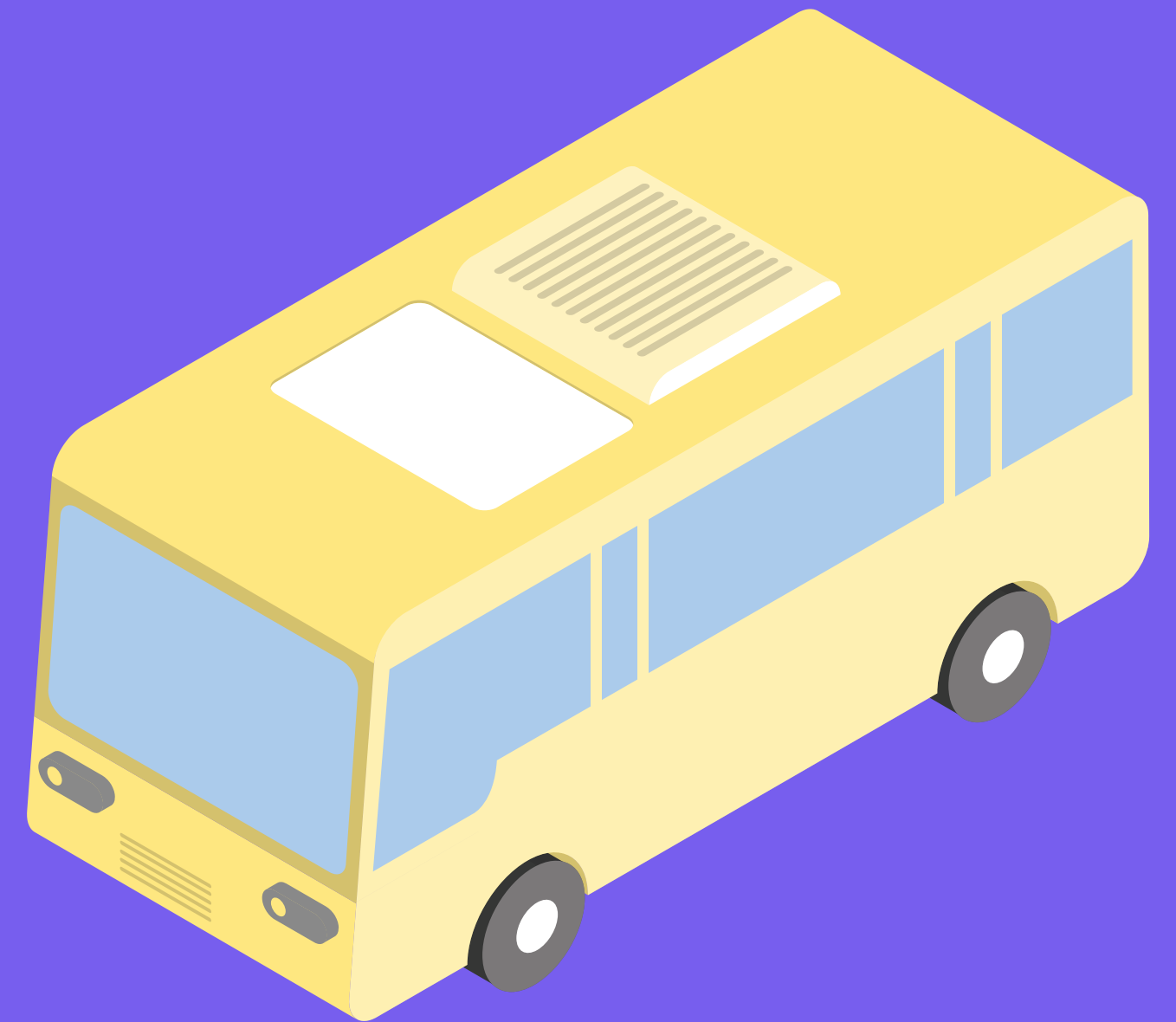


Hello, 반딧불버스

대구의 밤을 밝혀줄 심야버스

김동언, 김태형

안녕하세요,
반딧불버스입니다!



01

하나,



02

두울,



03

세엣,

WHY

기획 배경
반딧불버스의 의미

HOW

분석 방법
임시배차표

RESULT

반딧불버스의 기대효과

01

하나,

WHY

기획 배경
반딧불버스의 의미

02

두울,

HOW

분석 방법
임시배차표

03

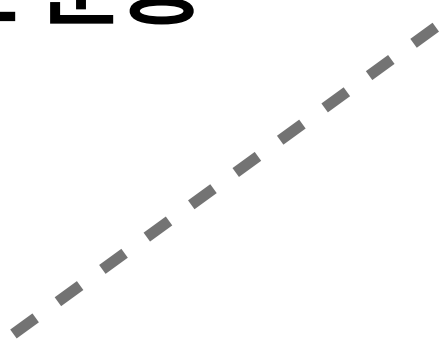
세엣,

RESULT

반딧불버스의 기대효과

why?

버스, 택시,
버스 배차, 버스 운행



(활용기법 - gensim - Word2Vec - Skip-gram)

	token	similar (%)
0	배차	37.36%
1	증차	34.34%
2	막차	34.09%
3	택시	33.67%
4	간격	31.91%
5	증설	31.7%
6	급행	31.56%
7	대중교통	31.33%
8	도착	31.3%
9	연장	31.18%
10	도서관	30.53%
11	증편	30.52%

민원데이터 분석 결과

```
df = df.iloc[:,0].apply(lambda x: re.sub('[^가-힣0-9]+', ' ', x))

# tokenize
okt = Okt()

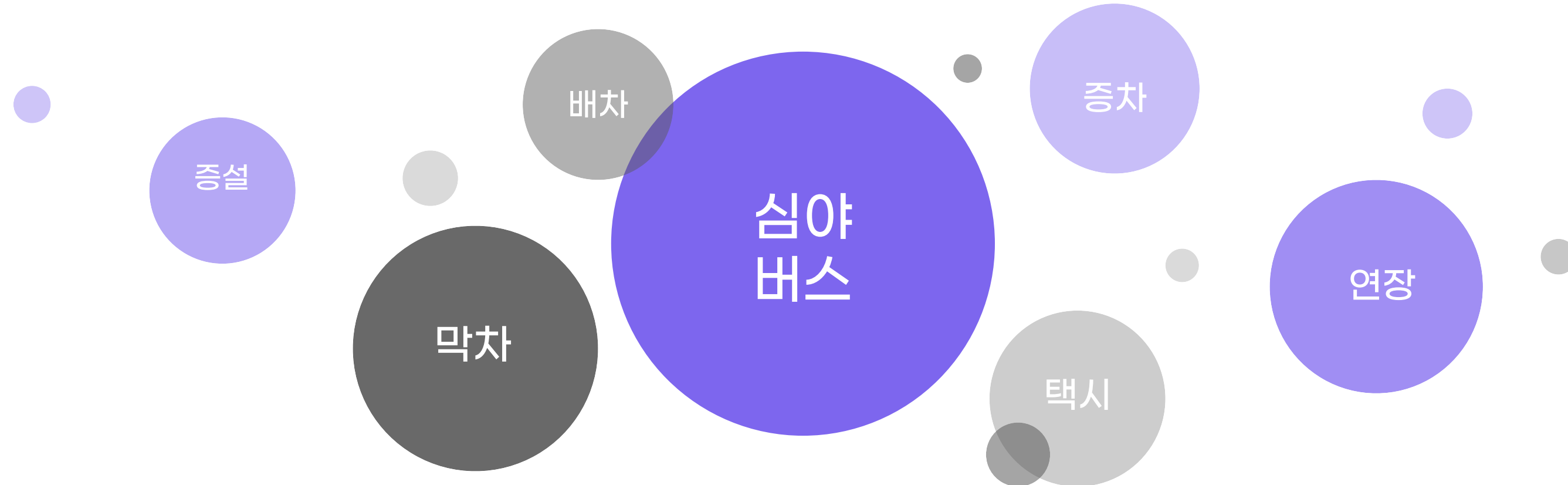
sol = []
for idx in result:
    sol.append(okt.nouns(idx))

gen_model = Word2Vec(df['tokens'], window = 3, size = 200, min_count = 5, sg = 1, iter = 200)

df = gen_model.wv.most_similar(df['tokens'], topn=30)

result = pd.DataFrame(df, columns = ['tokens', 'similar(%)']).reset_index(drop=True)
result['similar(%)'] = result['similar(%)'].apply(lambda x: str(round(x*100, 2)) + '%')
```

