

DOI: 10.32517/2221-1993-2025-24-4-65-74

**Е. М. Ганичева***Вологодский государственный университет, г. Вологда, Россия*

# ПРИМЕНЕНИЕ ЭЛЕКТРОННОЙ ТАБЛИЦЫ ДЛЯ РЕШЕНИЯ ЗАДАЧ АНАЛИЗА ДАННЫХ

## *Аннотация*

В статье показано, как с помощью инструментов табличного процессора можно решать задачи анализа данных. Согласно федеральной рабочей программе по учебному предмету «Информатика», необходимо сформировать у обучающихся представление о задачах анализа данных, таких как прогнозирование, классификация, кластеризация, анализ отклонений. На начальном этапе наиболее доступным программным средством для решения этих задач является табличный процессор. С использованием его инструментов можно решать задачи описательной статистики, визуализации данных, корреляционного и линейного регрессионного анализа данных. В статье приведен алгоритм решения задания 27 из контрольно-измерительных материалов единого государственного экзамена по информатике.

Материал статьи может быть полезен выпускникам при подготовке к экзамену по информатике, учителям информатики для организации работы по формированию умений использования электронной таблицы для решения задач анализа данных.

**Ключевые слова:** обучение информатике, анализ данных, задача кластеризации, табличный процессор, единый государственный экзамен.

## *Информация об авторе*

**Ганичева Елена Михайловна**, канд. пед. наук, доцент кафедры математики и информатики, Вологодский государственный университет, г. Вологда, Россия; *e-mail*: emg-ca@mail.ru

## 1. Введение

Основной целью изучения учебного предмета «Информатика» на уровне среднего общего образования является обеспечение развития информационных компетенций обучающегося, его готовности к жизни в условиях развивающегося информационного общества. В рамках курса информатики учащиеся имеют возможность получить представление о сквозных цифровых технологиях: дополненной и виртуальной реальности, робототехнике, технологии беспроводной связи, больших данных, машинном обучении [9, 18]. При этом весьма непростой задачей для учителя является выбор инструментов, с помощью которых можно продемонстрировать возможности той или иной технологии [1–3, 6, 10, 12]. Для решения задач линейной

регрессии, классификации, кластеризации можно использовать возможности электронной таблицы [13, 17].

Один из вариантов задачи анализа данных представлен в новой версии контрольно-измерительных материалов (задание 27) ЕГЭ 2025 года по информатике [4]. Задание 27 проверяет умение выполнять последовательность решения задач анализа данных: сбор первичных данных, очистка и оценка качества данных, выбор и построение модели, преобразование данных, интерпретация результатов. Выпускникам предлагается решить задачу для двух наборов данных (файл А и файл Б), отличающихся объемом: в файле А около 200 строк, в файле Б около 20 000 строк.

Для решения задачи можно использовать разные инструменты. В случае большого объема данных (файл Б, 20 000 строк) для решения задачи удобнее составить про-

**E. M. Ganicheva**

Vologda State University, Vologda, Russia

## USING A SPREADSHEET TO SOLVE DATA ANALYSIS PROBLEMS

### *Abstract*

The article demonstrates how spreadsheet tools can be used to solve data analysis problems. According to the Federal Curriculum for the subject informatics students must be familiar with data analysis tasks such as forecasting, classification, clustering, and variance analysis. At the initial stage, a spreadsheet is the most accessible software tool for solving these problems. Using its tools, one can solve problems involving descriptive statistics, data visualization, correlation, and linear regression analysis. The article provides an algorithm for solving Task 27 from the Unified State Exam in informatics.

The content of the article may be useful for graduates preparing for the informatics exam and for informatics teachers to organize work on developing the skills in using spreadsheets to solve data analysis problems.

**Keywords:** computer science training, data analysis, clustering task, table processor, unified state exam.

### *Information about the author*

**Elena M. Ganicheva**, Candidate of Sciences (Education), Associate Professor at the Department of Mathematics and Informatics, Vologda State University, Vologda, Russia; *e-mail*: emg-ca@mail.ru

грамму на языке программирования. Решение задачи в случае обработки данных меньшего объема вполне можно выполнить с использованием электронной таблицы.

## 2. Разбор заданий по анализу данных

### 2.1. Корреляционный анализ данных

В научных и прикладных исследованиях часто требуется установить связь между результативными и факторными переменными (урожайностью какой-либо культуры и количеством осадков, ростом и весом человека в однородных группах по полу и возрасту, частотой пульса и температурой тела и т. д.). В этом помогает корреляционный анализ. Это метод обработки статистических данных, заключающийся в изучении коэффициентов корреляции между переменными. Для того чтобы установить статистическую взаимосвязь между признаками, сравнивают коэффициенты корреляции между одной парой или множеством пар признаков.

#### Задача 1.

В школе города М проводилось исследование здоровья и благополучия подростков. Участие в данном исследовании было анонимным. Все полученные данные вносились в таблицу (рис. 1) для дальнейшего анализа.

Цели данного исследования:

- получить представление о состоянии здоровья подростков;
- определить факторы, вместе и по отдельности влияющие на здоровье мальчиков и девочек в раннем подростковом возрасте.

Задания.

1. Рассчитайте коэффициенты корреляции в ячейках H2:J2 с помощью встроенной функции КОРРЕЛ().

2. Для наглядности выполните визуализацию данных:

- постройте диаграммы разброса (*Вставка, Все диаграммы, Точечная*);

- выполните форматирование диаграмм (самостоятельно выберите стиль, палитру, макет).
3. Интерпретируйте полученный результат:
- выявлена ли корреляция;
  - если выявлена, то какая (прямая или обратная).
4. Сформулируйте ответ и запишите его над диаграммой.

Ответы представлены на рисунке 2.

### 2.2. Линейный регрессионный анализ данных

Корреляционный анализ позволяет выявить линейную связь между признаками и построить математическую модель линейной зависимости. В случае прямой или обратной корреляции точки на диаграмме образуют линию, похожую на прямую. Зная уравнение прямой и полученные параметры (коэффициент при  $x$  и свободный коэффициент), можно построить прямую, подставить интересующие нас  $x$  и  $y$  и сделать прогноз — предположить, что ждет нас в будущем.

Цель регрессионного анализа — предсказание ожидаемого среднего значения результирующей переменной с помощью уравнения регрессии. Например, анализ данных о продажах и закупках помогает выявить закономерности в покупках в определенные дни недели или время суток. Информация, полученная с помощью регрессионного анализа, помогает прогнозировать периоды времени, когда продукция компании будет пользоваться высоким спросом. Линейная регрессия важна и в спорте. Например, можно заинтересоваться вопросом о том, зависит ли количество матчей, выигранных баскетбольной командой за сезон, от среднего количества очков за матч.

