

Лабораторная работа №4

Вариант №8

Вид классов: `classification`

Random state: `36`

Class sep: `1.2`

Для всех:

```
n_features = 2
```

```
n_redundant = 0
```

```
n_informative = 2
```

```
n_clusters_per_class = 1
```

```
n_classes = 4
```

```
n_samples = 100
```

```
In [1]: from sklearn.datasets import make_classification
```

Загрузка выборки согласно варианту №7

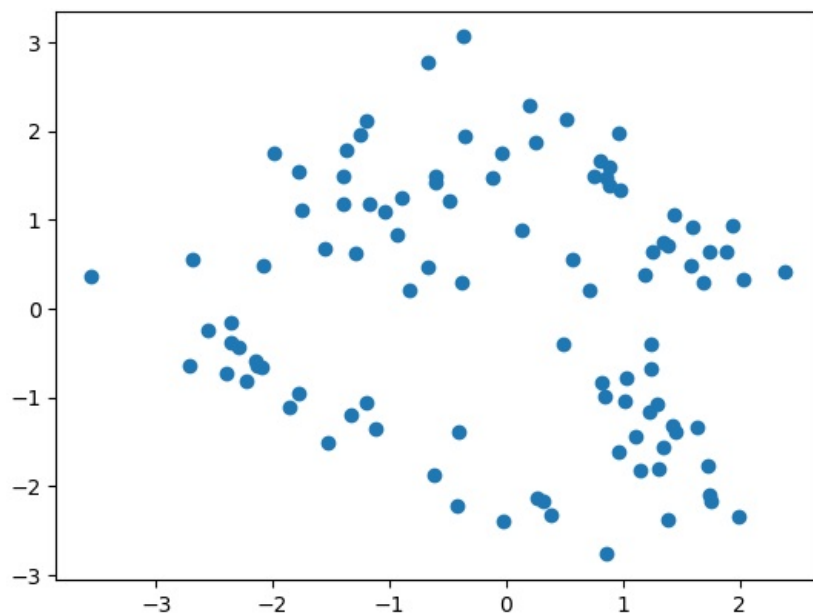
```
In [2]: X, y = make_classification(n_samples=100,  
                                n_features=2,  
                                n_redundant=0,  
                                n_informative=2,  
                                n_clusters_per_class=1,  
                                n_classes=4,  
                                random_state=36,  
                                class_sep=1.2)
```

Отображение выборки на графике

```
In [3]: import matplotlib.pyplot as plt
```

```
In [4]: plt.scatter(X[:, 0], X[:, 1])
```

```
Out[4]: <matplotlib.collections.PathCollection at 0x14c57e450>
```



Иерархическая кластеризация выборки

```
In [5]: from scipy.cluster.hierarchy import linkage, dendrogram
```

