

청소년 팝업 심터

청소년 전문상담센터 접근성 개선



팝콘 조 - 조지현, 김미화, 조혜진

목 차

분석의 필요성 및 목적

- 주제선정이유
- 청소년들의 위험신호
- 분석목적

위치선정

- 청소년들의
마음 들여다보기
- 위치선정과정

분석결과 및 향후 과제

- 위치선정 결과
- 팝업쉼터 제안



데이터 수집 및
가공



분석의 필요성 및
목적



위치선정



분석결과 및
향후 과제

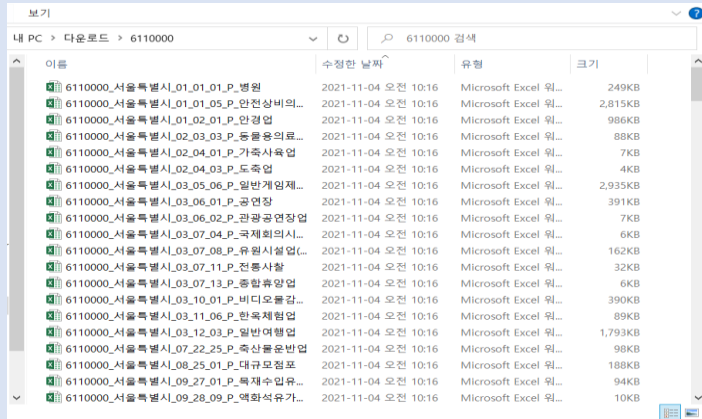


데이터 수집 및 가공

출처	자료명	활용
통계청	사망원인 통계	청소년 사망원인 통계 (선그래프)
	국민건강영양조사:스트레스인지율 추이	성인과 청소년의 스트레스 인지율 T검정/ 상자그림
	청소년건강행태조사:스트레스 인지율	청소년의 남학생 여학생의 스트레스 인지율 T검정/ 상자그림
청소년데이터플랫폼	여성가족부_청소년상담지원현황	청소년 상담내용 그래프(막대)
	청소년 고민 순위 정보	청소년 고민내용(워드클라우드/감정분석)
	청소년 관심키워드 분석 데이터	청소년 관심 키워드 (워드클라우드)
여성가족부	청소년 종합실태 조사	코로나19 이후 삶의 변화, 청소년 사망원인
서울열린데이터광장	서울시 청소년인구 통계	구별 청소년 수 그래프(막대)
	서울시 주요 공원현황	구로구 공원과의 거리계산
	서울특별시 학원 교습소정보	구별 청소년 수와 학원과의 상관관계
서울 정보 포털 광장	자치구 청소년상담복지센터 인력 소요 조사	청소년상담복지센터 상담 인력 현황 및 대기일수
공공데이터포털	보건복지부_서울시 년도별 자살자 수	청소년과 유소년의 자살율 그래프(선)
LOCALDATA	서울시 업종별 데이터	구별 청소년 수와 PC방, 코인노래방, 독서실의 상관관계
Do it! 쉽게 배우는 R 텍스트 마이닝	knu_sentiment_lexicon	감정분석을 위한 감정사전(군산대학교)

데이터 수집 및 가공

데이터 수집 예시 PC방, 코인노래방, 학원 데이터 수집



이름	수정된 날짜	유형	크기
6110000_서울특별시.01.01.01_P_병원	2021-11-04 오전 10:16	Microsoft Excel 워...	249KB
6110000_서울특별시.01.01.05_P_안전상비의...	2021-11-04 오전 10:16	Microsoft Excel 워...	2,815KB
6110000_서울특별시.01.02.01_P_안전염	2021-11-04 오전 10:16	Microsoft Excel 워...	986KB
6110000_서울특별시.02.03.03_P_동충여의로...	2021-11-04 오전 10:16	Microsoft Excel 워...	88KB
6110000_서울특별시.02.04.01_P_가족사육업	2021-11-04 오전 10:16	Microsoft Excel 워...	7KB
6110000_서울특별시.02.04.03_P_도축업	2021-11-04 오전 10:16	Microsoft Excel 워...	4KB
6110000_서울특별시.03.05.06_P_일반게임제...	2021-11-04 오전 10:16	Microsoft Excel 워...	2,935KB
6110000_서울특별시.03.06.01_P_공연장	2021-11-04 오전 10:16	Microsoft Excel 워...	391KB
6110000_서울특별시.03.06.02_P_관광공연장업	2021-11-04 오전 10:16	Microsoft Excel 워...	7KB
6110000_서울특별시.03.07.04_P_국제회의시...	2021-11-04 오전 10:16	Microsoft Excel 워...	6KB
6110000_서울특별시.03.07.08_P_유원시설업(...	2021-11-04 오전 10:16	Microsoft Excel 워...	162KB
6110000_서울특별시.03.07.11_P_전통사찰	2021-11-04 오전 10:16	Microsoft Excel 워...	32KB
6110000_서울특별시.03.07.13_P_종합류영업	2021-11-04 오전 10:16	Microsoft Excel 워...	6KB
6110000_서울특별시.03.10.01_P_비디오물업...	2021-11-04 오전 10:16	Microsoft Excel 워...	390KB
6110000_서울특별시.03.11.06_P_한옥체험업	2021-11-04 오전 10:16	Microsoft Excel 워...	89KB
6110000_서울특별시.03.12.03_P_일반여행업	2021-11-04 오전 10:16	Microsoft Excel 워...	1,793KB
6110000_서울특별시.07.22.25_P_축산물운반업	2021-11-04 오전 10:16	Microsoft Excel 워...	98KB
6110000_서울특별시.08.25.01_P_대규모점포	2021-11-04 오전 10:16	Microsoft Excel 워...	188KB
6110000_서울특별시.09.27.01_P_목재수입업...	2021-11-04 오전 10:16	Microsoft Excel 워...	94KB
6110000_서울특별시.09.28.09_P_역화석유가...	2021-11-04 오전 10:16	Microsoft Excel 워...	10KB

원본 (LOCALDATA - 서울시 업종별 데이터)

PC방 - 서울 상권 청소년게임제공업

코인노래방 - 서울 상권 노래연습장

독서실 - 서울특별시 학원 교습소정보 / 서울 상권 일반음식점 / 서울 상권 휴게음식점

데이터 가공 예시 청소년 연도별 사망원인 / 코로나19 이후 삶의 변화 (reshape2 패키지 사용)

```
# 데이터 행렬 바꾸는 패키지
library(reshape2)
reason <- read_excel("청소년 사망원인.xlsx")

# melt(데이터, id.var=) - 열을 행으로 바꾸는 명령어
# id.vars=c(기준 열)

data_melt <- melt(reason, id.vars=c('연도'),
                  variable.name="사망원인",
                  value.name='value') # 열 이름 변경
```