Уравнение линейной регрессии имеет вид:

$$y = ax + b,$$

где:

$y$  — зависимая переменная;  
 $x$  — независимая переменная;  
 $a$  и  $b$  — коэффициенты регрессии.

|    | A             | B   | C      | D          | E          | F     | G | H                                        | I                                       | J                                        |
|----|---------------|-----|--------|------------|------------|-------|---|------------------------------------------|-----------------------------------------|------------------------------------------|
|    | id<br>ученика | пол | группа | рост,<br>м | вес,<br>кг | ИМТ   |   | коэффициент<br>корреляции<br>(рост, вес) | коэффициент<br>корреляции<br>(вес, ИМТ) | коэффициент<br>корреляции<br>(рост, ИМТ) |
| 1  |               |     |        |            |            |       |   |                                          |                                         |                                          |
| 2  | 1710_52       | муж | 3      | 1,38       | 68         | 35,71 |   |                                          |                                         |                                          |
| 3  | 1710_33       | муж | 3      | 1,39       | 34         | 17,60 |   |                                          |                                         |                                          |
| 4  | 1710_37       | жен | 2      | 1,39       | 72         | 37,27 |   |                                          |                                         |                                          |
| 5  | 1710_57       | жен | 2      | 1,39       | 52         | 26,91 |   |                                          |                                         |                                          |
| 6  | 1710_68       | жен | 3      | 1,39       | 48         | 24,84 |   |                                          |                                         |                                          |
| 7  | 1710_02       | муж | 1      | 1,39       | 50         | 25,83 |   |                                          |                                         |                                          |
| 8  | 1710_34       | муж | 1      | 1,40       | 26         | 13,27 |   |                                          |                                         |                                          |
| 9  | 1710_39       | жен | 2      | 1,40       | 25         | 12,76 |   |                                          |                                         |                                          |
| 10 | 1710_70       | муж | 2      | 1,40       | 26         | 13,27 |   |                                          |                                         |                                          |
| 11 | 1710_15       | жен | 1      | 1,41       | 65         | 32,76 |   |                                          |                                         |                                          |
| 12 | 1710_46       | жен | 3      | 1,41       | 31         | 15,59 |   |                                          |                                         |                                          |
| 13 | 1710_51       | муж | 1      | 1,41       | 37         | 18,61 |   |                                          |                                         |                                          |
| 14 | 1710_73       | жен | 1      | 1,41       | 30         | 15,09 |   |                                          |                                         |                                          |
| 15 | 1710_96       | жен | 3      | 1,41       | 70         | 35,21 |   |                                          |                                         |                                          |
| 16 | 1710_56       | муж | 2      | 1,42       | 55         | 27,28 |   |                                          |                                         |                                          |
| 17 | 1710_07       | жен | 2      | 1,43       | 52         | 25,53 |   |                                          |                                         |                                          |
| 18 | 1710_05       | муж | 3      | 1,43       | 43         | 21,05 |   |                                          |                                         |                                          |
| 19 | 1710_27       | муж | 2      | 1,43       | 72         | 35,21 |   |                                          |                                         |                                          |
| 20 | 1710_62       | жен | 2      | 1,43       | 48         | 23,47 |   |                                          |                                         |                                          |
| 21 | 1710_90       | муж | 3      | 1,43       | 79         | 38,63 |   |                                          |                                         |                                          |

Рис. 1. Данные для задачи 1

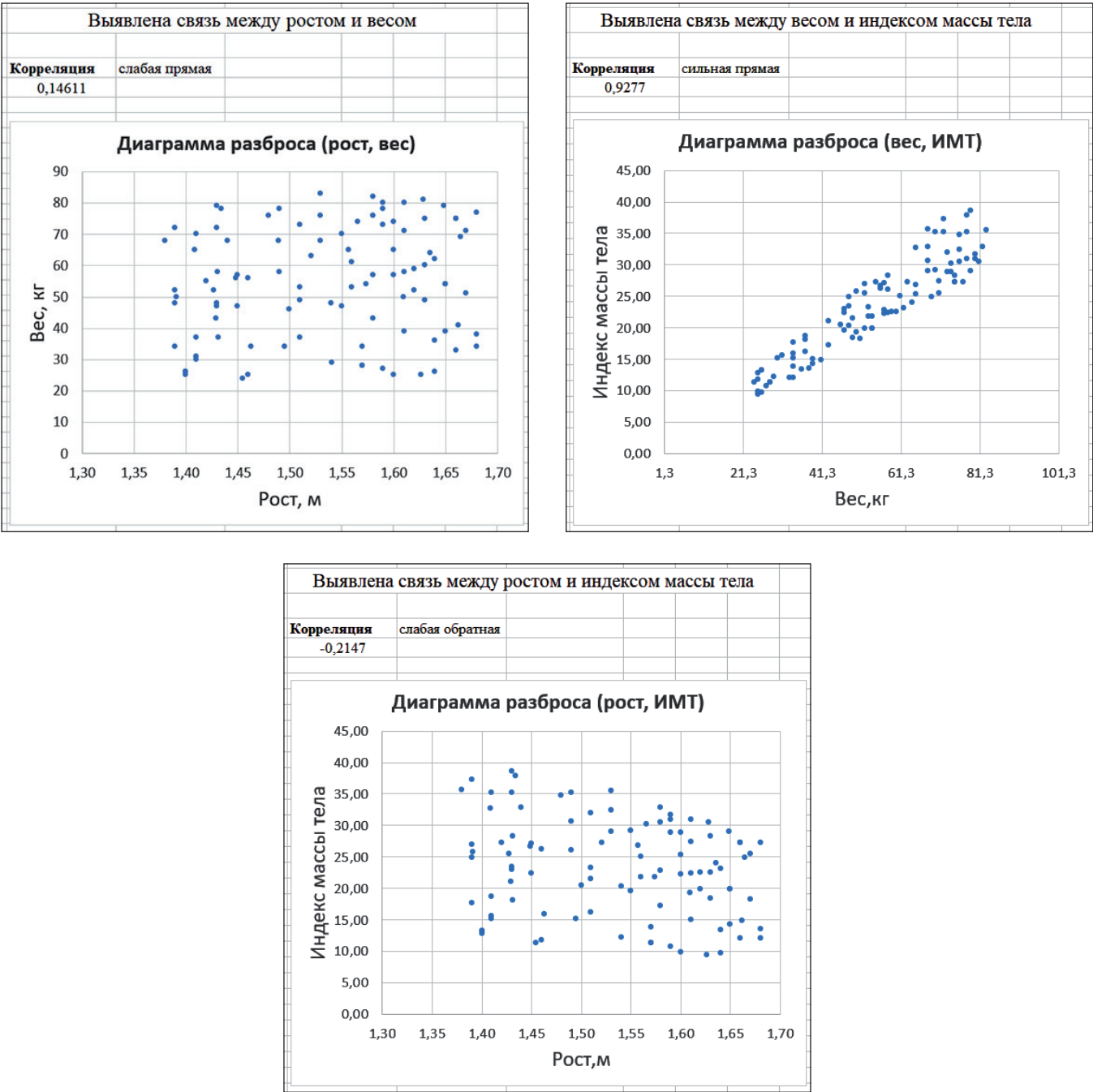


Рис. 2. Ответы для задачи 1

Коэффициент  $a$  характеризует наклон прямой, это самый важный коэффициент. Коэффициент  $b$  определяет точку пересечения с осью  $OY$  (ордината).