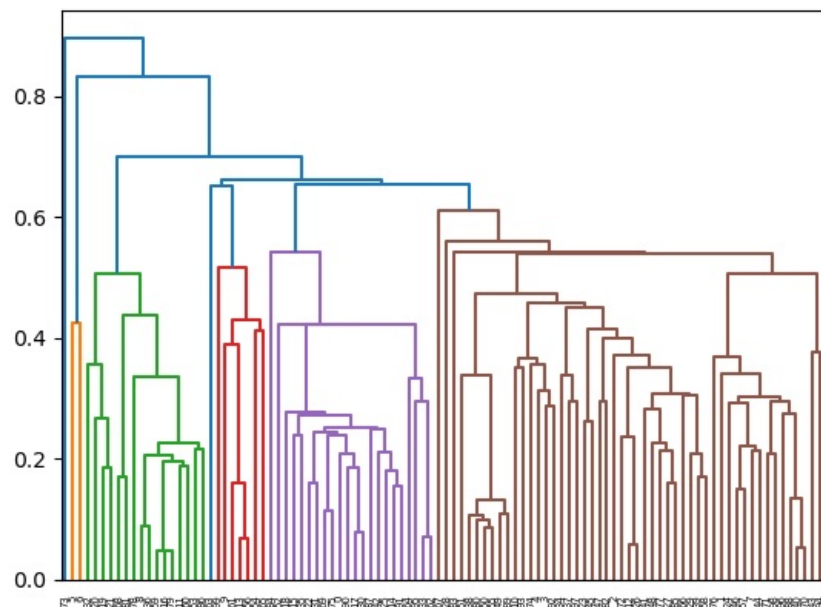
Расстояние ближайшего соседа (single)

```
In [6]: mergings_single = linkage(X, method='single')
mergings_single
```

```
Out[6]: array([[1.60000000e+01, 7.90000000e+01, 4.61700718e-02, 2.00000000e+00],
 [5.90000000e+01, 1.00000000e+02, 4.82382520e-02, 3.00000000e+00],
 [4.00000000e+01, 7.00000000e+01, 5.11405618e-02, 2.00000000e+00],
 [1.20000000e+01, 2.60000000e+01, 5.78546254e-02, 2.00000000e+00],
 [1.30000000e+01, 5.60000000e+01, 6.88632557e-02, 2.00000000e+00],
 [3.30000000e+01, 6.20000000e+01, 7.01583969e-02, 2.00000000e+00],
 [1.70000000e+01, 3.00000000e+01, 7.86585329e-02, 2.00000000e+00],
 [5.00000000e+01, 5.50000000e+01, 8.59038033e-02, 2.00000000e+00],
 [8.00000000e+00, 3.60000000e+01, 8.74654507e-02, 2.00000000e+00],
 [8.00000000e+01, 1.07000000e+02, 9.82539902e-02, 3.00000000e+00],
 [3.80000000e+01, 1.09000000e+02, 1.04612512e-01, 4.00000000e+00],
 [4.90000000e+01, 8.90000000e+01, 1.07241587e-01, 2.00000000e+00],
 [6.90000000e+01, 7.50000000e+01, 1.13102713e-01, 2.00000000e+00],
 [1.10000000e+02, 1.11000000e+02, 1.13180811e-01, 6.00000000e+00],
 [8.80000000e+01, 1.02000000e+02, 1.13359495e-01, 3.00000000e+00],
 [4.60000000e+01, 5.70000000e+01, 1.14845934e-01, 2.00000000e+00],
 [1.40000000e+01, 9.10000000e+01, 1.15353922e-01, 2.00000000e+00],
 [2.20000000e+01, 3.10000000e+01, 1.15815607e-01, 2.00000000e+00],
 [2.70000000e+01, 8.50000000e+01, 1.15829344e-01, 2.00000000e+00],
 [6.10000000e+01, 1.04000000e+02, 1.15912784e-01, 3.00000000e+00],
 [3.90000000e+01, 5.80000000e+01, 1.16834230e-01, 2.00000000e+00],
 [6.80000000e+01, 8.10000000e+01, 1.16999922e-01, 2.00000000e+00],
 [2.50000000e+01, 1.16000000e+02, 1.17887173e-01, 3.00000000e+00],
 [1.90000000e+01, 2.10000000e+01, 1.18491149e-01, 2.00000000e+00],
 [9.00000000e+01, 1.06000000e+02, 1.18513941e-01, 3.00000000e+00],
```

```
[1.10000000e+01, 6.00000000e+01, 1.86558882e-01, 2.00000000e+00],
[1.01000000e+02, 1.25000000e+02, 1.94738613e-01, 5.00000000e+00],
[1.08000000e+02, 1.26000000e+02, 2.04109976e-01, 7.00000000e+00],
[7.10000000e+01, 9.80000000e+01, 2.06745536e-01, 2.00000000e+00],
[0.00000000e+00, 1.24000000e+02, 2.06935827e-01, 4.00000000e+00],
[2.90000000e+01, 1.20000000e+02, 2.08117311e-01, 3.00000000e+00],
[4.20000000e+01, 1.22000000e+02, 2.10499144e-01, 4.00000000e+00],
[7.70000000e+01, 1.18000000e+02, 2.12314011e-01, 3.00000000e+00],
[7.00000000e+00, 4.40000000e+01, 2.13509573e-01, 2.00000000e+00],
[5.30000000e+01, 8.60000000e+01, 2.16317638e-01, 2.00000000e+00],
[1.15000000e+02, 1.33000000e+02, 2.20759246e-01, 4.00000000e+00],
[4.80000000e+01, 1.32000000e+02, 2.24782521e-01, 4.00000000e+00],
[1.27000000e+02, 1.34000000e+02, 2.26441925e-01, 9.00000000e+00],
[7.20000000e+01, 1.03000000e+02, 2.34703789e-01, 3.00000000e+00],
[1.50000000e+01, 3.50000000e+01, 2.37419614e-01, 2.00000000e+00],
[1.12000000e+02, 1.29000000e+02, 2.39218218e-01, 6.00000000e+00],
[1.17000000e+02, 1.40000000e+02, 2.43635255e-01, 8.00000000e+00],
[8.70000000e+01, 1.31000000e+02, 2.49083280e-01, 5.00000000e+00],
[1.41000000e+02, 1.42000000e+02, 2.49909254e-01, 1.30000000e+01],
[2.30000000e+01, 4.50000000e+01, 2.60918487e-01, 2.00000000e+00],
[2.00000000e+01, 1.23000000e+02, 2.65939550e-01, 3.00000000e+00],
[1.39000000e+02, 1.43000000e+02, 2.71195854e-01, 1.50000000e+01],
[9.60000000e+01, 1.14000000e+02, 2.74671575e-01, 4.00000000e+00],
[1.80000000e+01, 1.46000000e+02, 2.77275690e-01, 1.60000000e+01],
[4.10000000e+01, 1.36000000e+02, 2.80684946e-01, 5.00000000e+00],
[3.00000000e+00, 9.20000000e+01, 2.87599093e-01, 2.00000000e+00],
[2.40000000e+01, 1.35000000e+02, 2.92289409e-01, 5.00000000e+00],
[9.50000000e+01, 1.05000000e+02, 2.93946078e-01, 3.00000000e+00],
[4.70000000e+01, 8.20000000e+01, 2.95311525e-01, 2.00000000e+00],
[3.70000000e+01, 9.70000000e+01, 2.95371350e-01, 2.00000000e+00],
[1.28000000e+02, 1.47000000e+02, 2.95567032e-01, 6.00000000e+00],
[1.51000000e+02, 1.55000000e+02, 3.02471733e-01, 1.10000000e+01],
[6.60000000e+01, 1.30000000e+02, 3.04141006e-01, 4.00000000e+00],
[1.49000000e+02, 1.57000000e+02, 3.07845208e-01, 9.00000000e+00],
[4.00000000e+00, 1.50000000e+02, 3.12959059e-01, 3.00000000e+00],
[6.40000000e+01, 1.52000000e+02, 3.32407298e-01, 4.00000000e+00],
[7.80000000e+01, 1.37000000e+02, 3.34931329e-01, 1.00000000e+01],
[3.40000000e+01, 1.54000000e+02, 3.36704966e-01, 3.00000000e+00],
[5.10000000e+01, 1.13000000e+02, 3.38139955e-01, 7.00000000e+00],
[1.00000000e+00, 1.56000000e+02, 3.39115843e-01, 1.20000000e+01],
[1.00000000e+01, 9.30000000e+01, 3.49780562e-01, 2.00000000e+00],
[1.38000000e+02, 1.58000000e+02, 3.50960196e-01, 1.20000000e+01],
[3.20000000e+01, 1.45000000e+02, 3.55225349e-01, 4.00000000e+00],
[7.40000000e+01, 1.59000000e+02, 3.56148347e-01, 4.00000000e+00],
[1.65000000e+02, 1.68000000e+02, 3.66099720e-01, 6.00000000e+00],
[7.60000000e+01, 1.64000000e+02, 3.68137533e-01, 1.30000000e+01],
[2.00000000e+00, 1.66000000e+02, 3.70047569e-01, 1.30000000e+01],
[4.30000000e+01, 9.40000000e+01, 3.75083314e-01, 2.00000000e+00],
[9.00000000e+00, 1.19000000e+02, 3.87521122e-01, 4.00000000e+00],
[1.53000000e+02, 1.71000000e+02, 3.98258187e-01, 1.50000000e+01],
[5.40000000e+01, 6.50000000e+01, 4.11978836e-01, 2.00000000e+00],
[1.44000000e+02, 1.74000000e+02, 4.14732336e-01, 1.70000000e+01],
[1.48000000e+02, 1.60000000e+02, 4.20953957e-01, 2.00000000e+01],
[5.20000000e+01, 1.77000000e+02, 4.21597881e-01, 2.10000000e+01],
[5.00000000e+00, 6.00000000e+00, 4.24065464e-01, 2.00000000e+00],
[1.73000000e+02, 1.75000000e+02, 4.29478776e-01, 6.00000000e+00],
[1.21000000e+02, 1.61000000e+02, 4.37820451e-01, 1.20000000e+01],
[1.62000000e+02, 1.76000000e+02, 4.50137926e-01, 2.00000000e+01],
[1.69000000e+02, 1.82000000e+02, 4.58288161e-01, 2.60000000e+01],
[1.63000000e+02, 1.83000000e+02, 4.72191243e-01, 3.30000000e+01],
[1.70000000e+02, 1.72000000e+02, 5.04807725e-01, 1.50000000e+01],
[1.67000000e+02, 1.81000000e+02, 5.05705507e-01, 1.60000000e+01],
[9.90000000e+01, 1.80000000e+02, 5.16692019e-01, 7.00000000e+00],
[1.84000000e+02, 1.85000000e+02, 5.38657339e-01, 4.80000000e+01],
[8.30000000e+01, 1.88000000e+02, 5.40505699e-01, 4.90000000e+01],
[8.40000000e+01, 1.78000000e+02, 5.42531944e-01, 2.20000000e+01],
[2.80000000e+01, 1.89000000e+02, 5.58934710e-01, 5.00000000e+01],
[6.70000000e+01, 1.91000000e+02, 6.11461373e-01, 5.10000000e+01],
[6.30000000e+01, 1.87000000e+02, 6.51393101e-01, 8.00000000e+00],
[1.90000000e+02, 1.92000000e+02, 6.54227694e-01, 7.30000000e+01],
[1.93000000e+02, 1.94000000e+02, 6.61650913e-01, 8.10000000e+01],
[1.86000000e+02, 1.95000000e+02, 7.00424204e-01, 9.70000000e+01],
[1.79000000e+02, 1.96000000e+02, 8.31875208e-01, 9.90000000e+01],
[7.30000000e+01, 1.97000000e+02, 8.97116304e-01, 1.00000000e+02]]
```

```
In [7]: dendrogram(mergings_single)
plt.show()
```



