

# 서울시 자치구별 특성에 따른 음식물쓰레기 감량 정책 제언

군집 분석을 중심으로

5조

17101982 장세정  
17101946 김준영

16101334 한승구  
19101967 강채원

# 목차

## 1. 연구목적

자치구별 특성에 따른  
군집별 음식물 쓰레기  
감량 정책 제시

## 2. 기존 연구 조사 및 본 연구의 독창성

배출량에 영향을 미치는 변수를  
정의해둔 기존의 연구들.  
자치구별 분석의 독창성

## 3. 연구 내용

사용 데이터셋  
DTW군집화, 회귀를 통한  
유의미한 변수 추출

## 4. 연구 결과

군집별 예측 결과와  
그에 따른 맞춤 정책 제언

## 5. 한계점 및 향후 연구

이번 연구의 한계점과  
이후 진행해 볼 연구

# 1. 연구 목적

---



목표:

자치구별 특성에 따른  
군집별 음식물 쓰레기  
감량 정책 제시

---

1. 자치구별 특성 파악

---

2. 군집화

---

3. 맞춤 정책 제언

---

## 2. 기존 연구조사 및 본 연구의 독창성 및 우수성

### 기존 연구

전국 지역 혹은 광역시 위주로 연구 진행

대부분이 다중회귀분석을 통한 예측

### 본 연구의 독창성

서울특별시의 자치구를 군집화

시계열을 이용한 군집분석

기존 사용 변수 이외의 다양한 변수 데이터 사용

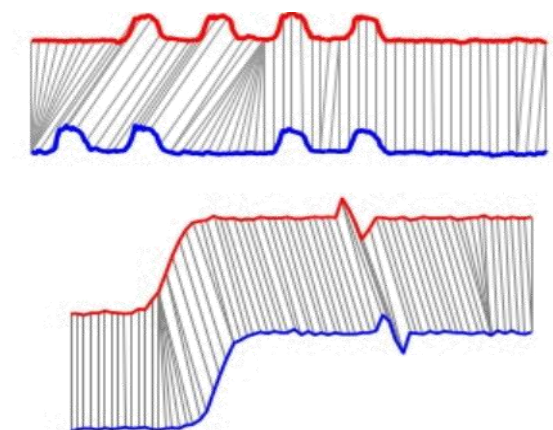
## 3-1. 연구 내용 - DTW 군집화

### 연구의 목적

- 자치구별 배출량 데이터의 분포를 살펴보니 **자치구별로 증감패턴이 다양했음.**  
-> **자치구별로 특성이 다르기 때문**이라고 예상함.
- 자치구별로 효율적 관리를 위해 **배출량 패턴이 비슷한 것끼리 군집화한 후, 군집별 중요 변수를 추출해서 군집별 특성을 파악**해보고자 함.

### 활용할 군집화 알고리즘 - DTW Algorithm

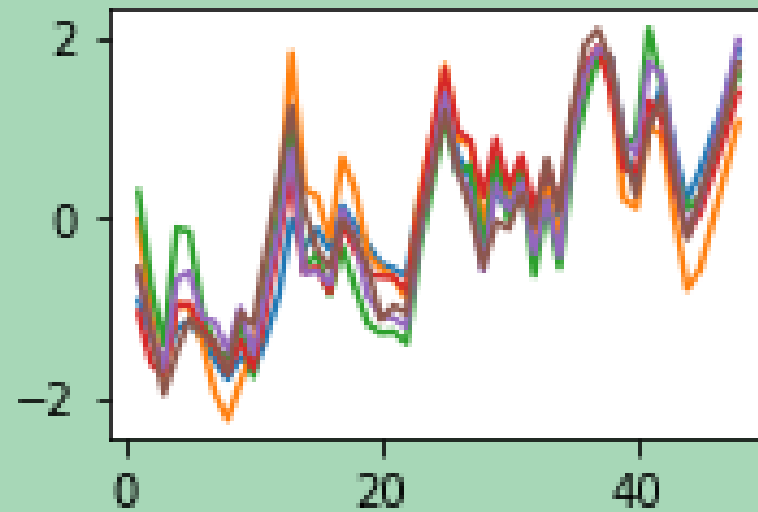
두 시계열 간의 거리를 최소화하는 방향으로 움직이면서 시계열 데이터의 패턴 시작점과 속도가 다르더라도 군집화 하는데 유용한 알고리즘



# 3-1. 연구 내용 – DTW 군집화

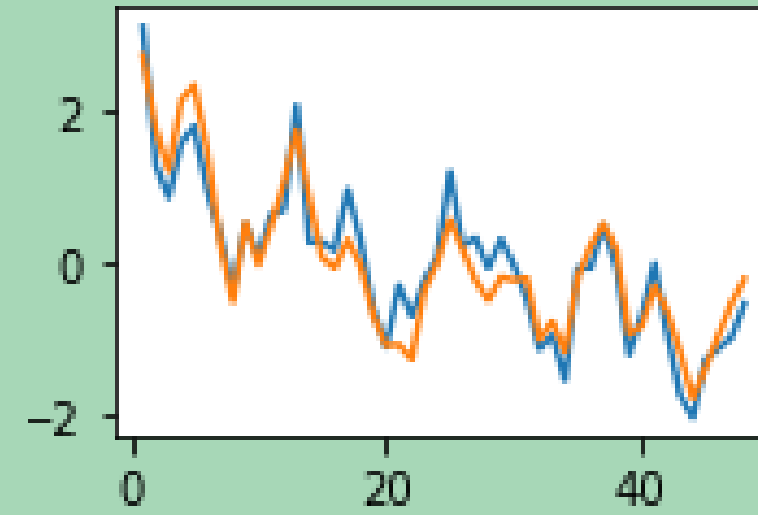
## 1) 배출량 패턴이 비슷한 자치구끼리 군집화

변화가 큰 증가 패턴의 군집



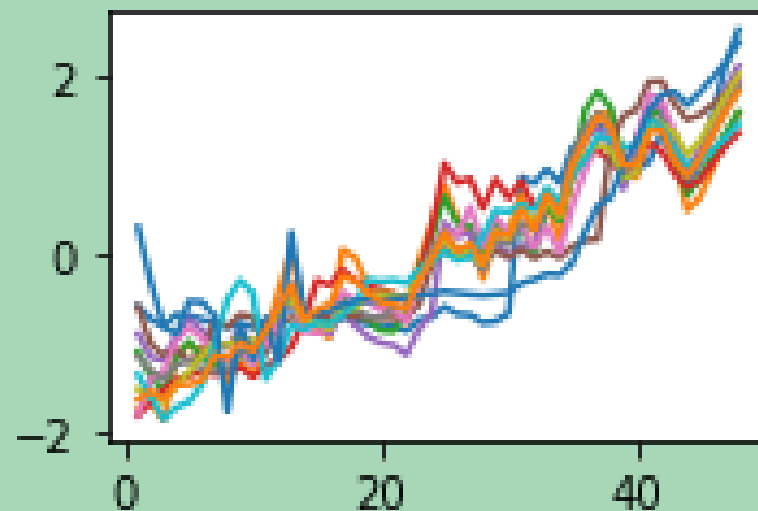
종로구, 중구, 중랑구,  
노원구, 동작구, 서초구

감소하는 패턴 군집



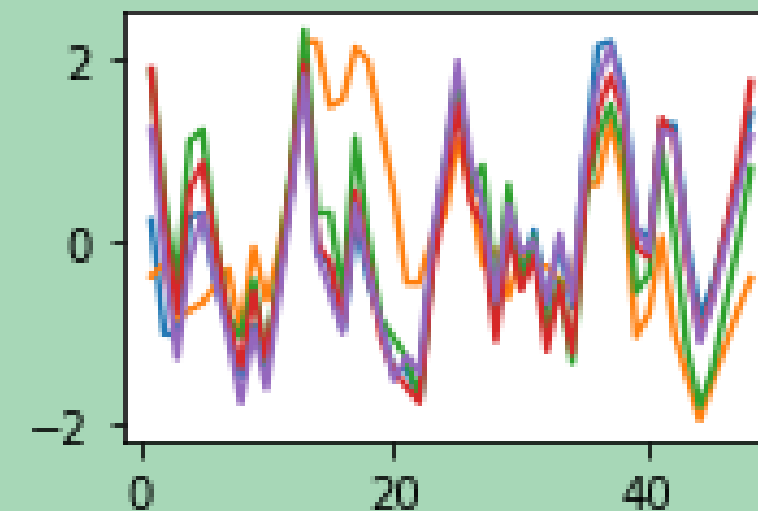
도봉구, 마포구

꾸준한 증가 패턴의 군집



용산구, 성동구, 광진구,  
동대문구, 성북구, 강북구,  
은평구, 강서구, 구로구,  
관악구, 강남구, 강동구

변화가 크지만 유지 패턴 군집



서대문구, 양천구, 금천구,  
영등포구, 송파구

# 3-1. 연구 내용 – DTW 군집화

## 2) 군집의 중요 변수 추출 후 군집별 특성 파악

자치구명 변수 정의

시군구	변수명	서대문구	Sdm
종로구	Jongno	마포구	Mapo
중구	Jung	양천구	Yangcheon
용산구	Yongsan	강서구	Gangseo
성동구	Sd	구로구	Guro
광진구	Gwangjin	금천구	Geumcheon
동대문구	Ddm	영등포구	Ydp
중랑구	Jungnang	동작구	Dongjak
성북구	Sb	관악구	Gwanak
강북구	Gangbuk	서초구	Seocho
도봉구	Dobong	강남구	Gangnam
노원구	Nowon	송파구	Songpa
은평구	Ep	강동구	Gangdong

분석에 활용할 변수 정의

총인구	Population	1년생존율	Survival1	주거지역	live
세대수	HouseholdNum	소득수준	Income	사업지역	company
세대당인구	HouseholdPer	개업수	OpenNun	공업지역	factory
65세이상	Old	개업률	OpenPer	녹지지역	green
18세이하	Young	배출량	Outcome		
1인가구비율	Single	공휴일	Holiday		
폐업수	CloseNum	용도면적	Area		
폐업률	ClosePer	승용차	Car		
점포수	Store	승합차	Van		
임대시세	Rental	화물차	FreightCar		
5년생존율	Survival4	이륜차	TwoCar		
3년생존율	Survival3	특수차	SpecialCar		

추후 특성분석에 활용할 변수 분류

인구특성	상업특성	도시특성
총인구수	폐업수	자동차등록대수
세대수	점포수	특수차
노인	임대시세	
청소년	생존율	
1인가구	개업수	

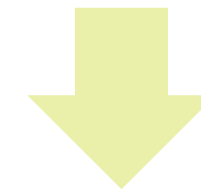
# 3-1. 연구 내용 - DTW 군집화

## 2) 군집의 중요 변수 추출 후 군집별 특성 파악

### OLS 분석 결과 예시

OLS Regression Results						
Dep. Variable:	Outcome		R-squared:	0.658		
Model:	OLS		Adj. R-squared:	0.608		
Method:	Least Squares		F-statistic:	13.17		
Date:	Wed, 01 Dec 2021		Prob (F-statistic):	3.02e-08		
Time:	23:34:43		Log-Likelihood:	-40.349		
No. Observations:	48		AIC:	94.70		
Df Residuals:	41		BIC:	107.8		
Df Model:	6					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	-8.327e-17	0.088	-9.51e-16	1.000	-0.177	0.177
CloseNum	0.4698	0.175	2.679	0.011	0.116	0.824
Store	-0.3076	0.225	-1.366	0.179	-0.762	0.147
Rental	1.2540	0.607	2.064	0.045	0.027	2.481
Survival5	0.8137	0.372	2.190	0.034	0.063	1.564
Income	-1.9320	1.286	-1.502	0.141	-4.530	0.665
Car	-0.1307	0.380	-0.344	0.733	-0.897	0.636
Omnibus:	0.487		Durbin-Watson:	1.216		
Prob(Omnibus):	0.784		Jarque-Bera (JB):	0.609		
Skew:	0.040		Prob(JB):	0.737		
Kurtosis:	2.454		Cond. No.	20.8		