Для вычисления коэффициентов регрессии в электронной таблице имеются встроенные функции НАКЛОН и ОТРЕЗОК.

Функция НАКЛОН(известные\_значения\_y; известные\_значения\_x) возвращает наклон линии линейной регрессии для точек данных в аргументах. Наклон определяется как частное от деления расстояния по вертикали на расстояние по горизонтали между двумя любыми точками прямой. Иными словами, наклон — это скорость изменения значений вдоль прямой.

Функция ОТРЕЗОК(известные\_значения\_y; известные\_значения\_x) вычисляет точку пересечения

линии с осью  $Y$ , используя значения аргументов. Точка пересечения находится на оптимальной линии регрессии, проведенной через точки, заданные аргументами. Функция используется для определения значения зависимой переменной при нулевом значении независимой переменной.

**Задача 2.**

Аналитическое агентство «Автостат» собрало данные путешественников о движении по туристическим маршрутам. Данные внесены в таблицу (рис. 3).

Цели данного исследования:

- выявить связь пройденного маршрута с временем;
- построить линейную математическую модель зависимости.

|    | A      | B                                          | C                                        | D | E                         | F      | G       |
|----|--------|--------------------------------------------|------------------------------------------|---|---------------------------|--------|---------|
|    | запись | время (ч)<br>(независимая<br>переменная X) | путь (км)<br>(зависимая<br>переменная Y) |   | коэффициент<br>корреляции | наклон | отрезок |
| 1  |        |                                            |                                          |   |                           |        |         |
| 2  | 1      | 1                                          | 65                                       |   |                           |        |         |
| 3  | 2      | 19                                         | 1235                                     |   |                           |        |         |
| 4  | 3      | 19                                         | 1235                                     |   |                           |        |         |
| 5  | 4      | 12                                         | 780                                      |   |                           |        |         |
| 6  | 5      | 4                                          | 260                                      |   |                           |        |         |
| 7  | 6      | 1                                          | 65                                       |   |                           |        |         |
| 8  | 7      | 4                                          | 260                                      |   |                           |        |         |
| 9  | 8      | 13                                         | 845                                      |   |                           |        |         |
| 10 | 9      | 12                                         | 780                                      |   |                           |        |         |
| 11 | 10     | 12                                         | 780                                      |   |                           |        |         |
| 12 | 11     | 14                                         | 910                                      |   |                           |        |         |
| 13 | 12     | 19                                         | 1235                                     |   |                           |        |         |
| 14 | 13     | 15                                         | 975                                      |   |                           |        |         |
| 15 | 14     | 11                                         | 715                                      |   |                           |        |         |
| 16 | 15     | 20                                         | 1300                                     |   |                           |        |         |
| 17 | 16     | 3                                          | 195                                      |   |                           |        |         |
| 18 | 17     | 4                                          | 260                                      |   |                           |        |         |
| 19 | 18     | 19                                         | 1235                                     |   |                           |        |         |
| 20 | 19     | 16                                         | 1040                                     |   |                           |        |         |
| 21 | 20     | 20                                         | 1300                                     |   |                           |        |         |
| 22 | 21     | 8                                          | 520                                      |   |                           |        |         |
| 23 | 22     | 18                                         | 1170                                     |   |                           |        |         |
| 24 | 23     | 0                                          | 0                                        |   |                           |        |         |
| 25 | 24     | 12                                         | 780                                      |   |                           |        |         |
| 26 | 25     | 2                                          | 130                                      |   |                           |        |         |

Рис. 3. Данные для задачи 2

**Задания.**

1. Определите наличие зависимости с помощью корреляционного анализа:

- рассчитайте коэффициент корреляции в ячейке E2 с помощью встроенной функции КОРРЕЛ();
- охарактеризуйте полученную зависимость (прямая или обратная, сильная или слабая).

2. Рассчитайте коэффициенты линейной регрессии:

- в ячейке F2 (a) с помощью функции НАКЛОН();
- в ячейке G2 (b) с помощью функции ОТРЕЗОК().

3. Подумайте: как интерпретировать полученные показатели? Какой формулой можно описать нашу зависимость?

4. Постройте диаграмму разброса:

- постройте диаграммы разброса (*Вставка, Все диаграммы, Точечная*);
- выполните форматирование диаграммы (самостоятельно выберите стиль, палитру, макет);
- кликните левой кнопкой мыши на точки диаграммы;
- перейдите в контекстное меню и выберите команду *Добавить линию тренда*;
- в настройках линии тренда отметьте пункты *Показать уравнение* и *Поместить на диаграмме величину достоверности аппроксимации*.

Ответ представлен на рисунке 4.