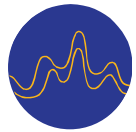


	연도	고의적 자해 (자살)	안전 사고
1	2010	926	935
2	2011	931	818
3	2012	822	731
4	2013	793	652

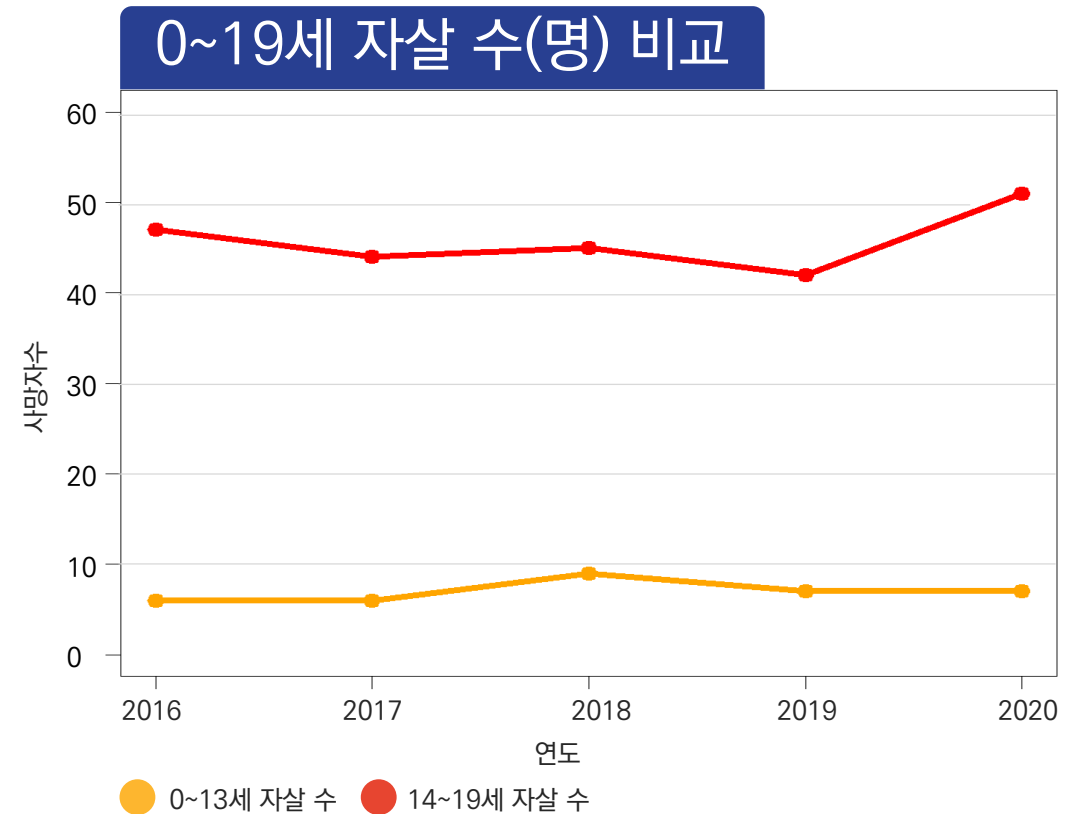
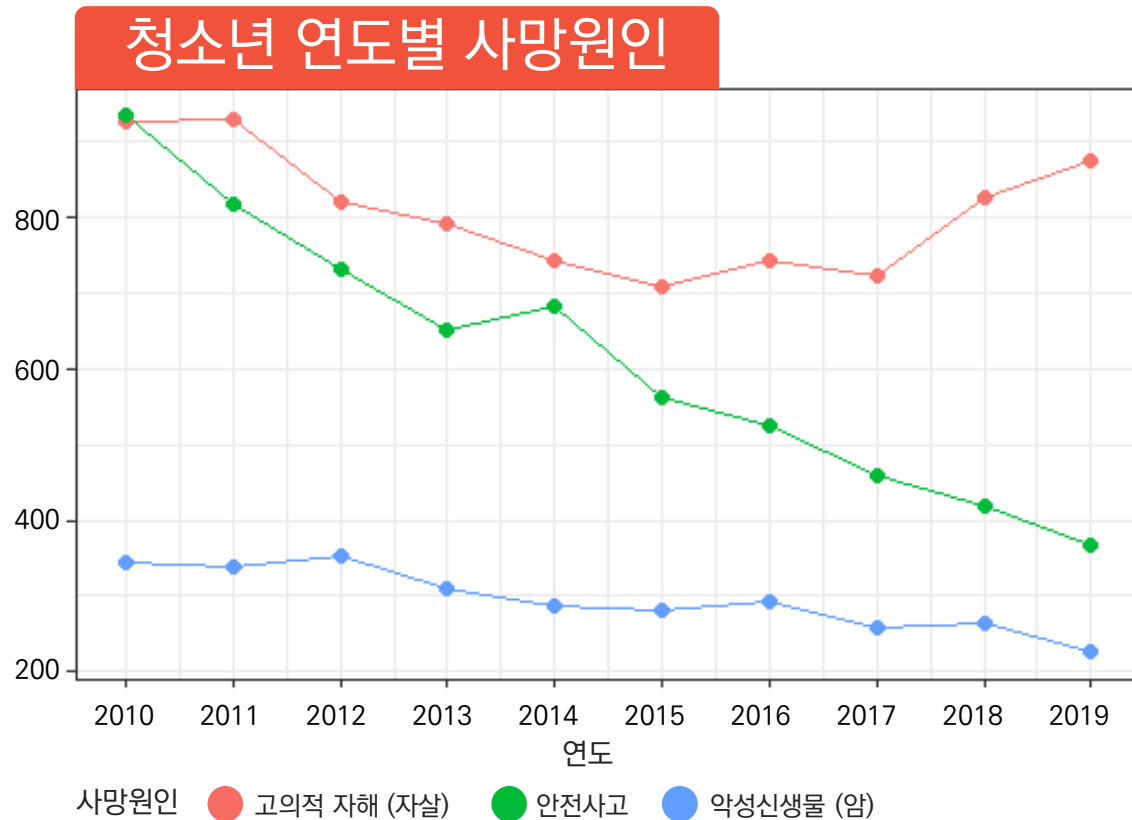


```
> str(data_melt)
'data.frame': 30 obs. of 3
 $ 연도      : num  2010 2011 2012
 $ 사망원인  : Factor w/ 3 levels
 $ value     : num  926 931 822
> head(data_melt)
  연도 사망원인 value
1 2010 고의적 자해(자살) 926
2 2011 고의적 자해(자살) 931
3 2012 고의적 자해(자살) 822
4 2013 고의적 자해(자살) 793
5 2014 고의적 자해(자살) 742
6 2015 고의적 자해(자살) 708
>
```

| 분석의 필요성 및 목적

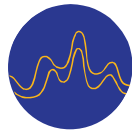


청소년들의 위험신호



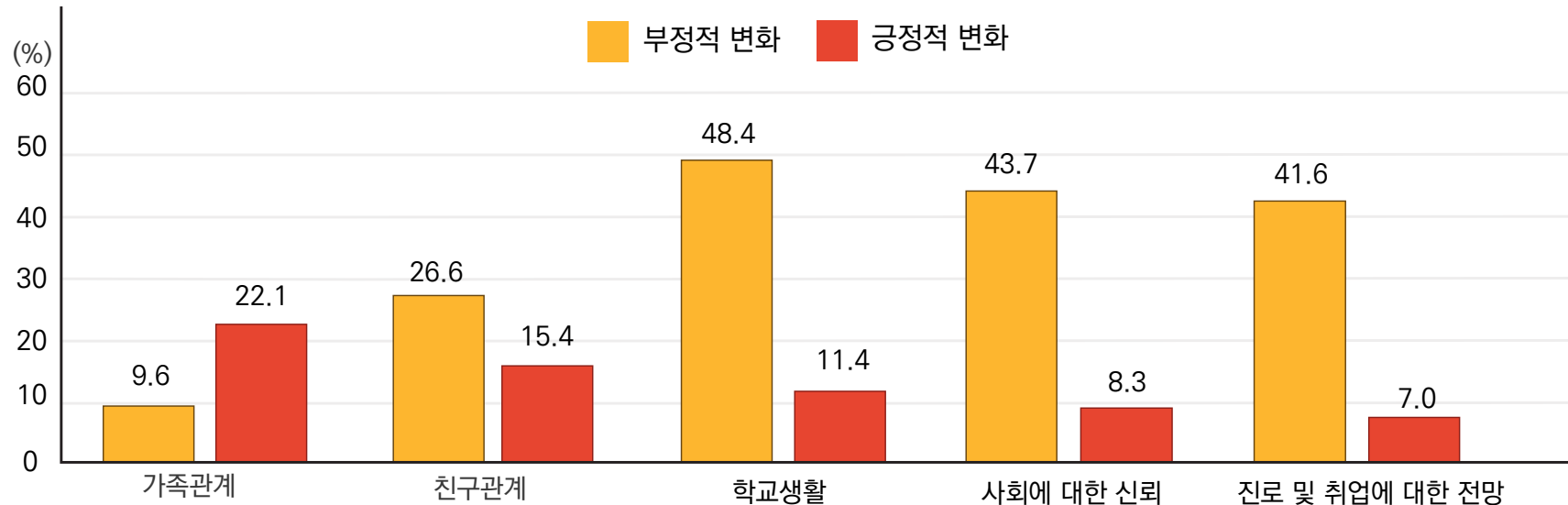
청소년의 사망 원인의 **1위 - 자살**
최근에 자살에 의한 사망자 수가 **증가**하는 추세

0 ~ 13세의 자살 명 수보다,
19세 이하의 자살의 명 수가 현저히 **높다**.



청소년들이 느끼는 코로나 19로 인한 삶의 변화

9~24세 청소년들이 느끼는 코로나19로 인한 삶의 변화 영역별 비교

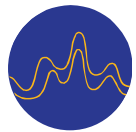


코로나19로 청소년들의 삶의 변화는 전반적으로 부정적으로 변했다가 많았다.

