전종설2

```
library(dplyr)
library(corrplot)
library(car)
library(leaps)
library(psych)
```

```
df <- read.csv(file=, fileEncoding = "cp949")

str(df)
```

```
## 'data.frame':    20 obs. of  10 variables:
##  $ 법정동                    : chr  "강동구 천호동" "강북구 수유동" "강서구 화곡동" "관악구 
##  $ 인구수                    : int  85764 102828 219202 224184 245562 91273 115338 10
##  $ 급속충전기.대.           : int  5 1 2 7 4 6 7 4 3 10 ...
##  $ 완속충전기.대.           : int  2 0 0 6 11 10 10 0 2 16 ...
##  $ 전기차.등록.수           : int  210 185 399 439 444 252 3313 234 271 462 ...
##  $ 전기차.충전소.개수       : int  7 1 2 13 15 16 17 4 5 26 ...
##  $ 전기차.한대당.급속.충전소: num  0.02381 0.00541 0.00501 0.01595 0.00901 ...
##  $ 전기차.한대당.충전소     : num  0.03333 0.00541 0.00501 0.02961 0.03378 ...
##  $ 인구수.한명당.급속.충전기: num  5.83e-05 9.72e-06 9.12e-06 3.12e-05 1.63e-05 6.57e-0
##  $ 친환경차_전기차.비율     : num  0.879 0.964 0.913 0.936 0.943 ...
```

```
df_cor <- df%>%
  select(-법정동,-전기차.충전소.개수)
corr_df <- cor(df_cor)
corrplot(corr_df, method = 'number', order = 'hclust', type = 'lower', diag = FALSE)
```


```
set.seed(123)
model <- lm(급속충전기.대.~., df_cor)

summary(model)
```

```
##
## Call:
## lm(formula = 급속충전기.대. ~ ., data = df_cor)
##
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -1.4611 -0.3343 -0.1123  0.4122  1.1885 
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)   
## (Intercept)                6.138e+00  7.576e+00   0.810  0.43355   
## 인구수                    9.638e-06  7.605e-06   1.267  0.22908   
## 완속충전기.대.            3.587e-01  1.135e-01   3.159  0.00823 **
## 전기차.등록.수            5.204e-04  7.469e-04   0.697  0.49923   
## 전기차.한대당.급속.충전소  2.172e+02  1.160e+02   1.872  0.08577 . 
## 전기차.한대당.충전소      -6.511e+01  2.698e+01  -2.414  0.03270 * 
## 인구수.한명당.급속.충전기  3.500e+04  4.648e+04   0.753  0.46600   
## 친환경차_전기차.비율      -8.007e+00  7.822e+00  -1.024  0.32617   
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8285 on 12 degrees of freedom
## Multiple R-squared:  0.9611, Adjusted R-squared:  0.9385 
## F-statistic: 42.41 on 7 and 12 DF,  p-value: 1.576e-07
```

```
vif(model)
```

```
##                    인구수            완속충전기.대.            전기차.등록.수 
##                  4.719085                  8.821511                  7.254413 
## 전기차.한대당.급속.충전소      전기차.한대당.충전소 인구수.한명당.급속.충전기 
##                 51.525154                 15.823227                 40.737862 
##      친환경차_전기차.비율 
##                  1.406167 
```

```
df2 <- df_cor%>%
  select(인구수, 급속충전기.대.,완속충전기.대., 전기차.등록.수, 친환경차_전기차.비율)
```

```
## best subset selection
response <- df2$급속충전기.대.
predictors <- df2[, c("인구수","완속충전기.대.","전기차.등록.수","친환경차_전기차.비율")]
bestsub.model <- regsubsets(response ~ ., data = data.frame(response, predictors), nvma
best_subset <- summary(bestsub.model)$which[which.min(summary(bestsub.model)$bic), ]
print(best_subset)
```