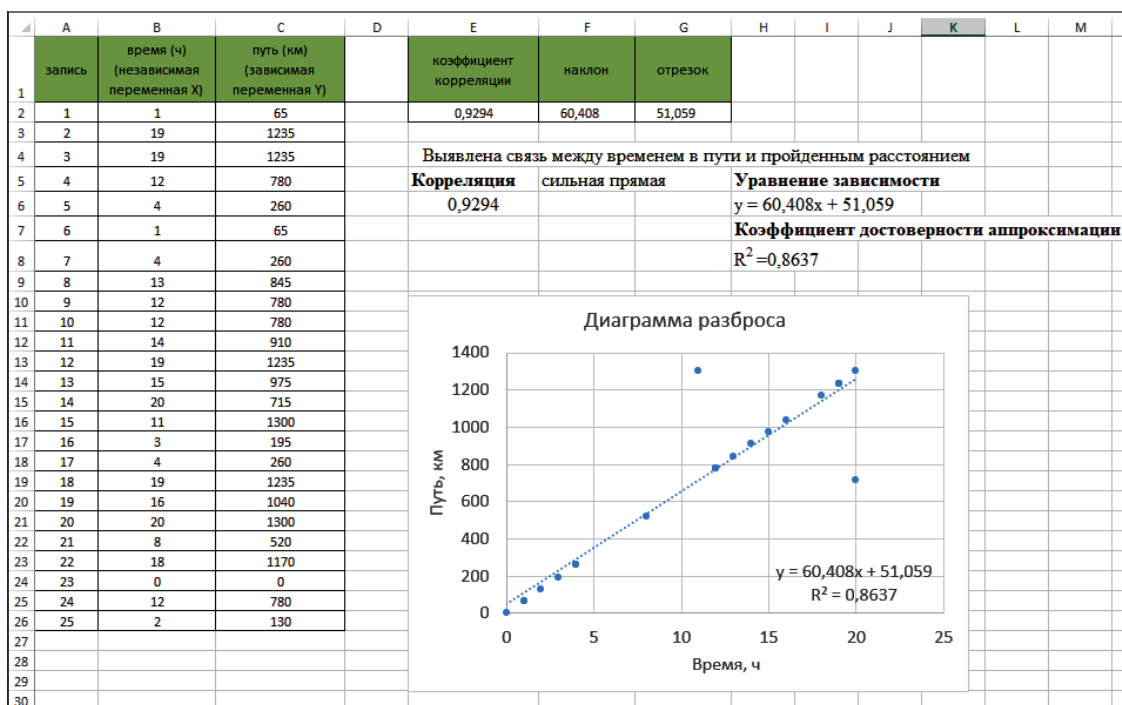


Рис. 4. Ответ для задачи 2

### 2.3. Задача кластеризации, решаемая в табличном процессоре

В задачах данного типа необходимо определить количество кластеров, найти центроид в каждом кластере и дать ответ на поставленный в задаче вопрос, выполнив действия с координатами кластеров.

Само задание формулируется следующим образом [4]:

Фрагмент звездного неба спроецирован на плоскость с декартовой системой координат. Ученый решил провести кластеризацию полученных точек, являющихся изображениями звезд, т. е. разбить их множество на  $N$  непересекающихся непустых подмножеств (кластеров), таких, что точки каждого подмножества лежат внутри квадрата со стороной длиной  $H$ , причем эти квадраты между собой не пересекаются. Стороны квадрата на обязательно параллельны координатным осям. Гарантируется, что такое разбиение существует и единственно для заданных размеров квадрата некоторого множества звезд по их расположению на карте звездного неба. Кластер звезд — это набор звезд (точек) на графике, лежащий внутри прямоугольника высотой  $H$  и шириной  $W$ . Каждая звезда обязательно принадлежит только одному из кластеров.

Истинный центр кластера, или центроид, — это одна из звезд на графике, сумма расстояний от которой до всех остальных звезд 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:

$$d(A, B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

*Входные данные.*

В файле А хранятся данные о звездах двух 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 звезд не превышает 1000.

В файле Б хранятся данные о звездах трех кластеров, где  $H = 3$ ,  $W = 3$  для каждого кластера. Известно, что количество звезд 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.

*Выходные данные.*

В ответе запишите четыре числа: в первой строке сначала целую часть произведения  $P_x \times 10\,000$ , затем целую часть произведения  $P_y \times 10\,000$  для файла А, во второй строке — аналогичные данные для файла Б.

Задача может быть решена на языке программирования как для задания А, так и для задания Б. Однако для многих обучающихся сложно понять суть задачи, поэтому полезно рассмотреть способ решения задачи с использованием данных из файла А в электронной таблице.

*Решение.*

*Идея решения:*

1. Выполним визуализацию данных, построив точечную диаграмму.

2. По внешнему виду диаграммы определим условия разделения кластеров.
3. Выполним сортировку данных согласно условию разделения так, чтобы в верхней части диапазона находились строки с данными, для которых выполнено условие разделения.
4. Для каждого из кластеров создадим новый лист электронной таблицы, перенесем на него строки с координатами точек, относящихся к этому кластеру, и подготовим таблицы для определения центроида.
5. Определим координаты центроидов всех кластеров.
6. Выполним действия с координатами.

*Описание алгоритма решения задачи.*

Исходные данные о координатах точек (звезд) представлены в файле электронной таблицы в виде двух столбцов данных. Для решения задачи надо определить количество кластеров. Это удобно сделать, построив точечную диаграмму.

*Способ для электронной таблицы Excel:* выделить данные (диапазон A1:B201) — *Вставка, Диаграмма, Точечная.*

*Способ для электронной таблицы Libre Calc:* выделить данные (диапазон A1:B201) — *Вставка, Диаграмма, X-Y разброс.*

В результате получим диаграмму, представленную на рисунке 5.

На рисунке 5 явно видны два кластера точек. Надо определить, как задать границу разделения этих кластеров. В нашем примере такой границей может стать прямая линия, заданная, например, уравнением  $y = 10$ . Назовем кластер, расположенный в нижней части диаграммы, первым, а кластер, расположенный выше, вторым. Разделение может быть и более сложным [11], например, можно использовать уравнение прямой в отрезках или наборы условий.

Для того чтобы найти координаты центроида, необходимо разделить массив значений координат точек на два множества: точки верхнего и точки нижнего кластеров. Это можно сделать, выполнив сортировку данных по второму столбцу (координата  $y$ ) по возрастанию (рис. 6).

После сортировки точки, координаты  $y$  которых имеют значения, меньшие 10, будут расположены в строках с 1-й по 74-ю. Заметим, что в 75-й строке координата  $y$  будет больше 10 (рис. 7):

Для вычисления расстояний между точками первого кластера добавим новый лист, переименуем его в *Кластер 1* и скопируем данные строк с 1-й по 74-ю на лист *Кластер 1*. Вставим значения в ячейки диапазона A3:B76.

Этот же массив данных вставим, начиная с ячейки C1, с транспонированием. (транспонирование — это изменение расположения элементов: столбец станет строкой) (рис. 8).

Как выполнить вставку с транспонированием элементов столбца?

*Способ для электронной таблицы Excel:* выделить данные — *Копировать* — установить курсор в ячейку C1 — *Вставка* — *с транспонированием.*

*Способ для электронной таблицы Libre Calc:* выделить данные — *Копировать* — установить курсор

