

УТВЕРЖДЕНА

Приказом Ректора АНО ВО
«Центральный университет»
Ивашкевич Е.В.
от «19» января 2024 г. № 0119.37

**Рабочая программа дисциплины (модуля)
«Python для анализа данных»
дополнительной профессиональной программы – программы
профессиональной переподготовки «Академия data science»**

Траектория: Продуктовый менеджмент

**Москва
2024**

Содержание

| | |
|--|----------|
| 1. Краткая характеристика дисциплины (модуля) | 3 |
| 2. Тематический план | 4 |
| 3. Содержание дисциплины (модуля) | 4 |
| 4. Учебно-методическое обеспечение | 5 |
| 5. Материально-техническое обеспечение | 5 |
| 6. Методические и оценочные материалы | 7 |

1. Краткая характеристика дисциплины (модуля)

Изучение дисциплины (модуля) «Python для анализа данных» помогает слушателям освоить один из самых популярных и универсальных языков программирования, широко используемых в науке, бизнесе и IT-индустрии. Освоение основ Python позволяет эффективно автоматизировать задачи, анализировать данные и создавать программные продукты, что значительно расширяет профессиональные возможности.

Цель изучения дисциплины (модуля): формирование у слушателей знаний базовых концепций программирования на языке Python для разработки простых программ и решения прикладных задач.

Задачи изучения дисциплины (модуля):

- формирование знания объектов и методов pandas для работы с датафреймами, объектов и методов Numpy для работы с многомерными массивами;
- формирование знания объектов и метод matplotlib и seaborn для визуализации;
- формирование знания базовых методов анализа данных, базовых методов решения прикладных задач, основных методов и библиотек для работы с API, консолью, git, cron;
- формирование умения обрабатывать данные в pandas и numpy;
- формирование умения проводить вычисления из линейной алгебры с помощью многомерных объектов numpy;
- формирование умения строить различные виды графиков с помощью библиотек matplotlib и seaborn;
- формирование умения подключаться к API;
- формирование умения настраивать автоматическую рассылку метрик через телеграм-бота;
- формирование умения решать прикладные задачи с использованием python
- формирование умения использовать основные средства разработки: jupyter notebook, jupyterhub;
- формирование навыка решать прикладные задачи с использованием python;
- формирование навыка анализировать данные с использованием python
- формирование навыка подготавливать и обрабатывать данные с помощью python;
- формирование навыка выбирать корректный тип визуализации и строить визуализации средствами matplotlib и seaborn;
- формирование навыка настраивать подключения к внешним источникам.

2. Тематический план

| № п/п | Наименование раздела дисциплины (модуля) | Трудоемкость, академические часы | | | | ТКУ (текущий контроль успеваемости) |
|---|--|----------------------------------|-----------|----------|------------------------|--|
| | | <i>Очная форма</i> | | | | |
| | | Аудиторная работа | | Контроль | Самостоятельная работа | |
| Лекции | Практические занятия | | | | | |
| 1 | Введение и настройка среды разработки | 6 | 6 | | 18 | Домашнее задание |
| 2 | Pandas | 6 | 7 | | 18 | Домашнее задание |
| 3 | NumPy | 7 | 7 | | 18 | Домашнее задание |
| 4 | API and Virtual machine | 7 | 7 | | 20 | Домашнее задание |
| 5 | Визуализация | 7 | 7 | | 20 | Домашнее задание Проект |
| | <i>Зачет с оценкой</i> | | | 4 | | |
| | Итого: | 33 | 34 | 4 | 94 | |
| Объем дисциплины (модуля) (в ак. ч.) | | 165 | | | | |

3. Содержание дисциплины (модуля)