심야버스를 기획해보자!

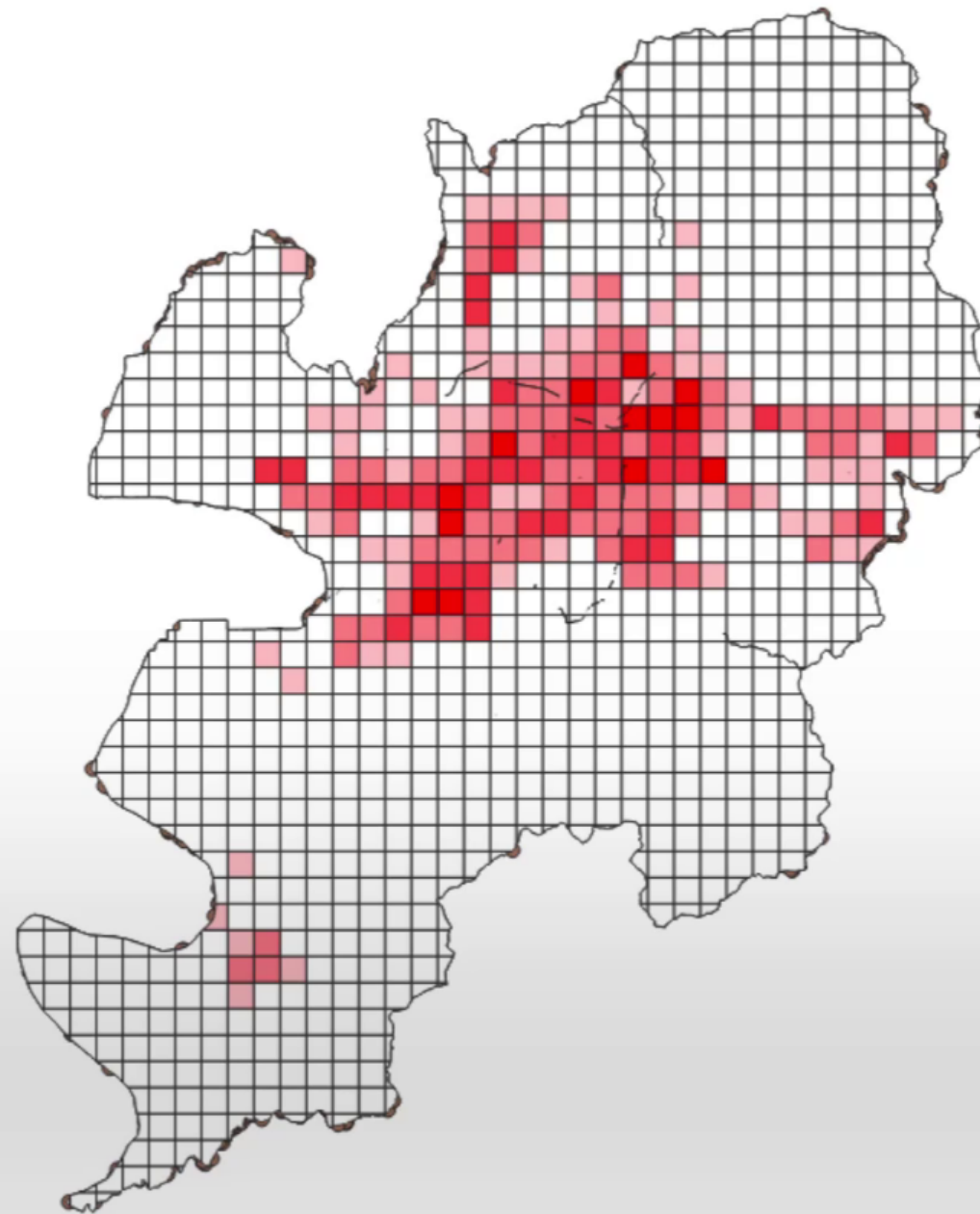


민원데이터 분석 결과 다음과 같은 키워드에서 아이디어를 가져와 심야버스를 기획!

