DEMOGRAPHICS.R

Qian

Sun Oct 18 21:22:59 2015

```
setwd("/Users/Qian/Desktop/chsi_dataset")
degraphi <- read.csv("DEMOGRAPHICS.csv")
# DATA CLEANING
#is.na(degraphi)
# there is no missing value in the dataset
head(degraphi)
```

```
##   State_FIPS_Code County_FIPS_Code CHSI_County_Name CHSI_State_Name
## 1               1                1          Autauga         Alabama
## 2               1                3          Baldwin         Alabama
## 3               1                5          Barbour         Alabama
## 4               1                7             Bibb         Alabama
## 5               1                9           Blount         Alabama
## 6               1               11          Bullock         Alabama
##   CHSI_State_Abbr Strata_ID_Number
## 1              AL               29
## 2              AL               16
## 3              AL               51
## 4              AL               42
## 5              AL               28
## 6              AL               75
##                                      Strata_Determining_Factors
## 1                    frontier status, population size, poverty, age
## 2                    frontier status, population size, poverty, age
## 3 frontier status, population size, poverty, age, population density
## 4                    frontier status, population size, poverty, age
## 5                    frontier status, population size, poverty, age
## 6 frontier status, population size, poverty, age, population density
##   Number_Counties Population_Size Min_Population_Size Max_Population_Size
## 1              37           48612               28447               55936
## 2              27          162586              118395              277035
## 3              33           28414               27269               43226
## 4              53           21516                8134               24778
## 5              