|  |
| --- |
|  |
| МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования  **"МИРЭА - Российский технологический университет"**  **РТУ МИРЭА** |
| Институт информационных технологий (ИТ) |
| Кафедра прикладной математики (ПМ) |

|  |  |
| --- | --- |
| **ОТЧЕТ**  **ПО ПРАКТИЧЕСКОЙ РАБОТЕ №2** | |
| **по дисциплине** |  |
| **«Технологии и инструментарий анализа больших данных»** | |
| Выполнил студент группы ИКБО-20-19 | Московка А.А. |
| Принял семинарист кафедры ИиППО | Десятников А.А. |

|  |  |  |
| --- | --- | --- |
| Практическая работы выполнены | « » 2022г. |  |
| «Зачтено» | « » 2022 г. |  |

Москва 2022

**Постановка задачи**

1. Напишите программу, которая считывает с консоли числа (по одному в строке) до тех пор, пока сумма введённых чисел не будет равна 0 и после этого выводит сумму квадратов всех считанных чисел.
2. Напишите программу, которая выводит последовательность чисел, длинною N, где каждое число повторяется столько раз, чему оно равно.

На вход программе передаётся неотрицательное целое число N. Например, если N = 7, то программа должна вывести 1 2 2 3 3 3 4

Вывод элементов списка через пробел – print(\*list).

1. Матрицу произвольного размера вытянуть в один вектор, не применяя встроенные методы Python.

Для создания матрицы можно использовать np.random.rand(кол-во строк, кол-во столбцов).

1. Даны два списка:

А = [1, 2, 3, 4, 2, 1, 3, 4, 5, 6, 5, 4, 3, 2]

В = [‘a’, ’b’, ’c’, ’c’, ’c’, ’b’, ’a’, ’c’, ’a’, ’a’, ’b’, ’c’, ’b’, ’a’]

Создать словарь, в котором ключи – это содержимое списка В, а значения для ключей словаря – это сумма всех элементов списка А в соответствии с буквой, содержащийся на той же позиции в списке В.

Пример Скриншот результатаа программы: {‘a’ : 10, ‘b’ : 15, ‘c’ : 6}.

1. Скачать и загрузить данные о стоимости домов в калифорнии, используя библиотеку sklearn.
2. Используя метод pd.concat([датафрейм1, датафрейм2], axis = 1), добавить к данным столбец, содержащий информацию о медианной стоимости дома (.target и .target\_names).
3. Использовать метод info().
4. Узнать, есть ли пропущенные значения, используя isna().sum().
5. Вывести записи, где средний возраст домов в районе более 50 лет и население более 2500 человек, используя метод loc().
6. Узнать максимальное и минимальное значения медианной стоимости дома (max(), min()).
7. Используя метод apply(), вывести на экран название признака и его среднее значение.
8. Составить отчет о проделанной работе. В отчете должен быть представлен код и результаты его выполнения с выводами.