OLS분석을 통해 각 군집별 중요 변수 추출 진행 후 결과가  
자치구별 중요 변수 결과와 **상이했음**



배출량에 영향 미치는 변수가 **자치구별로 각기 다를 수 있으므로**  
배출량 패턴으로 묶은 군집에 대한 OLS로 유의미한 변수 찾기는  
**우리의 목적에 맞지 않다고 판단**



# 3-1. 연구 내용 - DTW 군집화

## 2) 군집의 중요 변수 추출 후 군집별 특성 파악

### OLS 분석 결과 예시

OLS Regression Results			
Dep. Variable:	Outcome	R-squared:	0.658
Model:	OLS	Adj. R-squared:	0.608
Method:	Least Squares	F-statistic:	13.17
Date:	Wed, 01 Dec 2021	Prob (F-statistic):	3.02e-08
Time:	23:34:43	Log-Likelihood:	-40.349
No. Observations:	48	AIC:	94.70
Df Residuals:	41	BIC:	107.8
Df Model:	6		
Covariance Type:	nonrobust		
	coef	std err	
Intercept	-8.327e-17	0.088	
CloseNum	0.4698	0.175	
Store	-0.3076	0.225	
Rental	1.2540	0.607	
Survival5	0.8137	0.372	
Income	-1.9320	1.286	
Car	-0.1307	0.380	
Omnibus:		0.41	
Prob(Omnibus):		0.71	
Skew:		0.0	
Kurtosis:		2.41	

OLS분석을 통해 각 군집별 중요 변수 추출 진행 후 결과가 자치구별 중요 변수 결과와 **상이했음**

### 본 연구의 목표

- 1) 배출량 패턴이 자치구별로 왜 다른지, 어떻게 다른지, 어떤 변수 때문인지, 배출량에 어떤 영향을 미치는지 파악
- 2) 공통점이 있는 자치구끼리 군집화
- 3) 군집별로 효율적 관리, 감량 정책 제언

이 다를 수 있으므로 의미한 변수 찾기는

# 3-1. 연구 내용 – 자치구별 변수 추출

## Stepwise model 코드

```

1 # 단계적선택법 모델
2 def Stepwise_model(X,y):
3     Stepmodels = pd.DataFrame(columns=["AIC","model"])
4     tic = time.time()
5     predictors = []
6     Smodel_before = processSubset(X,y,predictors + ['const'])['AIC']
7     for i in range(1, len(X.columns.difference(['const'])) + 1):
8         Forward_result = forward(X=X, y=y, predictors = predictors)
9         print('forward')
10        Stepmodels.loc[i] = Forward_result
11        predictors = Stepmodels.loc[i]["model"].model.exog_names
12        predictors = [k for k in predictors if k != 'const']
13        Backward_result = backward(X=X,y=y, predictors = predictors)
14        if Backward_result['AIC'] < Forward_result['AIC']:
15            Stepmodels.loc[i] = Backward_result
16            predictors = Stepmodels.loc[i]["model"].model.exog_names
17            Smodel_before = Stepmodels.loc[i]["AIC"]
18            predictors = [k for k in predictors if k != 'const']
19            print('backward')
20        if Stepmodels.loc[i]['AIC'] > Smodel_before:
21            break
22        else:
23            Smodel_before = Stepmodels.loc[i]["AIC"]
24    toc = time.time()
25    print("Total elapsed time:", (toc - tic), "seconds.")
26    return Stepmodels['model'][len(Stepmodels['model'])-1],predictors

```

## 변수 선별



종로구	OpenPer, Store, OpenNum, Survival3, Rental, FreightCar, Old
중구	HouseholdNum, OpenPer, Single, Young, Survival3
용산구	Store, Young, HouseholdNum, Population, Single, FreightCar
성동구	Old, FreightCar, Young
광진구	Old, Survival3
동대문구	Young, Population, Old
중랑구	FreightCar, OpenNum, HouseholdNum
성북구	Old, Survival3, Rental, OpenPer
강북구	FreightCar, Population, Young, HouseholdNum, Store, Old, OpenPer
도봉구	Young, FreightCar, OpenPer, OpenNum, Survival3
노원구	Young, Population, Store, Rental, Single, HouseholdNum, Old
은평구	HouseholdNum, FreightCar, Store, Young, Population
서대문구	Store, OpenPer, OpenNum
마포구	Young, Rental, Single, OpenNum, OpenPer
양천구	Store
강서구	Old, Rental
구로구	Old, FreightCar, HouseholdNum
금천구	Population
영등포구	Store, OpenNum
동작구	FreightCar
관악구	HouseholdNum, OpenPer, Rental
서초구	Rental, HouseholdNum, Single, Young, Store, Population
강남구	Survival3, HouseholdNum
송파구	Survival3, OpenPer
강동구	Old, Store, Rental, OpenNum

Stepwise를 통한 각 자치구별 변수 선별

=> 다중공선성 문제를 해결하기 위해 VIF 검증과 OLS 분석을 추가로 진행

# 3-1. 연구 내용 - 자치구별 변수 추출

## VIF - OLS 변수추출 코드

```

1 def olsVif(df, dfColumns, features):
2     """
3     ols와 vif를 return 해주는 함수
4     ols는 객체를 받아 summary()를 통해 확인할 수 있다.
5     vif는 return 값으로 받아볼 수 있다.
6     """
7     # scaler = StandardScaler()
8     # df = scaler.fit_transform(df)
9     # df = pd.DataFrame(df, columns = dfColumns)
10
11     y, X = dmatrices('Outcome ~' + features, data = df, return_type = 'dataframe')
12
13     vif = pd.DataFrame()
14     vif['VIF Factor'] = [variance_inflation_factor(X.values, i) for i in range(X.shape[1])]
15     vif['features'] = X.columns
16
17     result = sm.OLS(y, X).fit()
18
19     return result, vif

```

Stepwise를 통해 선별한 변수 중  
**VIF < 10** 이고 **P-value < 0.05** 인 변수를  
 최종변수로 선정

## OLS 결과 예시

Sb

### OLS Regression Results

```

=====
Dep. Variable:      Outcome    R-squared:      0.893
Model:              OLS        Adj. R-squared:  0.886
Method:             Least Squares    F-statistic: 123.0
Date:               Wed, 08 Dec 2021    Prob (F-statistic): 2.07e-21
Time:               15:38:09    Log-Likelihood: -14.373
No. Observations:   48    AIC: 36.75
Df Residuals:       44    BIC: 44.23
Df Model:            3
Covariance Type:    nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	-2.949e-17	0.049	-5.99e-16	1.000	-0.099	0.099
Population	-0.3460	0.107	-3.224	0.002	-0.562	-0.130
Old	0.8954	0.055	16.155	0.000	0.784	1.007
Rental	-0.3778	0.102	-3.717	0.001	-0.583	-0.173

```

=====
Omnibus:      0.790    Durbin-Watson:      0.857
Prob(Omnibus): 0.674    Jarque-Bera (JB):      0.844
Skew:         0.165    Prob(JB):      0.656
Kurtosis:     2.440    Cond. No.      4.25
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

	VIF Factor	features
0	1.000000	Intercept
1	4.754330	Population
2	1.268490	Old
3	4.265842	Rental

# 3-1. 연구 내용 - 자치구별 변수 추출

자치구별로 추출된 유의미한 변수

	Population	Household	Old	Young	Single	Store	Rental	Survival3	OpenNum	OpenPer	FreightCar	Cnt
Jongno	0	0	1	0	0	0	0	0	1	0	0	2
Jung	0	0	0	0	0	0	0	0	0	1	0	1
Yongsan	1	1	0	0	0	1	0	0	0	0	0	3
Sd	0	0	0	1	0	0	0	0	0	0	1	2
Gwangjin	0	0	1	0	0	0	0	0	0	0	0	1
Ddm	0	0	1	0	0	0	0	0	0	0	0	1
Jungnang	0	0	0	0	0	0	0	0	1	0	0	1
Sb	0	0	1	0	0	0	0	1	0	1	0	3
Gangbuk	1	0	0	0	0	1	0	0	0	0	0	2
Dobong	0	0	0	0	0	0	0	1	0	0	0	1
Nowon	0	0	0	0	1	0	0	0	0	0	0	1
Ep	1	0	0	0	1	0	0	0	0	0	0	2
Sdm	0	0	0	0	0	1	0	0	0	0	0	1
Mapo	0	0	0	1	0	0	1	0	0	0	0	2
Yangcheo	0	0	0	0	0	1	0	0	0	0	0	1
Gangseo	0	0	1	0	0	0	1	0	0	0	0	2
Guro	0	0	0	0	0	0	0	0	0	0	1	1
Geumcheo	0	0	0	0	0	0	0	0	0	0	0	0
Ydp	0	0	0	0	0	1	0	0	0	0	0	1
Dongjak	0	0	0	0	0	0	0	0	0	0	1	1
Gwanak	0	0	0	0	1	0	0	0	0	0	0	1
Seocho	0	0	0	0	1	0	0	0	0	0	0	1
Gangnam	0	0	0	0	0	0	0	1	0	0	0	1
Songpa	0	0	0	0	0	0	0	1	0	0	0	1
Gangdong	0	0	0	0	0	1	1	0	0	0	0	2

파랑 : 양의 상관관계 (+)

빨강 : 음의 상관관계 (-)

-> 유의미한 변수들로 군집화 진행

-> 변수가 추출되지 않은 자치구

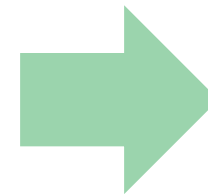
## 3-2. 연구 내용 - 분석과정 중 의문점

문제점 : 추출된 변수의 수가 적어 군집별 해석이 어려움

- 1) 혹시 **코로나의 영향이 분석에 영향**을 끼치지 않았을까?  
 -> 2017.07 ~ 2019.12 기간의 데이터로 똑같이 분석진행 해봄

코로나 이전 데이터로 OLS 분석 후 변수 추출 결과

	Population	Household	Old	Young	Single	Store	Rental	Survival	FreightCar	Cnt
Jongno	0	0	1	0	0	0	0	0	0	1
Jung	0	0	0	1	0	0	0	0	1	2
Yongsan	0	0	0	0	0	0	0	0	0	0
Sd	0	0	1	0	0	0	0	0	0	1
Gwangjin	1	0	0	0	0	0	0	0	0	1
Ddm	0	1	0	0	0	0	0	1	0	2
Jungnang	0	0	0	0	0	0	0	0	1	1
Sb	0	0	1	0	0	0	1	0	0	2
Gangbuk	0	0	0	0	0	0	0	0	1	1
Dobong	0	0	0	1	0	0	0	1	1	3
Nowon	0	0	0	0	1	0	0	0	0	1
Ep	0	0	0	0	0	1	0	0	1	2
Sdm	0	0	0	0	0	0	0	0	0	0
Mapo	0	0	0	0	0	0	0	1	1	2
Yangcheo	0	0	0	0	0	1	1	0	0	2
Gangseo	0	0	0	0	0	0	0	1	1	2
Guro	1	1	0	0	0	0	0	0	0	2
Geumcheo	0	0	0	0	0	0	0	0	0	0
Ydp	0	0	0	0	0	0	0	0	0	0
Dongjak	0	0	0	0	0	0	0	0	1	1
Gwanak	0	0	0	1	0	0	0	0	0	1
Seocho	0	0	0	0	0	0	0	0	1	1
Gangnam	0	0	0	0	0	0	0	0	0	0
Songpa	0	0	0	0	0	0	0	0	0	0
Gangdong	0	0	1	0	0	1	0	0	0	2
Cnt	2	2	4	3	1	3	2	4	9	30



유의미한 변수 추출이 잘 진행이 안됨  
 역시 뚜렷한 결과 해석이 어려움

코로나19상황이 변수 추출에 영향을 미치지 않는다고 판단함



## 3-2. 연구 내용 - 분석과정 중 의문점

2) 사례연구를 참고하여 변수특성에 따라 나누어 회귀 분석을 진행해보면 다를까?