39           55725               29009               53844
## 6              37           11055                6228               19495
##   Population_Density Min_Population_Density Max_Population_Density Poverty
## 1                 82                    40                   141    10.4
## 2                102                    39                   457    10.2
## 3                 32                    14                    41    22.1
## 4                 35                     9                    66    16.8
## 5                 86                    30                   229    11.9
## 6                 18                    15                    22    26.2
##   Min_Poverty Max_Poverty Age_19_Under Min_Age_19_Under Max_Age_19_Under
## 1         9.5        12.9         26.9             23.7             32.3
## 2         9.7        12.9         23.5             21.3             25.4
## 3        18.0        24.6         24.3             23.5             32.2
## 4        12.5        16.4         24.6             24.4             32.4
```

```
## 5         9.4        13.4         24.5             21.8             26.1
## 6        17.0        24.9         24.7             22.3             28.6
##   Age_19_64 Min_Age_19_64 Max_Age_19_65 Age_65_84 Min_Age_65_84
## 1      62.3          58.8          64.1       9.8           7.3
## 2      60.3          55.3          62.0      14.5          11.8
## 3      62.5          54.9          62.5      11.6           9.6
## 4      63.3          55.8          63.2      10.9           9.2
## 5      62.1          61.0          66.2      12.1           8.8
## 6      63.2          55.2          63.8      10.0          11.0
##   Max_Age_65_85 Age_85_and_Over Min_Age_85_and_Over Max_Age_85_and_Over
## 1          12.0             0.9                 0.8                 2.1
## 2          19.5             1.8                 