**Программный код с комментариями и результаты его выполнения**

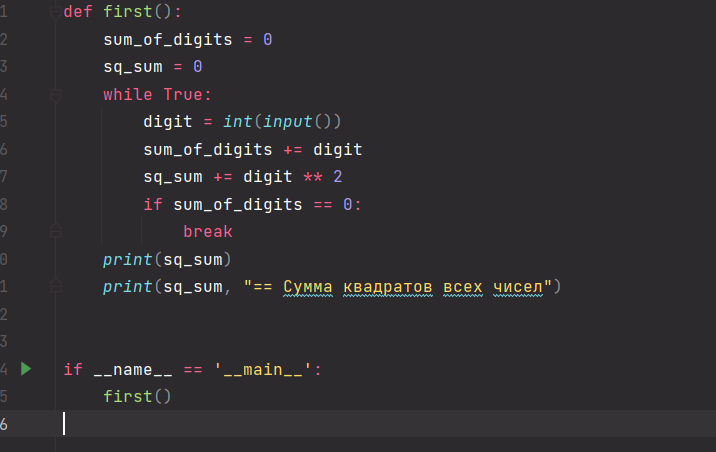
****

Рисунок 1 – Скриншот программного кода первой программы

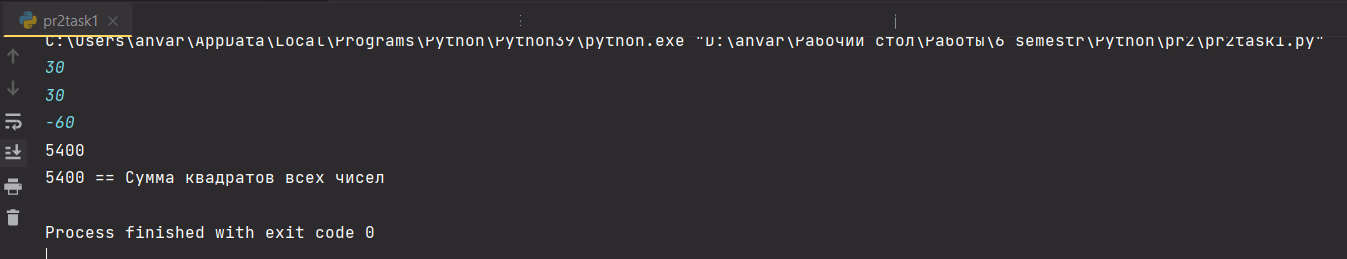
****

Рисунок 2 – Скриншот результата выполнения программы

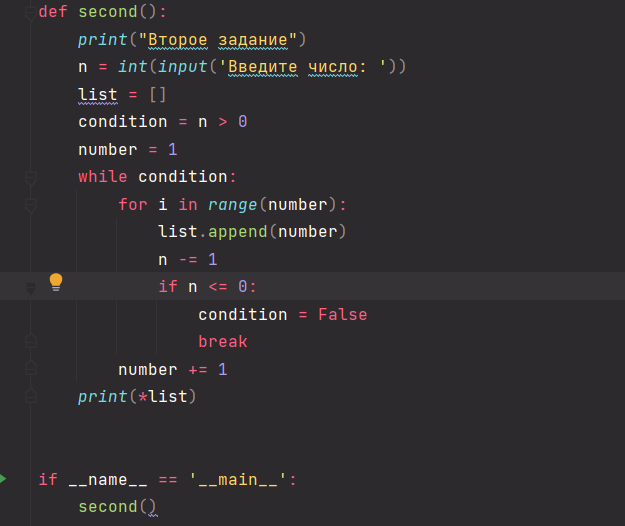
