정동 ->구 매핑)  
# 행정동 코드에서 "자치구 코드(앞 5자리)" 뽑기  
df['자치구코드'] = df['행정동_코드'].astype(str).str[:5]
```

행정동 데이터에서 자치구 단위 분석이 가능하도록
상위 지역구 정보를 생성 하였다.

전원 대비 15% P 상승

② 자치구 코드를 자치구명 , 권역으로 매핑하여 분석 단위 통일

```
▶ # 자치구코드 → 자치구명 매핑  
gu_map = {  
    '11110': '종로구', '11140': '중구', '11170': '용산구',  
    '11200': '성동구', '11215': '광진구', '11230': '동대문구',  
    '11260': '중랑구', '11290': '성북구', '11305': '강북구',  
    '11320': '도봉구', '11350': '노원구', '11380': '은평구',  
    '11410': '서대문구', '11440': '마포구', '11470': '양천구',  
    '11500': '강서구', '11530': '구로구', '11545': '금천구',  
    '11560': '행동구', '11590': '동작구', '11620': '관악구',  
    '11650': '서초구', '11680': '강남구', '11710': '송파구',  
    '11740': '강동구'  
}  
  
df['구'] = df['자치구코드'].map(gu_map)
```

숫자로 된 자치구 코드를 텍스트 구 이름으로 변환

```
▶ # 구 → 권역 만들기 (분석 그룹용)  
region_map = {  
    '종로구': '도심권', '중구': '도심권', '용산구': '도심권',  
    '은평구': '서북권', '서대문구': '서북권', '마포구': '서북권',  
    '성북구': '동북권', '강북구': '동북권', '도봉구': '동북권',  
    '노원구': '동북권', '동대문구': '동북권', '중랑구': '동북권',  
    '성동구': '동북권', '광진구': '동북권',  
  
    '강남구': '동남권', '서초구': '동남권', '송파구': '동남권',  
    '강동구': '동남권',  
  
    '영등포구': '서남권', '동작구': '서남권', '관악구': '서남권',  
    '금천구': '서남권', '구로구': '서남권', '강서구': '서남권',  
    '양천구': '서남권'  
}  
  
df['권역'] = df['구'].map(region_map)
```

서울시. 통계청등에서 쓰는 공식적인 <서울 5대 권역 구분 기준> 사용하여
각 자치구가 속하는 5개의 권역 정보를 맵핑

03 데이터 전처리

지역 단위 재구성 및 파생 변수 생성

전체 대비 15% P 상승

```
▶ # 소비 항목 비중(0~1) 만들기  
df['교육비중'] = df['교육_지출_총금액'] / df['총지출']  
df['여가비중'] = df['여가_문화_지출_총금액'] / df['총지출']  
df['의료비중'] = df['의료비_지출_총금액'] / df['총지출']  
df['교통비중'] = df['교통_지출_총금액'] / df['총지출']  
  
ratio_cols = ['교육비중', '여가비중', '의료비중', '교통비중']
```



총지출 대비 항목별 지출 비율을
0~1로 표준화하여 지역간 비교

각 자치구마다 총지출 규모가 다르기 때문에
총지출을 비율 형태로 표준화하였다.

③ 소비 항목 비중 변수 생성

03 데이터 전처리

이상치 탐색

베이비파크에서

권역을 기준으로 데이터를 분리하여 각 그룹별로 이상치 탐색을 수행

```
❶ # 동북권 이상치 탐색(IQR 방식)
sub = df[df['권역'] == '동북권']

outlier_idx = []

# 교육비중
q1 = sub['교육비중'].quantile(.25)
q3 = sub['교육비중'].quantile(.75)
iqr = q3 - q1
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr
outlier_idx.extend(list(sub[(sub['교육비중'] < lower) | (sub['교육비중'] > upper)].index))

# 여가비중
q1 = sub['여가비중'].quantile(.25)
q3 = sub['여가비중'].quantile(.75)
iqr = q3 - q1
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr
outlier_idx.extend(list(sub[(sub['여가비중'] < lower) | (sub['여가비중'] > upper)].index))

# 의료비중
q1 = sub['의료비중'].quantile(.25)
q3 = sub['의료비중'].quantile(.75)
iqr = q3 - q1
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr
outlier_idx.extend(list(sub[(sub['의료비중'] < lower) | (sub['의료비중'] > upper)].index))

# 교통비중
q1 = sub['교통비중'].quantile(.25)
q3 = sub['교통비중'].quantile(.75)
iqr = q3 - q1
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr
outlier_idx.extend(list(sub[(sub['교통비중'] < lower) | (sub['교통비중'] > upper)].index))

outlier_idx = list(set(outlier_idx))

print("동북권 이상치 구 목록:")
print(sub.loc[outlier_idx, '구'].tolist())
```

같은 방식으로
5개의 각 권역의 이상치를 탐색한다.

```
❷ # 서남권 이상치 탐색(IQR 방식)
sub = df[df['권역'] == '서남권']

outlier_idx = []

# 교육비중
q1 = sub['교육비중'].quantile(.25)
q3 = sub['교육비중'].quantile(.75)
iqr = q3 - q1
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr
outlier_idx.extend(list(sub[(sub['교육비중'] < lower) | (sub['교육비중'] > upper)].index))

# 여가비중
q1 = sub['여가비중'].quantile(.25)
q3 = sub['여가비중'].quantile(.75)
iqr = q3 - q1
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr
outlier_idx.extend(list(sub[(sub['여가비중'] < lower) | (sub['여가비중'] > upper)].index))

# 의료비중
q1 = sub['의료비중'].quantile(.25)
q3 = sub['의료비중'].quantile(.75)
iqr = q3 - q1
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr
outlier_idx.extend(list(sub[(sub['의료비중'] < lower) | (sub['의료비중'] > upper)].index))

# 교통비중
q1 = sub['교통비중'].quantile(.25)
q3 = sub['교통비중'].quantile(.75)
iqr = q3 - q1
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr
outlier_idx.extend(list(sub[(sub['교통비중'] < lower) | (sub['교통비중'] > upper)].index))

outlier_idx = list(set(outlier_idx))

