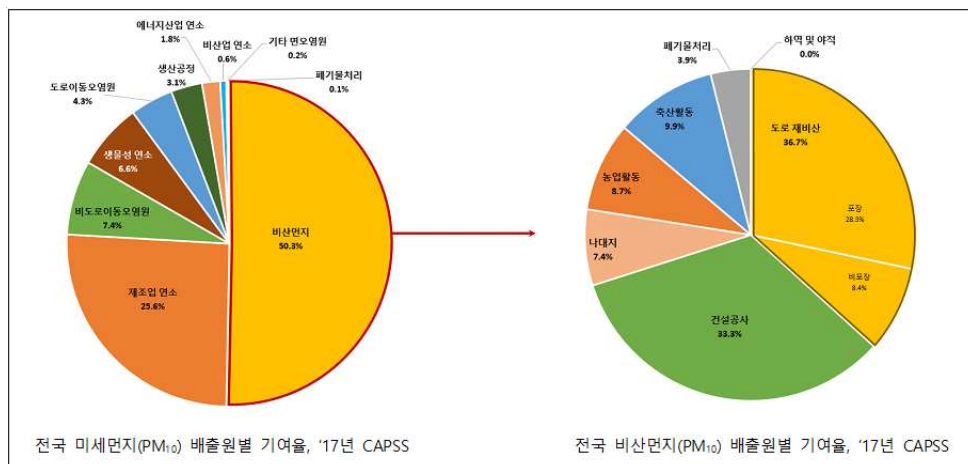


[도로 재비산먼지와 버스정류장 데이터를 활용한 셸터형 버스정류장 입지선정] 분석결과 보고서

□ 분석 개요

- (배경) 미세먼지의 총 발생량의 가장 큰 부분을 비산 먼지가 차지하며, 총 비산먼지 중 재비산먼지 기여도가 가장 높으므로 재비산먼지 관리 는 효과적인 미세먼지 관리에 반드시 필요함. 또한, 교통량이 많고 인 구밀도가 높은 곳이 재비산먼지 위험도가 높으며, 차량 이용자 보다는 도로에 직접 노출되는 사람에게 더 위험도가 높음



- (목표) 재비산먼지 위험도가 높을 것으로 예상되는 도로와 이용인구가 많아 더 큰 효과를 가져올 수 있는 버스정류장에 데이터를 분석하여 재비산먼지로부터 시민을 보호할 수 있는 '스마트셸터' 또는 '친환경 버스정류장'에 우선적인 설치 및 입지 선정 제언
- (데이터) 도로 재비산먼지 측정정보, 서울시 버스정류장별 승하차 인원, 서울시 버스정류장 위·경도 위치 데이터, 서울시 도로 데이터

- ▶ 도로 재비산먼지 측정정보
 - 한국환경공단(환경데이터포털)
- ▶ 서울시 버스정류장별 승하차 인원 데이터
 - 서울 열린데이터 광장
- ▶ 서울시 버스정류장 위·경도 위치 데이터
 - 서울 열린데이터 광장
- ▶ 서울시 도로(도로명) shp 데이터
 - 국토교통부

□ 분석 내용

- ① **(분석항목1)** 서울시 버스정류장별 승하차 인원 분석
 - 서울시 버스정류장별 승하차 데이터를 정제한 후 EDA와 통계 분석을 통해 이용 인원이 많은 버스정류장을 선별
 - ※ EDA(Exploratory Data Analysis) : 데이터를 다양한 각도에서 관찰하고 분석하여 자료를 직관적으로 이해하는 과정
- ② **(분석항목2)** 항목1에서 정리된 버스정류장과 서울시 버스정류장 위·경도 위치 데이터를 결합하여 선별된 버스정류장에 위치 데이터를 정비
- ③ **(분석항목3)** 도로 재비산면지 데이터 분석
 - 서울시 도로 재비산면지 1년 데이터를 분석하고 결합하여 필요한 통계적 데이터 추출
 - ※ 재비산면지는 온도와 습도의 영향이 크므로 4계절 데이터를 모두 활용
- ④ **(분석항목4)** 서울시 전체 도로(도로명 주소) 데이터와 재비산면지 위험도가 높은 도로 데이터를 결합하여 서울시 도로 재비산면지 위험도로 데이터를 만든 후 SHP 시각화
- ⑤ **(분석항목5)** 항목2에서 나온 선별된 최종 버스정류장 데이터와 항목4에서 정리된 서울시 도로 재비산면지 위험도로 데이터를 결합하여 최종 도로 재비산면지 위험도 높은 버스정류장을 선정