시간별 유동인구 변화량



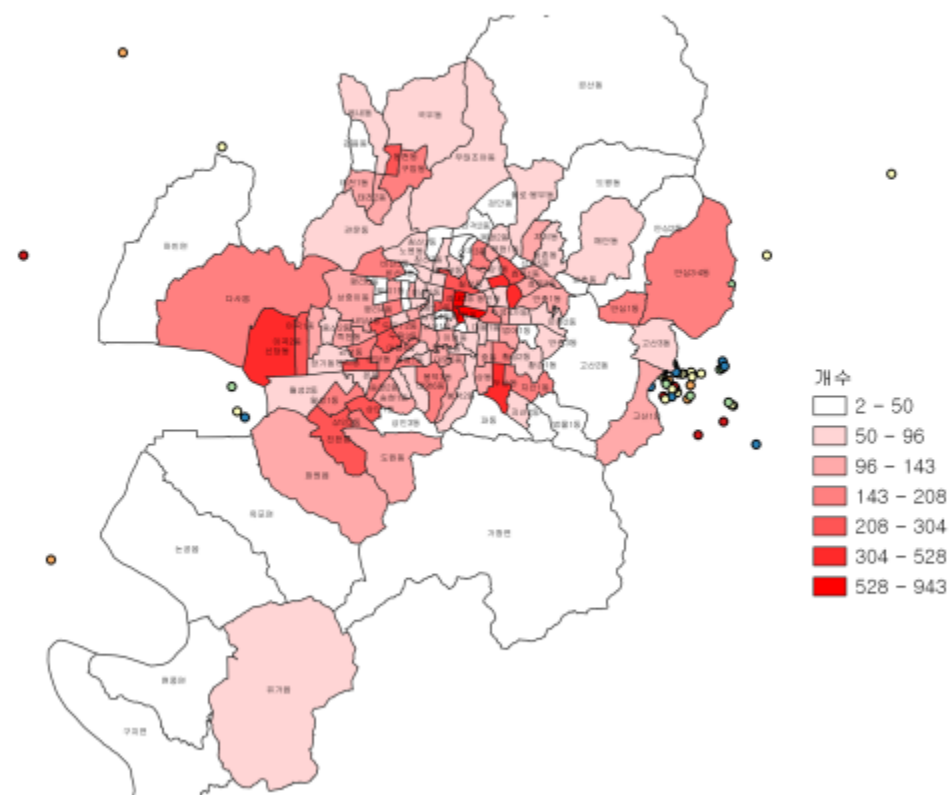
21시



지역별 택시 이용수

택시데이터를 활용해 21시~02시까지 어느지역에서 많이 택시를 이용했는지 시각화 해보고 이를 통해 심야버스의 필요 지역을 파악.
또 각 시간별로 택시 이용자 수의 데이터도 수치로 확인.

21시~02시 지역별 택시 이용수



21시

3110

22시

3215

23시

3627

00시

2998

01시

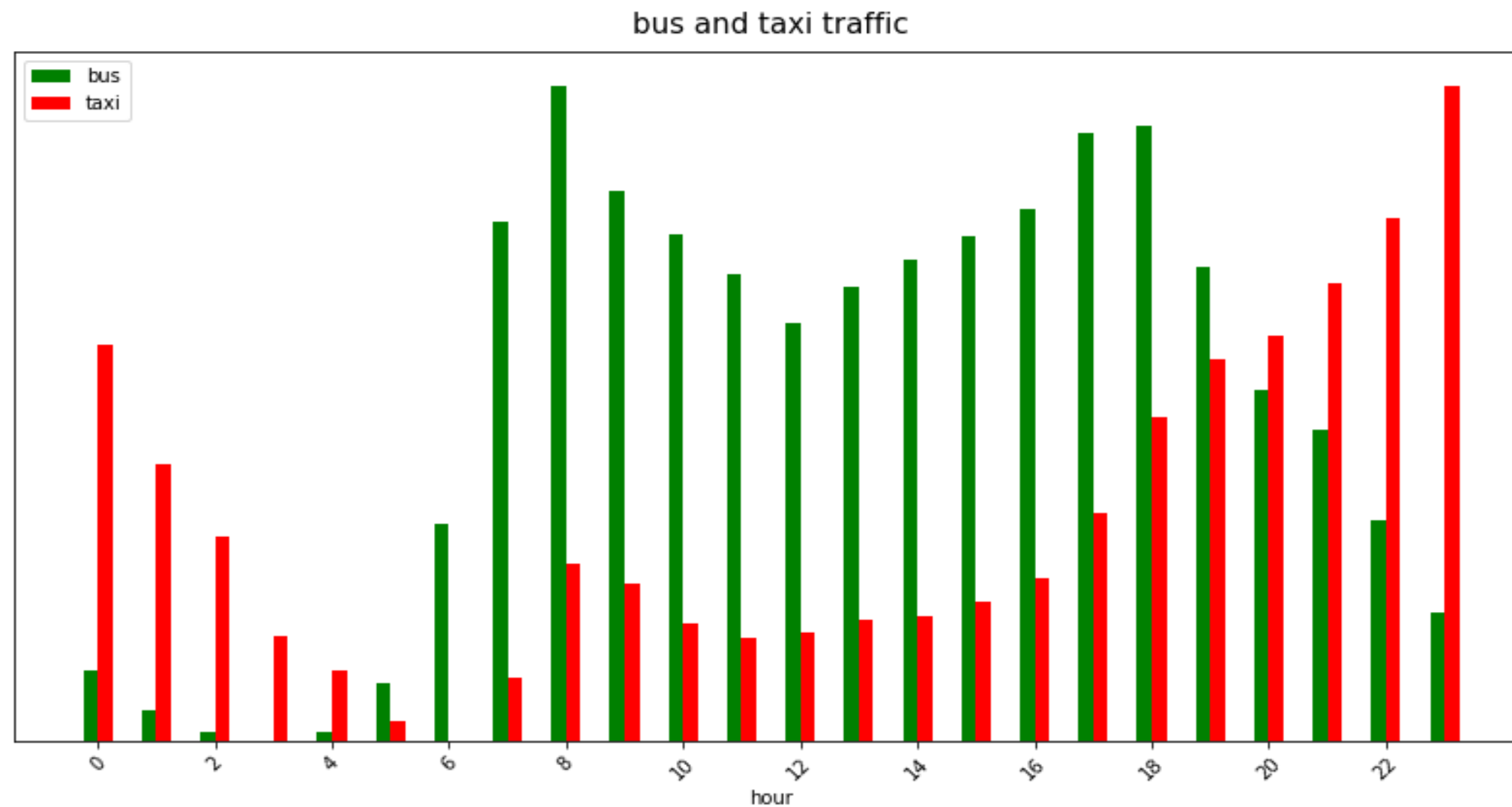
2388

02시

1842

버스와 택시 운행량

버스의 운행이 줄어들면서 택시의 수요가 높아진 것을 확인
05시부터 버스가 운행하자 택시를 이용하는 승객이 줄어든 것을 확인



BUS ↓
TAXI ↑



민원데이터 분석결과

'가로등'을 키워드

```
gen_model.wv.most_similar('가로등', topn = 20 )
```

```
[('신호등', 0.5138801336288452),
 ('멀등', 0.4949066638946533),
 ('불빛', 0.49401572346687317),
 ('고장', 0.4724493622779846),
 ('보안등', 0.44558054208755493),
 ('멀등원', 0.44086316227912903),
 ('밤길', 0.4349859356880188),
 ('전주', 0.4191935062408447),
 ('전지', 0.41051000356674194),
 ('전구', 0.40936267375946045),
 ('밤', 0.4019755721092224),
 ('우범', 0.392807275056839),
 ('두운', 0.39139091968536377),
 ('조명', 0.3911404013633728),
 ('밀', 0.3880215883255005),
 ('타트', 0.3851237893104553),
 ('전봇대', 0.38386544585227966),
 ('암흑', 0.38213232159614563),
 ('돌다리', 0.38156354427337646),
 ('교체', 0.3806523084640503)]
```

**멀등,고장,밤길,
암흑,밤,보안등**

위와 같이 가로등에 대한 불만을 확인하고
버스의 불빛이 가로등의 역할을 보조하지 않을까 생각

왜 반딧불버스인가?



반딧불이는 어두운 밤 산 속에서 길 잃은 사람들에게 길을 밝혀 안전한 길을 만들어주기도 하고 말레이시아 코타키나발루에서는 반딧불투어를 만들어 멋진 장관을 연출한다고 합니다.

대구의 반딧불버스 또한 대구의 밤을 비춰 안전한 귀갓길을 제공하고 밝은 도로의 빛으로 좀 더 건강하고 화려한 대구가 되길 응원하는 마음으로 작명하였습니다.

