**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 2

|  |
| --- |
| Выполнил:  Богдан Александр Анатольевич  студент 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. Создал общедоступный репозиторий на github.
2. Выполнила клонирование репозитория.
3. Дополнил файл gitignore.
4. Организовал репозиторий в соответствии с моделью ветвления git-flow.

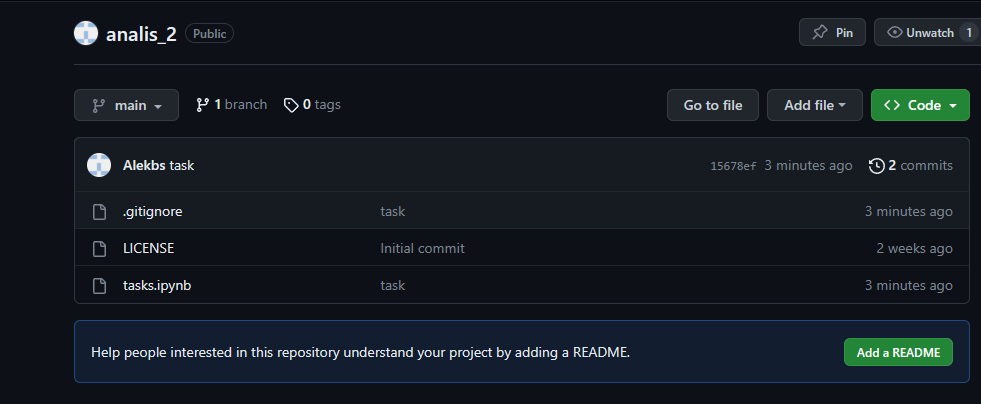


Рисунок 1. Создание репозитория

1. Решила следующие задачи:

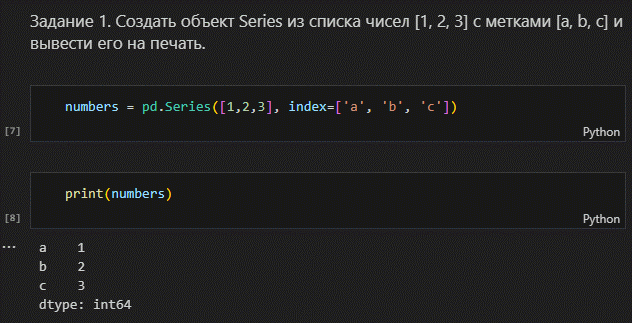
Задание 1. Создать объект Series из списка чисел [1, 2, 3] с метками [a, b, c] и вывести его на печать.

Рисунок 2. Выполнение задания 1

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

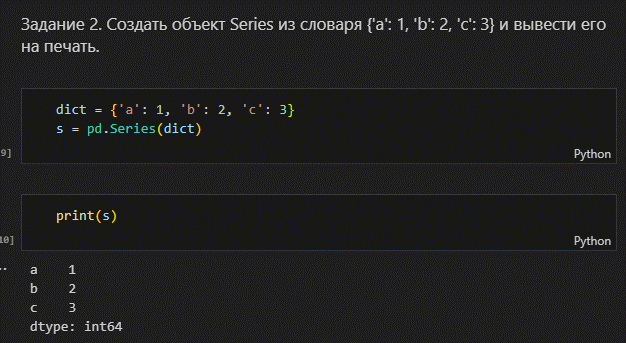


Рисунок 3. Выполнение задания 2

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

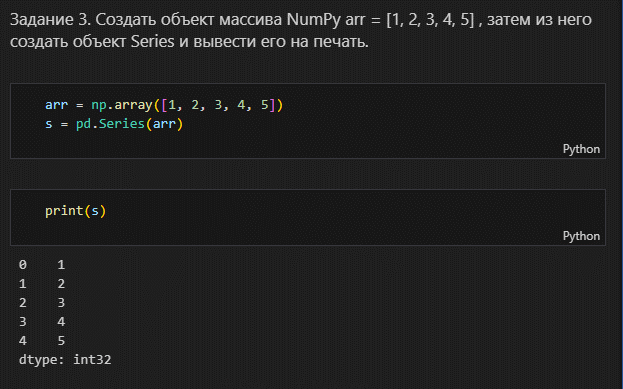


Рисунок 4. Выполнение задания 3

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

Изображение выглядит как текст, снимок экрана, программное обеспечение, Шрифт

Автоматически созданное описание

Рисунок 5. Выполнение задания 4

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

Изображение выглядит как текст, снимок экрана, программное обеспечение, Шрифт

Автоматически созданное описание

Рисунок 6. Выполнение задания 5

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

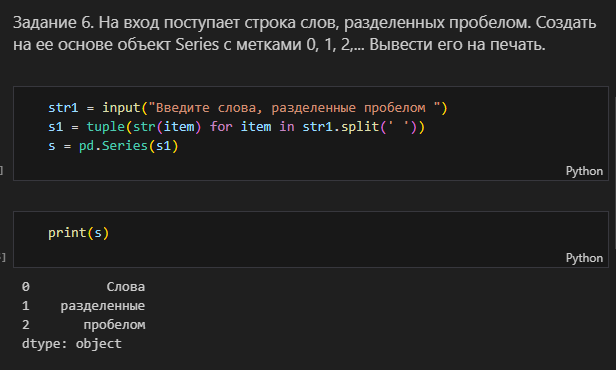


Рисунок 7. Выполнение задания 6

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,...]