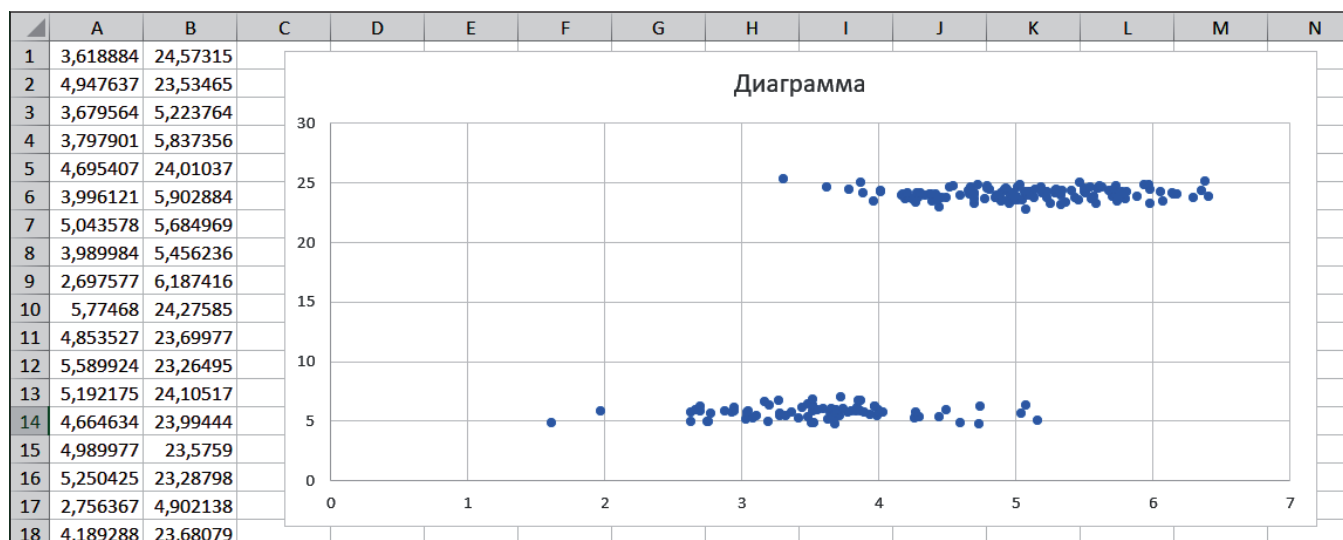


Рис. 5. Диаграмма точек кластеров

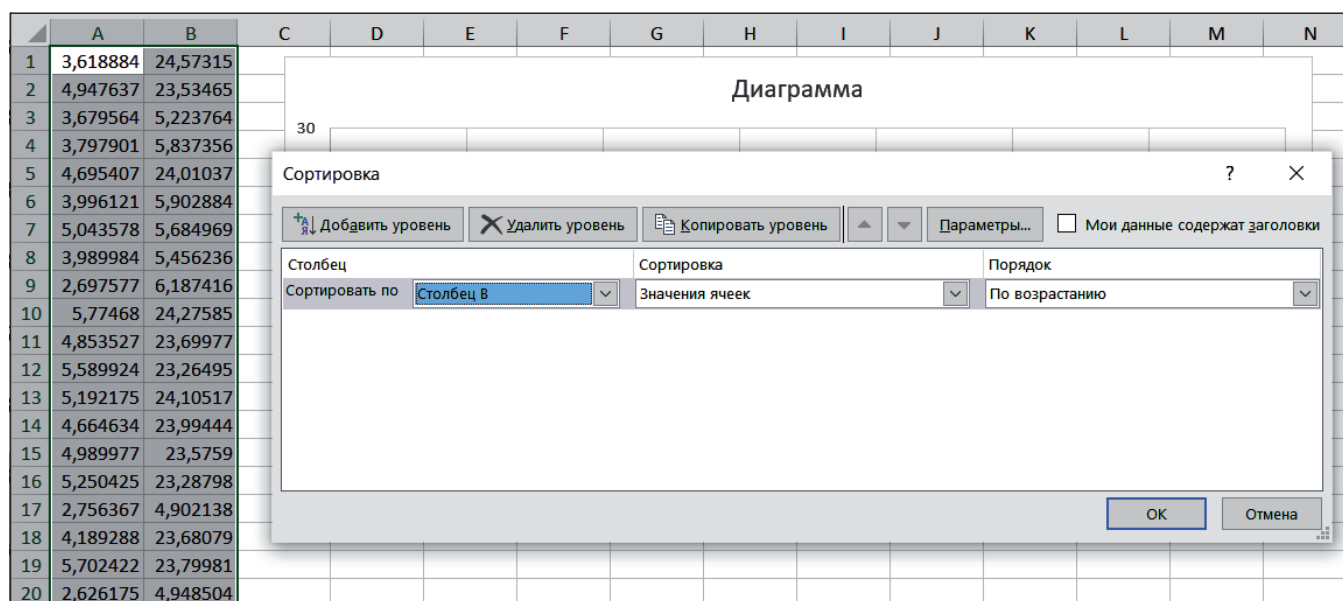


Рис. 6. Окно параметров сортировки

|    |          |          |
|----|----------|----------|
| 72 | 3,870719 | 6,702293 |
| 73 | 3,517272 | 6,84483  |
| 74 | 3,719871 | 6,984127 |
| 75 | 5,076741 | 22,79163 |
| 76 | 4,444314 | 22,97295 |
| 77 | 5,333727 | 23,19304 |

Рис. 7. Фрагмент результата сортировки

|   | A        | B        | C       | D        | E        | F        | G        |
|---|----------|----------|---------|----------|----------|----------|----------|
| 1 |          |          | 3,68409 | 4,731548 | 3,509011 | 3,525933 | 4,595204 |
| 2 |          |          | 4,70573 | 4,789812 | 4,841977 | 4,84847  | 4,876607 |
| 3 | 3,68409  | 4,70573  |         |          |          |          |          |
| 4 | 4,731548 | 4,789812 |         |          |          |          |          |
| 5 | 3,509011 | 4,841977 |         |          |          |          |          |
| 6 | 3,525933 | 4,84847  |         |          |          |          |          |
| 7 | 4,595204 | 4,876607 |         |          |          |          |          |
| 8 | 1,609716 | 4,895914 |         |          |          |          |          |

Рис. 8. Результат вставки данных на лист **Кластер 1**

в ячейку C1 — *Правка* — *Вставить как* — *транспонировать*. При пояснении этих действий можно привести аналогию, например, с таблицей умножения или таблицей квадратов чисел.

В ячейку C3 запишем формулу для вычисления расстояния между точками:

$$d(A, B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

При этом координаты точки  $(x_1, y_1)$  находятся в ячейках A3 и B3, координаты точки  $(x_2, y_2)$  находятся в ячейках C1 и C2 (рис. 9).

Для того чтобы формулу можно было скопировать на весь диапазон, зададим абсолютную адресацию. Не-

обходимо зафиксировать имена столбцов — A и B и номера строк — 1 и 2. Формула примет вид, как показано на рисунке 10.

Значение в ячейке C3 будет равно нулю, а в остальных ячейках первой строки получим расстояния от первой точки до всех 73 точек кластера. Во второй строке — расстояния от второй до всех остальных точек кластера, и т. д.

Удобно скопировать формулу из ячейки C3 во все ячейки столбца C3:C76, а затем продолжить копирование всего столбца по горизонтали.

