

№1 (7.7 11.03. Решение задач №27 (ч.1) 7-8)

Учёный решил провести кластеризацию некоторого множества звёзд по их расположению на карте звёздного неба. Кластер звёзд – это набор звёзд (точек) на графике. Каждая звезда обязательно принадлежит только одному из кластеров. Истинный центр кластера, или центроид, – это одна из звёзд на графике, сумма расстояний от которой до всех остальных звёзд кластера минимальна.

Под расстоянием понимается расстояние Евклида между двумя точками $A(x_1, y_1)$ и $B(x_2, y_2)$ на плоскости, которое вычисляется по формуле: $d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

В файле А хранятся данные о звёздах двух кластеров. В каждой строке записана информация о расположении на карте одной звезды: сначала координата x , затем координата y (в условных единицах). Известно, что количество звёзд не превышает 1000. В файле Б хранятся данные о звёздах трёх кластеров. Известно, что количество звёзд не превышает 10 000. Структура хранения информации о звездах в файле Б аналогична файлу А.

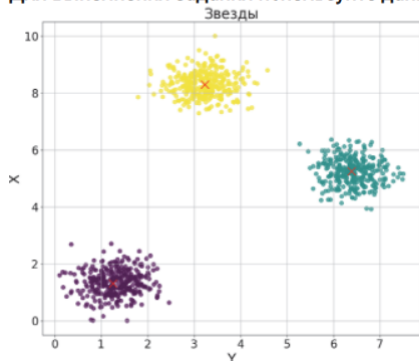
Для каждого файла определите координаты центра каждого кластера, затем вычислите два числа: P_x – среднее арифметическое абсцисс центров кластеров, и P_y – среднее арифметическое ординат центров кластеров.

В ответе запишите четыре числа: в первой строке сначала целую часть произведения $P_x \times 10000$, затем целую часть произведения $P_y \times 10000$ для файла А, во второй строке – аналогичные данные для файла В.

Возможные данные одного из файлов иллюстрированы графиком.

Внимание! График приведён в иллюстративных целях для произвольных значений, не имеющих отношения к заданию.

Для выполнения задания используйте данные из прилагаемого файла.



ОТВЕТ:

5317 32715

53088 78127

№2 (7.7 11.03. Решение задач №27 (ч.1) 11-12)

Фрагмент звёздного неба спроецирован на плоскость с декартовой системой координат. Учёный решил провести кластеризацию полученных точек, являющихся изображениями звёзд, то есть разбить их множество на N непересекающихся непустых подмножеств (кластеров), таких что точки каждого подмножества лежат внутри прямоугольника со сторонами длиной H и W , причём эти прямоугольники между собой не пересекаются. Стороны прямоугольников не обязательно параллельны координатным осям. Гарантируется, что такое разбиение существует и единственно для заданных размеров прямоугольников.

Точкой наведения назовём точку, в единичной окрестности от которой находится наибольшее количество точек кластера. Если таких точек несколько, то выбирается точка с наибольшей координатой x . Расстояние между двумя точками на плоскости $A(x_1, y_1)$ и $B(x_2, y_2)$ вычисляется по формуле: $d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.

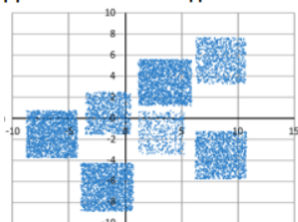
В файле А хранятся данные о звёздах четырёх кластеров, где $H=3$, $W=3$ для каждого кластера. В каждой строке записана информация о расположении на карте одной звезды: сначала координата x , затем координата y . Значения даны в условных единицах. Известно, что количество звёзд не превышает 1000.

В файле В хранятся данные о звёздах семи кластеров, где $H=5$, $W=5$ для каждого кластера. Известно, что количество звёзд не превышает 10 000. Структура хранения информации о звездах в файле В аналогична файлу А.

Для каждого файла определите координаты точки наведения каждого кластера, затем вычислите два числа: P_x – среднее арифметическое абсцисс точек наведения кластеров, и P_y – среднее арифметическое ординат точек наведения кластеров. В ответе запишите четыре числа: в первой строке сначала целую часть произведения $|P_x| \times 100000$, затем целую часть произведения $|P_y| \times 100000$ для файла А, во второй строке – аналогичные данные для файла Б.

Внимание! График приведён в иллюстративных целях для произвольных значений, не имеющих отношения к заданию.

Для выполнения задания используйте данные из прилагаемого файла.



ОТВЕТ:

171553 32527

157853 13516

№3 (7.7 11.03. Решение задач №27 (ч.1) 13-14)

Учёный решил провести кластеризацию некоторого множества звёзд по их расположению на карте звёздного неба. Кластер звёзд – это набор звёзд (точек) на графике. Каждая звезда обязательно принадлежит только одному из кластеров.

Истинный центр кластера, или центроид, – это одна из звёзд на графике, сумма расстояний от которой до всех остальных звёзд кластера минимальна.

Под расстоянием понимается расстояние Евклида между двумя точками $A(x_1, y_1)$ и $B(x_2, y_2)$ на плоскости, которое вычисляется по формуле: $d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

В файле А хранятся данные о звёздах двух кластеров. В каждой строке записана информация о расположении на карте одной звезды: сначала координата x , затем координата y (в условных единицах). Известно, что количество звёзд не превышает 1000. В файле Б хранятся данные о звёздах трёх кластеров. Известно, что количество звёзд не превышает 10 000. Структура хранения информации о звёздах в файле Б аналогична файлу А.

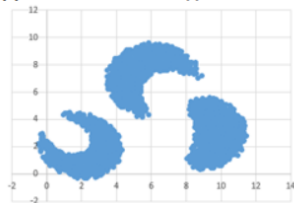
Для каждого файла определите координаты центра каждого кластера, затем вычислите два числа: P_x – среднее арифметическое абсцисс центров кластеров, и P_y – среднее арифметическое ординат центров кластеров.

В ответе запишите четыре числа: в первой строке сначала целую часть произведения $P_x \times 100000$, затем целую часть произведения $P_y \times 100000$ для файла А, во второй строке – аналогичные данные для файла Б.

Возможные данные одного из файлов иллюстрированы графиком.

Внимание! График приведён в иллюстративных целях для произвольных значений, не имеющих отношения к заданию.

Для выполнения задания используйте данные из прилагаемого файла.



ОТВЕТ:

398800 348577

596884 400991