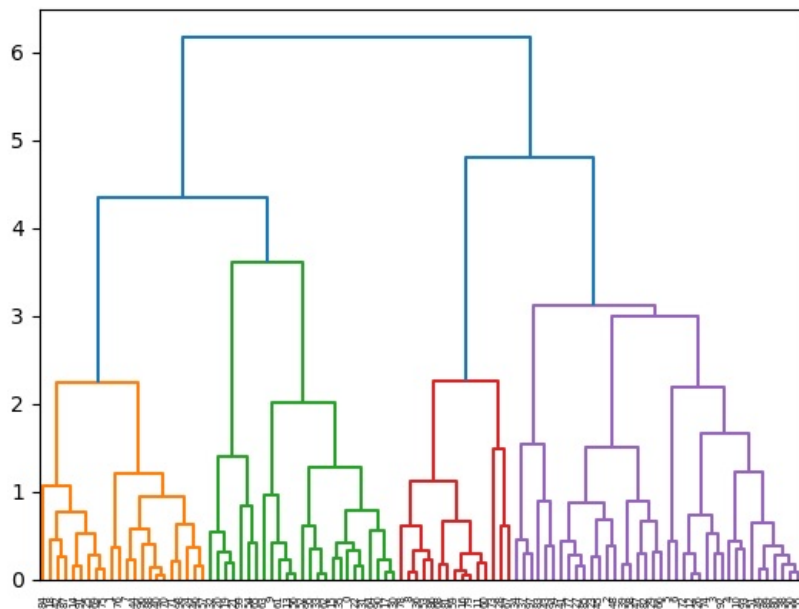
Расстояние дальнего соседа (complete)

```
In [8]: mergings_complete = linkage(X, method='complete')
mergings_complete
```

```
Out[8]: array([[1.60000000e+01, 7.90000000e+01, 4.61700718e-02, 2.00000000e+00],
 [4.00000000e+01, 7.00000000e+01, 5.11405618e-02, 2.00000000e+00],
 [1.20000000e+01, 2.60000000e+01, 5.78546254e-02, 2.00000000e+00],
 [1.30000000e+01, 5.60000000e+01, 6.88632557e-02, 2.00000000e+00],
 [3.30000000e+01, 6.20000000e+01, 7.01583969e-02, 2.00000000e+00],
 [1.70000000e+01, 3.00000000e+01, 7.86585329e-02, 2.00000000e+00],
 [5.00000000e+01, 5.50000000e+01, 8.59038033e-02, 2.00000000e+00],
 [8.00000000e+00, 3.60000000e+01, 8.74654507e-02, 2.00000000e+00],
 [5.90000000e+01, 1.00000000e+02, 9.07472797e-02, 3.00000000e+00],
 [4.90000000e+01, 8.90000000e+01, 1.07241587e-01, 2.00000000e+00],
 [6.90000000e+01, 7.50000000e+01, 1.13102713e-01, 2.00000000e+00],
 [8.80000000e+01, 1.01000000e+02, 1.34292668e-01, 3.00000000e+00],
 [4.60000000e+01, 5.70000000e+01, 1.48459346e-01, 2.00000000e+00],
 [1.40000000e+01, 9.10000000e+01, 1.53539228e-01, 2.00000000e+00],
 [2.20000000e+01, 3.10000000e+01, 1.58156070e-01, 2.00000000e+00],
 [2.70000000e+01, 8.50000000e+01, 1.58293444e-01, 2.00000000e+00],
 [3.90000000e+01, 5.80000000e+01, 1.68342306e-01, 2.00000000e+00],
 [3.80000000e+01, 1.06000000e+02, 1.68541667e-01, 3.00000000e+00],
 [6.80000000e+01, 8.10000000e+01, 1.6999224e-01, 2.00000000e+00],
 [1.90000000e+01, 2.10000000e+01, 1.84911495e-01, 2.00000000e+00],
 [1.10000000e+01, 6.00000000e+01, 1.86558882e-01, 2.00000000e+00],
 [7.10000000e+01, 9.80000000e+01, 2.06745536e-01, 2.00000000e+00],
 [7.00000000e+00, 4.40000000e+01, 2.13509573e-01, 2.00000000e+00],
 [9.00000000e+01, 1.05000000e+02, 2.14127186e-01, 3.00000000e+00],
 [5.30000000e+01, 8.60000000e+01, 2.16317638e-01, 2.00000000e+00],
 [6.10000000e+01, 1.03000000e+02, 2.22720948e-01, 3.00000000e+00],
 [1.50000000e+01, 3.50000000e+01, 2.37419614e-01, 2.00000000e+00],
 [4.20000000e+01, 8.70000000e+01, 2.49083280e-01, 2.00000000e+00],
 [2.30000000e+01, 4.50000000e+01, 2.60918487e-01, 2.00000000e+00],
 [8.00000000e+01, 1.17000000e+02, 2.66426819e-01, 4.00000000e+00],
 [7.70000000e+01, 1.15000000e+02, 2.72834727e-01, 3.00000000e+00],
 [2.50000000e+01, 1.10000000e+02, 2.76879078e-01, 3.00000000e+00],
 [3.00000000e+00, 9.20000000e+01, 2.87599093e-01, 2.00000000e+00],
 [7.20000000e+01, 1.02000000e+02, 2.88198463e-01, 3.00000000e+00],
 [4.70000000e+01, 8.20000000e+01, 2.95311525e-01, 2.00000000e+00],
 [3.70000000e+01, 9.70000000e+01, 2.95371350e-01, 2.00000000e+00],
 [1.08000000e+02, 1.20000000e+02, 2.97303728e-01, 5.00000000e+00],
 [2.00000000e+01, 1.19000000e+02, 3.03833324e-01, 3.00000000e+00],
 [2.90000000e+01, 6.60000000e+01, 3.04141006e-01, 2.00000000e+00],
 [0.00000000e+00, 1.14000000e+02, 3.31587384e-01, 3.00000000e+00],
 [1.07000000e+02, 1.24000000e+02, 3.33316821e-01, 4.00000000e+00],
 [1.00000000e+01, 9.30000000e+01, 3.49780562e-01, 2.00000000e+00],
 [9.50000000e+01, 1.04000000e+02, 3.51109554e-01, 3.00000000e+00],
 [2.40000000e+01, 1.12000000e+02, 3.53217215e-01, 3.00000000e+00],
 [1.00000000e+00, 7.60000000e+01, 3.68137533e-01, 2.00000000e+00],
 [2.00000000e+00, 4.80000000e+01, 3.70047569e-01, 2.00000000e+00],
 [1.09000000e+02, 1.29000000e+02, 3.73593294e-01, 6.00000000e+00],
 [4.30000000e+01, 9.40000000e+01, 3.75083314e-01, 2.00000000e+00],
 [9.60000000e+01, 1.11000000e+02, 3.97795792e-01, 4.00000000e+00],
 [9.00000000e+00, 1.25000000e+02, 4.09144649e-01, 4.00000000e+00],
 [5.40000000e+01, 6.50000000e+01, 4.11978836e-01, 2.00000000e+00],
 [1.26000000e+02, 1.39000000e+02, 4.19870328e-01, 5.00000000e+00],
 [5.00000000e+00, 6.00000000e+00, 4.24065464e-01, 2.00000000e+00],
 [4.10000000e+01, 1.30000000e+02, 4.25788720e-01, 4.00000000e+00],
 [4.00000000e+00, 1.41000000e+02, 4.28073425e-01, 3.00000000e+00],
 [1.80000000e+01, 1.27000000e+02, 4.46178959e-01, 3.00000000e+00],
```

```
[3.40000000e+01, 1.35000000e+02, 4.51613166e-01, 3.00000000e+00],
[7.40000000e+01, 1.32000000e+02, 5.11800996e-01, 3.00000000e+00],
[1.13000000e+02, 1.31000000e+02, 5.26546798e-01, 5.00000000e+00],
[3.20000000e+01, 1.37000000e+02, 5.38170017e-01, 4.00000000e+00],
[6.40000000e+01, 1.23000000e+02, 5.48501450e-01, 4.00000000e+00],
[1.22000000e+02, 1.48000000e+02, 5.63582283e-01, 6.00000000e+00],
[5.20000000e+01, 1.42000000e+02, 6.02801555e-01, 4.00000000e+00],
[7.80000000e+01, 1.40000000e+02, 6.06146132e-01, 5.00000000e+00],
[2.80000000e+01, 6.70000000e+01, 6.11461373e-01, 2.00000000e+00],
[1.21000000e+02, 1.43000000e+02, 6.15271468e-01, 5.00000000e+00],
[5.10000000e+01, 1.46000000e+02, 6.34465467e-01, 7.00000000e+00],
[1.34000000e+02, 1.38000000e+02, 6.61636437e-01, 4.00000000e+00],
[1.18000000e+02, 1.36000000e+02, 6.61846436e-01, 7.00000000e+00],
[1.28000000e+02, 1.45000000e+02, 6.73099165e-01, 4.00000000e+00],
[1.33000000e+02, 1.57000000e+02, 7.28083022e-01, 6.00000000e+00],
[1.55000000e+02, 1.58000000e+02, 7.66528328e-01, 8.00000000e+00],
[1.51000000e+02, 1.60000000e+02, 7.82511188e-01, 9.00000000e+00],
[9.90000000e+01, 1.50000000e+02, 8.31233855e-01, 3.00000000e+00],
[1.53000000e+02, 1.69000000e+02, 8.68782328e-01, 8.00000000e+00],
[8.30000000e+01, 1.47000000e+02, 8.81163241e-01, 3.00000000e+00],
[1.16000000e+02, 1.67000000e+02, 8.81721401e-01, 6.00000000e+00],
[1.61000000e+02, 1.65000000e+02, 9.35090217e-01, 1.00000000e+01],
[6.30000000e+01, 1.49000000e+02, 9.53233632e-01, 5.00000000e+00],
[8.40000000e+01, 1.71000000e+02, 1.06015086e+00, 9.00000000e+00],
[1.63000000e+02, 1.68000000e+02, 1.11366408e+00, 1.20000000e+01],
[1.44000000e+02, 1.77000000e+02, 1.19969328e+00, 1.30000000e+01],
[1.54000000e+02, 1.66000000e+02, 1.22550936e+00, 1.00000000e+01],
[1.62000000e+02, 1.72000000e+02, 1.27143591e+00, 1.30000000e+01],
[1.59000000e+02, 1.73000000e+02, 1.39929379e+00, 7.00000000e+00],
[7.30000000e+01, 1.64000000e+02, 1.48789287e+00, 3.00000000e+00],
[1.74000000e+02, 1.76000000e+02, 1.50184640e+00, 1.40000000e+01],
[1.56000000e+02, 1.75000000e+02, 1.53795858e+00, 6.00000000e+00],
[1.70000000e+02, 1.82000000e+02, 1.65120487e+00, 1.60000000e+01],
[1.78000000e+02, 1.83000000e+02, 2.01117697e+00, 1.80000000e+01],
[1.52000000e+02, 1.88000000e+02, 2.18538625e+00, 1.80000000e+01],
[1.79000000e+02, 1.81000000e+02, 2.23326880e+00, 2.20000000e+01],
[1.80000000e+02, 1.85000000e+02, 2.25879519e+00, 1.50000000e+01],
[1.86000000e+02, 1.90000000e+02, 2.99903057e+00, 3.20000000e+01],
[1.87000000e+02, 1.93000000e+02, 3.10920040e+00, 3.80000000e+01],
[1.84000000e+02, 1.89000000e+02, 3.61152716e+00, 2.50000000e+01],
[1.91000000e+02, 1.95000000e+02, 4.35299476e+00, 4.70000000e+01],
[1.92000000e+02, 1.94000000e+02, 4.79367466e+00, 5.30000000e+01],
[1.96000000e+02, 1.97000000e+02, 6.17998796e+00, 1.00000000e+02]]
```

```
In [9]: dendrogram(mergings_complete)
plt.show()
```