코로나가 안전하게 종식된 후 반딧불버스로 대구시민들이 대구의 자부심을 갖을 수 있기를 기원합니다.

01

하나,

WHY

기획 배경
반딧불버스의 의미

02

두울,

HOW

분석 방법
임시배차표

03

세엣,

RESULT

반딧불버스의 기대효과

elbow # k-means++

평가방법

elbow method :

데이터들의 inertia를 가지고 군집의 수를 평가하는 metric

inertia가 낮을수록 군집간의 inertia가 낮은것으로 평가되어 군집을 더 잘 분류했다고 본다.

RMSE 와 silhouette score를 metric으로 하여 평가를 해본 결과

elbow method를 사용하는것이 좋다고 판단해 평가지표로 사용

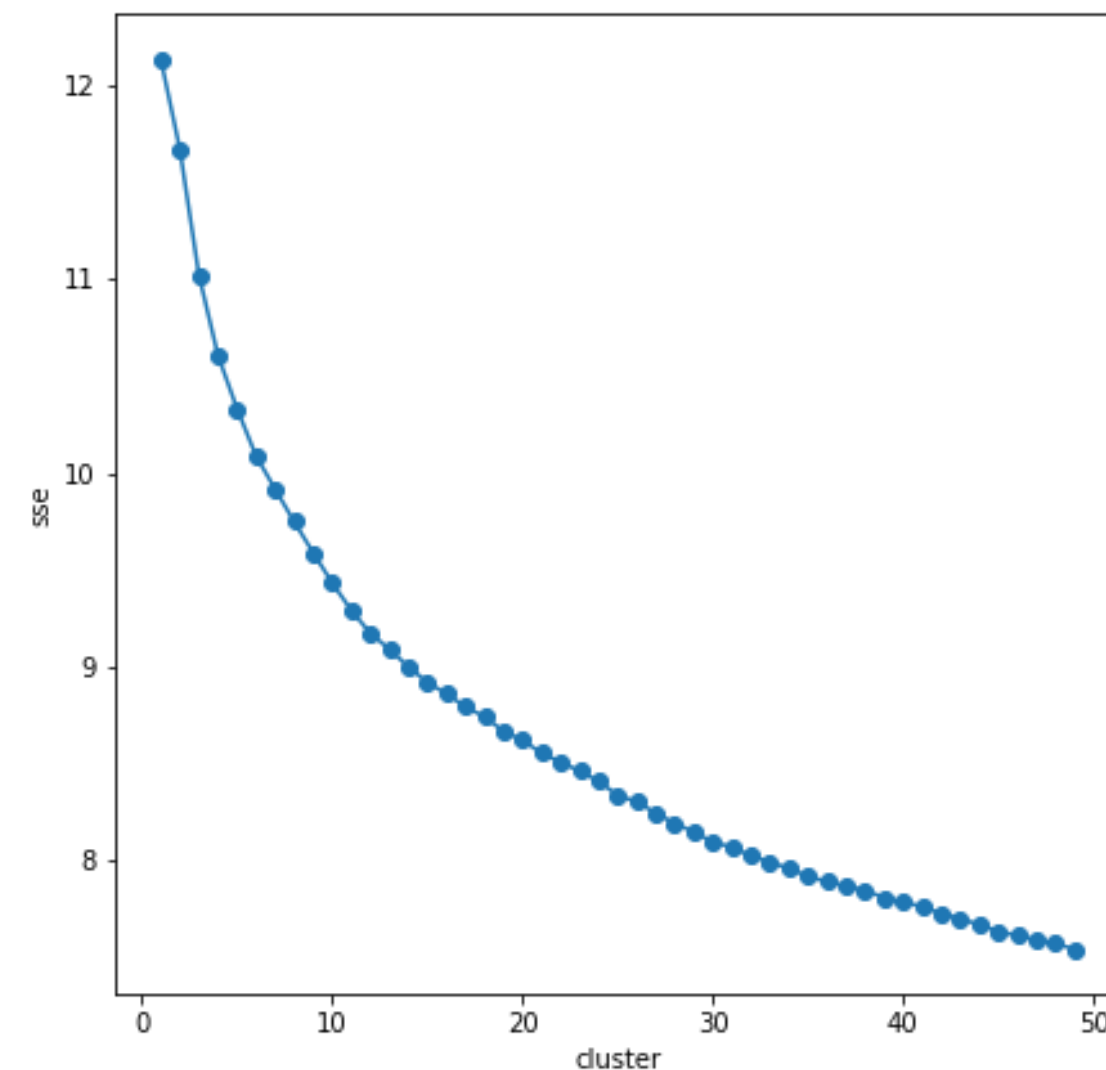
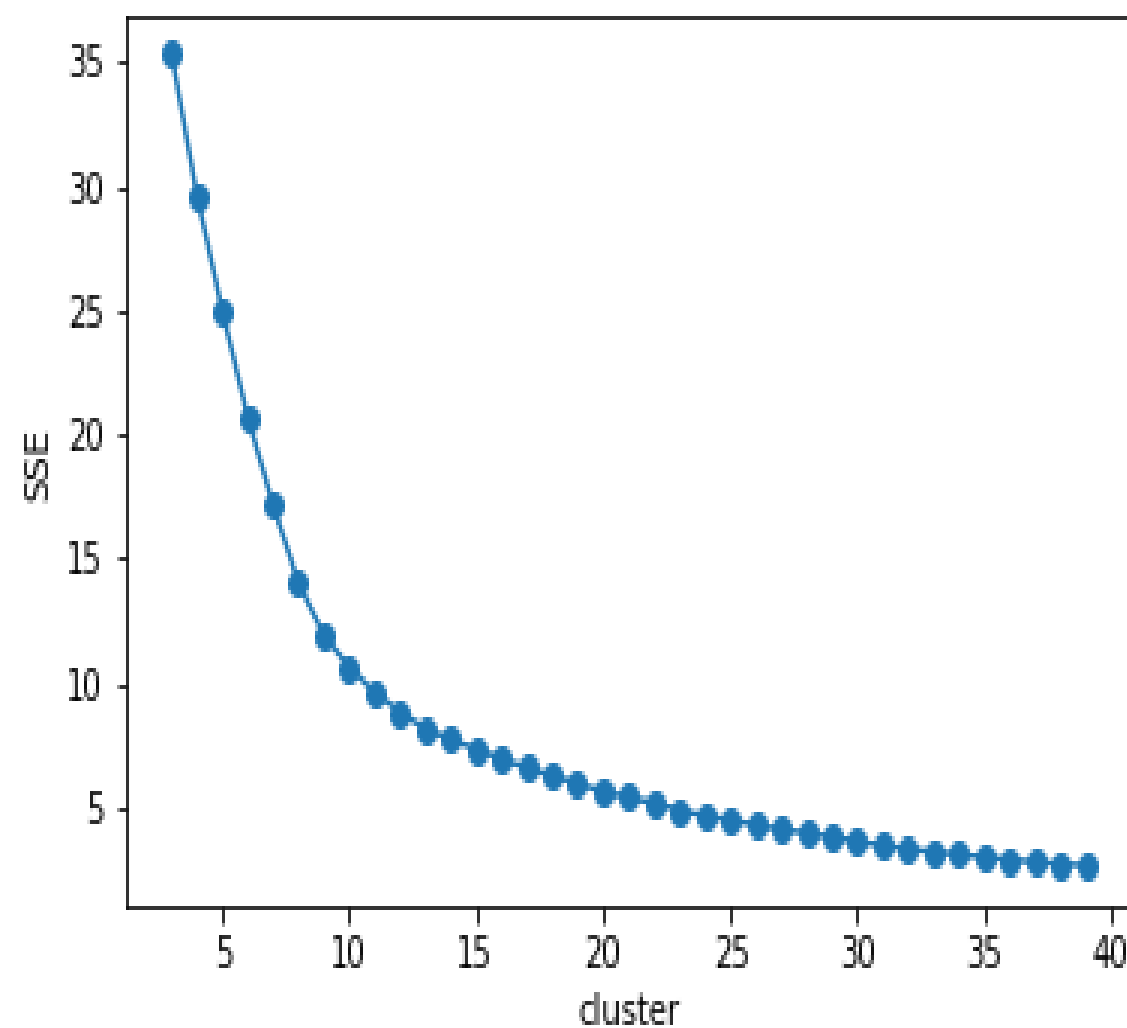
```
def k_means(df):
    sse = []
    for i in range(1, 50):
        km = KMeans(n_clusters = i, init = 'k-means++', random_state = 0).fit(df)
        sse.append(km.inertia_)
    plt.figure(figsize = (7,7))
    plt.plot(range(1, 50), sse, marker = 'o')
    plt.title('bus cluster', pad = 15, size = 15)
    plt.xlabel('cluster', size = 10)
    plt.ylabel('sse', size = 10)
    plt.savefig('k-means clustering.png')
    plt.show()
k_means(loc)
```

```
def k_means(df):
    sse = []
    for i in range(1, 50):
        km = KMeans(n_clusters = i, init = 'k-means++', random_state = 0).fit(df)
        sse.append(km.inertia_)
    plt.figure(figsize = (7,7))
    plt.plot(range(1, 50), np.log(sse), marker = 'o')
    plt.title('bus cluster', pad = 15, size = 15)
    plt.xlabel('cluster', size = 10)
    plt.ylabel('sse', size = 10)
    plt.savefig('k-means clustering with log scale.png')
    plt.show()
k_means(loc)
```

