

지역분산분석

```
load("../../refinedata/analysis/analysis_total_Fixed.rda")  
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.6.3
```

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##   filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
library(multcomp)
```

```
## Warning: package 'multcomp' was built under R version 3.6.3
```

```
## Loading required package: mvtnorm
```

```
## Loading required package: survival
```

```
## Warning: package 'survival' was built under R version 3.6.3
```

```
## Loading required package: TH.data
```

```
## Warning: package 'TH.data' was built under R version 3.6.3
```

```
## Loading required package: MASS
```

```
##  
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':  
##  
##   select
```

```
##  
## Attaching package: 'TH.data'
```

```
## The following object is masked from 'package:MASS':  
##  
##   geyser
```

```
library(RColorBrewer)  
library(plotly)
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 3.6.3
```

```
##  
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:MASS':
##
##      select
```

```
## The following object is masked from 'package:stats':
##
## filter
```

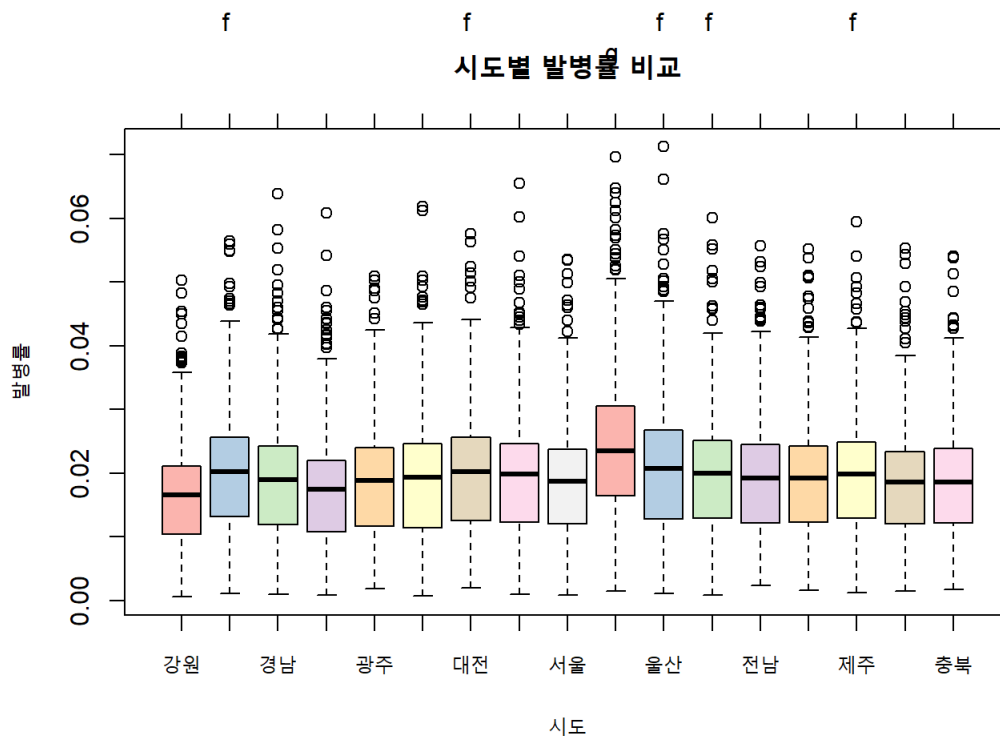
```
## The following object is masked from 'package:graphics':
##
## layout
```

```
analysis_total <- analysis_total_Fixed
analysis_total$시도 <- factor(analysis_total$시도)
fit <- aov(formula=발병률~시도,data=analysis_total)
summary(fit)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## 시도           16 0.0509 0.003182   35.07 <2e-16 ***
## Residuals    18615 1.6893 0.000091
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05),
      col=brewer.pal(17, 'Pastell1'),
      main = "시도별 발병률 비교"
    )
```

[illegible]



#시도별 발병률의 비교
 #차이는 유의미한 것으로 나타난다
 #유난이 높은 곳은 세종으로 나타났다, 나머지는 비슷한 양상을 보인다
 #세종에는 극단치 값을 보이는 것이 있다 이를 통해 특정 날짜의 세종에 발병률이 높아진 원인이 있을것 같다

```
## 시도 발병률
## 1 강원 0.01596549
## 2 경북 0.01681319
## 3 서울 0.01789104
## 4 충남 0.01808048
## 5 광주 0.01821688
## 6 경남 0.01833162
## 7 충북 0.01838652
## 8 대구 0.01839525
## 9 전북 0.01858002
## 10 부산 0.01874255
## 11 전남 0.01892690
## 12 인천 0.01919658
## 13 제주 0.01932978
## 14 경기 0.01963415
## 15 대전 0.01966008
## 16 울산 0.02034245
## 17 세종 0.02408904
```

```
fit <- aov(formula=PM25~시도,data=analysis_total)
summary(fit)
```

```
##          Df Sum Sq Mean Sq F value Pr(>F)
## 시도      16   66072     4129   27.77 <2e-16 ***
## Residuals 18615 2768511     149
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

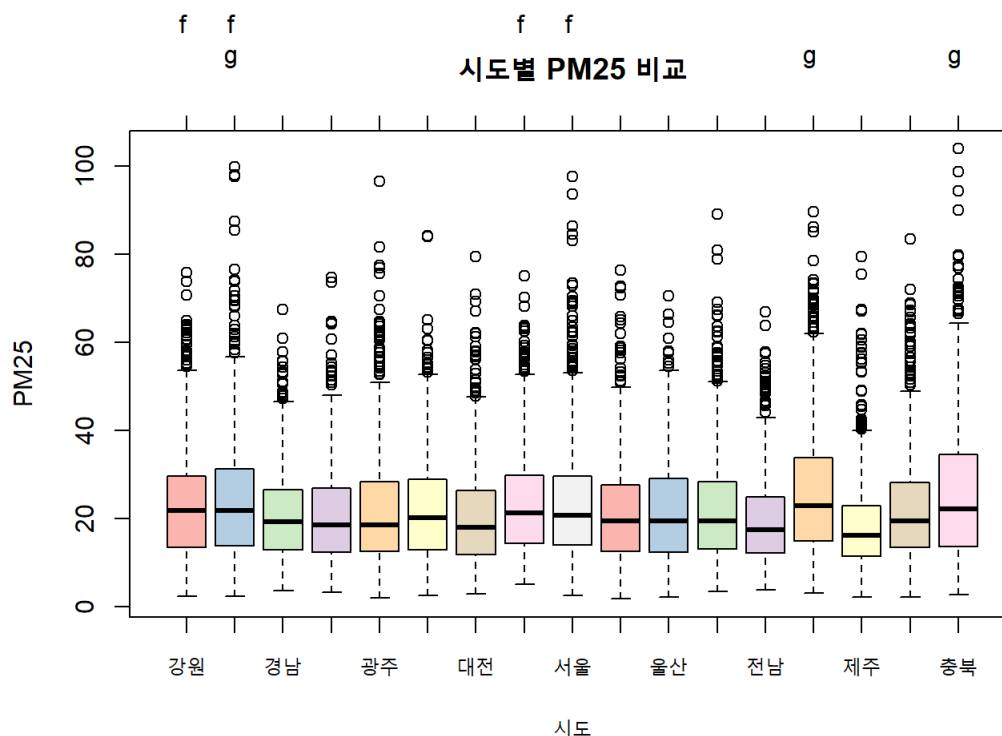
```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05),
      col=brewer.pal(17,'Pastell1'),
      main = "시도별 PM25 비교"
    )
```

```
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfunction("adjusted", ...): Completion with error > absens
```