Расстояние Уорда (Ward)

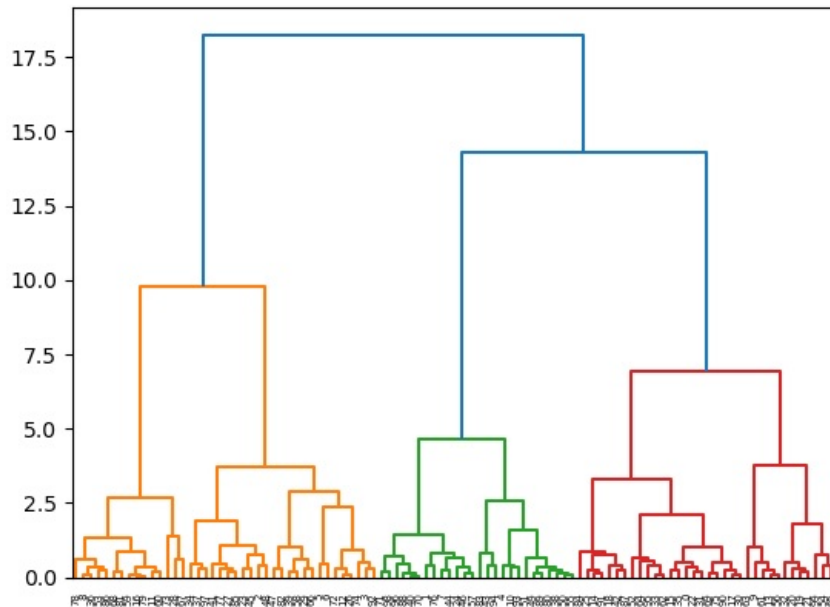
```
In [10]: mergings_ward = linkage(X, method='ward')
mergings_ward
```

```
Out[10]: array([[1.60000000e+01, 7.90000000e+01, 4.61700718e-02, 2.00000000e+00],
 [4.00000000e+01, 7.00000000e+01, 5.11405618e-02, 2.00000000e+00],
 [1.20000000e+01, 2.60000000e+01, 5.78546254e-02, 2.00000000e+00],
 [1.30000000e+01, 5.60000000e+01, 6.88632557e-02, 2.00000000e+00],
 [3.30000000e+01, 6.20000000e+01, 7.01583969e-02, 2.00000000e+00],
 [1.70000000e+01, 3.00000000e+01, 7.86585329e-02, 2.00000000e+00],
 [5.90000000e+01, 1.00000000e+02, 7.95661568e-02, 3.00000000e+00],
 [5.00000000e+01, 5.50000000e+01, 8.59038033e-02, 2.00000000e+00],
 [8.00000000e+00, 3.60000000e+01, 8.74654507e-02, 2.00000000e+00],
```

