

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное  
образовательное учреждение высшего образования  
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»**

**Кафедра инфокоммуникаций**

**Лабораторная работа 1**

**Введение в Pandas. Создание объекта Series в  
Pandas**

Выполнил студент группы ИВТ-б-о-21-1

Пентухов С. А. «    »                      20    г.

Подпись студента\_\_\_\_\_

Работа защищена «    »                      20    г.

Проверил Воронкин Р.А.\_\_\_\_\_

(подпись)

Ставрополь, 2023 г

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

### Ход работы

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

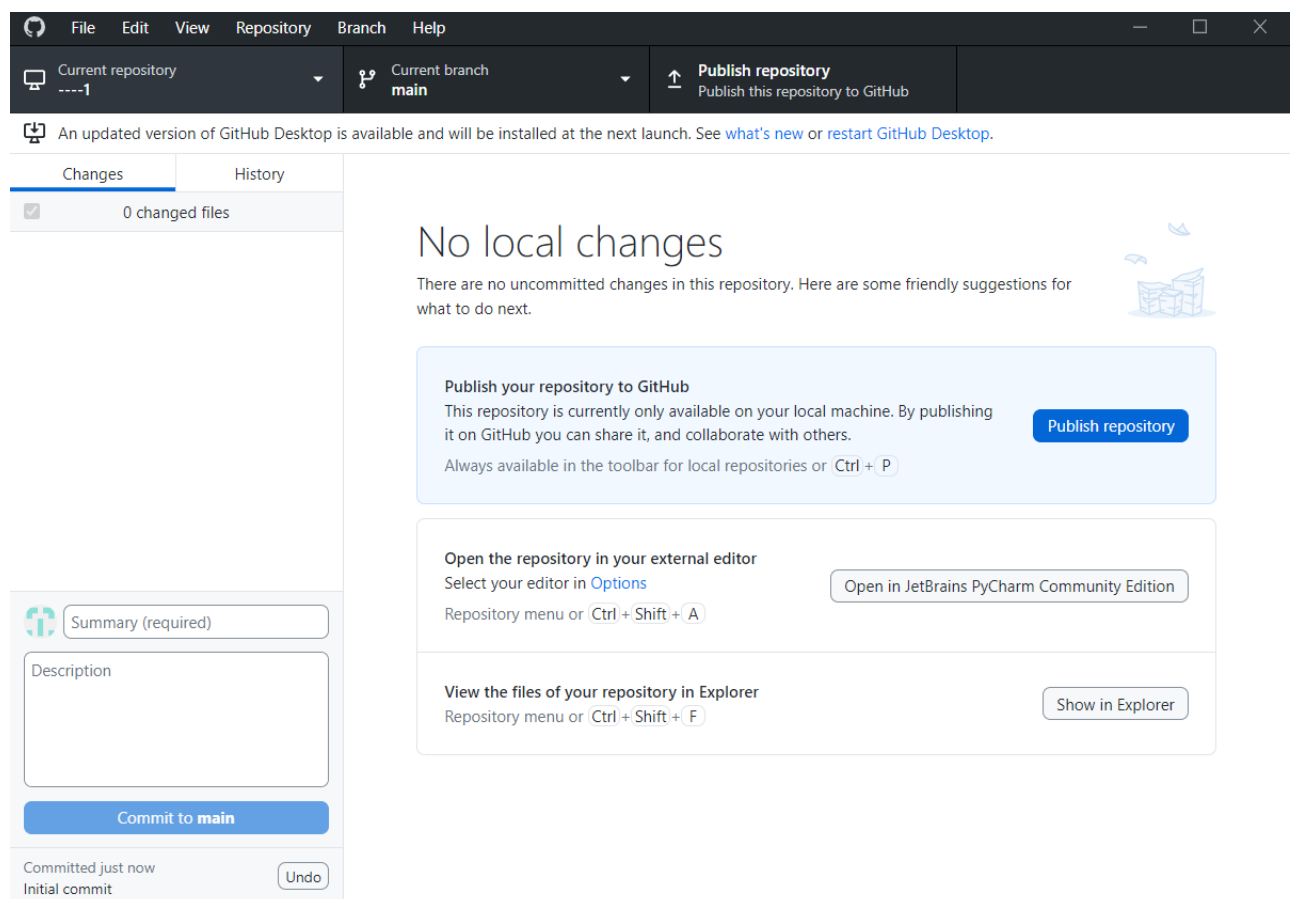


Рис. 1 Организация репозитория

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

```
In [3]: import pandas as pd
import numpy as np
```

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

```
In [2]: s = pd.Series([1,2,3], index=['a', 'b', 'c'])
print(s)
```

```
a    1
b    2
c    3
dtype: int64
```

Рис. 2 Выполнение задания 1

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

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

```
In [3]: my_list = {'a': 1, 'b': 2, 'c': 3}
s = pd.Series(my_list)
print(s)

a    1
b    2
c    3
dtype: int64
```

Рис. 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.

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

```
In [4]: my_array = np.array([1, 2, 3, 4, 5])
s = pd.Series(my_array)
print(s)

0    1
1    2
2    3
3    4
4    5
dtype: int32
```

Рис. 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.

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

```
In [5]: s = pd.Series(0, index=['a', 'b', 'c'])
print(s)

a    0
b    0
c    0
dtype: int64
```

Рис. 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 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.

```
In [7]: str1 = input("Введите символы, разделенные пробелом ")
str2 = input("Введите целые числа ")
s1 = tuple(str(item) for item in str1.split(' '))
s2 = tuple(int(item) for item in str2.split(' '))
my_list = [s1, s2]
s = pd.Series(my_list)
print(s)

Введите символы, разделенные пробелом a b c d i f
Введите целые числа 2 5 7 9 0 1
0    (a, b, c, d, i, f)
1    (2, 5, 7, 9, 0, 1)
dtype: object
```

Рис. 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.