□ 전처리 방법

- 서울시 버스정류장별 승하차 인원 데이터
 - 일별 시간대별 데이터를 활용하여 버스정류장별 1년간 총 승하차 인원 산출 (계절별 날씨별 승하차 인원이 상이할 수 있으므로 1년 평균 데이터를 활용하기 위함)
 - ※ R-studio 활용
- CSV 수치 데이터를 SHP 지도 데이터로 처리
 - 버스정류장 위·경도 위치 데이터, 서울시 도로(도로명) 데이터 (QGIS 분석을 위한 전처리)
 - ※ QGIS : 지리정보체계를 활용하여 데이터 시각화, 분석을 제공하는 툴



□ 분석 방법

○ EDA와 통계적 분석

- ① 버스정류장별 승하차 인원(연간) 데이터를 EDA와 통계적 분석을 통해 후보 버스정류장 선별

기준 승객수	$\geq 5,000,000$ (오백만 명)	$\geq 4,000,000$ (사백만명)	$\geq 3,000,000$ (삼백만명)	$\geq 1,000,000$ (일백만명)
정류장 수	5 (개)	11 (개)	25 (개)	500 (개)
승객 수	26,993,522 (명)	53,558,515 (명)	101,518,297 (명)	810,143,844 (명)
전체 대비 승객수(%)	0.96 %	1.91 %	3.62 %	28.9 %

- 연간 이용 승객 수 1,000,000(일백만)명 이상의 정류장 선별(500개 정류장), 이는 서울시 전체 버스정류장 승객 수에 약 30%에 해당함.

- ② 도로 재비산먼지(일별) 데이터를 종합하여 연간 평균 도로 재비산먼지 농도 데이터를 만들고 이에 대한 통계적 분석

- 도로 재비산먼지 분류 기준(한국환경공단)에 따른 도로 재비산먼지 위험 도로(도로명 분류)를 선별 → 2,620개 도로 중 32개 도로 선별

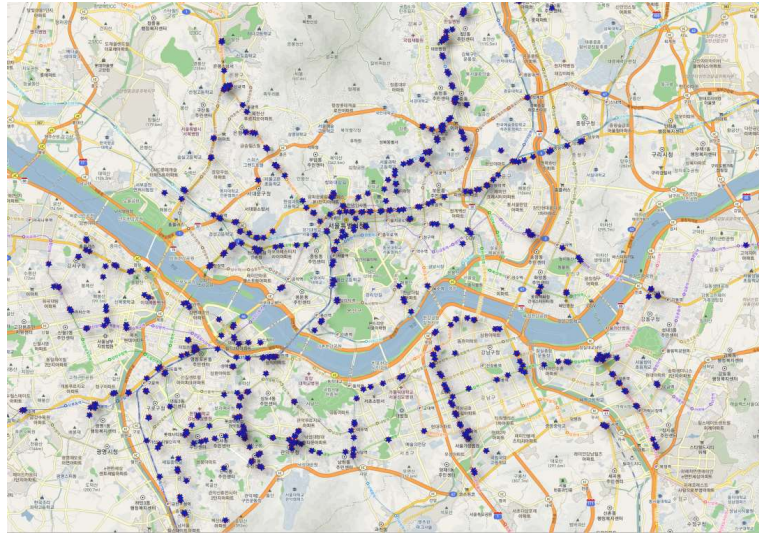
※ 도로 재비산먼지 보통 기준의 농도 100 이상의 도로

도로 재비산먼지 등급 ($\mu\text{g}/\text{m}^3$)				
매우좋음	좋음	보통	나쁨	매우나쁨
0~50	51~100	101~150	151~200	201~