| №п/п | Наименование раздела дисциплины (модуля) | Содержание дисциплины (модуля) по темам |
|------|--|---|
| 1 | Введение и настройка среды разработки | Среда разработки: Jupyter Notebook, Google Colab, JupyterHub. Командная строка |
| 2 | Pandas | Основные структуры данных: DataFrame, Series. Базовые методы DataFrame (head, tail, info, describe, dtypes, shape, len, проверка на null). Импорт/экспорт данных: файлы, базы данных. Слайсинг и фильтрация DataFrame: по колонкам, по индексам, создание масок, регулярные выражения для работы со строками. Сортировка. Соединение и преобразование DataFrame: merge, append, pivot, groupby. Синтаксический сахарок: lambda functions, map(). Факторный и когортный анализ. Основные продуктовые метрики и как с ними работать |
| 3 | NumPy | Устройство NumPy и парадигма осей. Векторные вычисления: скаляр, вектор, матрица. Операции с осями. Практические задачи на NumPy: линейная алгебра, упрощение вычислений, временные ряды и прогнозирование |
| 4 | API and Virtual machine | Концепция API: with token, without token. Подключение к API, трансфер данных. Проектирование архитектуры потоков данных. Работа на виртуальной машине: Cron, command line |
| 5 | Визуализация | Основные типы визуализаций: plot, scatter, bar. Выбор визуализации под конкретную задачу. Библиотеки: Matplotlib, Seaborn, Plotly |

4. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый слушатель в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Слушателям обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Чернышев, С. А. Основы программирования на Python : учебник для вузов / С. А. Чернышев. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 349 с. — (Высшее образование). — ISBN 978-5-534-17139-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/567821>.

Дополнительная литература:

1. Федоров, Д. Ю. Программирование на python : учебное пособие для вузов / Д. Ю. Федоров. — 6-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 187 с. — (Высшее образование). — ISBN 978-5-534-19666-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/556864>.

5. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

| № | Наименование портала (издания, курса, документа) | Ссылка |
|----|--|---|
| 1 | Катастрофы, стихийные бедствия, аварии, эпидемии. Солнечная и геомагнитная активность. /ежедневный обзор | http://www.disasters.chat.ru |
| 2 | Каталог по безопасности жизнедеятельности | http://www.eun.chat.ru |
| 3 | Научная электронная библиотека eLibrary.ru библиотека | https://elibrary.ru/defaultx.asp |
| 4 | База данных для IT-специалистов | https://habr.com |
| 5 | База данных ScienceDirect | https://www.sciencedirect.com |
| 6 | Официальный сайт Министерства науки и высшего образования Российской Федерации | https://minobrnauki.gov.ru/ |
| 7 | Федеральный портал «Российское образование» | https://www.edu.ru/ |
| 8 | Информационная система "Единое окно доступа к образовательным ресурсам" | http://window.edu.ru/ |
| 9 | Единая коллекция цифровых образовательных ресурсов | http://school-collection.edu.ru/ |
| 10 | Федеральный центр информационно - образовательных ресурсов | http://fcior.edu.ru/ |

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

| Наименование ПО | Производство | Лицензионное / свободно распространяемое |
|---|---------------|--|
| Операционные системы: | | |
| Microsoft Imagine (Windows Client, Server) | зарубежное | лицензионное |
| Браузеры: | | |
| Яндекс.Браузер | отечественное | свободно распространяемое |
| Google Chrome | зарубежное | свободно распространяемое |
| Офисные приложения: | | |
| Microsoft Imagine (Visio, OneNote) | зарубежное | лицензионное |
| TeXstudio | зарубежное | свободно распространяемое |
| Adobe Acrobat Reader | зарубежное | свободно распространяемое |
| Программное обеспечение для планирования и учета времени: | | |
| Toggle app | зарубежное | свободно распространяемое |
| Системы управления проектами: | | |
| Microsoft Imagine (Project) | зарубежное | лицензионное |
| Системы управления базами данных: | | |
| Microsoft Imagine (SQL Server) | зарубежное | лицензионное |
| Системы резервного копирования (backup): | | |
| Acronis Backup Advanced for HyperV | зарубежное | лицензионное |
| Справочно-правовые системы: | | |
| КонсультантПлюс: справочно-правовая система | отечественное | лицензионное |
| Средства антивирусной защиты: | | |
| Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition | отечественное | лицензионное |
| Пакеты программных средств и библиотек: | | |
| AutoPsy | зарубежное | свободно распространяемое |
| Interactive Disassembler (IDA) | зарубежное | свободно распространяемое |
| Системы управления библиографической информацией: | | |

| | | |
|--------------------------|------------|---------------------------|
| Zotero | зарубежное | свободно распространяемое |
| Сервисы и службы: | | |
| Bind | зарубежное | свободно распространяемое |
| Docker | зарубежное | свободно распространяемое |

6. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Python для анализа данных» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, практические занятия, домашние задания, проект, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Участие в семинаре (практическом занятии) – активная работа слушателя на семинаре, его ответы на вопросы преподавателя и участие в дискуссии.

