PCA

```
# 표준 패키지 state.x77
# 미국 50개주에 대한 정보가 들어있음
str(state.x77)
```

```
##   num [1:50, 1:8] 3615 365 2212 2110 21198 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : chr [1:50] "Alabama" "Alaska" "Arizona" "Arkansas" ...
##   ..$ : chr [1:8] "Population" "Income" "Illiteracy" "Life Exp" ...
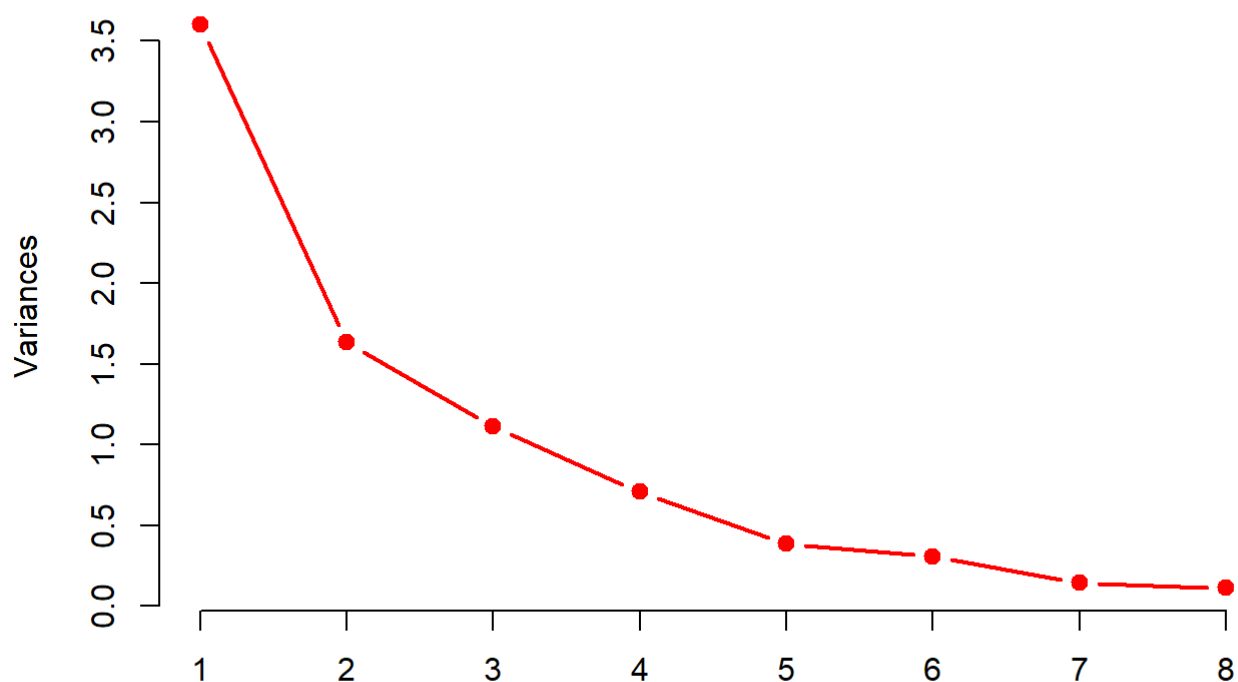```

```
head(state.x77)
```

```
##            Population Income Illiteracy Life Exp Murder HS Grad Frost    Area
## Alabama          3615   3624        2.1    69.05   15.1    41.3    20   50708
## Alaska            365   6315        1.5    69.31   11.3    66.7   152  566432
## Arizona          2212   4530        1.8    70.55    7.8    58.1    15  113417
## Arkansas         2110   3378        1.9    70.66   10.1    39.9    65   51945
## California      21198   5114        1.1    71.71   10.3    62.6    20  156361
## Colorado         2541   4884        0.7    72.06    6.8    63.9   166  103766
```

```
# 측정단위 통일하기 위해 scale=TRUE로 지정
pca <- prcomp(state.x77, scale=TRUE)

# x축 성분, y축 분산
plot(pca, type="l", pch=19, lwd=2, col="red", main="scree plot")
```

scree plot


```
round(pca$rotation, 3)
```

```
##                PC1     PC2     PC3     PC4     PC5     PC6     PC7     PC8
## Population   0.126   0.411  -0.656  -0.409   0.406  -0.011  -0.062  -0.219
## Income      -0.299   0.519  -0.100  -0.088  -0.638   0.462   0.009   0.060
## Illiteracy   0.468   0.053   0.071   0.353   0.004   0.387  -0.620  -0.339
## Life Exp    -0.412  -0.082  -0.360   0.443   0.327   0.219  -0.256   0.527
## Murder       0.444   0.307   0.108  -0.166  -0.128  -0.325  -0.295   0.678
## HS Grad     -0.425   0.299   0.050   0.232  -0.099  -0.645  -0.393  -0.307
## Frost       -0.357  -0.154   0.387  -0.619   0.217   0.213  -0.472   0.028
## Area        -0.033   0.588   0.510   0.201   0.499   0.148   0.286   0.013
```

