Разработка программного обеспечения на языке Python

Обзорная панель

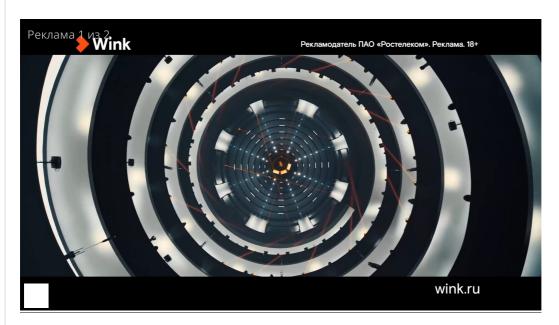
Мои курсы Разработка ПО на языке Python Анализ данных и машинное обучение

<u>Лекция 5. Анализ данных на Python</u>

Лекция 5. Анализ данных на Python

Посмотрите видеоуроки и ответьте на контрольные вопросы после лекции

Библиотека Pandas. Dataframe



Библиотека Pandas позволяет работать с различными форматами данных: csv, excel, sql, html и многие другие. В Pandas есть метод с именем .read_ для чтения содержимого файла. Все доступные форматы можно посмотреть через нижнее подчеркивание.

Рассмотрим работу методов библиотеки Pandas на примере данных о об цветках ириса разных сортов. Данные содержатся в файле iris. Это известный как набор данных о цветках ириса, представляет собой тип набора данных для многомерного анализа.

Здесь мы импортировали данные, записали в датафрейм df. Аргумент sep указывает на разделитель столбцов. Наша цель - загрузить данные и проанализировать их, чтобы сделать выводы. Каждый объект описывается 5 переменными: длина и ширина лепестка и чашелистика, а также вид.

import pandas as pd

В результате получим:

	sep_length	sep_width	petal_l	petal_w	class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa



© 2010-2023 Центр обучающих систем
Сибирского федерального университета, sfu-kras.ru

Разработано на платформе moodle
Вeta-version (3.9.1.5.w3)

Политика конфиденциальности

Соглашение о Персональных данных

Политика допустимого использования

Контакты +7(391) 206-27-05

info-ms@sfu-kras.ru

Скачать мобильное приложение

Инструкции по работе в системе