### 인구 특성

	Population	Household	Old	Young	Single	Cnt
Jongno	0	0	0	1	0	1
Jung	0	0	1	0	0	1
Yongsan	1	1	0	0	0	2
Sd	0	0	1	0	0	1
Gwangjin	0	0	1	0	0	1
Ddm	0	0	1	0	0	1
Jungnang	0	0	0	0	1	1
Sb	0	0	1	0	0	1
Gangbuk	0	0	1	0	0	1
Dobong	0	0	0	1	0	1
Nowon	0	0	0	0	1	1
Ep	0	0	1	0	0	1
Sdm	0	0	1	0	0	1
Mapo	0	0	1	0	0	1
Yangcheo	0	1	0	0	0	1
Gangseo	0	0	1	0	0	1
Guro	0	0	1	0	0	1
Geumcheong	0	0	0	0	0	0
Ydp	0	0	0	0	0	0
Dongjak	0	0	0	1	0	1
Gwanak	0	0	1	0	0	1
Seocho	0	0	1	0	0	1
Gangnam	1	1	0	0	0	2
Songpa	0	0	1	0	0	1
Gangdong	0	0	1	1	0	2
Cnt	2	3	15	4	2	26

### 상업 특성

	CloseNum	Store	Rental	Survival5	Survival3	Survival1	OpenNum	Cnt
Jongno	0	1	0	1	0	1	1	4
Jung	1	0	0	1	1	0	0	3
Yongsan	1	1	0	1	1	1	0	5
Sd	0	1	1	0	1	0	1	4
Gwangjin	1	1	0	1	1	1	1	6
Ddm	1	1	1	0	0	1	1	5
Jungnang	1	1	0	1	1	1	0	5
Sb	0	1	0	1	1	1	1	5
Gangbuk	1	0	1	1	0	1	0	4
Dobong	1	0	1	1	1	0	0	4
Nowon	0	1	1	0	1	0	0	3
Ep	0	1	1	0	0	1	1	4
Sdm	0	1	0	0	0	0	0	1
Mapo	1	1	0	1	1	0	0	4
Yangcheo	0	1	0	0	0	0	0	1
Gangseo	1	1	0	0	0	0	0	2
Guro	0	1	0	1	1	0	0	3
Geumcheong	0	0	0	0	0	0	0	0
Ydp	0	1	0	0	0	0	0	1
Dongjak	1	1	1	0	0	1	0	4
Gwanak	0	0	1	1	0	1	0	3
Seocho	0	0	0	0	1	0	0	1
Gangnam	0	0	0	0	1	0	0	1
Songpa	1	0	0	0	1	0	0	2
Gangdong	1	0	1	1	1	0	1	5
Cnt	12	16	9	12	14	10	7	80

### 도시 특성

	TotalCar	FreightCar	SpecialCar	Cnt
Jongno	0	1	0	1
Jung	0	1	0	1
Yongsan	0	1	0	1
Sd	1	1	1	3
Gwangjin	1	1	1	3
Ddm	1	1	1	3
Jungnang	1	0	1	2
Sb	1	0	0	1
Gangbuk	0	1	0	1
Dobong	1	1	0	2
Nowon	1	0	1	2
Ep	1	1	0	2
Sdm	0	1	0	1
Mapo	0	1	0	1
Yangcheo	0	0	0	0
Gangseo	1	1	0	2
Guro	0	1	1	2
Geumcheong	0	0	0	0
Ydp	0	0	0	0
Dongjak	0	1	0	1
Gwanak	0	1	0	1
Seocho	1	0	1	2
Gangnam	1	1	0	2
Songpa	0	0	1	1
Gangdong	1	1	1	3
Cnt	12	17	9	38

## 3-2. 연구 내용 - 분석과정 중 의문점

2) 사례연구를 참고하여 변수특성에 따라 나누어 회귀 분석을 진행해보면 다를까?

인구 특성

	Population	Household	Old	Young	Single	Cnt
Jongno	0	0	0	1	0	1
Jung	0	0	1	0	0	1
Yongsan	1	1	0	0	0	2
Sd	0	0	1	0	0	1
Gwangjin	0	0	1	0	0	1
Ddm	0	0	1	0	0	1
Jungnang	0	0	0	0	1	1
Sb	0	0	1	0	0	1
Gangbuk	0	0	1	0	0	1
Dobong	0	0	0	1	0	1
Nowon	0	0	0	0	1	1
Ep						
Sdm						
Mapo						
Yang						
Gang						
Guro						
Geum						
Ydp						
Dongj						
Gwana						
Seoch						
Gangnam	1	1	0	0	0	2
Songpa	0	0	1	0	0	1
Gangdong	0	0	1	1	0	2
Cnt	2	3	15	4	2	26

상업 특성

	CloseNum	Store	Rental	Survival5	Survival3	Survival1	OpenNum	Cnt
Jongno	0	1	0	1	0	1	1	4
Jung	1	0	0	1	1	0	0	3
Yongsan	1	1	0	1	1	1	0	5
Sd	0	1	1	0	1	0	1	4
Gwangjin	1	1	0	1	1	1	1	6
Ddm	1	1	1	0	0	1	1	5
Jungnang	1	1	0	1	1	1	0	5
Sb	0	1	0	1	1	1	1	5
Gangbuk	1	0	1	1	0	1	0	4
Dobong	1	0	1	1	1	0	0	4
Nowon	0	1	1	0	1	0	0	3
Seoch	0	0	0	0	1	0	0	1
Gangnam	0	0	0	0	1	0	0	1
Songpa	1	0	0	0	1	0	0	2
Gangdong	1	0	1	1	1	0	1	5
Cnt	12	16	9	12	14	10	7	80

도시 특성

	TotalCar	FreightCar	SpecialCar	Cnt
Jongno	0	1	0	1
Jung	0	1	0	1
Yongsan	0	1	0	1
Sd	1	1	1	3
Gwangjin	1	1	1	3
Ddm	1	1	1	3
Jungnang	1	0	1	2
Sb	1	0	0	1
Gangbuk	0	1	0	1
Dobong	1	1	0	2
Nowon	1	0	1	2
Seoch	1	0	1	2
Gangnam	1	1	0	2
Songpa	0	0	1	1
Gangdong	1	1	1	3
Cnt	12	17	9	38

모든 자치구의 특성을 추출하기 위해 변수들을 특성에 따라 나누어 각각 OLS 분석을 진행해보았으나, 여전히 **유의미한 변수가 추출되지 않은 자치구들이 존재함**

## 3-2. 연구 내용 - 최종 변수 선정

따라서, 모든 변수를 고려해서 OLS를 진행한 후의 변수추출 결과가 가장 적합하다고 판단하여 처음에 시도했던 방식대로 군집화 진행하기로 결정

최종으로 선정된 자치구별로 유의미한 변수

	Population	Household	Old	Young	Single	Store	Rental	Survival3	OpenNum	OpenPer	FreightCar	Cnt
Jongno	0	0	1	0	0	0	0	0	1	0	0	2
Jung	0	0	0	0	0	0	0	0	0	1	0	1
Yongsan	1	1	0	0	0	1	0	0	0	0	0	3
Sd	0	0	0	1	0	0	0	0	0	0	1	2
Gwangjin	0	0	1	0	0	0	0	0	0	0	0	1
Ddm	0	0	1	0	0	0	0	0	0	0	0	1
Jungnang	0	0	0	0	0	0	0	0	1	0	0	1
Sb	0	0	1	0	0	0	0	1	0	1	0	3
Gangbuk	1	0	0	0	0	1	0	0	0	0	0	2
Dobong	0	0	0	0	0	0	0	1	0	0	0	1
Nowon	0	0	0	0	1	0	0	0	0	0	0	1
Ep	1	0	0	0	1	0	0	0	0	0	0	2
Sdm	0	0	0	0	0	1	0	0	0	0	0	1
Mapo	0	0	0	1	0	0	1	0	0	0	0	2
Yangcheo	0	0	0	0	0	1	0	0	0	0	0	1
Gangseo	0	0	1	0	0	0	1	0	0	0	0	2
Guro	0	0	0	0	0	0	0	0	0	0	1	1
Geumcheo	0	0	0	0	0	0	0	0	0	0	0	0
Ydp	0	0	0	0	0	1	0	0	0	0	0	1
Dongjak	0	0	0	0	0	0	0	0	0	0	1	1
Gwanak	0	0	0	0	1	0	0	0	0	0	0	1
Seocho	0	0	0	0	1	0	0	0	0	0	0	1
Gangnam	0	0	0	0	0	0	0	1	0	0	0	1
Songpa	0	0	0	0	0	0	0	1	0	0	0	1
Gangdong	0	0	0	0	0	1	1	0	0	0	0	2

군집화 기준

- 1) 추출된 변수가 같아야 함.
- 2) 변수의 상관관계 부호가 같아야 함.

모든 자치구가 적어도 한 군집에 포함되어야 하는가? **NO!**

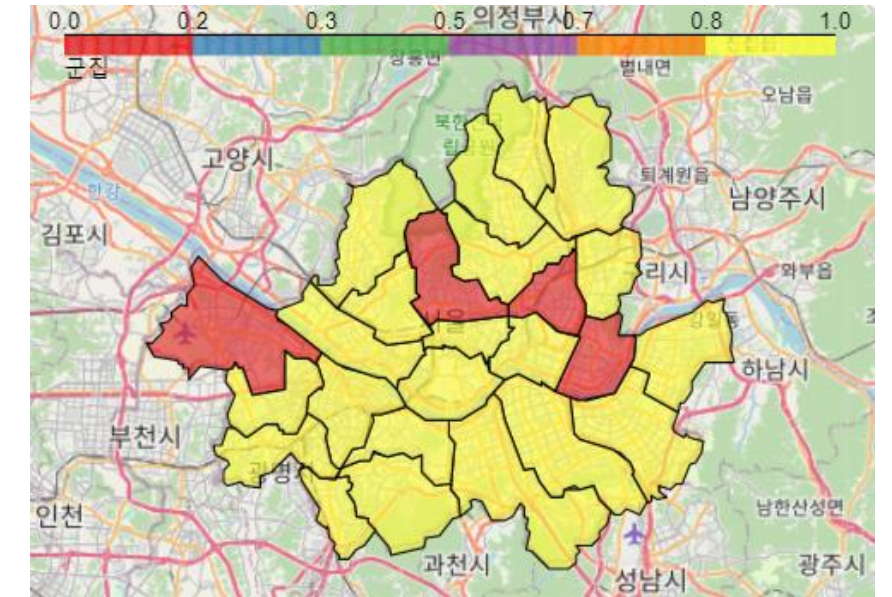
본 연구의 목적 : 자치구들을 군집별로 묶어 효율적으로  
관리하기위해 군집별 맞춤 제언  
-> 군집으로 묶이지 않는 자치구들을 억지로 군집화하지 않음