[illegible]

```
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
```

```
## Warning in brewer.pal(17, "Pastell1"): n too large, allowed maximum for palette Pastell1 is 9
## Returning the palette you asked for with that many colors
```



```
#시도별 PM25 비교
#차이는 유의미한 것으로 나타난다
#PM25가 높은 곳으로는 충북과 전북, PM25가 낮은 곳으로는 제주,전남으로 나타났다
#충북에서 극단치가 보인다
#발병률이 높았던 세종은 PM25에서는 높은 순위를 보이지는 않고 6번째에 위치한다.
```

```
aggregate(formula=PM25~시도,data=analysis_total,FUN=mean) %>% arrange(PM25)
```

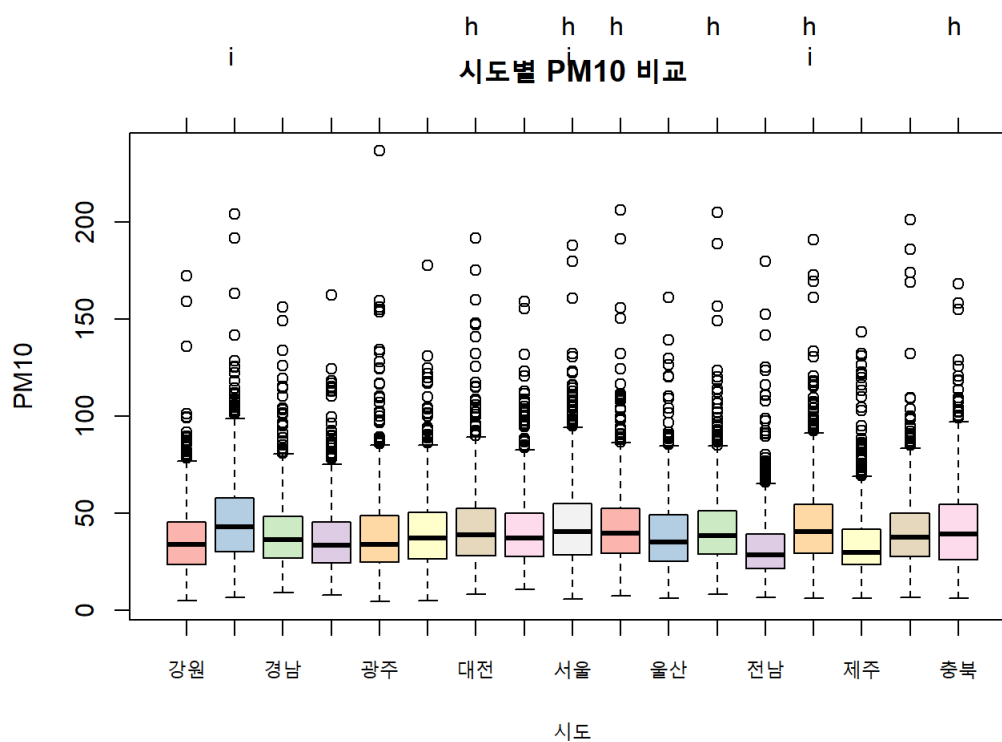
```
## 시도 PM25
## 1 제주 18.39358
## 2 전남 19.53340
## 3 대전 20.02411
## 4 경북 20.75651
## 5 경남 20.76171
## 6 세종 21.01550
## 7 광주 21.72012
## 8 인천 21.88465
## 9 울산 21.93485
## 10 충남 21.98443
## 11 대구 22.00011
## 12 강원 22.90361
## 13 서울 22.99522
## 14 부산 23.44712
## 15 경기 24.13786
## 16 충북 25.30047
## 17 전북 25.71422
```

```
fit <- aov(formula=PM10~시도,data=analysis_total)
summary(fit)
```



```
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
```

```
## Warning in brewer.pal(17, "Pastell"): n too large, allowed maximum for palette Pastell is 9
## Returning the palette you asked for with that many colors
```



```
# 시도별 PM10 비교
# 차이는 유의미한 것으로 나타난다
# PM10이 높은 곳으로는 전북과 경기, PM25가 낮은 곳으로는 제주, 전남으로 나타났다
# PM10과 PM25는 비슷한 양상을 보인다
# 발병률이 최고로 높았던 세종은 PM10에서 또한 중간에 위치한다
```

```
aggregate(formula=PM10~시도, data=analysis_total, FUN=mean) %>% arrange(PM10)
```

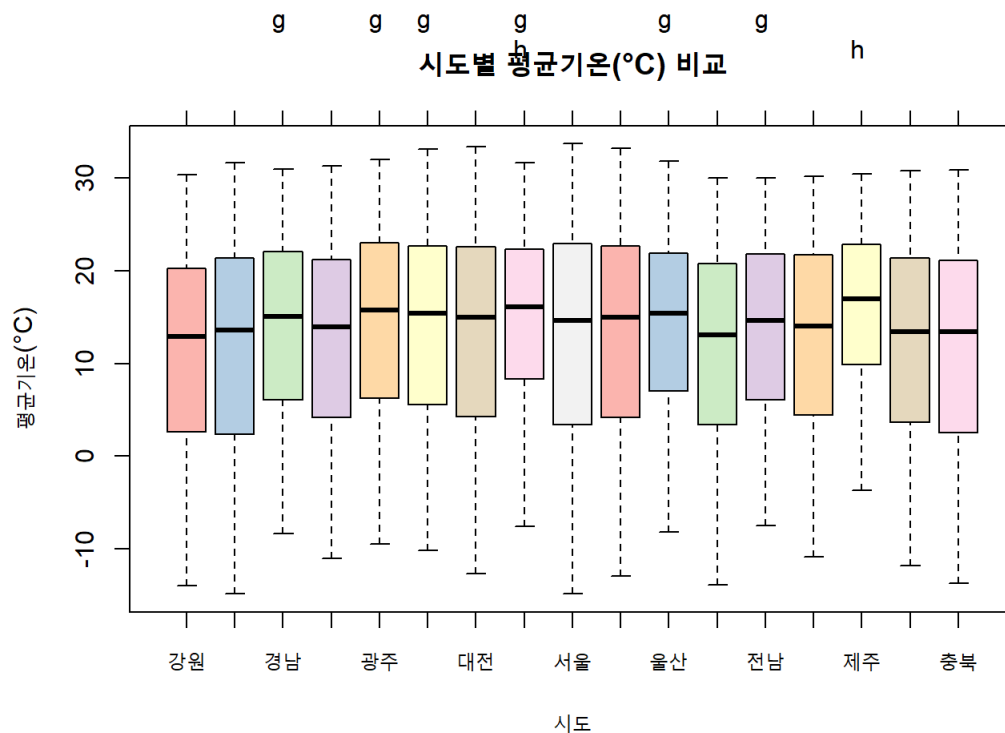

##	시도	PM10
## 1	전남	32.57905
## 2	제주	34.62154
## 3	강원	36.44890
## 4	경북	36.81324
## 5	광주	38.85021
## 6	울산	38.96866
## 7	경남	39.82094
## 8	대구	40.27707
## 9	충남	40.44254
## 10	부산	40.94183
## 11	인천	41.70554
## 12	대전	41.89636
## 13	충북	41.94077
## 14	세종	42.47925
## 15	서울	43.54443
## 16	전북	44.01496
## 17	경기	45.98849

```
fit <- aov(formula=`평균기온(°C)`~시도,data=analysis_total)
summary(fit)
```

```
##              Df    Sum Sq Mean Sq F value Pr(>F)
## 시도          16    30226   1889.1    19.14 <2e-16 ***
## Residuals    18615  1836832     98.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05),
      col=brewer.pal(17, 'Pastell1'),
      main = "시도별 평균기온(°C) 비교")
```

