

지역 군집화

```
analysis_total <- analysis_total %>%
  group_by(시도) %>%
  summarise(
    `평균기온(°C)` = mean(`평균기온(°C)`, na.rm=TRUE),
    `최저기온(°C)` = mean(`최저기온(°C)`, na.rm=TRUE),
    `최고기온(°C)` = mean(`최고기온(°C)`, na.rm=TRUE),
    `평균 풍속(m/s)` = mean(`평균 풍속(m/s)`, na.rm=TRUE),
    `평균 현지기압(hPa)` = mean(`평균 현지기압(hPa)`, na.rm=TRUE),
    `일 최심신적설(cm)` = mean(`일 최심신적설(cm)`, na.rm=TRUE),
    `일강수량(mm)` = mean(`일강수량(mm)`, na.rm=TRUE),
    `강수 계속시간(hr)` = mean(`강수 계속시간(hr)`, na.rm=TRUE),
    SO2 = mean(SO2, na.rm=TRUE),
    CO = mean(CO, na.rm=TRUE),
    O3 = mean(O3, na.rm=TRUE),
    NO2 = mean(NO2, na.rm=TRUE),
    PM10 = mean(PM10, na.rm=TRUE),
    PM25 = mean(PM25, na.rm=TRUE),
    발병률 = sum(발병률)
  )

analysis_total_1 <- analysis_total %>% select(-`최저기온(°C)`, -`최고기온(°C)`)

analysis_total_1 <- analysis_total_1[-1]

rownames(analysis_total_1) <- analysis_total$시도
```

```
## Warning: Setting row names on a tibble is deprecated.
```

```
df <- scale(analysis_total_1)

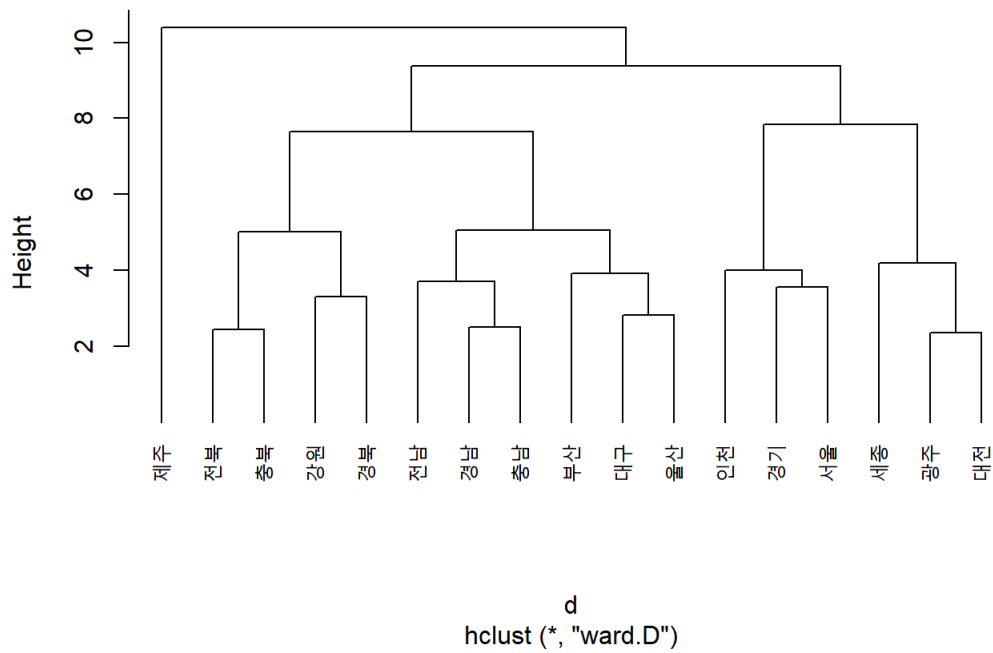
d <- dist(df)

fit <- hclust(d, method='ward')
```

```
## The "ward" method has been renamed to "ward.D"; note new "ward.D2"
```

```
plot(fit, hang=-1)
```

Cluster Dendrogram



```
# k=5 군집개수로 수행도의 군집들을 자름!!
clusters <- cutree(fit, k = 8)
```

```
fit$clusters <- clusters
```

```
library(cluster)
```

```
# (4) 차원축소 후, 군집결과 시각화
```

```
clusplot(
  analysis_total_1,
  fit$cluster,
  color = T,
  shade = T,
  labels = 2,
  lines = 0
)
```

CLUSPLOT(analysis_total_1)

