Please refer to the csv file “grad\_admissions” that contains data about applications to universities in the United States. The file has several columns given as:

1. GRE Scores (out of 340)
2. TOEFL Scores (out of 120)
3. University Rating (out of 5)
4. Statement of Purpose and Letter of Recommendation Strength (out of 5)
5. Undergraduate GPA (out of 10)
6. Research Experience (either 0 or 1)

Please answer the following:

1. Use K means to find clusters in the data set, please also interpret the clusters formed. There is no right and wrong answer as long as you can explain the clusters. Provide your output, and a summary of your results and analysis below.

Cluster 0: High GRE score, high TOEFL score, high university ratio, high SOP, high LOR, high CGPA, did research

Cluster 1: Low GRE score, low TOEFL score, low university ratio, low SOP, low LOR, low CGPA, no research

Cluster 2: Medium GRE score, medium TOEFL score, medium university ratio, medium SOP, medium LOR, medium CGPA, with research

Cluster 3: Medium GRE score, medium TOEFL score, medium university ratio, medium SOP, medium LOR, medium CGPA, no research

Cluster 0:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ﻿ | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research |
| count | 115 | 115 | 115 | 115 | 115 | 115 | 115 |
| mean | 328.791304 | 113.556522 | 4.37391304 | 4.39565217 | 4.29130435 | 9.24547826 | 1 |
| std | 6.73364901 | 3.94966422 | 0.64155608 | 0.49778303 | 0.56577306 | 0.34150144 | 0 |
| min | 310 | 102 | 2 | 3 | 2.5 | 8 | 1 |
| 25% | 324 | 111 | 4 | 4 | 4 | 9.06 | 1 |
| 50% | 329 | 113 | 4 | 4.5 | 4.5 | 9.23 | 1 |
| 75% | 334 | 117 | 5 | 5 | 4.5 | 9.455 | 1 |
| max | 340 | 120 | 5 | 5 | 5 | 9.92 | 1 |

Cluster 1:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ﻿ | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research |
| count | 89 | 89 | 89 | 89 | 89 | 89 | 89 |
| mean | 305.505618 | 101.146067 | 1.94382022 | 2.18539326 | 2.51123596 | 7.94797753 | 0 |
| std | 7.58156196 | 3.48878255 | 0.68023309 | 0.62332513 | 0.59818631 | 0.37026459 | 0 |
| min | 290 | 92 | 1 | 1 | 1 | 7.2 | 0 |
| 25% | 299 | 99 | 2 | 2 | 2 | 7.65 | 0 |
| 50% | 305 | 101 | 2 | 2 | 2.5 | 7.97 | 0 |
| 75% | 312 | 104 | 2 | 2.5 | 3 | 8.22 | 0 |
| max | 321 | 107 | 4 | 4 | 4 | 9 | 0 |

Cluster 2:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ﻿ | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research |
| count | 92 | 92 | 92 | 92 | 92 | 92 | 92 |
| mean | 313.347826 | 107.043478 | 3.08695652 | 3.60869565 | 3.5923913 | 8.55630435 | 0 |
| std | 8.13273778 | 4.16134235 | 0.84713057 | 0.67051762 | 0.65435711 | 0.33633355 | 0 |
| min | 298 | 97 | 1 | 2 | 2 | 7.66 | 0 |
| 25% | 308 | 104 | 3 | 3 | 3 | 8.33 | 0 |
| 50% | 314 | 107 | 3 | 3.5 | 3.5 | 8.56 | 0 |
| 75% | 318.25 | 110 | 3 | 4 | 4 | 8.76 | 0 |
| max | 339 | 120 | 5 | 5 | 5 | 9.7 | 0 |

Cluster 3:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ﻿ | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research |
| count | 104 | 104 | 104 | 104 | 104 | 104 | 104 |
| mean | 316.288462 | 106.298077 | 2.64423077 | 3.15384615 | 3.20673077 | 8.47875 | 1 |
| std | 8.39980707 | 4.72111515 | 0.63732251 | 0.66492179 | 0.67769805 | 0.42693008 | 0 |
| min | 293 | 95 | 1 | 1 | 1.5 | 6.8 | 1 |
| 25% | 311.75 | 104 | 2 | 3 | 3 | 8.195 | 1 |
| 50% | 318 | 107 | 3 | 3.5 | 3 | 8.5 | 1 |
| 75% | 323.25 | 110 | 3 | 3.5 | 3.5 | 8.7725 | 1 |
| max | 329 | 116 | 4 | 4.5 | 5 | 9.32 | 1 |

1. Use Hierarchical clustering (any method that you find appropriate), please also interpret the clusters formed. Provide your output, and a summary of your results and analysis below.