При этом нули на диагонали помогут ориентироваться при определении диапазона копирования. Нуль должен оказаться в правом нижнем углу диапазона (рис. 11).

|    | A        | B        | C                            | D        | E        | F        | G        | H        | I        | J        | K        | L        | M        |
|----|----------|----------|------------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1  |          |          | 3,68409                      | 4,731548 | 3,509011 | 3,525933 | 4,595204 | 1,609716 | 2,756367 | 3,190141 | 2,626175 | 2,747839 | 5,159421 |
| 2  |          |          | 4,70573                      | 4,789812 | 4,841977 | 4,84847  | 4,876607 | 4,895914 | 4,902138 | 4,932063 | 4,948504 | 4,953834 | 5,034498 |
| 3  | 3,68409  | 4,70573  | =КОРЕНЬ((A3-C1)^2+(B3-C2)^2) |          |          |          |          |          |          |          |          |          |          |
| 4  | 4,731548 | 4,789812 |                              |          |          |          |          |          |          |          |          |          |          |
| 5  | 3,509011 | 4,841977 |                              |          |          |          |          |          |          |          |          |          |          |
| 6  | 3,525933 | 4,84847  |                              |          |          |          |          |          |          |          |          |          |          |
| 7  | 4,595204 | 4,876607 |                              |          |          |          |          |          |          |          |          |          |          |
| 8  | 1,609716 | 4,895914 |                              |          |          |          |          |          |          |          |          |          |          |
| 9  | 2,756367 | 4,902138 |                              |          |          |          |          |          |          |          |          |          |          |
| 10 | 3,190141 | 4,932063 |                              |          |          |          |          |          |          |          |          |          |          |
| 11 | 2,626175 | 4,948504 |                              |          |          |          |          |          |          |          |          |          |          |
| 12 | 2,747839 | 4,953834 |                              |          |          |          |          |          |          |          |          |          |          |
| 13 | 5,159421 | 5,034498 |                              |          |          |          |          |          |          |          |          |          |          |
| 14 | 3,033077 | 5,148418 |                              |          |          |          |          |          |          |          |          |          |          |
| 15 | 3,629783 | 5,180207 |                              |          |          |          |          |          |          |          |          |          |          |
| 16 | 3,679564 | 5,223764 |                              |          |          |          |          |          |          |          |          |          |          |
| 17 | 3,082168 | 5,242039 |                              |          |          |          |          |          |          |          |          |          |          |

Рис. 9. Окно аргументов функции

|   | A        | B        | C                                    | D        | E        | F        | G        |
|---|----------|----------|--------------------------------------|----------|----------|----------|----------|
| 1 |          |          | 3,68409                              | 4,731548 | 3,509011 | 3,525933 | 4,595204 |
| 2 |          |          | 4,70573                              | 4,789812 | 4,841977 | 4,84847  | 4,876607 |
| 3 | 3,68409  | 4,70573  | =КОРЕНЬ((\$A3-C\$1)^2+(\$B3-C\$2)^2) |          |          |          |          |
| 4 | 4,731548 | 4,789812 |                                      |          |          |          |          |

Рис. 10. Формула для вычисления расстояния

| BO       | BP       | BQ       | BR       | BS       | BT       | BU       | BV       | BW       | BX       |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 0,603617 | 0,354591 | 1,36317  | 0,470813 | 0,837519 | 0,668355 | 0,815875 | 0,688715 | 0,849323 | 0,95864  |
| 0,509433 | 0,272005 | 1,438614 | 0,393074 | 0,750311 | 0,648199 | 0,734701 | 0,669983 | 0,790123 | 0,919047 |
| 0,664852 | 0,372611 | 1,247628 | 0,468276 | 0,850049 | 0,567094 | 0,810245 | 0,585552 | 0,796884 | 0,875372 |
| 0,311937 | 0,602527 | 2,141142 | 0,619537 | 0,56361  | 1,071181 | 0,657868 | 1,093547 | 0,921914 | 1,159335 |
| 0,291186 | 0,181768 | 1,640258 | 0,289067 | 0,573244 | 0,685941 | 0,584002 | 0,709409 | 0,713479 | 0,892554 |
| 0,516518 | 0,828782 | 2,38145  | 0,822561 | 0,654572 | 1,260067 | 0,765678 | 1,281156 | 1,050758 | 1,296084 |
| 1,540194 | 1,224961 | 0,347336 | 1,270974 | 1,62516  | 1,001461 | 1,543684 | 0,996682 | 1,375571 | 1,277618 |
| 0,774128 | 0,460805 | 1,105239 | 0,524029 | 0,901079 | 0,471001 | 0,841068 | 0,484931 | 0,766693 | 0,797177 |
| 0,769751 | 0,454544 | 1,105509 | 0,510416 | 0,884088 | 0,437376 | 0,820844 | 0,451055 | 0,738389 | 0,764099 |
| 0        | 0,316126 | 1,873718 | 0,309932 | 0,343786 | 0,759531 | 0,402649 | 0,781799 | 0,632211 | 0,860496 |
| 0,316126 | 0        | 1,55761  | 0,124231 | 0,484341 | 0,510308 | 0,46275  | 0,533816 | 0,543679 | 0,712038 |
| 1,873718 | 1,55761  | 0        | 1,594118 | 1,933739 | 1,278209 | 1,844897 | 1,269195 | 1,649683 | 1,516793 |
| 0,309932 | 0,124231 | 1,594118 | 0        | 0,382832 | 0,453496 | 0,345908 | 0,476323 | 0,424847 | 0,60956  |
| 0,343786 | 0,484341 | 1,933739 | 0,382832 | 0        | 0,685569 | 0,114641 | 0,702132 | 0,402453 | 0,647992 |
| 0,759531 | 0,510308 | 1,278209 | 0,453496 | 0,685569 | 0        | 0,583066 | 0,023511 | 0,374376 | 0,329027 |
| 0,402649 | 0,46275  | 1,844897 | 0,345908 | 0,114641 | 0,583066 | 0        | 0,598277 | 0,287976 | 0,533693 |
| 0,781799 | 0,533816 | 1,269195 | 0,476323 | 0,702132 | 0,023511 | 0,598277 | 0        | 0,381106 | 0,319665 |
| 0,632211 | 0,543679 | 1,649683 | 0,424847 | 0,402453 | 0,374376 | 0,287976 | 0,381106 | 0        | 0,245866 |
| 0,860496 | 0,712038 | 1,516793 | 0,60956  | 0,647992 | 0,329027 | 0,533693 | 0,319665 | 0,245866 | 0        |

Рис. 11. Результат копирования формулы для вычисления расстояния

Следующим шагом станет вычисление суммы расстояний по строке (рис. 12).

Далее находим минимальную сумму расстояний от одной точки до всех остальных точек кластера (рис. 13).

Поиск нужного значения в столбце ВУ можно выполнить разными способами: визуально, применить средства поиска или условное форматирование (рис. 14).

