

# assignment

*sn0wfree*

11/16/2016

## Contents

```
#assig<-read.csv( "States facts.csv" ,header=TRUE)
all<-read.csv( "all.csv" ,header=1)
head(all)
```

```
##      state  pop2010  intuse2010  faceuse2010  pop2015  landaream2
## 1  Alaska   705813    593193      336440    740405    570641
## 2  Arizona  6665093   5230474   2448140   6826302    113594
## 3  Arkansas 2919815   1949869    989820   2977410     52035
## 4  California 37350092 29758896  16673720 39141723    155779
## 5  Colorado 5077553   4058749   2369420   5450364    103642
## 6 Connecticut 3555261  3074229   1398220   3590615     4842
##      landareak2  vehiclesper1000  ginicoeff  povrate1  npov000s  rovrates2  povrate3
## 1  1477953.4          960      0.422    0.114      81      0.121    0.125
## 2  294207.1          660      0.455    0.182     1195    0.213    0.188
## 3  134770.0          700      0.458    0.187     539    0.191    0.165
## 4  403465.8          840      0.471    0.164    6253    0.155    0.238
## 5  268431.5          340      0.457    0.121     632    0.124    0.137
## 6  12540.7          860      0.486    0.108     376    0.106    0.125
##      pop2010x  densitym2  murd  gunmurd  gunownpc  murdpop2010x  v_relig  emitco2
## 1  710231      1.264    31      19      0.617    4.36e-05     28      35
## 2  6392017     57.050   352     232    0.323    5.51e-05     35     93
## 3  2915918     56.430   130      93    0.579    4.46e-05     54     69
## 4  37253956    244.200 1811    1257    0.201    4.86e-05     35    359
## 5  5029196     49.330   117      65    0.343    2.33e-05     33     91
## 6  3574097     741.400  131      97    0.166    3.67e-05     31     35
##      pop2014  emitrate  obesead  overwead  obesech  phigh  pbatdeg  paddeg  ninjail
## 1  736732     47.17    0.273    0.645    0.111  0.914    0.266    0.090    5100
## 2  6731484    13.79    0.233    0.595    0.122  0.842    0.256    0.093    55200
## 3  2966369    23.13    0.281    0.647    0.164  0.824    0.189    0.061    22800
## 4  38802500    9.26    0.231    0.594    0.132  0.806    0.299    0.107    218800
## 5  5355866    16.95    0.210    0.550    0.099  0.893    0.359    0.127    32100
## 6  3596677     9.77    0.208    0.587    0.123  0.886    0.356    0.155    17600
##      njailper100k  pcrenew  pcrenewnohydro  renelec  renelecnohydro  totalelec
## 1      940      29.00          3.57    1759          217    6073
## 2     1090      9.50          3.70    10734         4199    113351
## 3     1010      9.30          2.69    5159         1496    55682
## 4      750     30.50         23.50    60359        46498    197994
## 5      790     17.80         15.00    9354         7862    52515
## 6      620      3.42          2.20    1289          829    37649
##      pcunion  nunion  workpop  right2work  nbillionaires  collectivism
## 1     19.6    60000    304000          0          0          48
## 2      5.2   138000    2661000          1          9          49
## 3      5.1   58000    1155000          1          5          54
## 4     15.9  2486000   15657000          0        124          60
```

```
## 5      8.4 194000 2310000      0      10      36
## 6     17.0 269000 1587000      0      12      50
##   lnfirmsspercapita landwatertotalarea waterarea unemploymentrates
## 1      -1.964075      663267.26 91316.00      6.5
## 2      -2.285662      113998.30  363.73      6.1
## 3      -2.113070      53178.62  1110.45      5.2
## 4      -2.067482      163695.57  7736.23      6.2
## 5      -1.899379      104093.57   376.04      3.9
## 6      -1.994490      5543.33   698.53      5.6
```

```
variables_name<-variable.names(all)
colnames(all)
```

```
## [1] "state"           "pop2010"         "intuse2010"
## [4] "faceuse2010"     "pop2015"         "landaream2"
## [7] "landareak2"      "vehiclesper1000"  "ginicoeff"
## [10] "povrate1"        "npov000s"        "rovrage2"
## [13] "povrate3"        "pop2010x"        "densitym2"
## [16] "murd"            "gunmurd"         "gunownpc"
## [19] "murdpop2010x"    "v_relig"         "emitco2"
## [22] "pop2014"         "emitrate"        "obesead"
## [25] "overwead"        "obesech"         "phigh"
## [28] "pbatdeg"         "paddeg"          "ninjail"
## [31] "njailper100k"    "pcrenew"         "pcrenewnohydro"
## [34] "renelec"         "renelecnohydro"  "totalelec"
## [37] "pcunion"         "nunion"          "workpop"
## [40] "right2work"      "nbillionaires"   "collectivism"
## [43] "lnfirmsspercapita" "landwatertotalarea" "waterarea"
## [46] "unemploymentrates"
```