```
##          (Intercept)               인구수       완속충전기.대. 
##                 TRUE                FALSE                 TRUE 
##       전기차.등록.수 친환경차_전기차.비율 
##                FALSE                FALSE 
```

```
summary(bestsub.model)
```

```
## Subset selection object
## Call: regsubsets.formula(response ~ ., data = data.frame(response, 
##     predictors), nvmax = 4)
## 4 Variables  (and intercept)
##                      Forced in Forced out
## 인구수                   FALSE      FALSE
## 완속충전기.대.           FALSE      FALSE
## 전기차.등록.수           FALSE      FALSE
## 친환경차_전기차.비율     FALSE      FALSE
## 1 subsets of each size up to 4
## Selection Algorithm: exhaustive
##          인구수 완속충전기.대. 전기차.등록.수 친환경차_전기차.비율
## 1  ( 1 ) " "    "*"            " "            " "                 
## 2  ( 1 ) "*"    "*"            " "            " "                 
## 3  ( 1 ) "*"    "*"            " "            "*"                 
## 4  ( 1 ) "*"    "*"            "*"            "*"                 
```

```
#Performance measures
cbind(
  RSS = summary(bestsub.model)$rss,
  Cp     = summary(bestsub.model)$cp,
  r2     = summary(bestsub.model)$rsq,
  Adj_r2 = summary(bestsub.model)$adjr2,
  BIC    =summary(bestsub.model)$bic
)
```

```
##           RSS        Cp        r2    Adj_r2        BIC
## [1,] 108.0277 -0.382789 0.4904355 0.4621264 -7.4925122
## [2,] 105.7772  1.291861 0.5010512 0.4423513 -4.9178373
## [3,] 104.3850  3.090596 0.5076181 0.4152965 -2.1870842
## [4,] 103.7583  5.000000 0.5105741 0.3800605  0.6882171
```

```
predictors1 <- df2[, c("완속충전기.대.")]
predictors2 <- df2[, c("인구수","완속충전기.대.")]
predictors3 <- df2[, c("인구수","완속충전기.대.","친환경차_전기차.비율")]
predictors4 <- df2[, c("인구수","완속충전기.대.","전기차.등록.수","친환경차_전기차.비율")]
scaled_predictors1 <- scale(predictors1)
scaled_predictors2 <- scale(predictors2)
scaled_predictors3 <- scale(predictors3)
scaled_predictors4 <- scale(predictors4)
library(psych)
kmo_result1 <- KMO(scaled_predictors1)
print(kmo_result1)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = scaled_predictors1)
## Overall MSA =  NaN
## MSA for each item = 
## [1] NaN
```

```
set.seed(123)
kmo_result2 <- KMO(scaled_predictors2)
print(kmo_result2)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = scaled_predictors2)
## Overall MSA =  0.5
## MSA for each item = 
##         인구수 완속충전기.대. 
##            0.5            0.5 
```

```
set.seed(123)
kmo_result3 <- KMO(scaled_predictors3)
print(kmo_result3)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = scaled_predictors3)
## Overall MSA =  0.55
## MSA for each item = 
##               인구수       완속충전기.대. 친환경차_전기차.비율 
##                 0.56                 0.54                 0.58 
```

```
set.seed(123)
kmo_result4 <- KMO(scaled_predictors4)
print(kmo_result4)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = scaled_predictors4)
## Overall MSA =  0.6
## MSA for each item = 
##               인구수       완속충전기.대.       전기차.등록.수 
##                 0.59                 0.61                 0.59 
## 친환경차_전기차.비율 
##                 0.59 
```

```
lm2 <- lm(급속충전기.대.~인구수+완속충전기.대., data=df2)

lm3 <- lm(급속충전기.대.~인구수+완속충전기.대.
          +친환경차_전기차.비율, data=df2)

lm4 <- lm(급속충전기.대.~인구수+완속충전기.대.