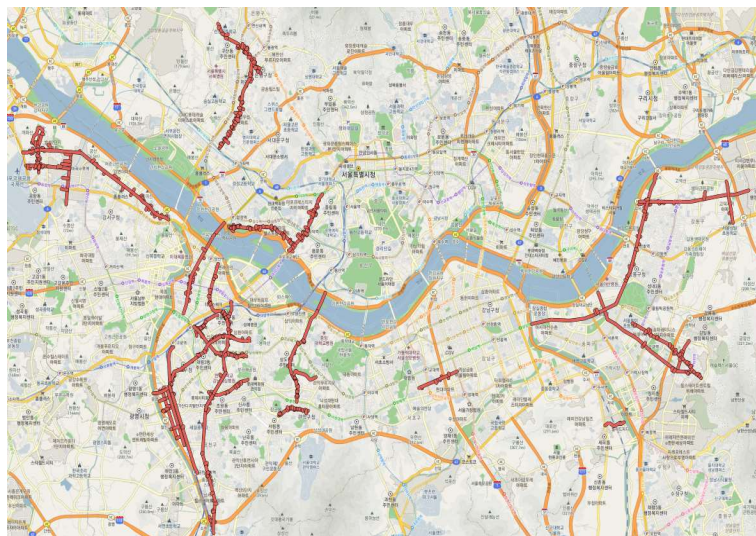
○ QGIS 분석

- ① 위에서 선별된 500개의 정류장 데이터와 버스정류장 위·경도 위치 데이터와 결합하여 SHP 파일로 변환

- 버스정류장 ARS-ID를 기준으로 INNER_JOIN 후 QGIS에서 선별된 정류장 위치 표시



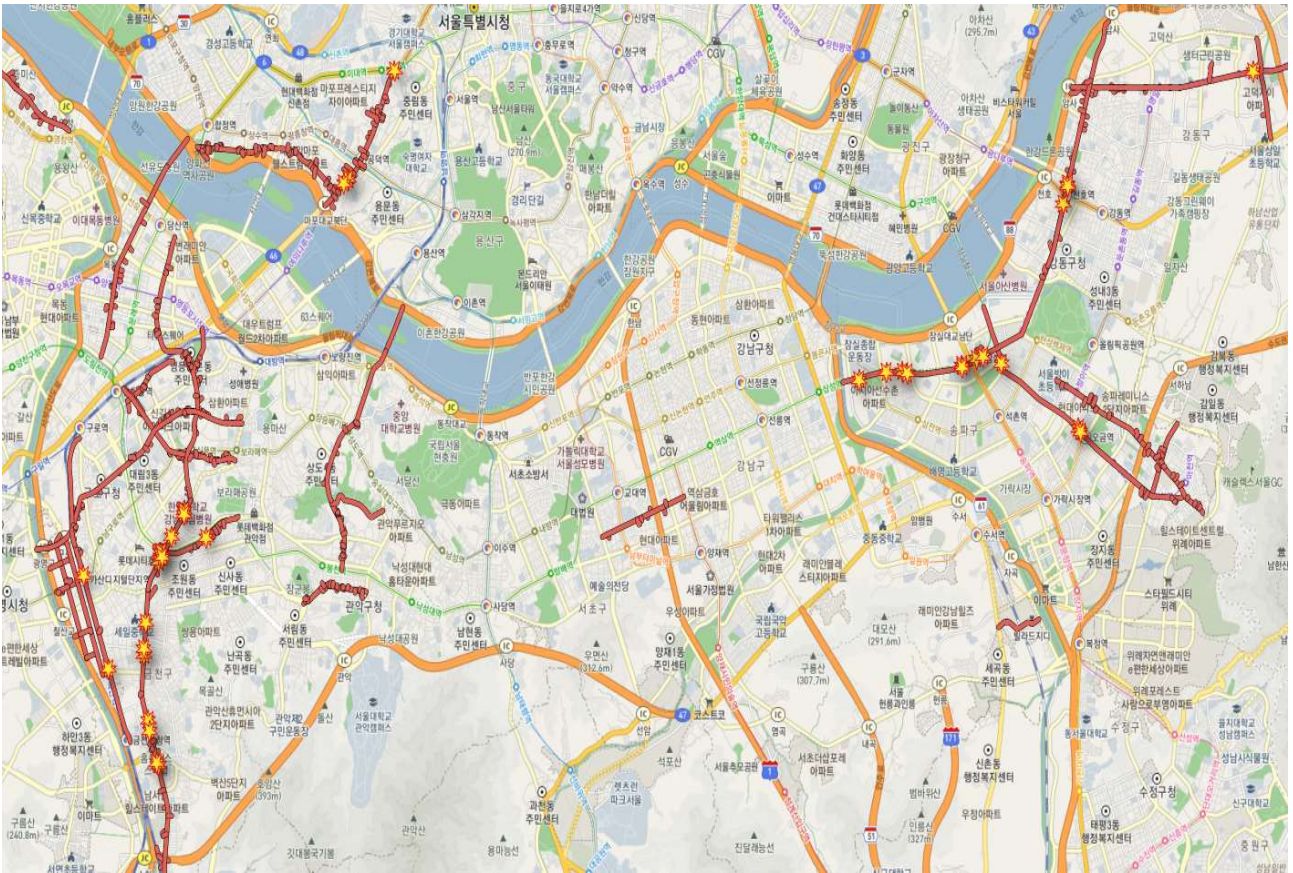
- ② 도로 재비산먼지 위험 지역으로 선정된 도로 데이터를 shp 파일 변환 뒤 시각화



