

2020 응용통계학과 분석 공모전

서울시  
**자동차 극장**  
최적입지 추천

식스틴

조형래 · 박주형 · 김동우 · 김재홍

# CONTENTS

---

## 1. 프로젝트 소개

프로젝트 배경, 분석 방향

## 2. 요인 설정

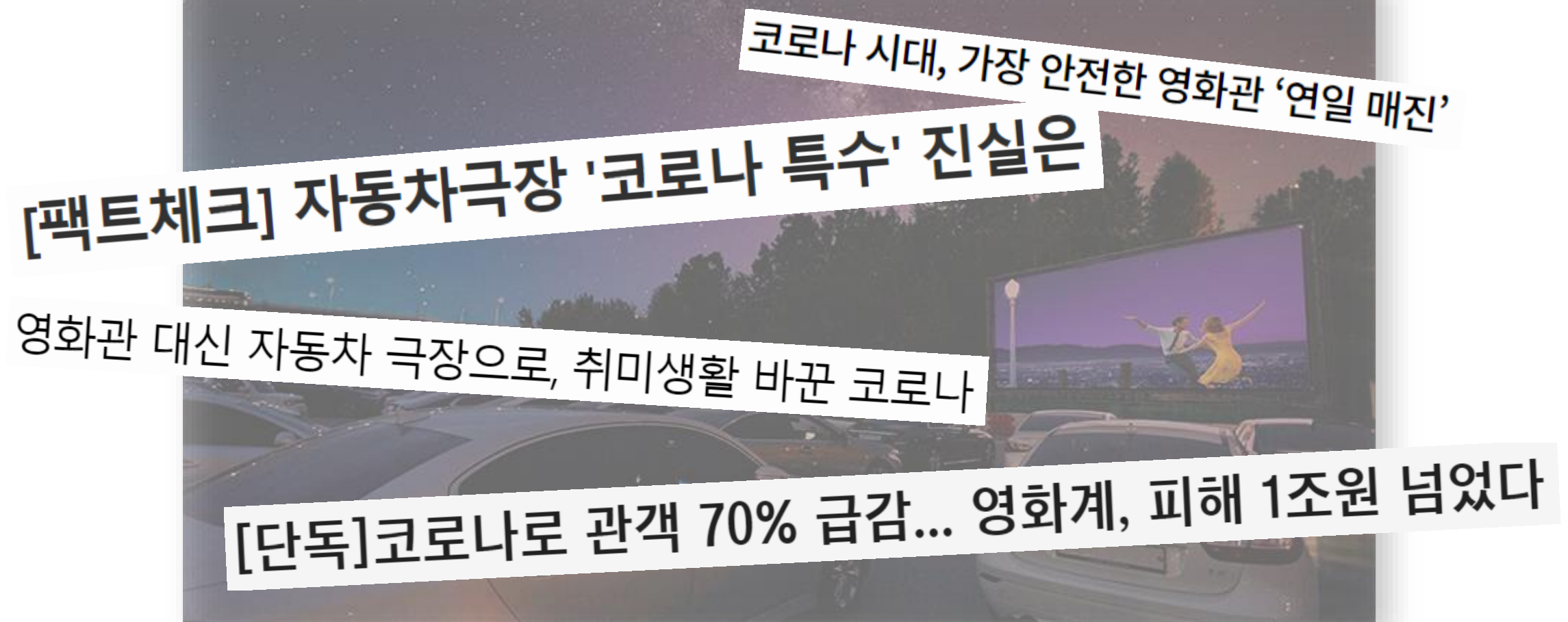
요인 설명, 유의미한 요인

## 3. 분석 과정 및 결과

다중회귀분석을 통한 모형 도출

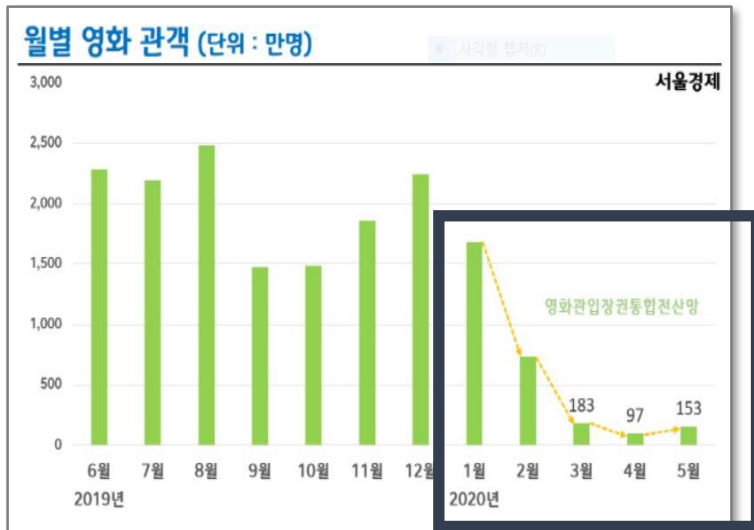
## 4. 결론

최종입지 추천 및 활용방안



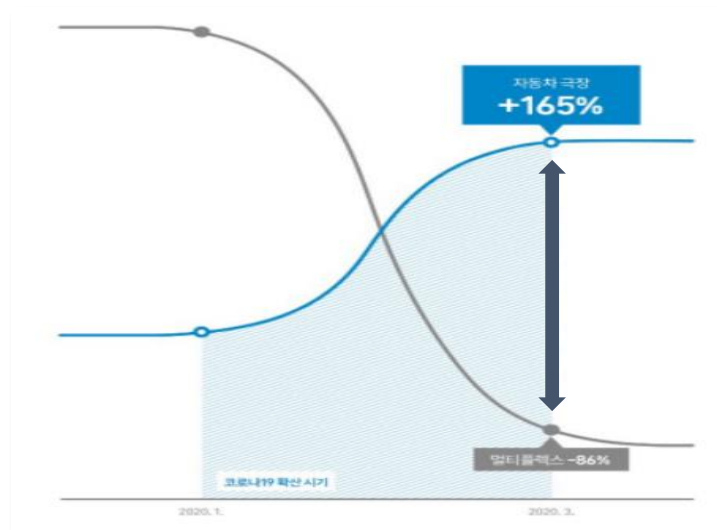
코로나 19로 인한 **영화관** 이용 부담  