          +전기차.등록.수
          +친환경차_전기차.비율, data=df2)
```

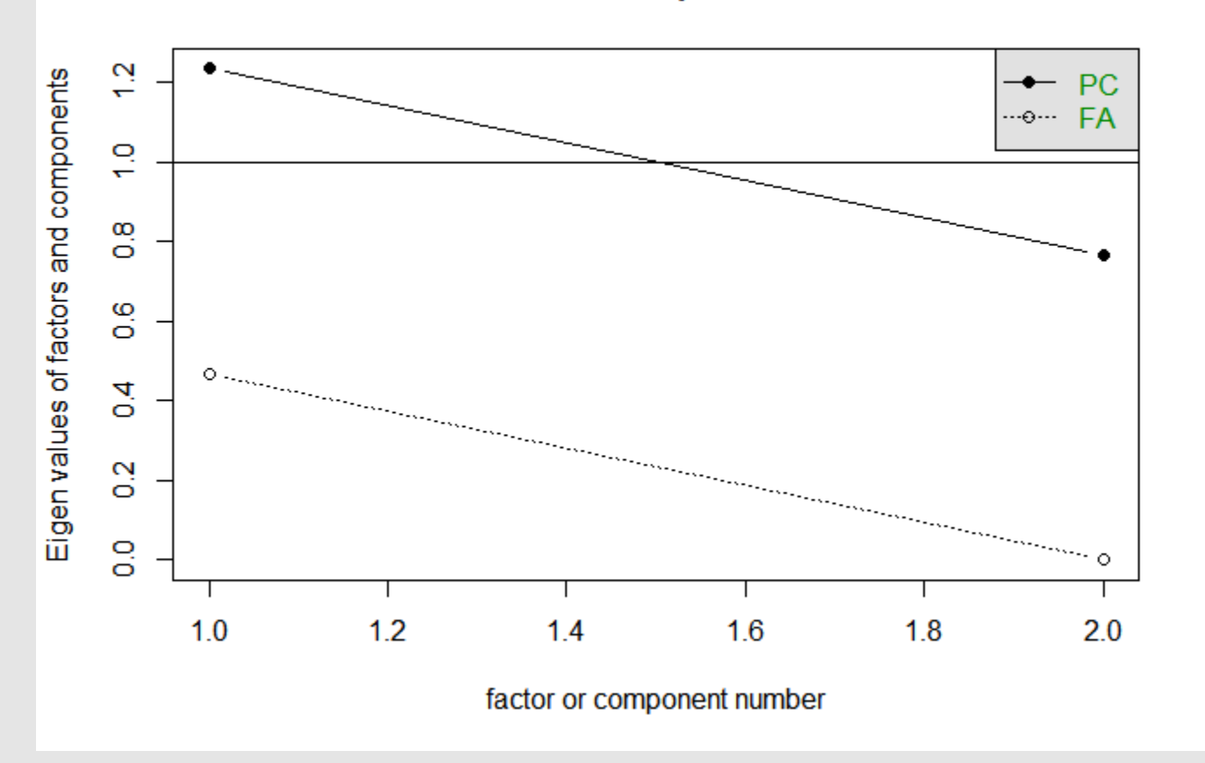
```
anova(lm2,lm3,lm4)
```

```
## Analysis of Variance Table
##
## Model 1: 급속충전기.대. ~ 인구수 + 완속충전기.대.
## Model 2: 급속충전기.대. ~ 인구수 + 완속충전기.대. + 친환경차_전기차.비율
## Model 3: 급속충전기.대. ~ 인구수 + 완속충전기.대. + 전기차.등록.수 + 
##     친환경차_전기차.