[4.90000000e+01, 8.90000000e+01, 1.07241587e-01, 2.00000000e+00],
[6.90000000e+01, 7.50000000e+01, 1.13102713e-01, 2.00000000e+00],
[4.60000000e+01, 5.70000000e+01, 1.48459346e-01, 2.00000000e+00],
[8.80000000e+01, 1.01000000e+02, 1.51821067e-01, 3.00000000e+00],
[1.40000000e+01, 9.10000000e+01, 1.53539228e-01, 2.00000000e+00],
[3.80000000e+01, 1.07000000e+02, 1.54186769e-01, 3.00000000e+00],
[2.20000000e+01, 3.10000000e+01, 1.58156070e-01, 2.00000000e+00],
[2.70000000e+01, 8.50000000e+01, 1.58293444e-01, 2.00000000e+00],
[3.90000000e+01, 5.80000000e+01, 1.68342306e-01, 2.00000000e+00],
[6.80000000e+01, 8.10000000e+01, 1.69999224e-01, 2.00000000e+00],
[1.90000000e+01, 2.10000000e+01, 1.84911495e-01, 2.00000000e+00],
[1.10000000e+01, 6.00000000e+01, 1.86558882e-01, 2.00000000e+00],
[7.10000000e+01, 9.80000000e+01, 2.06745536e-01, 2.00000000e+00],
[7.00000000e+00, 4.40000000e+01, 2.13509573e-01, 2.00000000e+00],
[5.30000000e+01, 8.60000000e+01, 2.16317638e-01, 2.00000000e+00],
[8.00000000e+01, 1.14000000e+02, 2.18570442e-01, 4.00000000e+00],
[6.10000000e+01, 1.03000000e+02, 2.19932139e-01, 3.00000000e+00],
[9.00000000e+01, 1.05000000e+02, 2.26617847e-01, 3.00000000e+00],
[1.50000000e+01, 3.50000000e+01, 2.37419614e-01, 2.00000000e+00],
[4.20000000e+01, 8.70000000e+01, 2.49083280e-01, 2.00000000e+00],
[2.50000000e+01, 1.13000000e+02, 2.57432492e-01, 3.00000000e+00],
[2.30000000e+01, 4.50000000e+01, 2.60918487e-01, 2.00000000e+00],
[7.70000000e+01, 1.16000000e+02, 2.67067557e-01, 3.00000000e+00],
[3.00000000e+00, 9.20000000e+01, 2.87599093e-01, 2.00000000e+00],
[4.70000000e+01, 8.20000000e+01, 2.95311525e-01, 2.00000000e+00],
[3.70000000e+01, 9.70000000e+01, 2.95371350e-01, 2.00000000e+00],
[7.20000000e+01, 1.02000000e+02, 3.01629622e-01, 3.00000000e+00],
[2.90000000e+01, 6.60000000e+01, 3.04141006e-01, 2.00000000e+00],
[2.00000000e+01, 1.19000000e+02, 3.11921360e-01, 3.00000000e+00],
[0.00000000e+00, 1.15000000e+02, 3.23317862e-01, 3.00000000e+00],
[1.00000000e+01, 9.30000000e+01, 3.49780562e-01, 2.00000000e+00],
[1.09000000e+02, 1.24000000e+02, 3.51253976e-01, 6.00000000e+00],
[1.06000000e+02, 1.20000000e+02, 3.53267428e-01, 5.00000000e+00],
[1.08000000e+02, 1.23000000e+02, 3.54008337e-01, 4.00000000e+00],
[2.40000000e+01, 1.11000000e+02, 3.64394851e-01, 3.00000000e+00],
[1.00000000e+00, 7.60000000e+01, 3.68137533e-01, 2.00000000e+00],
[2.00000000e+00, 4.80000000e+01, 3.70047569e-01, 2.00000000e+00],
[9.50000000e+01, 1.04000000e+02, 3.71681867e-01, 3.00000000e+00],
[4.30000000e+01, 9.40000000e+01, 3.75083314e-01, 2.00000000e+00],
[1.80000000e+01, 1.28000000e+02, 4.04092590e-01, 3.00000000e+00],
[5.40000000e+01, 6.50000000e+01, 4.11978836e-01, 2.00000000e+00],
[4.00000000e+00, 1.39000000e+02, 4.13201033e-01, 3.00000000e+00],
[5.00000000e+00, 6.00000000e+00, 4.24065464e-01, 2.00000000e+00],
[9.60000000e+01, 1.12000000e+02, 4.25435672e-01, 4.00000000e+00],
[3.40000000e+01, 1.34000000e+02, 4.27163176e-01, 3.00000000e+00],
[4.10000000e+01, 1.31000000e+02, 4.40111082e-01, 4.00000000e+00],
[9.00000000e+00, 1.25000000e+02, 4.76139480e-01, 4.00000000e+00],
[7.40000000e+01, 1.32000000e+02, 4.81265890e-01, 3.00000000e+00],
[1.27000000e+02, 1.38000000e+02, 4.83717448e-01, 5.00000000e+00],
[1.10000000e+02, 1.26000000e+02, 5.13982302e-01, 5.00000000e+00],
[3.20000000e+01, 1.37000000e+02, 5.18295628e-01, 4.00000000e+00],
[6.40000000e+01, 1.46000000e+02, 5.49647693e-01, 4.00000000e+00],
[7.80000000e+01, 1.42000000e+02, 5.97721106e-01, 5.00000000e+00],
[2.80000000e+01, 6.70000000e+01, 6.11461373e-01, 2.00000000e+00],
[1.17000000e+02, 1.36000000e+02, 6.13537176e-01, 4.00000000e+00],
[5.20000000e+01, 1.60000000e+02, 6.31940297e-01, 5.00000000e+00],
[1.22000000e+02, 1.43000000e+02, 6.38032364e-01, 5.00000000e+00],
[5.10000000e+01, 1.40000000e+02, 6.43048891e-01, 7.00000000e+00],
[1.21000000e+02, 1.52000000e+02, 7.01595824e-01, 6.00000000e+00],
[1.30000000e+02, 1.45000000e+02, 7.36528468e-01, 4.00000000e+00],
[1.29000000e+02, 1.48000000e+02, 7.5522273e-01, 6.00000000e+00],
[9.90000000e+01, 1.49000000e+02, 7.62914186e-01, 3.00000000e+00],
[8.30000000e+01, 1.47000000e+02, 8.15782374e-01, 3.00000000e+00],
[1.44000000e+02, 1.65000000e+02, 8.28587610e-01, 7.00000000e+00],
[8.40000000e+01, 1.69000000e+02, 8.65686951e-01, 7.00000000e+00],
[1.18000000e+02, 1.41000000e+02, 8.78734555e-01, 7.00000000e+00],
[1.35000000e+02, 1.56000000e+02, 9.14424563e-01, 6.00000000e+00],
[1.33000000e+02, 1.63000000e+02, 1.00854830e+00, 6.00000000e+00],
[6.30000000e+01, 1.55000000e+02, 1.01739222e+00, 5.00000000e+00],
[1.57000000e+02, 1.58000000e+02, 1.02340067e+00, 1.00000000e+01],
[1.54000000e+02, 1.68000000e+02, 1.04684830e+00, 8.00000000e+00],
[1.61000000e+02, 1.74000000e+02, 1.32346910e+00, 1.20000000e+01],
[7.30000000e+01, 1.62000000e+02, 1.37397243e+00, 3.00000000e+00],
[1.67000000e+02, 1.72000000e+02, 1.44309332e+00, 1.30000000e+01],
[1.50000000e+02, 1.66000000e+02, 1.57228950e+00, 1.00000000e+01],
[1.59000000e+02, 1.70000000e+02, 1.80732095e+00, 7.00000000e+00],
[1.53000000e+02, 1.79000000e+02, 1.91354883e+00, 1.10000000e+01],
[1.64000000e+02, 1.78000000e+02, 2.12087664e+00, 1.50000000e+01],
[1.51000000e+02, 1.75000000e+02, 2.37678395e+00, 8.00000000e+00],
[1.71000000e+02, 1.83000000e+02, 2.5558849e+00, 1.30000000e+01],
[1.80000000e+02, 1.81000000e+02, 2.64725074e+00, 1.50000000e+01],
[1.76000000e+02, 1.87000000e+02, 2.87700067e+00, 1.40000000e+01],
[1.73000000e+02, 1.86000000e+02, 3.28810096e+00, 2.20000000e+01],
[1.85000000e+02, 1.90000000e+02, 3.70398026e+00, 2.50000000e+01],
[1.77000000e+02, 1.84000000e+02, 3.74722389e+00, 1.20000000e+01],
[1.82000000e+02, 1.88000000e+02, 4.65440256e+00, 2.60000000e+01],
[1.91000000e+02, 1.93000000e+02, 6.90126715e+00, 3.40000000e+01],
[1.89000000e+02, 1.92000000e+02, 9.77245930e+00, 4.00000000e+01],
[1.94000000e+02, 1.95000000e+02, 1.42960671e+01, 6.00000000e+01],

```
[1.96000000e+02, 1.97000000e+02, 1.82581285e+01, 1.00000000e+02]])
```

```
In [11]: dendrogram(mergings_ward)
plt.show()
```