[illegible]



시도별 평균기온 (°C) 비교
 # 차이는 유의미한 것으로 나타난다
 # 평균기온이 높은 곳은 제주, 평균기온이 낮은 곳은 강원으로 예상과 같은 결과가 나왔다
 # 이상치 값은 존재하지 않는다
 # 대체로 다른 변수들보다 집단간의 차이가 적은 편에 속한다

```
aggregate(formula=`평균기온(°C)`~시도,data=analysis_total,FUN=mean) %>% arrange(`평균기온(°C)`)
```

```
## 시도 평균기온(°C)
## 1 강원 11.62339
## 2 인천 12.06524
## 3 경기 12.12680
## 4 충북 12.16904
## 5 충남 12.61407
## 6 경북 12.95783
## 7 전북 13.15654
## 8 서울 13.22318
## 9 세종 13.68266
## 10 대전 13.70593
## 11 전남 14.17382
## 12 경남 14.20308
## 13 대구 14.42709
## 14 울산 14.60931
## 15 광주 14.80557
## 16 부산 15.36907
## 17 제주 16.49498
```

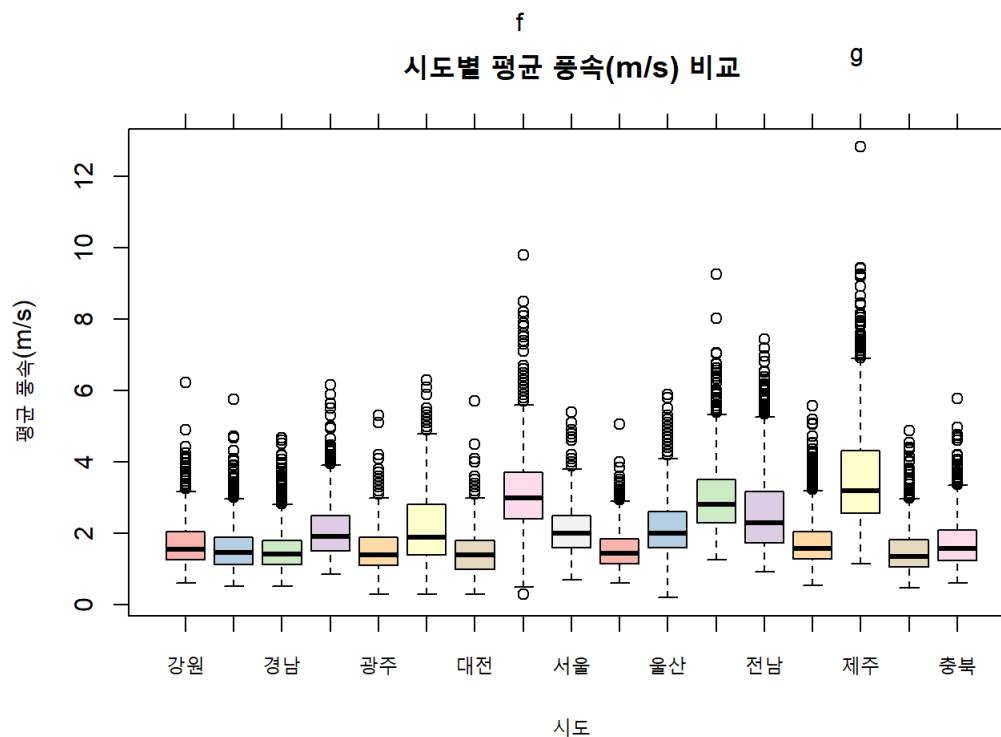
```
fit <- aov(formula=`평균 풍속(m/s)`~시도,data=analysis_total)
summary(fit)
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## 시도 16 7146 446.6 601.8 <2e-16 ***
## Residuals 18615 13815 0.7
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot(cld(tuk, level = .05),
     col=brewer.pal(17,'Pastell1'),
     main = "시도별 평균 풍속(m/s) 비교")
```

```
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
```

```
## Warning in brewer.pal(17, "Pastell"): n too large, allowed maximum for palette Pastell is 9
## Returning the palette you asked for with that many colors
```



```
#시도별 평균 풍속(m/s) 비교
#차이는 유의미한 것으로 나타난다
#다른 변수들에 비해서 집단간의 차이가 뚜렷하다
#섬인 제주가 바람이 많이 불며 이상치 또한 많이 존재한다
```

```
aggregate(formula=`평균 풍속(m/s)`~시도,data=analysis_total,FUN=mean) %>% arrange(`평균 풍속(m/s)`)
```

```
## 시도 평균 풍속(m/s)
## 1 충남 1.518525
## 2 대전 1.524726
## 3 경남 1.562283
## 4 세종 1.562409
## 5 경기 1.566195
## 6 광주 1.582026
## 7 강원 1.724909
## 8 충북 1.733467
## 9 전북 1.761739
## 10 서울 2.071809
## 11 경북 2.092668
## 12 울산 2.170438
## 13 대구 2.177920
## 14 전남 2.573816
## 15 인천 3.003923
## 16 부산 3.152007
## 17 제주 3.575593
```

