

```
1 C:\Users\User\Desktop\ergasia\.venv\Scripts\python.exe C:\Users\User\Desktop\ergasia\apomakrismeno_test.py
```

```
2  
3 #####
```

```
4 Running experiment for K = 7
```

```
5 #####
```

```
6
```

```
7 Loading and Preprocessing Data (Brute Force)...
```

```
8 Creating Shingles (Brute Force)...
```

```
9 Starting Brute Force Jaccard computation...
```

```
10 Finished Brute Force Jaccard computation.
```

```
11 Brute Force total time: 9429.46 sec
```

```
12
```

```
13
```

```
14 Loading and preprocessing data (LSH) with threshold=0.5, num_perm=64...
```

```
15 Generating MinHash signatures and shingles (LSH)...
```

```
16 Building LSH index...
```

```
17 Finding nearest neighbors with refined brute force Jaccard ranking...
```

```
18 Nearest neighbor search completed.
```

```
19
```

```
20 Final Results Table:
```

```
21
```

```
22           Type BuildTime QueryTime TotalTime fraction of the true K most similar documents that are reported by
```

```
LSH method as well Parameters
```

```
23 Brute-Force-Jaccard 0.00 9429.460251 9429.460251
```

```
100%
```

```
K=7
```

```
24 . LSH-Jaccard 7363.49 194.170000 7557.660000
```

```
7% K=7, Perm=64, threshold=0.5
```

```
25
```

```
26
```

```
27 Process finished with exit code 0
```

```
28
```