비율
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     17 105.78                            
## 2     16 104.39  1   1.39219 0.2013 0.6601
## 3     15 103.76  1   0.62667 0.0906 0.7676
```

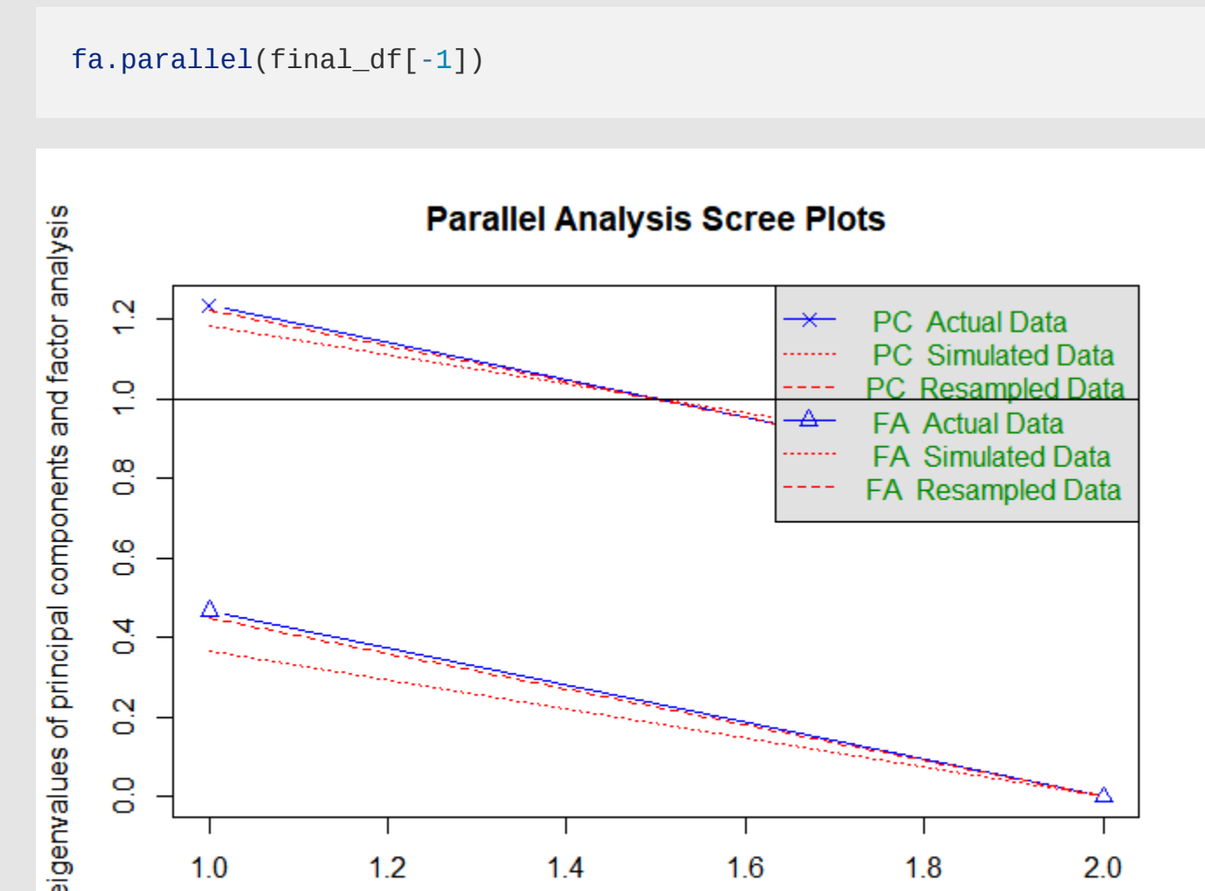
```
## 최종 model 선택
final_df <- df%>%
  select("법정동","인구수","완속충전기.대.")
str(final_df)
```