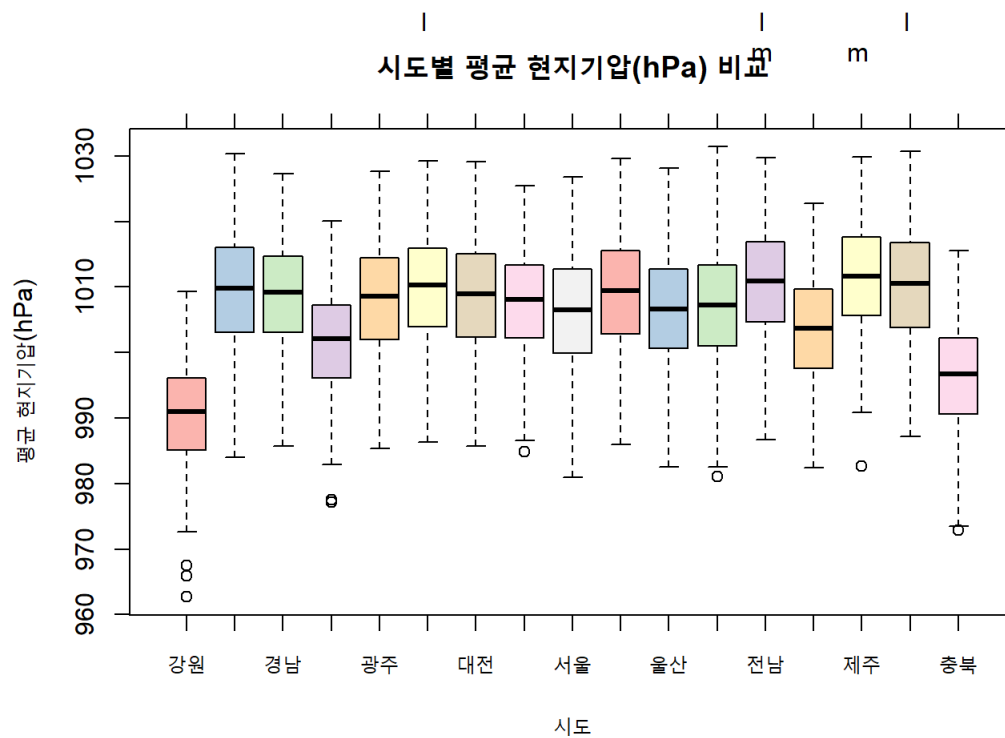
```
fit <- aov(formula=`평균 현지기압(hPa) ~시도`,data=analysis_total)
summary(fit)
```

```
##              Df    Sum Sq Mean Sq F value Pr(>F)
## 시도          16    534273    33392   571.2 <2e-16 ***
## Residuals    18615  1088236         58
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05),
      col=brewer.pal(17, 'Pastell'),
      main = "시도별 평균 현기압 (hPa) 비교")
```

[illegible]

```
## Warning in brewer.pal(17, "Pastell"): n too large, allowed maximum for palette Pastell is 9
## Returning the palette you asked for with that many colors
```



#시도별 평균 현지기압(hPa) 비교
 #차이는 유의미한 것으로 나타난다
 #다른 변수들에 비해서 이상치는 적은 편이다
 #산이 많은 강원이 현지기압이 낮은 편이며 이상치 또한 많이 존재하는 편이다

```
aggregate(formula=~평균 현지기압(hPa)~시도,data=analysis_total,FUN=mean) %>% arrange(~평균 현지기압(hPa)~)
```

```
## 시도 평균 현지기압(hPa)
## 1 강원 990.6346
## 2 충북 996.4020
## 3 경북 1001.5924
## 4 전북 1003.6024
## 5 서울 1006.3178
## 6 울산 1006.5622
## 7 인천 1007.1614
## 8 부산 1007.7126
## 9 광주 1008.2412
## 10 대전 1008.7540
## 11 경남 1008.8731
## 12 세종 1009.2316
## 13 경기 1009.5797
## 14 대구 1009.9084
## 15 충남 1010.3554
## 16 전남 1010.7189
## 17 제주 1011.5058
```

```
fit <- aov(formula=~일 최심신적설(cm)~시도,data=analysis_total)
summary(fit)
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## 시도 16 21.5 1.3410 11.3 <2e-16 ***
## Residuals 18615 2208.4 0.1186
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

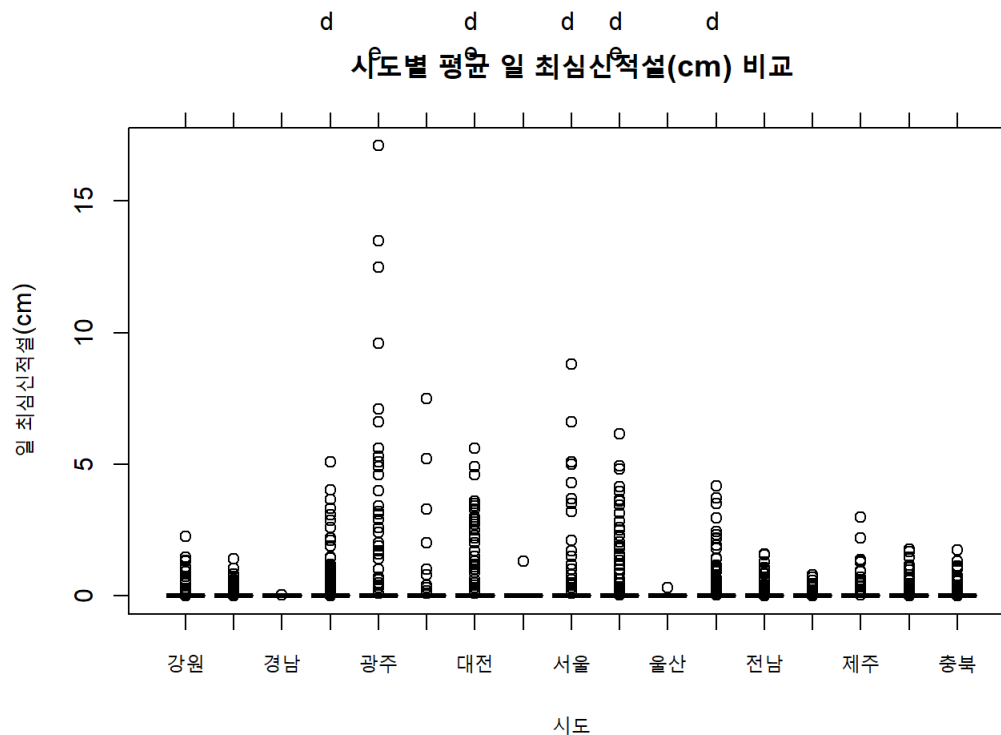
```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05),
      col=brewer.pal(17,'Pastell'),
      main = "시도별 평균 일 최심신적설(cm) 비교")
```

```
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
```

[illegible]

```
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
```

```
## Warning in brewer.pal(17, "Pastell1"): n too large, allowed maximum for palette Pastell1 is 9
## Returning the palette you asked for with that many colors
```

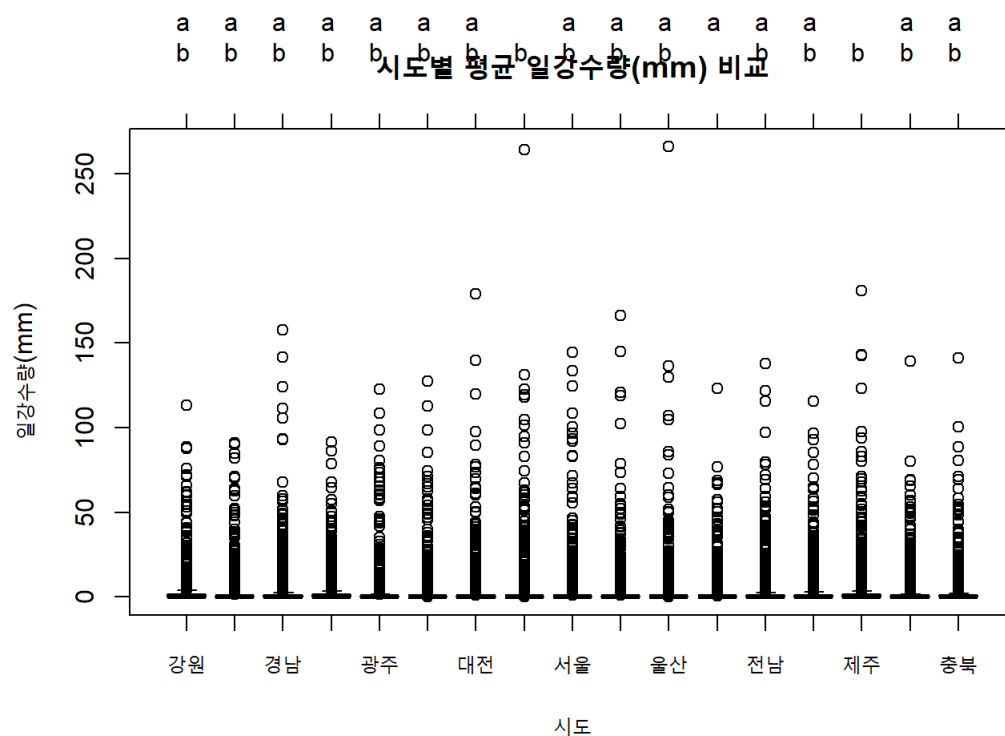