print("서남권 이상치 구 목록:")
print(sub.loc[outlier_idx, '구'].tolist())
```

동북권 이상치 탐색

서남권 이상치 탐색

03 데이터 전처리

이상치 탐색 결과 및 처리

- IQR 기준에서 일부 값이 이상치로 탐지되었으나, 이는 변수 특성상 작은 차이도 이상치로 오인될 수 있는 통계적 현상
 - 소비비중 변수는 0~1 범위의 안정적인 변수로 실제 극단적 이상치는 거의 없었다.

< 이상치 제거 후, 데이터 왜곡 가능성 >

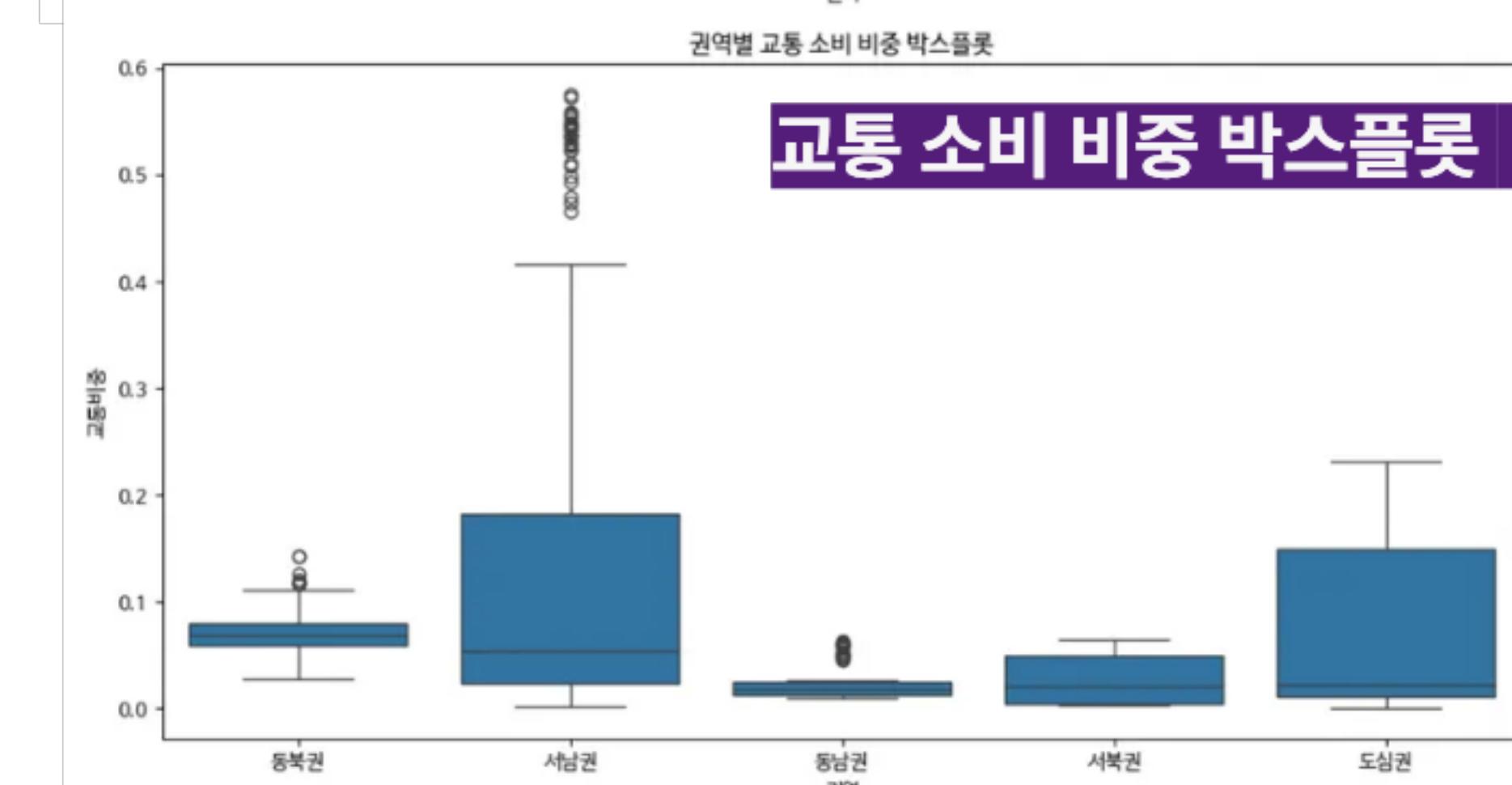
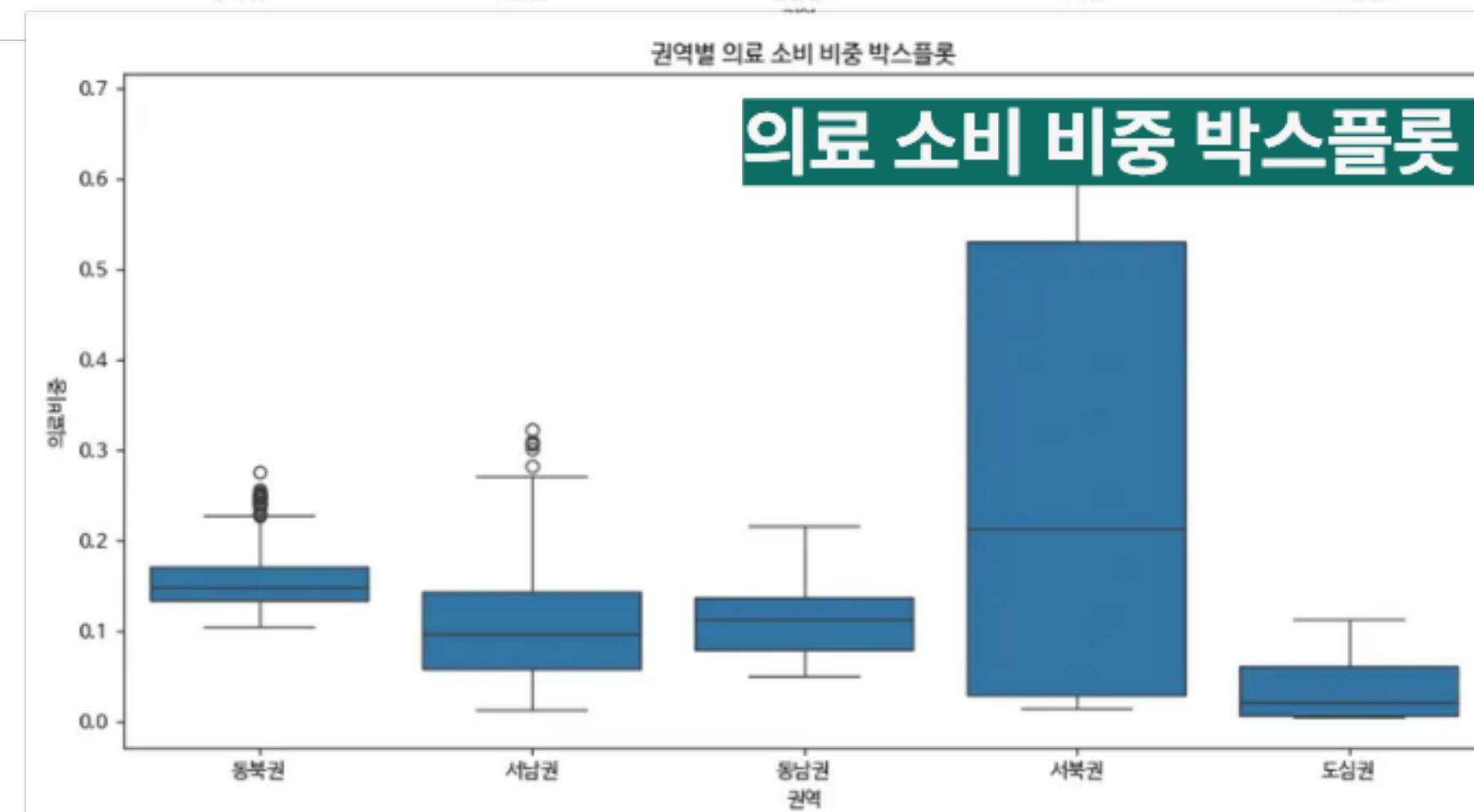
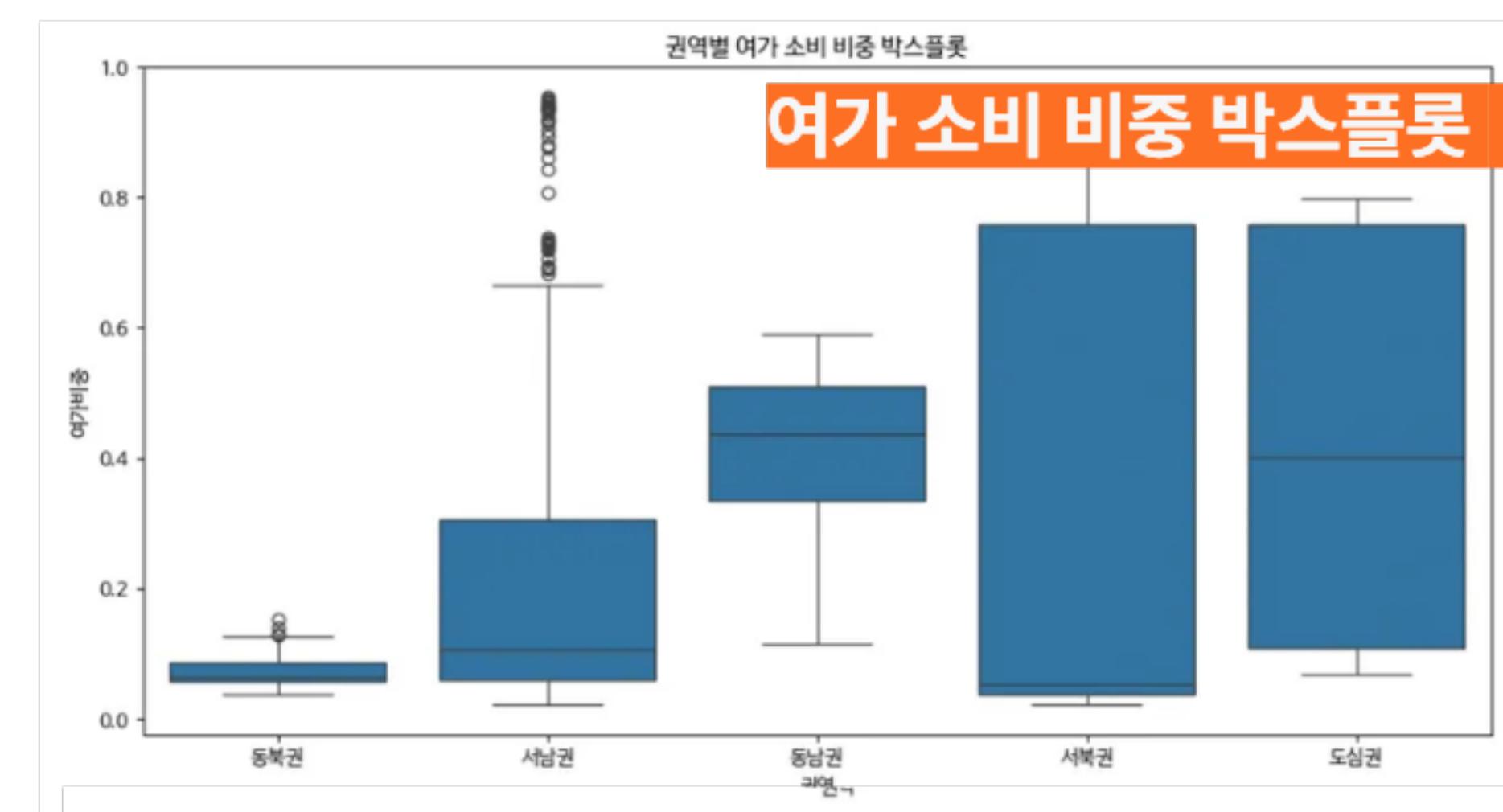
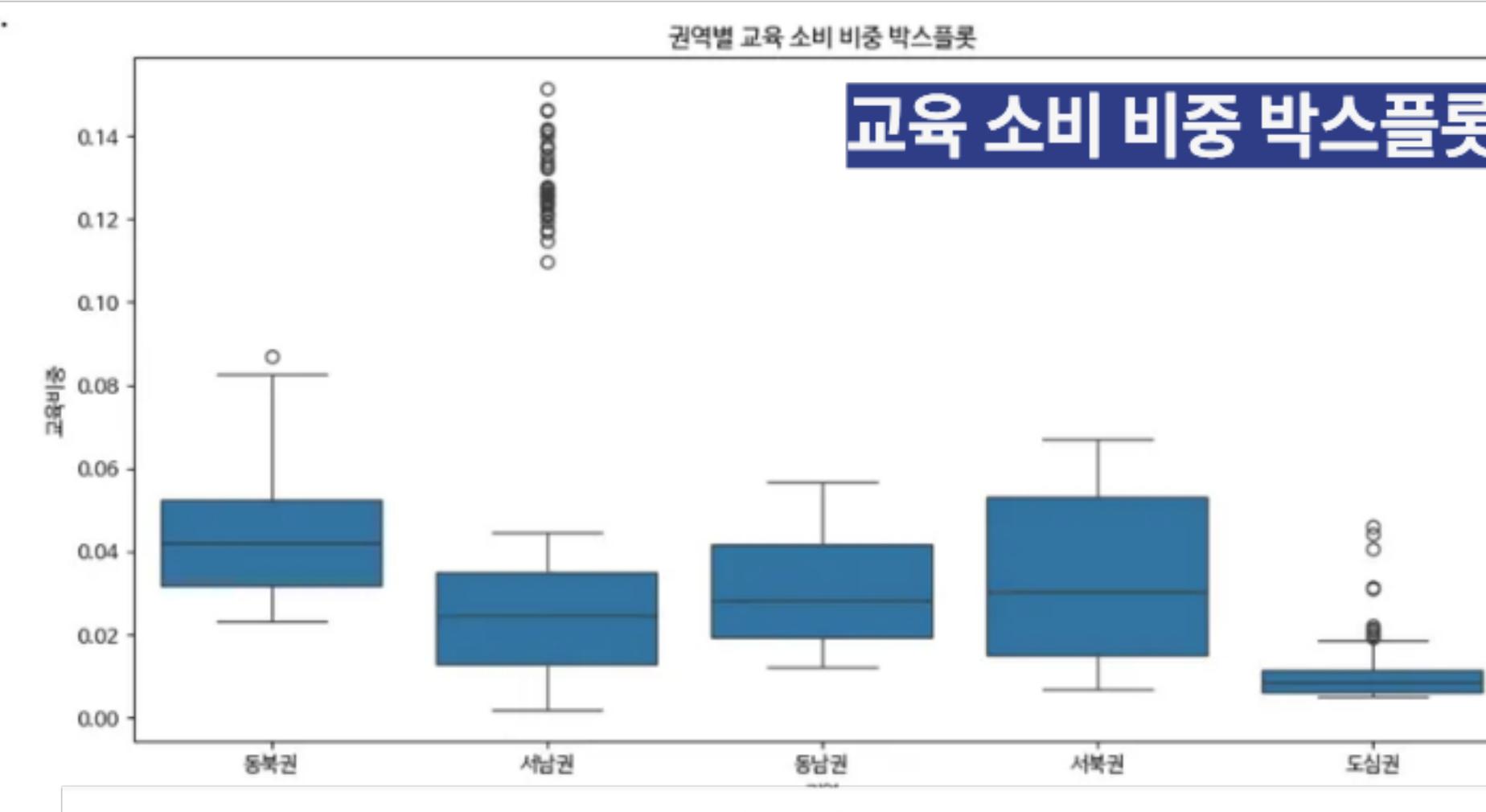
자치구 단위 집계 데이터 특성상 이상치를 제거하면, 중요한 지역 정보가 손실되고, 오히려 데이터 왜곡 가능성이 있다.
소비 비중 변수가 갖는 구조적 특성과 연구 목적을 고려해야 한다.



이상치 제거 A red 'X' mark is placed next to the text '이상치 제거' to indicate that this action is generally not recommended or is incorrect.

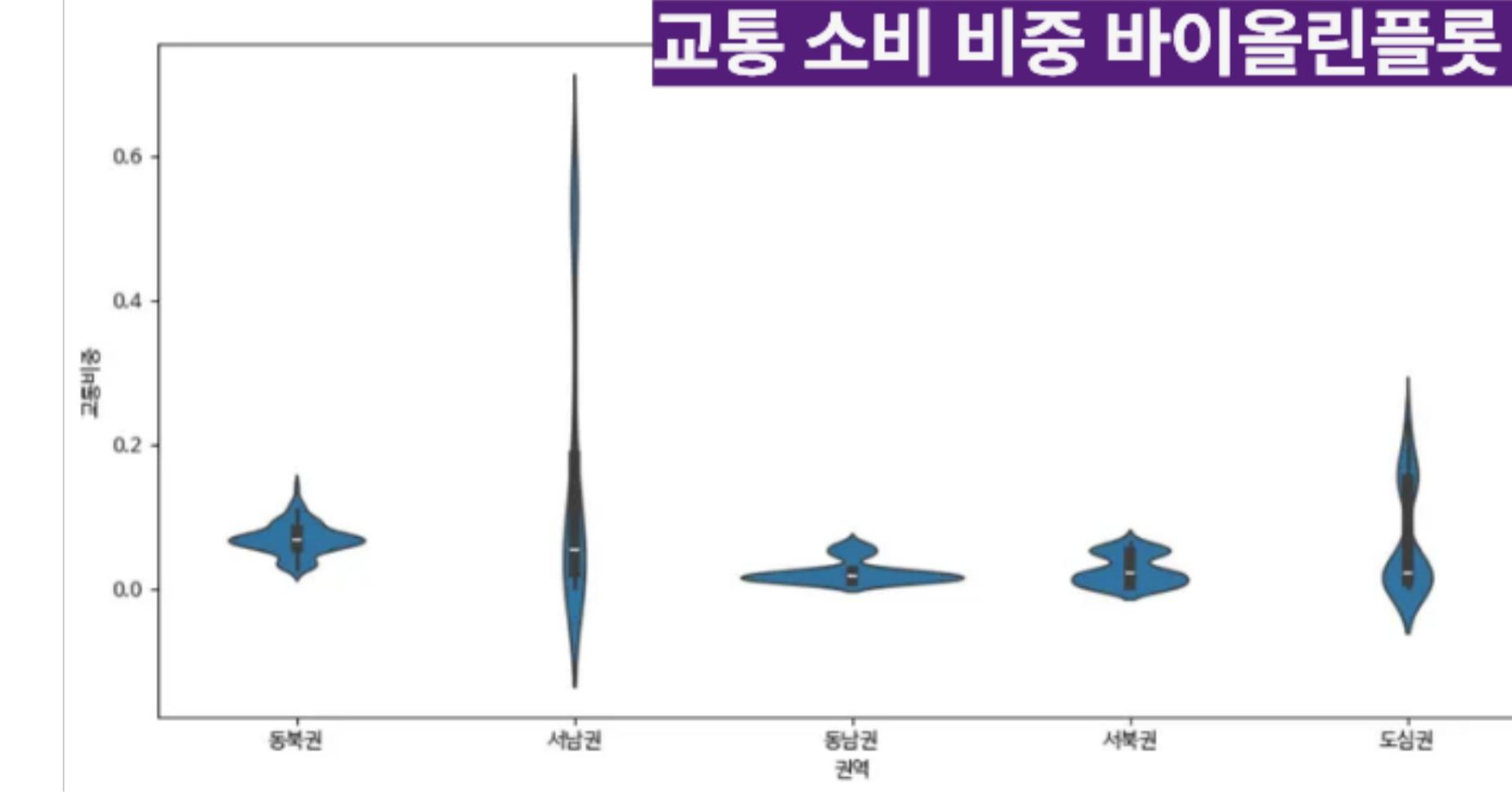
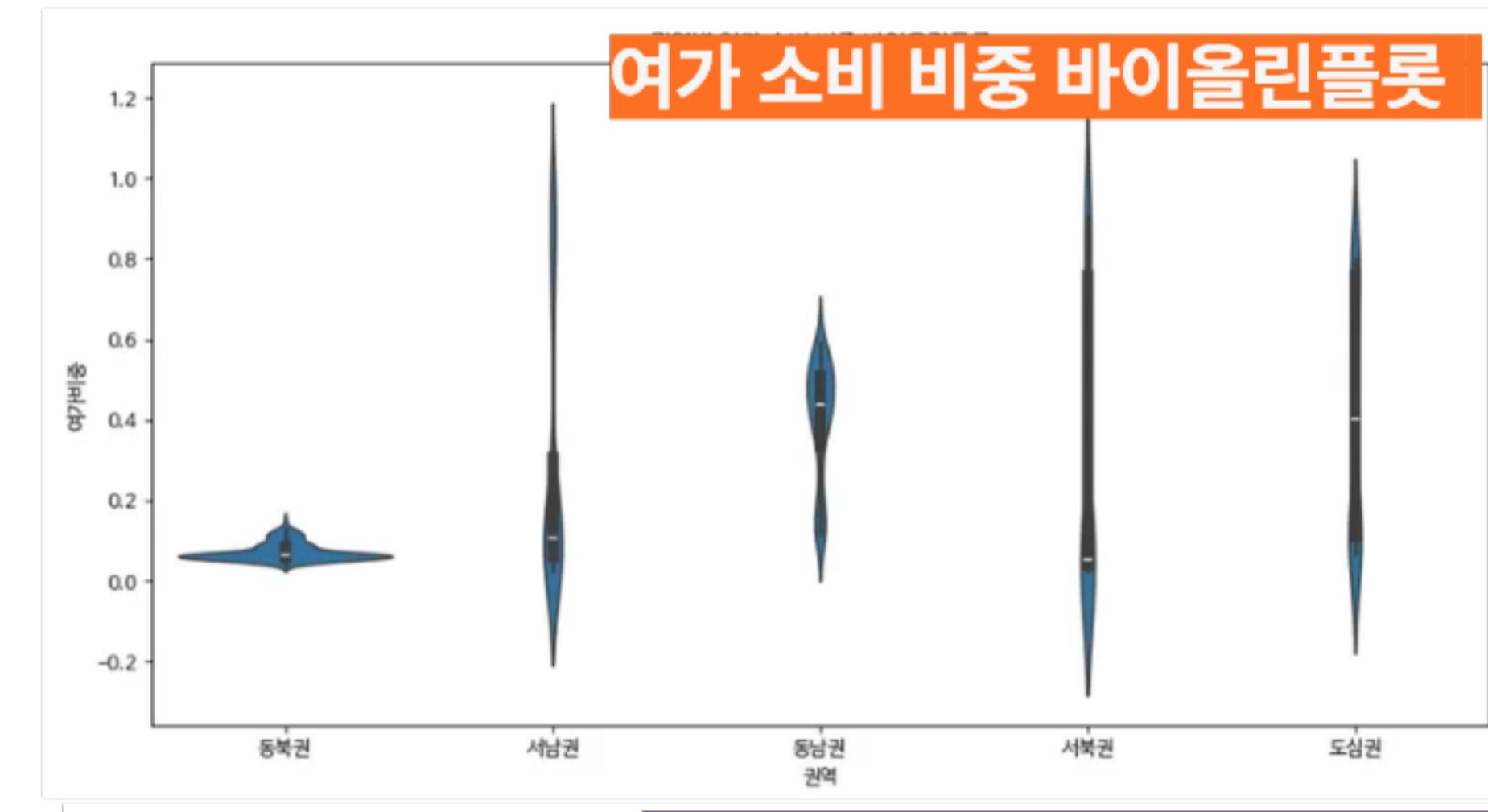
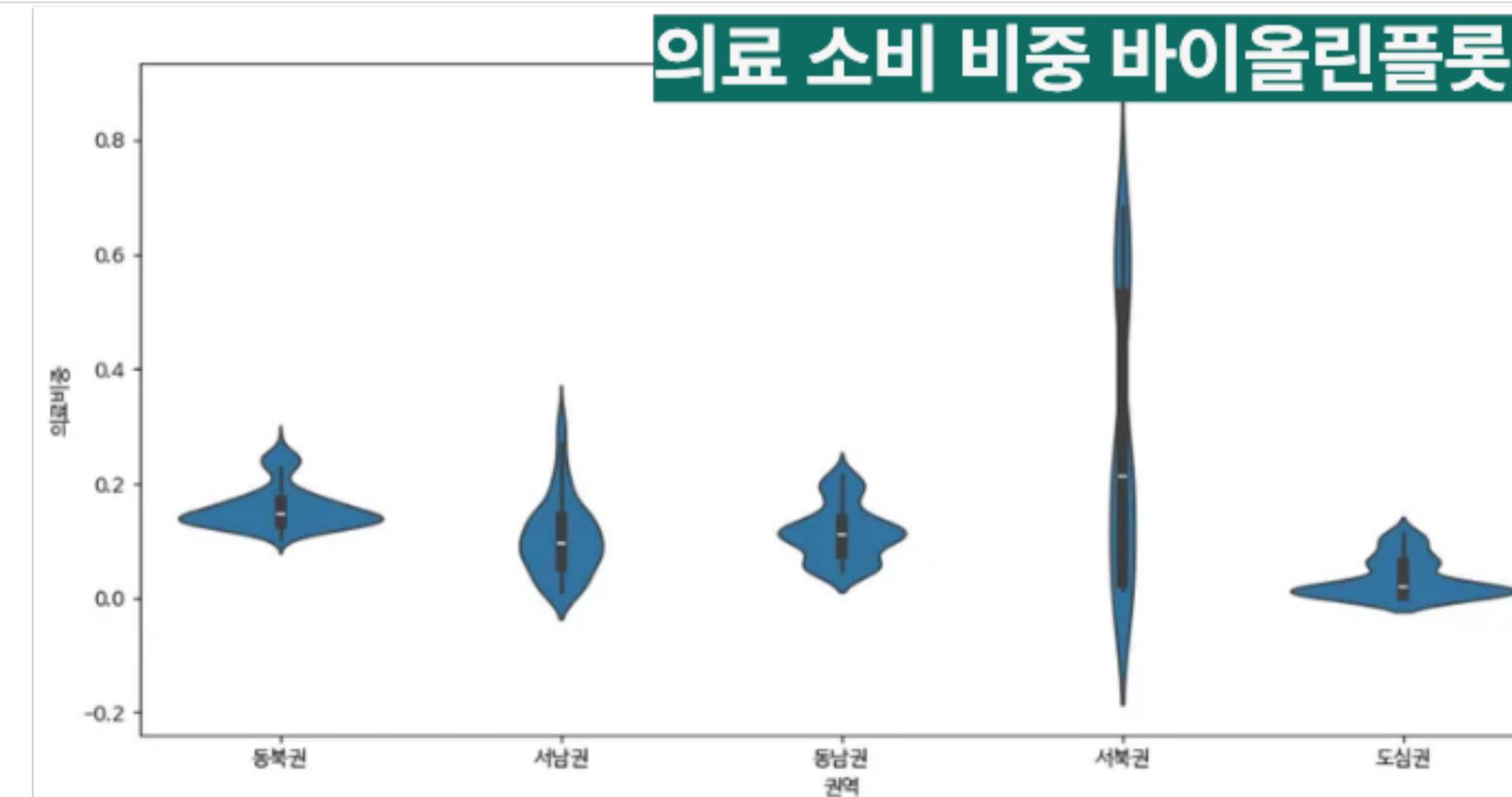
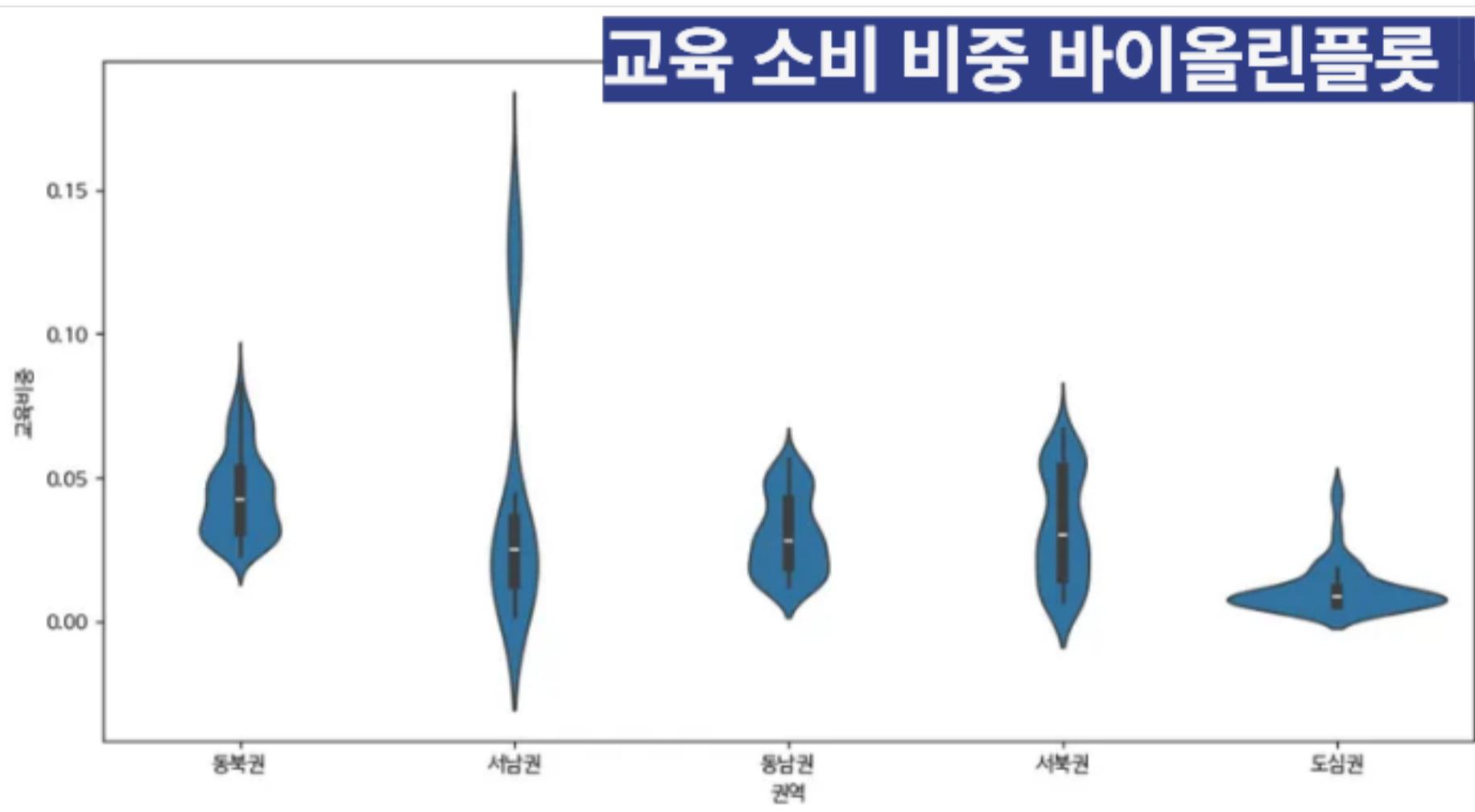
04 데이터 시각화 1: Box plot

일부 권역에서는 교육/여가/의료/교통 비중에서 뚜렷한 차이가 나타남



04 데이터 시각화 2: Violin Plot

특정 권역에서 분포가 한쪽으로 치우치거나
넓게 퍼져 있는 양상이 나타남



05 데이터 분석 : ANOVA

ANOVA란



집단 간 평균 차이가 통계적으로 유의한지 검정하는 방법으로,
하나의 독립변수(집단)에 따라 종속변수의 평균이 달라지는지를 비교한다.



05 데이터 분석 : ANOVA

ANOVA 가설 설정

H0 : 귀무가설

