МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ» (ПГУ)

ОБРАБОТКА ДАННЫХ В PANDAS

Методические указания

Составители:
А. Г. Петренко, С. Н. Катков

ПЕНЗА 2026

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ» (ПГУ)

ОБРАБОТКА ДАННЫХ В PANDAS

Методические указания

Составители:
А. Г. Петренко, С. Н. Катков

Пенза
Издательство ПГУ
2026

УДК 004(075.8)
ББК 65ся7
О-23

Рецензент
кандидат экономических наук
заведующий кафедрой «Информационные системы и цифровые технологии»
Пензенского казачьего института технологий (филиал)
Московского государственного университета технологий и управления
имени К. Г. Разумовского (Первый казачий университет)
А. А. Тусков

Обработка данных в Pandas : метод. указания / сост.:
О-23 А. Г. Петренко, С. Н. Катков. – Пенза : Изд-во ПГУ, 2026. – 52 с.

Рассматриваются способы и методы Pandas, связанные с очисткой и подготовкой наборов данных для последующего анализа и построения моделей.

Издание подготовлено на кафедре «Цифровая экономика» ПГУ и предназначено для обучающихся по направлениям подготовки 38.03.01 «Экономика», 38.03.05 «Бизнес-информатика» и другим смежным направлениям.

УДК 004(075.8)
ББК 65ся7

© Пензенский государственный
университет, 2026

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 5
1. ОЧИСТКА И ПОДГОТОВКА ДАННЫХ 6
1.1. Дубликаты 7
1.1.1. Удаление дубликатов 8
1.2. Работа с пропущенными данными 9
1.2.1. Обнаружение пропущенных значений 9
1.2.2. Обработка пропущенных значений 10
1.3. Нетипичные данные (выбросы) 15
1.4. Переименование меток строк и столбцов во фрейме данных 16
1.4.1. Переименование столбцов 16
1.4.2. Перевод наименований столбцов в строчное написание 18
1.4.3. Переименование меток строк (индексов) 18
1.5. Преобразование данных в нужный формат 19
1.5.1. Перевод данных в нижний регистр и удаление пробелов 19
1.5.2. Приведение данных к единому формату даты 20
1.5.3. Обработка смешанных форматов дат 22
1.6. Преобразование категориальных данных 22
1.7. Задание 24
Список литературы 24
2. ВЫБОР И ИЗМЕНЕНИЕ ДАННЫХ 25
2.1. Выбор данных из фрейма данных Pandas 25
2.1.1. Выделение одного столбца с помощью [] 25
2.1.2. Выделение двух или более столбцов с помощью [[]] 25
2.1.3. Выделение одной строки с помощью [] 26
2.1.4. Выделение двух или более строк с помощью [] 26
2.1.5. Использование loc[] и iloc[] для извлечения строк и столбцов 27

Учебное издание

Обработка данных в Pandas

Составители:
Петренко Александр Григорьевич,
Катков Сергей Николаевич

Редактор *В. В. Чувашова*
Технический редактор *М. А. Борисов*
Компьютерная верстка *М. А. Борисова*

Подписано в печать 22.09.2026.
Формат 60×84 1/16. Усл. печ. л. 3,02.
Заказ № 363. Тираж 10.

Издательство ПГУ.
440026, г. Пенза, ул. Красная, 40.
Тел./факс: (8412) 66-60-49, 66-67-77; e-mail: iic@pnzgu.ru