```
# 시도별 평균 일 최심신적설 (cm) 비교
# 차이는 유의미한 것으로 나타난다
# 0인 값이 대부분이며 광주에 이상치가 많이 존재한다
# 겨울에 춥고 눈이 많이 온다던 강원은 생각과 달리 낮은 편이다
```

```
fit <- aov(formula=`일강수량(mm)`~시도,data=analysis_total)
summary(fit)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## 시도          16   3917    244.8   1.786  0.027 *
## Residuals    18615 2551280    137.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05),
      col=brewer.pal(17,'Pastell1'),
      main = " 시도별 평균 일강수량(mm) 비교" )
```

```
# 시도별 평균 일강수량(mm) 비교
# 차이는 유의미한 것으로 나타난다
# 0인 값이 대부분이다
# 값의 편차가 큰 편이다
# 예상과 갖게 섬나라인 제주에서 강수량이 많다
```

```
aggregate(formula=`일강수량(mm)`~시도,data=analysis_total,FUN=mean) %>% arrange(`일강수량(mm)`)
```

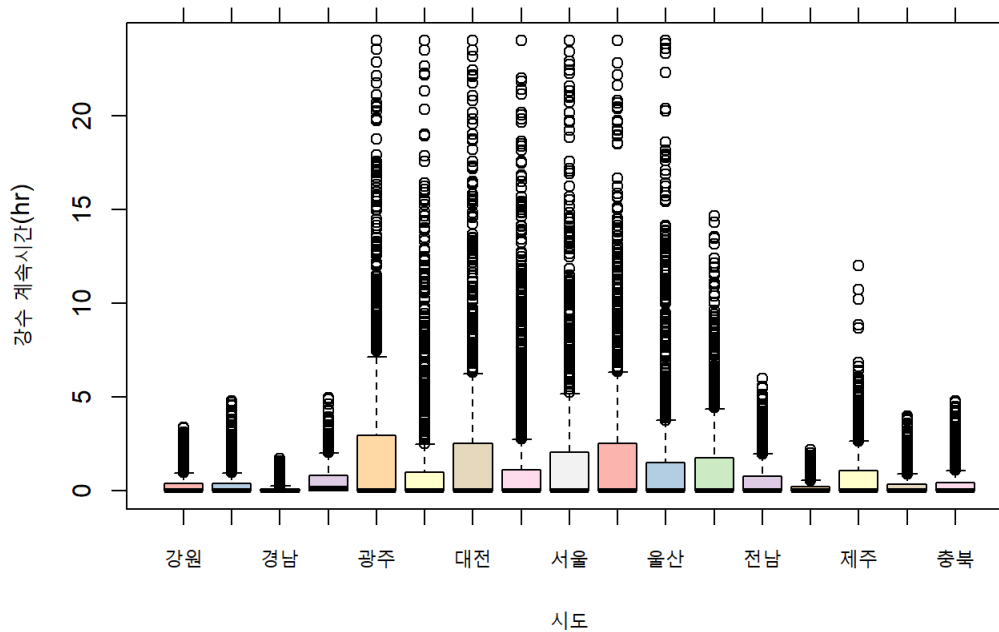
```
## 시도 일강수량(mm)
## 1 인천 2.405809
## 2 대구 2.909307
## 3 충남 3.025791
## 4 경북 3.103076
## 5 경기 3.127482
## 6 서울 3.201642
## 7 충북 3.274599
## 8 전북 3.341440
## 9 강원 3.361638
## 10 세종 3.430109
## 11 울산 3.450182
## 12 광주 3.509854
## 13 전남 3.521248
## 14 대전 3.556569
## 15 경남 3.917238
## 16 부산 4.154380
## 17 제주 4.444434
```

```
fit <- aov(formula=`강수 계속시간(hr)`~시도,data=analysis_total)
summary(fit)
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## 시도 16 12717 794.8 99.01 <2e-16 ***
## Residuals 18615 149430 8.0
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot(cld(tuk, level = .05),
     col=brewer.pal(17,'Pastell'),
     main = "시도별 평균 강수 계속시간(hr) 비교" )
```


시도별 평균 강수 계속시간(hr) 비교



#시도별 평균 강수 계속시간(hr) 비교
 #차이는 유의미한 것으로 나타난다
 #0인 값이 대부분이다
 #값의 편차가 강수량에 비해 적은 편이다.
 #제주는 강수량이 많은 편에 비해 비가 오는 시간은 적은 편이다
 #지역간의 편차가 큰편이다

```
aggregate(formula=`강수 계속시간(hr)`~시도,data=analysis_total,FUN=mean) %>% arrange(`강수 계속시간(hr)`)
```