BUT, 문화 생활(영화)에 대한 욕구는 여전!

## 코로나19로 인한 **영화** 관객 ↓



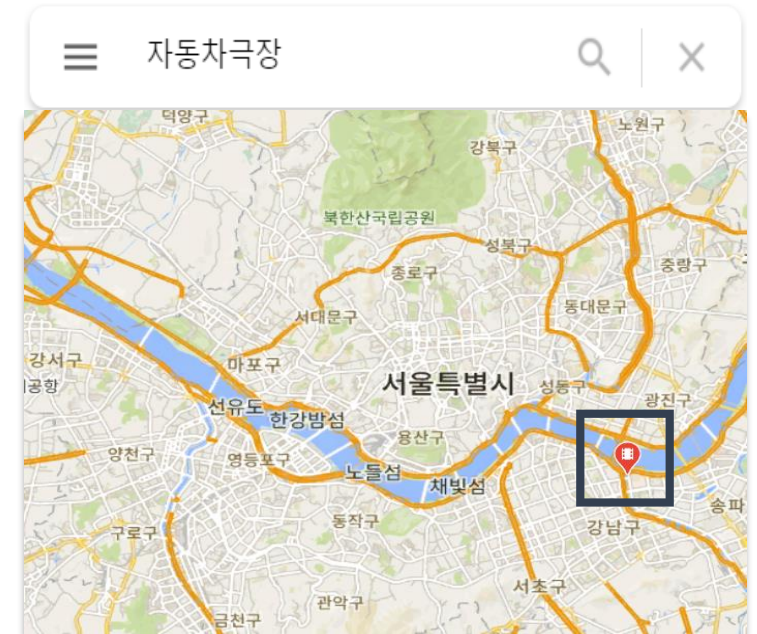
출처: 서울경제

## 자동차 극장 vs 영화관 수요



출처: SK텔레콤

## 서울시 **자동차 극장** 현황



코로나 19로부터 안전한, “서울시 **자동차 극장** 필요성”

데이터 수집 및 처리

분석 및 모델링

결과 및 입지선정

영화관  
좌석수

주차장  
면 수

도로면적

GRDP

20-34세  
인구

토지매매  
실거래가

※ 영화관 좌석 수와 관련된 14가지 요인 수집

선형성 확인  
(다중공선성 문제)