****

Рисунок 3 – Скриншот программного кода второй программы

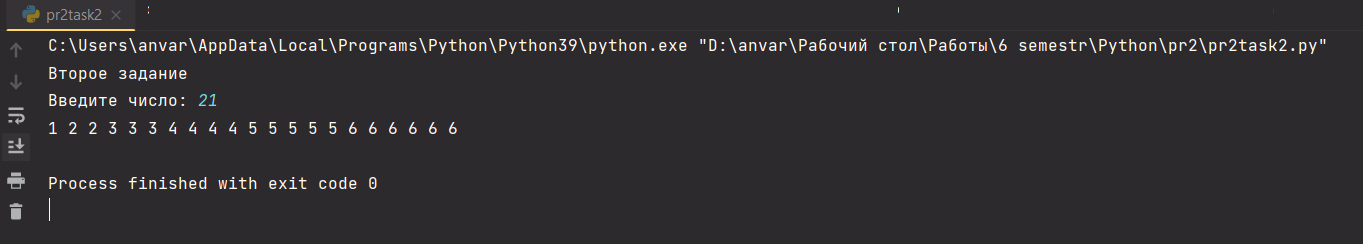
****

Рисунок 4 – Скриншот результата выполнения программы

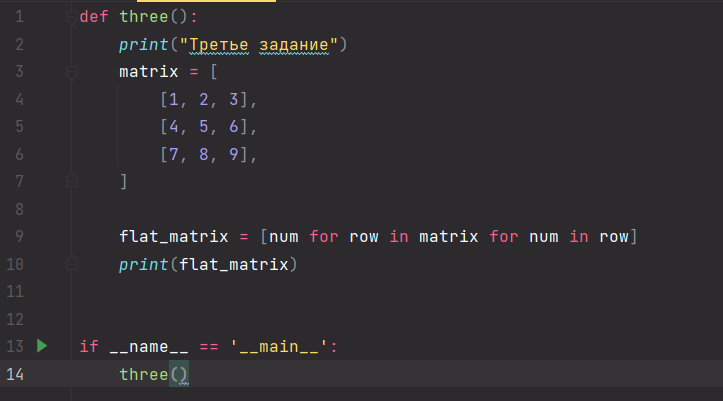
****

Рисунок 5 – Скриншот программного кода третьей программы

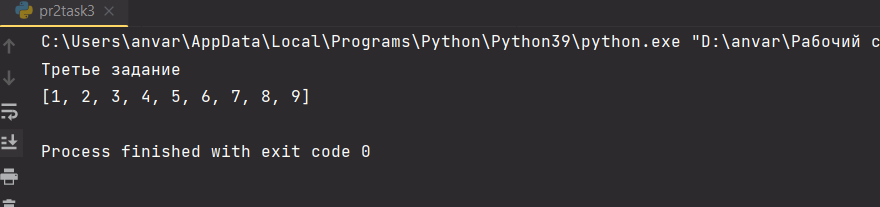
****

Рисунок 6 – Скриншот результата выполнения программы