Для успешного участия в семинаре слушателям рекомендуется заранее ознакомиться с темой обсуждения, прочитать необходимые материалы и подготовить вопросы. Важно активно слушать и вовлекаться в дискуссию, высказывая свои мнения и аргументируя их. При ответах на вопросы преподавателя стоит быть уверенным, четким и логичным, опираясь на изученный материал. Также полезно поддерживать диалог с однокурсниками, чтобы обогатить обсуждение и расширить свои знания.

Проект – это целенаправленная деятельность, имеющая определенные цели, задачи и временные рамки, в результате которой создается уникальный продукт или услуга.

Для успешной подготовки проекта рекомендуется следует выполнять следующие рекомендации:

- четко определите цель и задачи проекта, чтобы понимать, какой результат вы хотите достичь;
- составьте план работы, разбив проект на этапы с указанием сроков выполнения каждого из них;
- используйте разнообразные источники информации и инструменты для исследования темы, чтобы обеспечить качественную основу для вашего проекта;
- регулярно проверяйте прогресс и вносите коррективы в план, если это необходимо, чтобы оставаться на правильном пути к завершению проекта.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Самостоятельная работа – работа слушателей, направленная на углубленное

изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы слушатели взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи слушателя включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме **зачета с оценкой**.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом:

| Десятибалльная оценка | Пятибалльная оценка | Оценка за зачет | Общая характеристика результата обучения по дисциплине (модулю) |
|------------------------------|----------------------------|------------------------|--|
| 10 | Отлично | Зачтено | Слушатель полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Слушатель хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами. |
| 9 | Отлично | Зачтено | |
| 8 | Отлично | Зачтено | |
| 7 | Хорошо | Зачтено | Слушатель обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его |
| 6 | Хорошо | Зачтено | |

| Десятибалльная оценка | Пятибалльная оценка | Оценка за зачет | Общая характеристика результата обучения по дисциплине (модулю) |
|-----------------------|---------------------|-----------------|--|
| | | | ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Слушатель хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами. |
| 5 | Удовлетворительно | Зачтено | Слушатель обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Слушатель способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования. |
| 4 | Удовлетворительно | Зачтено | |
| 3 | Не сдан | Не зачтено | Слушатель не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы. |
| 2 | Не сдан | Не зачтено | |
| 1 | Не сдан | Не зачтено | |

Дисциплина (модуль) «Python для анализа данных» оценивается следующим образом:

| Активность | Вес | Количество | Описание |
|------------------|-----|------------|---|
| Домашние задания | 50% | 5 | Письменная работа с набором задач, которые нужно решить за ограниченное время |
| Проект | 50% | 1 | Защита проекта |

Формула расчёта итоговой оценки по дисциплине (модулю) «Python для анализа данных»: $\langle 0,8 \times \text{среднее за домашние задания} + 0,2 \times \text{проект} \rangle$.

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание: Python для анализа данных и синтаксические конструкции

1. История и особенности языка Python

Напишите краткий обзор (150-200 слов) о том, когда и почему был создан Python, а также о его основных особенностях.

2. Установка и настройка окружения

Установите Python и настройте среду разработки (например, PyCharm или VS Code). Сделайте скриншоты процесса установки и настройки, а затем напишите краткое описание (50-100 слов) о том, как вы это сделали.

3. Основные синтаксические конструкции

Напишите программу, которая запрашивает у пользователя его имя и возраст, а затем выводит сообщение в формате: "Привет, [имя]! Тебе [возраст] лет."

4. Числовые типы

Создайте программу, которая запрашивает у пользователя два числа (целое и дробное), выполняет над ними арифметические операции (сложение, вычитание, умножение, деление) и выводит результаты.

5. Логические типы и операторы

Напишите программу, которая запрашивает у пользователя два булевых значения (например, True или False), и выводит результат логических операций (AND, OR, NOT) над ними.