이상치 확인 및 제거

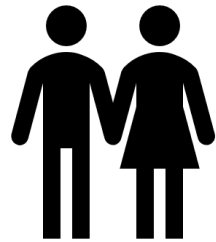
다중 회귀분석

변수선택법을 통한  
최적 모형 도출

입지 분석 및  
입지 선정

자동차 극장  
입지 타당성 평가

## 1. 인구 요인



주민등록 인구	자치구
생활 인구	자치구
20-34세 인구	자치구
상존인구	길 / 건물 단위
선호 인구 비율	영상 / 문화예술

※ 출처 : 서울시 열린 데이터 광장

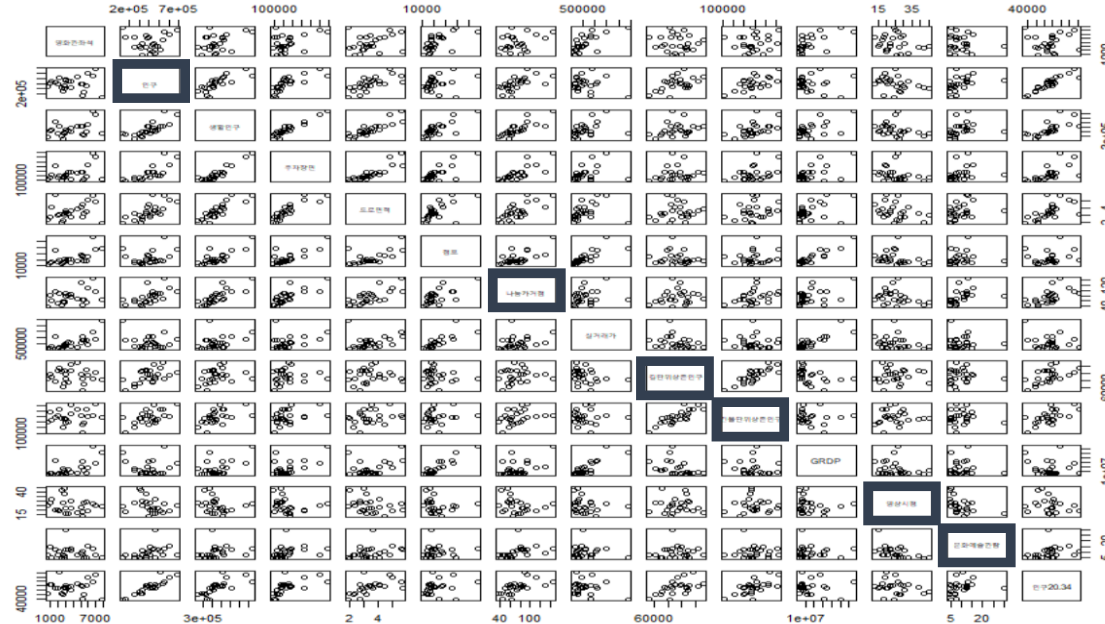
## 2. 지역 특성 요인



도로 면적	단위: $m^2$
점포 수	소매점
실거래가	건물 / 토지
나눔카 거점 수	자치구
총 생산(GRDP)	자치구
주차장 면 수	자치구
영화관 좌석 수	자치구

※ 출처 : 서울시 열린 데이터 광장

## 선형성 확인



종속 변수와 독립 변수간 선형성 확인 결과

**‘주민등록 인구’, ‘상존 인구’** (길/건물 단위)  
**‘나눔카 거점 수’, ‘선호인구 비율’** (영상 시청/문화예술)

6개 요인 제거

## 독립 변수

주민등록 인구  
 주차장 면 수  
 도로 면적  
 소매점 점포 수

상존 인구  
 토지 매매 실거래가  
 선호 인구 비율

## 종속 변수

영화관 좌석 수

## 다중공선성 문제

요인	생활 인구	주차장 면 수	도로 면적	점포 수	실거래가	GRDP	20-35세 인구
VIF	21.60	13.43	4.84	37.25	6.07	17.62	5.27



VIF > 10 이면서, 가장 높은

**‘점포 수’, ‘생활 인구’** 2개 요인 제거

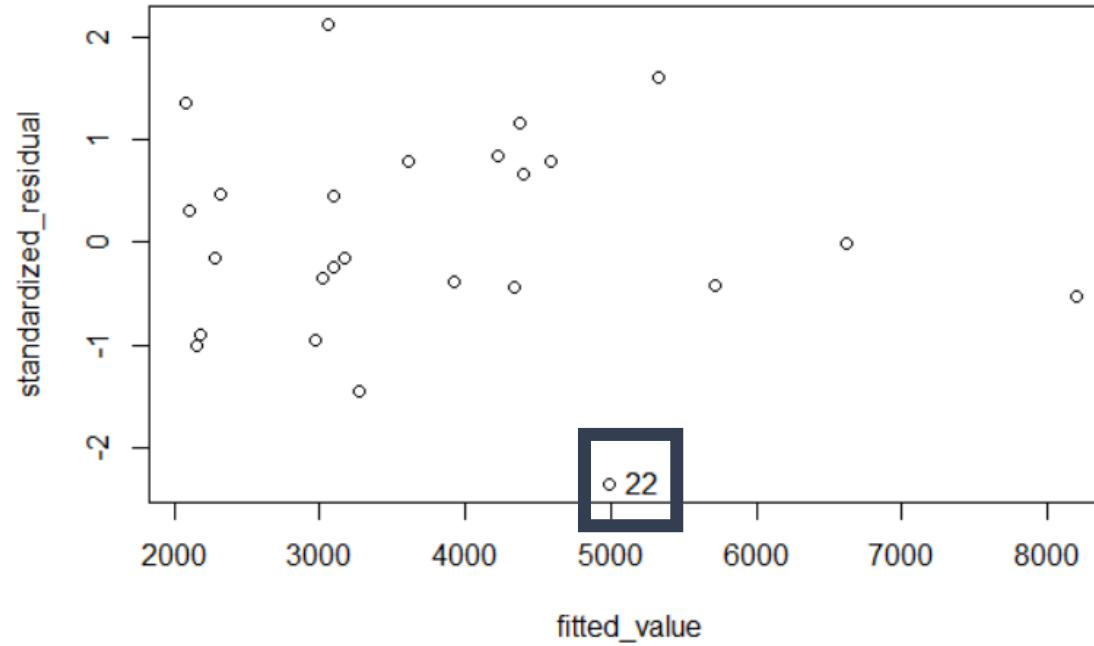
요인	주차장 면 수	도로 면적	실거래가	GRDP	20-34세 인구
VIF	7.68	4.21	4.37	7.18	2.69

“다중공선성” 문제 해결

1차 회귀 모형:  $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \epsilon_i$



### 이상치(Outlier) 제거



산점도 (fitted value & standardized residual) 확인 결과

**22**번 관측치를 이상치(Outlier)로 판단

## 단계적 변수선택법

```
> reg = lm(영화관좌석~. , data=final)
> step(reg, direction = 'both', scope=~1))
Start: AIC=343.51
영화관좌석 ~ 주차장면 + 도로면적 + 실거래가 + GRDP + 인구20.34
```

	Df	Sum of Sq	RSS	AIC
- 인구20.34	1	234642	24173865	341.75
- 주차장면	1	234816	24174039	341.75
- GRDP	1	341667	24280890	341.85
- 도로면적	1	1684639	25623862	343.14
<none>			23939223	343.51
- 실거래가	1	13988333	37927556	352.56

```
Step: AIC=341.75
영화관좌석 ~ 주차장면 + 도로면적 + 실거래가 + GRDP
```