k 클러스터링
k -means++

k -means++

default vs log_scale



Where bus station?

```

km = KMeans(n_clusters = 20, init = 'k-means++', random_state = 0).fit(loc)
center = km.cluster_centers_

df['X좌표'] = df['X좌표'].apply(lambda x : np.round(x, 4))
df['Y좌표'] = df['Y좌표'].apply(lambda x : np.round(x, 4))

df_sol = df[['X좌표', 'Y좌표']].copy()

def distance(x, y):
    return np.sqrt(pow(x[0] - y[0], 2) + pow(x[1] - y[1], 2))

def bus_station(center, origin, copy_file, k):
    bus_stop_li = []
    for cen in center:
        dis_li = []
        for i in range(len(origin)):
            dis_li.append(copy_file.values[i].tolist() + [distance(copy_file.values[i], cen)])
        dis_topn = pd.DataFrame(dis_li, columns = ['X좌표', 'Y좌표', 'distance']).sort_values(by = 'distance')[0:k]
        bus_stop_li.append(pd.merge(dis_topn.iloc[:, :2], origin)['정류소명'].values.tolist())
    return bus_stop_li

k = int(input())
col = []
for i in range(1, k+1):
    col.append(f'cluster {i}')
answer = pd.DataFrame(bus_station(center, df, df_sol, 5), columns = col )
answer

```

위에서 구한 평가방식에서 클러스터를 20개, 25개 설정.

중점을 구하고 중점간 유클리드 거리를 사용해서
중점 기준으로 가장 가까운 n개의 정류소명을 추출

elbow method를 metric 으로 활용해
k 값을 20~25개 사이로 설정

우리는 k를 20개 25개로 설정했고
정확도 향상을 위해 5개의 배치지도표를 만들었다.

버스배차 # 20개 # 25개

임시배차표

	cluster 1	cluster 2	cluster 3	cluster 4	cluster 5
0	각산푸르지오건너	각산푸르지오앞	동부고교앞	동부고교건너	각산비나리공원건너
1	한국방송통신대학교앞	한국방송통신대학교건너	계명문화대학건너	계명문화대학앞	성서종합시장건너
2	농산물도매시장앞	월배차량기지앞	행정1리건너	방천리환경자원사업소	연호지
3	효목네거리1	효목네거리3	효목네거리2	한전동대구지사	한전동대구지사건너
4	응암2리	응암2리건너	응암2리앞	과학마을청아람앞	달성화성파크드림앞
5	성당초교앞	성당포스코더샵앞	본리롯데캐슬	남부초등학교앞2	성당포스코더샵건너
6	대천초등학교건너	대천초등학교앞	동천역1	럭키아파트앞	동천역2
7	한샘교회건너	한샘교회	두산스포츠킨너	수성관광호텔건너	수성관광호텔앞
8	용화사앞	노이2리건너	노이2리앞	솔비식당입구맞은편	솔비식당입구
9	감문1리3	감문1리4	감문1리1	감문1리2	문양리(서부실)
10	방짜유기박물관앞	방짜유기박물관건너	서당마을건너	서당마을앞	도학동
11	산격초등학교건너	산격초등학교앞	산격4동행정복지센터앞	실내체육관건너	실내체육관앞
12	상수도사업본부	상수도사업본부건너	건들바위건너	상수도중남부사업소앞	건들바위앞
13	용연사	용연사시설집단지구앞	용연사시설집단지구건너	귀남지후계소1	귀남지후계소2
14	달서자동차검사소앞	달서자동차검사소건너	유천교(유천치안센터건너)	유천교(유천치안센터앞)	유천교1
15	서부소방서	북부정류장	서진중학교건너	평리네거리2	서대구중학교앞
16	세하앞	현풍경남은행앞	현풍경남은행건너	세하건너	제일풍경채건너
17	노변청구타운앞	고산2동행정복지센터건너	노변청구타운건너	고산2동행정복지센터앞	노변중학교건너
18	동북주유소건너	동북주유소앞	화전동1	화전동2	대원사앞(공산동)
19	옥분리(구진터)	주리(삼진식당)	주2리1	주2리(최정사입구)	주1리1