‘학교생활’(부정적 48.4%, 긍정적 11.4%)

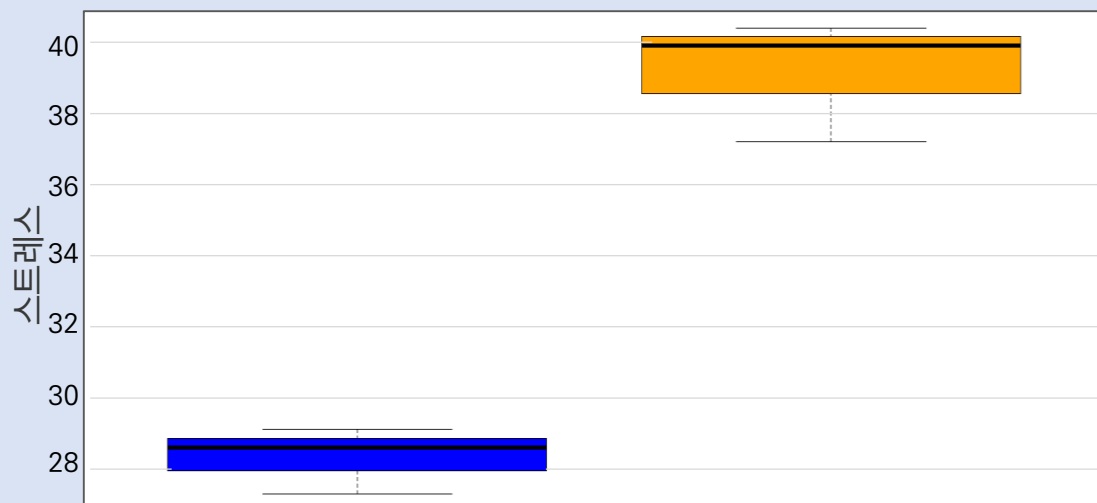
‘사회에 대한 신뢰’(부정적 43.7%, 긍정적 8.3%)

‘가족관계’(긍정적 22.1%, 부정적 9.6%)



청소년들이 진짜 성인보다 스트레스가 높은지?

청소년, 어른 스트레스 비교



성인 (19세 이상)

청소년 (중, 고등학생)

```
> t.test(data=ad, 스트레스~구별, var.equal=T)
```

Two Sample t-test

data: 스트레스 by 구별

t = -9.5921, df = 4, p-value = 0.0006602

alternative hypothesis: true difference in means between group
성인 and group 청소년 is not equal to 0

95 percent confidence interval:

-13.969064 -7.697602

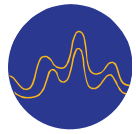
sample estimates:

mean in group 성인	mean in group 청소년
28.33333	39.16667

분석결과

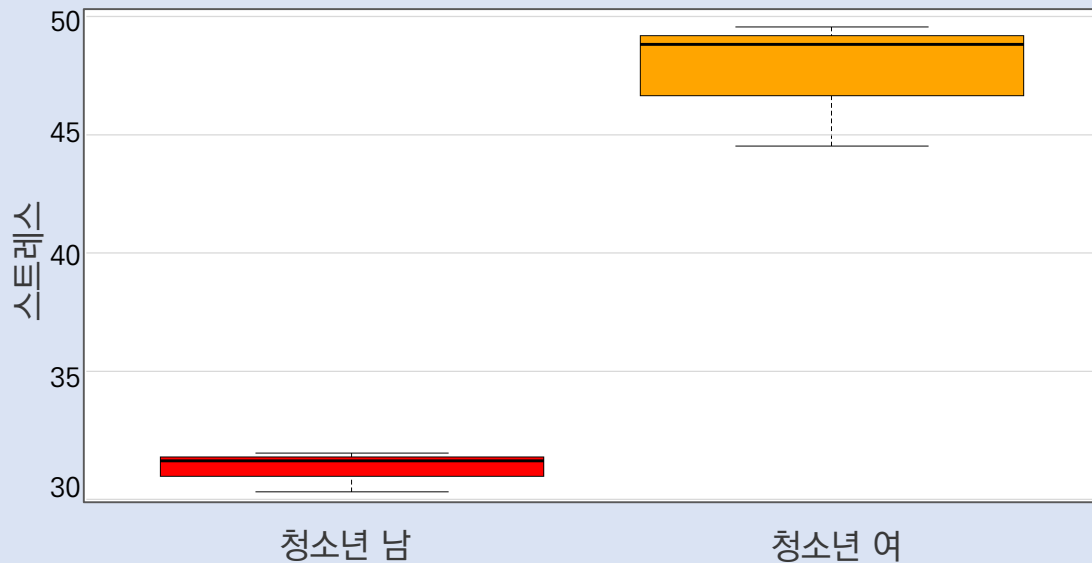
2017 ~ 2019년의 성인과 청소년의 스트레스 인지율을 T검정 실시.

- P값이 0.00066 정도로 0.05보다 작아서 두 집단의 평균이 유의미한 차이가 있다.
- 청소년의 스트레스인지율의 평균이 성인보다 더 높았다.



남학생, 여학생의 스트레스 인지율 비교

청소년 남.여 스트레스 비교



```
> t.test(data=young, 스트레스~성별, var.equal=T)
```

Two Sample t-test

data: 스트레스 by 성별

t = -10.123, df = 4, p-value = 0.000536

alternative hypothesis: true difference in means between group

남학생 and group 여학생 is not equal to 0

95 percent confidence interval:

-20.72822 -11.80512

sample estimates:

mean in group 남학생 mean in group 여학생

31.36667

47.63333

분석결과

2017 ~ 2019년의 청소년 남학생, 여학생의 스트레스 인지율을 T검정 실시.

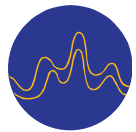
- P값이 0.000536으로 0.05보다 작아서 두 집단의 평균이 유의미한 차이가 있다.
- 여학생의 스트레스 인지율의 평균이 남학생보다 더 높았다.



정신건강 진료



아동 청소년(10세~19세) 정신진료가 점점 **증가**하고 있는 추세.
하지만 상담대기일 수가 최대 **60일까지** 오래 **기다려야** 한다.



분석방향

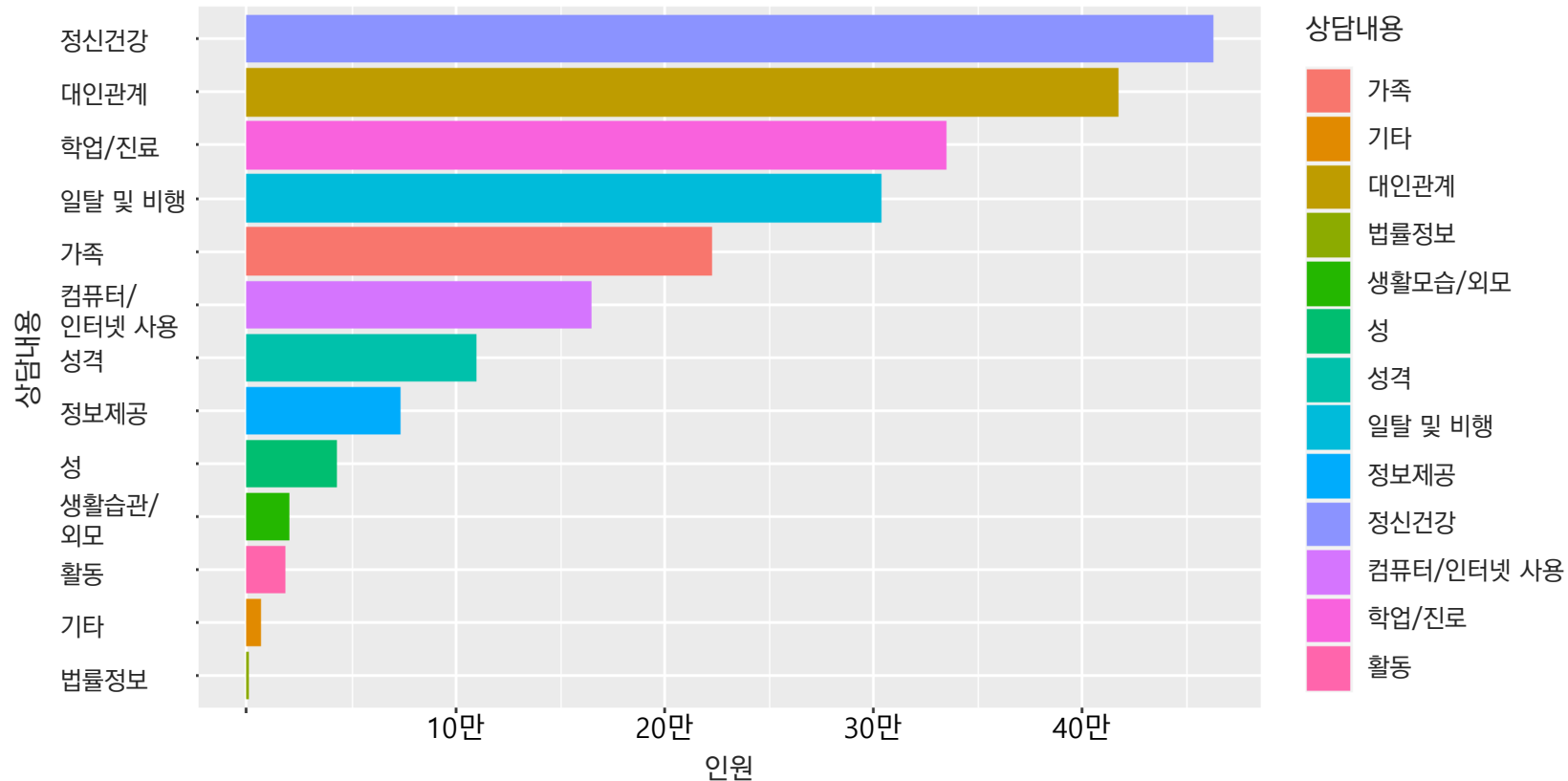
서울시만 해도 구마다 하나의 청소년상담센터가 있지만,
청소년들이 상담을 받고 싶어도 대기시간이 오래 걸리고,
상담을 받기까지 어려움이 많음을 알 수 있었음.