- ③ 인구 수로 선별된 버스정류장 데이터와 위험지역으로 선정된 도로 데이터를 결합하여 최종 도로 재비산먼지 위험 버스정류장 데이터를 산출하고 이를 shp 파일 변화 후 맵핑

□ 분석 결과(요약)

- 도로 재비산먼지 위험에 노출되어 있으며 연간 승하차 인원수 일백만 명 이상인 버스 정류장을 최종 선정하고 이를 맵핑
 - ⇒ 32개 도로, 35개 버스 정류장



○ 최종 선정된 개선 필요 버스정류장 목록은 첨부.

□ 정책 제언

○ 현재 미세먼지(재비산먼지)로부터 시민들의 건강을 보호하기 위해 각 지자체에서는 쉼터 형태의 버스정류장을 설치할 진행하려 하지만 예산 문제로 실제 설치, 시행되기엔 어려움을 겪고 있음

⇒ 도로 재비산먼지에 노출이 많고 승하차 인원이 많은 정류장을 선제적으로 개선하여 더 효과적이고 빠른 행정에 도움

○ 쉼터형 개선된 버스정류장 설치로 실제 미세먼지에 가장 큰 부분을 차지하며 도로에서는 더욱 해로운 재비산먼지에 대하여 시민들의 알 권리를 위한 홍보 역할 기대

○ 서울시의 경우 인구가 가장 많아 예산 대비 효과가 클 것으로 예상하여 이를 대상으로 정함. 하지만 도로 재비산먼지 농도의 경우 도로 상태가 나쁠 수 있는 경기도와 같은 서울 이외 지역이 더 높으므로 향후 같은 분석 형태를 활용할 경우 효과가 높을 것으로 예상됨.

□ 첨부 목록

- R-studio 분석코드
- QGIS 분석 화면
- 최종 선정된 35개 버스정류장 목록 (서울시)