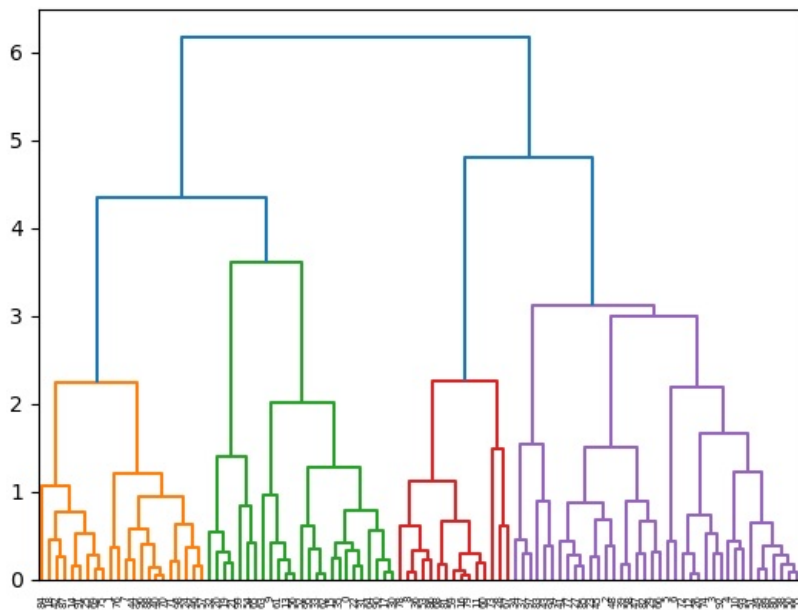
Выбор лучшего разбиения

```
In [12]: mergings_complete = linkage(X, method='complete')
mergings_complete
```

```
Out[12]: array([[1.60000000e+01, 7.90000000e+01, 4.61700718e-02, 2.00000000e+00],
 [4.00000000e+01, 7.00000000e+01, 5.11405618e-02, 2.00000000e+00],
 [1.20000000e+01, 2.60000000e+01, 5.78546254e-02, 2.00000000e+00],
 [1.30000000e+01, 5.60000000e+01, 6.88632557e-02, 2.00000000e+00],
 [3.30000000e+01, 6.20000000e+01, 7.01583969e-02, 2.00000000e+00],
 [1.70000000e+01, 3.00000000e+01, 7.86585329e-02, 2.00000000e+00],
 [5.00000000e+01, 5.50000000e+01, 8.59038033e-02, 2.00000000e+00],
 [8.00000000e+00, 3.60000000e+01, 8.74654507e-02, 2.00000000e+00],
 [5.90000000e+01, 1.00000000e+02, 9.07472797e-02, 3.00000000e+00],
 [4.90000000e+01, 8.90000000e+01, 1.07241587e-01, 2.00000000e+00],
 [6.90000000e+01, 7.50000000e+01, 1.13102713e-01, 2.00000000e+00],
 [8.80000000e+01, 1.01000000e+02, 1.34292668e-01, 3.00000000e+00],
 [4.60000000e+01, 5.70000000e+01, 1.48459346e-01, 2.00000000e+00],
 [1.40000000e+01, 9.10000000e+01, 1.53539228e-01, 2.00000000e+00],
 [2.20000000e+01, 3.10000000e+01, 1.58156070e-01, 2.00000000e+00],
 [2.70000000e+01, 8.50000000e+01, 1.58293444e-01, 2.00000000e+00],
 [3.90000000e+01, 5.80000000e+01, 1.68342306e-01, 2.00000000e+00],
 [3.80000000e+01, 1.06000000e+02, 1.68541667e-01, 3.00000000e+00],
 [6.80000000e+01, 8.10000000e+01, 1.69999224e-01, 2.00000000e+00],
 [1.90000000e+01, 2.10000000e+01, 1.84911495e-01, 2.00000000e+00],
 [1.10000000e+01, 6.00000000e+01, 1.86558882e-01, 2.00000000e+00],
 [7.10000000e+01, 9.80000000e+01, 2.06745536e-01, 2.00000000e+00],
 [7.00000000e+00, 4.40000000e+01, 2.13509573e-01, 2.00000000e+00],
 [9.00000000e+01, 1.05000000e+02, 2.14127186e-01, 3.00000000e+00],
 [5.30000000e+01, 8.60000000e+01, 2.16317638e-01, 2.00000000e+00],
 [6.10000000e+01, 1.03000000e+02, 2.22720948e-01, 3.00000000e+00],
 [1.50000000e+01, 3.50000000e+01, 2.37419614e-01, 2.00000000e+00],
 [4.20000000e+01, 8.70000000e+01, 2.49083280e-01, 2.00000000e+00],
 [2.30000000e+01, 4.50000000e+01, 2.60918487e-01, 2.00000000e+00],
 [8.00000000e+01, 1.17000000e+02, 2.66426819e-01, 4.00000000e+00],
 [7.70000000e+01, 1.15000000e+02, 2.72834727e-01, 3.00000000e+00],
 [2.50000000e+01, 1.10000000e+02, 2.76879078e-01, 3.00000000e+00],
 [3.00000000e+00, 9.20000000e+01, 2.87599093e-01, 2.00000000e+00],
 [7.20000000e+01, 1.02000000e+02, 2.88198463e-01, 3.00000000e+00],
 [4.70000000e+01, 8.20000000e+01, 2.95311525e-01, 2.00000000e+00],
 [3.70000000e+01, 9.70000000e+01, 2.95371350e-01, 2.00000000e+00],
 [1.08000000e+02, 1.20000000e+02, 2.97303728e-01, 5.00000000e+00],
 [2.00000000e+01, 1.19000000e+02, 3.03833324e-01, 3.00000000e+00],
 [2.90000000e+01, 6.60000000e+01, 3.04141006e-01, 2.00000000e+00],
 [0.00000000e+00, 1.14000000e+02, 3.31587384e-01, 3.00000000e+00],
 [1.07000000e+02, 1.24000000e+02, 3.33316821e-01, 4.00000000e+00],
 [1.00000000e+01, 9.30000000e+01, 3.49780562e-01, 2.00000000e+00],
 [9.50000000e+01, 1.04000000e+02, 3.51109554e-01, 3.00000000e+00],
 [2.40000000e+01, 1.12000000e+02, 3.53217215e-01, 3.00000000e+00],
 [1.00000000e+00, 7.60000000e+01, 3.68137533e-01, 2.00000000e+00],
 [2.00000000e+00, 4.80000000e+01, 3.70047569e-01, 2.00000000e+00],
 [1.09000000e+02, 1.29000000e+02, 3.73593294e-01, 6.00000000e+00],
 [4.30000000e+01, 9.40000000e+01, 3.75083314e-01, 2.00000000e+00],
 [9.60000000e+01, 1.11000000e+02, 3.97795792e-01, 4.00000000e+00],
 [9.00000000e+00, 1.25000000e+02, 4.09144649e-01, 4.00000000e+00],
 [5.40000000e+01, 6.50000000e+01, 4.11978836e-01, 2.00000000e+00],
```

```
[1.26000000e+02, 1.39000000e+02, 4.19870328e-01, 5.00000000e+00],
[5.00000000e+00, 6.00000000e+00, 4.24065464e-01, 2.00000000e+00],
[4.10000000e+01, 1.30000000e+02, 4.25788720e-01, 4.00000000e+00],
[4.00000000e+00, 1.41000000e+02, 4.28073425e-01, 3.00000000e+00],
[1.80000000e+01, 1.27000000e+02, 4.46178959e-01, 3.00000000e+00],
[3.40000000e+01, 1.35000000e+02, 4.51613166e-01, 3.00000000e+00],
[7.40000000e+01, 1.32000000e+02, 5.11800996e-01, 3.00000000e+00],
[1.13000000e+02, 1.31000000e+02, 5.26546798e-01, 5.00000000e+00],
[3.20000000e+01, 1.37000000e+02, 5.38170017e-01, 4.00000000e+00],
[6.40000000e+01, 1.23000000e+02, 5.48501450e-01, 4.00000000e+00],
[1.22000000e+02, 1.48000000e+02, 5.63582283e-01, 6.00000000e+00],
[5.20000000e+01, 1.42000000e+02, 6.02801555e-01, 4.00000000e+00],
[7.80000000e+01, 1.40000000e+02, 6.06146132e-01, 5.00000000e+00],
[2.80000000e+01, 6.70000000e+01, 6.11461373e-01, 2.00000000e+00],
[1.21000000e+02, 1.43000000e+02, 6.15271468e-01, 5.00000000e+00],
[5.10000000e+01, 1.46000000e+02, 6.34465467e-01, 7.00000000e+00],
[1.34000000e+02, 1.38000000e+02, 6.61636437e-01, 4.00000000e+00],
[1.18000000e+02, 1.36000000e+02, 6.61846436e-01, 7.00000000e+00],
[1.28000000e+02, 1.45000000e+02, 6.73099165e-01, 4.00000000e+00],
[1.33000000e+02, 1.57000000e+02, 7.28083022e-01, 6.00000000e+00],
[1.55000000e+02, 1.58000000e+02, 7.66528328e-01, 8.00000000e+00],
[1.51000000e+02, 1.60000000e+02, 7.82511188e-01, 9.00000000e+00],
[9.90000000e+01, 1.50000000e+02, 8.31233855e-01, 3.00000000e+00],
[1.53000000e+02, 1.69000000e+02, 8.68782328e-01, 8.00000000e+00],
[8.30000000e+01, 1.47000000e+02, 8.81163241e-01, 3.00000000e+00],
[1.16000000e+02, 1.67000000e+02, 8.81721401e-01, 6.00000000e+00],
[1.61000000e+02, 1.65000000e+02, 9.35090217e-01, 1.10000000e+01],
[6.30000000e+01, 1.49000000e+02, 9.53233632e-01, 5.00000000e+00],
[8.40000000e+01, 1.71000000e+02, 1.06015086e+00, 9.00000000e+00],
[1.63000000e+02, 1.68000000e+02, 1.11366408e+00, 1.20000000e+01],
[1.44000000e+02, 1.77000000e+02, 1.19969328e+00, 1.30000000e+01],
[1.54000000e+02, 1.66000000e+02, 1.22550936e+00, 1.00000000e+01],
[1.62000000e+02, 1.72000000e+02, 1.27143591e+00, 1.30000000e+01],
[1.59000000e+02, 1.73000000e+02, 1.39929379e+00, 7.00000000e+00],
[7.30000000e+01, 1.64000000e+02, 1.48789287e+00, 3.00000000e+00],
[1.74000000e+02, 1.76000000e+02, 1.50184640e+00, 1.40000000e+01],
[1.56000000e+02, 1.75000000e+02, 1.53795858e+00, 6.00000000e+00],
[1.70000000e+02, 1.82000000e+02, 1.65120487e+00, 1.60000000e+01],
[1.78000000e+02, 1.83000000e+02, 2.01117697e+00, 1.80000000e+01],
[1.52000000e+02, 1.88000000e+02, 2.18538625e+00, 1.80000000e+01],
[1.79000000e+02, 1.81000000e+02, 2.23326880e+00, 2.20000000e+01],
[1.80000000e+02, 1.85000000e+02, 2.25879519e+00, 1.50000000e+01],
[1.86000000e+02, 1.90000000e+02, 2.99903057e+00, 3.20000000e+01],
[1.87000000e+02, 1.93000000e+02, 3.10920040e+00, 3.80000000e+01],
[1.84000000e+02, 1.89000000e+02, 3.61152716e+00, 2.50000000e+01],
[1.91000000e+02, 1.95000000e+02, 4.35299476e+00, 4.70000000e+01],
[1.92000000e+02, 1.94000000e+02, 4.79367466e+00, 5.30000000e+01],
[1.96000000e+02, 1.97000000e+02, 6.17998796e+00, 1.00000000e+02]]
```

```
In [13]: dendrogram(mergings_complete)
plt.show()
```



```
In [14]: import numpy as np

def update_cluster_centers(X, c):
    centers = np.zeros((4, 2))
    for i in range(1, 5):
        ix = np.where(c == i)
        centers[i - 1, :] = np.mean(X[ix, :], axis=1)
    return centers
```

```
In [15]: from scipy.cluster.hierarchy import fcluster
```