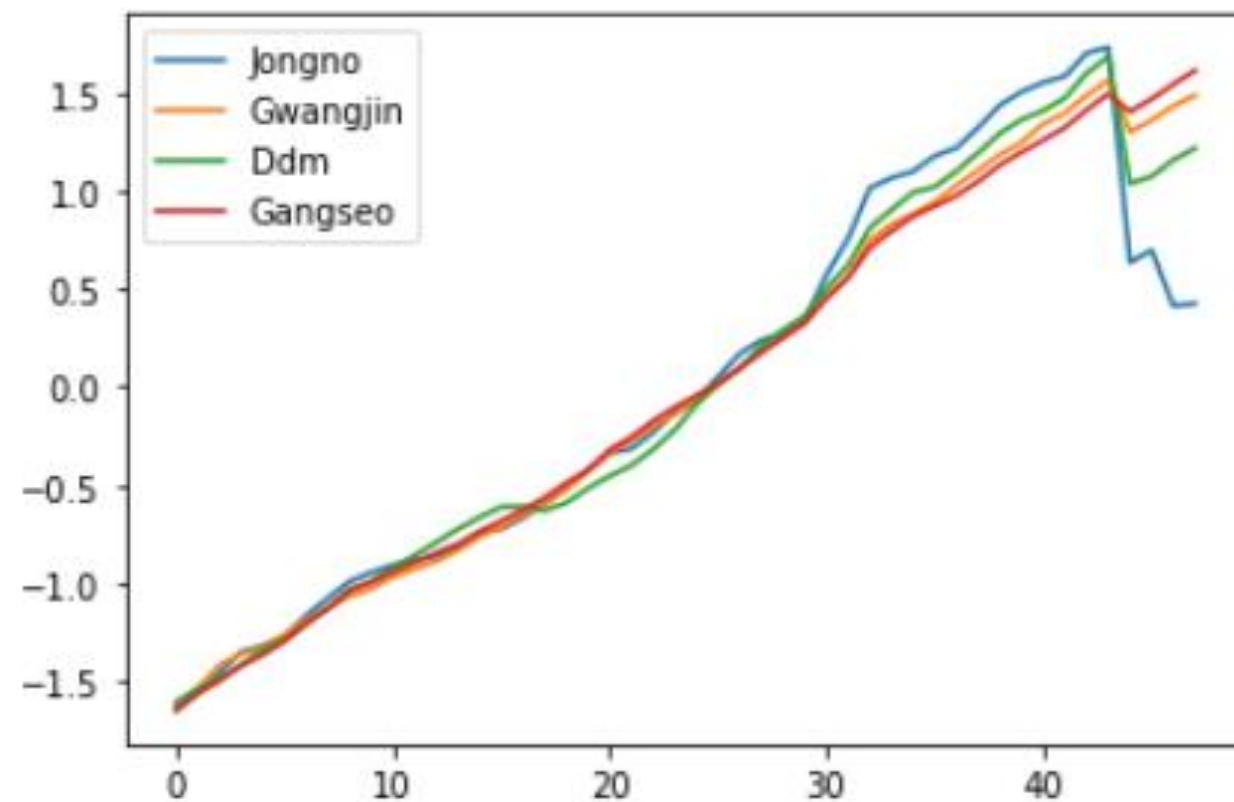
# 3-3. 연구 내용 – 군집0 데이터 파악

군집 0 : 종로구, 광진구, 동대문구, 강서구

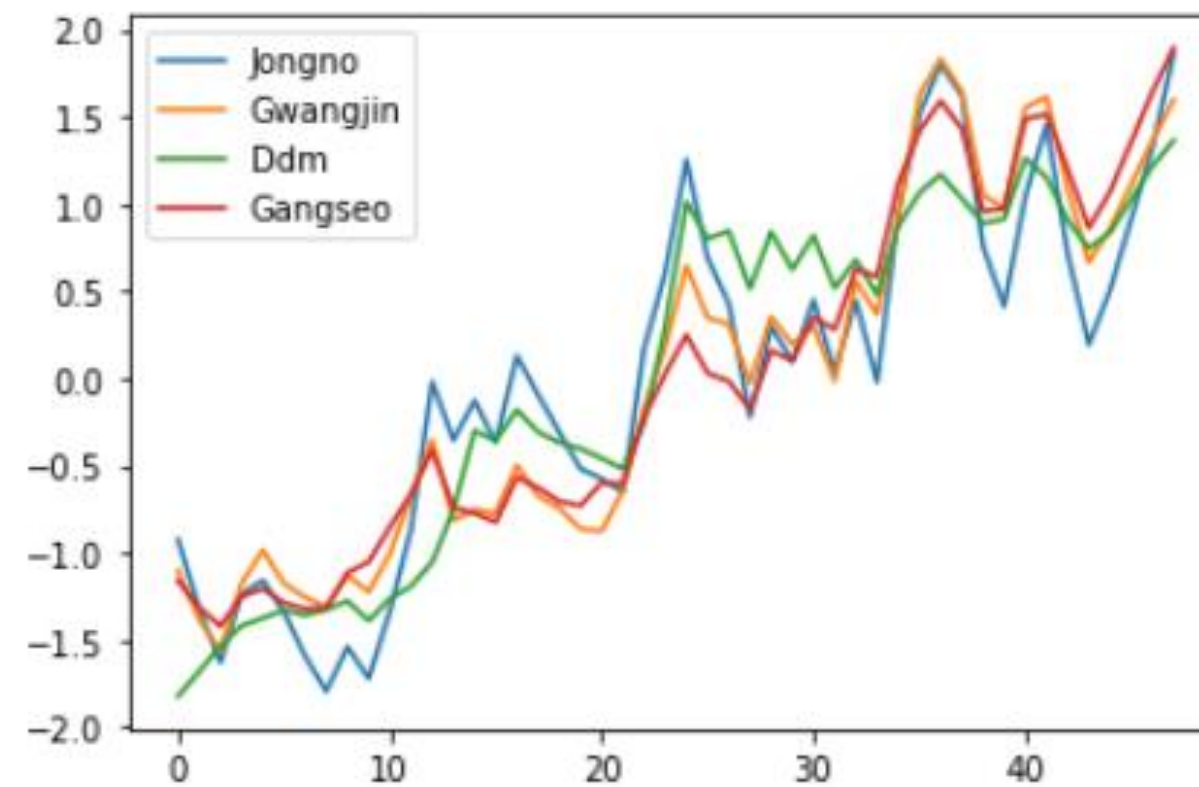


군집 0 지도 시각화

“ 65세이상 인구수(+) ”



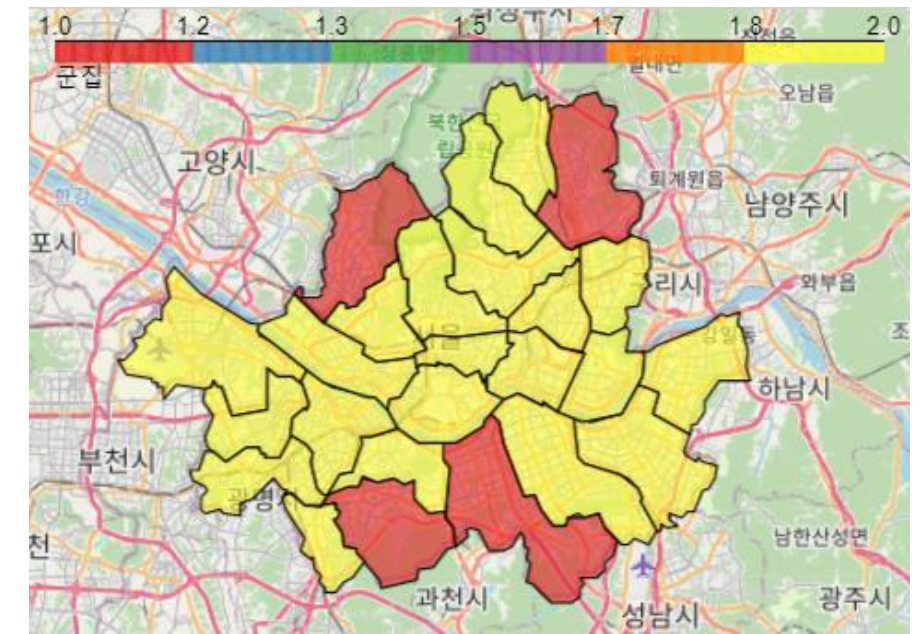
음식물쓰레기 배출량





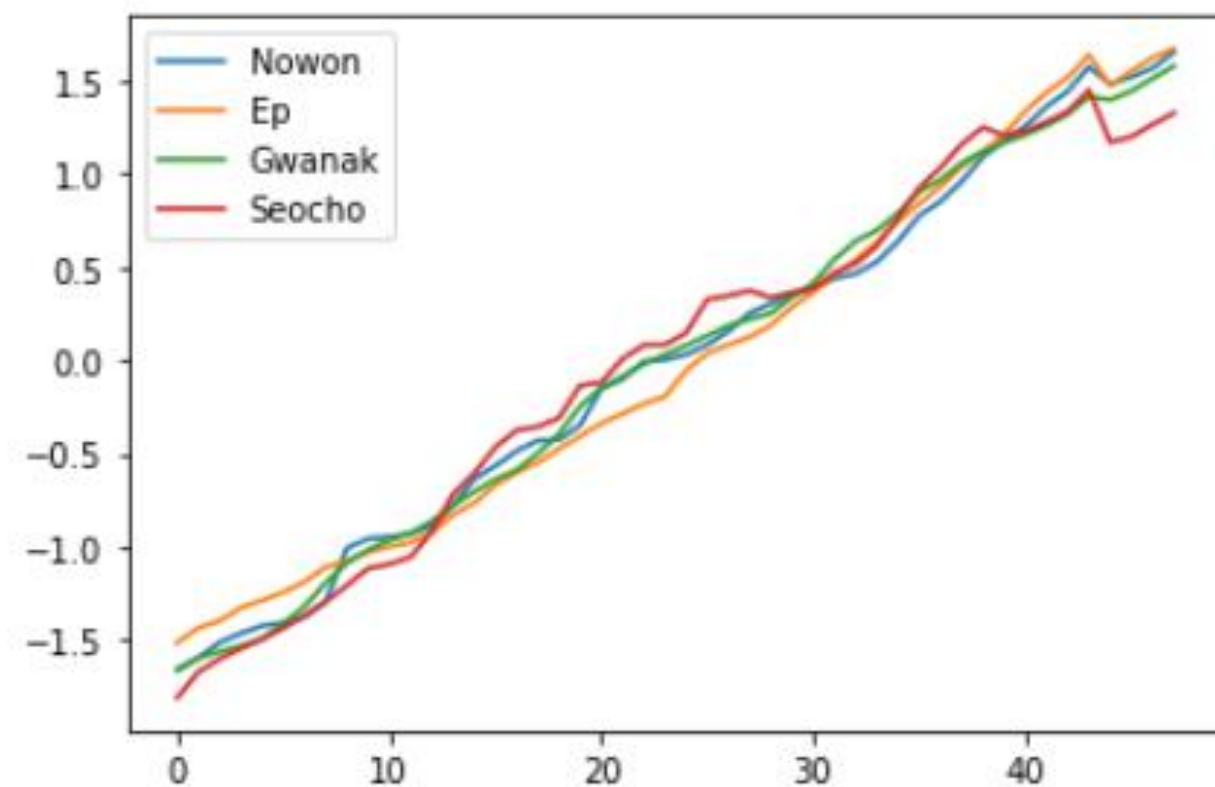
# 3-3. 연구 내용 – 군집1 데이터 파악

군집 1 : 노원구, 은평구, 관악구, 서초구

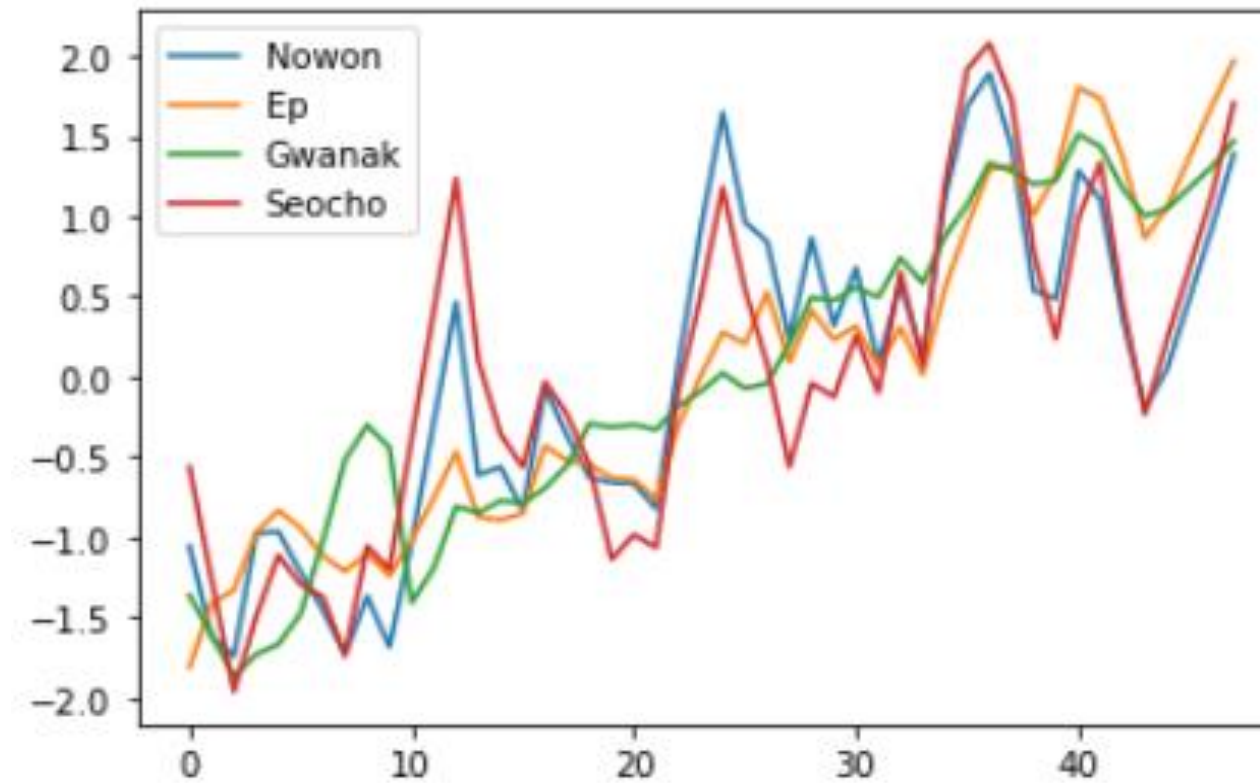


군집 1 지도 시각화

“ 1인가구 비율(+) ”