1.9                 3.4
## 3          12.6             1.6                 1.2                 1.8
## 4          13.3             1.2                 1.0                 2.1
## 5          13.5             1.3                 1.2                 2.1
## 6          15.1             2.2                 1.8                 2.6
##   White Min_White Max_White Black Min_Black Max_Black Native_American
## 1  80.7      80.7      98.5  17.3       0.4      17.3             0.5
## 2  88.4      83.5      96.3   9.9       1.0      14.1             0.5
## 3  52.2      48.6      97.0  46.8       0.7      50.7             0.4
## 4  76.8      63.9      97.6  22.5       0.3      35.4             0.3
## 5  97.1      78.1      97.6   1.5       0.6      20.5             0.5
## 6  27.8      30.0      97.1  71.4       1.1      69.6             0.4
##   Min_Native_American Max_Native_American Asian Min_Asian Max_Asian
## 1                 0.1                 1.1   0.6       0.2       2.2
## 2                 0.1                 1.1   0.4       0.4       3.3
## 3                 0.2                 7.1   0.3       0.2       2.3
## 4                 0.1                 2.6   0.1       0.1       1.5
## 5                 0.1                 0.9   0.2       0.3       2.2
## 6                 0.1                 1.6   0.2       0.1       0.5
##   Hispanic Min_Hispanic Max_Hispanic
## 1      1.7          0.8         19.2
## 2      2.3          0.8         13.7
## 3      3.1          1.0         67.7
## 4      1.4          1.2         46.9
## 5      6.3          0.8          6.3
## 6      5.9          0.7         37.8
```

```
class(degraphi)
```

```
## [1] "data.frame"
```

```
summary(degraphi$Population_Size)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      62   11210   25240   94370   64040 9935000
```

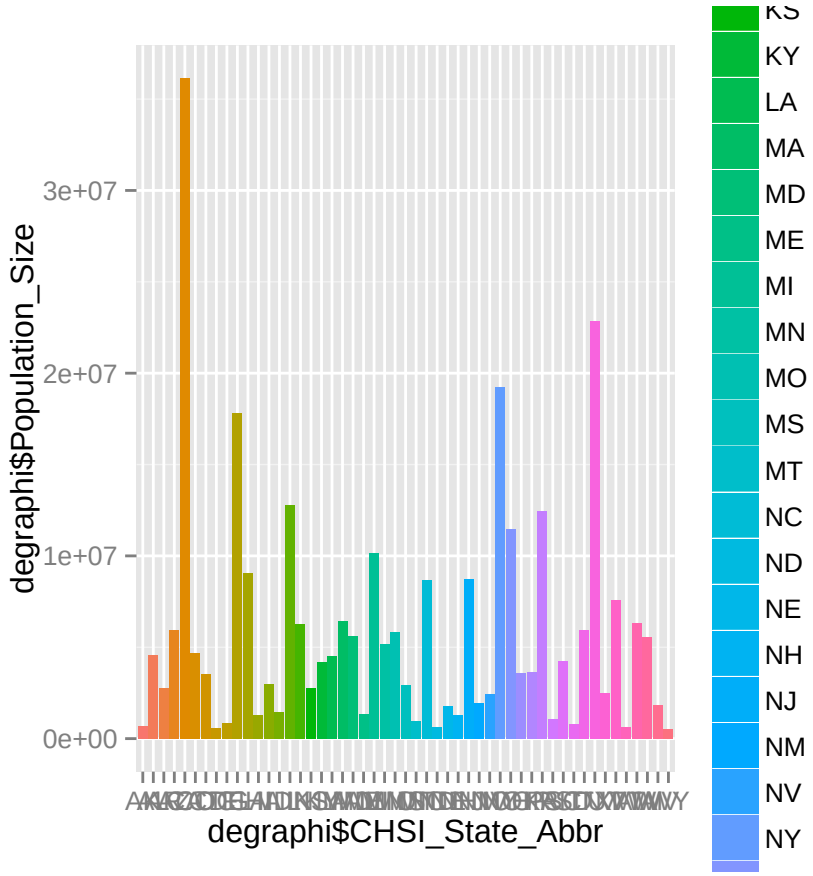
```
