**Министерство науки и высшего образования Российской Федерации**

**Федеральное государственное автономное образовательное учреждение**

**высшего образования**

**«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»**

Институт цифрового развития

Кафедра инфокоммуникаций

**ОТЧЕТ**

**ПО ЛАБОРАТОРНОЙ РАБОТЕ №1**

**по дисциплине «Анализ данных»**

# «Введение в Pandas. Создание объекта Series в Pandas»

**Вариант 5**

|  |  |
| --- | --- |
|  | Выполнила:  Диченко Дина Алексеевна  студентка 3 курса, группы ИВТ-б-о-21-1  направление подготовки Информатика и вычислительная техника, очная форма обучения  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  (подпись) |
|  | Проверил:  Воронкин Роман Александрович  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  (подпись) |
|  |  | |

Отчет защищен с оценкой \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Дата защиты\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Ставрополь, 2023 г

**Цель**: исследование назначения и способов установки пакета Pandas, исследование способов создания основного объекта Series пакета Pandas.

**Порядок выполнения:**

1. Создала общедоступный репозиторий на github.
2. Выполнила клонирование репозитория.
3. Дополнила файл gitignore.
4. Организовала репозиторий в соответствии с моделью ветвления git-flow.
5. Решила следующие задачи:

Задание 1. Создать объект Series из списка чисел [1, 2, 3] с метками [a, b, c] и вывести его на печать.

Задание 2. Создать объект Series из словаря {'a': 1, 'b': 2, 'c': 3} и вывести его на

печать.

Задание 3. Создать объект массива NumPy arr = [1, 2, 3, 4, 5] , затем из него создать объект Series и вывести его на печать.

Задание 4. Создать объект Series с заполнением по умолчанию значением 0 и метками [a, b, c] и вывести его на печать.

Задание 5. В первой строке ввода с клавиатуры поступает строка с символами разлепленными пробелом, во второй строка с целыми числами. На их основе создать список кортежей и из него объект Series, вывести его на печать.

Задание 6. На вход поступает строка слов, разделенных пробелом. Создать на ее основе объект Series с метками 0, 1, 2,... Вывести его на печать.

Задание 7. На вход поступает последовательность из чисел 0 и 1, разделенных пробелом. Создать на ее основе объект Series, содержащий булевы значения True (1) , False (0) c метками [a, b, c, d,...]

Задание 8. Создать объект Series из списка дат ['2022-01-01', '2022-01-02', '2022 01-03'] , вывести его на печать. Даты в объекте Series должны быть представлены объектами datetime стандартной библиотеки языка программирования Python.

Задание 9. На вход подаются два целых числа n и m через пробел. На их основе создать объект Series с числами от n до m (включительно) с шагом 10 и метками [1, 3, 5, 7,...] . Вывести его на печать.

1. Самостоятельно придумала и решила задачу, где использовался бы объект Series библиотеки Pandas.

**Контрольные вопросы:**

1. Какое назначение библиотеки Pandas?

Pandas - это библиотека на языке Python для работы с данными, которая предоставляет широкий набор инструментов для анализа и манипуляции табличных данных. Pandas использует основной объект - DataFrame - для хранения и работы с данными, а также предоставляет многочисленные функции для их обработки.

2. Каким образом можно загрузить данные в Pandas DataFrame?

3. Какие типы данных могут содержать столбцы в Pandas DataFrame?

4. Какие недостатки могут быть у библиотеки Pandas?

- Потребление памяти. Pandas может быть не очень эффективной в использовании памяти для больших наборов данных. Это может привести к тому, что операции на больших данных могут занимать много времени и памяти.

- Сложность. Pandas имеет множество функций и методов, что делает ее мощной, но возможно, сложной в использовании, особенно для новичков в анализе данных.

- Производительность. Несмотря на то, что Pandas является достаточно быстрой библиотекой, некоторые операции, особенно на больших наборах данных, могут быть довольно медленными.

- Ограниченная поддержка параллелизма. Pandas не имеет полной поддержки параллельных вычислений, что может замедлить операции на больших наборах данных.

- Поддержка типов данных. Pandas может быть не очень гибкой в работе с некоторыми типами данных, например, с данными, которые имеют смешанные типы или с данными, которые требуют высокой точности.

5. Какая команда используется для установки библиотеки Pandas в Python?

6. Как проверить, установлена ли библиотека Pandas на компьютере?

7. Какие операционные системы поддерживаются библиотекой Pandas?

8. Как создать объект Series из списка значений?

9. Как создать объект Series с явным заданием индексов?

10. Как создать объект Series из словаря?

11. Как создать объект Series с заполнением по умолчанию?

12. Как создать объект Series из массива NumPy?

13. Как создать объект Series из списка кортежей?

**Вывод:** в результате выполнения работы были исследованы назначения и способов установки пакета Pandas, исследованы способы создания основного объекта Series пакета Pandas.