	Df	Sum of Sq	RSS	AIC
- 주차장면	1	74371	24248237	339.82
- GRDP	1	168637	24342503	339.91
- 도로면적	1	1518592	25692457	341.21
<none>			24173865	341.75
- 실거래가	1	13820689	37994555	350.60

```
Step: AIC=339.82
영화관좌석 ~ 도로면적 + 실거래가 + GRDP
```

	Df	Sum of Sq	RSS	AIC
- GRDP	1	103150	24351387	337.92
<none>			24248237	339.82
- 도로면적	1	7570990	31819226	344.34
- 실거래가	1	14276462	38524699	348.93

```
Step: AIC=337.92
영화관좌석 ~ 도로면적 + 실거래가
```

	Df	Sum of Sq	RSS	AIC
<none>			24351387	337.92
- 도로면적	1	8553790	32905177	343.15
- 실거래가	1	46958463	71309849	361.71

```
Call:
lm(formula = 영화관좌석 ~ 도로면적 + 실거래가, data = final)
```

```
Coefficients:
(Intercept)      도로면적      실거래가
  1.175e+02    6.393e+02    2.808e-03
```

## 단계적 변수선택법을 통한 모형 설정

Step: AIC=337.92

영화관좌석 ~ 도로면적 + 실거래가

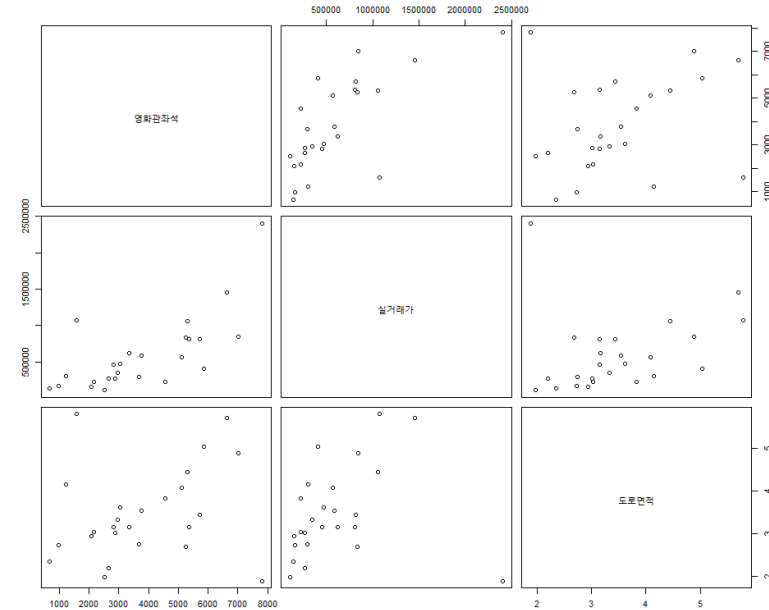
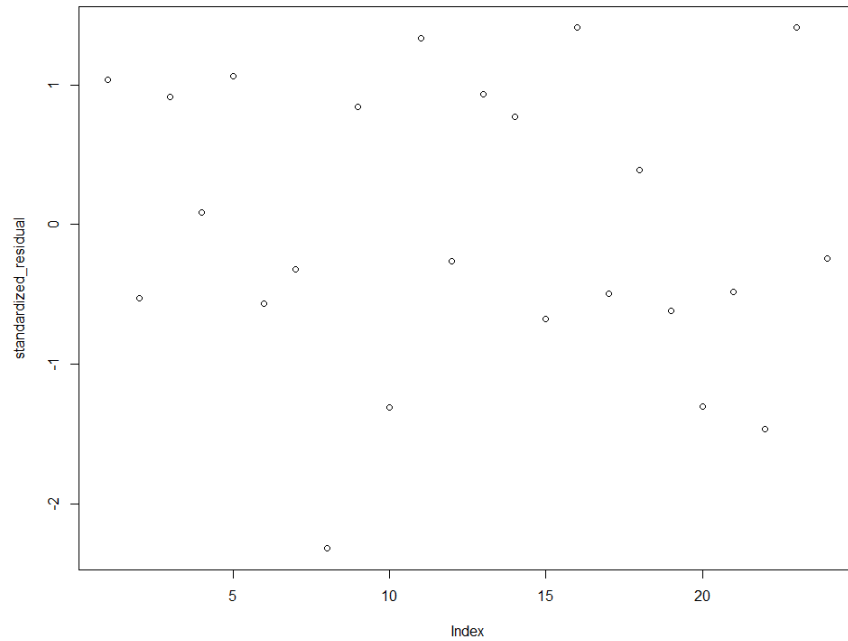
	Df	Sum of Sq	RSS	AIC
<none>			24351387	337.92
- 도로면적	1	8553790	32905177	343.15
- 실거래가	1	46958463	71309849	361.71