summary(degraphi$Poverty)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.40    9.80   12.60   13.35   16.20   36.20
```

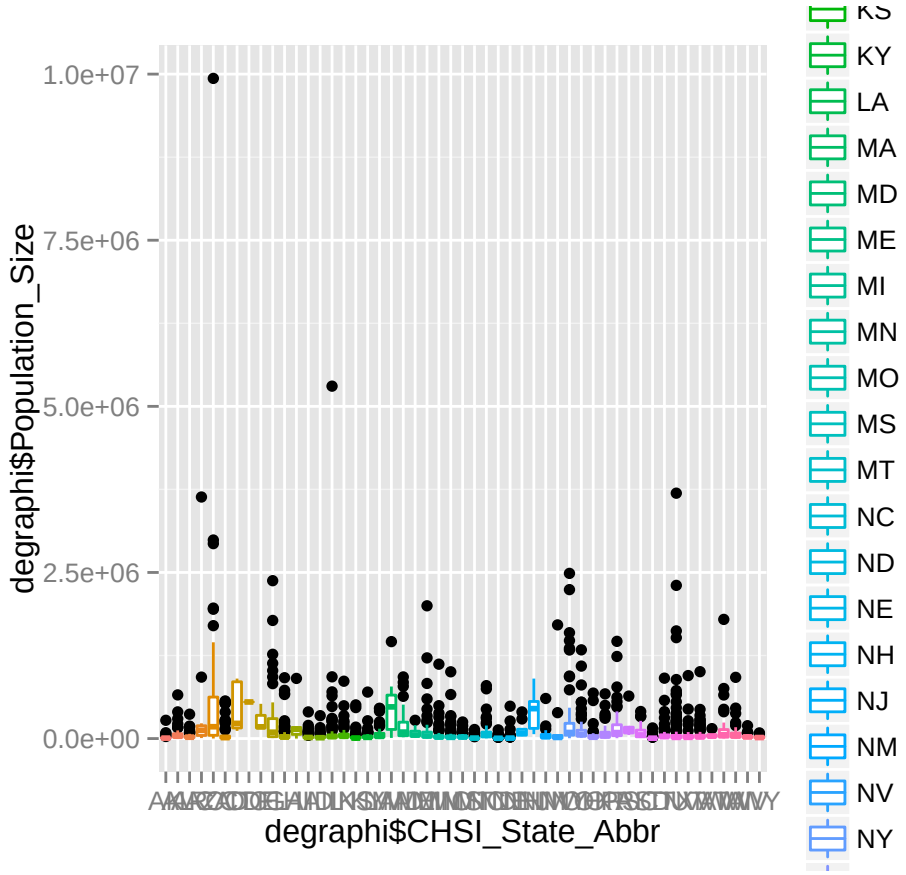
```r
# The % of age group 0-85 should covers 90% or above 90% of whole population.
#degraphi$Age_19_Under+degraphi$Age_19_64+degraphi$Age_65_84+degraphi$Age_85_and_Over > 90
# Age 19 to 64 should have the largest proportion.
#degraphi$Age_19_64 > degraphi$Age_19_Under
#degraphi$Age_19_64 > degraphi$Age_85_and_Over
#degraphi$Age_19_64 > degraphi$Age_65_84
# Age data is ok, no need to clean
# Next, check for race data
# the % of race, white, Black, Asian, Hispanic should be over 50% of whole population
#degraphi$White+degraphi$Black+degraphi$Asian+degraphi$Hispanic < 50
#degraphi$CHSI_County_Name[degraphi$White+degraphi$Black+degraphi$Asian+degraphi$Hispanic < 50]
#degraphi$CHSI_State_Name[degraphi$White+degraphi$Black+degraphi$Asian+degraphi$Hispanic < 50]
# Most of states are in the border, so it is reasonable to have many other races(like Latino) of people
#degraphi$CHSI_County_Name[degraphi$CHSI_State_Name=="Nebraska" & degraphi$White+degraphi$Black+degraph
# the race data has no need to clean
# Then we look at poverty data
summary(degraphi$Poverty)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.40    9.80   12.60   13.35   16.20   36.20
```

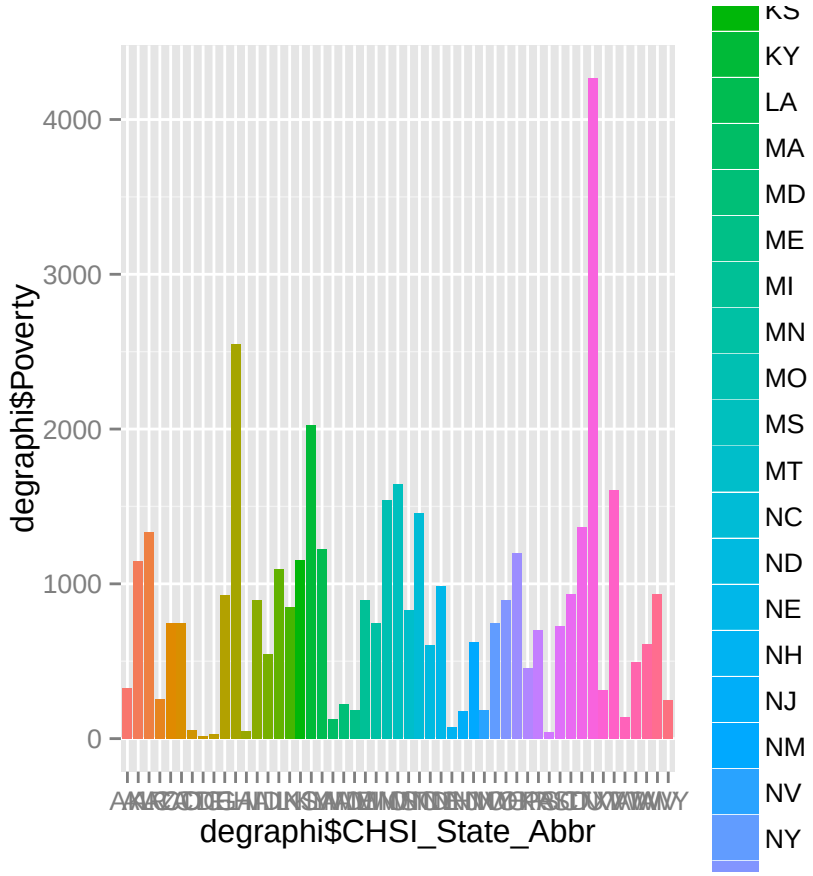
```r
#There is a extreme value of minimum poverty, -2222
#since this is a percentage data, it shouldn't be -2222. we clean it to NA.
degraphi$Poverty[degraphi$Poverty==-2222.2] <- 0


#EDA

library(ggplot2)
qplot(degraphi$CHSI_State_Abbr,degraphi$Population_Size,stat = "identity",geom = "bar",fill=factor(degra
```



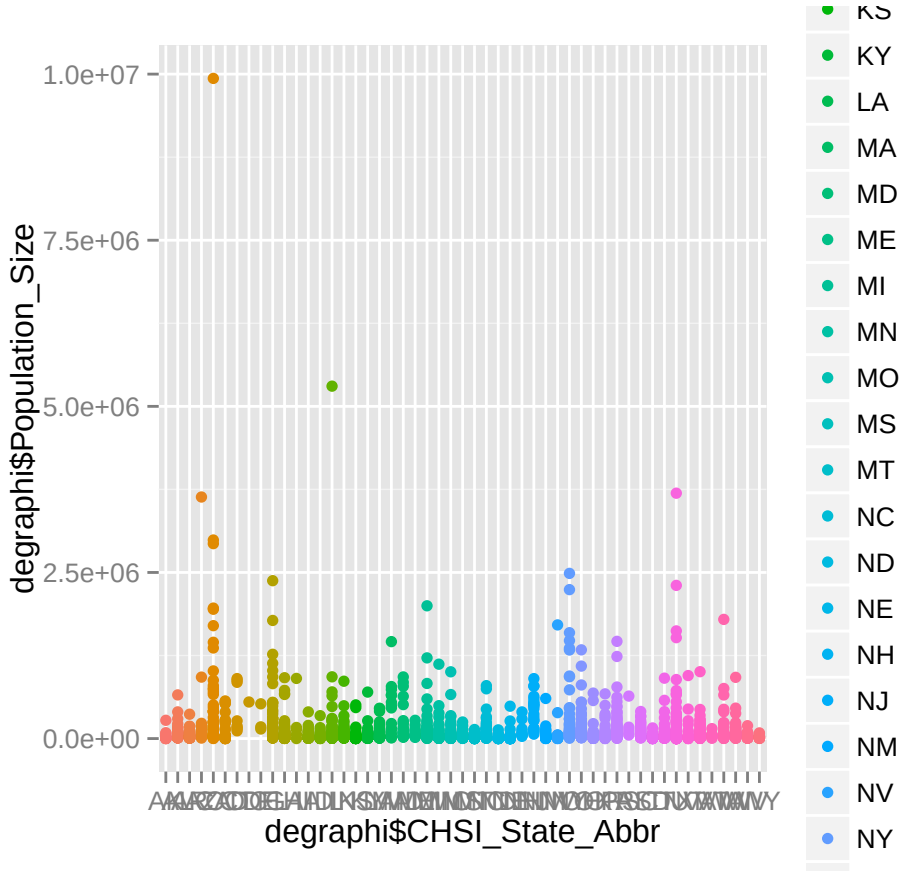

```
ggplot(data=degraphi, aes(degraphi$CHSI_State_Abbr,degraphi$Population_Size))+ geom_boxplot((aes(color=
```



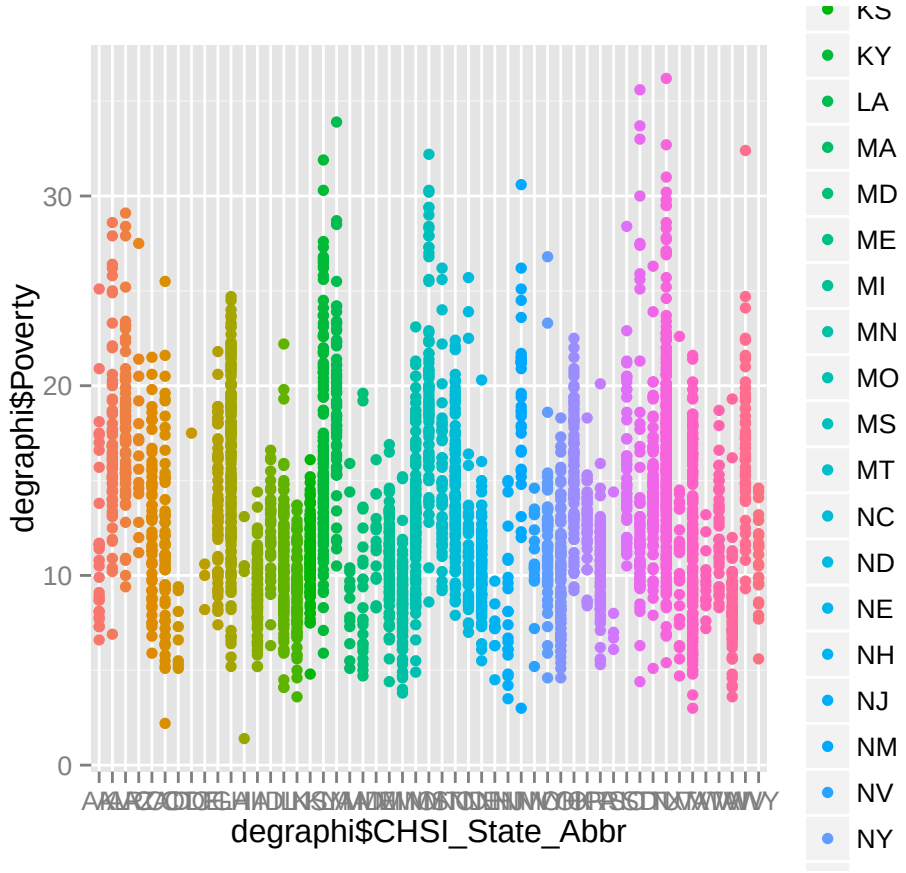

```