서울 5대 권역의 상권 간 소비 항목 비중에는 구조적 차이가 없다.
즉, 권역이 달라져도 소비 항목 비중의 평균은 동일하다.

H1 : 대립가설

서울 5대 권역의 상권 간 소비 항목 비중에는 유의미한 차이가 존재한다.
즉, 특정 권역은 특정 소비 항목에서 상대적으로 더 높은 비중을 보인다.



05 데이터 분석 : ANOVA 수행

교육 소비 비중 ANOVA

OLS Regression Results									
Dep. Variable:	교육비중	R-squared:	0.132						
Model:	OLS	Adj. R-squared:	0.126						
Method:	Least Squares	F-statistic:	24.45						
Date:	Sat, 06 Dec 2025	Prob (F-statistic):	7.33e-19						
Time:	15:04:59	Log-Likelihood:	1419.8						
No. Observations:	650	AIC:	-2838.						
Df Residuals:	645	BIC:	-2807.						
Df Model:	4								
Covariance Type:	nonrobust								
	coef	std err	t	P> t	[0.025	0.975]			
Intercept	0.003	3.720	0.000	0.005	0.018				
C(권역)	0.004	4.475	0.000	0.010	0.026				
C(권역) [T.동남권]	0.004	9.011	0.000	0.025	0.048				
C(권역) [T.서남권]	0.004	8.333	0.000	0.024	0.038				
C(권역) [T.서북권]	0.004	4.773	0.000	0.012	0.029				
Notes:									
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified									
sum_sq	df	F	PR(>F)						
Intercept	0.010347	1.0	13.840667	2.163280e-04					
C(권역)	0.073116	4.0	24.449974	7.327357e-19					
Residual	0.482210	645.0	NaN	NaN					

1) 교육비중 ANOVA

