Big Data Analysis

Assignment # 6 Report



Hussam-Ul-Hussain (18L-1827)

May 8, 2019

**Question # 1**

We took count of married and single people who have “tertiary”, “secondary” and “primary” education.

Married People count = 27214

Single People count = 12790

Since, the count of married people is greater than single people, therefore, Married people are more educated than Single people.

**Question # 2**

|  |  |  |
| --- | --- | --- |
| **K** | **Convergence Distance** | **Results** |
| 5 | 40 | **Finished iteration** (delta = 33.69146695321666)  **Clusters**  DenseVector(43.40948525159051, 2.7385772122614225, 0.44592249855407745, 0.8426836321573163, 8825.071717755927)  DenseVector(44.632075471698116, 2.9465408805031448, 0.4276729559748428, 0.8301886792452831, 21336.37735849057)  DenseVector(40.33199512708345, 2.6747881942521734, 0.523866216291046, 0.7007309374826957, 407.99036491500084)  DenseVector(52.56666666666666, 3.933333333333333, 0.5666666666666667, 0.9666666666666667, 56590.799999999996)  DenseVector(43.219925883694415, 2.8104332953249718, 0.4787628278221209, 0.786060433295325, 3294.27665336374) |
| 3 | 60 | **Finished iteration** (delta = 35.35790227445056)  **Clusters**  DenseVector(43.443652253909846, 2.77736890524379, 0.4588316467341306, 0.8183072677092916, 6356.404093836247)  DenseVector(45.637540453074436, 3.110032362459547, 0.4498381877022654, 0.8187702265372169, 25501.38511326861)  DenseVector(40.631552991073626, 2.6897223455146224, 0.5195541746806727, 0.7092518617152438, 642.898579671549) |
| 7 | 120 | **Finished iteration** (delta = 107.83068356928521)  **Clusters**  DenseVector(53.565217391304344, 3.8260869565217392, 0.6956521739130435, 0.9565217391304348, 61275.086956521736)  DenseVector(43.2455552692605, 2.7840762690028344, 0.4632826591084772, 0.7884565833548055, 3892.321566606545)  DenseVector(40.04116672863081, 2.6511305640990974, 0.5341535133673573, 0.696893906107412, 208.86504207929158)  DenseVector(43.62932662051605, 2.774071743234739, 0.486469477658905, 0.828193832599119, 7474.941472624292)  DenseVector(43.36968576709797, 2.813308687615527, 0.4066543438077634, 0.8207024029574861, 13605.735674676525)  DenseVector(45.247191011235955, 3.0168539325842696, 0.4606741573033708, 0.8651685393258427, 25737.55056179775)  DenseVector(42.09062898427539, 2.7961113472163195, 0.4792817679558011, 0.7393752656183595, 1601.6676583085423) |

**Question # 3**

|  |  |  |
| --- | --- | --- |
| **K** | **Convergence Distance** | **Results** |
| 3 | 50 | [43.42570373788648,6365.629672358099,16.139132441162896,276.5922934933087,2.662898015689894,39.92362713428703,0.7007383479464697]  [40.63444093867088,643.8847120883455,15.768339577992508,256.2109297968843,2.774773220272136,40.284657858410576,0.5677381187142576]  [45.637540453074436,25501.38511326861,16.13915857605178,255.9708737864078,2.7443365695792883,32.64401294498382,0.5436893203883495] |
| 5 | 60 | [43.2136592379583,3360.302803738318,16.12293314162473,273.980589503954,2.618116462976276,41.298346513299784,0.6940330697340044]  [40.34448243533888,413.94562400044117,15.733662383499697,254.22249489880326,2.7956212430375555,40.020735675288144,0.5536039265427674]  [45.05882352941177,22100.723183391005,16.432525951557093,256.9861591695502,2.7024221453287196,33.54671280276817,0.5640138408304498]  [53.07142857142857,57830.03571428571,11.857142857142856,251.35714285714283,3.464285714285714,20.392857142857142,0.21428571428571427]  [43.38015540944411,9085.638971906754,16.02570233114166,278.14465032875074,2.679617453676031,40.94202032277346,0.6957561267184698] |
| 7 | 70 | [40.08579479207151,240.50369218810727,15.71693224510947,251.7278144837414,2.815487757481539,39.25799974089908,0.5331649177354579]  [43.3727959697733,16022.546599496221,15.79345088161209,263.05541561712846,2.7632241813602016,35.3727959697733,0.5516372795969773]  [42.32601707127813,1825.5094778849352,15.909876953774527,269.87817315153535,2.63806673317814,43.650703913091675,0.678084469571001]  [43.42112756264237,4468.780466970387,16.200740318906604,276.9114464692483,2.675683371298405,39.86531890660592,0.6899202733485194]  [55.888888888888886,65827.05555555555,10.944444444444445,253.66666666666666,2.5555555555555554,32.27777777777778,0.3333333333333333]  [43.436220472440944,8694.992913385826,16.177952755905512,278.20472440944883,2.651968503937008,41.166141732283464,0.7385826771653543]  [45.90598290598291,29172.076923076926,16.36752136752137,256.92307692307696,2.7264957264957266,39.05128205128206,0.6153846153846154] |