Call:

lm(formula = 영화관좌석 ~ 도로면적 + 실거래가, data = final)

$$2\text{차 회귀 모형: } y_i = \beta_0 + \beta_2 x_2 + \beta_3 x_3 + \epsilon_i$$

#### 회귀 분석 기본 가정



회귀 분석 기본 가정 확인 결과

**독립성, 등분산성, 정규성, 선형성 만족**

## 최종 모형 도출

```
> reg = lm(영화관좌석 ~ 도로면적 + 실거래가, data=final)
> summary(reg)
```

```
Call:
lm(formula = 영화관좌석 ~ 도로면적 + 실거래가, data = final)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-2385.0  -600.6  -259.0   958.5  1396.8
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.175e+02  8.273e+02   0.142  0.8884
도로면적     6.393e+02  2.354e+02   2.716  0.0129 *
실거래가     2.808e-03  4.413e-04   6.364 2.62e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1077 on 21 degrees of freedom
Multiple R-squared:  0.721,    Adjusted R-squared:  0.6944
F-statistic: 27.13 on 2 and 21 DF,  p-value: 1.509e-06
```

Multiple R-squared: **0.721**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	117.5	827.3	0.142	0.8884
도로면적	639.3	235.4	2.716	0.0129 *
실거래가	0.0028	0.00044	6.364	2.62e-06 ***

최종 모형:  $\hat{y}_i = 117.5 + 639.3x_1 + 0.0028x_2$

※  $x_1$ : 도로면적,  $x_2$ : 실거래가

도로면적 넓을수록, 실거래가 높을수록

**영화관 좌석 수 ↑**

## 영화관 좌석 추정

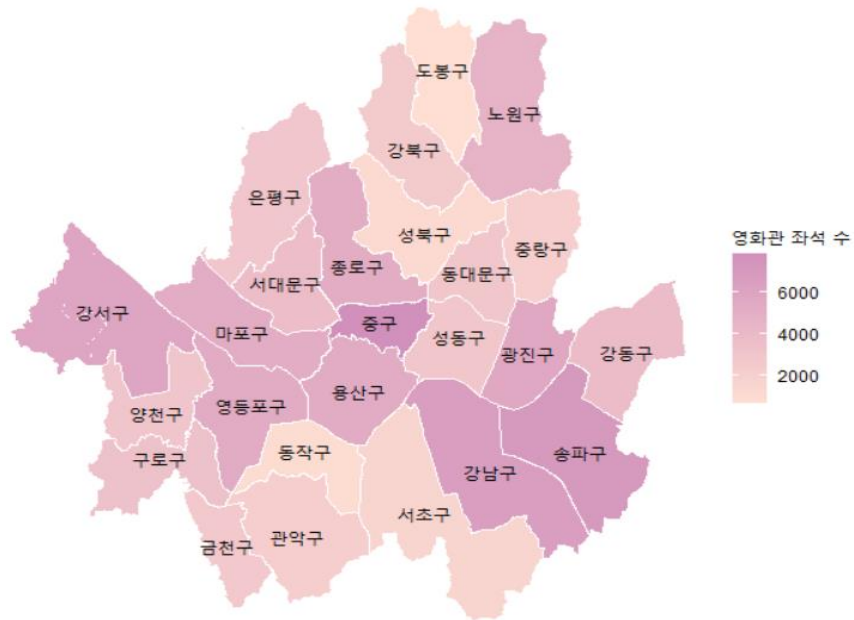
자치구	영화관 좌석	영화관 좌석추정	영화관 좌석 차이
종로구	5252	4183.146	1068.854
중구	7822	8082.927	-260.927
용산구	5361	4407	953.9997
강북구	2516	1679.318	836.6816
성북구	1238	3622.968	-2384.97
동대문구	2822	3414.753	-592.753
		⋮	
서초구	1601	6828.335	-5227.33
강남구	6619	7861.99	-1242.99
송파구	7017	5620.237	1396.763
강동구	3778	4035.097	-257.097

영화관 좌석 차이가 양수인 경우  
**과잉공급**으로 판단!

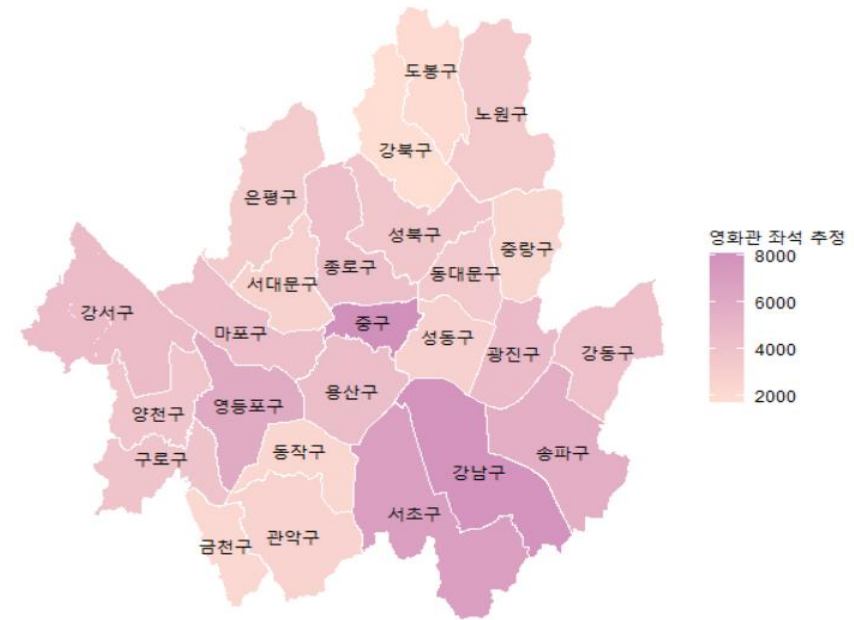
영화관 좌석 차이가 음수인 경우  
**과소공급**으로 판단!