```
## 시도 강수 계속시간(hr)
## 1 경남 0.1509372
## 2 전북 0.2096865
## 3 강원 0.3330567
## 4 충남 0.3567503
## 5 경기 0.3912974
## 6 충북 0.4323011
## 7 경북 0.5497067
## 8 전남 0.6134607
## 9 제주 0.7790876
## 10 인천 1.3909246
## 11 대구 1.6971624
## 12 부산 1.9215055
## 13 울산 1.9307391
## 14 서울 2.0844617
## 15 세종 2.2267974
## 16 대전 2.2920894
## 17 광주 2.4352281
```

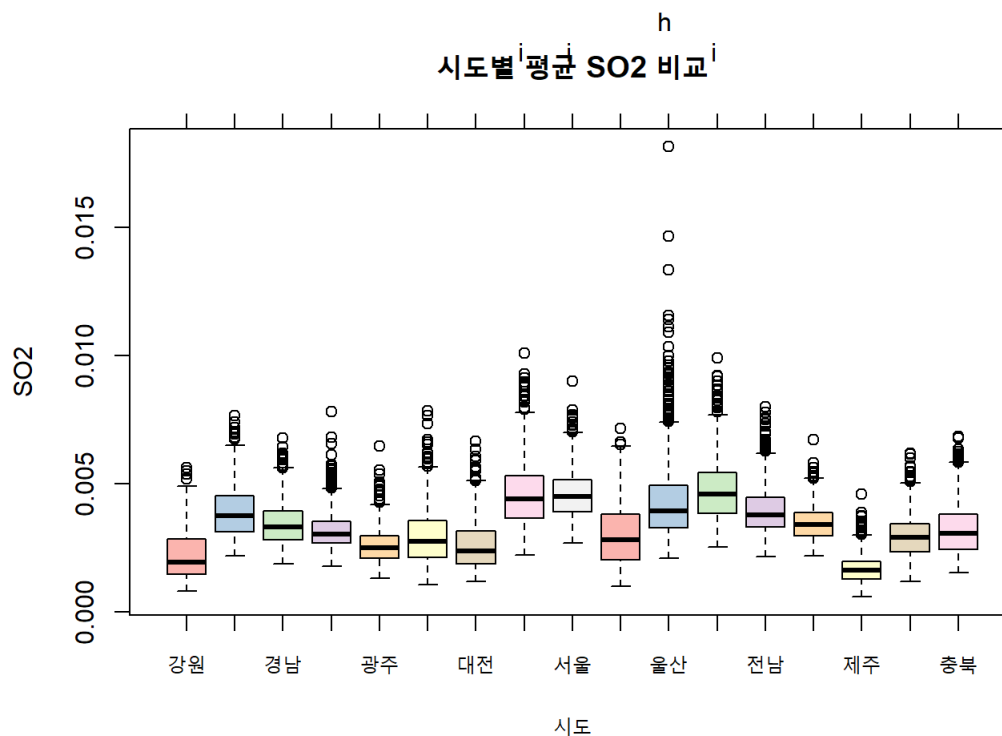
```
fit <- aov(formula=S02~시도,data=analysis_total)
summary(fit)
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## 시도 16 0.01378 0.0008611 886.9 <2e-16 ***
## Residuals 18615 0.01807 0.0000010
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05), col=brewer.pal(17,'Pastell1'),
      main = "시도별 평균 SO2 비교" )
```

```
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfunction("adjusted", ...): Completion with error > abseps
```

```
## Warning in brewer.pal(17, "Pastell1"): n too large, allowed maximum for palette Pastell1 is 9
## Returning the palette you asked for with that many colors
```



```
#시도별 평균 SO2 비교
#차이는 유의미한 것으로 나타난다
#지역간의 편차가 크다
#공장지대인 울산이 다른 지역에 비해 높은 양상을 보인다
#울산과 붙어있는 부산, 수도권인 서울과 인천이 높은 양상을 보인다
#미세먼지의 요소들간의 비슷한 양상을 보인다(SO2,NO2,O3,CO)
```

```
aggregate(formula=SO2~시도,data=analysis_total,FUN=mean) %>% arrange(SO2)
```

```
##   시도      SO2
## 1  제주 0.001680595
## 2  강원 0.002184327
## 3  대전 0.002579741
## 4  광주 0.002600924
## 5  대구 0.002914837
## 6  충남 0.002951149
## 7  세종 0.003017727
## 8  경북 0.003169494
## 9  충북 0.003193873
##10  경남 0.003434274
##11  전북 0.003462246
##12  경기 0.003893332
##13  전남 0.003960595
##14  울산 0.004332363
##15  부산 0.004605508
##16  서울 0.004610812
##17  인천 0.004744645
```

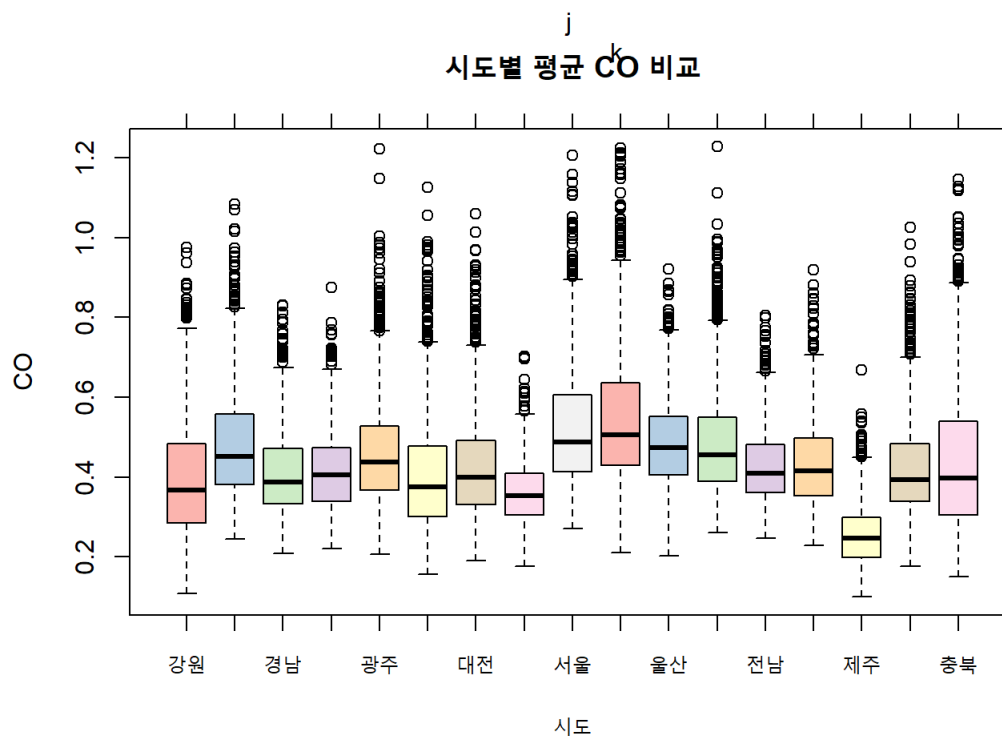
```
fit <- aov(formula=CO~시도,data=analysis_total)
summary(fit)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## 시도          16   77.5    4.841   281.1 <2e-16 ***
## Residuals 18615  320.6    0.017
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05), col=brewer.pal(17, 'Pastell1') ,
      main = "시도별 평균 CO 비교" )
```

```
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
```

```
## Warning in brewer.pal(17, "Pastell1"): n too large, allowed maximum for palette Pastell1 is 9
## Returning the palette you asked for with that many colors
```



#시도별 평균 CO 비교
 #차이는 유의미한 것으로 나타난다
 #지역간의 편차가 크다
 #공장지대인 울산이 다른 지역에 비해 높은 양상을 보인다
 #울산과 붙어있는 부산, 수도권인 서울과 인천이 높은 양상을 보인다
 #미세먼지의 요소들간의 비슷한 양상을 보인다(SO2,NO2,O3,CO)