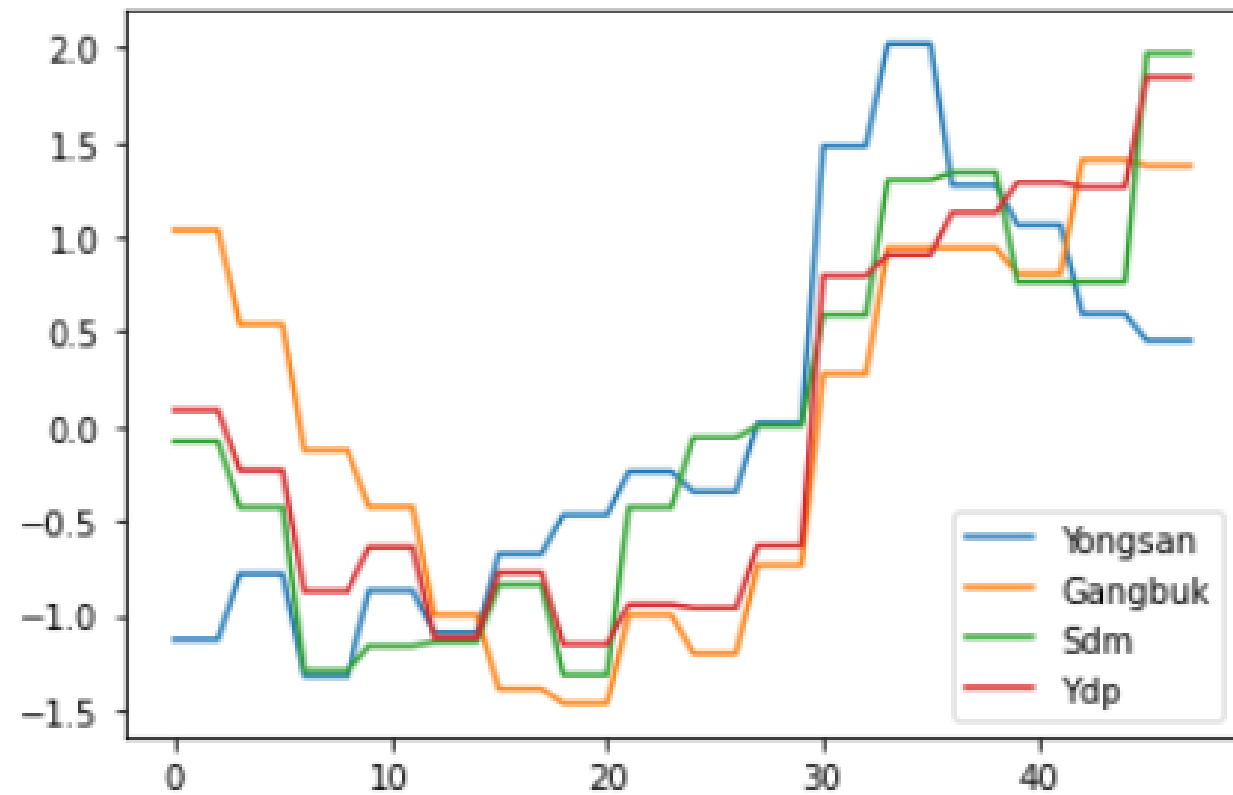
음식물쓰레기 배출량



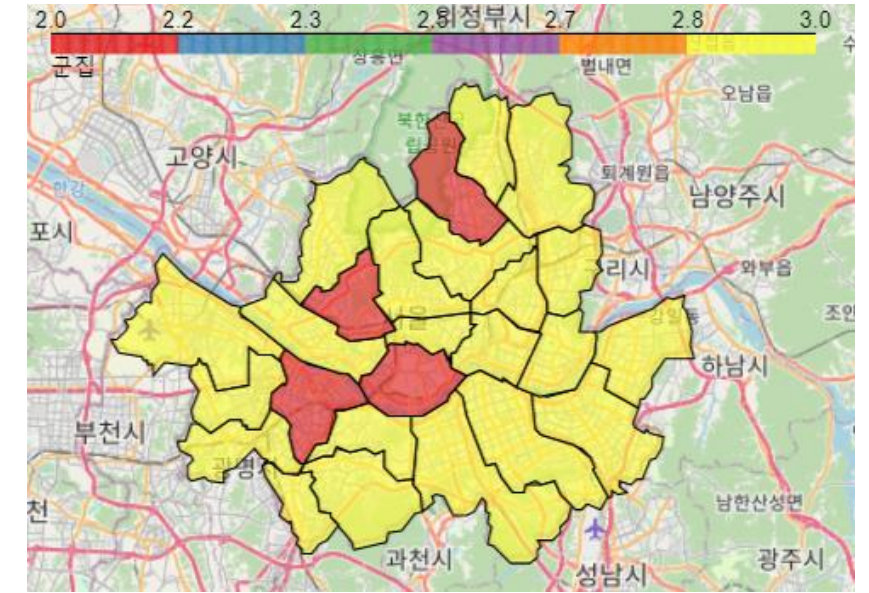
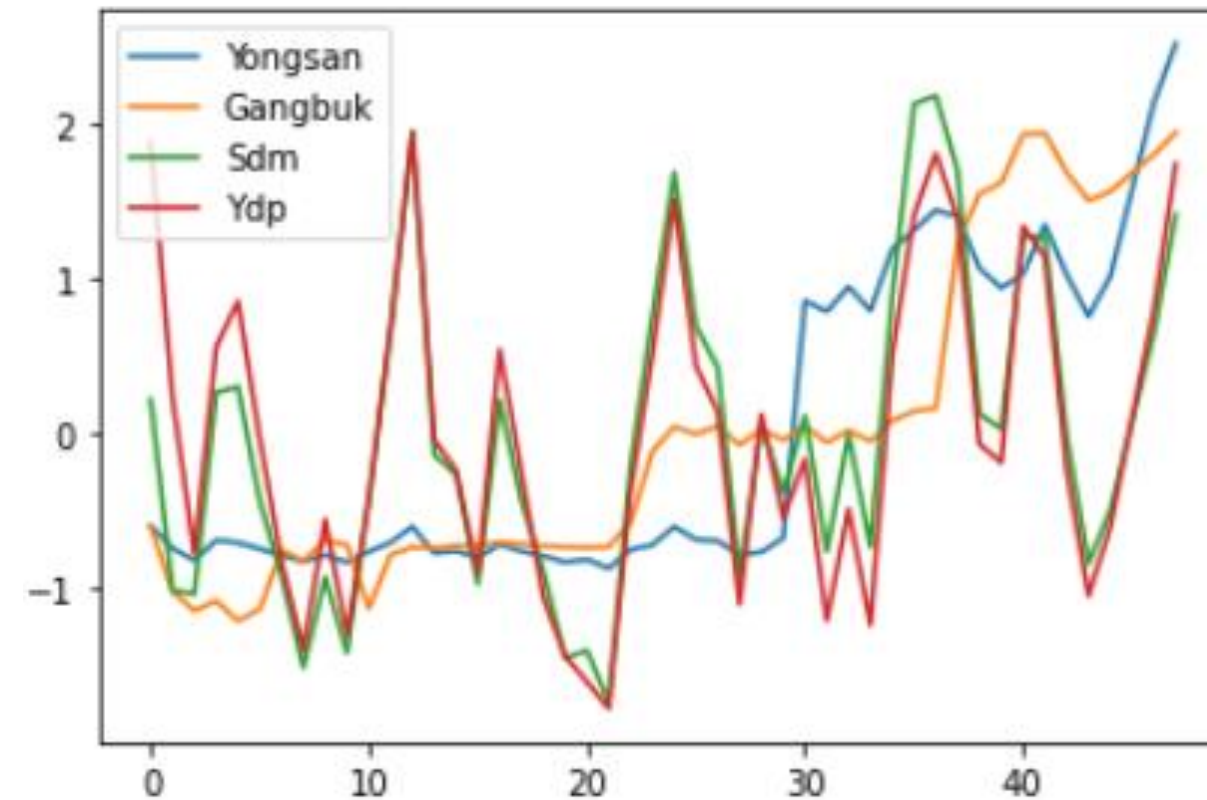
# 3-3. 연구 내용 – 군집2 데이터 파악

군집 2 : 용산구, 강북구, 서대문구, 영등포구

“ 점포수(+) ”