```
import statsmodels.api as sm
import statsmodels.formula.api as smf

EduFit = smf.ols(formula='교육비중 ~ C(권역)', data=df).fit()
print(EduFit.summary())

print(sm.stats.anova_lm(EduFit, typ=3))
```

여가 소비 비중 ANOVA

OLS Regression Results									
Dep. Variable:	여가비중	R-squared:	0.214						
Model:	OLS	Adj. R-squared:	0.209						
Method:	Least Squares	F-statistic:	43.98						
Date:	Sat, 06 Dec 2025	Prob (F-statistic):	1.17e-32						
Time:	15:14:53	Log-Likelihood:	-9.7127						
No. Observations:	650	AIC:	29.43						
Df Residuals:	645	BIC:	51.81						
Df Model:	4								
Covariance Type:	nonrobust								
	coef	std err	t	P> t	[0.025	0.975]			
Intercept	0.3911	0.028	14.010	0.000	0.336	0.446			
C(권역)	0.008	0.037	0.185	0.853	-0.066	0.879			
C(권역) [T.동남권]	0.008	0.033	-9.743	0.000	-0.383	-0.255			
C(권역) [T.서남권]	0.008	0.033	-3.420	0.001	-0.180	-0.049			
C(권역) [T.서북권]	0.008	0.039	-2.425	0.016	-0.173	-0.018			
Notes:									
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified									
sum_sq	df	F	PR(>F)						
Intercept	11.933369	1.0	196.292969	3.965989e-39					
C(권역)	18.694651	4.0	43.979298	1.167468e-32					
Residual	39.211914	645.0	NaN	NaN					

2) 여가비중 ANOVA