청소년들의 상담에 대한 인식을 낮추기 위해 팝업쉼터를 제안.
많은 청소년들이 도움을 받을 수 있도록 청소년 동선 파악 후, 팝업 쉼터의 위치선정 예정.

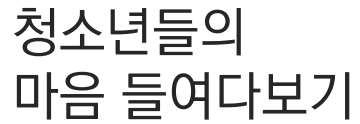
| 위치 선정

청소년 상담내용 그래프



청소년 상담 내용

- 1위 - 정신건강 (스트레스, 우울감, 자살생각 등)
- 2위 - 대인관계
- 3위 - 학업 / 진로



청소년들의 관심키워드/ 고민내용

관심키워드 (워드클라우드)

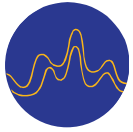


고민내용 (워드클라우드)



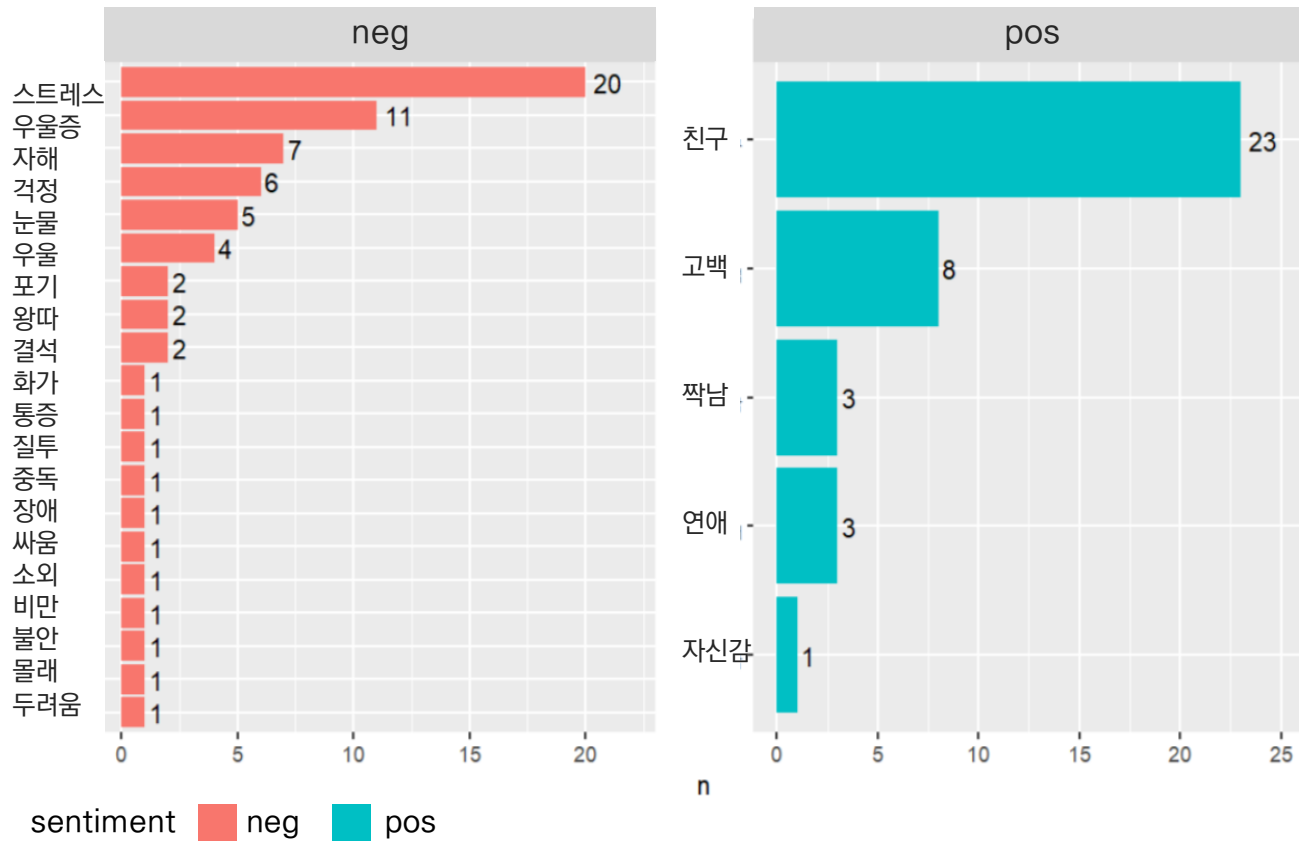
* 고민내용은 친구가 빈도수가 많아 빠고 워드클라우드 실행

2019~2020년 까지 월별 청소년들의 관심키워드 / 고민내용 1위부터 30위까지



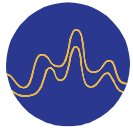
청소년들의 고민내용(감정분석)

고민내용 (감정분석)



우울증
스트레스
자해
걱정
눈물
왕따
결석
포기
소외
불안
통증
중독
장애
싸움
비만
불안
몰래
두려움

고민 내용으로 감정분석결과,
부정단어는 **20개**, 긍정단어는
고작 **5개** 나옴.
부정적인 단어 중에서 **스트레스가**
제일 많았음.



팝업 심터란?

청소년들이 조금 더 가볍게 상담을 받을 수 있도록 도와주는 팝업형식의 상담장소.
조금 더 심도 있는 상담에 대해서는 전문상담센터에 연계.