﻿Cluster 0: High GRE score, high TOEFL score, high university ratio, high SOP, high LOR, high CGPA, did research

Cluster 1: Low GRE score, low TOEFL score, low university ratio, low SOP, low LOR, low CGPA, low research

Cluster 2: Medium GRE score, medium/low TOEFL score, medium/low university ratio, medium/low SOP, medium/low LOR, medium/low CGPA, with research

Cluster 3: Medium GRE score, medium/high TOEFL score, medium/high university ratio, medium/high SOP, medium/high LOR, medium/high CGPA, no research

Cluster 0:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ﻿ | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research |
| count | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| mean | 327.842857 | 113.121429 | 4.15 | 4.23928571 | 4.125 | 9.18842857 | 1 |
| std | 6.57800065 | 3.80357245 | 0.7766836 | 0.58293433 | 0.65734642 | 0.34839255 | 0 |
| min | 310 | 102 | 2 | 3 | 2.5 | 8 | 1 |
| 25% | 324 | 110 | 4 | 4 | 3.5 | 9 | 1 |
| 50% | 327 | 113 | 4 | 4.5 | 4 | 9.165 | 1 |
| 75% | 332.25 | 116 | 5 | 4.5 | 4.5 | 9.4325 | 1 |
| max | 340 | 120 | 5 | 5 | 5 | 9.92 | 1 |

Cluster 1:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ﻿ | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research |
| count | 123 | 123 | 123 | 123 | 123 | 123 | 123 |
| mean | 306.471545 | 102.211382 | 2.08130081 | 2.49186992 | 2.73577236 | 8.05869919 | 0 |
| std | 7.58268649 | 3.81380588 | 0.68465016 | 0.79182229 | 0.67849038 | 0.39599213 | 0 |
| min | 290 | 92 | 1 | 1 | 1 | 7.2 | 0 |
| 25% | 300 | 99.5 | 2 | 2 | 2 | 7.8 | 0 |
| 50% | 306 | 102 | 2 | 2.5 | 2.5 | 8.1 | 0 |
| 75% | 313 | 105 | 2 | 3 | 3 | 8.325 | 0 |
| max | 321 | 110 | 4 | 5 | 4 | 9.12 | 0 |

Cluster 2:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ﻿ | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research |
| count | 79 | 79 | 79 | 79 | 79 | 79 | 79 |
| mean | 314.012658 | 104.772152 | 2.49367089 | 3.03797468 | 3.15822785 | 8.33721519 | 1 |
| std | 8.2423137 | 4.26659465 | 0.61755136 | 0.69693538 | 0.7229937 | 0.37055031 | 0 |
| min | 293 | 95 | 1 | 1 | 1.5 | 6.8 | 1 |
| 25% | 310.5 | 102 | 2 | 2.5 | 3 | 8.11 | 1 |
| 50% | 315 | 105 | 3 | 3 | 3 | 8.36 | 1 |
| 75% | 320 | 107 | 3 | 3.5 | 3.5 | 8.565 | 1 |
| max | 329 | 114 | 3 | 4.5 | 5 | 9.15 | 1 |

Cluster 3:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ﻿ | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research |
| count | 58 | 58 | 58 | 58 | 58 | 58 | 58 |
| mean | 315.896552 | 108.241379 | 3.46551724 | 3.79310345 | 3.75 | 8.67810345 | 0 |
| std | 7.65870599 | 4.21488734 | 0.75429803 | 0.64232638 | 0.683451 | 0.29159246 | 0 |
| min | 301 | 99 | 2 | 2.5 | 2 | 8 | 0 |
| 25% | 312 | 105 | 3 | 3.5 | 3.5 | 8.485 | 0 |
| 50% | 315.5 | 108 | 3 | 3.75 | 3.5 | 8.67 | 0 |
| 75% | 321.75 | 110 | 4 | 4.5 | 4 | 8.8 | 0 |
| max | 339 | 120 | 5 | 5 | 5 | 9.7 | 0 |