```
# 표준화된 성분 * 성분적재값
round(scale(state.x77) %*% pca$rotation, 3)
```

```
##                    PC1    PC2    PC3    PC4    PC5    PC6    PC7    PC8
## Alabama          3.790 -0.235  0.229  0.383 -0.248 -0.434  0.057  0.535
## Alaska          -1.053  5.456  4.241  0.576  0.109  1.001  0.302 -0.118
## Arizona          0.867  0.745  0.077  1.719 -0.560 -0.305  0.131 -0.525
## Arkansas         2.382 -1.288  0.223  0.623  0.647  0.259  0.034  0.488
## California       0.241  3.510 -2.806 -0.070  0.969 -0.651  0.045 -0.257
## Colorado        -2.062  0.506  0.511 -0.110  0.002 -0.266 -0.644  0.375
## Connecticut     -1.899 -0.243 -0.663  0.170 -0.789  1.078 -0.538 -0.073
## Delaware        -0.425 -0.508  0.219 -0.193 -1.307 -0.157  0.292 -0.240
## Florida          1.172  1.135 -1.282  0.488 -0.676 -0.341  0.411  0.257
## Georgia          3.294  0.110  0.387 -0.456 -0.477  0.105  0.036  0.200
## Hawaii          -0.487  0.125 -1.377  2.950 -1.068  0.175 -0.808  0.263
## Idaho           -1.423 -0.611  0.434  0.406  0.408 -0.649  0.133  0.232
## Illinois         0.119  1.282 -0.804 -1.583 -0.334  0.045 -0.135  0.151
## Indiana         -0.471 -0.245 -0.301 -0.656 -0.045 -0.232  0.220  0.173
## Iowa            -2.322 -0.537 -0.293  0.205  0.246  0.093  0.129 -0.020
## Kansas          -1.902 -0.077 -0.177  0.614  0.111 -0.135  0.136  0.324
## Kentucky         2.129 -1.064  0.252 -0.349  0.333  0.394  0.148  0.564
## Louisiana        4.241 -0.346  0.229  0.879 -0.227 -0.044 -0.442 -0.374
## Maine           -0.960 -1.702  0.721 -0.545  0.505 -0.481  0.254 -0.742
## Maryland        -0.203  0.389 -0.339 -0.662 -1.468  0.216  0.192  0.208
## Massachusetts   -1.196 -0.219 -1.020  0.290 -0.256  0.160 -0.262 -0.598
## Michigan         0.182  0.847 -0.554 -1.183 -0.068 -0.230 -0.257  0.549
## Minnesota       -2.434 -0.365 -0.272  0.067  0.628  0.489 -0.098  0.102
## Mississippi      4.032 -1.051  0.920  0.176  0.150 -0.423 -0.116 -0.448
## Missouri         0.311 -0.148  0.007 -0.554  0.179 -0.218  0.420  0.607
## Montana         -1.379 -0.034  1.340 -0.245  0.349 -0.412  0.374 -0.275
## Nebraska        -2.181 -0.548  0.056  0.456  0.352  0.031  0.052  0.094
## Nevada          -1.127  1.133  1.891 -1.505 -1.240 -1.101 -0.449  0.244
## New Hampshire   -1.671 -1.312  0.438 -0.480 -0.052 -0.172 -0.275 -0.339
## New Jersey      -0.650  0.281 -0.974 -0.616 -0.781  0.732 -0.074 -0.395
## New Mexico       1.322 -0.294  1.362  0.707  0.709 -0.276 -1.168 -0.358
## New York         1.050  1.894 -2.198 -1.267  0.349  0.001 -0.491 -0.243
## North Carolina   2.694 -0.517  0.143 -0.561  0.108  0.404  0.215  0.103
## North Dakota    -2.418 -0.782  0.277 -0.136  0.031  1.655 -0.058  0.247
## Ohio            -0.268  0.417 -1.033 -1.148  0.350 -0.147  0.015 -0.117
## Oklahoma         0.074 -0.647 -0.071  0.607  0.409 -0.182  0.331  0.148
## Oregon          -1.325  0.228 -0.499  1.350 -0.189 -0.458  0.924  0.052
## Pennsylvania     0.077  0.269 -1.070 -1.289  0.588  0.198  0.114 -0.571
## Rhode Island    -0.741 -1.461 -0.228  0.297 -0.237  1.294  0.007 -0.154
## South Carolina   3.711 -0.910  0.749 -0.316 -0.363  0.261  0.046 -0.453
## South Dakota    -2.013 -1.315  0.537 -0.157  0.751  0.347  0.346 -0.020
## Tennessee        2.218 -0.651 -0.017  0.003  0.145  0.141  0.075  0.419
## Texas            2.414  2.327 -0.286  0.804  1.725  0.532 -0.001  0.114
## Utah            -2.263 -0.534  0.220  0.851  0.769 -1.062 -0.488  0.171
## Vermont         -1.369 -1.509  0.446 -0.359  0.310 -0.626 -0.348  0.278
## Virginia         0.994  0.185 -0.211 -0.324 -0.672  0.317  0.074  0.125
## Washington      -1.340  0.512 -0.852  1.237 -0.654 -0.764  0.819 -0.278
## West Virginia    1.507 -1.602  0.492 -0.342  0.162  0.187  0.553 -0.334
## Wisconsin       -1.758 -0.636 -0.425 -0.112  0.611  0.415 -0.009  0.041
## Wyoming         -1.484  0.042  1.354 -0.639 -0.296 -0.764 -0.224 -0.131
```