qplot(degraphi$CHSI_State_Abbr,degraphi$Poverty,stat = "identity",geom = "bar",fill=factor(degraphi$CHSI
```



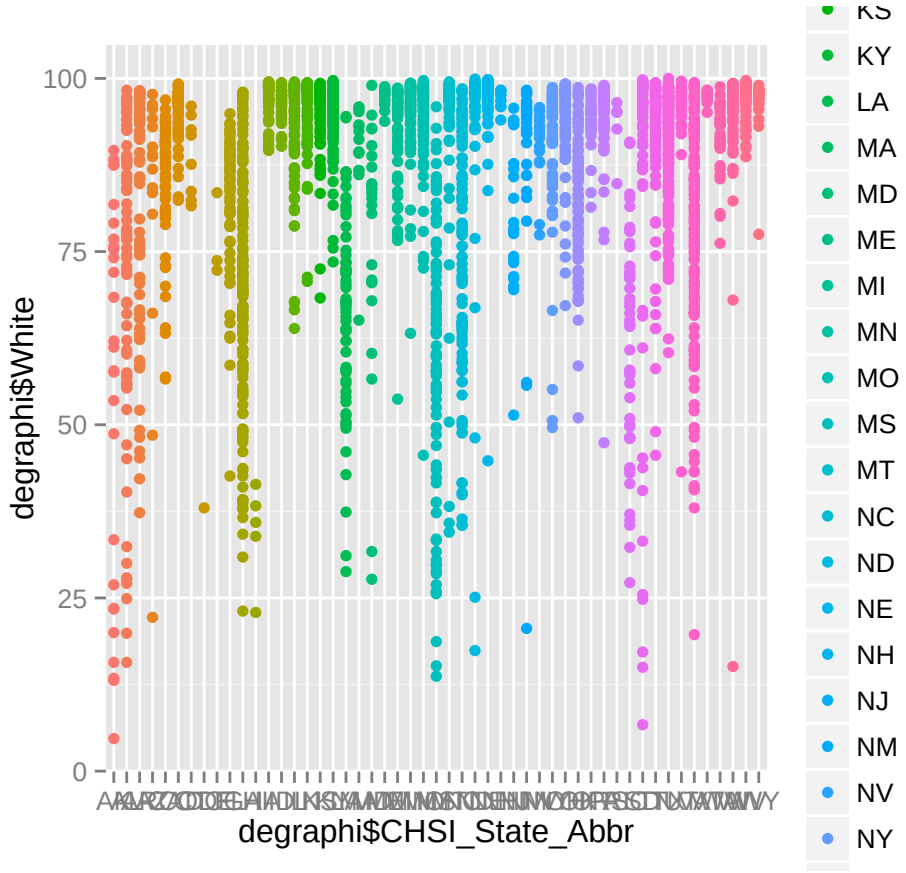

```
ggplot(data=degraphi,aes(x = degraphi$CHSI_State_Abbr,y=degraphi$Population_Size))+geom_point(aes(color=
```



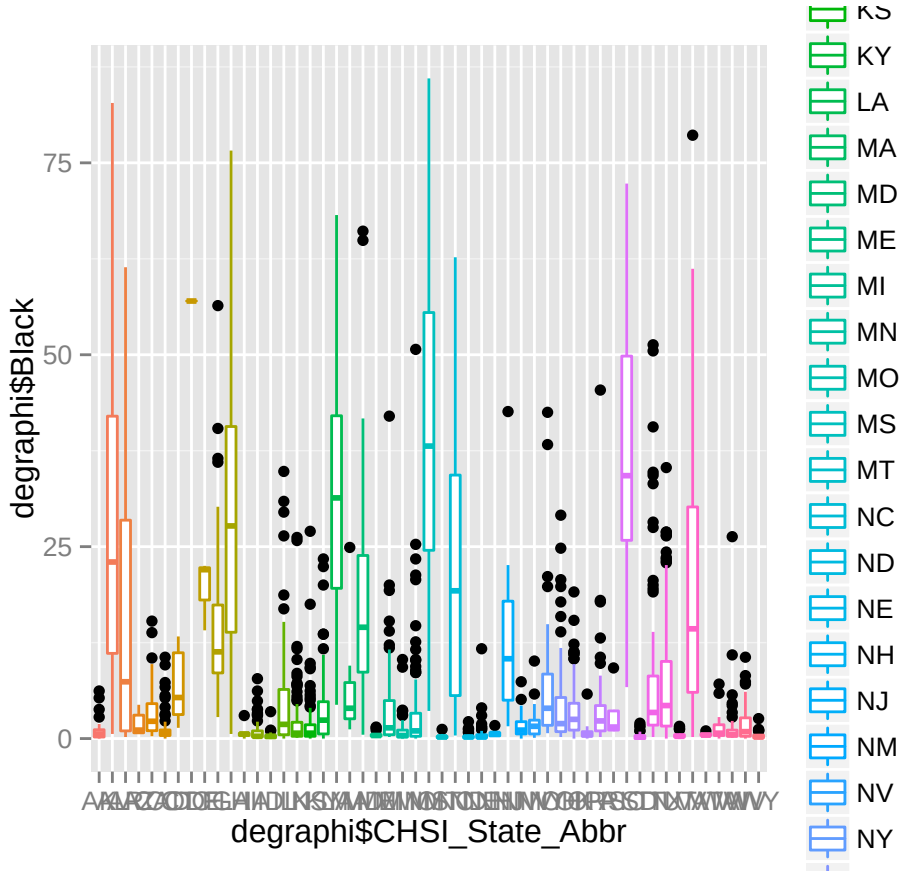

```
ggplot(data=degraphi,aes(x = degraphi$CHSI_State_Abbr,y=degraphi$Poverty))+geom_point(aes(color=degraphi
```




```
ggplot(data=degraphi,aes(x = degraphi$CHSI_State_Abbr,degraphi$White))+ geom_point(aes(color=degraphi$CH
```




```
ggplot(data=degraphi, aes(degraphi$CHSI_State_Abbr,degraphi$Black))+ geom_boxplot((aes(color=degraphi$CH
```




```
# CLUSTER K MEANS
require(useful)
```

```
## Loading required package: useful
```

```
data2 <- cbind.data.frame(degraphi$Population_Size,degraphi$State_FIPS_Code)
class(data2)
```

```
## [1] "data.frame"
```

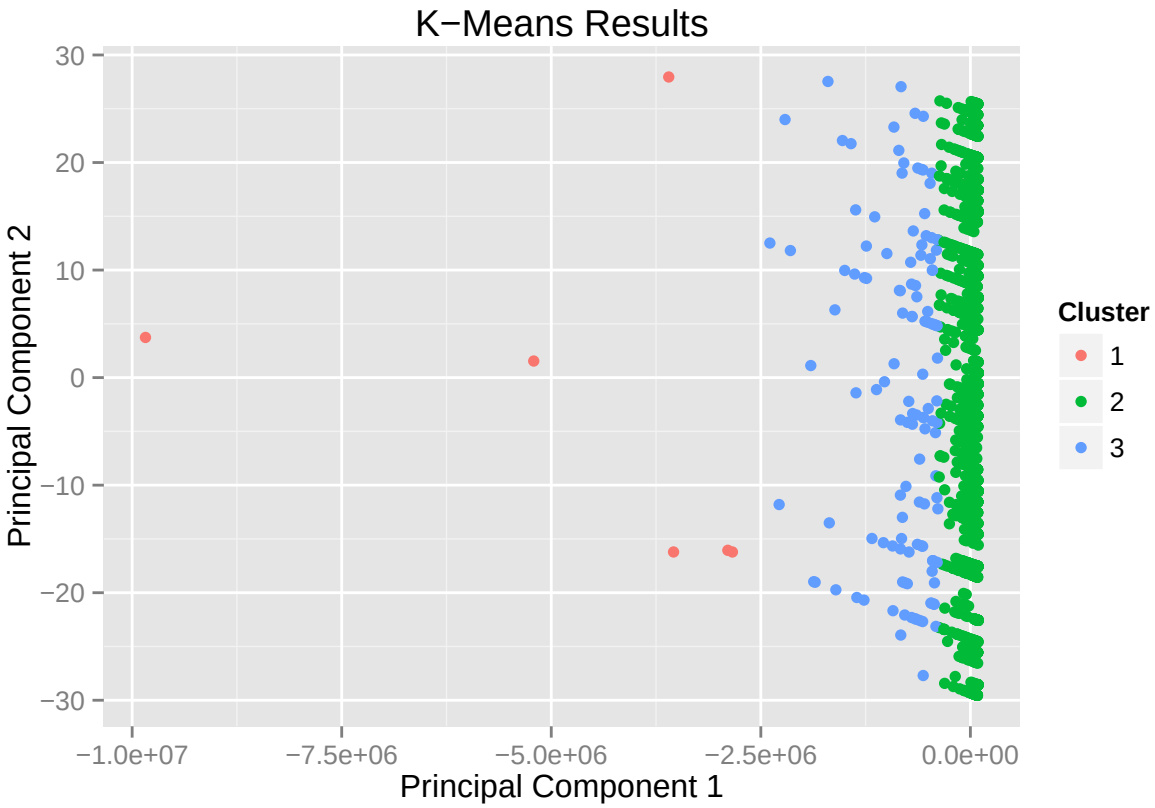
```
set.seed(50)
data2k<- kmeans(data2, centers=3)
data2k
```

```
## K-means clustering with 3 clusters of sizes 6, 3011, 124
##
## Cluster means:
##   degraphi$Population_Size degraphi$State_FIPS_Code
## 1                4748211.67                  14.50000
## 2                  51667.34                  30.48854
## 3                 906054.74                  26.60484
##
## Clustering vector:
##    [1] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
```

```
##   [35] 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##   [69] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1
##  [103] 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [137] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [171] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 3 2 2 3 2 2 2 2 3 2 2 2 1 2
##  [205] 2 2 2 2 2 2 2 2 2 1 2 2 3 3 2 3 1 3 3 2 3 2 3 2 2 2 2 2 2 3 2 2 2 2
##  [239] 2 3 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 3 2 2 2 2 2 2 2
##  [273] 