Правило выделения ячеек задаем равным полученному минимальному значению суммы расстояний (рис. 15).

| BX       | BY                          | BZ          |
|----------|-----------------------------|-------------|
| 3,719871 | Сумма                       | Минимальная |
| 6,984127 | расстояний                  | сумма       |
| 2,278679 | =СУММ(СЗ:BX3)               |             |
| 2,416301 | СУММ(число1; [число2]; ...) |             |
| 2,152503 |                             |             |

Рис. 12. Формула для вычисления суммы расстояний

| BX       | BY         | BZ                         | CA |
|----------|------------|----------------------------|----|
| 3,719871 | Сумма      | Минимальная                |    |
| 6,984127 | расстояний | сумма                      |    |
| 2,278679 | 91,2853144 | =МИН(BY3:BY76)             |    |
| 2,416301 | 120,106613 | МИН(число1; [число2]; ...) |    |
| 2,152503 | 82,6373214 |                            |    |

Рис. 13. Формула для определения минимальной суммы расстояний

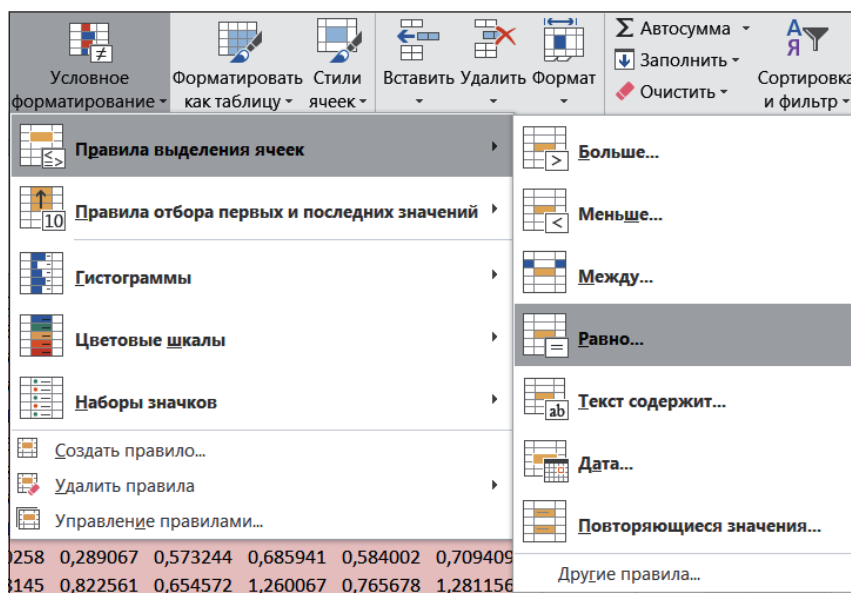


Рис. 14. Панель Условное форматирование

| BJ       | BK       | BL       | BM       | BN       | BO       | BP       | BQ         | BR       | BS       | BT       | BU       | BV       | BW       | BX       | BY         | BZ          |
|----------|----------|----------|----------|----------|----------|----------|------------|----------|----------|----------|----------|----------|----------|----------|------------|-------------|
| 3,444692 | 2,697577 | 4,740487 | 3,973541 | 3,971426 | 3,202424 | 3,518517 | 5,076123   | 3,486241 | 3,171447 | 3,855513 | 3,272513 | 3,870719 | 3,517272 | 3,719871 | Сумма      | Минимальная |
| 6,135052 | 6,187416 | 6,21559  | 6,228388 | 6,262624 | 6,296594 | 6,301153 | 6,304981   | 6,421117 | 6,638982 | 6,684361 | 6,693096 | 6,702293 | 6,84483  | 6,984127 | расстояний | сумма       |
| 1,449232 | 1,780057 | 1,84273  | 1,549926 | 1,583187 | 1,662183 | 1,603992 | 2,120227   | 1,72676  | 2,000067 | 1,986043 | 2,029537 | 2,005267 | 2,145595 | 2,278679 | 91,2853144 | 54,48790025 |
| 1,861631 | 2,46786  | 1,425806 | 1,626061 | 1,657396 | 2,146768 | 1,937935 | 1,553856   | 2,052303 | 2,419369 | 2,087283 | 2,398181 | 2,097286 | 2,386957 | 2,416301 | 120,106613 |             |
| 1,294674 | 1,571188 | 1,844816 | 1,462164 | 1,49401  | 1,486576 | 1,459207 | 2,14388    | 1,579305 | 1,828436 | 1,874684 | 1,866166 | 1,895154 | 2,00287  | 2,152503 | 82,6373214 |             |
| 1,289144 | 1,574468 | 1,828704 | 1,450698 | 1,482665 | 1,48382  | 1,452701 | 2,12709    | 1,573148 | 1,825265 | 1,865239 | 1,861952 | 1,885613 | 1,996378 | 2,144445 | 82,2764631 |             |
| 1,705099 | 2,306341 | 1,346842 | 1,487876 | 1,519915 | 1,98902  | 1,785661 | 1,507162   | 1,901397 | 2,265623 | 1,953232 | 2,24703  | 1,964181 | 2,244067 | 2,282071 | 109,766382 |             |
| 2,214181 | 1,688615 | 3,397539 | 2,713514 | 2,728657 | 2,120997 | 2,370279 | 3,741851   | 2,41818  | 2,340362 | 2,870914 | 2,44842  | 2,893984 | 2,727095 | 2,968736 | 160,747171 |             |
| 1,412044 | 1,286622 | 2,379472 | 1,800125 | 1,824086 | 1,464061 | 1,593146 | 2,710947   | 1,685234 | 1,785755 | 2,093905 | 1,86385  | 2,117153 | 2,086391 | 2,294127 | 96,0127191 |             |
| 1,518209 | 1,707022 | 1,874371 | 1,762959 | 1,89655  | 1,940539 | 2,119335 | 81,2863891 | 1,705375 | 1,776243 | 2,127081 | 1,860472 | 2,150504 | 2,095258 | 2,31083  | 99,8205752 |             |
| 1,642606 | 1,737575 | 2,054668 | 1,816677 | 2,077972 | 2,041541 | 2,250985 | 94,0659322 | 2,173073 | 2,554684 | 2,102908 | 2,512244 | 2,107675 | 2,444168 | 2,423502 | 136,082054 |             |
| 1,35097  | 1,496973 | 1,742274 | 1,563125 | 1,765267 | 1,764159 | 1,959978 | 75,5231668 | 1,249184 | 1,529083 | 1,520997 | 1,554501 | 1,541037 | 1,66842  | 1,806168 | 66,8220603 |             |
| 1,21286  | 1,503671 | 1,471157 | 1,524673 | 1,490835 | 1,62917  | 1,760825 | 65,5847009 |          |          |          |          |          |          |          |            |             |

Рис. 15. Окно аргументов условного форматирования

|    |          |          |          |          |
|----|----------|----------|----------|----------|
| 34 | 2,626477 | 5,692607 | 1,446539 | 2,290495 |
| 35 | 3,367221 | 5,710764 | 1,053803 | 1,646069 |
| 36 | 3,776316 | 5,733491 | 1,031891 | 1,342758 |
| 37 | 3,648817 | 5,733546 | 1,028422 | 1,436294 |
| 38 | 3,897581 | 5,751782 | 1,067617 | 1,273141 |
| 39 | 4,268581 | 5,759235 | 1,204784 | 1,0743   |
| 40 | 4,028229 | 5,764213 | 1,113022 | 1,201713 |

Рис. 16. Координаты центроида первого кластера

|   | A                            | B        | C        |
|---|------------------------------|----------|----------|
| 1 |                              |          |          |
| 2 |                              |          |          |
| 3 | Координаты первого центроида | 3,648817 | 5,733546 |
| 4 | Координаты второго центроида | 5,069665 | 24,05706 |
| 5 |                              | 43592,41 | 148953   |
| 6 |                              |          |          |
| 7 | Ответ:                       | 43592    | 148953   |

Рис. 17. Результат решения задачи 3

После нахождения нужного значения в столбце ВУ по строке определим координаты центроида. В нашем случае координаты центроида находятся в 37-й строке (рис. 16).