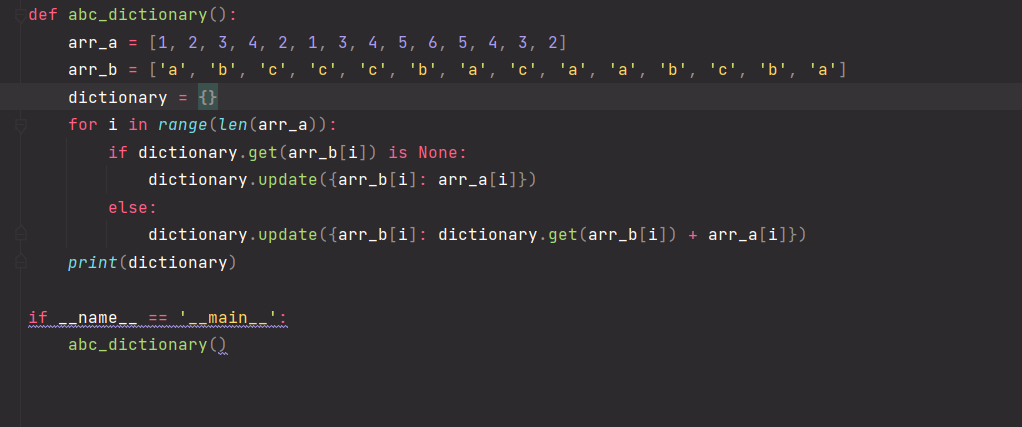


Рисунок 7 – Скриншот программного кода четвертой программы

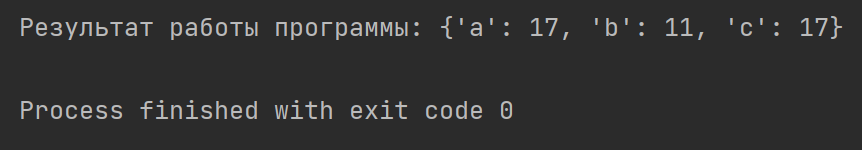


Рисунок 8 – Скриншот результата выполнения программы

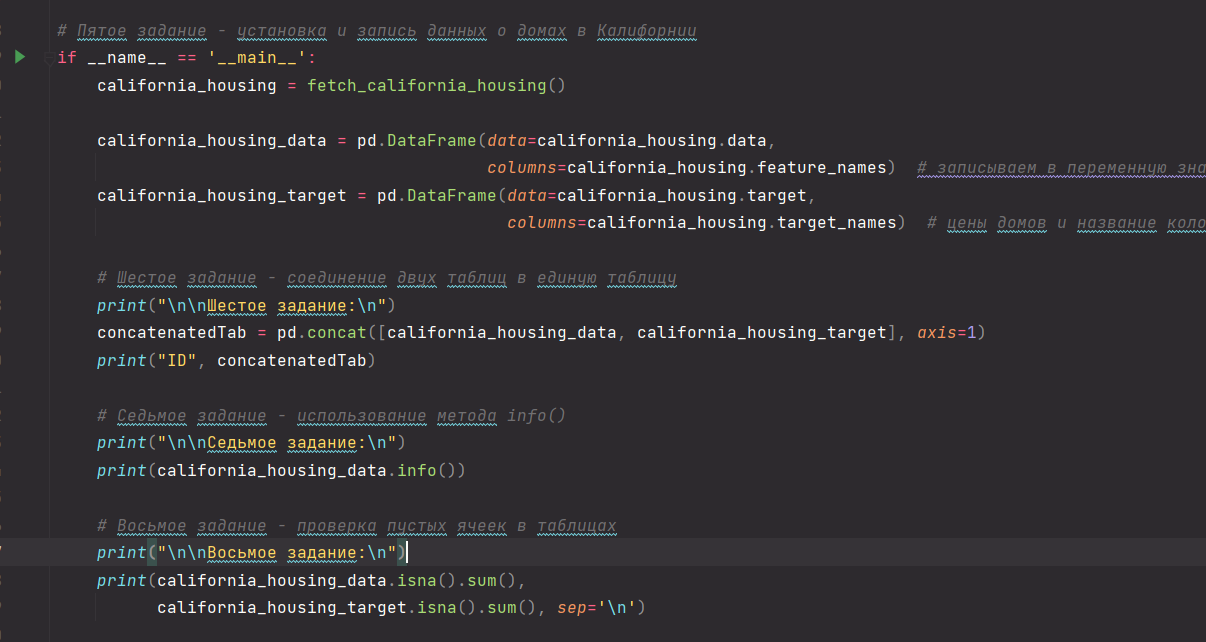


Рисунок 9 – Скриншот части программного кода пятой программы (Часть 1)

