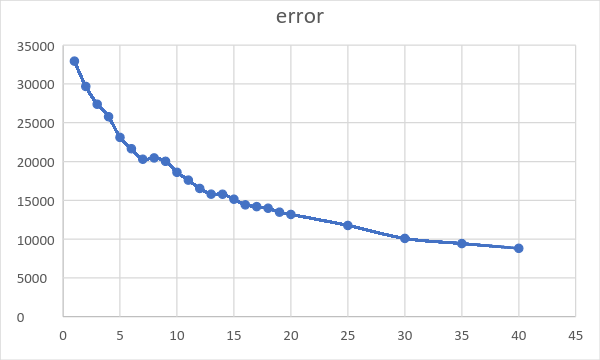
Results of PCA alone:

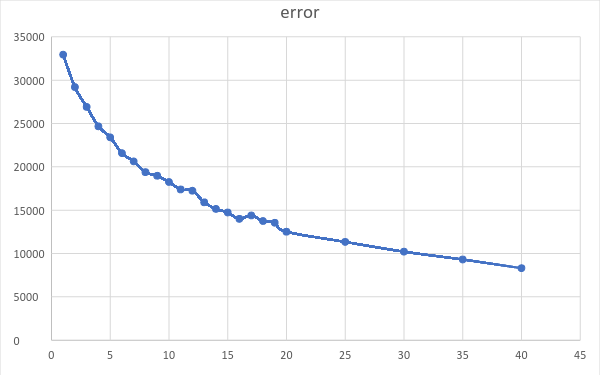
Below are the outcomes of the simple-K-means algorithm along with the graph. There is no difference of the sum of squared error after splitting the data into training and testing data sets. The results below have the seed value set to 10.

|  |  |
| --- | --- |
| clusters | Error |
| 1 | 32941.21268 |
| 2 | 29678.93322 |
| 3 | 27387.66419 |
| 4 | 25778.14111 |
| 5 | 23111.54135 |
| 6 | 21653.84523 |
| 7 | 20297.41842 |
| 8 | 20464.33993 |
| 9 | 20026.09097 |
| 10 | 18607.36865 |
| 11 | 17602.2734 |
| 12 | 16531.2821 |
| 13 | 15793.12456 |
| 14 | 15785.36579 |
| 15 | 15144.89644 |
| 16 | 14414.50536 |
| 17 | 14187.11193 |
| 18 | 13980.61198 |
| 19 | 13474.32648 |
| 20 | 13179.37527 |
| 25 | 11758.46829 |
| 30 | 10090.10642 |
| 35 | 9425.838205 |
| 40 | 8810.831591 |



Below are the results of seed value of 20.

|  |  |
| --- | --- |
| clusters | Error |
| 1 | 32941.21268 |
| 2 | 29208.65663 |
| 3 | 26918.62999 |
| 4 | 24678.37713 |
| 5 | 23406.52777 |
| 6 | 21580.5268 |
| 7 | 20628.32363 |
| 8 | 19387.26152 |
| 9 | 18971.88968 |
| 10 | 18253.03539 |
| 11 | 17391.64599 |
| 12 | 17242.26089 |
| 13 | 15908.66791 |
| 14 | 15150.85193 |
| 15 | 14729.89124 |
| 16 | 13998.54873 |
| 17 | 14389.87753 |
| 18 | 13751.41417 |
| 19 | 13545.96901 |
| 20 | 12513.20641 |
| 25 | 11335.60326 |
| 30 | 10208.63445 |
| 35 | 9305.176092 |
| 40 | 8304.874964 |



Results of PCA alone with normalization:

Getting the same values as PCA alone.