	cluster 1	cluster 2	cluster 3	cluster 4	cluster 5
0	기업은행죽전동지점앞	기업은행죽전동지점건너	죽전119안전센터건너	죽전119안전센터앞	대구의료원라파엘헬빙센터건너
1	월배차량기지건너	청남타운건너	월배차량기지앞	다사역(3번출구)	방천리환경자원사업소
2	대구소방본부앞	대구소방본부건너	대구콘서트하우스앞	대구역북편네거리2	대구콘서트하우스건너
3	성요셉요양병원건너	성요셉요양병원앞	대동공업건너	대동공업	평광3차아파트건너
4	범어4동행정복지센터앞	범어4동행정복지센터건너	만촌육교1	만촌육교2	만촌역(1번출구)
5	각산역(1번출구)	동부고교건너	동부고교앞	각산역(2번출구)	동호초등학교건너
6	대천초등학교건너	대천초등학교앞	동천역1	동천역2	한양공작맨션앞
7	현내1리	무등2리	무등2리입구건너	국립식량과학원대구시험지	연젓마을입구
8	응암1리1	응암1리2	구지농공단지1	구지농공단지2	응암3리1
9	방짜유기박물관건너	방짜유기박물관앞	서당마을건너	서당마을앞	도학동
10	경창산업	한국산업인력공단건너	한국산업인력공단앞	유림물산	삼보
11	대구가톨릭대학병원앞	대구가톨릭대학병원건너	대명1동행정복지센터건너	대명초등학교건너	대명1동행정복지센터앞
12	옥분리(구진터)	주리(삼진식당)	주1리1	주1리2	주2리1
13	산림조합중앙회앞	산림조합중앙회건너	동구청역(4번출구)	동구청역(3번출구)	동대구역지하도2
14	비산초등학교앞	비산초등학교건너	대평중학교앞	대평중학교건너	만평네거리
15	영남고등학교건너	성모여성병원앞	롯데백화점상인점건너	성모여성병원건너	영남고등학교앞
16	동곡초등학교건너	동곡초등학교앞	하빈농협동곡지점2	하빈농협동곡지점1	하목정(종점)
17	봉우청구새마을앞	봉우청구새마을건너	원불교앞	이시아폴리스더샵1단지	파군재삼거리2
18	달성군청앞	달성군청건너	옥포농협강림지소앞	옥포농협강림지소건너	경서중학교앞
19	수성고등학교앞	수성고등학교건너	불교한방병원건너2	불교한방병원앞2	불교한방병원건너1
20	정대1리(배깎매남)	정대1리(안매남)	정대리(대구미술광장)	구삼폭포건너	구삼폭포앞
21	수성동삼익맨션앞	중동치안센터건너	중동치안센터앞	삼환맨션	희망교1
22	명곡미래빌3단지건너	명곡미래빌3단지앞	명곡미래빌5단지앞	명곡미래빌5단지건너	명곡미래빌2단지건너
23	고산역화성파크드림앞	고산역화성파크드림건너	대구농업마이스터고앞	고산정수사업소건너	신매역(3번출구)
24	연호지	율하역(2번출구)	신평동2	신평동1	행정1리건너

01

하나,

WHY

기획 배경
반딧불버스의 의미

02

두울,

HOW

분석 방법
임시배차표

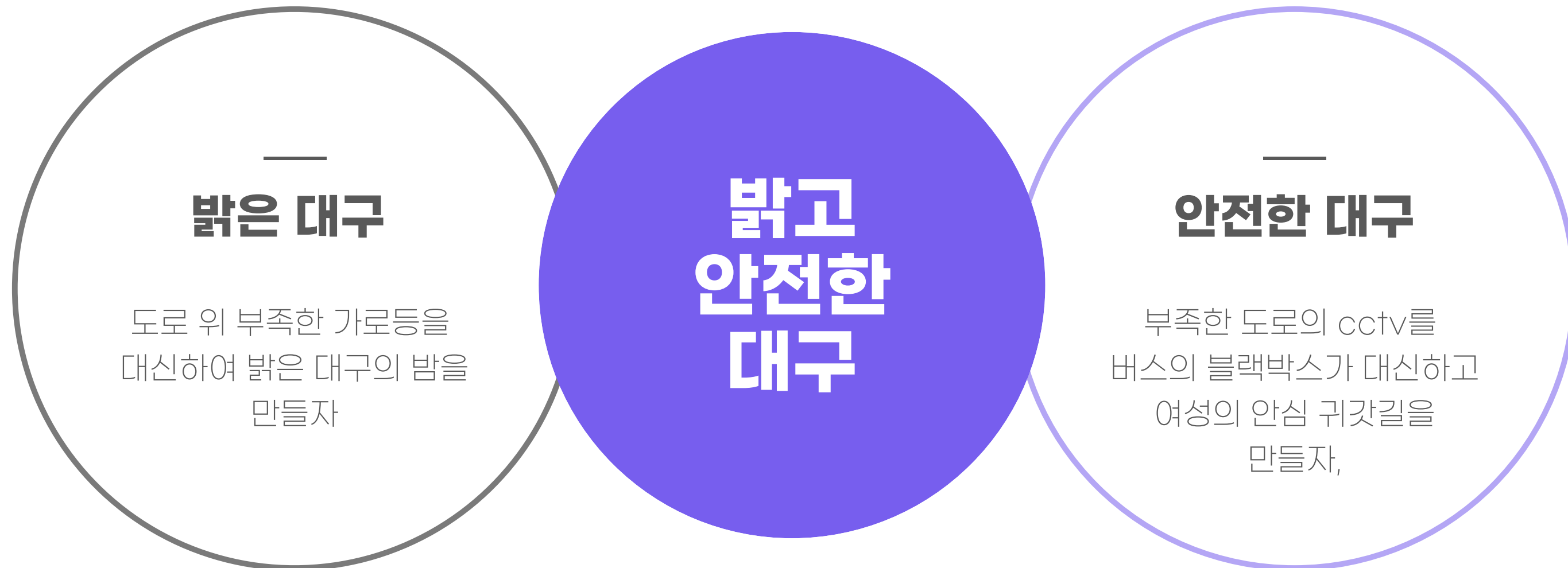
03

세엣,

RESULT

반딧불버스의 기대효과

Advantage



Advantage

목표 : 심야버스를 통한 대구 시민 및 관광객 편의 증대

- 1 도로의 밝은 빛
- 2 부족한 도로의 cctv 충원
- 3 여성 안심 귀가길



Advantage - (1)