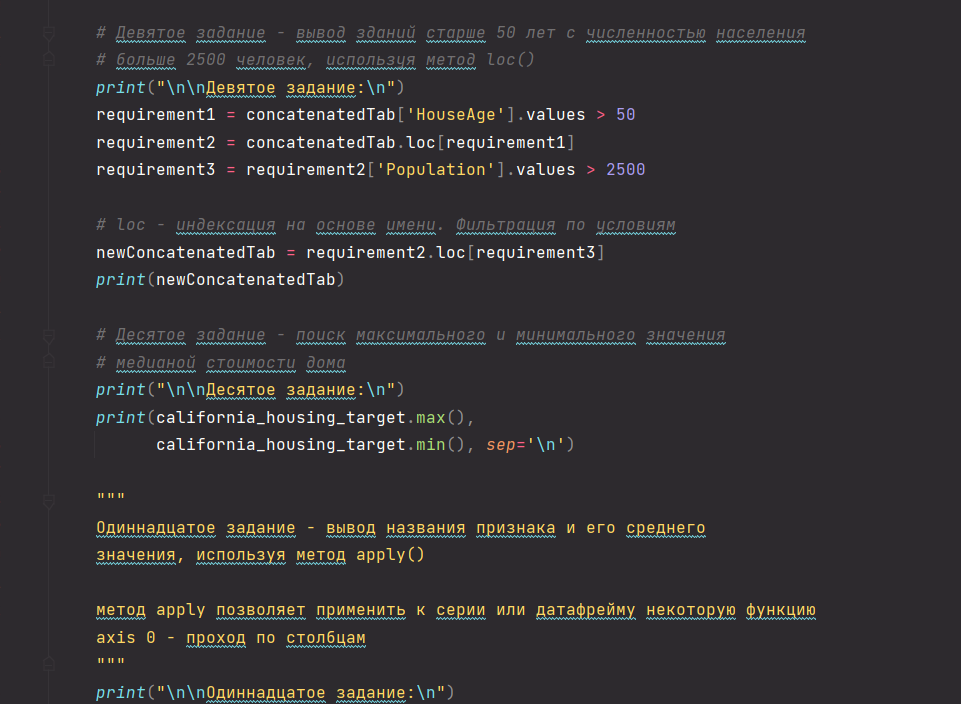


Рисунок 10 – Скриншот части программного кода пятой программы (Часть 2)

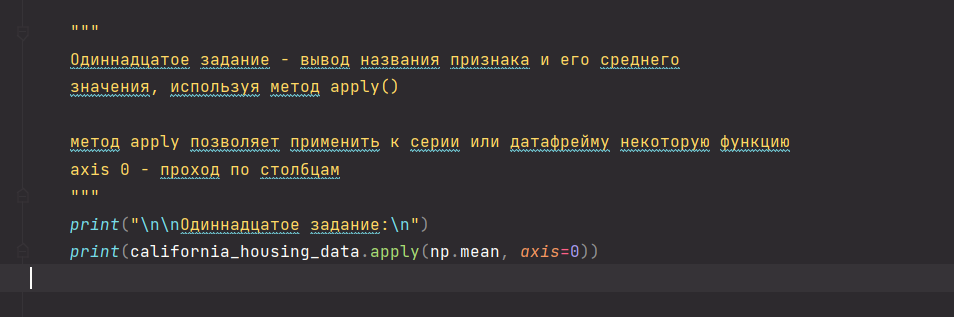


Рисунок 10 – Скриншот части программного кода пятой программы (Часть 3)