```
LeisureFit = smf.ols(formula='여가비중 ~ C(권역)', data=df).fit()
print(LeisureFit.summary())

print(sm.stats.anova_lm(LeisureFit, typ=3))
```

의료 소비 비중 ANOVA

OLS Regression Results									
Dep. Variable:	의료비중	R-squared:	0.305						
Model:	OLS	Adj. R-squared:	0.308						
Method:	Least Squares	F-statistic:	70.62						
Date:	Sat, 06 Dec 2025	Prob (F-statistic):	1.34e-49						
Time:	15:23:07	Log-Likelihood:	609.64						
No. Observations:	650	AIC:	-1209.						
Df Residuals:	645	BIC:	-1187.						
Df Model:	4								
Covariance Type:	nonrobust								
	coef	std err	t	P> t	[0.025	0.975]			
Intercept	0.0347	0.011	3.226	0.001	0.014	0.056			
C(권역)	0.0080	0.014	5.620	0.000	0.052	0.188			
C(권역) [T.동남권]	0.1235	0.013	9.779	0.000	0.089	0.148			
C(권역) [T.서남권]	0.0726	0.013	5.645	0.000	0.047	0.096			
C(권역) [T.서북권]	0.2379	0.015	15.623	0.000	0.208	0.268			
Omnibus:	168.138	Durbin-Watson:	1.865						
Prob(Omnibus):	0.000	Jarque-Bera (JB):	809.840						
Skew:	1.006	Prob(JB):	1.40e-176						
Kurtosis:	8.005	Cond. No.	7.54						
Notes:									
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified									
sum_sq	df	F	PR(>F)						
Intercept	0.094087	1.0	10.406594	1.319143e-03					
C(권역)	2.553820	4.0	78.617236	1.335241e-49					
Residual	5.831486	645.0	NaN	NaN					

3) 의료비중 ANOVA

```
MedicalFit = smf.ols(formula='의료비중 ~ C(권역)', data=df).fit()
print(MedicalFit.summary())