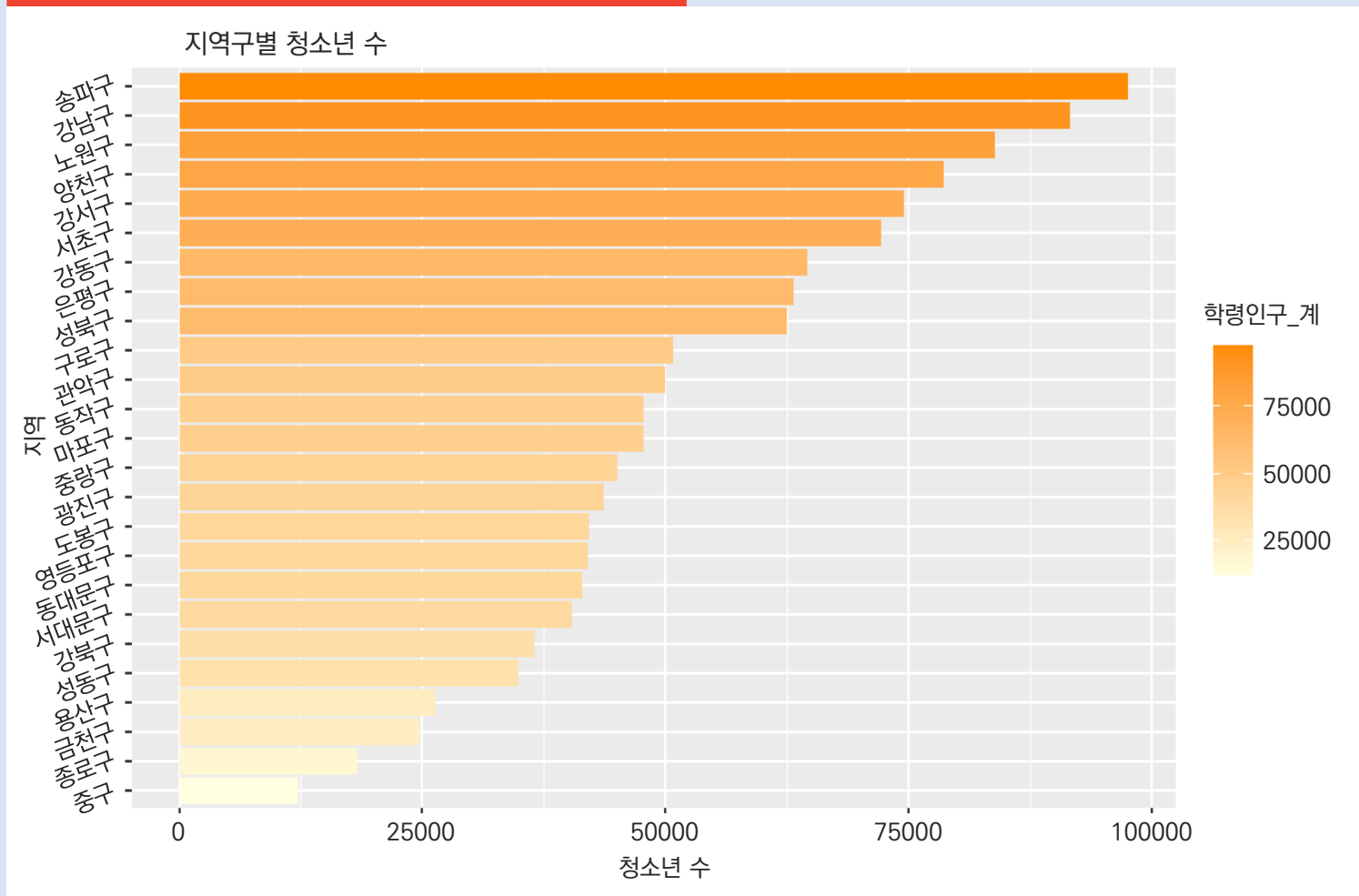
공원에 팝업심터를 여는 이유 - 여유 있는 공간 / 심리적 안정 공간(녹지)

팝업심터 참고 이미지



구별 청소년 수 파악

구별 청소년 수 그래프(막대)



분석결과

역시 학구열이 높은 구가 청소년의 수가 많았다.

상위 5개 구 - 송파구, 강남구, 노원구, 양천구, 노원구

하위 5개 구 - 중구, 종로구, 금천구, 용산구, 성동구

중간 순위 10위인 구로구를 가지고 팝업센터의 위치를 알아볼 예정.

구별 청소년 수와 PC방, 코인노래방, 독서실, 학원의 상관관계

산점도와 상관계수 관계



```
> cor(d[, -1])
```

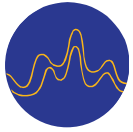
	학령인구_계	노래방수	PC방수	독서실	학원
학령인구_계	1.0000000	0.6166803	0.2803907	0.8393913	0.8754596
노래방수	0.6166803	1.0000000	0.1251447	0.6942880	0.4535770
PC방수	0.2803907	0.1251447	1.0000000	0.1811784	0.1259048
독서실	0.8393913	0.6942880	0.1811784	1.0000000	0.8596299
학원	0.8754596	0.4535770	0.1259048	0.8596299	1.0000000

학생들이 많이 있을 것 같은 곳

(코인노래방, pc방, 독서실, 학원)와 학생들 수
와의 상관관계 계산함.

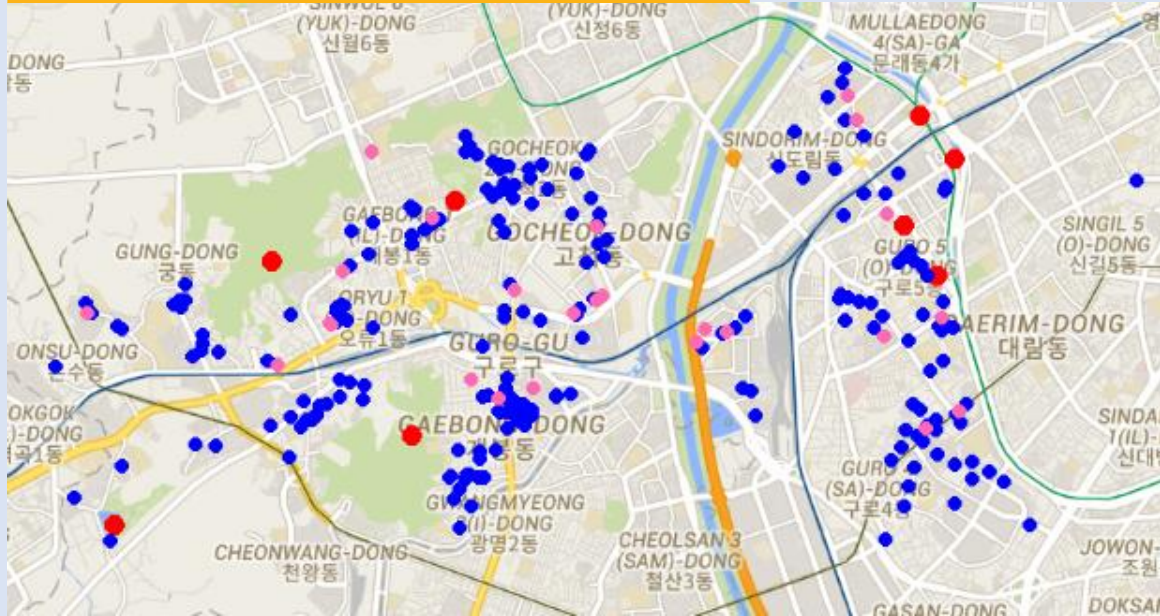
학원(0.87), 독서실(0.84), 노래방은(0.61) 학
생수와 상당한 상관관계가 있음을 볼 수 있다.

학원과 독서실의 위치를 가지고 팝업점터의 위
치를 선정할 예정.

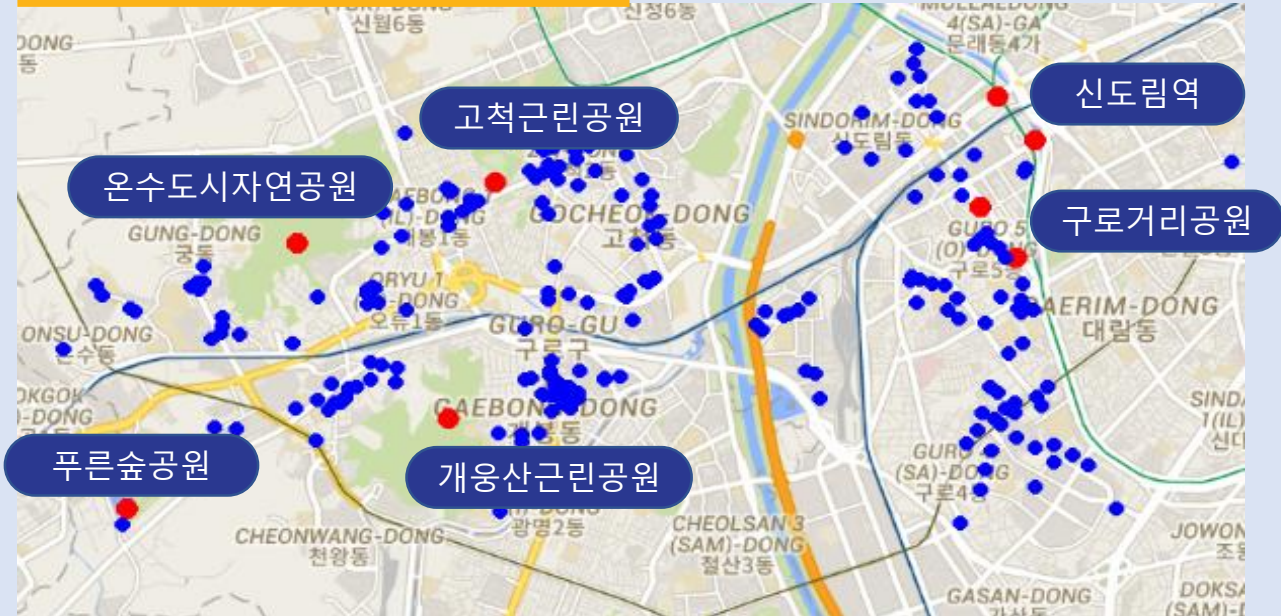


위치선정

학원, 독서실, 공원 지도위치



학원, 공원 지도위치



분석결과

학원과 독서실로 팝업 쉼터의 위치를 선정하려고 했지만, **학원과 독서실이 별로 차이가 없었음.**
그래서 **학원만 가지고** 공원과의 거리계산을 통해 팝업쉼터의 위치를 선정할 예정.