```
round(pca$x, 3)
```

```
##                   PC1     PC2     PC3     PC4     PC5     PC6     PC7     PC8
## Alabama         3.790  -0.235   0.229   0.383  -0.248  -0.434   0.057   0.535
## Alaska         -1.053   5.456   4.241   0.576   0.109   1.001   0.302  -0.118
## Arizona         0.867   0.745   0.077   1.719  -0.560  -0.305   0.131  -0.525
## Arkansas        2.382  -1.288   0.223   0.623   0.647   0.259   0.034   0.488
## California      0.241   3.510  -2.806  -0.070   0.969  -0.651   0.045  -0.257
## Colorado       -2.062   0.506   0.511  -0.110   0.002  -0.266  -0.644   0.375
## Connecticut    -1.899  -0.243  -0.663   0.170  -0.789   1.078  -0.538  -0.073
## Delaware       -0.425  -0.508   0.219  -0.193  -1.307  -0.157   0.292  -0.240
## Florida         1.172   1.135  -1.282   0.488  -0.676  -0.341   0.411   0.257
## Georgia         3.294   0.110   0.387  -0.456  -0.477   0.105   0.036   0.200
## Hawaii         -0.487   0.125  -1.377   2.950  -1.068   0.175  -0.808   0.263
## Idaho          -1.423  -0.611   0.434   0.406   0.408  -0.649   0.133   0.232
## Illinois        0.119   1.282  -0.804  -1.583  -0.334   0.045  -0.135   0.151
## Indiana        -0.471  -0.245  -0.301  -0.656  -0.045  -0.232   0.220   0.173
## Iowa           -2.322  -0.537  -0.293   0.205   0.246   0.093   0.129  -0.020
## Kansas         -1.902  -0.077  -0.177   0.614   0.111  -0.135   0.136   0.324
## Kentucky        2.129  -1.064   0.252  -0.349   0.333   0.394   0.148   0.564
## Louisiana       4.241  -0.346   0.229   0.879  -0.227  -0.044  -0.442  -0.374
## Maine          -0.960  -1.702   0.721  -0.545   0.505  -0.481   0.254  -0.742
## Maryland       -0.203   0.389  -0.339  -0.662  -1.468   0.216   0.192   0.208
## Massachusetts  -1.196  -0.219  -1.020   0.290  -0.256   0.160  -0.262  -0.598
## Michigan        0.182   0.847  -0.554  -1.183  -0.068  -0.230  -0.257   0.549
## Minnesota      -2.434  -0.365  -0.272   0.067   0.628   0.489  -0.098   0.102
## Mississippi     4.032  -1.051   0.920   0.176   0.150  -0.423  -0.116  -0.448
## Missouri        0.311  -0.148   0.007  -0.554   0.179  -0.218   0.420   0.607
## Montana        -1.379  -0.034   1.340  -0.245   0.349  -0.412   0.374  -0.275
## Nebraska       -2.181  -0.548   0.056   0.456   0.352   0.031   0.052   0.094
## Nevada         -1.127   1.133   1.891  -1.505  -1.240  -1.101  -0.449   0.244
## New Hampshire  -1.671  -1.312   0.438  -0.480  -0.052  -0.172  -0.275  -0.339
## New Jersey     -0.650   0.281  -0.974  -0.616  -0.781   0.732  -0.074  -0.395
## New Mexico      1.322  -0.294   1.362   0.707   0.709  -0.276  -1.168  -0.358
## New York        1.050   1.894  -2.198  -1.267   0.349   0.001  -0.491  -0.243
## North Carolina  2.694  -0.517   0.143  -0.561   0.108   0.404   0.215   0.103
## North Dakota   -2.418  -0.782   0.277  -0.136   0.031   1.655  -0.058   0.247
## Ohio           -0.268   0.417  -1.033  -1.148   0.350  -0.147   0.015  -0.117
## Oklahoma        0.074  -0.647  -0.071   0.607   0.409  -0.182   0.331   0.148
## Oregon         -1.325   0.228  -0.499   1.350  -0.189  -0.458   0.924   0.052
## Pennsylvania    0.077   0.269  -1.070  -1.289   0.588   0.198   0.114  -0.571
## Rhode Island   -0.741  -1.461  -0.228   0.297  -0.237   1.294   0.007  -0.154
## South Carolina  3.711  -0.910   0.749  -0.316  -0.363   0.261   0.046  -0.453
## South Dakota   -2.013  -1.315   0.537  -0.157   0.751   0.347   0.346  -0.020
## Tennessee       2.218  -0.651  -0.017   0.003   0.145   0.141   0.075   0.419
## Texas           2.414   2.327  -0.286   0.804   1.725   0.532  -0.001   0.114
## Utah           -2.263  -0.534   0.220   0.851   0.769  -1.062  -0.488   0.171
## Vermont        -1.369  -1.509   0.446  -0.359   0.310  -0.626  -0.348   0.278
## Virginia        0.994   0.185  -0.211  -0.324  -0.672   0.317   0.074   0.125
## Washington     -1.340   0.512  -0.852   1.237  -0.654  -0.764   0.819  -0.278
## West Virginia   1.507  -1.602   0.492  -0.342   0.162   0.187   0.553  -0.334
## Wisconsin      -1.758  -0.636  -0.425  -0.112   0.611   0.415  -0.009   0.041
## Wyoming        -1.484   0.042   1.354  -0.639  -0.296  -0.764  -0.224  -0.131
```

