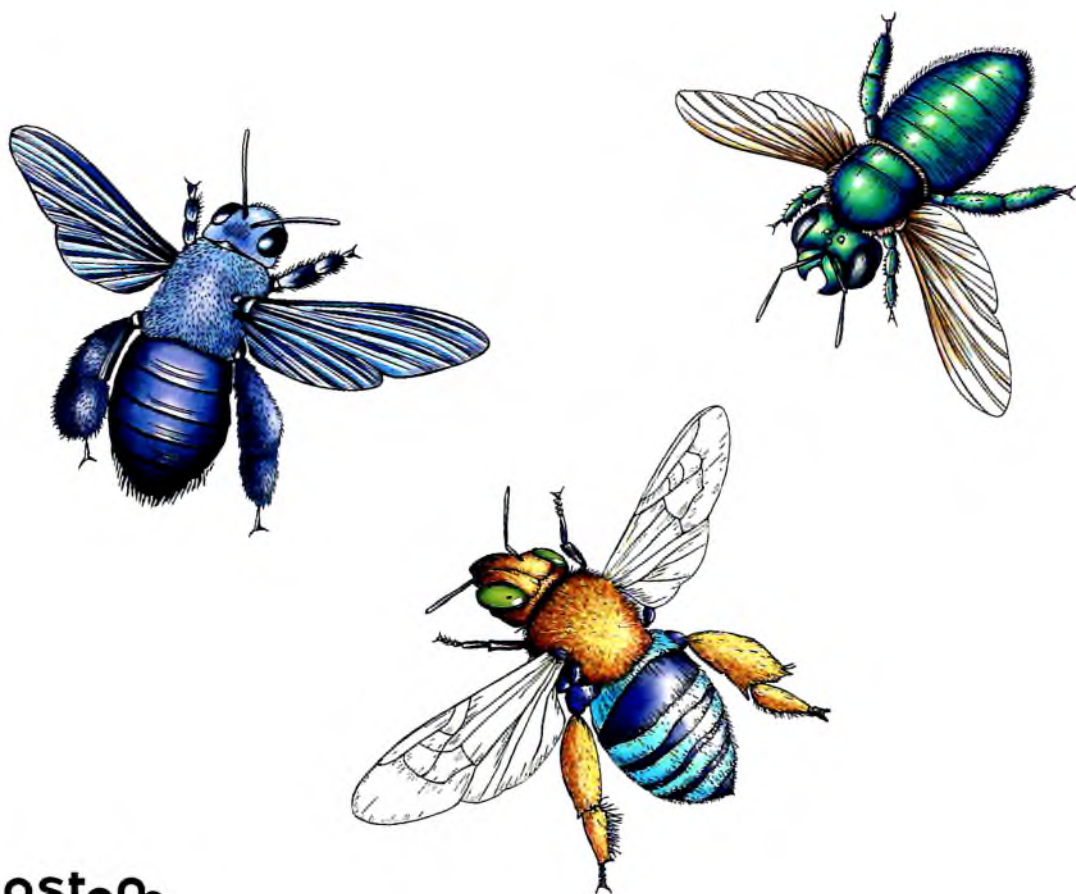


Визуализация данных с помощью Python и JavaScript

Анализ и преобразование данных



Оглавление

Предисловие	10
Второе издание	14
Принятые в книге обозначения	15
Использование примеров кода	16
Благодарности	17
Введение	18
Для кого эта книга?	19
Почему именно Python и JavaScript?	22
Чему вы научитесь	25
Предварительные сведения	26
Тулчейн для визуализации данных	27
Как пользоваться этой книгой	30
Немного контекста	31
Резюме	33
Рекомендуемые книги	34

Раздел I. Базовый пакет инструментов

Глава 1. Подготовка окружения	36
Сопутствующий код	36
Python	36
Установка дополнительных библиотек	37
JavaScript	39
Базы данных	41
Интегрированная среда разработки	43
Резюме	44
Глава 2. Обучающий мостик между Python и JavaScript	45
Сходство и различия	45
Взаимодействие с кодом	46
Строим мост	49
Примеры различий	77
Шпаргалка	90
Резюме	92
Глава 3. Чтение и запись данных с помощью Python	94
Просто ли это?	94
Передача данных	95
Работа с системными файлами	96
CSV, TSV и табличные форматы данных	97
JSON	100
SQL	105

MongoDB	116
Работа с датами, временем и сложными типами данных	122
Резюме	123
Глава 4. Основы веб-разработки	125
Общая картина	125
Одностраничные приложения	126
Настройка инструментов	126
Создание веб-страницы	130
Chrome DevTools	140
Базовая страница с плейсхолдерами	142
Позиционирование и изменение размера контейнеров с помощью Flex	146
Масштабируемая векторная графика	155
Резюме	169

Раздел II. Получение данных

Глава 5. Получение данных из интернета с помощью Python	173
Получение данных из интернета с помощью библиотеки Requests	173
Получение файлов данных с помощью Requests	174
Использование Python для получения данных через web API	177
Доступ к web API с помощью библиотек	183
Скрейпинг данных	189
Получение объекта BeautifulSoup	191
Выбор тегов	192
Резюме	202
Глава 6. Эффективный скрейпинг с помощью Scrapy	203
Установка Scrapy	204
Постановка целей	205
Работа с XPath в Scrapy	207
Первый паук Scrapy	214
Скрейпинг биографических страниц лауреатов	221
Цепочка запросов и извлечение данных	224
Конвейеры Scrapy	229
Скрейпинг текста и изображений с помощью конвейера	231
Резюме	239