위치선정 – 공원과의 거리계산

공원과 학원과의 근거리계산 결과

구로거리공원, 개웅산근린공원, 고척근린공원이 팝업선택에 적합한 위치로 나옴.

```
## 거리계산
install.packages("geosphere")
library(geosphere)

# 거리계산
mat <- distm(edu_add[,c('lon', 'lat')], gong[,c('lon', 'lat')], fun = distVincentyEllipsoid)

# max.col(-mat) - which.min(mat)
edu_add$locality <- gong$공원명[max.col(-mat)]

# 최소거리 추출 / apply 반복
edu_add$near_dist <- apply(mat, 1, min)

table(edu_add$locality)
```

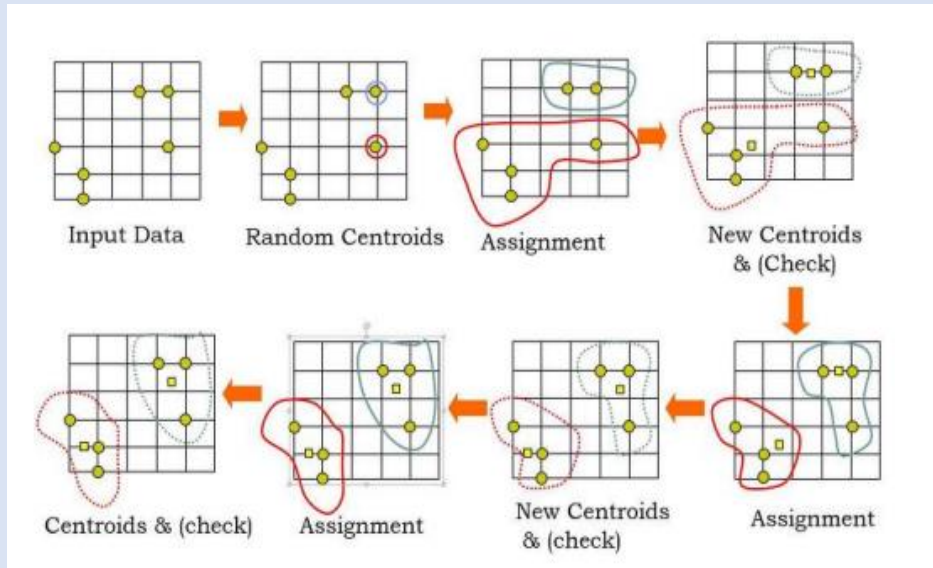
	V1	V2	V3	V4	V5	V6	V7
공원명	개웅산근린 공원	고척근린공 원	구로거리공 원	신도림역남 측	신도림역북 측	온수도시자 연공원	푸른수목원
학원수	88	81	162	4	23	28	10

	개웅산	고척	온수	푸른	신도림역	구로거리	
	V1	V2	V3	V4	V5	V6	V7
1	3491.6077	3691.5219	4739.6310	5673.564	2511.2902	2236.1588	1710.8069
2	831.5549	1836.4367	1041.5293	1640.170	4950.5713	5061.0237	4562.5599
3	3567.5938	3713.4485	4788.0411	5761.437	2411.2205	2127.8029	1613.4100
4	2489.1364	2557.0164	3618.0618	4736.788	2344.3258	2240.5182	1639.4498
5	2159.7379	983.8233	2373.3519	4202.802	2473.8288	2666.0769	2293.7563
6	3969.2135	3964.2949	5113.4190	6186.693	2175.4982	1844.6348	1417.2373
7	2477.8791	2235.7098	3412.4950	4740.144	2046.3718	2014.1664	1440.9646
8	3551.8946	2517.3826	3927.1310	5716.134	923.8837	1183.8163	1049.8911
9	3261.2258	2879.6362	4142.1987	5523.791	1453.4734	1318.1036	716.0012
10	1578.3360	184.5634	1265.1987	3242.991	3581.0754	3776.5606	3385.0042
11	1318.1842	901.9126	575.4726	2550.730	4292.4793	4470.3893	4048.2102
12	2109.1980	1053.8178	2417.8661	4192.659	2442.1118	2616.4513	2221.4442
13	572.3311	2024.5609	2178.2126	2671.260	4156.2478	4170.3820	3605.1371

k-평균 군집 (k-means clustering)

머신러닝의 비지도학습

- 군집의 수만큼(k개) 초기값을 지정
 - 각 개체를 가까운 초기값으로 군집을 형성
 - 각 군집의 평균을 재계산하여 초기값을 갱신
 - 위의 과정을 반복하여 k개의 최종 군집을 형성
- 단, 평균을 사용하기 때문에 이상값에 영향을 받을 수 있음