```
## 'data.frame':    20 obs. of  3 variables:
##  $ 법정동        : chr  "강동구 천호동" "강북구 수유동" "강서구 화곡동" "관악구 봉천동" ...
##  $ 인구수        : int  85764 102828 219202 224184 245562 91273 115338 103061 101795 
##  $ 완속충전기.대.: int  2 0 0 6 11 10 10 0 2 16 ...
```

```
scree(final_df[-1])
```

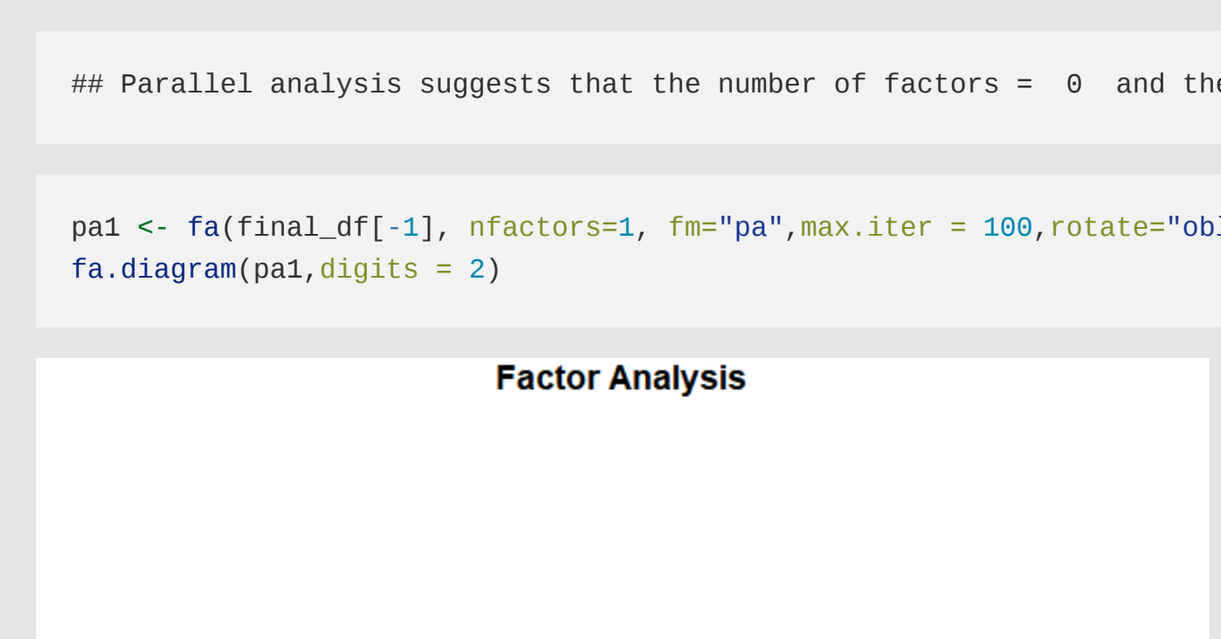



```
fa.parallel(final_df[-1])
```




```
## Parallel analysis suggests that the number of factors =  0  and the number of compo
```

```
pa1 <- fa(final_df[-1], nfactors=1, fm="pa",max.iter = 100,rotate="oblimin")
fa.diagram(pa1,digits = 2)
```




```
factor_loadings <- pa1$loadings
factor_loadings
```

```
##
## Loadings:
##                PA1  
## 인구수         0.483
## 완속충전기.대. 0.483
##
##                  PA1
## SS loadings    0.466
## Proportion Var 0.233
```

```
final_df$충전환경 <- final_df$인구수 * 0.48 + final_df$완속충전기.대. * 0.48
```

```
final_df$급속충전기결핍률 <- 1-(df0$급속충전기.대./df0$인구수)
final_df$공공자전거거치대결핍률 <- 1-(df0$공공자전거.거치대수.LCD...QR./df0$인구수)
final_df$택배함결핍률 <- 1-(df0$택배함.개수/df0$인구수)
final_df$물류창고결핍률 <- 1-(df0$물류창고.개수/df0$인구수)
str(final_df)
```

```
## 'data.frame':    20 obs. of  8 variables:
##  $ 법정동                : chr  "강동구 천호동" "강북구 수유동" "강서구 화곡동" "관악구 봉천
##  $ 인구수                : int  85764 102828 219202 224184 245562 91273 115338 10306
##  $ 완속충전기.대.        : int  2 0 0 6 11 10 10 0 2 16 ...
##  $ 충전환경              : num  41168 49357 105217 107611 117875 ...
##  $ 급속충전기결핍률      : num  1 1 1 1 1 ...
##  $ 공공자전거거치대결핍률: num  0.998 0.998 0.999 0.998 0.998 ...
##  $ 택배함결핍률          : num  1 1 1 1 1 ...
##  $ 물류창고결핍률        : num  1 1 1 1 1 ...
```

```
final <- final_df %>%
  select(-완속충전기.대.,-인구수)
str(final)
```