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

```
In [8]: str1 = input("Введите слова, разделенные пробелом ")
s1 = tuple(str(item) for item in str1.split(' '))
s = pd.Series(s1)
print(s)
```

Введите слова, разделенные пробелом Снова седая ночь  
0 Снова  
1 седая  
2 ночь  
dtype: object

Рис. 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) с метками [a, b, c, d,...]

**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) с метками [a, b, c, d,...]**

```
In [4]: s =input()
s = s.replace("1", "True")
s = s.replace("0", "False")
st = s.split(' ')
st = pd.Series(st)
st = st.rename(lambda x: chr(97 + x))
print(st)
```

```
1 2 3 4 0
a      True
b         2
c         3
d         4
e      False
dtype: object
```

Рис. 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

**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.**

```
In [5]: s = pd.Series(
        data = pd.date_range('2022-01-01', periods=3))
print(s)
```

```
0    2022-01-01
1    2022-01-02
2    2022-01-03
dtype: datetime64[ns]
```

Рис. 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  $[1, 3, 5, 7, \dots]$ . Вывести его на печать.

```
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.

In [6]: a = int(input("Начальное число: "))
b = int(input("Конечное число: "))
s = []

sq = 0
for i in range(10):
    s.append(a + sq)
    sq += 10
    if s[i] > b:
        break

st = pd.Series(s)
st = st.rename(lambda x: int(1 + 2*x))
print(st)

Начальное число: 11
Конечное число: 32
1    11
3    21
5    31
7    41
dtype: int64
```

Рис. 10 Выполнение задания 9

Задание 10. Возьмем данные о погоде в разных городах в течение недели. У нас есть три города: Москва, Санкт-Петербург и Казань. Необходимо вычислить среднюю температуру для каждого города за неделю.

```
In [7]: weather_data = {'Москва': [15, 18, 20, 17, 16, 19, 18],
                        'Санкт-Петербург': [12, 14, 13, 10, 15, 11, 12],
                        'Казань': [22, 25, 23, 20, 24, 25, 23]}
moscow_weather = pd.Series(weather_data['Москва'])
average_temperature_moscow = moscow_weather.mean()
print("Средняя температура в Москве:", average_temperature_moscow)
print(moscow_weather)

Средняя температура в Москве: 17.571428571428573
0    15
1    18
2    20
3    17
4    16
5    19
6    18
dtype: int64
```

Рис. 11 Выполнение задания 10



```
In [10]: spb_weather = pd.Series(weather_data['Санкт-Петербург'])
average_temperature_spb = spb_weather.mean()
print("Средняя температура в Санкт-Петербурге:", average_temper
print(spb_weather)
```

Средняя температура в Санкт-Петербурге: 12.428571428571429

|   |    |
|---|----|
| 0 | 12 |
| 1 | 14 |
| 2 | 13 |
| 3 | 10 |
| 4 | 15 |
| 5 | 11 |
| 6 | 12 |

dtype: int64

Рис. 12 Выполнение задания 10

```
In [11]: kazan_weather = pd.Series(weather_data['Казань'])
average_temperature_kazan = kazan_weather.mean()
print("Средняя температура в Казани:", average_temperature_kaza
print(kazan_weather)
```

Средняя температура в Казани: 23.142857142857142

|   |    |
|---|----|
| 0 | 22 |
| 1 | 25 |
| 2 | 23 |
| 3 | 20 |
| 4 | 24 |
| 5 | 25 |
| 6 | 23 |

dtype: int64

Рис. 13 Выполнение задания 10

Вывод: Научился исследовать способы создания основного объекта Series пакета 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?

Каждый столбец может иметь разные типы данных, например, числа, строки или даты.

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?

`pip install pandas.`

6. 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`

7. Какие операционные системы поддерживаются библиотекой Pandas?

Mac OS Linux, Windows

8. Как создать объект Series из списка значений?

```
# создание объекта Series из списка
my_list = [1, 2, 3, 4, 5]
s = pd.Series(my_list)
print(s)
#0    1
#1    2
#2    3
#3    4
#4    5
#dtype: int64
```

9. Как создать объект Series с явным заданием индексов?

```
import pandas as pd

# создание объекта Series из скалярного значения
s = pd.Series(5, index=['a', 'b', 'c'])

print(s)
```

10. Как создать объект Series из словаря?

```
# создание объекта Series из словаря Python
my_dict = {'a': 1, 'b': 2, 'c': 3, 'd': 4, 'e': 5}
s = pd.Series(my_dict)
print(s)
#a    1
#b    2
#c    3
#d    4
#e    5
#dtype: int64
```

11. Как создать объект Series с заполнением по умолчанию?

```
s = pd.Series(0, index=['a', 'b', 'c'])
```

12. Как создать объект Series из массива NumPy?

```
my_array = np.array([1, 2, 3, 4, 5])
s = pd.Series(my_array)
```

13. Как создать объект Series из списка кортежей?

```
str1 = input("Введите символы, разделенные пробелом ")
str2 = input("Введите целые числа ")
s1 = tuple(str(item) for item in str1.split(' '))
s2 = tuple(int(item) for item in str2.split(' '))
my_list = [s1, s2]
s = pd.Series(my_list)
```