군집코드

```
# k-means 비지도학습
data <- edu_add[,c('lon', 'lat')]

# 군집개수 3개
fit <- kmeans(x=data, centers = 3)
fit

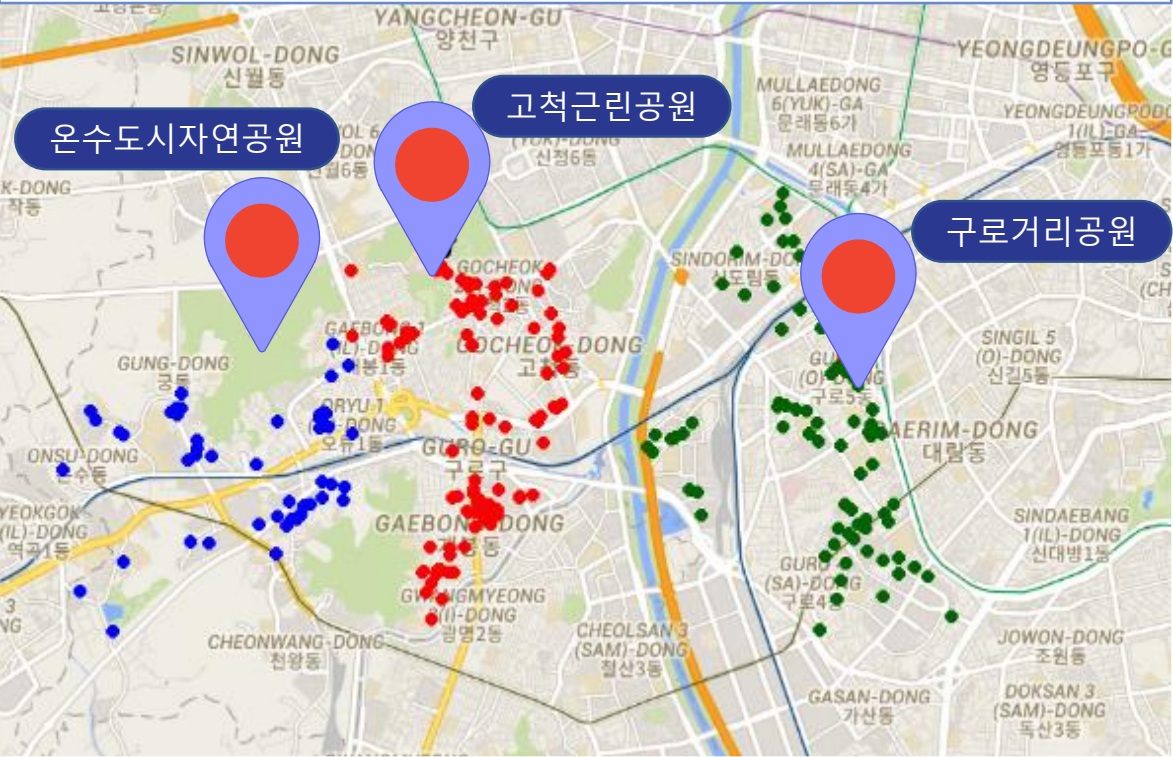
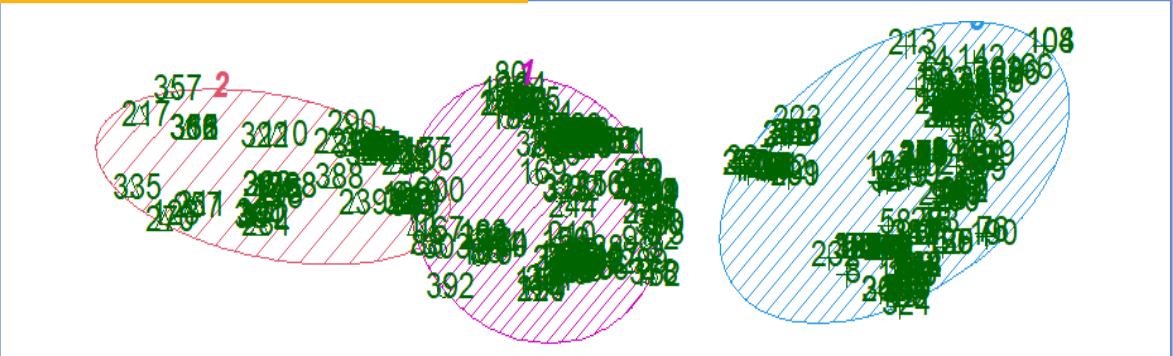
# 이상값 제거
data <- data[-c(98,284),]

fit$cluster           # 군집번호
fit$centers           # 군집 중심점 좌표

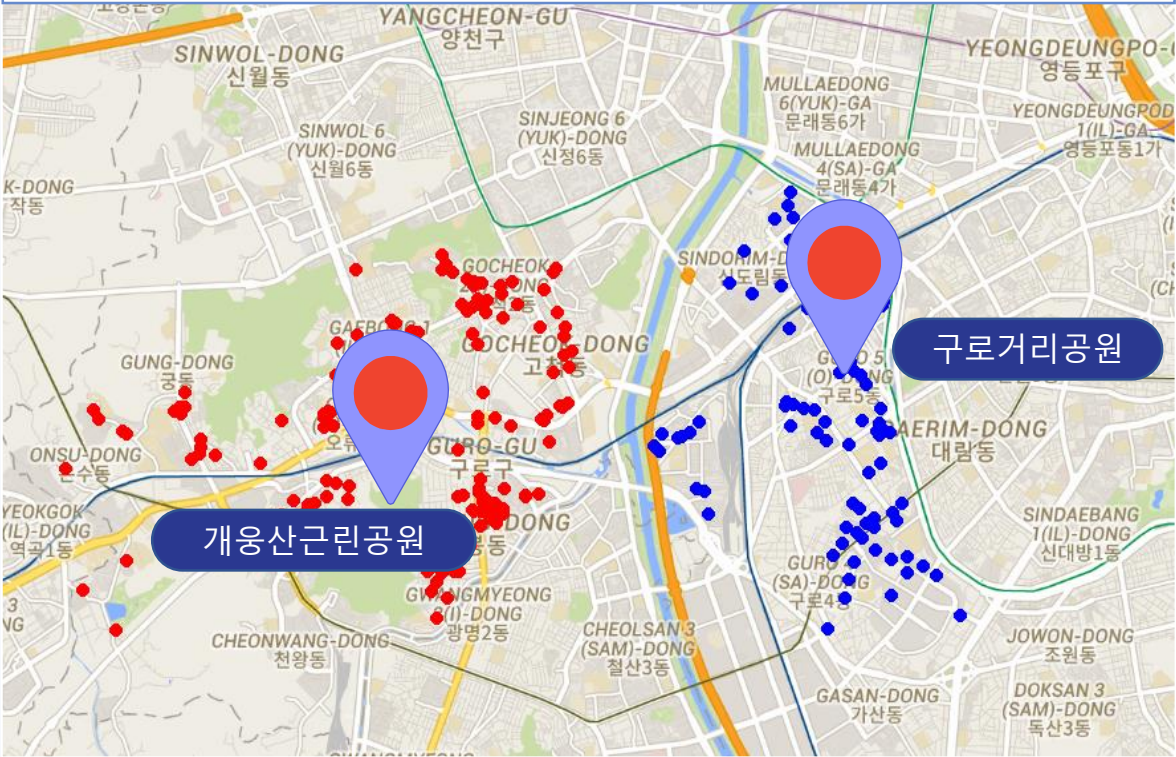
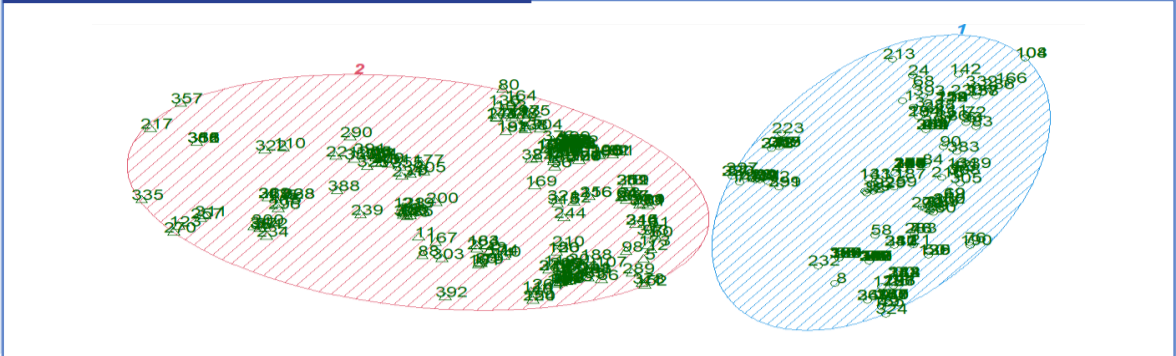
# 군집 시각화
clusplot(data, fit$cluster, color = TRUE,
          shade = TRUE, labels = 2, lines = 0)
```


위치선정 – K-means 사용 | 학원 위치로 군집을 나눈 후, 군집의 중심과 가까운 공원 도출

3그룹으로 군집하기



2그룹으로 군집하기



위치선정 – K-means 결과

	V1	V2	V3	V4	V5	V6	V7
공원명	개웅산근린공원	고척근린공원	구로거리공원	신도림역남측	신도림역북측	온수도시자연공원	푸른수목원
학원수	88	81	162	4	23	28	10

K-means로 3개의 군집

```
> won1 <- gong$공원명[max.col(-mat3)]  
> won1
```

```
[1] "온수도시자연공원" "고척근린공원" "구로거리공원"
```

온수도시자연공원 -> 주위 학원이 28개이기 때문에 청소년들의 접근성이 낮다는 이유로 3개의 위치 선정 하지 않기로 판단.

K-means로 2개의 군집

```
> won <- gong$공원명[max.col(-mat2)]  
> won
```

```
[1] "구로거리공원" "개웅산근린공원"
```

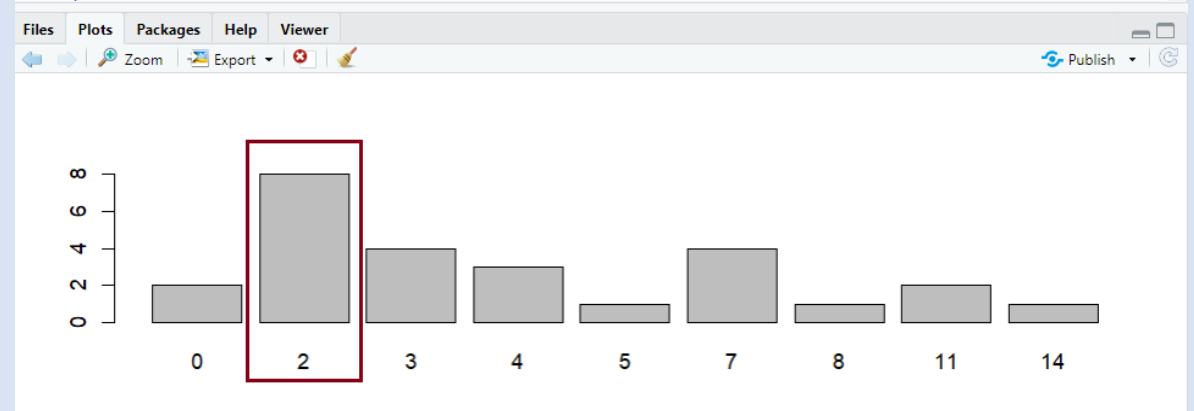
구로거리공원과 개웅산근린공원 도출.
주위 학원 수가 많은 공원들이 결과로 나와 위의 공원들을 팝업센터 위치로 결정.