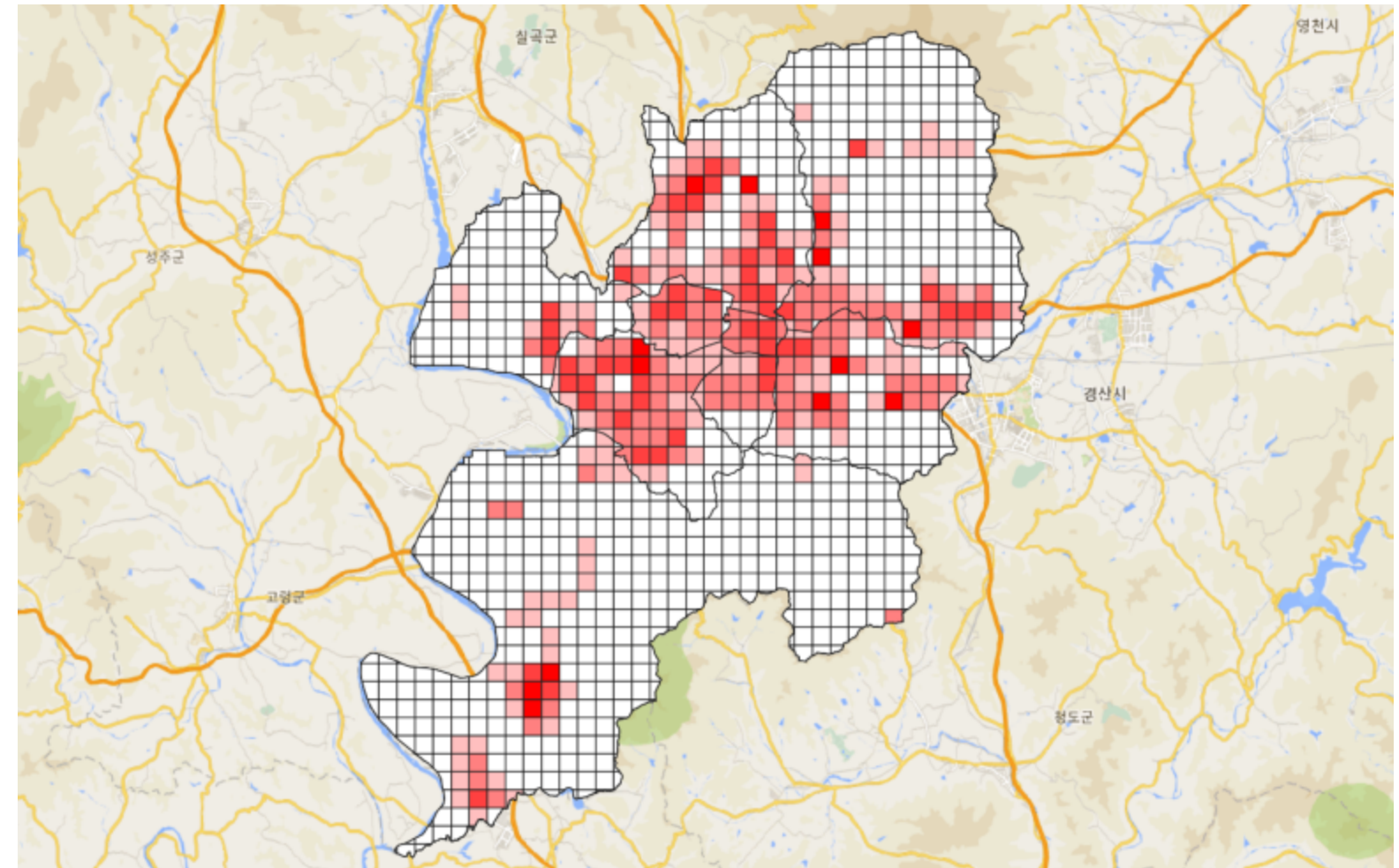
1 도로의 밝은 빛

```
gen_model.wv.most_similar('가로등', topn = 20 )
```

```
[('신호등', 0.5138801336288452),
 ('모퉁이', 0.4949066638946533),
 ('불빛', 0.49401572346687317),
 ('고상', 0.4724493622779846),
 ('보안등', 0.44558054208755493),
 ('모퉁이', 0.44086316227912903),
 ('밤길', 0.4349859356880188),
 ('전주', 0.4191935062408447),
 ('전지', 0.41051000356674194),
 ('전구', 0.40936267375946045),
 ('밤', 0.4019755721092224),
 ('무법', 0.392807275056839),
 ('두문', 0.39139091968536377),
 ('조명', 0.3911404013633728),
 ('밀', 0.3880215883255005),
 ('타트', 0.3851237893104553),
 ('전봇대', 0.38386544585227966),
 ('암흑', 0.38213232159614563),
 ('홀다리', 0.38156354427337646),
 ('교체', 0.3806523084640503)]
```

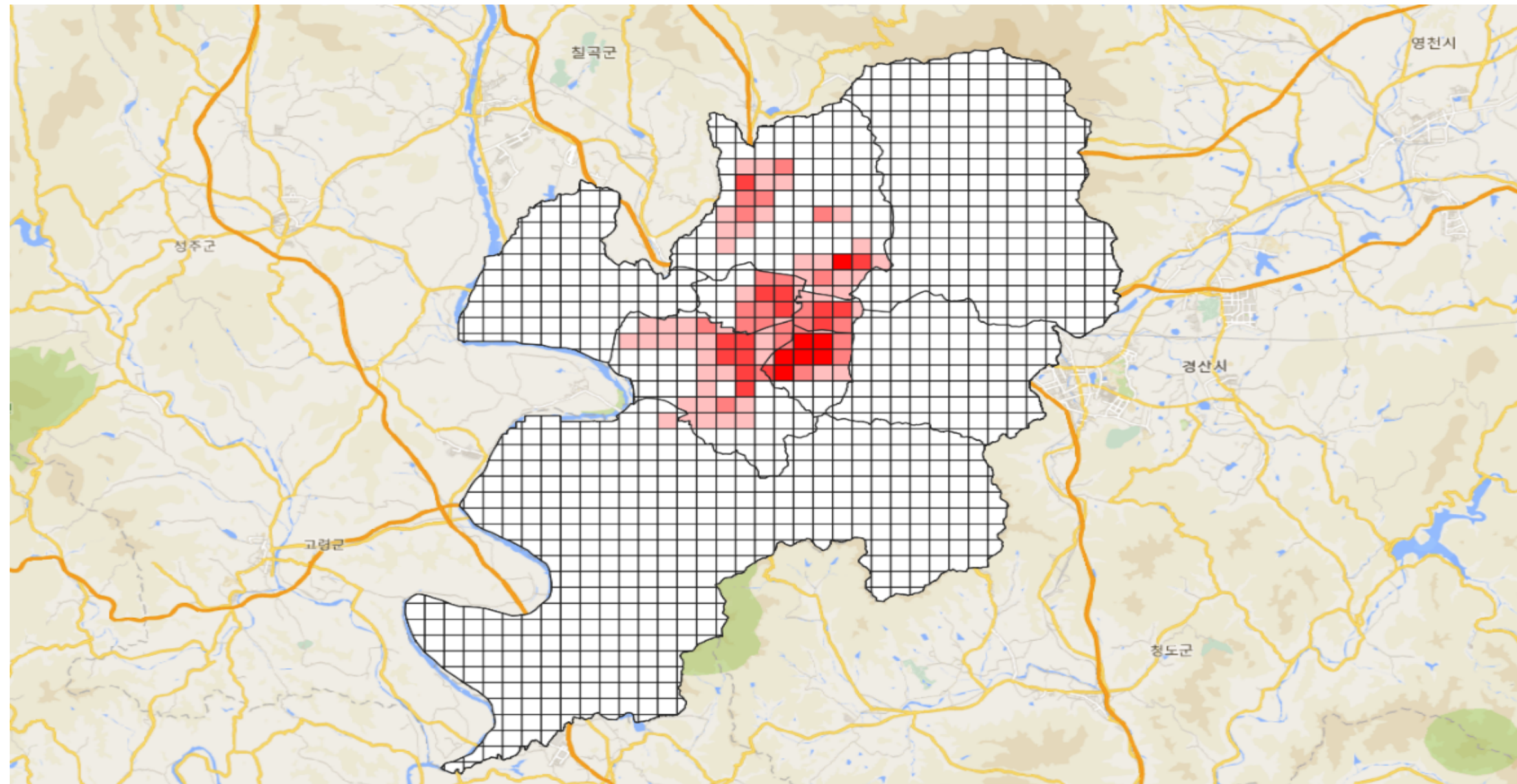
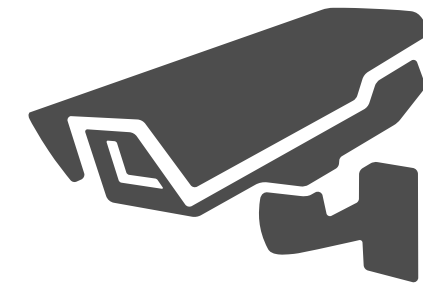


