

Quiz 3

Data released from the US department of Commerce, Bureau of the Census is available in R

```
>data(state)
```

```
>?state
```

Focus on the data {Population, Income, Illiteracy, Life Exp, Murder, HS Grad, Frost, Area}.

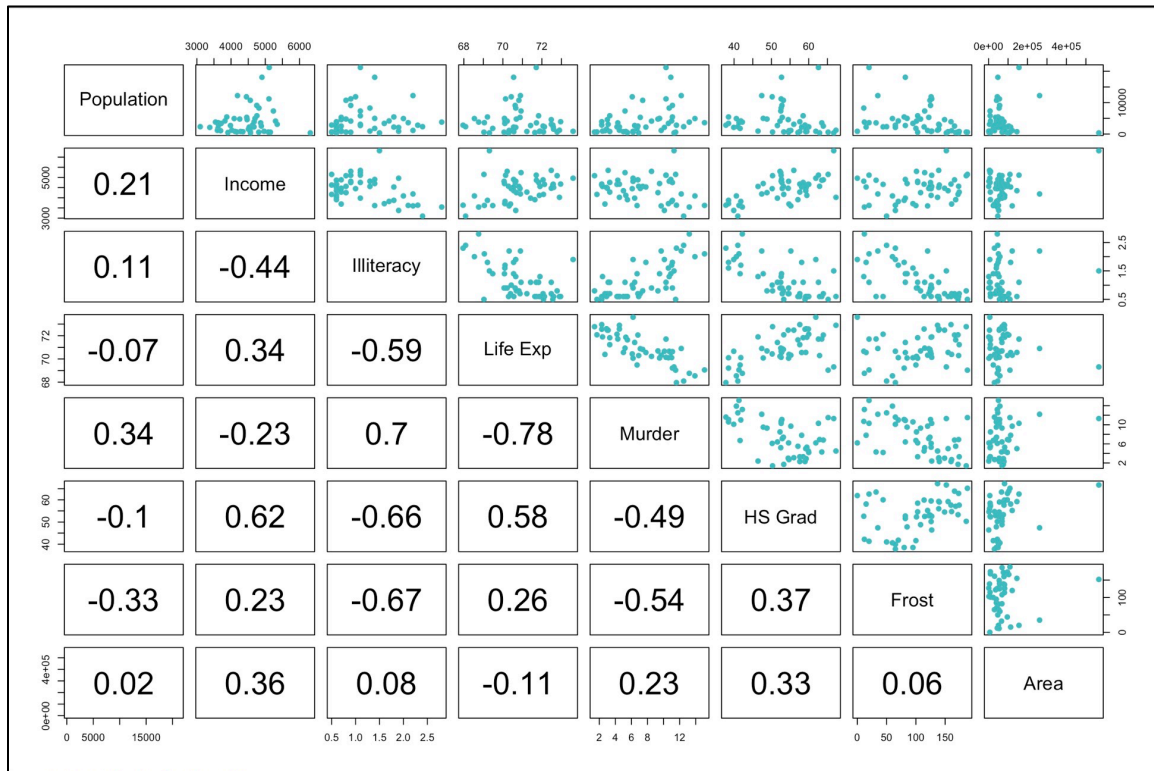
Cluster this data using at least two methods learned in STA 546 (e.g., k-means, hierarchical clustering, SOM, PCA, etc). Keep the class labels (region, or state name) in mind, but do not use them in the modeling. Report your detailed findings.

Summary of Data:

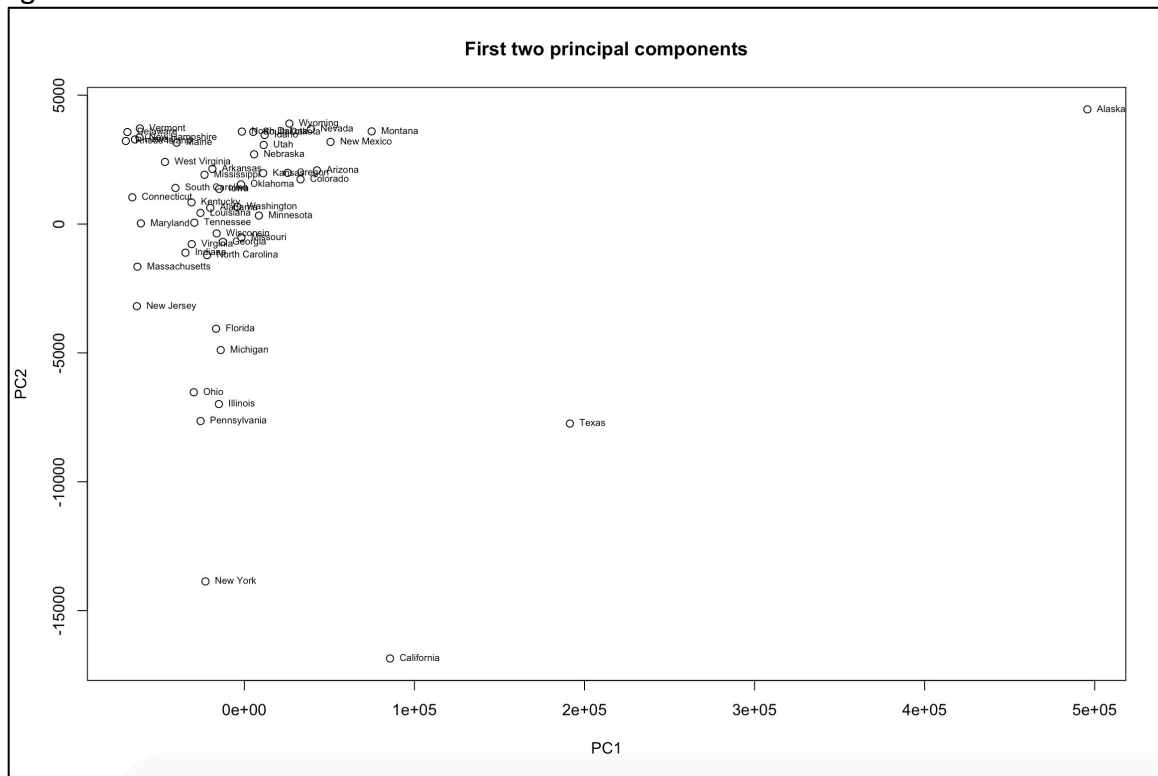
```
> summary(stData)
```

Population	Income	Illiteracy	Life Exp	Murder	HS Grad	Frost	Area
Min. : 365	Min. :3098	Min. :0.500	Min. :67.96	Min. : 1.400	Min. :37.80	Min. : 0.00	Min. : 1049
1st Qu.: 1080	1st Qu.:3993	1st Qu.:0.625	1st Qu.:70.12	1st Qu.: 4.350	1st Qu.:48.05	1st Qu.: 66.25	1st Qu.: 36985
Median : 2838	Median :4519	Median :0.950	Median :70.67	Median : 6.850	Median :53.25	Median :114.50	Median : 54277
Mean : 4246	Mean :4436	Mean :1.170	Mean :70.88	Mean : 7.378	Mean :53.11	Mean :104.46	Mean : 70736
3rd Qu.: 4968	3rd Qu.:4814	3rd Qu.:1.575	3rd Qu.:71.89	3rd Qu.:10.675	3rd Qu.:59.15	3rd Qu.:139.75	3rd Qu.: 81162
Max. :21198	Max. :6315	Max. :2.800	Max. :73.60	Max. :15.100	Max. :67.30	Max. :188.00	Max. :566432

Pairs Co-relation:

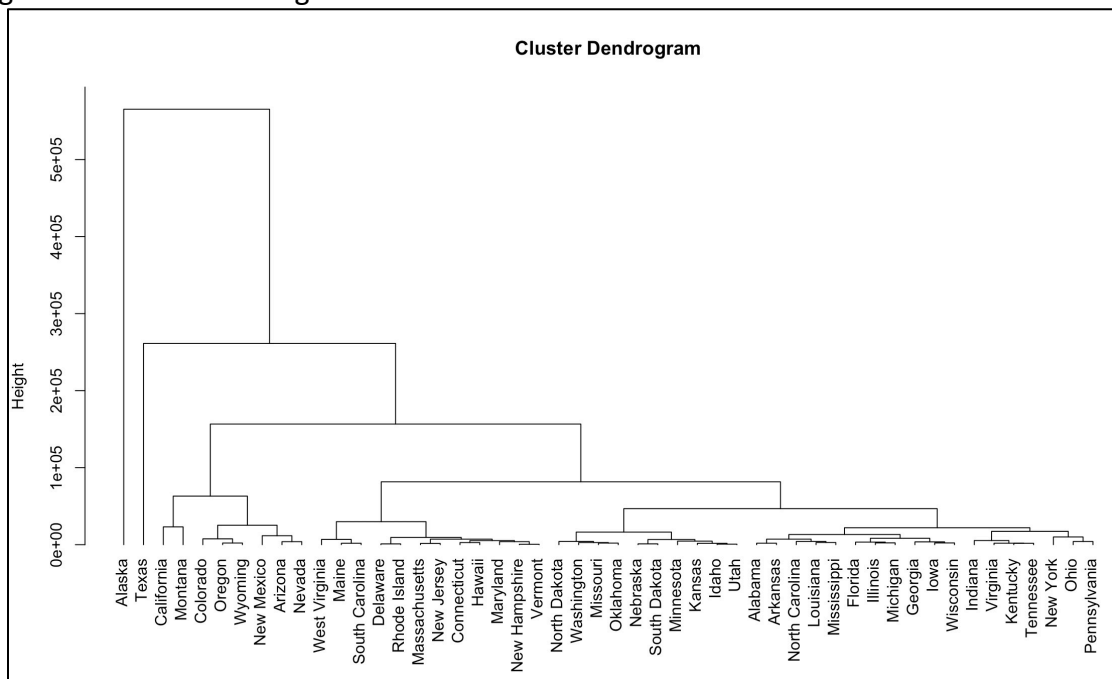


Using PCA:



Texas, California, New York seems like outlier in Principal Components Plot

Using Hierarchical clustering:



Alabama	Alaska	Arizona	Arkansas	California	Colorado	Connecticut	Delaware	Florida
1	2	3	1	3	3	4	4	1
Georgia	Hawaii	Idaho	Illinois	Indiana	Iowa	Kansas	Kentucky	Louisiana
1	4	1	1	1	1	1	1	1
Maine	Maryland	Massachusetts	Michigan	Minnesota	Mississippi	Missouri	Montana	Nebraska
4	4	4	1	1	1	1	3	1
Nevada	New Hampshire	New Jersey	New Mexico	New York	North Carolina	North Dakota	Ohio	Oklahoma
3	4	4	3	1	1	1	1	1
Oregon	Pennsylvania	Rhode Island	South Carolina	South Dakota	Tennessee	Texas	Utah	Vermont
3	1	4	4	1	1	5	1	4
Virginia	Washington	West Virginia	Wisconsin	Wyoming				
1	1	4	1	3				

Using Kmean

Alabama	Alaska	Arizona	Arkansas	California	Colorado	Connecticut	Delaware	Florida
3	1	5	3	5	5	4	4	3
Georgia	Hawaii	Idaho	Illinois	Indiana	Iowa	Kansas	Kentucky	Louisiana
3	4	3	3	3	3	3	3	3
Maine	Maryland	Massachusetts	Michigan	Minnesota	Mississippi	Missouri	Montana	Nebraska
4	4	4	3	3	3	3	5	3
Nevada	New Hampshire	New Jersey	New Mexico	New York	North Carolina	North Dakota	Ohio	Oklahoma
5	4	4	5	3	3	3	3	3
Oregon	Pennsylvania	Rhode Island	South Carolina	South Dakota	Tennessee	Texas	Utah	Vermont
5	3	4	4	3	3	2	3	4
Virginia	Washington	West Virginia	Wisconsin	Wyoming				
3	3	4	3	5				