```
# 설명력이 큰 상위 2개만 추출해서 데이터의 복잡성 줄이기
round(pca$x[, c(1,2)], 3)
```

```
##                 PC1    PC2
## Alabama         3.790 -0.235
## Alaska         -1.053  5.456
## Arizona         0.867  0.745
## Arkansas        2.382 -1.288
## California      0.241  3.510
## Colorado       -2.062  0.506
## Connecticut    -1.899 -0.243
## Delaware       -0.425 -0.508
## Florida         1.172  1.135
## Georgia         3.294  0.110
## Hawaii         -0.487  0.125
## Idaho          -1.423 -0.611
## Illinois        0.119  1.282
## Indiana        -0.471 -0.245
## Iowa           -2.322 -0.537
## Kansas         -1.902 -0.077
## Kentucky        2.129 -1.064
## Louisiana       4.241 -0.346
## Maine          -0.960 -1.702
## Maryland       -0.203  0.389
## Massachusetts  -1.196 -0.219
## Michigan        0.182  0.847
## Minnesota      -2.434 -0.365
## Mississippi     4.032 -1.051
## Missouri        0.311 -0.148
## Montana        -1.379 -0.034
## Nebraska       -2.181 -0.548
## Nevada         -1.127  1.133
## New Hampshire  -1.671 -1.312
## New Jersey     -0.650  0.281
## New Mexico      1.322 -0.294
## New York        1.050  1.894
## North Carolina  2.694 -0.517
## North Dakota   -2.418 -0.782
## Ohio           -0.268  0.417
## Oklahoma        0.074 -0.647
## Oregon         -1.325  0.228
## Pennsylvania    0.077  0.269
## Rhode Island   -0.741 -1.461
## South Carolina  3.711 -0.910
## South Dakota   -2.013 -1.315
## Tennessee       2.218 -0.651
## Texas           2.414  2.327
## Utah           -2.263 -0.534
## Vermont        -1.369 -1.509
## Virginia        0.994  0.185
## Washington     -1.340  0.512
## West Virginia   1.507 -1.602
## Wisconsin      -1.758 -0.636
## Wyoming        -1.484  0.042
```

```
# 주성분분석의 시각화
biplot(pca, cex=c(0.5, 0.75), main="Biplot")
```

Biplot