

KMeans Clustering

Name: Sri Harish Pinnimti

UTA ID: 1001865949

Net ID: sxp5949

This PDF contains the output for the KMeans Clustering assignment with the given datasets.

To run the KMeans Clustering algorithm, run the following command:

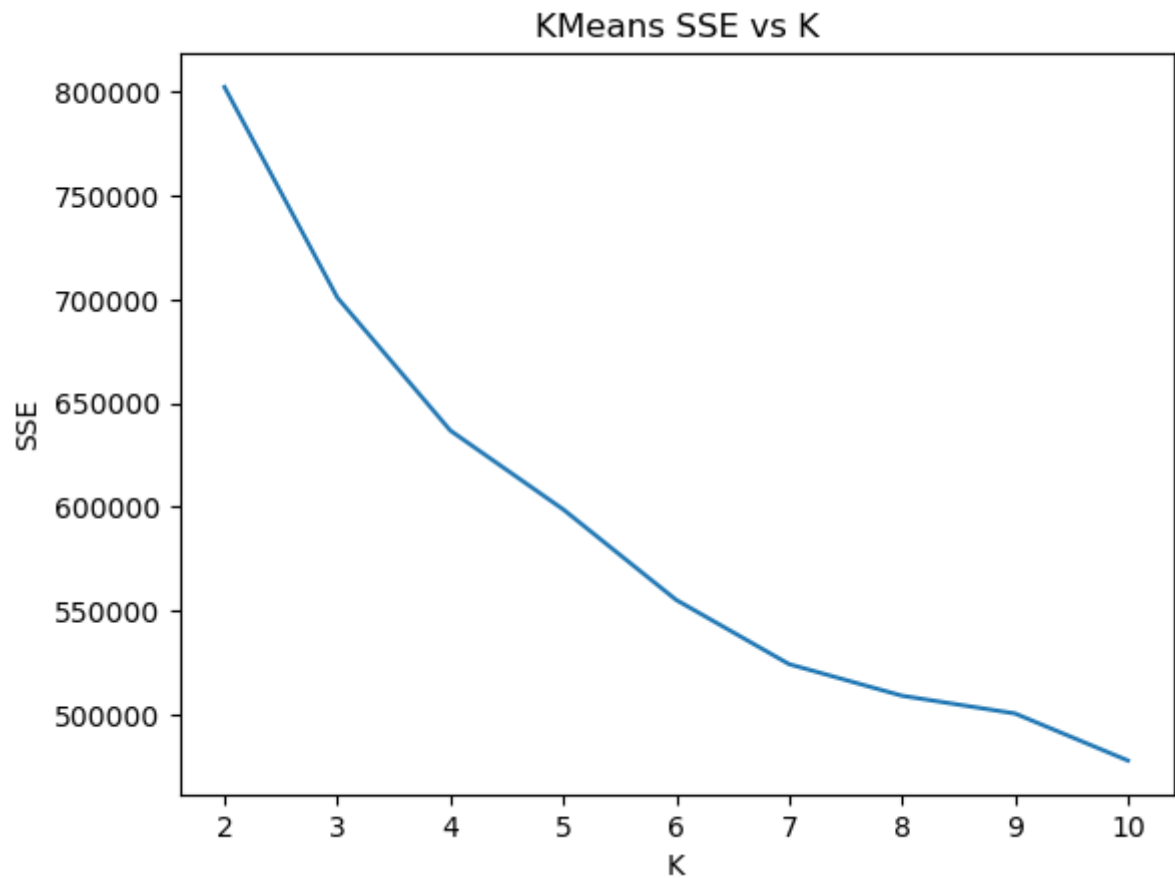
```
python3 P3.py <filepath>
```

```
In [1]: # Import the P3.py file
        from P3 import *
```

Pendigits Dataset

```
In [2]: print("Pendigits Dataset")
        run("UCI_datasets/pendigits_training.txt")
```

```
Pendigits Dataset
For k = 2 After 20 iterations: SSE error = 802390.5237
For k = 3 After 20 iterations: SSE error = 700857.4281
For k = 4 After 20 iterations: SSE error = 636815.2687
For k = 5 After 20 iterations: SSE error = 598796.5830
For k = 6 After 20 iterations: SSE error = 555188.7321
For k = 7 After 20 iterations: SSE error = 524211.8400
For k = 8 After 20 iterations: SSE error = 508982.2051
For k = 9 After 20 iterations: SSE error = 500438.8373
For k = 10 After 20 iterations: SSE error = 477759.0605
```



Satellite Dataset

```
In [3]: print("Satellite Dataset")
run("UCI_datasets/satellite_training.txt")
```