첨부1

R-Studio 분석 코드

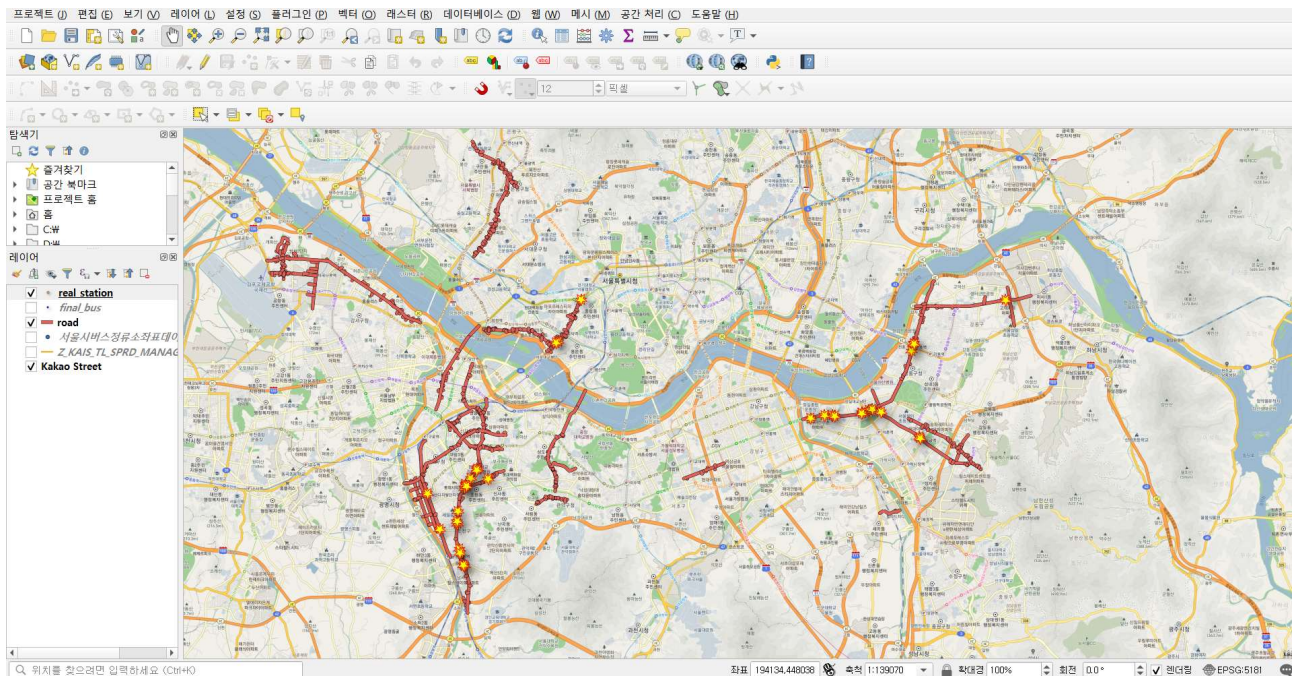
```
19 # EDA
20 pp <- bus_pop %>% group_by(버스정류장ARS번호) %>% summarise(sum(total))
21 pp <- as.data.frame(pp)
22
23 head(pp)
24 names(pp) <- c("station", "pop")
25
26 pp %>% arrange(-pop)
27
28 quantile(pp$pop)
29 boxplot(pp$pop)
30
31 pp[2,2] <- 5452887
32 head(pp)
33
34 pp %>% filter(pop > 5000000) %>% sum()
35 pp %>% filter(pop > 4000000) %>% sum()
36 pp %>% filter(pop > 3000000) %>% sum()
37 b <- pp %>% filter(pop > 1000000) %>% sum()
38 pp %>% filter(pop > 0 & pop <= 3000000) %>% sum()
39 c <- pp %>% filter(pop > 1000000)
40 a <- sum(pp$pop)
```

```
47 # 정류장 위치 정보
48 bus_location <- read.csv("d:/rwork/bus_location.csv")
49 head(bus_location)
50 bus_location <- bus_location[-1]
51
52 head(bus_location)
53
54 final_bus <- inner_join(c, bus_location, by = c('station' = 'ARS.ID'))
55
56 write.csv(final_bus, "final_bus.csv")
```

```
188 y <- str_sub(m12$도로명, -5)
189 m12$도로명 <- str_replace(m12$도로명, y, "")
190 m12$도로명 <- gsub("[()]", "", m12$도로명)
191 head(m12)
192
193 road_dust <- rbind(m1, m2, m3, m4, m5, m6, m7, m8, m9, m10, m11, m12)
194 road_dust
195
196 brt <- road_dust %>% group_by(도로명) %>% summarise(mean = mean(농도)) %>% arrange(-mean)
197
198 road_dust$농도 <- as.numeric(road_dust$농도)
199 str(road_dust)
200
201 brt <- as.data.frame(brt)
202 head(brt)
203 top_brt <- brt %>% filter(mean >= 100)
204 top_brt
205
206 ### 서울시 도로 데이터
207 seoul_road <- read.csv("d:/rwork/seoul_road.csv")
208 head(seoul_road)
209
210 select <- inner_join(seoul_road, top_brt, by = c("RN" = "도로명"))
211 prp <- as.data.frame(unique(select$RN))
212
213 write.csv(select, "d:/select.csv")
214
```


첨부2

QGIS 최종 분석 화면



첨부3

최종 선정된 버스정류장 목록

station	pop	정류소명
17014	5688625	구로디지털단지역환승센터
21001	5206011	구로디지털단지역
24146	3118849	잠실역.롯데월드
17013	2933632	구로디지털단지역
24138	2618050	잠실역.롯데월드
17015	2366650	구로디지털단지역환승센터
21225	2276858	신대방역
24132	2228247	잠실역8번출구
18003	2018932	금천우체국
14001	1944074	마포역
24134	1821318	잠실역1번.11번출구
18004	1807404	금천우체국
24158	1758428	잠실새내역4번출구
24154	1633832	잠실종합운동장
13044	1575330	아현역
19010	1555746	강남성심병원.대림성모병원
25162	1441326	천호역현대백화점.노동권익센터
14002	1407955	마포역
18009	1309308	시흥사거리
18002	1276445	문성초등학교
18010	1244989	시흥사거리
25161	1243523	천호역현대백화점.노동권익센터
18007	1222650	금천구청.금천경찰서
24249	1222405	오금동현대아파트
20002	1208889	신대방경남아파트
24101	1152601	천호역.풍납시장
18503	1148338	가산디지털단지역
24369	1118497	송파경찰서.오금역
25131	1066890	상일동역4번출구.센트럴푸르지오
24142	1062182	잠실엘스아파트앞
24157	1058571	종합운동장사거리
24215	1056381	송파구청.방이맛골
18008	1033190	금천구청.금천경찰서
19009	1024011	서울영림초등학교
18511	1004196	독산역

※ station : 정류장 ARS_ID 코드

pop : 연간 이용객 수