Раздел III. Очистка и исследование данных с помощью pandas

Глава 7. Введение в NumPy	243
Массив NumPy	244
Создание функций для работы с массивами	251
Резюме	253
Глава 8. Знакомство с библиотекой pandas	254
Почему pandas оптимальна для визуализации данных	254
Зачем разработали pandas	254
Классификация данных и измерения	255
DataFrame	256

Создание и сохранение структур DataFrame	261
Создание DataFrame из Series	272
Резюме	275
Глава 9. Очистка данных с помощью pandas	276
Чистая правда о грязных данных	276
Проверка качества данных	278
Индексы и отбор данных с помощью pandas	282
Очистка данных	286
Полная функция для очистки данных	306
Добавление столбца <code>born_in</code>	307
Сохранение очищенных наборов данных	312
Резюме	313
Глава 10. Визуализация данных с помощью Matplotlib	314
Pyplot и объектно-ориентированная библиотека Matplotlib	314
Запуск интерактивной сессии	315
Создание интерактивных графиков с помощью глобального состояния pyplot	316
Фигуры и объектно-ориентированная Matplotlib	322
Типы графиков	327
Seaborn	336
Резюме	345
Глава 11. Анализ данных с помощью pandas	347
Начало исследования	348
Построение графиков с помощью pandas	350
Гендерные диспропорции	352
Национальные тренды	360
Возраст и ожидаемая продолжительность жизни лауреатов	373
Нобелевская «диаспора»	380
Резюме	382

Раздел IV. Передача данных

Глава 12. Передача данных	385
Передача данных	386
Доставка файлов данных	391
Динамическое обновление данных с помощью Flask API	396
Использование динамической или статической доставки	398
Резюме	399
Глава 13. RESTful Data с помощью Flask	400
Инструменты для работы с RESTful	400
Создание базы данных	401
Flask RESTful для работы с данными	402
Добавление маршрутов RESTful API	405
Расширение API с помощью MethodView	411
Пагинация возвращаемых данных	414
Удаленное развертывание API на Heroku	418
Резюме	422

Раздел V. Визуализация данных с помощью D3 и Plotly

Глава 14. Перенос диаграмм в интернет с помощью Matplotlib и Plotly	425
Создание статических диаграмм с помощью Matplotlib	425
Построение диаграмм с помощью Plotly	430
Из Notebook в веб-формат с помощью Plotly	444
Создание нативных JavaScript-диаграмм с помощью Plotly	448
Интерактивная визуализация Plotly с помощью JavaScript и HTML	454
Резюме	459
Глава 15. Разработка концепции визуализации Нобелевской премии	460
Для кого эта визуализация?	460
Выбор визуальных элементов	461
Строка меню	462
Распределение премии по годам	463
Карта, показывающая выборку стран нобелевских лауреатов	464
Столбчатая диаграмма, показывающая количество лауреатов по странам	465
Список выбранных лауреатов	466
Визуализация целиком	468
Резюме	469
Глава 16. Создание визуализации	470
Предварительные сведения	471
HTML-каркас	473
Стили CSS	477
Движок JavaScript	481
Запуск приложения для визуализации данных о нобелевских лауреатах	496
Резюме	497
Глава 17. Введение в D3 на примере столбчатой диаграммы	498
Формулирование задачи	499
Работа с выборкой	499
Добавление элементов DOM	503
Использование D3	510
Шкалы в D3: от данных к их визуальному представлению	510
Привязка данных к элементам DOM — главное преимущество D3	516
Обновление DOM при изменении данных	516
Сборка столбчатой диаграммы	520
Оси и метки	523
Переходы	529
Резюме	534
Глава 18. Визуализация отдельных премий	535
Создание структуры	535
Шкалы	536
Оси	537
Метки номинаций	538
Вложенные данные	540
Добавление лауреатов с помощью вложенных объединений данных	543
Добавим немного блеска!	547
Резюме	549

Глава 19. Картографирование с помощью D3	550
Доступные карты	550
Форматы данных для картографирования в D3	551
Библиотека D3-geo, проекции и пути	556
Соединение элементов воедино	562
Обновление карты	565
Добавление индикаторов показателей	569
Готовая карта	572
Создание простой всплывающей подсказки	573
Резюме	578
Глава 20. Визуализация данных отдельных лауреатов	579
Создание списка лауреатов	580
Создание биографического блока	584
Резюме	587
Глава 21. Строка меню	589
Создание HTML-элементов с помощью D3	590
Создание строки меню	590
Резюме	601
Глава 22. Заключение	602
Подведение итогов	602
Дальнейшее развитие	605
Заключительные замечания	607
Приложение А. Паттерн enter/exit библиотеки D3	608
Метод enter	609
Доступ к привязанным данным	613
 Об авторе	 616
Послесловие	617
Алфавитный указатель	618