## 영화관 좌석 추정 시각화

영화관 좌석 수



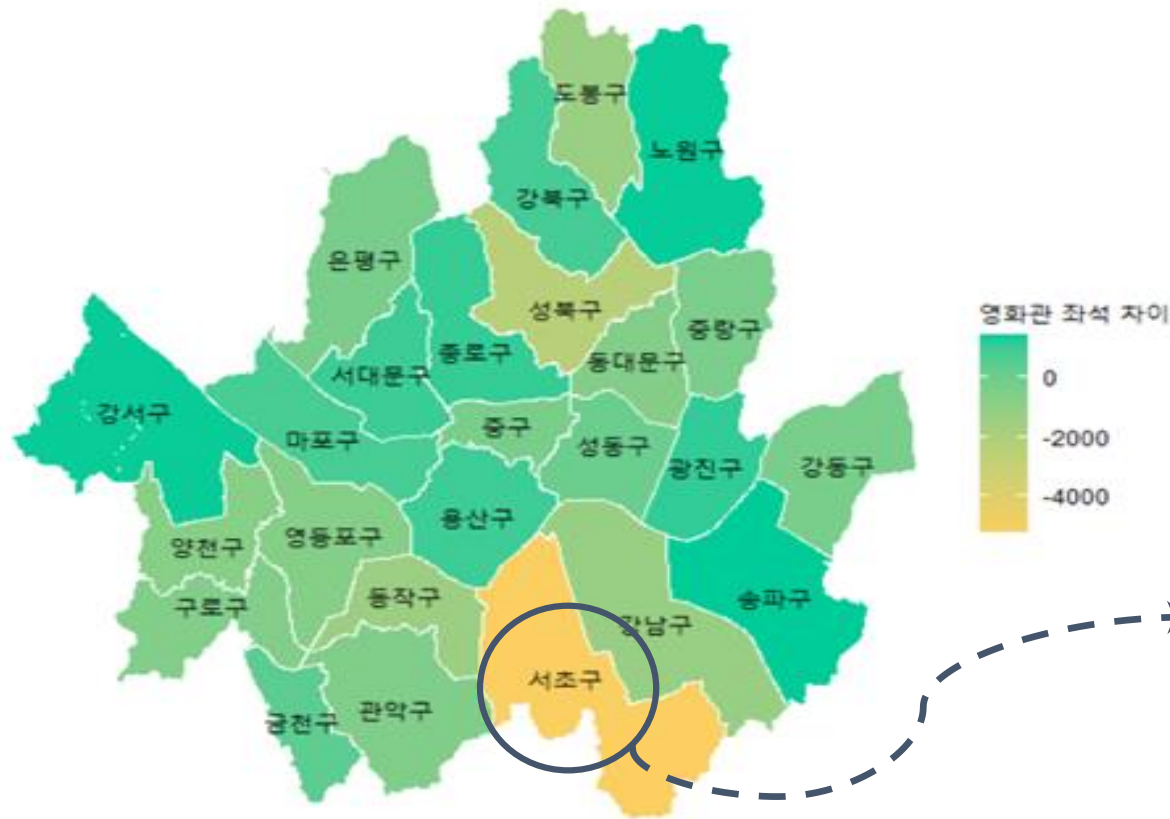
영화관 좌석 추정



구별 상이한 영화관 좌석 수  
좌석 수와 추정 값은 대체로 비례

## 영화관 좌석 추정 시각화

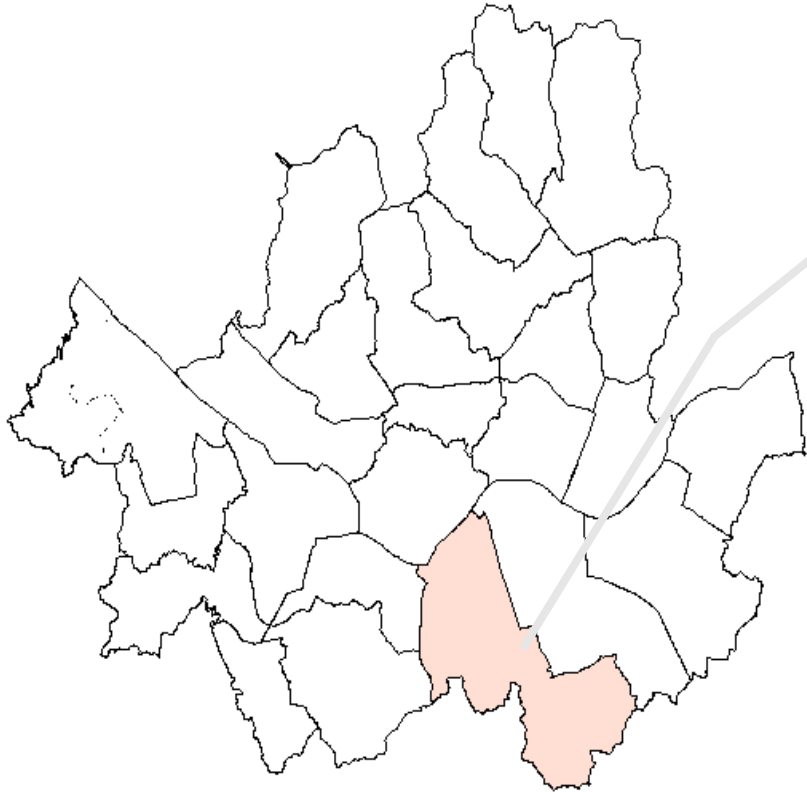
영화관 좌석 추정 값과 차이

영화관 & 자동차 극장  
모두 **영화 상영** 의 장소**코로나 19** 상황 **보다 안전한, 자동차 극장 설치!**

추정 값보다 약 5227석이나 적은

**‘서초구’** 내에 설치 추천!

### 추가적인 고려사항



**서초구** 내에

자동차 극장의 최적입지를 추천한다면?!