```
rownames(all)<-all$state
str(all)
```

```
## 'data.frame':   49 obs. of  46 variables:
## $ state          : Factor w/ 49 levels "Alaska","Arizona",...: 1 2 3 4 5 6 7 8 9 10 ...
## $ pop2010        : int  705813 6665093 2919815 37350092 5077553 3555261 894424 18732783 9932505 1308789 ...
## $ intuse2010     : int  593193 5230474 1949869 29758896 4058749 3074229 719500 14764418 7597608 1081506 ...
## $ faceuse2010    : int  336440 2448140 989820 16673720 2369420 1398220 216140 7839520 4841900 597100 ...
## $ pop2015        : int  740405 6826302 2977410 39141723 5450364 3590615 945913 20267012 10216414 1432002 ...
## $ landaream2     : int  570641 113594 52035 155779 103642 4842 1949 53625 57513 6423 ...
## $ landareak2     : num  1477953 294207 134770 403466 268432 ...
## $ vehiclesper1000 : int  960 660 700 840 340 860 950 710 820 760 ...
## $ ginicoeff      : num  0.422 0.455 0.458 0.471 0.457 0.486 0.44 0.474 0.468 0.433 ...
## $ povrate1       : num  0.114 0.182 0.187 0.164 0.121 0.108 0.13 0.166 0.184 0.115 ...
## $ npov000s       : int  81 1195 539 6253 632 376 118 3231 1298 158 ...
## $ rovrage2       : num  0.121 0.213 0.191 0.155 0.124 0.106 0.124 0.146 0.185 0.126 ...
## $ povrate3       : num  0.125 0.188 0.165 0.238 0.137 0.125 0.139 0.195 0.182 0.173 ...
## $ pop2010x       : int  710231 6392017 2915918 37253956 5029196 3574097 897934 19687653 9920000 1360301 ...
## $ densitym2      : num  1.26 57.05 56.43 244.2 49.33 ...
## $ murd           : int  31 352 130 1811 117 131 48 987 527 24 ...
## $ gunmurd        : int  19 232 93 1257 65 97 38 669 376 7 ...
## $ gunownpc       : num  0.617 0.323 0.579 0.201 0.343 0.166 0.052 0.325 0.316 0.451 ...
## $ murdpop2010x   : num  4.36e-05 5.51e-05 4.46e-05 4.86e-05 2.33e-05 ...
## $ v_relig        : int  28 35 54 35 33 31 33 39 48 33 ...
## $ emitco2        : int  35 93 69 359 91 35 13 227 139 18 ...
## $ pop2014        : int  736732 6731484 2966369 38802500 5355866 3596677 935614 19893297 10097343 1419561 ...
```

```
## $ emitrate      : num  47.17 13.79 23.13 9.26 16.95 ...
## $ obesead       : num  0.273 0.233 0.281 0.231 0.21 0.208 0.259 0.233 0.275 0.207 ...
## $ overwead      : num  0.645 0.595 0.647 0.594 0.55 0.587 0.639 0.608 0.633 0.553 ...
## $ obesech       : num  0.111 0.122 0.164 0.132 0.099 0.123 0.228 0.144 0.164 0.133 ...
## $ phigh         : num  0.914 0.842 0.824 0.806 0.893 0.886 0.874 0.853 0.839 0.904 ...
## $ pbatdeg       : num  0.266 0.256 0.189 0.299 0.359 0.356 0.287 0.253 0.275 0.296 ...
## $ paddeg        : num  0.09 0.093 0.061 0.107 0.127 0.155 0.114 0.09 0.099 0.099 ...
## $ ninjail       : int   5100 55200 22800 218800 32100 17600 7000 154500 91600 5600 ...
## $ njailper100k   : int   940 1090 1010 750 790 620 960 990 1220 510 ...
## $ pcrenew       : num   29 9.5 9.3 30.5 17.8 3.42 1.86 2.2 5.9 13.7 ...
## $ pcrenewnohydro : num   3.57 3.7 2.69 23.5 15 2.2 1.86 2.1 3.53 12.6 ...
## $ renelec       : int  1759 10734 5159 60359 9354 1289 144 5215 7546 1358 ...
## $ renelecnohydro : int   217 4199 1496 46498 7862 829 144 4992 4525 1248 ...
## $ totalelec     : int  6073 113351 55682 197994 52515 37649 7729 237338 128259 9930 ...
## $ pcunion       : num   19.6 5.2 5.1 15.9 8.4 17 9.2 6.8 4 20.4 ...
## $ nunion        : int  60000 138000 58000 2486000 194000 269000 38000 546000 162000 119000 ...
## $ workpop       : int  304000 2661000 1155000 15657000 2310000 1587000 412000 7994000 4016000 583000 ...
## $ right2work     : int    0 1 1 0 0 0 0 1 1 0 ...
## $ nbillionaires  : int    0 9 5 124 10 12 0 44 9 1 ...
## $ collectivism   : int   48 49 54 60 36 50 55 54 60 91 ...
## $ lnfirmispcapita : num  -1.96 -2.29 -2.11 -2.07 -1.9 ...
## $ landwatertotalarea: num  663267 113998 53179 163696 104094 ...
## $ waterarea      : num   91316 364 1110 7736 376 ...
## $ unemploymentrates : num   6.5 6.1 5.2 6.2 3.9 5.6 4.9 5.4 5.9 3.6 ...
```