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [307] 3 3 2 2 3 2 2 2 2 3 2 3 2 2 2 2 3 3 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2
##  [341] 2 2 2 2 2 3 2 2 2 2 2 2 3 2 2 2 2 2 2 2 3 2 2 2 2 3 2 3 2 3 3 2 2 2
##  [375] 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [409] 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [443] 2 2 3 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [477] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [511] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [545] 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [579] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2
##  [613] 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 3 2 2 2 2
##  [647] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [681] 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [715] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 3 2 2 2 2
##  [749] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [783] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [817] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [851] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [885] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [919] 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [953] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [987] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1021] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2
## [1055] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1089] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1123] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1157] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1191] 2 2 3 3 2 2 2 2 2 2 2 2 2 2 3 3 2 2 2 2 2 2 3 2 2 3 2 3 2 2 2 3
## [1225] 2 3 3 3 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1259] 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 3
## [1293] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1327] 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1361] 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1395] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1429] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1463] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1497] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2
## [1531] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1565] 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1599] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1633] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1667] 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1701] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1735] 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1769] 2 2 2 2 2 3 2 3 2 2 3 2 3 2 2 3 3 3 3 2 2 2 3 2 3 2 2 2 2 2 2 2 2
## [1803] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2
## [1837] 2 2 2 2 3 2 2 2 2 2 2 2 2 3 2 2 2 3 2 3 3 2 2 2 2 2 2 2 2 3 2 2 2
```

```
## [1871] 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1905] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1939] 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [1973] 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2007] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2041] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 3 2 2 2 2 2 3 2 2
## [2075] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2
## [2109] 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2143] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2177] 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2
## [2211] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 3 2 2 2 3
## [2245] 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 3
## [2279] 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2313] 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2347] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2381] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2415] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2
## [2449] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2483] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2
## [2517] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2551] 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 3 2 2
## [2585] 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2619] 2 2 2 1 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2653] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2687] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2721] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2
## [2755] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2789] 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2823] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2
## [2857] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2891] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2925] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [2959] 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 3 2 2 2 3 2 2 2 2 2 2 2 2 2 2
## [2993] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [3027] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [3061] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2
## [3095] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [3129] 2 2 2 2 2 2 2 2 2 2 2 2 2
##
## Within cluster sum of squares by cluster:
## [1] 3.595909e+13 1.676923e+13 2.498356e+13
##  (between_SS / total_SS =  73.6 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"
## [5] "tot.withinss" "betweenss"    "size"         "iter"
## [9] "ifault"
```

```
plot(data2k,data2)
```


K−Means Results
Principal Component 2
30
20
10
0
−10
−20
−30
−1.0e+07
−7.5e+06
−5.0e+06
−2.5e+06
0.0e+00
Principal Component 1
Cluster
1
2
3