Домашнее задание: Условия и циклы

1. Оператор if, elif, else

Напишите программу, которая запрашивает у пользователя оценку (число от 1 до 100) и выводит соответствующий текст в зависимости от оценки (например, "Отлично", "Хорошо", "Удовлетворительно", "Неудовлетворительно").

2. Логические операторы

Создайте программу, которая запрашивает у пользователя два числа и выводит, является ли первое число больше второго, меньше или равно, используя логические операторы.

3. Цикл for

Напишите программу, которая выводит все четные числа от 1 до 50, используя цикл for.

4. Цикл while

Создайте программу, которая запрашивает у пользователя ввод чисел и суммирует их, пока пользователь не введет 0. В конце выведите сумму.

5. Операторы break и continue

Напишите программу, которая запрашивает у пользователя ввод чисел до тех пор, пока он не введет отрицательное число. Используйте оператор break для выхода из цикла.

Домашнее задание: Функции, строки и списки

1. Определение и вызов функций

Напишите функцию, которая принимает два числа и возвращает их произведение. Вызовите эту функцию и выведите результат.

2. Аргументы и параметры

Создайте функцию, которая принимает строку и число, а затем выводит строку заданное количество раз. Протестируйте функцию с различными аргументами.

3. Основные операции со строками

Напишите программу, которая запрашивает у пользователя строку и выводит её в верхнем регистре, а также количество символов в строке.

4. Создание и инициализация списков

Создайте список из 5 ваших любимых фруктов. Выведите его на экран и добавьте еще один фрукт в конец списка.

5. Методы списков

Напишите программу, которая создает список чисел, затем удаляет из него одно число по индексу, сортирует оставшиеся числа и выводит результат.

Примерные задания по проекту

Задание 1. Расширение отчетности

Текущую отчетность необходимо обогатить новыми метриками, пользуясь доступными данными.

- Придумай минимум две дополнительные метрики, которые можно рассчитать на имеющихся данных, и предоставь код на Python с вычислением данных метрик на имеющихся таблицах.

- Добавь эти метрики в текущую версию репортинга с телеграм-ботом.

В результате

- Предоставь `ipynb` файл с описанием и реализацией новых метрик.
- Обнови репортинг через телеграм таким образом, чтобы расчёт и визуализация новых метрик там появились.

Критерии

1. Описано (словами в маркдаунах) минимум две метрики, которые имеют отношение к бизнесу онлайн-кинотеатра (1 балл за каждую метрику, максимум 2)

2. Для данных метрик представлены примеры кода на Python для их расчёта в динамике (1 балл за каждую метрику, максимум 2)

2. Сравнительные значения данных метрик, а также графики рассылаются телеграм-ботом (1 балл)

Задание 2. Презентация результатов

Необходимо презентовать результаты вашей работы

- Сделай слайды `pptx` с демонстрацией результатов твоей работы, а также с демонстрацией

финансовых и продуктовых результатов онлайн-кинотеатра (не забудь использовать материалы из предыдущего задания).

- Сними видео (OBS/Loom/etc.) со скринкастом слайдов и устным объяснением содержимого на 5-7 минут.

В результате

- Предоставь слайды в формате `pptx`

- Предоставь видео с объяснением слайдов

Критерии

1. Длина видео незначительно отличается от установленного коридора 5-7 мин. Ваше лицо не обязано быть на видео, но с ним, конечно же, лучше. (1.5 балла)

2. Слайды оформлены с использованием визуализаций и должны объяснять выполненную работу и бизнес онлайн-кинотеатра. (1.5 балла)

3. Первое задание, наравне с другими метриками, учтено в презентации и в видео (0.5 балла)

Задание 3. Техническое задание для DWH

Общая рекомендация по сдаче задания

Необходимо обратиться к дата инженерам и запросить доп. данные

Придумай метрики, которые невозможно рассчитать на имеющихся таблицах, но которые были бы полезны для мониторинга

В результате

Предоставь текст (не более 2000 знаков) с описанием новых метрик и данных (витрин), которые нужны для их расчета

Критерии

1. В тексте указаны необходимые таблицы (0.75 балла)

- Набор колонок

- Тип данных

- Смысловое наполнение колонок

- Временной промежуток, который хранится в таблице

- Частоту обновления таблицы

2. В тексте названы и описаны метрики, которые будут строиться на новых данных, а также описан их бизнес-смысл и интерпретация (0.75 балла)