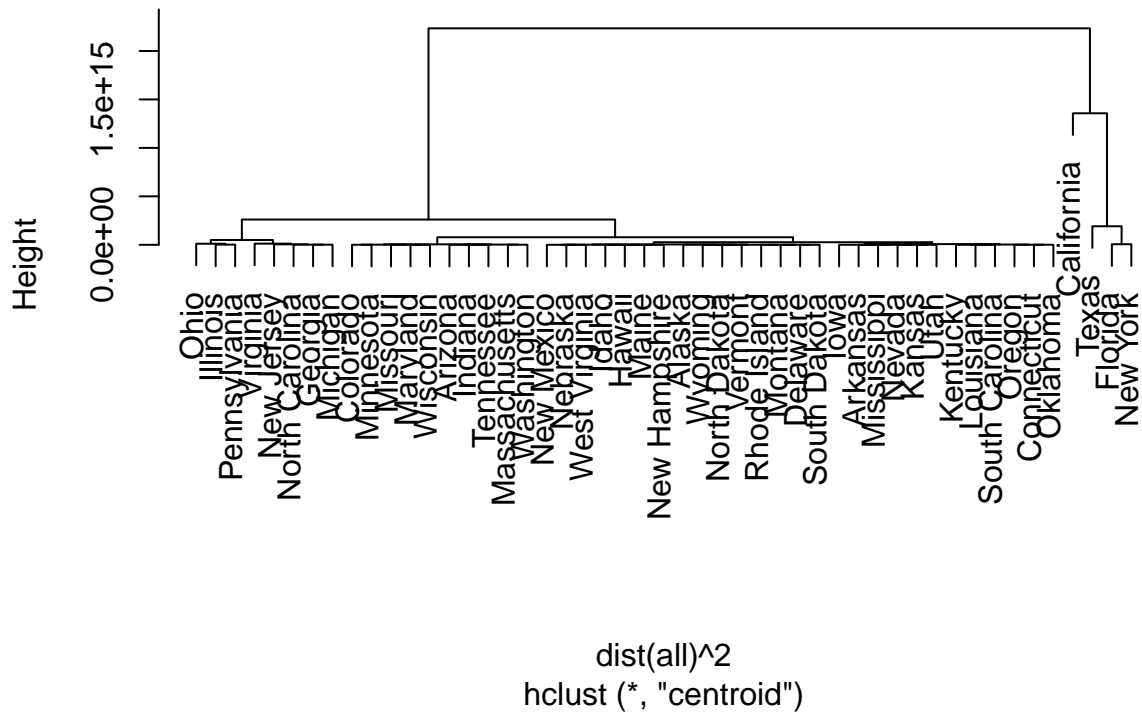
```
newall<-all
```

```
all$state<-NULL
```

```
hc<-hclust(dist(all)^2,method = "cen", members = NULL)
```

```
plot(hc)
```

## Cluster Dendrogram



```
data(UScitiesD)
str(UScitiesD)
```

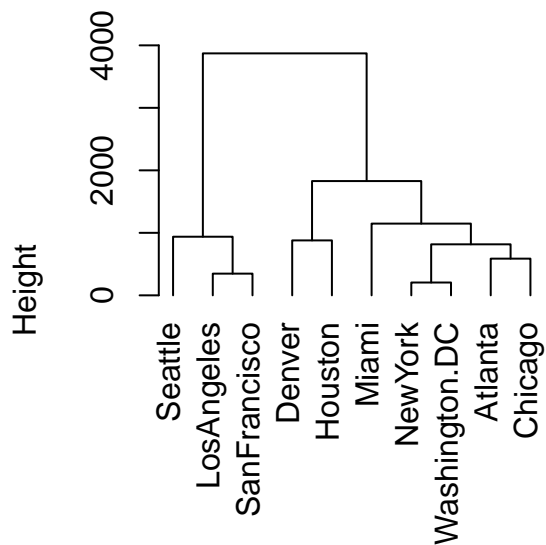
```
## Class 'dist' atomic [1:45] 587 1212 701 1936 604 748 2139 2182 543 920 ...
##   .. attr(*, "Labels")= chr [1:10] "Atlanta" "Chicago" "Denver" "Houston" ...
##   .. attr(*, "Size")= int 10
##   .. attr(*, "call")= language as.dist.default(m = t(cities.mat))
##   .. attr(*, "Diag")= logi FALSE
##   .. attr(*, "Upper")= logi FALSE
```

```
mds2 <- cmdscale(UScitiesD)
plot(mds2, type="n", axes=FALSE, ann=FALSE)
text(mds2, labels=rownames(mds2), xpd = NA)
```



```
hcity.D2 <- hclust(UScitiesD, "ward.D2")
opar <- par(mfrow = c(1, 2))
#plot(hcity.D, hang=-1)
plot(hcity.D2, hang=-1)
par(opar)
```

### Cluster Dendrogram



UScitiesD  
hclust (\*, "ward.D2")