기존 영화관과 공통적인 요인 외에



**차량 접근 용이성**



**토지 활용 가능 여부**



**주변 편의 시설**

추가적으로 고려

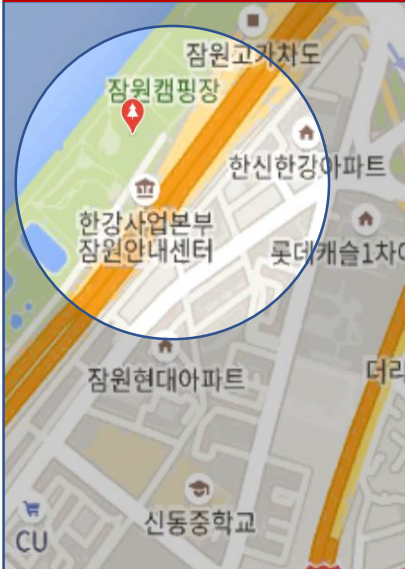


## 후보지

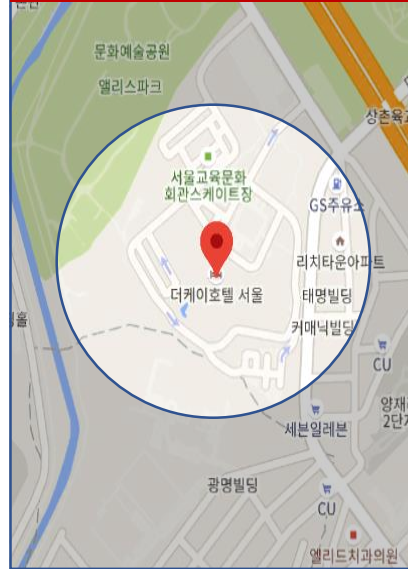
★ 수익성

★ 실현 가능성

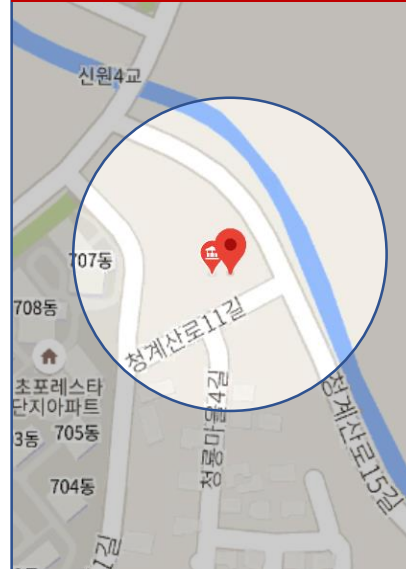
### 1. 잠원 한강 공원



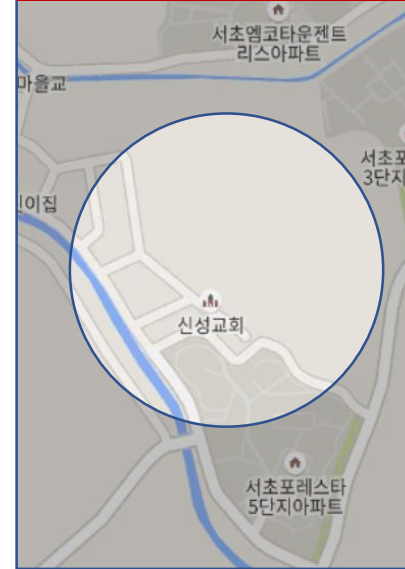
### 2. The K호텔



### 3. 서울 추모 공원



### 4. 용두리들



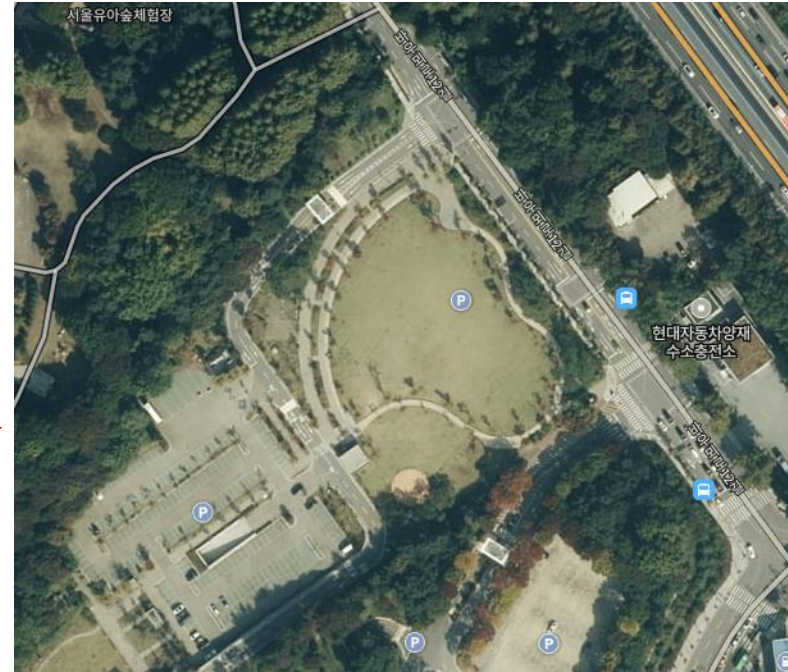
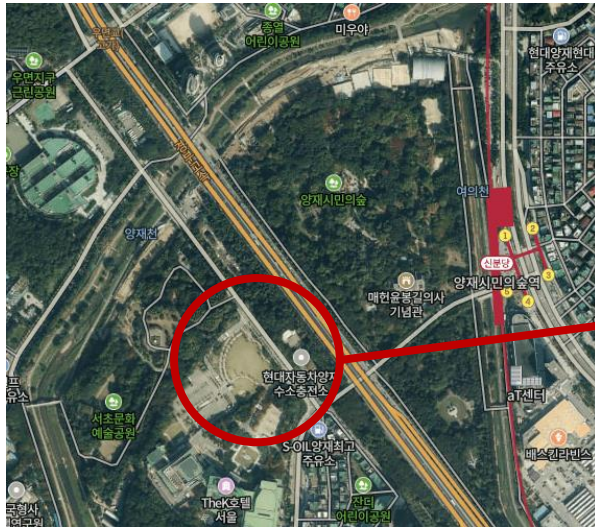
### 5. 내곡동 주민 센터



“5개의 후보지 선정”

### 최종 후보지 1

#### The K호텔



면적 약  $3,500 m^2$   
양재 IC 주변에 위치



양재 시민의 숲 사이에  
위치하여 어두움 조성

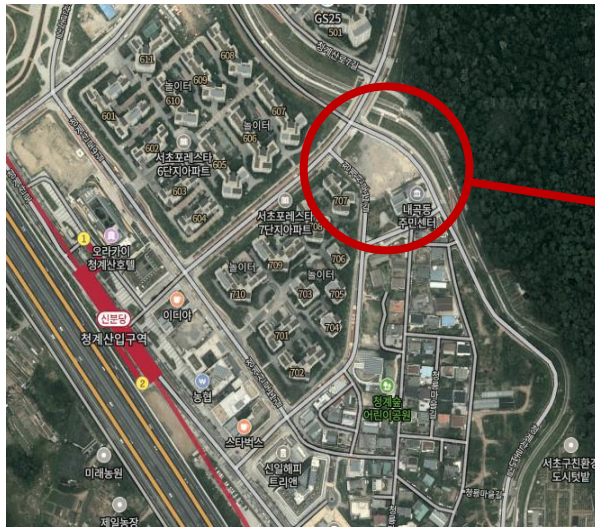


The K호텔과 협업하여  
사업 및 이벤트 기획으로  
경제적 효과 극대화



## 최종 후보지 2

내곡동 주민센터 주차장



면적 약  $2,200 m^2$   
양재 IC 주변에 위치



인릉산 앞에  
위치하여 어두움 조성



공공 복지 사업으로  
추진할 경우, 장소 협조 수월

### 결론 및 활용방안

☞ 코로나 19로 인한 침체된 영화 산업

