PRÁCTICA DIRIGIDA DE CLUSTER

**Caso 1: Violencia en USA**

A continuación, se muestra el conjunto de datos USA, el cual contiene variables sobre asesinato, asalta, y violación entre las principales ciudades de Estados unidos.

Realice el análisis Cluster, y realice las siguientes actividades:

1. Analice si existe correlación entre las variables

Existe correlación entre la cantidad de asaltos y la cantidad de asesinatos (r = 0.8)

1. A cuanto equivale la similaridad entre Colorado y Arizona

Distancia = 90.024811

1. Según el método de SDindex cuanto debería ser el número de cluster adecuado
2. Identifique las ciudades de California y Colorado a que Cluster pertenecen
3. Caracterice los cluster e identifique en que cluster se encuentra la mayor cantidad de Asesinatos.
4. Cual es la diferencia principal entre el cluster 1 y el cluster 3

| Estados | Asesinato | Asalto | Pob.Urbana | Violación |
| --- | --- | --- | --- | --- |
| Alabama | 13.2 | 236 | 58 | 21.2 |
| Alaska | 10 | 263 | 48 | 44.5 |
| Arizona | 8.1 | 294 | 80 | 31 |
| Arkansas | 8.8 | 190 | 50 | 19.5 |
| California | 9 | 276 | 91 | 40.6 |
| Colorado | 7.9 | 204 | 78 | 38.7 |
| Connecticut | 3.3 | 110 | 77 | 11.1 |
| Delaware | 5.9 | 238 | 72 | 15.8 |
| Florida | 15.4 | 335 | 80 | 31.9 |
| Georgia | 17.4 | 211 | 60 | 25.8 |
| Hawaii | 5.3 | 46 | 83 | 20.2 |
| Idaho | 2.6 | 120 | 54 | 14.2 |
| Illinois | 10.4 | 249 | 83 | 24 |
| Indiana | 7.2 | 113 | 65 | 21 |
| Iowa | 2.2 | 56 | 57 | 11.3 |
| Kansas | 6 | 115 | 66 | 18 |
| Kentucky | 9.7 | 109 | 52 | 16.3 |
| Louisiana | 15.4 | 249 | 66 | 22.2 |
| Maine | 2.1 | 83 | 51 | 7.8 |
| Maryland | 11.3 | 300 | 67 | 27.8 |
| Massachusetts | 4.4 | 149 | 85 | 16.3 |
| Michigan | 12.1 | 255 | 74 | 35.1 |
| Minnesota | 2.7 | 72 | 66 | 14.9 |
| Mississippi | 16.1 | 259 | 44 | 17.1 |
| Missouri | 9 | 178 | 70 | 28.2 |
| Montana | 6 | 109 | 53 | 16.4 |
| Nebraska | 4.3 | 102 | 62 | 16.5 |
| Nevada | 12.2 | 252 | 81 | 46 |
| New Hampshire | 2.1 | 57 | 56 | 9.5 |
| New Jersey | 7.4 | 159 | 89 | 18.8 |
| New Mexico | 11.4 | 285 | 70 | 32.1 |
| New York | 11.1 | 254 | 86 | 26.1 |
| North Carolina | 13 | 337 | 45 | 16.1 |
| North Dakota | 0.8 | 45 | 44 | 7.3 |
| Ohio | 7.3 | 120 | 75 | 21.4 |
| Oklahoma | 6.6 | 151 | 68 | 20 |
| Oregon | 4.9 | 159 | 67 | 29.3 |
| Pennsylvania | 6.3 | 106 | 72 | 14.9 |
| Rhode Island | 3.4 | 174 | 87 | 8.3 |
| South Carolina | 14.4 | 279 | 48 | 22.5 |
| South Dakota | 3.8 | 86 | 45 | 12.8 |
| Tennessee | 13.2 | 188 | 59 | 26.9 |
| Texas | 12.7 | 201 | 80 | 25.5 |
| Utah | 3.2 | 120 | 80 | 22.9 |
| Vermont | 2.2 | 48 | 32 | 11.2 |
| Virginia | 8.5 | 156 | 63 | 20.7 |
| Washington | 4 | 145 | 73 | 26.2 |
| West Virginia | 5.7 | 81 | 39 | 9.3 |
| Wisconsin | 2.6 | 53 | 66 | 10.8 |
| Wyoming | 6.8 | 161 | 60 | 15.6 |

**Caso 2: Data distritos**

Con la data distritos realice un cluster con y sin componentes principales y determine cual perminte indicar indicadores más eficientes.