K-means 최적의 군집 수(코드)

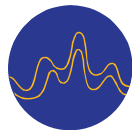
NbClust패키지를 사용

k-means의 최적의 군집 수를 계산 -> 2개의 군집

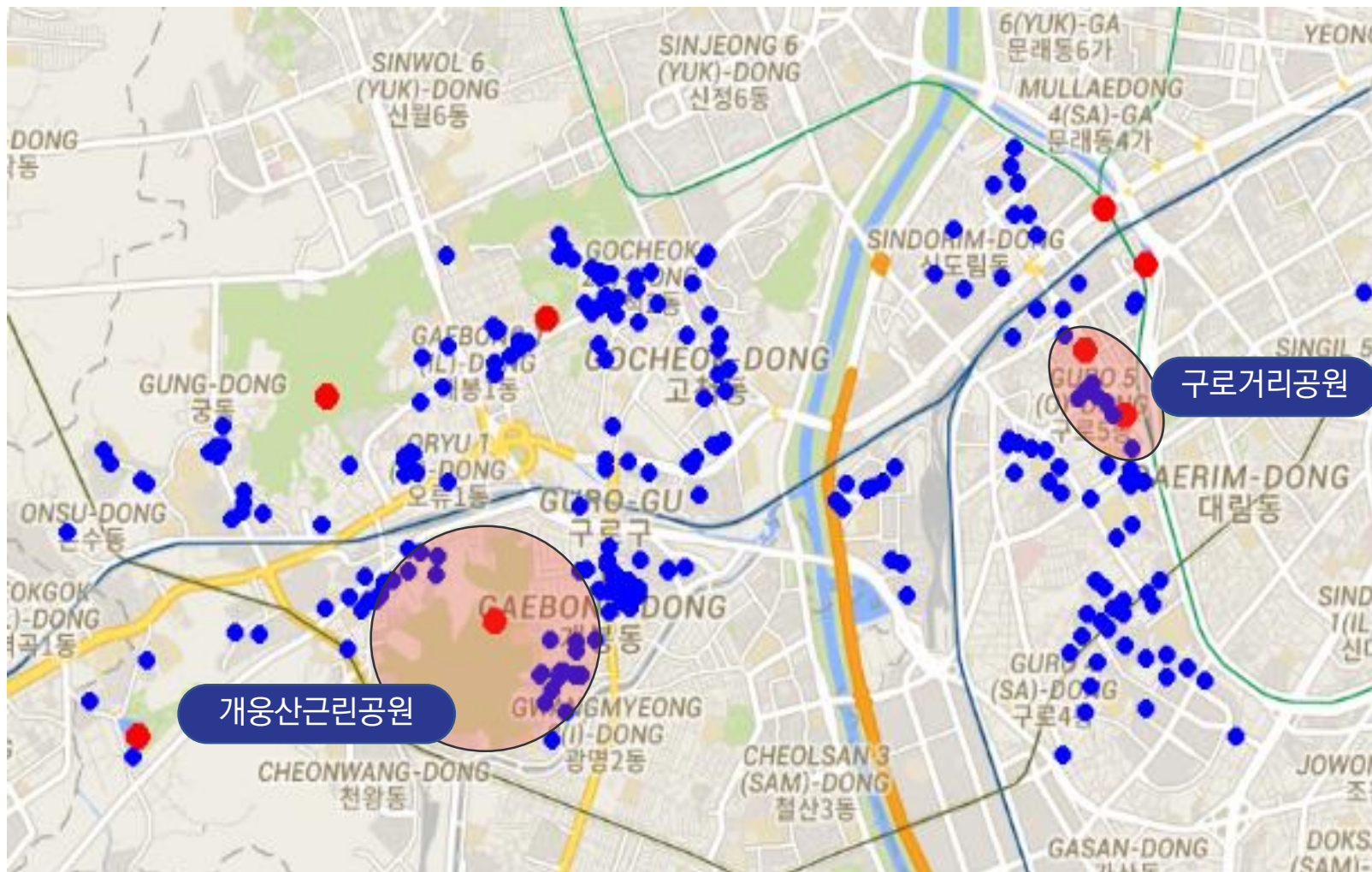
```
> nc <- NbClust(data1, min.nc = 2, max.nc = 15, method = "kmeans")  
*** : The Hubert index is a graphical method of determining the number of clusters.  
      In the plot of Hubert index, we seek a significant knee that corresponds to a  
      significant increase of the value of the measure i.e the significant peak in Hubert  
      index second differences plot.  
  
*** : The D index is a graphical method of determining the number of clusters.  
      In the plot of D index, we seek a significant knee (the significant peak in Dindex  
      second differences plot) that corresponds to a significant increase of the value of  
      the measure.  
  
*****  
* Among all indices:  
* 8 proposed 2 as the best number of clusters  
* 4 proposed 3 as the best number of clusters  
* 3 proposed 4 as the best number of clusters  
* 1 proposed 5 as the best number of clusters  
* 4 proposed 7 as the best number of clusters  
* 1 proposed 8 as the best number of clusters  
* 2 proposed 11 as the best number of clusters  
* 1 proposed 14 as the best number of clusters  
  
***** Conclusion *****  
  
* According to the majority rule, the best number of clusters is 2  
  
*****  
> par(mfrow=c(1,1))  
> barplot(table(nc$Best.n[1,]))
```



| 분석결과 및 향후 과제



청소년을 위한 팝업선정 위치



위치선정 결과

구로구
팝업센터의 최종 위치

구로거리공원,
개웅산근린공원

팝업쉼터의 내용 제안

기관과 연계하여 팝업쉼터의 프로그램 만들 예정

한국아동청소년
심리상담센터



사이버1388
청소년상담센터



청소년이동쉼터
여우별



대부분의 청소년 프로그램이 가출 청소년 위주인데
학업 스트레스나 대인관계에 힘들어 하는 청소년들을 위한 프로그램이 부족함.
또한, 인지하지 못한 청소년들도 도움을 받을 수 있는 프로그램 개설.



10대 부터 상담을 통해 스트레스, 우울감 등을 줄여가는 방법을 경험한다면 성인이 되어서도 도움을 청할 수 있을 것이다.

향후 과제

여학생들의 스트레스인지율이 높은 원인과 대책방안을 생각해볼 예정.

팝업심터를 공원으로 잡았는데 다른 기준으로 입지선정을 할 수 있을지
다른 위치선정 방법 연구

데이터 목록, 분석툴

데이터 목록

출처	자료명	활용
통계청	사망원인 통계	청소년 사망원인 통계 (선그래프)
	국민건강영양조사:스트레스인지율 추이	성인과 청소년의 스트레스 인지율 (T검정/ 상자그림)
	청소년건강행태조사:스트레스 인지율	청소년의 남학생 여학생의 스트레스 인지율 (T검정/상자그림)
청소년데이터플랫폼	여성가족부_청소년상담지원현황	청소년 상담내용 그래프(막대)
	청소년 고민 순위 정보	청소년 고민내용(워드클라우드/감정분석)
	청소년 관심키워드 분석 데이터	청소년 관심 키워드 (워드클라우드)
여성가족부	청소년 종합실태 조사	코로나19 이후 삶의 변화, 연도별 청소년 사망원인
서울열린데이터광장	서울시 청소년인구 통계	구별 청소년 수 그래프(막대)
	서울시 주요 공원현황	구로구 공원과의 거리계산
	서울특별시 학원 교습소정보	구별 청소년 수와 학원과의 상관관계
서울 정보 포털 광장	자치구 청소년상담복지센터 인력 소요 조사	청소년상담복지센터 상담 인력 현황 및 대기일수 (막대)
공공데이터포털	보건복지부_서울시 연도별 자살자 수	청소년과 유소년의 자살율 그래프(선)
LOCALDATA	서울시 업종별 데이터	구별 청소년 수와 PC방, 코인노래방, 독서실의 상관관계
Do it! 쉽게 배우는 R 텍스트 마이닝	knu_sentiment_lexicon	감정분석을 위한 감정사전(군산대학교)

분석툴



출처 사이트

<http://www.kccp.kr/>

<https://www.cyber1388.kr:447/>

<http://www.chungnamilbo.com/news/articleView.html?idxno=468051>

<http://www.thesegeye.com/news/newsview.php?ncode=1065587798323497>



THANK YOU