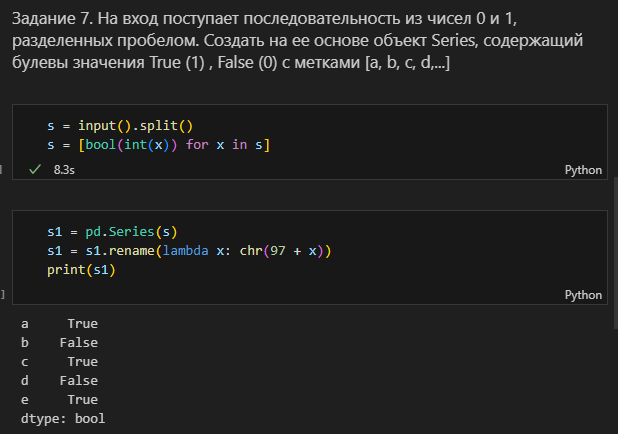


Рисунок 8. Выполнение задания 7

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.

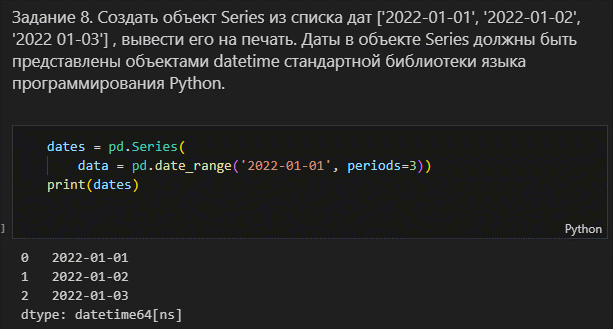


Рисунок 9. Выполнение задания 8

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.

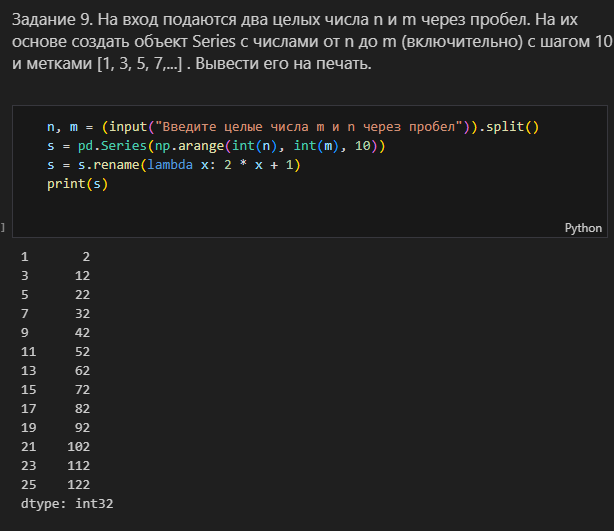


Рисунок 10. Выполнение задания 9

1. Самостоятельно придумал и решил задачу, где использовался

объект series

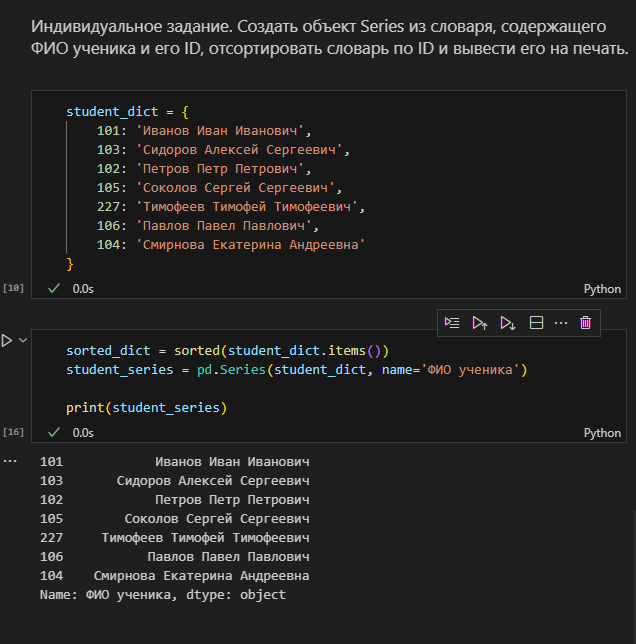


Рисунок 11. Выполнение индивидуального задания

# 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.

1. Каким образом можно загрузить данные в Pandas DataFrame?
2. Какие типы данных могут содержать столбцы в Pandas DataFrame? Каждый столбец может иметь разные типы данных, например, числа,

строки или даты.

1. 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.

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

pip install pandas.

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

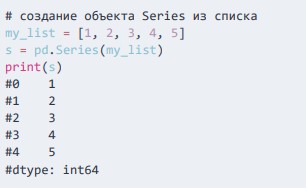
Чтобы убедиться, что Pandas установлен правильно, можно запустить интерпретатор Python и импортировать Pandas:

import pandas as pd

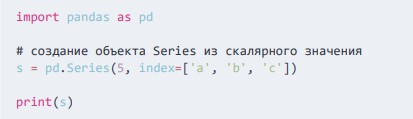
1. Какие операционные системы поддерживаются библиотекой Pandas?

Mac O,S Linux, Windows

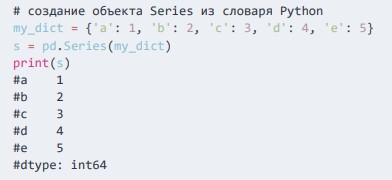
1. Как создать объект Series из списка значений?



1. Как создать объект Series с явным заданием индексов?



1. Как создать объект Series из словаря?



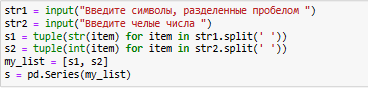
1. Как создать объект Series с заполнением по умолчанию?



1. Как создать объект Series из массива NumPy?



1. 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.