print(sm.stats.anova_lm(MedicalFit, typ=3))
```

교통 소비 비중 ANOVA

OLS Regression Results						
Dep. Variable:	교통비중	R-squared:	0.180			
Model:	OLS	Adj. R-squared:	0.175			
Method:	Least Squares	F-statistic:	35.32			
Date:	Sat, 06 Dec 2025	Prob (F-statistic):	1.07e-26			
Time:	15:28:38	Log-Likelihood:	531.4			

05 데이터 분석 : ANOVA 결과

교육 소비 비중 ANOVA

검정 통계량 (F) : 24.45
p-value : 0.000000≈ 0.000

↓
유의수준 0.05하에서 귀무가설 기각

권역에 따라 교육비중의 평균에 유의미한 차이가 있다.
이는 권역별 상권의 교육 소비 구조가 서로 다름을 의미

여가 소비 비중 ANOVA

검정 통계량 (F) : 43.98
p-value : 0.000000≈ 0.000

↓
유의수준 0.05하에서 귀무가설 기각

권역에 따라 여가비중의 평균에 유의미한 차이가 있다.
이는 권역별 상권의 여가 소비 구조가 서로 다름을 의미

의료 소비 비중 ANOVA

검정 통계량 (F) : 70.62
p-value : 0.000000≈ 0.000

↓
유의수준 0.05하에서 귀무가설 기각

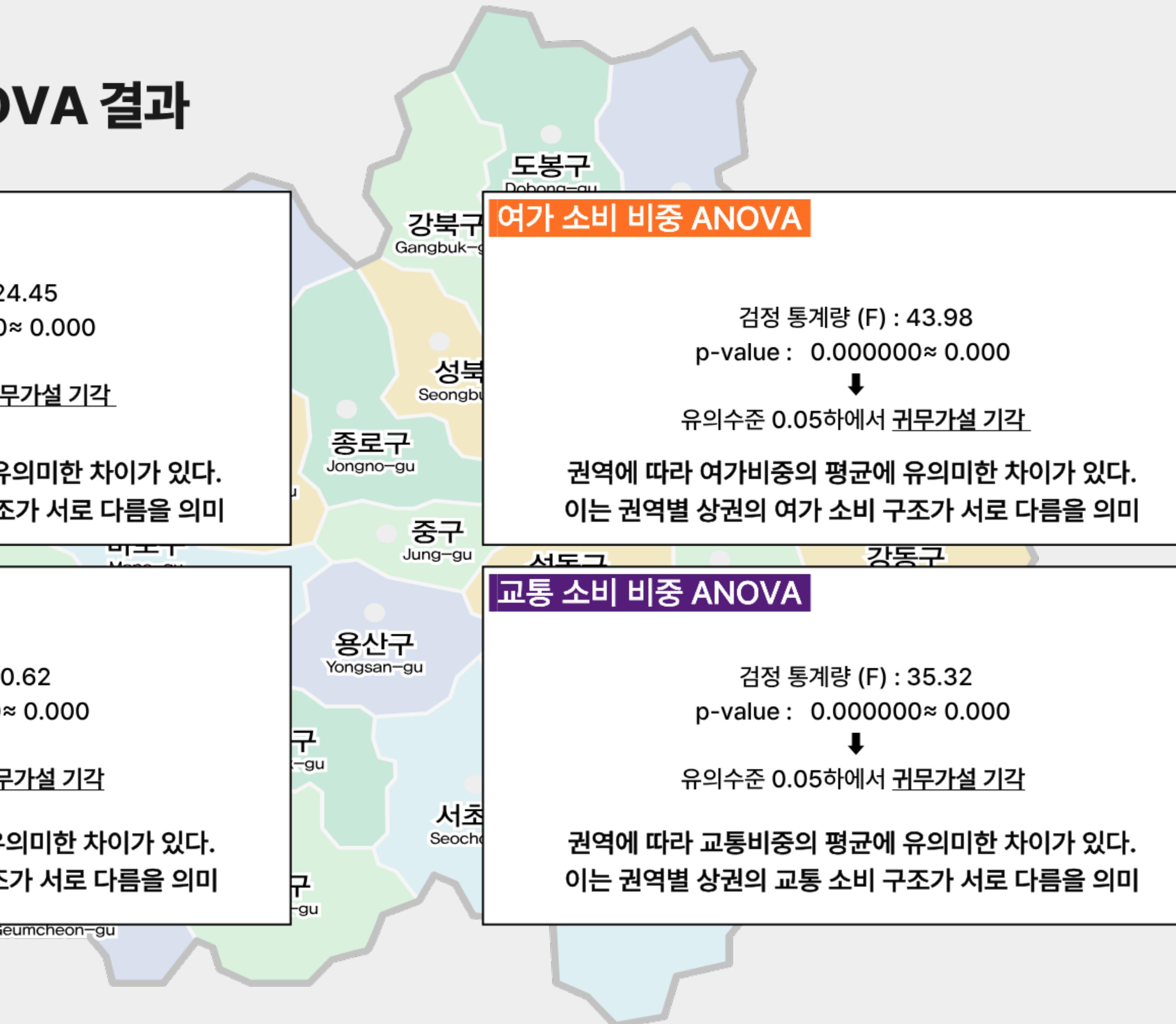
권역에 따라 의료비중의 평균에 유의미한 차이가 있다.
이는 권역별 상권의 의료 소비 구조가 서로 다름을 의미

교통 소비 비중 ANOVA

검정 통계량 (F) : 35.32