```
## 'data.frame':    20 obs. of  6 variables:
##  $ 법정동                : chr  "강동구 천호동" "강북구 수유동" "강서구 화곡동" "관악구 봉천
##  $ 충전환경              : num  41168 49357 105217 107611 117875 ...
##  $ 급속충전기결핍률      : num  1 1 1 1 1 ...
##  $ 공공자전거거치대결핍률: num  0.998 0.998 0.999 0.998 0.998 ...
##  $ 택배함결핍률          : num  1 1 1 1 1 ...
##  $ 물류창고결핍률        : num  1 1 1 1 1 ...
```

```
set.seed(123)

final_pca2 <- prcomp(final[-1],scale=T)
summary(final_pca2)
```

```
## Importance of components:
##                           PC1    PC2    PC3     PC4     PC5
## Standard deviation     1.4149 1.2011 0.8928 0.66912 0.55735
## Proportion of Variance 0.4004 0.2885 0.1594 0.08954 0.06213
## Cumulative Proportion  0.4004 0.6889 0.8483 0.93787 1.00000
```

```
loadings <- final_pca2$rotation
round(loadings,2)
```

```
##                         PC1   PC2   PC3   PC4   PC5
## 충전환경               0.50 -0.48  0.08 -0.32  0.65
## 급속충전기결핍률       0.46  0.22 -0.72  0.46  0.13
## 공공자전거거치대결핍률 0.52  0.40 -0.03 -0.64 -0.41
## 택배함결핍률           0.39 -0.60  0.15  0.32 -0.60
## 물류창고결핍률         0.36  0.45  0.68  0.43  0.18
```

```
final0 <- final%>%
  select(법정동)

final0$충전환경 <- final$충전환경*0.08
final0$급속충전기 <- final$급속충전기결핍률*0.72
final0$공공자전거거치 <- final$공공자전거거치대결핍률*0.03
final0$택배함 <- final$택배함결핍률*0.15
final0$물류창고 <- final$물류창고결핍률*0.68

final0$score <- final0$충전환경+final0$급속충전기+final0$공공자전거거치+final0$택배함+final0$
```

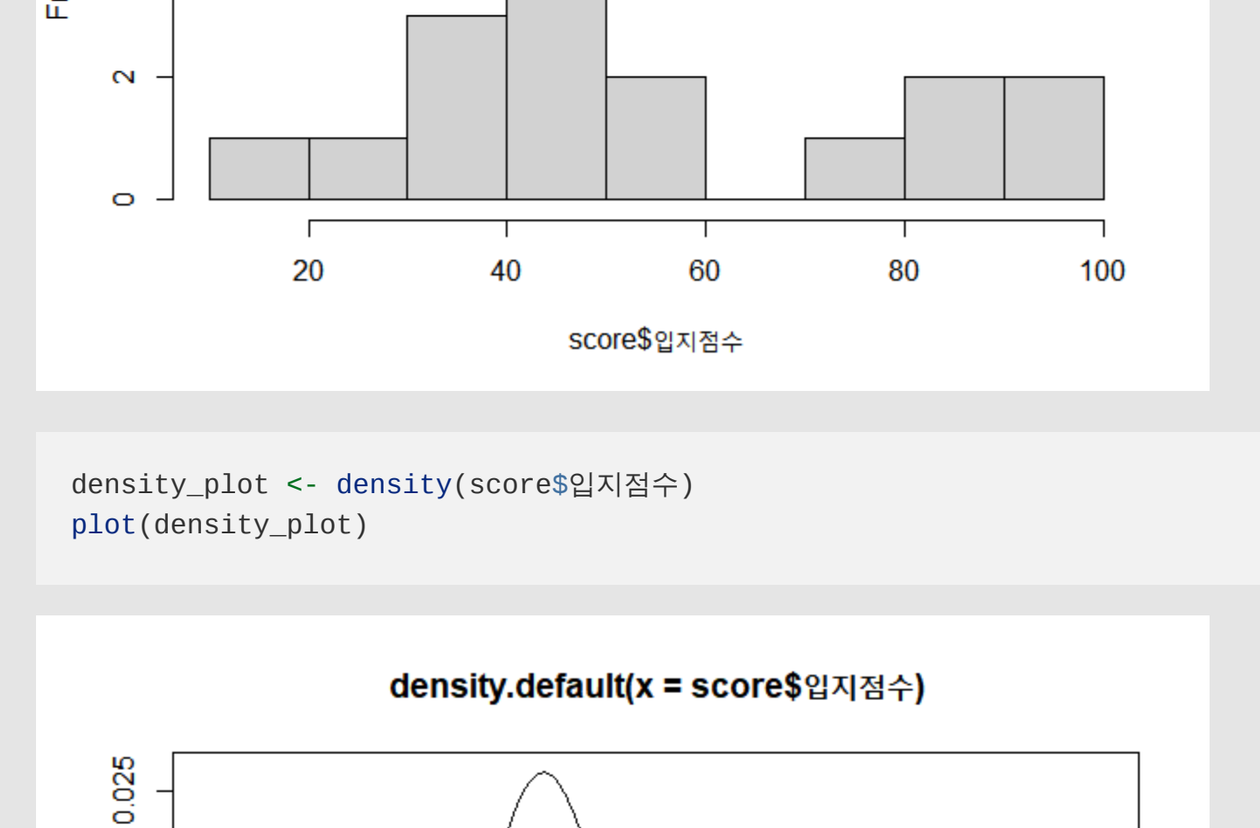
```
score <- final0%>%
  select("법정동","score")
score$입지점수 <- (score$score/max(score$score))*100
```

```
quantile(score$입지점수,probs=seq(0,1,0.01))
```

