

УДК 004.6(075)
ББК 32.97я7
Т45

*Печатается по решению редакционно-издательского совета
Казанского национального исследовательского технологического университета*

Рецензенты:
канд. пед. наук, доц. С. М. Куценко
канд. экон. наук, доц. О. С. Семичева

Титов А. Н.
Т45 Обработка данных в Python. Основы работы с библиотекой Pandas : учебно-методическое пособие / А. Н. Титов, Р. Ф. Тазиева; Минобрнауки России, Казан. нац. исслед. технол. ун-т. – Казань : Изд-во КНИТУ, 2022. – 116 с.

ISBN 978-5-7882-3164-8

Рассмотрены основы анализа и обработки данных в Python с использованием библиотеки Pandas. Приведены примеры работы с объектами Series и DataFrame: осуществление доступа к данным, внесение изменений в DataFrame, визуализация полученных результатов, а также сортировка, фильтрация и группировка данных. Для оценки уровня усвоения студентами пройденного материала предложены варианты самостоятельных работ.

Предназначено для бакалавров, обучающихся по направлениям подготовки 28.03.02 «Наноинженерия», 09.03.02 «Информационные системы и технологии», изучающих дисциплины «Обработка экспериментальных данных», «Алгоритмы и структуры данных», «Большие данные», «Теория информации, данные, знания», «Прикладная математика», «Вычислительная математика».

Подготовлено на кафедре информатики и прикладной математики.

УДК 004.6(075)
ББК 32.97я7

ISBN 978-5-7882-3164-8

© Титов А. Н., Тазиева Р. Ф., 2022
© Казанский национальный исследовательский
технологический университет, 2022

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	4
1. СТРУКТУРА ДАННЫХ В PANDAS	6
1.1. Объект Series	6
1.1.1. Создание объекта Series	6
1.1.2. Свойства объекта Series	9
1.1.3. Работа с объектами Series	9
1.2. Объект DataFrame	13
1.2.1. Создание объекта DataFrame	14
1.2.2. Импорт данных в объект DataFrame и сохранение в файл	17
1.2.3. Вывод на экран содержимого объекта DataFrame	19
1.2.4. Атрибуты объекта DataFrame	22
2. ОПЕРАЦИИ С ОБЪЕКТОМ DATAFRAME	27
2.1. Доступ к строкам и столбцам объекта DataFrame.....	27
2.2. Внесение изменений в объект DataFrame	30
2.2.1. Изменение содержимого объекта DataFrame	31
2.2.2. Добавление новых записей и столбцов в объект DataFrame	35
2.2.3. Удаление строк и столбцов.....	39
2.2.4. Разбивка и объединение объектов DataFrame	43
2.3. Итерирование по элементам строки объекта DataFrame.....	46
2.4. Предварительный анализ и обработка данных.....	50
2.5. Сортировка данных	56
2.6. Фильтрация данных.....	58
2.7. Группировка данных	63
3. ВИЗУАЛИЗАЦИЯ ДАННЫХ.....	82
3.1. Визуализация данных в библиотеке Pandas.....	82
3.2. Визуализация данных в библиотеке Matplotlib	91
3.3. Визуализация данных в библиотеке Seaborn.....	94
Задания для самостоятельной работы студентов.....	112
ЛИТЕРАТУРА.....	115