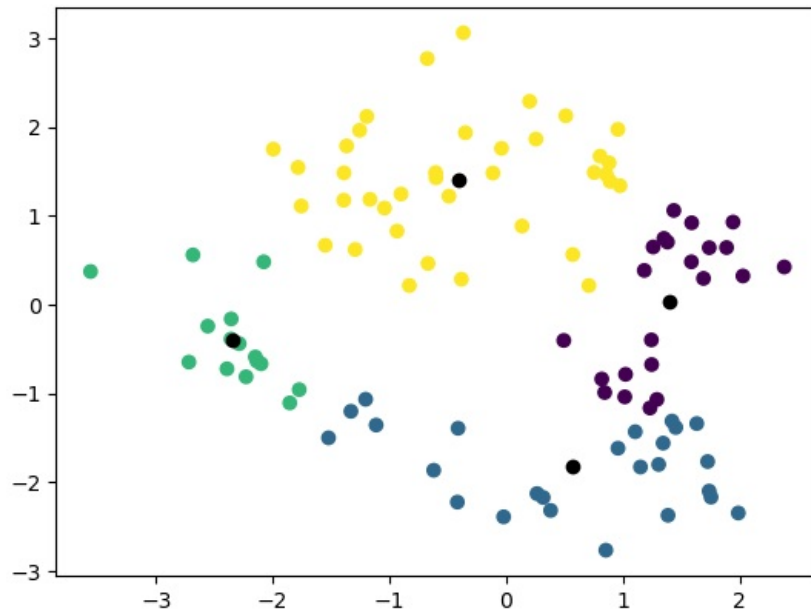


```
In [16]: T = fcluster(mergings_complete, 4, criterion='maxclust')
clusters = update_cluster_centers(X, T)
clusters
```

```
Out[16]: array([[ 1.39460673,  0.03669046],
 [ 0.56596802, -1.81710257],
 [-2.34856748, -0.39843064],
 [-0.41133245,  1.40697805]])
```

```
In [17]: plt.scatter(X[:, 0], X[:, 1], c=T)
plt.scatter(clusters[:, 0], clusters[:, 1], c='black')
```

```
Out[17]: <matplotlib.collections.PathCollection at 0x14c7c2390>
```



Вычисление характеристик

```
In [18]: from sklearn.metrics.pairwise import euclidean_distances
```

```
In [19]: sum_sq_dist = np.zeros(4)
for i in range(1, 5):
    ix = np.where(T == i)
    sum_sq_dist[i - 1] = np.sum(euclidean_distances(*X[ix, :], [clusters[i - 1]]) ** 2)
sum_sq_dist = np.sum(sum_sq_dist) / 4
sum_sq_dist
```

```
Out[19]: 25.843203994463714
```

```
In [20]: sum_avg_intercluster_dist = np.zeros(4)
for i in range(1, 5):
    ix = np.where(T == i)
    sum_avg_intercluster_dist[i - 1] = np.sum(euclidean_distances(*X[ix, :], [clusters[i - 1]]) ** 2) / len(*X[ix, :])
sum_avg_intercluster_dist = np.sum(sum_avg_intercluster_dist) / 4
sum_avg_intercluster_dist
```

```
Out[20]: 0.9399837208674937
```

```
In [21]: sum_intercluster_dist = np.sum(euclidean_distances(clusters, clusters))
sum_intercluster_dist
```

```
Out[21]: 34.64882792358676
```

Кластеризация выборки методом k-средних

```
In [34]: from sklearn.cluster import KMeans

import warnings
warnings.simplefilter(action='ignore', category=FutureWarning)
```

```
In [35]: models = []
predicted_values = []

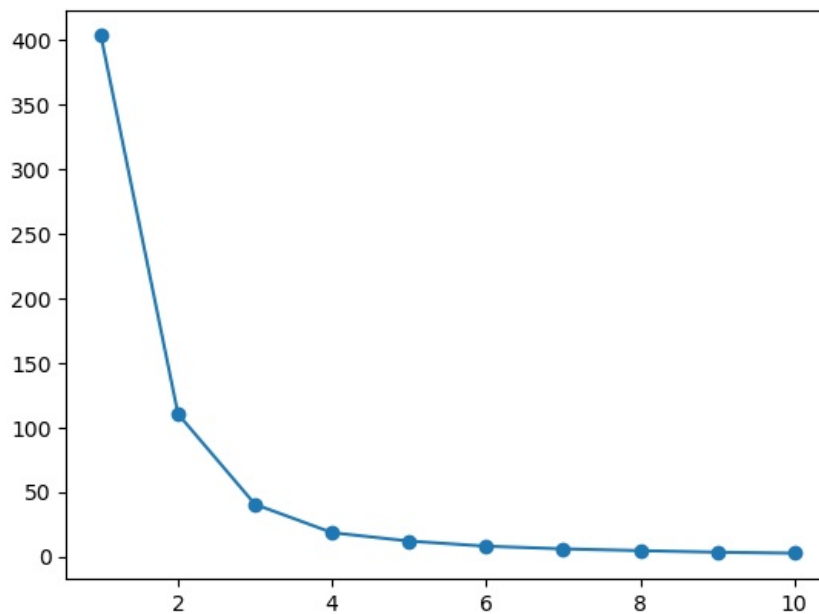
for k in range(1, 11):
    kmeans = KMeans(n_clusters=k)
    kmeans.fit(X)
    models.append(kmeans)
    predicted_values.append(kmeans.predict(X))
```

```
In [24]: sum_sq_dist_avg = []
for it, kmean in enumerate(models):
    sum_sq_dist_avg.append(kmean.inertia_ / (it + 1))
sum_sq_dist_avg
```

```
Out[24]: [403.07111622330194,
110.30529486555724,
40.34514738171166,
18.42231088866526,
11.920247809909446,
7.940148267162104,
5.895817325260838,
4.51018103443448,
3.291952979883026,
2.5322764756413925]
```

```
In [25]: plt.plot(range(1, 11), sum_sq_dist_avg, '-o')
```

```
Out[25]: [<matplotlib.lines.Line2D at 0x14dd301d0>]
```