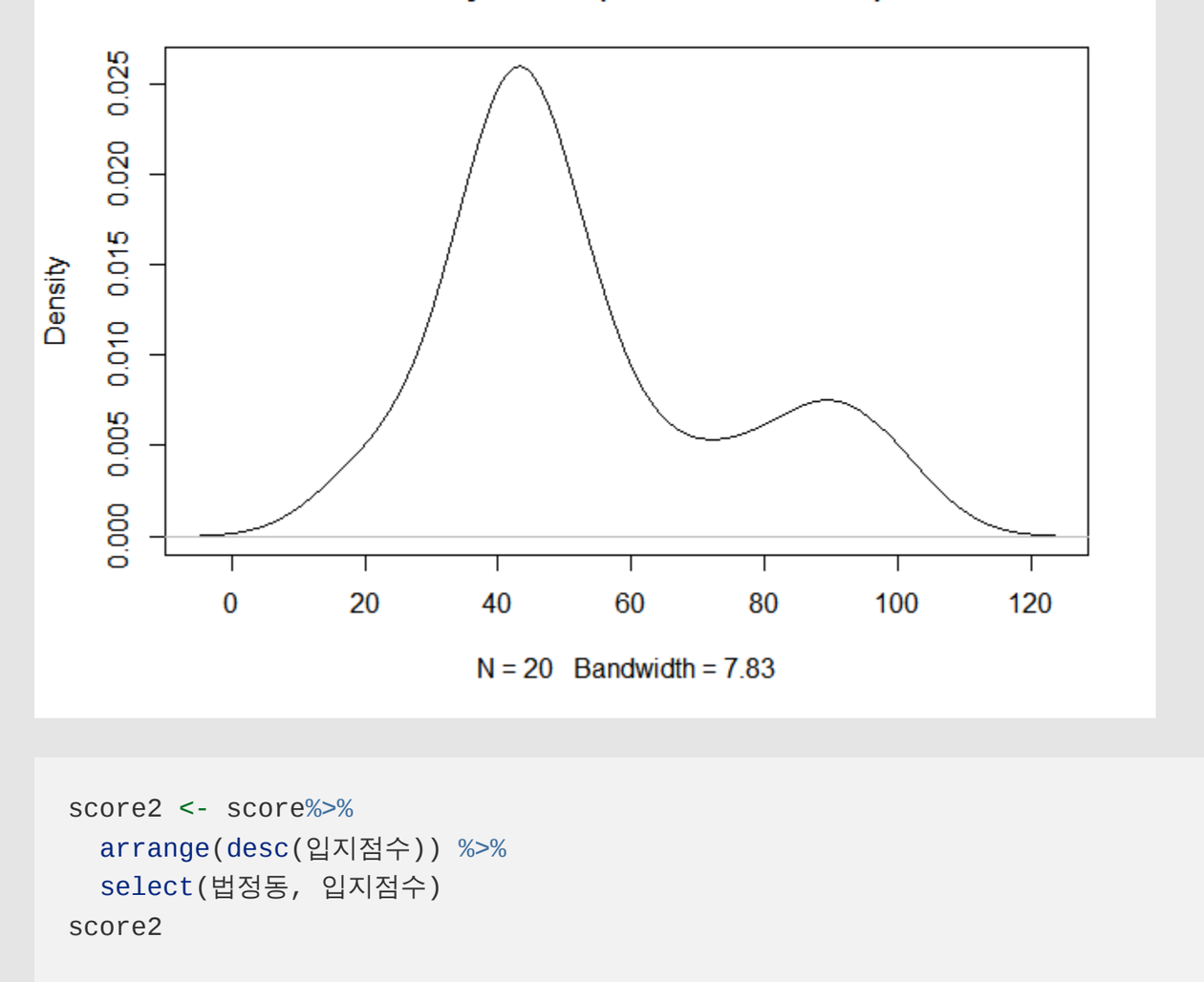
```
##        0%        1%        2%        3%        4%        5%        6%        7% 
##  18.73310  20.44207  22.15105  23.86002  25.56899  27.27797  28.73682  30.10635 
##        8%        9%       10%       11%       12%       13%       14%       15% 
##  31.47588  32.84541  34.21494  35.13791  35.56469  35.99147  36.41825  36.84502 
##       16%       17%       18%       19%       20%       21%       22%       23% 
##  37.28206  37.75758  38.23309  38.70861  39.18412  39.65964  40.00470  40.34252 
##       24%       25%       26%       27%       28%       29%       30%       31% 
##  40.68034  41.01816  41.35597  41.51722  41.59698  41.67673  41.75649  41.83624 
##       32%       33%       34%       35%       36%       37%       38%       39% 
##  41.89001  41.90803  41.92606  41.94408  41.96210  41.99223  42.08692  42.18160 
##       40%       41%       42%       43%       44%       45%       46%       47% 
##  42.27629  42.37097  42.46566  43.19352  43.99586  44.79821  45.60056  46.40291 
##       48%       49%       50%       51%       52%       53%       54%       55% 
##  46.73227  46.78572  46.83918  46.89263  46.94608  47.00230  47.06326  47.12422 
##       56%       57%       58%       59%       60%       61%       62%       63% 
##  47.18517  47.24613  47.32804  47.58804  47.84804  48.10803  48.36803  48.62803 
##       64%       65%       66%       67%       68%       69%       70%       71% 
##  49.73820  51.00779  52.27737  53.54696  54.81654  55.80219  56.58133  57.36047 
##       72%       73%       74%       