```
aggregate(formula=CO~시도,data=analysis_total,FUN=mean) %>% arrange(CO)
```

```
## 시도 CO
## 1 제주 0.2562050
## 2 부산 0.3613538
## 3 강원 0.3937574
## 4 대구 0.4081705
## 5 경남 0.4097785
## 6 경북 0.4160029
## 7 충남 0.4245540
## 8 대전 0.4278963
## 9 전남 0.4289108
## 10 전북 0.4328231
## 11 충북 0.4408928
## 12 광주 0.4662895
## 13 경기 0.4827654
## 14 울산 0.4857937
## 15 인천 0.4871533
## 16 서울 0.5255507
## 17 세종 0.5484376
```

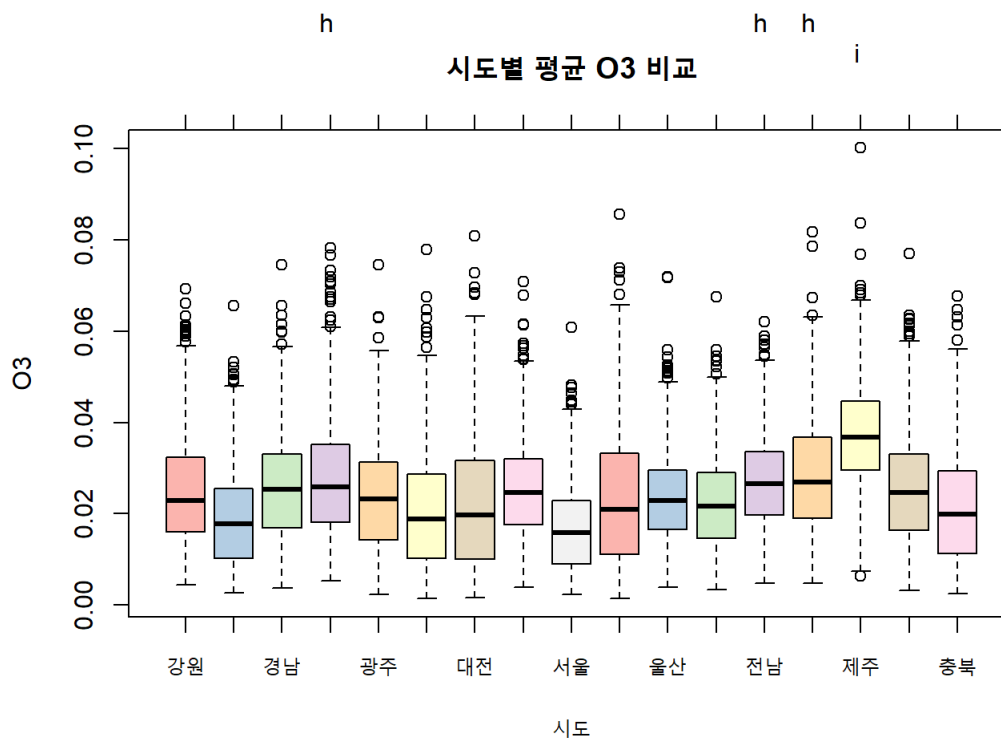
```
fit <- aov(formula=O3~시도,data=analysis_total)
summary(fit)
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## 시도 16 0.3578 0.022362 167.2 <2e-16 ***
## Residuals 18615 2.4899 0.000134
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05), col=brewer.pal(17,'Pastell1'),
      main = "시도별 평균 O3 비교" )
```

```
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
## Warning in RET$pfuction("adjusted", ...): Completion with error > abseps
```

```
## Warning in brewer.pal(17, "Pastell"): n too large, allowed maximum for palette Pastell is 9
## Returning the palette you asked for with that many colors
```



```
# 시도별 평균 O3 비교
# 차이는 유의미한 것으로 나타난다
# 지역간의 편차가 크다
# 공장지대인 울산이 다른 지역에 비해 높은 양상을 보인다
# 울산과 붙어있는 부산, 수도권인 서울과 인천이 높은 양상을 보인다
# 미세먼지의 요소들간의 비슷한 양상을 보인다(SO2, NO2, O3, CO)
```



```
aggregate(formula=O3~시도,data=analysis_total,FUN=mean) %>% arrange(O3)
```

```
##      시도      O3
## 1 서울 0.01668578
## 2 경기 0.01892358
## 3 대구 0.02059165
## 4 충북 0.02122055
## 5 대전 0.02201361
## 6 인천 0.02224103
## 7 세종 0.02309253
## 8 광주 0.02329249
## 9 울산 0.02362190
## 10 강원 0.02529897
## 11 부산 0.02546126
## 12 충남 0.02573987
## 13 경남 0.02591471
## 14 전남 0.02719429
## 15 경북 0.02768057
## 16 전북 0.02826032
## 17 제주 0.03700835
```

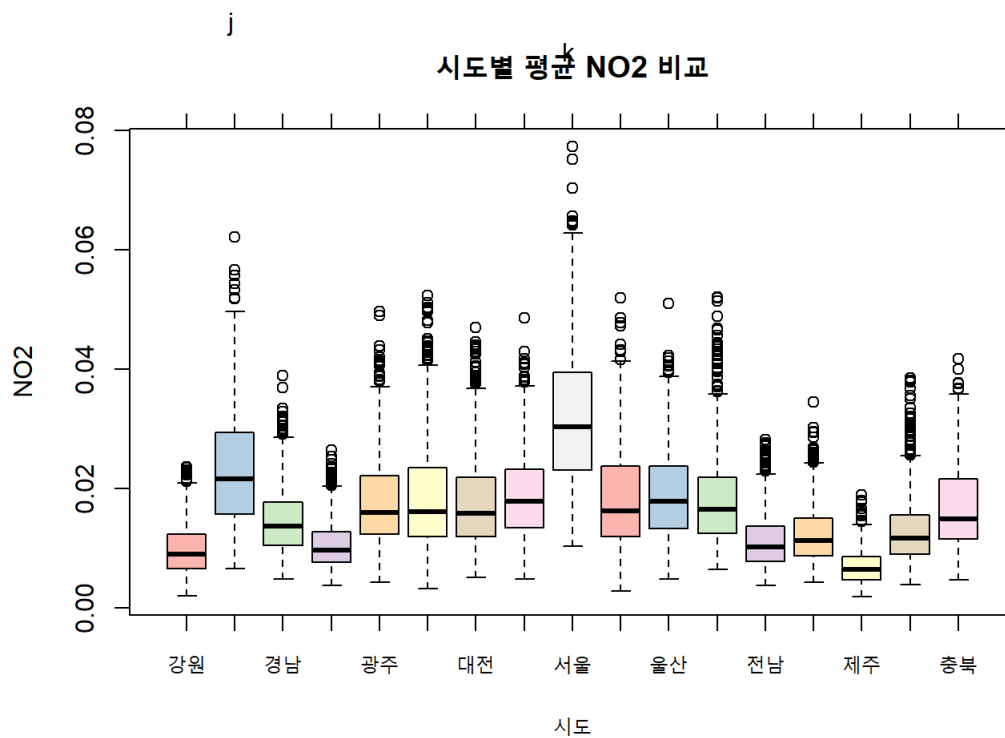
```
fit <- aov(formula=NO2~시도,data=analysis_total)
summary(fit)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## 시도          16 0.5989 0.03743   741.6 <2e-16 ***
## Residuals    18615 0.9396 0.00005
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tuk <- glht(fit, linfct=mcp(시도='Tukey'))
plot( cld(tuk, level = .05), col=brewer.pal(17,'Pastell') ,
      main = "시도별 평균 NO2 비교" )
```

```
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
## Warning in RET$pffunction("adjusted", ...): Completion with error > abseps
```