음식물쓰레기 배출량



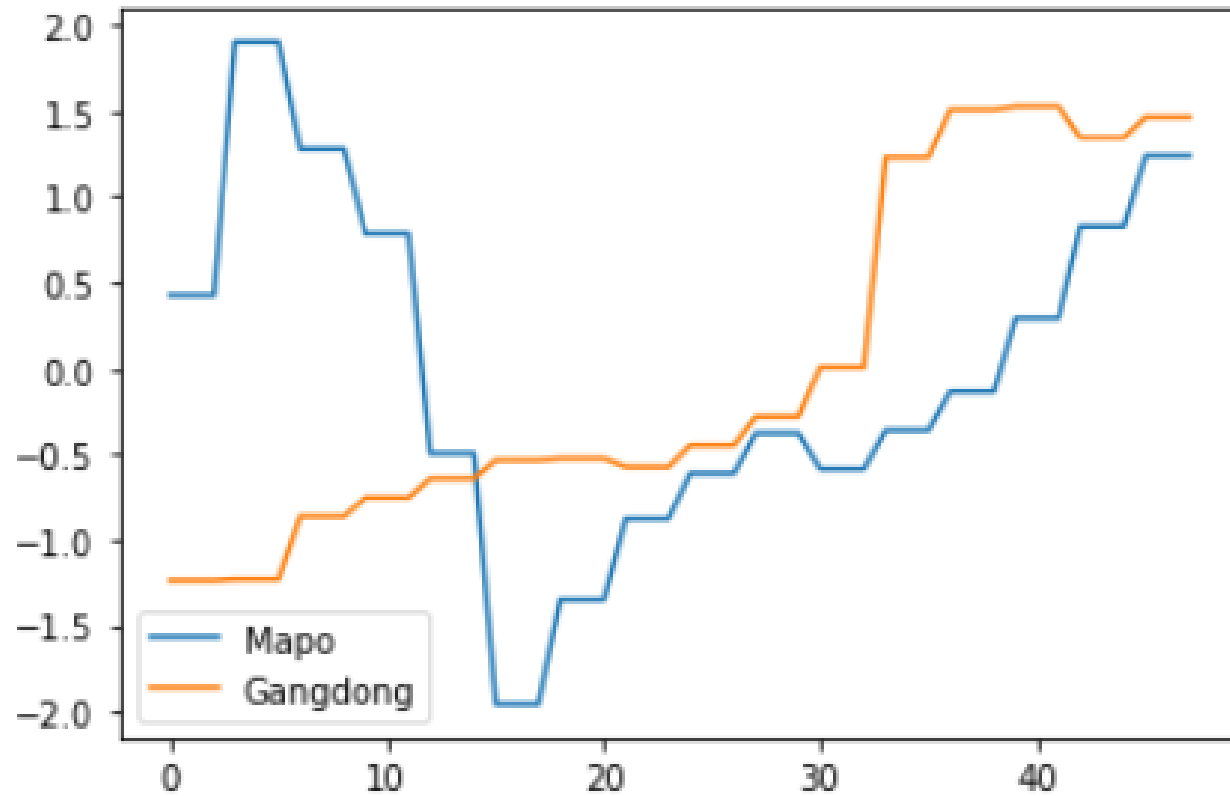
군집 2 지도 시각화



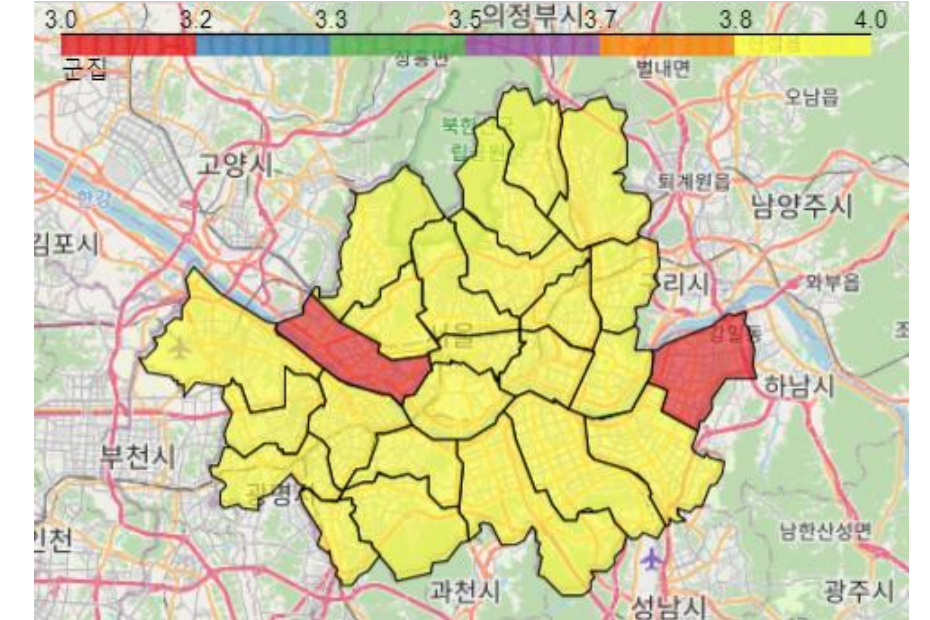
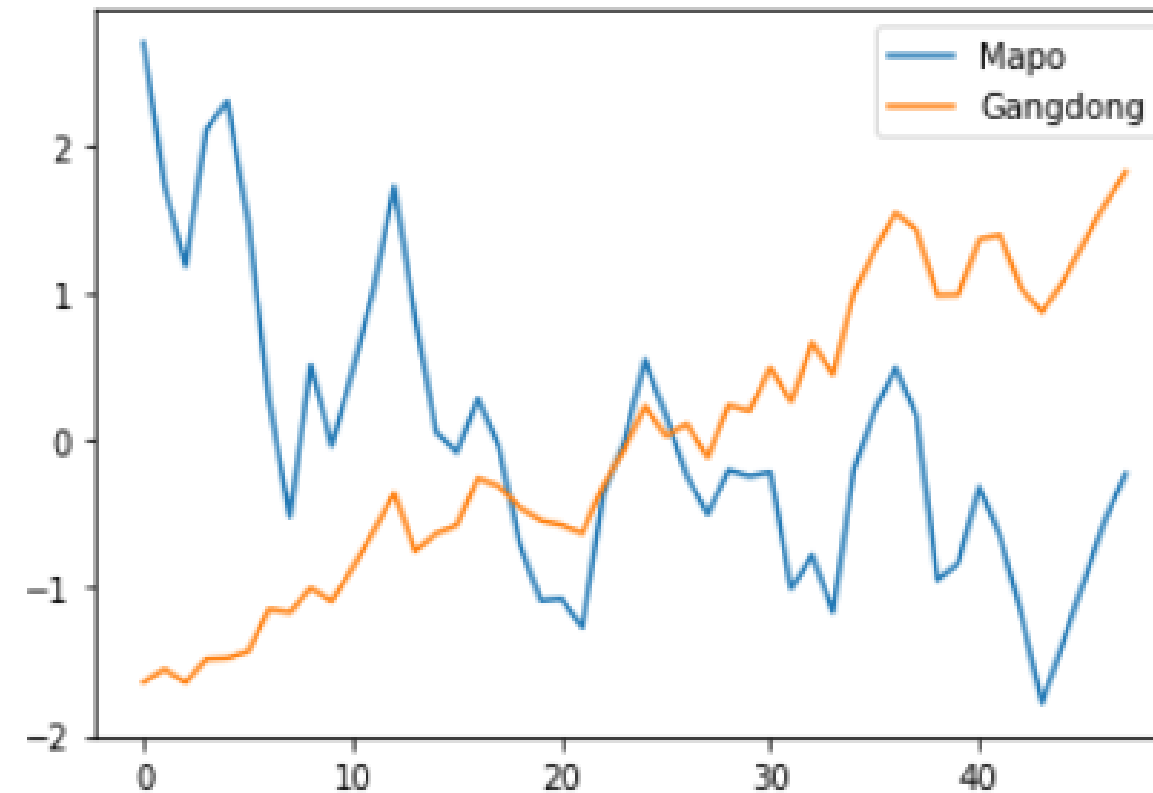
# 3-3. 연구 내용 – 군집3 데이터 파악

군집 3 : 마포구, 강동구

“ 임대시세(+) ”



음식물쓰레기 배출량



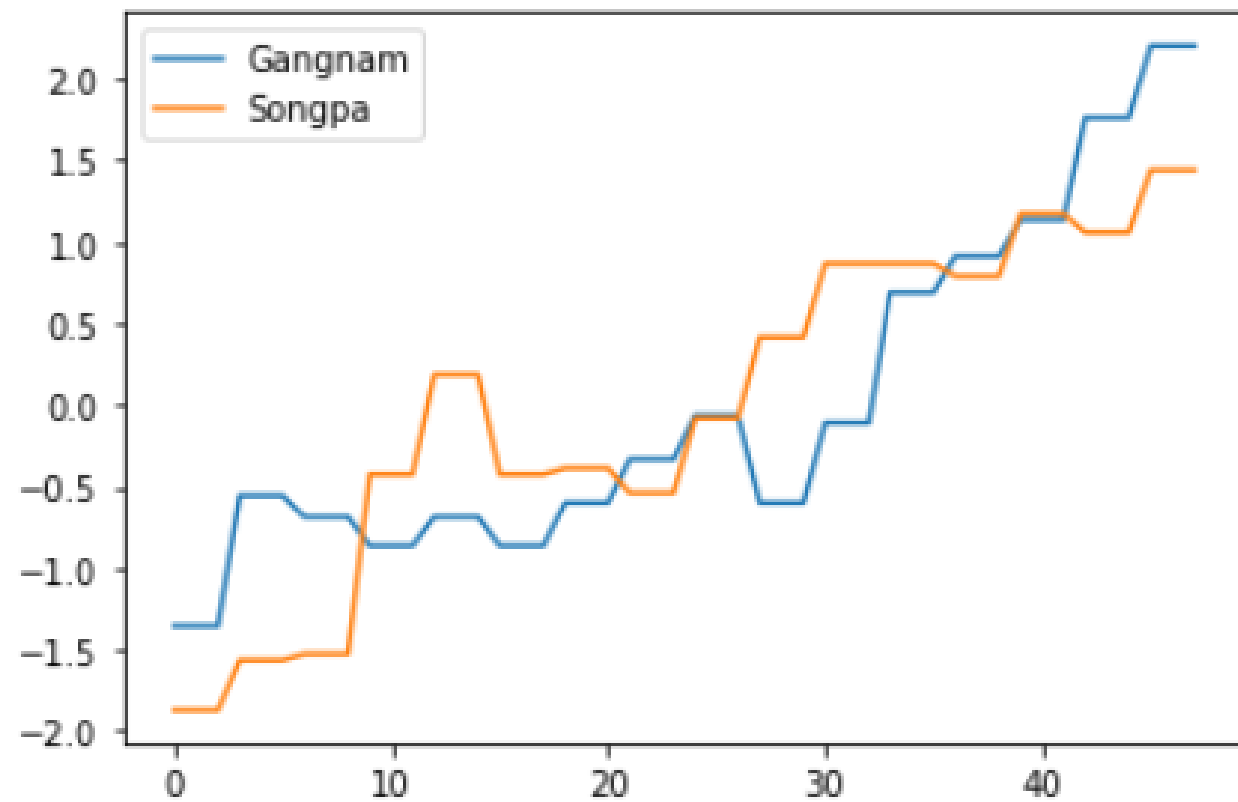
군집 3 지도 시각화

# 3-3. 연구 내용 – 군집4 데이터

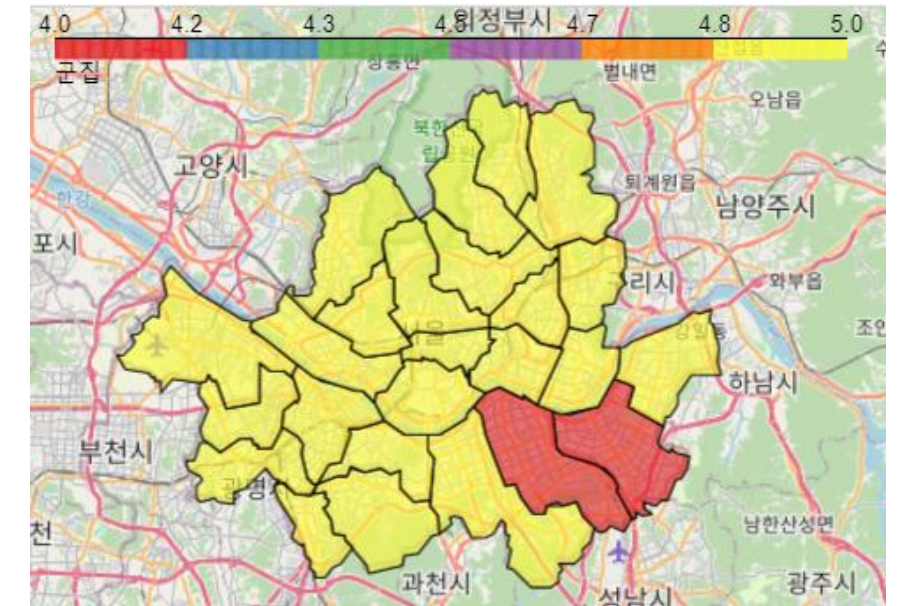
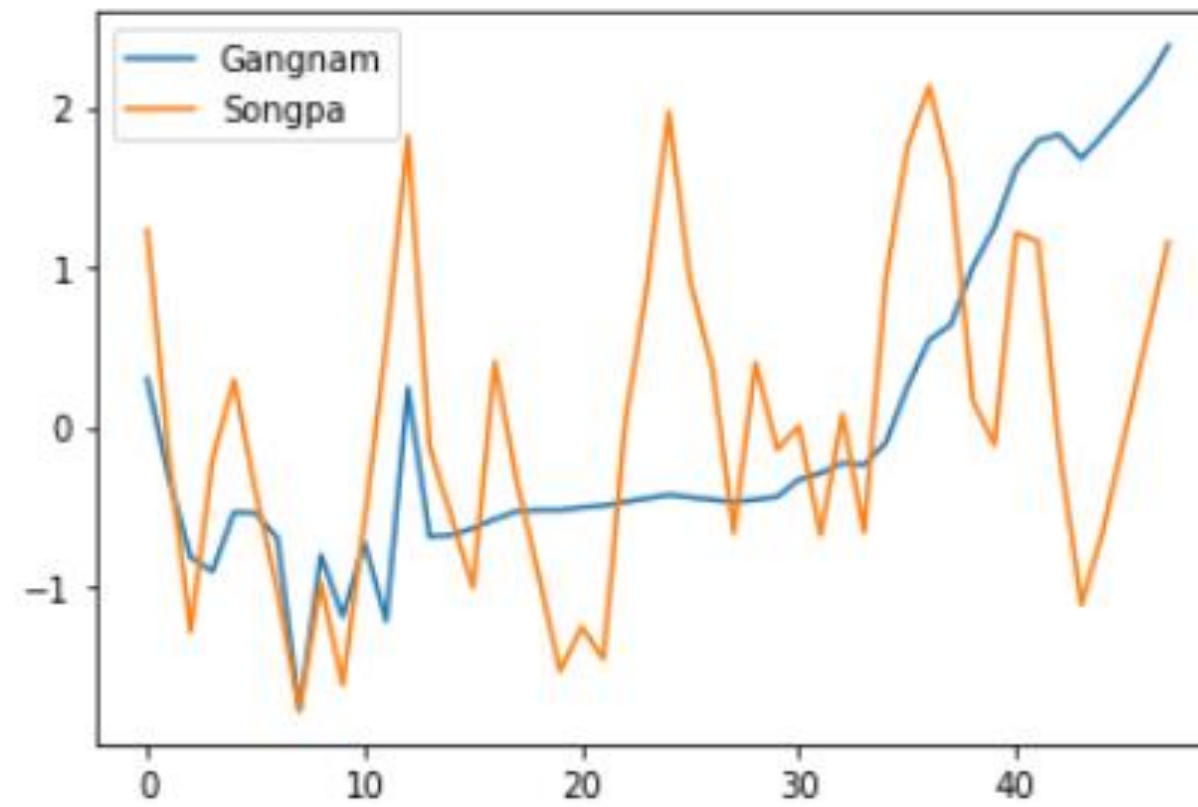
## 파악