75%       76%       77%       78%       79% 
##  58.13961  58.91875  60.12173  62.24301  64.36430  66.48559  68.60687  70.71481 
##       80%       81%       82%       83%       84%       85%       86%       87% 
##  72.58252  74.45023  76.31794  78.18565  80.05336  81.76907  83.44423  85.11940 
##       88%       89%       90%       91%       92%       93%       94%       95% 
##  86.79457  88.46974  89.46632  89.85218  90.23804  90.62389  91.00975  91.72936 
##       96%       97%       98%       99%      100% 
##  93.38349  95.03762  96.69175  98.34587 100.00000 
```

```
# 상위 10$
```

```
hist(score$입지점수)
```




```
density_plot <- density(score$입지점수)
plot(density_plot)
```




```
score2 <- score%>%
  arrange(desc(입지점수)) %>%
  select(법정동, 입지점수)
score2
```

```
##               법정동  입지점수
## 1      관악구 신림동 100.00000
## 2      관악구 봉천동  91.29407
## 3      강서구 화곡동  89.26324
## 4      노원구 상계동  80.44657
## 5      양천구 신정동  70.61651
## 6        양천구 목동  59.45185
## 7      중랑구 면목동  55.35111
## 8      양천구 신월동  48.66908
## 9        도봉구 창동  47.30067
## 10     구로구 구로동  46.97984
## 11     동작구 상도동  46.69851
## 12   영등포구 신길동  42.47563
## 13     금천구 독산동  41.97728
## 14     강북구 수유동  41.88242
## 15     금천구 시흥동  41.46265
## 16     동작구 사당동  39.68466
## 17     광진구 자양동  37.18196
## 18     강동구 천호동  34.93575
## 19       중구 신당동  27.72770
## 20 서대문구 남가좌동  18.73310
```

```
score <- score %>%
  filter(입지점수 >= 80) %>%
  arrange(desc(입지점수)) %>%
  select(법정동, 입지점수)
score
```

```
##          법정동  입지점수
## 1 관악구 신림동 100.00000
## 2 관악구 봉천동  91.29407
## 3 강서구 화곡동  89.26324
## 4 노원구 상계동  80.44657
```