

Практическая работа

Анализ датасета. Функции. Строковые данные

Задание 1 Создание своих функций и их применение

1. Определение функции. Аргументы функции.
2. Вызов функции.
3. Функция как объект**.
4. Функция `map()`**.
5. Лямбда-функции**.
6. Функция `filter()`**

Задание 2. Операции над строками

1. Строка - итерируемый объект. Индексация элементов строки.
2. Поиск подстроки в строке. Срез.
3. Методом `find()`.
4. Метод `count()`.
5. Методы `lower()` и `upper()`
6. Метод `replace()`.
7. Регулярные выражения. Модуль `re`**

Задание 3. Анализ данных текстового файла.

1. Загрузить файл `StudentsPerformance.csv`
2. Преобразование данных файла в список.
3. Преобразование данных файла в словарь.
4. Общий алгоритм анализа данных.
5. Написать отчет.
6. Выполнить задание.

Примечание. Информация в файле `StudentsPerformance.csv` содержит категории:

- *gender* — пол абитуриента;
- *race/ethnicity* — национальность (зашифрована в виде значений "Group A", "Group B" и т.д.);
- *parental level of education* — уровень образования родителей;
- *lunch* — насколько качественно пообедал абитуриент перед тестом;
- *test preparation course* — закончил ли абитуриент подготовительные курсы;
- *math score* — оценка по математике;
- *reading score* — оценка по чтению;
- *writing score* — оценка по письму.

Варианты заданий

1. Выполните задание на основе набора данных **StudentsPerformance.csv**:
У скольких девочек родители имеют степень магистра (parental level of education = master's degree)?

2. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Какой средний балл на экзамене по чтению набрали ученики, набравшие максимальный балл на экзамене по математике? Ответ округлите до трёх цифр после запятой.

3. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько процентов абитуриентов, получивших на экзамене по письму (writing score) оценку более 90, хорошо пообедали перед экзаменом (lunch = standard)? Ответ округлите до одного знака после запятой.

4. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько абитуриентов, относящихся к этнической группе С, закончили подготовительные курсы?

5. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько абитуриентов получили на экзамене по письму (writing score) оценку выше 90?

6. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько девочек, родители которых имеют степень магистра, набрали по математике больше 90 баллов?

7. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Какой средний балл экзамене по чтению (reading score) получили девочки? Ответ округлите до трёх знаков после запятой.

8. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько мальчиков закончили подготовительные курсы (test preparation course = completed)?

9. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Какой средний балл, полученный мальчиками на экзамене по чтению? Ответ округлите до трёх цифр после запятой.

10. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько мальчиков хорошо пообедали перед экзаменом (lunch = standard)?

-
11. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько мальчиков хорошо пообедали перед экзаменом (lunch = standard)?
-
12. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько абитуриентов получили на экзамене по чтению (reading score) оценку ниже среднего?
-
13. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько разных вариантов значений встречается в столбце "parental level of education"? Введите ответ в виде целого числа
-
14. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько процентов абитуриентов полноценно пообедали перед экзаменом? Иными словами, у скольких процентов абитуриентов значение столбца "lunch" = "standard"? Ответ округлите до одной цифры после запятой.
-
15. Выполните задание на основе набора данных **StudentsPerformance.csv**:
Сколько разных этнических групп встречается в файле?
-