☞ 회귀 분석을 통해  
영화관 좌석 수에 영향을 미치는 요인 파악  
☞ 구별 실거래가, 구별 도로면적

☞ 회귀 분석 모형을 통해 서울시 내  
영화관 추가 설치 필요 구역  
☞ 서초구

☞ 코로나 19 상황을 고려하여 기존의 영화관 대신  
☞ 자동차 극장 설치 제안

☞ 코로나 19 이후에도 지속적인  
자동차 극장 이용 기대!

### 참고 문헌

정민아, 「포스트 코로나 시대 영화관과 영화산업 전망」, 2020, 8쪽.

서충섭, “[서민들 일상의 변화] 언택트 시대, 선명해지는 양극화 대비해야”, 무등일보, 2020.10.07.

### 사용 데이터

1. 구별 서울시 나눔카 거점리스트 정보 - 서울 열린데이터 광장
2. 구별 1  $m^2$ 당 상업 업무용 건물 및 토지 매매 실거래가 - 국토교통부 실거래가 공개시스템
3. 구별 길 단위 상존인구 - 우리마을 상권분석
4. 구별 건물 단위 상존인구 - 우리마을 상권분석
5. 구별 소매점 점포 수 데이터 - 우리마을 상권분석
6. 구별 지역 내 총 생산(GRDP) - 서울 열린데이터 광장
7. 구별 주민등록 인구 - 서울 열린데이터 광장

### 사용 도구



RStudio



Excel

8. 구별 생활인구 - 서울 열린데이터 광장
9. 구별 주차장 면수 - 서울 열린데이터 광장
10. 구별 도로현황 (도로율) - 서울 열린데이터 광장
11. 구별 20-34세 주민등록 인구 - 서울 열린데이터 광장
12. 구별 영상시청 선호 인구비율 - 서울 열린데이터 광장
13. 구별 문화예술 선호 인구비율 - 서울 열린데이터 광장