가로등개수



Advantage - (2)

2 부족한 도로의 cctv 충원



Advantage - (3)

3 여성 안심 귀가 서비스

택시 이용에 두려워 하는
여성들에게 안전한 귀갓길을
제공 가능할 것으로 예상



출처 : 파주시청

결론

밝고 # 건강한 #대구

02

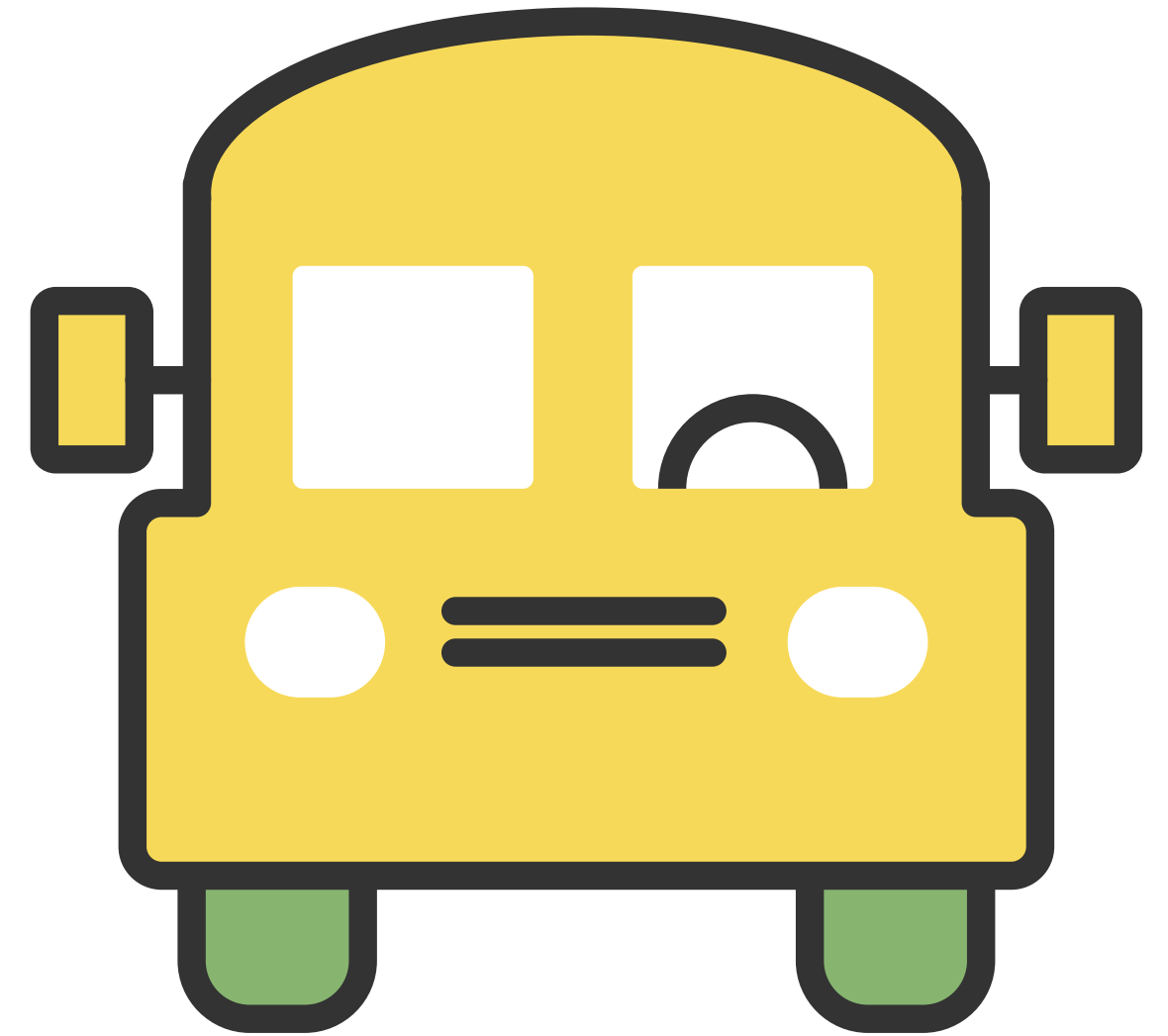
RMSE 와 silhouette score를 metric으로 하여 평가를 해본 결과 elbow method를 사용하였고 k=20,25개로 설정했다.

01

시간별 택시와 버스의 이용자 수의 대비를 통해 특정 시간 심야버스를 운행 시 시민들이 큰 만족을 느낄 것으로 예상됨.

03

임시배차표를 5개 만들었다.
22시부터 24시까지 시범운행을 한 다음
수요량과 경제적 효과를 검토후
타 행정구의 심야버스 정책을 벤치마킹.



활용한 데이터 목록

데이터 활용센터

6-1 통신사 유동인구 데이터

9 택시 승하차

10 대구시 버스데이터

19 북구 민원 상담자료

D-hub

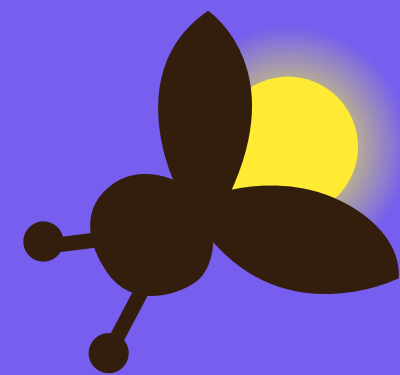
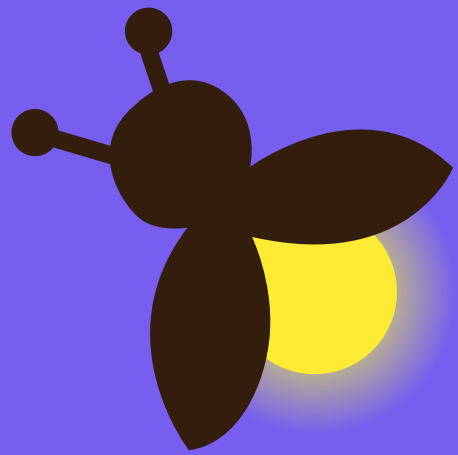
교통cctv

가로등 수

Hello, 반딧불버스

대구의 밤을 밝혀줄 심야버스

김동언 김태형



감사합니다,
반딧불버스였습니다!