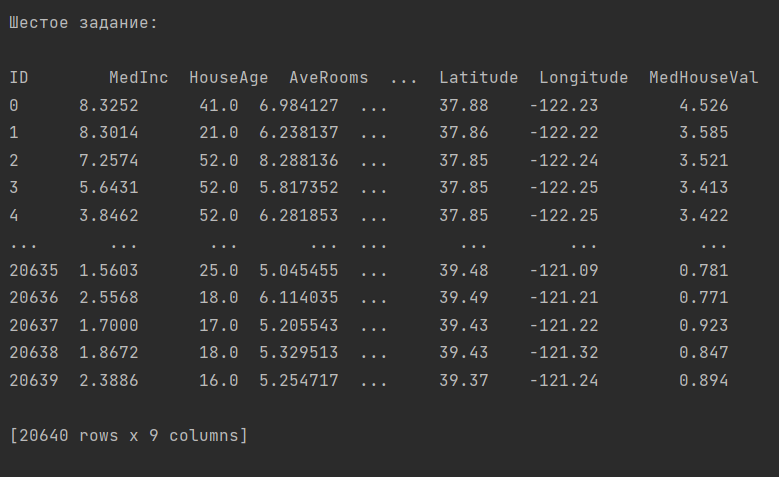


Рисунок 11 – Скриншот результата выполнения шестого задания

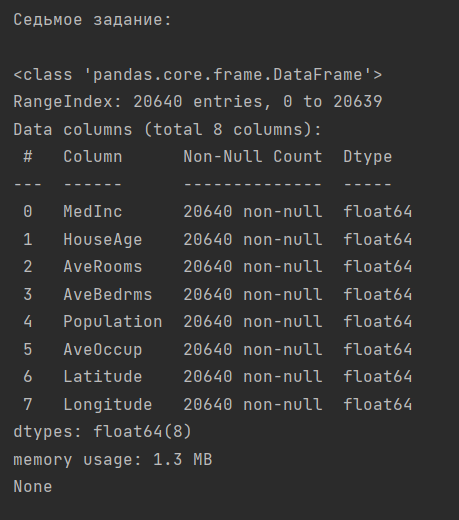


Рисунок 12 – Скриншот результата выполнения седьмого задания

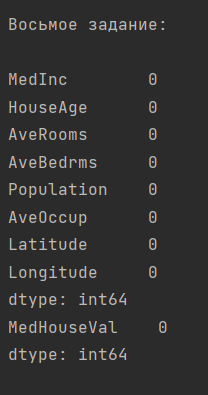


Рисунок 13 – Скриншот результата выполнения восьмого задания

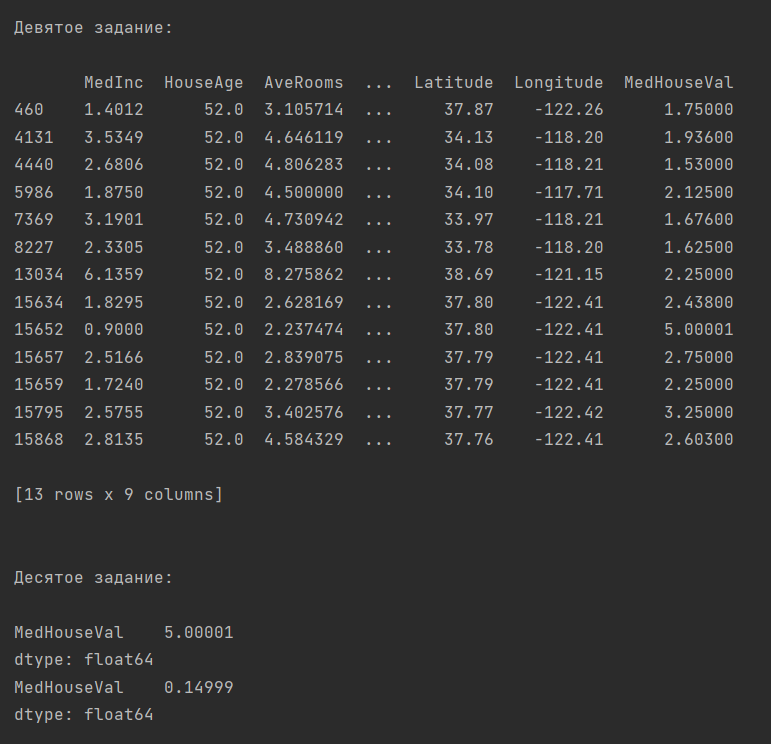


Рисунок 14 – Скриншот результата выполнения девятого и десятого заданий

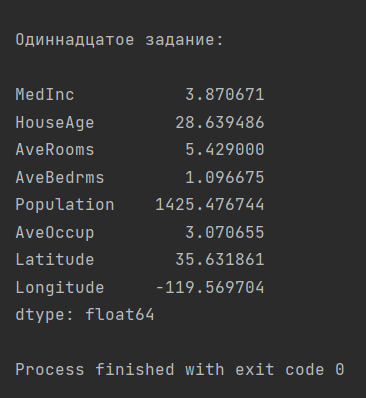


Рисунок 15 – Скриншот результата выполнения одиннадцатого задания

**Вывод**

В результате выполнения данной практической работы был получен ценный теоретический и практический опыт использования пакета scikit-learn, а также пакета с инструментами работы с особым типом данных DataFrame pandas, были написаны и протестированы программы, выполняющие требуемые задачи.