p-value : 0.000000≈ 0.000

↓
유의수준 0.05하에서 귀무가설 기각

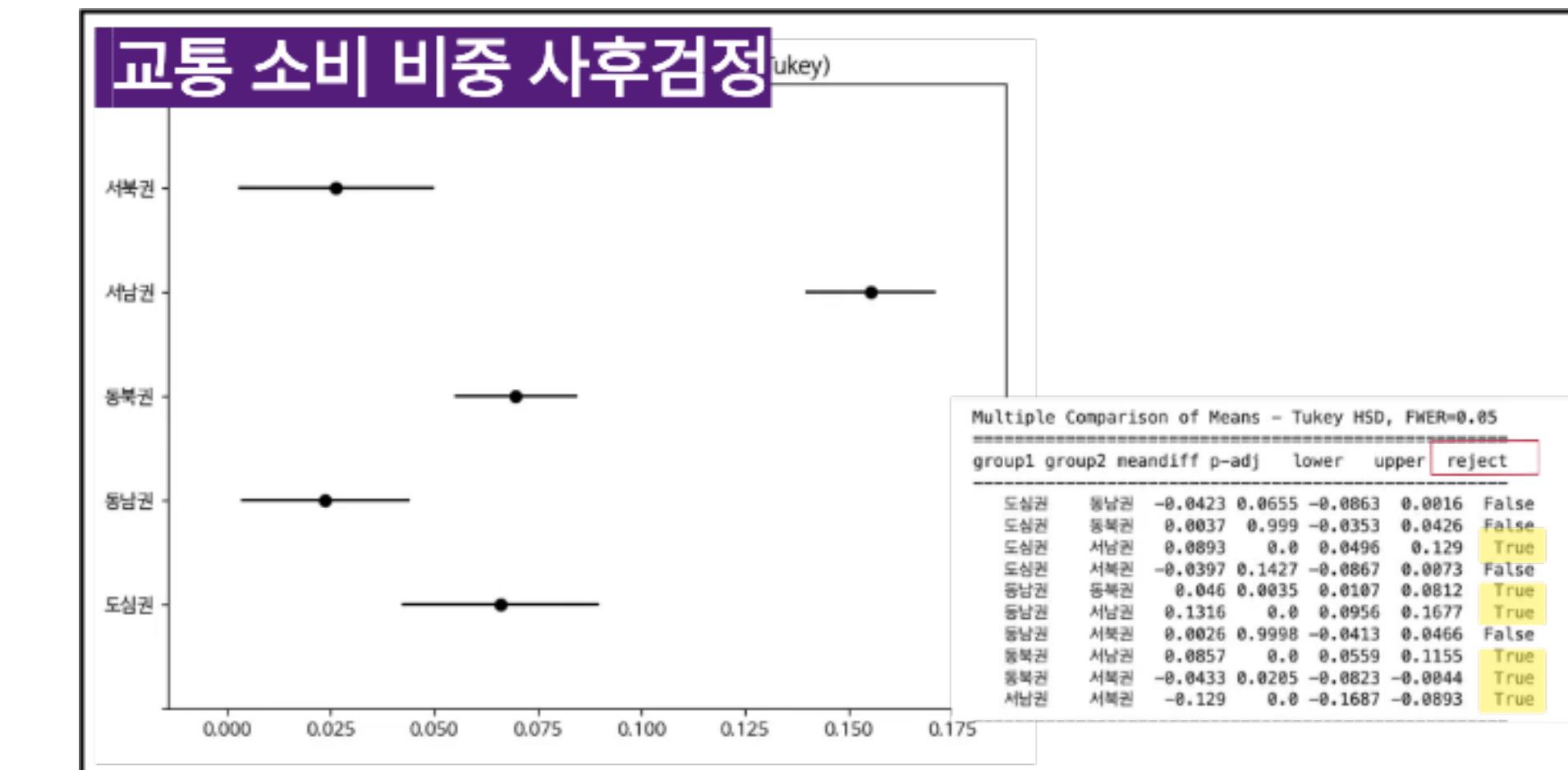
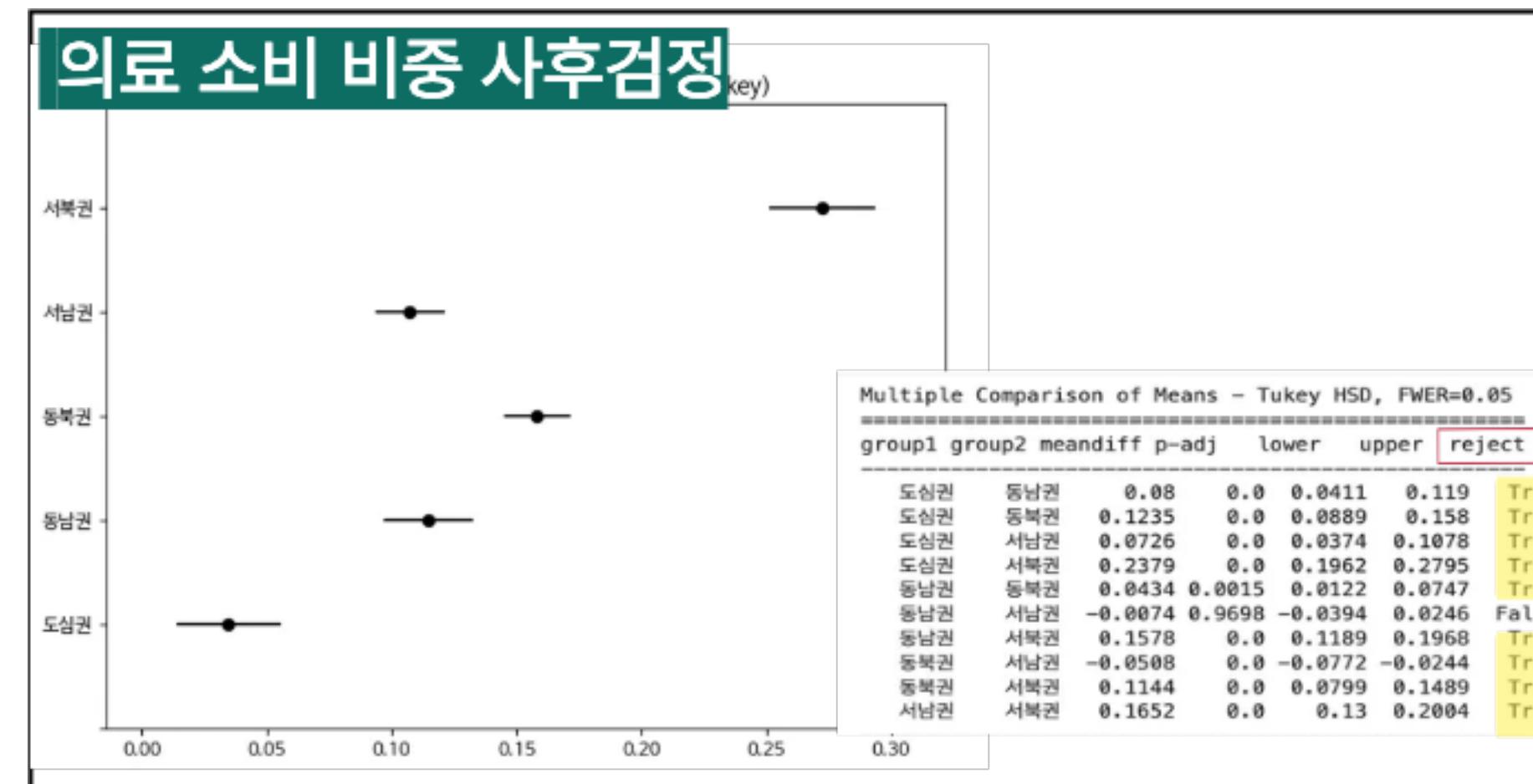
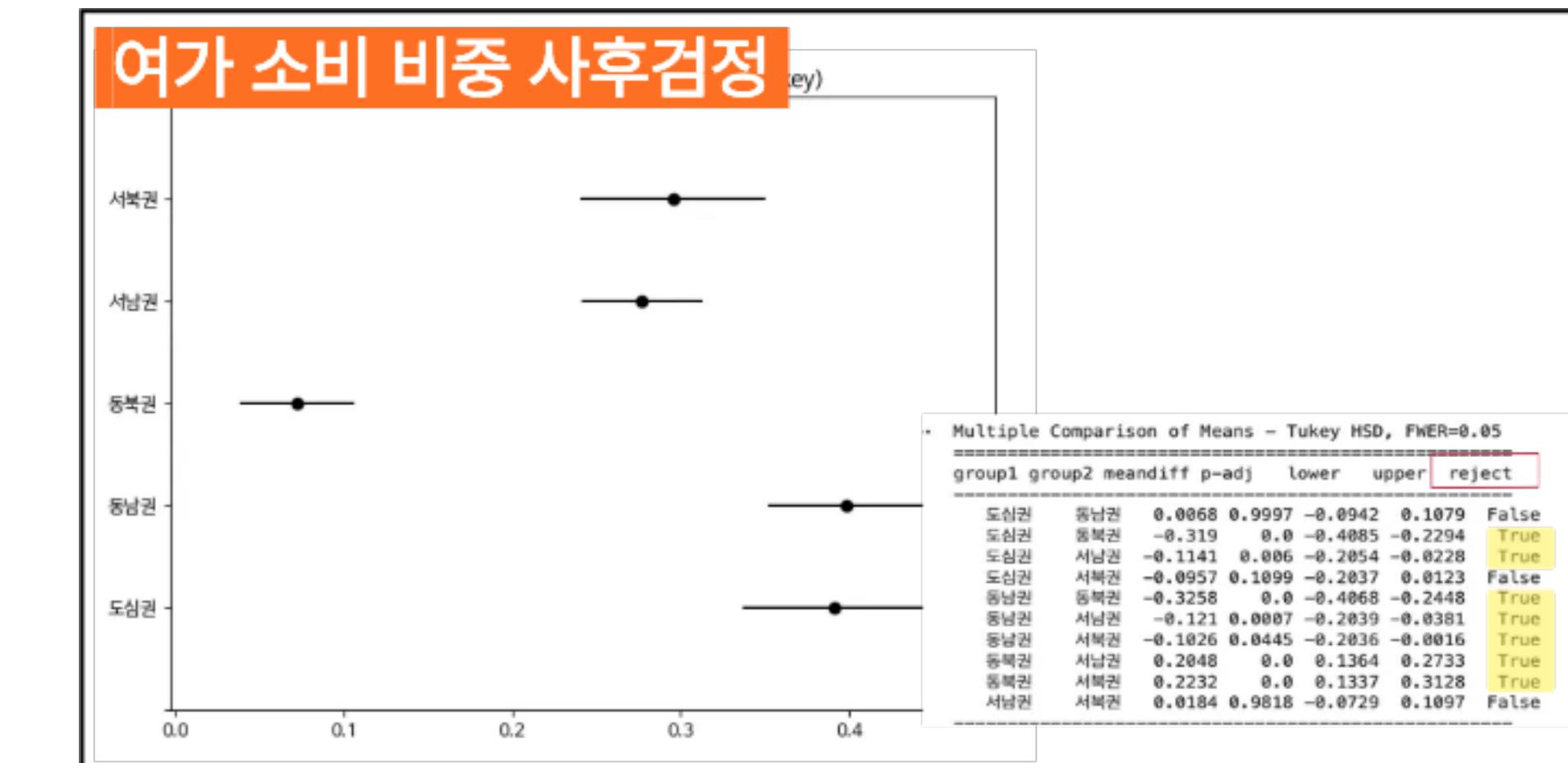
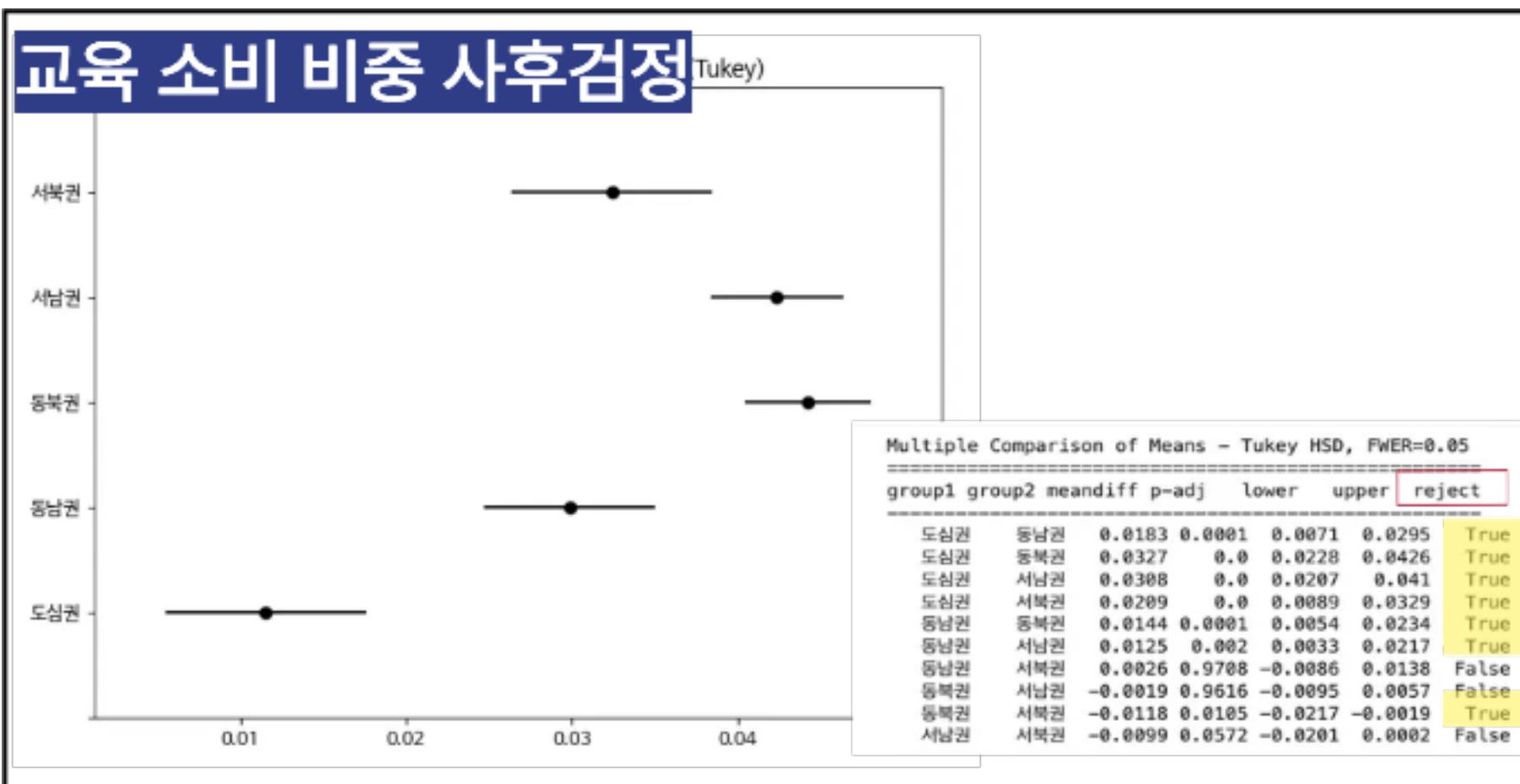
권역에 따라 교통비중의 평균에 유의미한 차이가 있다.
이는 권역별 상권의 교통 소비 구조가 서로 다름을 의미



05 데이터 분석 : ANOVA 결론



05 데이터 분석 : 사후검정 - Tukey HSD 수행

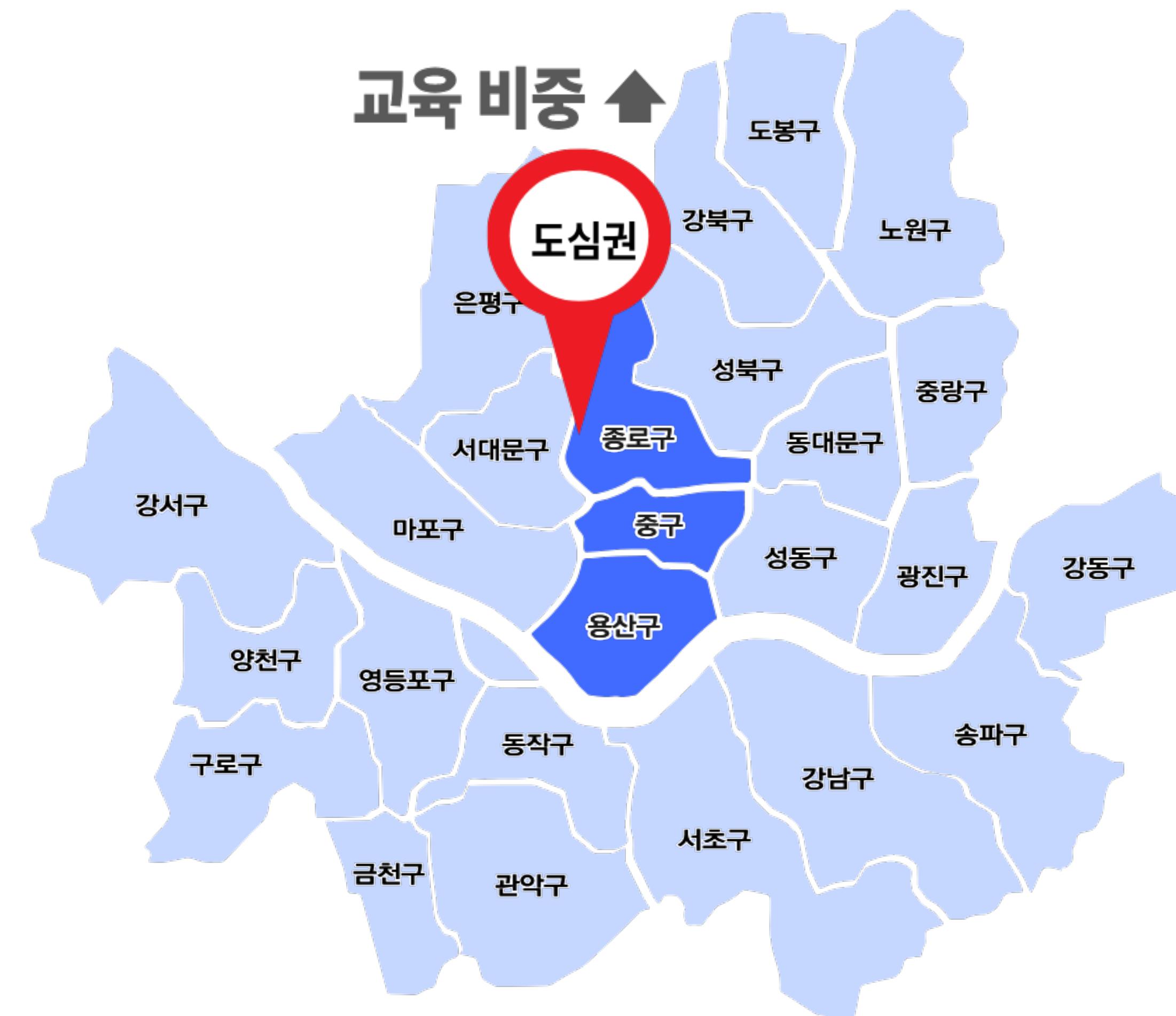


05 데이터 분석 : 사후검정 - Tukey HSD 결론

교육 소비 비중 사후검정

- 도심권이 모든 다른 권역보다 교육비중이 높음
- 서북권은 전체적으로 낮은 편
- 동남권, 동북권, 서남권도 서로 유의한 차이를 보임

도심권 > 기타 권역 순으로 높은 경향
권역 간 교육 소비 비중 차이는 뚜렷하다



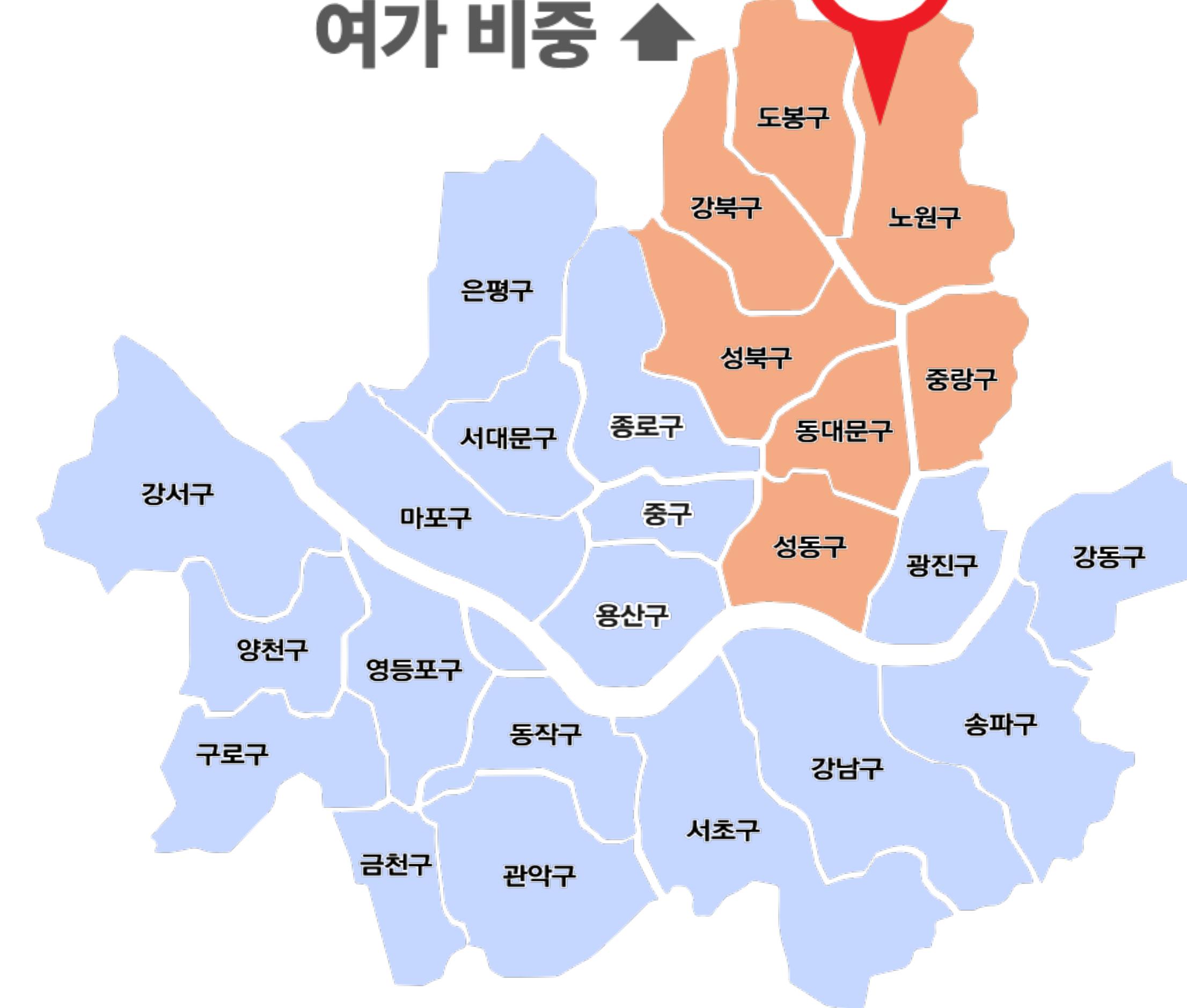
05 데이터 분석 : 사후검정 - Tukey HSD 결론

여가 소비 비중 사후검정

- 동북권이 여가비중에서 압도적으로 높음
- 도심권, 동남권은 상대적으로 낮음
(특히 동북권-도심권/ 동북권-동남권 큰 차이)
- 서북권은 중간 수준

동북권 > 서북권 > 기타 권역 순으로 높은 경향
권역 간 여가 소비 비중 차이는 뚜렷하다

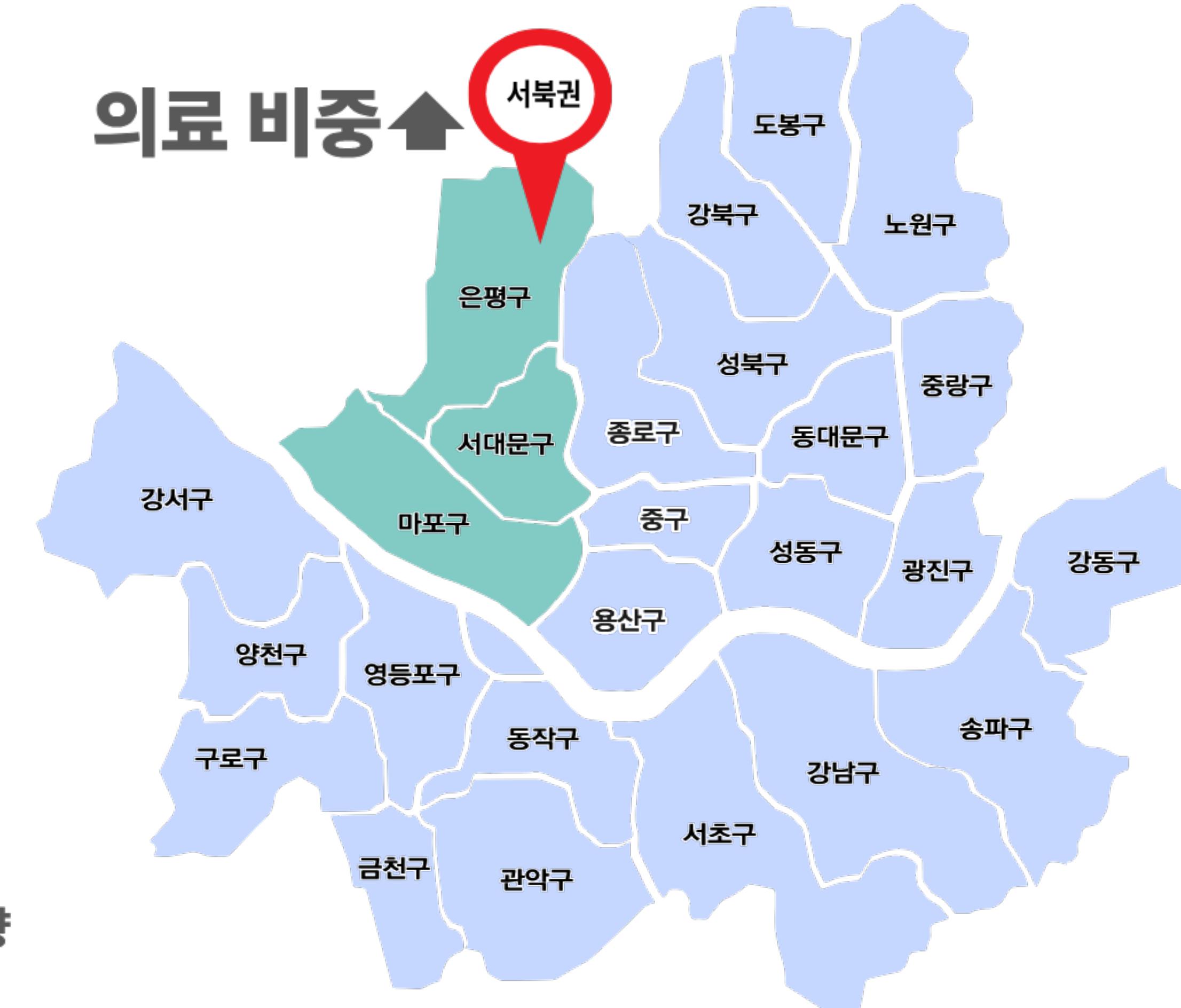
여가 비중 ↑



05 데이터 분석 : 사후검정 - Tukey HSD 결론

의료 소비 비중 사후검정