Satellite Dataset

For k = 2 After 20 iterations: SSE error = 334665.6434

For k = 3 After 20 iterations: SSE error = 259999.4592

For k = 4 After 20 iterations: SSE error = 235302.8459

For k = 5 After 20 iterations: SSE error = 228705.5949

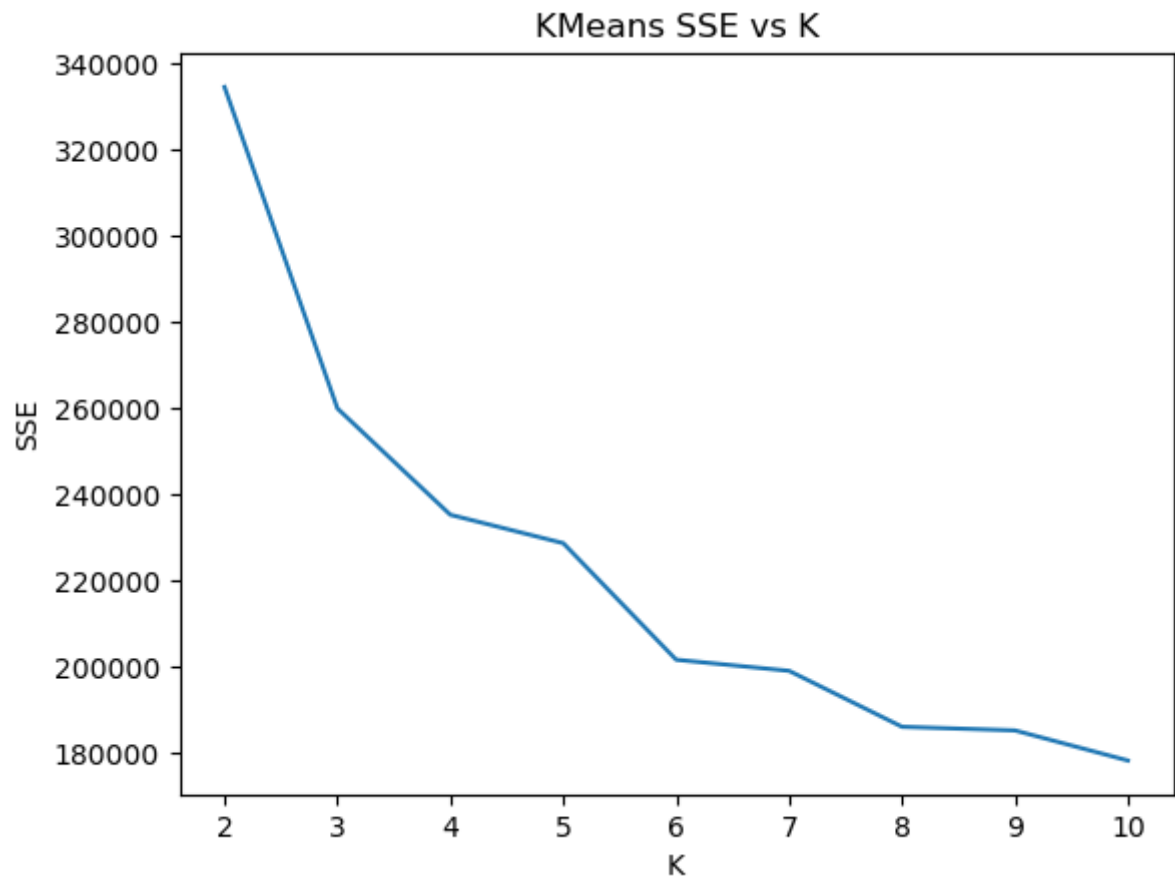
For k = 6 After 20 iterations: SSE error = 201628.5676

For k = 7 After 20 iterations: SSE error = 199047.3693

For k = 8 After 20 iterations: SSE error = 186047.6737

For k = 9 After 20 iterations: SSE error = 185205.1174

For k = 10 After 20 iterations: SSE error = 178212.5437



Yeast Dataset

```
In [4]: print("Yeast Dataset")  
run("UCI_datasets/yeast_training.txt")
```

```
Yeast Dataset  
For k = 2 After 20 iterations: SSE error = 226.5971  
For k = 3 After 20 iterations: SSE error = 211.1458  
For k = 4 After 20 iterations: SSE error = 191.5702  
For k = 5 After 20 iterations: SSE error = 181.3487  
For k = 6 After 20 iterations: SSE error = 175.6126  
For k = 7 After 20 iterations: SSE error = 170.3621  
For k = 8 After 20 iterations: SSE error = 165.0169  
For k = 9 After 20 iterations: SSE error = 163.9576  
For k = 10 After 20 iterations: SSE error = 160.4186
```

KMeans SSE vs K