Добавим новый лист *Результат* и вставим найденные значения координат. Координаты второго центроида найдем по такому же алгоритму и вставим найденные значения на лист *Результат*. Вычислим  $P_x$  — среднее арифметическое абсцисс центров кластеров и  $P_y$  — среднее арифметическое ординат центров кластеров, умножим полученные значения на 10 000 и запишем ответ (рис. 17).

Электронную таблицу можно применить и для визуализации данных, представленных в файле Б, что является необходимым для определения количества кластеров и условий принадлежности точек к конкретному кластеру.

### 3. Заключение

Подготовка к решению задач анализа данных позволяет сформировать навыки работы с инструментами электронной таблицы, такие как умение построить диаграмму, выполнить сортировку данных, знание разных способов вставки данных, умение записать формулу, применить встроенные функции, условное форматирование [5, 7, 10, 11, 15]. Уверенное владение инструментами электронной таблицы позволит выпускнику успешно применять их как для решения задач в практической деятельности, так и при выполнении заданий итоговой государственной аттестации [14, 16].

#### Список источников

1. Богатырева Ю. И., Гончаров К. Г., Родионова О. В. Направления решения проблемы недостаточной подготовки школьников к ЕГЭ по информатике в Тульской области // Педагогическая информатика. 2023. № 4. С. 38–48. EDN: HONUDZ.
2. Везилов Т. Т. Средства электронного обучения в подготовке к ЕГЭ по информатике и ИКТ // Педагогический журнал. 2018. Т. 8. № 6А. С. 225–232. EDN: ZAHILB.

3. Ганичева Е. М., Голубев О. Б. Работа над задачами по теме «Использование динамических (электронных) таблиц для выполнения учебных заданий» при подготовке обучающихся к ГИА по информатике: учебное пособие для подготовки к государственной итоговой аттестации выпускников основной и старшей школы. Вологда: Вологодский институт развития образования, 2022. 76 с. EDN: WHAOJQ.

4. Демоверсии, спецификации, кодификаторы ЕГЭ 2025 год. <https://fipi.ru/ege/demoversii-spezifikacii-kodifikatory>

5. Зубрилин А. А. Возможности табличных процессоров при подготовке к ЕГЭ по информатике // Информатика в школе. 2024. Т. 23. № 6. С. 78–86. EDN: GOOTWX. DOI: 10.32517/2221-1993-2024-23-6-78-86.

6. Зубрилин А. А., Хохлова А. Д. Возможности табличного процессора MS Excel при решении задач ЕГЭ по информатике (на примере задачи № 5) // Инновационное образовательное пространство педагогического вуза: проблемы и перспективы. Сборник статей по материалам Всероссийской научно-практической конференции (Саранск, 23 ноября 2023 года). Саранск: Мордовский государственный педагогический университет им. М. Е. Евсевьева, 2023. С. 80–86. EDN: ANLAKS.

7. Ильина Е. А., Засов В. О., Масальский Л. С., Логунова Т. В., Шишиморов А. П. О генерации визуальных материалов к ЕГЭ по информатике // Педагогическая информатика. 2023. № 2. С. 25–28. EDN: KMEJPH.

8. Кабанов А. М. Задание 27 | КЕГЭ по информатике 2025. [https://vkvideo.ru/video-205865487\\_456240432](https://vkvideo.ru/video-205865487_456240432).

9. Касторнова В. А. К вопросу о внедрении технологий искусственного интеллекта в школьное образование // Педагогическая информатика. 2022. № 1. С. 18–29. EDN: NDWHAN.

10. Морозовская Т. Д. Использование табличного процессора для подготовки и сдачи ЕГЭ по информатике и информационным технологиям // Инновационные технологии в образовательной деятельности. Материалы Всероссийской научно-методической конференции (Нижний Новгород, 02 февраля 2021 года). Н. Новгород: Нижегородский государственный технический университет им. Р. Е. Алексеева, 2021. С. 144–149. EDN: ALXHXG.

11. Поляков К. Ю. Кластеризация в задачах ЕГЭ по информатике // Информатика в школе. 2025. Т. 24. № 1. С. 69–76. EDN: EGOGQD. DOI: 10.32517/2221-1993-2025-24-1-69-76.

12. Самохвалова Е. А. Применение чат-бота с элементами искусственного интеллекта для изучения темы «Работа с табличными процессорами» // Информатика в школе. 2024. Т. 23. № 1. С. 60–69. EDN: AKHOMJ. DOI: 10.32517/2221-1993-2024-23-1-60-69.

13. Самылкина Н. Н., Салахова А. А. Обучение основам искусственного интеллекта и анализа данных в курсе информатики на уровне среднего общего образования. М.: Московский педагогический государственный университет, 2022. 242 с. EDN: BASMCW. DOI: 10.31862/9785426310643.

14. Сдам ГИА: Решу ЕГЭ. Образовательный портал для подготовки к экзаменам. Информатика. <https://inf-ege.sdamgia.ru>

15. Табачук Н. П., Бусарова У. А. Образовательный контент ЕГЭ по информатике в картах визуализации // Педагогическая информатика. 2023. № 2. С. 63–71. EDN: RFORQK.

16. Фирсова С. А. Особенности решения некоторых задач компьютерного ЕГЭ по информатике // Вестник Таганрогского института имени А. П. Чехова. 2021. № 2. С. 47–51. EDN: RMJWOD.

17. Хун Чжоу. Машинное обучение сквозь призму Excel: примеры и упражнения: пер. с англ. А. Ю. Гинько. М.: ДМК Пресс, 2025. 270 с.

18. Чикова О. А., Каменев Р. В., Витюнин М. А., Сартанов И. В. Методика решения прикладных задач на уроках информатики в профильных информационно-технологических классах // Педагогическая информатика. 2023. № 4. С. 58–75. EDN: EOEZEC.