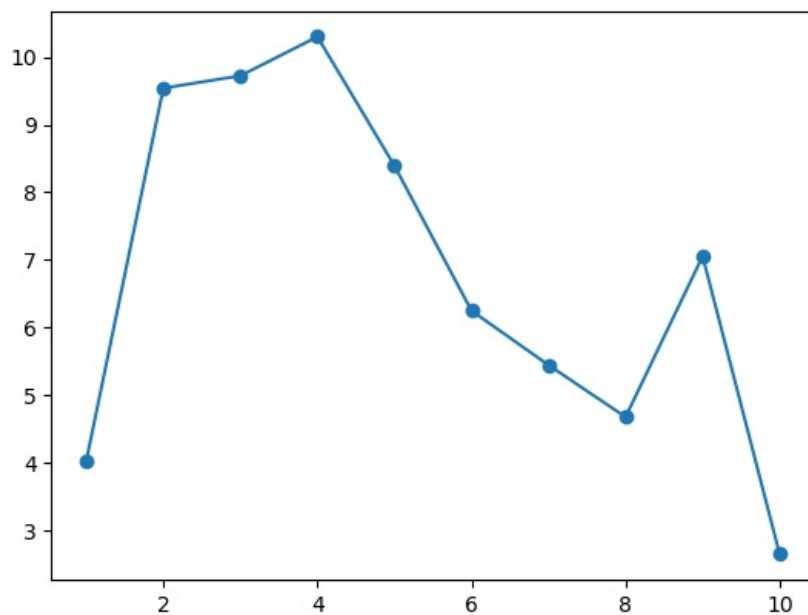
```
In [26]: # Средней суммы средних внутрикластерных расстояний
new_centers = [kmean.cluster_centers_ for kmean in models]

sum_avg_intercluster_dist_avg = []
for k, kmean in enumerate(models):
    intercluster_sum = np.zeros(4)
    for i in range(4):
        ix = np.where(predicted_values[k] == i)
        if len(ix[0]) == 0:
            intercluster_sum[i - 1] = 0
        else:
            intercluster_sum[i - 1] = np.sum(euclidean_distances(*X[ix, :], [kmean.cluster_centers_[i - 1]])) **
sum_avg_intercluster_dist_avg.append(np.sum(intercluster_sum) / (k + 1))
sum_avg_intercluster_dist_avg
```

```
Out[26]: [4.030711162233019,
9.539818208956952,
9.722873156214417,
10.299566956099865,
8.391771367689321,
6.249375960463264,
5.445107229284192,
4.677511901785664,
7.05546172286521,
2.649105068258345]
```

```
In [27]: plt.plot(range(1, 11), sum_avg_intercluster_dist_avg, '-o')
```

```
Out[27]: [<matplotlib.lines.Line2D at 0x14e157350>]
```



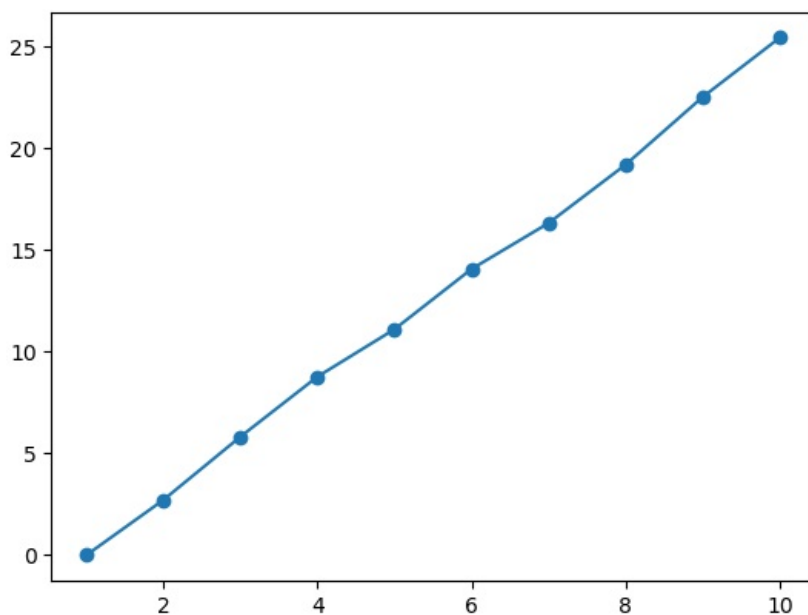
```
In [28]: # Средней суммы межкластерных расстояний от количества кластеров
sum_intercluster_dist_avg = []

for k, kmean in enumerate(models):
    value = np.sum(euclidean_distances(kmean.cluster_centers_, kmean.cluster_centers_))
    sum_intercluster_dist_avg.append(value / (k + 1))
sum_intercluster_dist_avg
```

```
Out[28]: [0.0,
2.7151689605865834,
5.824907427629434,
8.778857420489008,
11.11796453200029,
14.073958288905624,
16.333348130911535,
19.206850733209073,
22.532207165164277,
25.448612512468152]
```

```
In [29]: plt.plot(range(1, 11), sum_intercluster_dist_avg, '-o')
```

```
Out[29]: [<matplotlib.lines.Line2D at 0x14e1a7150>]
```



Составление сравнительной таблицы

```
In [30]: import pandas as pd

In [31]: columns = pd.MultiIndex.from_product([['Иерархический метод', 'Метод k-средних'],
                                              ['Сумма квадратов расстояний до центроида', 'Сумма средних внутрикластерных расстояний', 'Сумма межкластерных расстояний'])
df = pd.DataFrame(columns=columns)
df

Out[31]:
```

	Иерархический метод			Метод k-средних		
	Сумма квадратов расстояний до центроида	Сумма средних внутрикластерных расстояний	Сумма межкластерных расстояний	Сумма квадратов расстояний до центроида	Сумма средних внутрикластерных расстояний	Сумма межкластерных расстояний

```
In [32]: df['Иерархический метод', 'Сумма квадратов расстояний до центроида'] = [sum_sq_dist for _ in range(len(sum_sq_dist))]
df['Иерархический метод', 'Сумма средних внутрикластерных расстояний'] = [sum_avg_intercluster_dist for _ in range(len(sum_avg_intercluster_dist))]
df['Иерархический метод', 'Сумма межкластерных расстояний'] = [sum_intercluster_dist for _ in range(len(sum_intercluster_dist))]

df['Метод k-средних', 'Сумма квадратов расстояний до центроида'] = sum_sq_dist_avg
df['Метод k-средних', 'Сумма средних внутрикластерных расстояний'] = sum_avg_intercluster_dist_avg
df['Метод k-средних', 'Сумма межкластерных расстояний'] = sum_intercluster_dist_avg

df

Out[32]:
```

	Иерархический метод			Метод k-средних		
	Сумма квадратов расстояний до центроида	Сумма средних внутрикластерных расстояний	Сумма межкластерных расстояний	Сумма квадратов расстояний до центроида	Сумма средних внутрикластерных расстояний	Сумма межкластерных расстояний
0	25.843204	0.939984	34.648828	403.071116	4.030711	0.000000
1	25.843204	0.939984	34.648828	110.305295	9.539818	2.715169
2	25.843204	0.939984	34.648828	40.345147	9.722873	5.824907
3	25.843204	0.939984	34.648828	18.422311	10.299567	8.778857
4	25.843204	0.939984	34.648828	11.920248	8.391771	11.111796
5	25.843204	0.939984	34.648828	7.940148	6.249376	14.073958
6	25.843204	0.939984	34.648828	5.895817	5.445107	16.333348
7	25.843204	0.939984	34.648828	4.510181	4.677512	19.206851
8	25.843204	0.939984	34.648828	3.291953	7.055462	22.532207
9	25.843204	0.939984	34.648828	2.532276	2.649105	25.448613

```
In [33]: df.to_excel('result.xlsx')
```