Tae Coding
Introduction to Data Science: CS61
Summer 2018
Homework#9

Date Given: July 10, 2018   Due Date:
==

The following data is given for the 49 of America's largest cities.
Filename "RawDataUSCities.csv

- Percentage Black
- Percentage Hispanic
- Percentage Asian
- Media Age
- Unemployment rate
- Per Capita income

For example, Atlanta's demographic information is as follows: 67% black, 2% Hispanic, 1% Asian, has a median age of 31, a 5% unemployment rate, and a per-capita income of $22,000.

	A	B	C	D	E	F	G	H
1	City #	City	%age Black	%age Hispanic	%age Asian	Median Age	Unemployment rate	Per capita income(000's)
2	1	Albuquerque	3	35	2	32	5	18
3	2	Atlanta	67	2	1	31	5	22
4	3	Austin	12	23	3	29	3	19
5	4	Baltimore	59	1	1	33	11	22
6	5	Boston	26	11	5	30	5	24
7	6	Charlotte	32	1	2	32	3	20
8	7	Chicago	39	20	4	31	9	24
9	8	Cincinnati	38	1	1	31	8	21
10	9	Cleveland	47	5	1	32	13	22
11	10	Columbus	23	1	2	29	3	13
12	11	Dallas	30	21	2	30	9	22
13	12	Denver	13	23	2	34	7	23
14	13	Detroit	76	3	1	31	9	21
15	14	El Paso	3	69	1	29	11	13
16	15	Fort Worth	22	20	2	30	9	20
17	16	Fresno	9	30	13	28	13	16
18	17	Honolulu	1	5	71	37	5	24
19	18	Houston	28	28	4	30	7	22
20	19	Indianapolis	22	1	1	32	5	21
21	20	Jacksonville	25	3	2	32	7	19
22	21	Kansas City	30	4	1	33	6	21

1. Run a 3-cluster analysis on your U.S. City data using k-means algorithm before and after standardizing the data.

2. Run a 5-cluster analysis on your U.S. City data using k-means algorithm before and after standardizing the data.