```
## Warning in brewer.pal(17, "Pastell"): n too large, allowed maximum for palette Pastell is 9
## Returning the palette you asked for with that many colors
```



시도별 평균 NO2 비교
 # 차이는 유의미한 것으로 나타난다
 # 지역간의 편차가 크다
 # 공장지대인 울산이 다른 지역에 비해 높은 양상을 보인다
 # 울산과 붙어있는 부산, 수도권인 서울과 인천이 높은 양상을 보인다
 # 미세먼지의 요소들간의 비슷한 양상을 보인다(SO2, NO2, O3, CO)

```
aggregate(formula=NO2~시도,data=analysis_total,FUN=mean) %>% arrange(NO2)
```

```
##      시도      NO2
## 1  제주 0.006759712
## 2  강원 0.009687837
## 3  경북 0.010572085
## 4  전남 0.011172208
## 5  전북 0.012258530
## 6  충남 0.012873303
## 7  경남 0.014667995
## 8  충북 0.016796769
## 9  대전 0.017574609
## 10 광주 0.017852193
## 11 인천 0.018095602
## 12 세종 0.018251362
## 13 대구 0.018541864
## 14 부산 0.018742499
## 15 울산 0.018929079
## 16 경기 0.023199353
## 17 서울 0.031954832
```

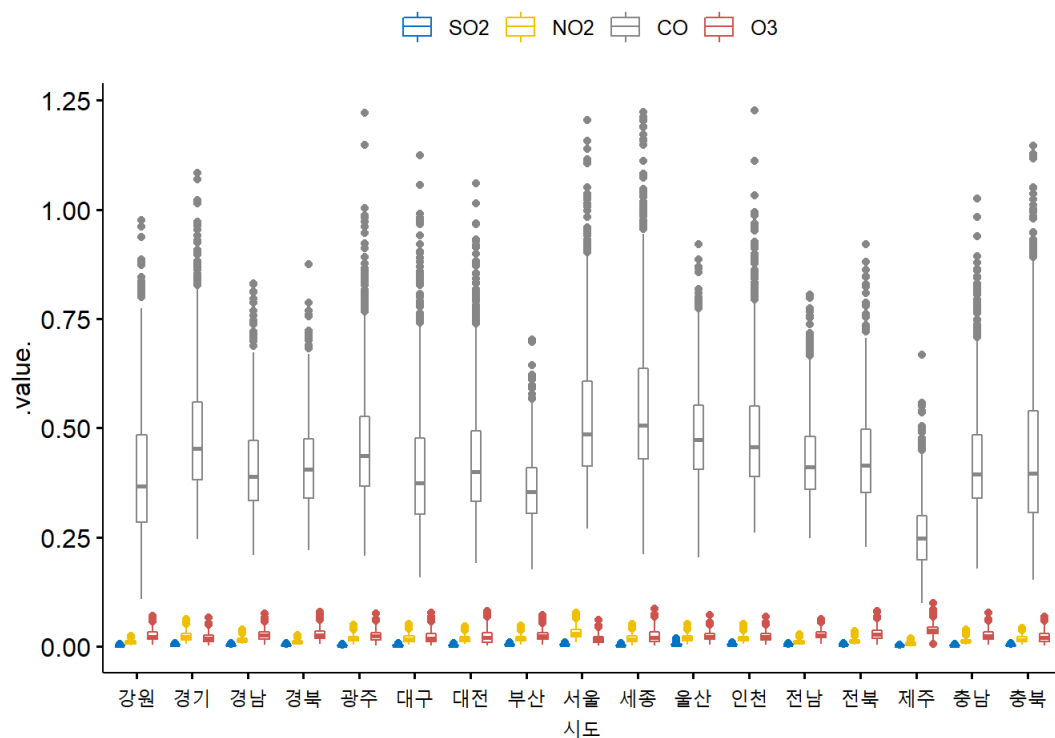
```
fit <- manova(cbind(analysis_total$SO2,analysis_total$NO2,analysis_total$O3,analysis_total$CO)~analysis_total$시도)
summary(fit)
```

```
##              Df Pillai approx F num Df den Df      Pr(>F)
## analysis_total$시도 16 1.1487   468.74      64 74460 < 2.2e-16 ***
## Residuals      18615
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
library(ggpubr)
```

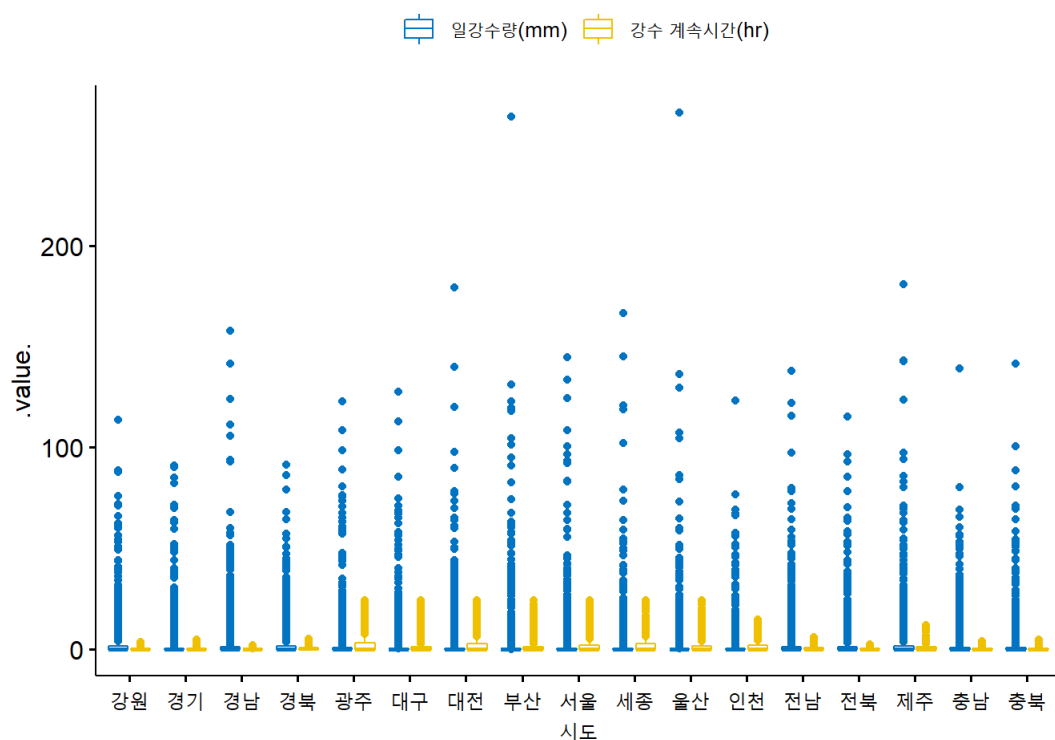
```
## Loading required package: magrittr
```

```
ggboxplot(
  analysis_total, x = "시도", y = c("SO2", "NO2", "CO", "O3"),
  merge = TRUE, palette = "jco"
)
```



#시각적인 결과와 위의 통계와는 달라보인다

```
ggboxplot(
  analysis_total, x = "시도", y = c("일강수량(mm)", "강수 계속시간(hr)"),
  merge = TRUE, palette = "jco"
)
```



#시각적인 결과와 위의 통계와는 달라보인다