의료 소비 비중은 가장 강한 차이가 나타남
-서북권이 의료비중 가장 유의하게 높음
-동남권도 동북권, 서남권보다 유의하게 높음
-도심권도 전체적으로 높은 편



서북권 > 동남권 > 도심권 > 기타 권역 순으로 높은 경향
권역 간 의료 소비 비중 차이는 뚜렷하다

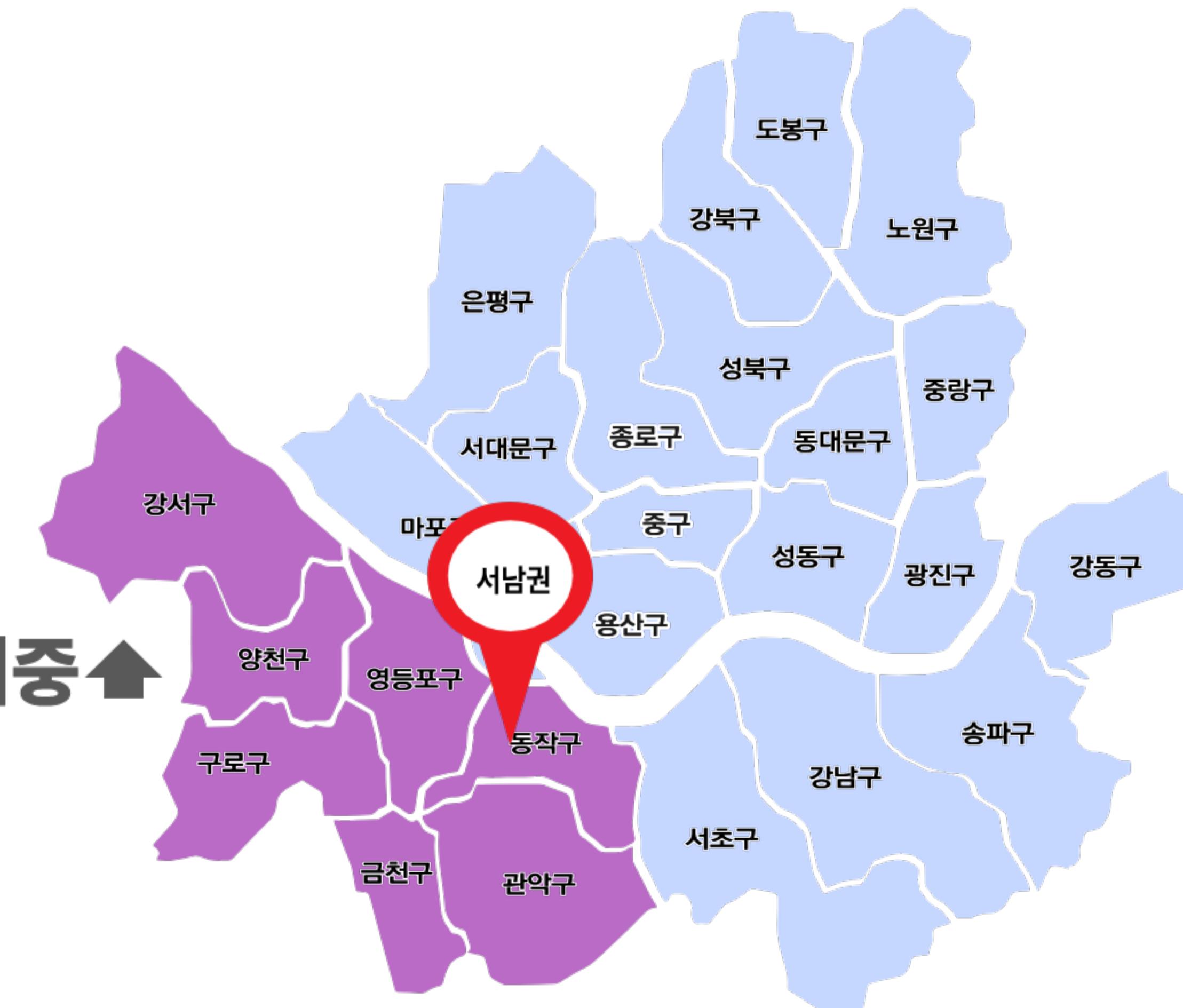
05 데이터 분석 : 사후검정 - Tukey HSD 결론

교통 소비 비중 사후검정

- 교통비중은 다른 항목보다 차이가 비교적 작은 편
- 서남권이 가장 유의하게 높음
 - 동남권은 교통비중이 높음(동북권, 서남권보다 높음)
 - 도심권은 중간 수준

교통 비중 ↑

서남권 > 동남권 > 기타 권역 순으로 높은 경향
권역 간 교통 소비 비중 차이는 뚜렷하다



06 결론 및 시사점

데이터 분석 결과

ANOVA 분석을 통해 서울 5개 권역 간 소비 항목에서 유의한 차이

Tukey HSD 사후 검정을 통한 권역별 소비 패턴 특징 파악

ANOVA 결과와 Tukey HSD 결과는 서로 일관

교육비중은 도심권이 가장 높다
여가비중은 동북권이 가장 높다
의료비중은 서북권, 교통비중은 서남권이 두드러진다.

결론

본 연구에서는 서울시 5대 권역의 소비 항목 비중 차이를 분석한 결과 – 일부 항목에서 권역 간 유의한 차이가 나타났다.

지역 인구 구성 / 상권 규모 / 문화 / 여가 인프라등의 차이가 소비 패턴에 반영되었음을 확인할 수 있다.

정책 및 시사점

권역별 소비 특성을 고려한
지역 맞춤형 소비 전략 수립 가능

생활 인프라의 균형 발전
정책 근거로 활용

인구·상권·문화 인프라 차이를
반영한 서비스 재배치 필요

지역별 취약항목에 대한
투자 우선순위 설정 가능