군집 4 : 강남구, 송파구

“ 신생 음식점 3년 생존율(+) ”



음식물쓰레기 배출량

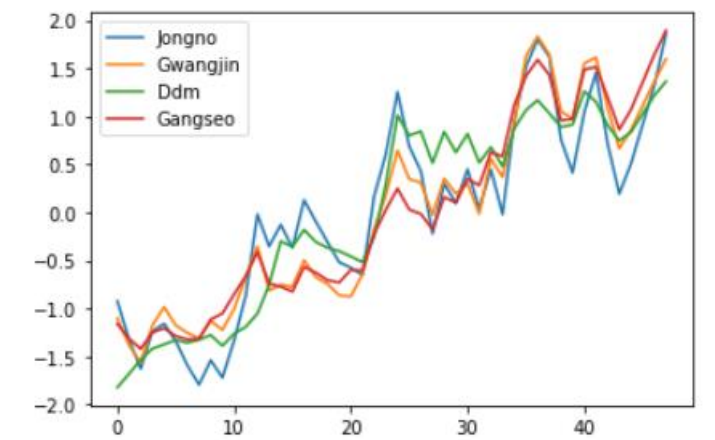


군집 4 지도 시각화

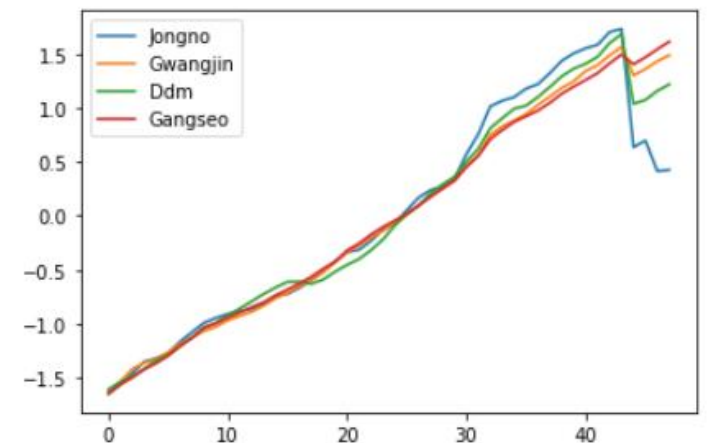
## 4. 연구 결과 – 군집0 맞춤 정책 제언

1. 지상파 공익광고, 캠페인 등을 활성화

2. 감자, 배추와 같은 식재료 손질된 반 가공 제품 판매 촉진



65세이상 인구수



배출량



국민이 하나도  
같이가치  
광진

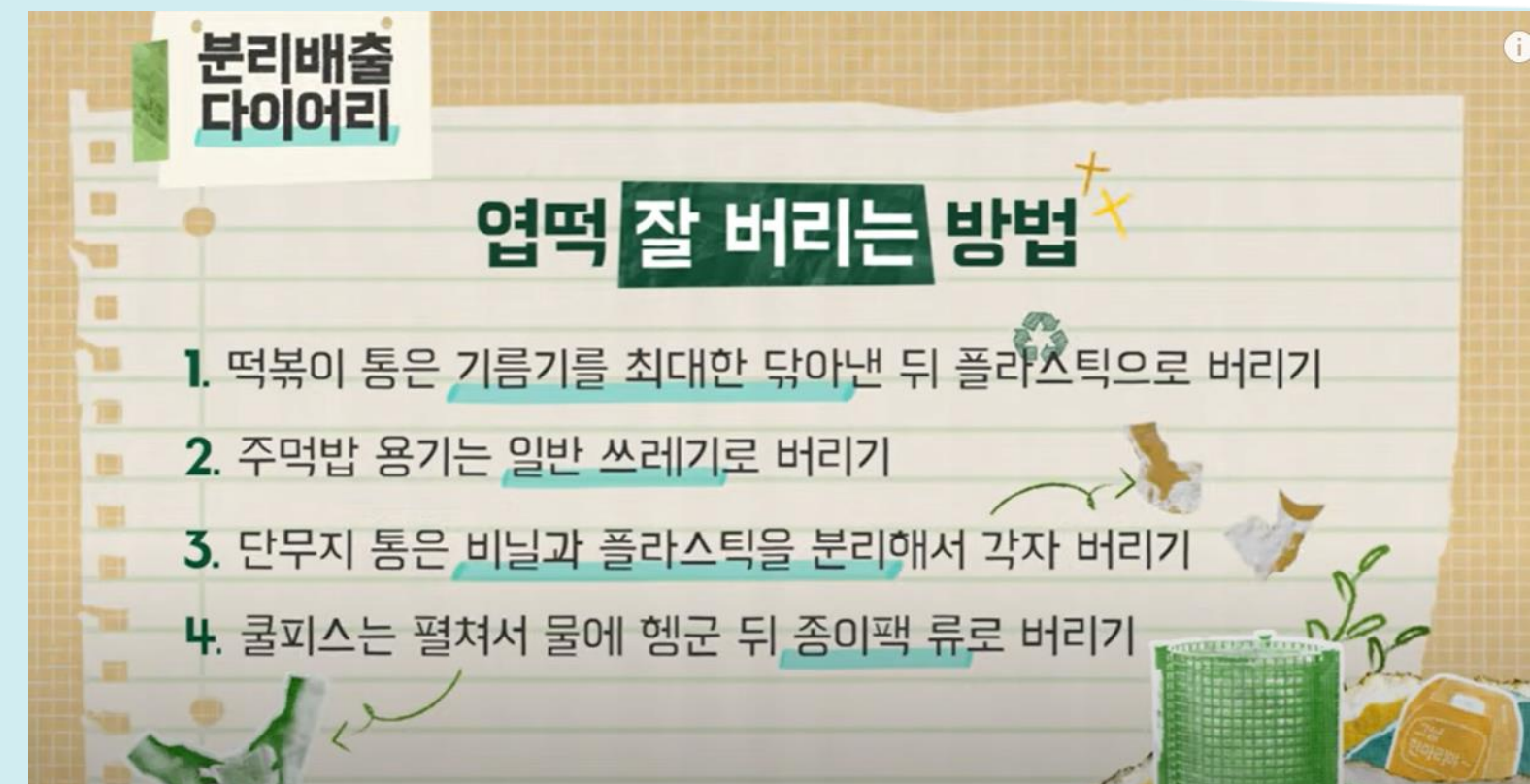
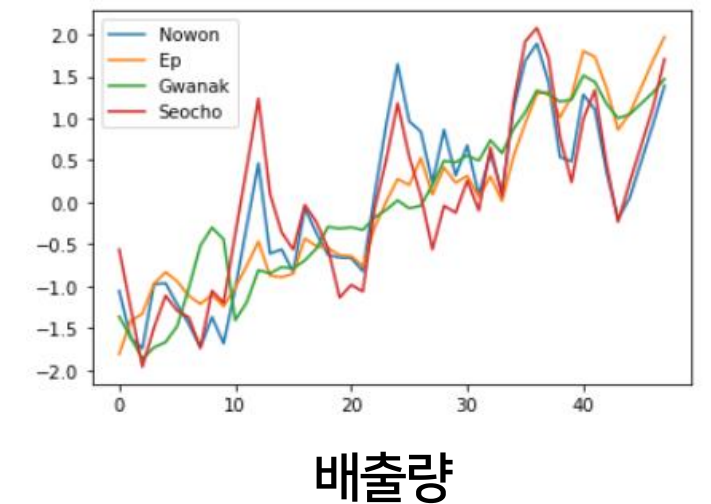
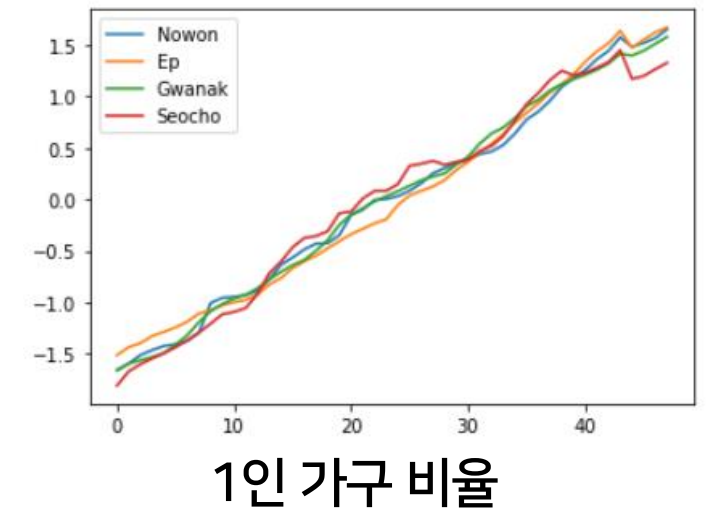




## 4. 연구 결과 – 군집1 맞춤 정책 제언

1. 1인 가구 밀집 지역 음식물 감량기 우선 설치

2. 음식물 쓰레기 분리배출 방법 배달 서비스와 협력해 알리기

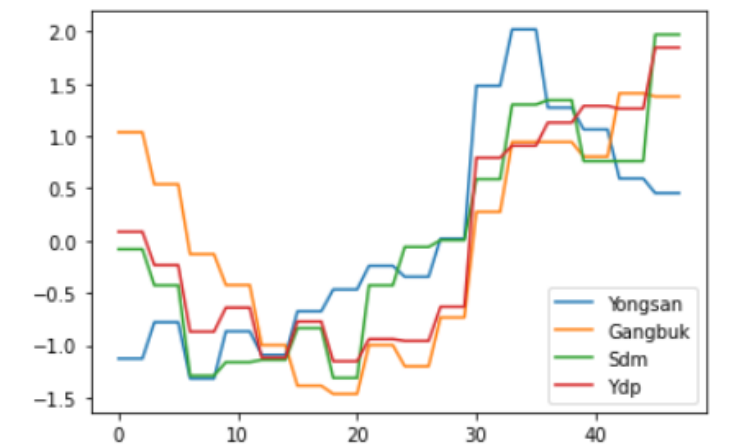




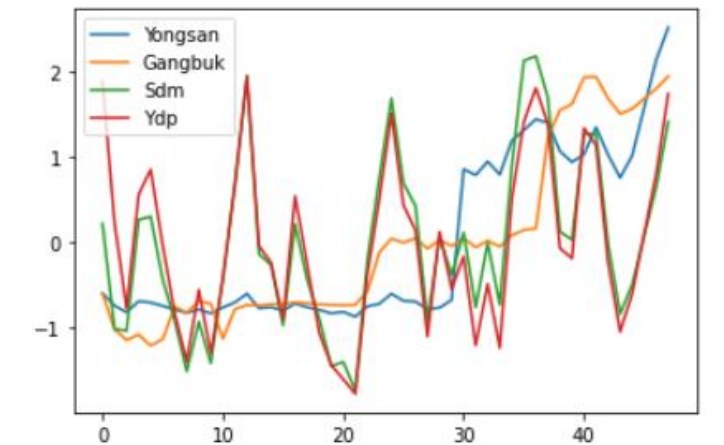
# 4. 연구 결과 – 군집2 맞춤 정책 제언

1. 점포 반찬 캠페인 – 참여 업소에 여러 지원

2. 푸드마켓 확대



점포 수



배출량



강북구

## 푸드뱅크란?

기업이나 가정으로부터  
식품, 생활용품 등을 기부받아  
사회복지시설 및 저소득가정에  
기부물품을 전달하는  
“식품나눔은행”입니다



## 푸드마켓이란?

편의점 형태의 이용자가 직접  
방문하여 필요한 식품 등을  
선택할 수 있는 “나눔공간”입니다.

