

Weekly Exercise 4: Solutions for validating your LOF code!

Dataset: Class example dataset.

File name: class_example_dataset.csv

Value for k: 3

```
LOF[0] = 2.3111659487059772
LOF[1] = 2.901337923383965
LOF[2] = 2.7192136676924727
LOF[3] = 1.0
LOF[4] = 1.0
LOF[5] = 1.0
LOF[6] = 1.0
```

Value for k: 4

```
LOF[0] = 1.2077141496973964
LOF[1] = 1.0684998099980942
LOF[2] = 0.9636526192690476
LOF[3] = 1.0047916991293748
LOF[4] = 1.03622273887712
LOF[5] = 0.981857516592909
LOF[6] = 0.9818575165929087
```

Dataset: Artificial example dataset.

File name: class_example_dataset.csv

Value for k: 4

```
LOF[0] = 1.056893349870319
LOF[1] = 1.4689258225389958
LOF[2] = 1.0823130765793778
LOF[3] = 1.4599117849321113
LOF[4] = 1.0183831411877478
LOF[5] = 0.944778614517102
LOF[6] = 1.6134885302766089
LOF[7] = 1.0459472042634712
LOF[8] = 1.1358273701950414
LOF[9] = 0.9689959922001229
LOF[10] = 1.1323573613574582
LOF[11] = 0.9794678383811398
LOF[12] = 1.5483697150877322
LOF[13] = 1.2789303519048583
LOF[14] = 1.0460556744659266
LOF[15] = 1.0060835850748058
LOF[16] = 1.027266449763105
LOF[17] = 1.061508477098397
LOF[18] = 0.9723964951879704
LOF[19] = 0.9866756744894671
```

Value for k: 5

```
LOF[0] = 1.075148043222013
LOF[1] = 1.4236650353095646
LOF[2] = 1.1320010424996756
LOF[3] = 1.403048629622052
LOF[4] = 1.0330487568928615
LOF[5] = 0.9721288397984098
LOF[6] = 1.6015074112052121
LOF[7] = 1.0314890314004175
LOF[8] = 1.1004659249563982
LOF[9] = 0.9580488947646348
LOF[10] = 1.1304960691218002
LOF[11] = 0.9653990607588974
LOF[12] = 1.4764577387853817
LOF[13] = 1.2747110443336653
LOF[14] = 1.0651501323526504
LOF[15] = 0.9877349524409432
LOF[16] = 1.0053691641048474
```

```
LOF[17] = 1.0557945390931913  
LOF[18] = 1.0701600067599166  
LOF[19] = 0.9935123743458353
```

Notes:

You can also validate using the following code. Read the csv file to variable data. The only difference is that the numbers will be negative. Scikit learn has a unified scoring API that always maximizes the score, so scores which need to be minimized are negated in order for the unified scoring API to work correctly.

```
import numpy as np  
from sklearn.neighbors import LocalOutlierFactor  
  
clf = LocalOutlierFactor(n_neighbors=3)  
clf.fit_predict(data)  
  
print(clf.negative_outlier_factor_)
```