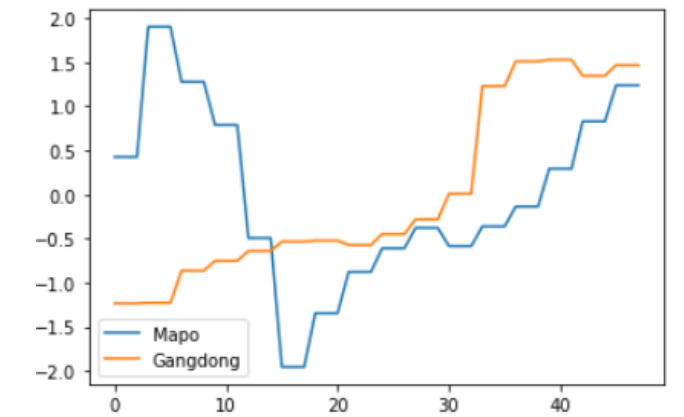
## 이동푸드마켓이란?

푸드마켓의 이용혜택을 받지  
못하는 복지사각지대를 해소하기  
위해 복지소외계층을 방문하는  
“찾아가는 푸드마켓”  
입니다.

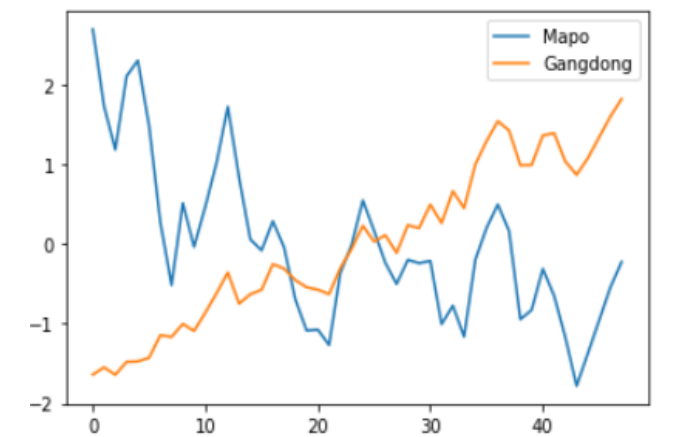
## 4. 연구 결과 – 군집3 맞춤 정책 제언

1. 가정용 음식물 감량기/분쇄기 구비 촉진

2. 누진세 적용 세율 세분화



임대시세



배출량



마포구



더불어 행복한 강동

강동구

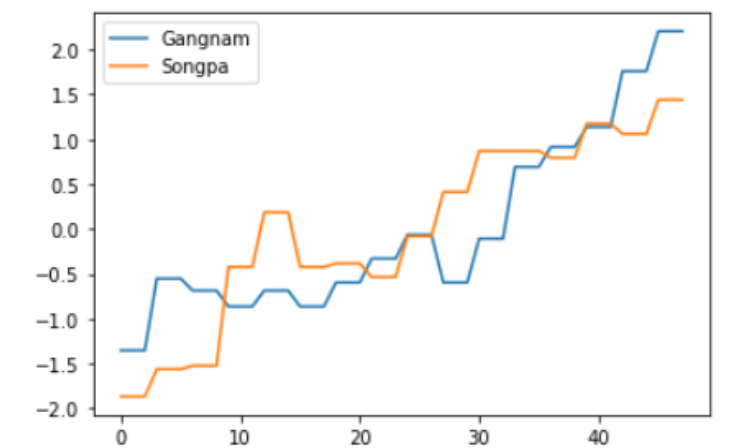
## 4. 연구 결과 – 군집4 맞춤 정책 제언

1. 배달 앱 업무 협약 활성화로 잔반 최소화 강구

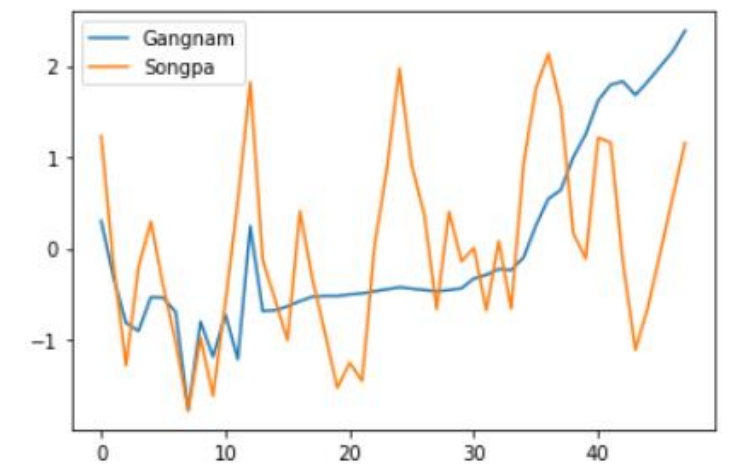
2. 환경 우수업소 선정



강남구



신생 음식점 3년 생존율



배출량

## 4. 연구 목적 대비 연구 결과

### 1. 우리의 차별점인 군집화를 통해 얻을 수 있는 효과

- 1) 동일한 양상을 가지지만 서로 다른 시기에 놓여져 있는 자치구들을 보다 쉽게 예측 가능
- 2) 자치구별 특성에 따라 서로 다른 정책을 강구하지 않아도 되며
- 3) 군집화된 자치구별 특성파악을 통해 효율적인 맞춤 전략 제시 가능하다.
- 4) 군집별 관리를 통해 더 효율적이고 유동적인 계획, 협력 가능

### 2. 연구의 기대효과,활용효과

음식물쓰레기 감량기 우선 설치지역 선정 등 앞으로 활용될 대책에 기초 자료로 활용 가능



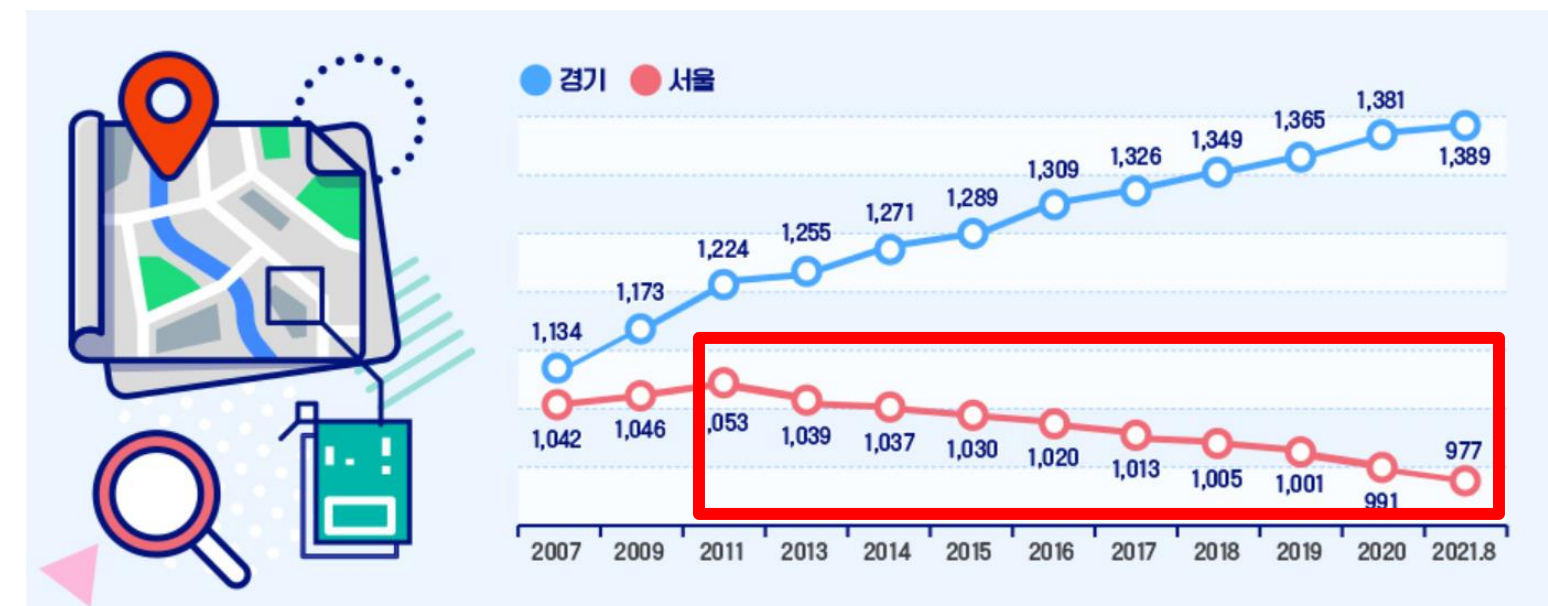
## 4. 연구 목적 대비 연구 결과

### 3. 연구 목적 대비 왜 분석이 제대로 이뤄지지 않았는지에 대한 고찰

- 음식물쓰레기 배출량이 다양한 요인으로 결정될 것이라고 가정했으나,  
분석결과로 추출된 유의미한 변수는 다양하지 않았음
- 변수들 간의 다중공선성이 높아서 군집별 특성분석을 하려했던 목적대비 군집별 특성 해석 및 결론 도출이 어려웠음
- 사례연구에서의 결과와 다른 결과가 도출되는 변수가 존재했음

Ex) 서울시 자치구의 인구수가 대체적으로 감소, 배출량과 음의 상관관계

서울시 인구수 변화 그래프



## 5. 본 연구의 한계점



### 변수

1. 변수들 간의 다중공선성이 높아 계획만큼 분석에 변수 활용하지 못함
2. 군집별 핵심변수 부족으로 AutoML을 통한 예측 진행이 어려웠음



### 데이터 부족

1. 음식물쓰레기의 정확한 배출지를 알 수 없음
2. 자치구별 월별로 다양한 데이터 수집의 어려움



### 제언

군집화가 되지 않은 자치구에는 제언을 하지 못함

## 6. 추후 연구 예정



변수

데이터의 부족과 다중공선성을 해결한 후 군집별 음식물쓰레기 배출량 예측



데이터

공공데이터 이외의 음식점 매출 데이터와 같은 생활밀착형 데이터를 활용하여 좀 더 나은 연구



제언

앞서 제시한 군집별 맞춤 제언의 효과를 분석



# 감사합니다

질문이 있다면 말씀해주세요.