

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Центр сопровождения образовательных инициативных проектов

УТВЕРЖДЕНО:

Руководитель сетевой ОПОП
В.В. Кашпур

Рабочая программа дисциплины

Python для анализа данных

по направлению подготовки

09.04.03 Прикладная информатика

Направленность (профиль) подготовки:
«Дата-аналитика для бизнеса»

Форма обучения
Очная

Квалификация
Магистр

Год приема
2023

Томск – 2023

1. Цель и планируемые результаты освоения дисциплины (модуля)

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-2 Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач;

ОПК-3 Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями;

ОПК-5 Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем;

ПК-1 Способен управлять получением, хранением, передачей, обработкой больших данных.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-2.1 Владеет необходимыми методами алгоритмизации и программирования для решения профессиональных задач;

ИОПК-2.2 Знает современные подходы, методы применения современных интеллектуальных технологий для решения профессиональных задач;

ИОПК-2.3 Использует методы современных интеллектуальных технологий для решения профессиональных задач;

ИОПК-3.1 Осуществляет сбор, обработку и анализ научно-технической информации, необходимой для решения профессиональных задач;

ИОПК-3.2 Умеет работать с различными видами информации с помощью различных средств информационных и коммуникационных технологий;

ИОПК-5.1 Владеет современными инструментальными, технологическими и методическими средствами проектирования программно-аппаратного обеспечения информационных и автоматизированных систем;

ИОПК-5.2 Выбирает и использует методы проектирования информационных систем, необходимые для решения поставленных задач;

ИОПК-5.3 Использует современные информационно-коммуникационные технологии и программные средства на всех этапах жизненного цикла программных систем;

ИПК-1.1 Осуществляет мониторинг и оценку производительности обработки данных в организации, разработку предложений по повышению производительности обработки больших данных.

2. Задачи освоения дисциплины

- Получить представление о средах разработки и научиться работать с Jupyter Notebook;

- научиться работать с основными типами и структурами данных языка Python;

- научиться понимать логику и применять циклы при работе с разными типами данных;

- научиться использовать условные конструкции;

- научиться создавать пользовательские функции;

- освоить основы объектно-ориентированного программирования и научиться писать собственный класс;

- научиться использовать основные библиотеки, используемые при анализе данных;

- освоить основы git.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к Блоку 1 «Дисциплины (модули)».

Дисциплина относится к обязательной части образовательной программы.

4. Семестр освоения и форма промежуточной аттестации по дисциплине

Семестр 1, зачет с оценкой.

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются компетенции, сформированные в ходе освоения образовательных программ предшествующего уровня образования.

6. Язык реализации

Русский

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 4 з.е., 144 часа, из которых:

– практические занятия: 56 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Базовые понятия.

Python для аналитики данных. Документация. Среды для работы с Python. Jupyter Notebook. Visual Studio Code.

Тема 2. Переменные и типы данных.

Переменные: имена, значения, типы, преобразование типов. Числа: основные типы, арифметические операции. Строки: литералы, операции, индексы и срезы, методы строк, f-строки. Списки: создание, операции, индексы и срезы, методы, сортировка, списки и строки. Цикл for. Вложенный список: список списков. Словари: создание, обращение к элементам словаря, методы, операции со словарями. Список словарей: на пути к json и датафреймам. Кортежи. Множества. Файлы.

Тема 3. Условные конструкции и цикл while.

Логический тип данных, логические выражения и операторы сравнения. Условные конструкции: if-elif-else. Цикл while.

Тема 4. Функции.

Синтаксис пользовательской функции. Параметры и аргументы. Локальные и глобальные переменные. Анонимные функции.

Тема 5. Объектно-ориентированное программирование.

Инкапсуляция, полиморфизм, наследование. Объекты. Классы.

Тема 6. Основные библиотеки для анализа данных.

Краткое содержание темы. Работа с библиотеками и модулями. Получение данных: обзор библиотек os, sql alchemy, pandahouse, beautiful soup, selenium, scrapy. Работа с данными: pandas, re. Визуализация данных: обзор библиотек matplotlib, seaborn.

Тема 7. Git

Краткое содержание темы. Размещение проектов в репозитории. Работа с локальным репозиторием с помощью консольного Git.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости синхронных занятий, проведения тестов по пройденному материалу, выполнения практических домашних заданий и фиксируется в форме контрольной точки не менее одного раза в семестр.

Оценочные материалы текущего контроля размещены на сайте ТГУ в разделе «Информация об образовательной программе» – <https://www.tsu.ru/sveden/education/eduop/>.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Для получения **зачета с оценкой** необходимо выполнение трех самостоятельных практических работ (проектов):

Проект 1: Основы языка программирования Python (Темы 1-4)

Проект 2: Написание классов (Тема 5)

Проект 3: Работа с библиотеками (Темы 6-7)

Оценка проектов осуществляется по пятибалльной системе.

Итоговая оценка складывается из результатов текущего контроля по дисциплине из расчета:

выполнение текущих работ (тестов и практических домашних заданий)

от 80% и более – от 4,1 до 5,0 баллов,

79 – 60% - от 3,1 до 4,0 баллов,

59 – 50 % - от 2,9 до 3,0 баллов,

менее 50 % - до 2 баллов,

и результата выполнения самостоятельных практических работ (проектов).

Результаты зачета с оценкой определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно» и являются средним арифметическим от общего количества баллов за текущий контроль и оценок за выполнение трех практических проектов.

Оценка «отлично» - 4,6 – 5,0 баллов;

«хорошо» - 3,6 – 4,5 баллов;

«удовлетворительно» - 2,6 – 3,5 баллов

менее 2,5 баллов – «неудовлетворительно».

Оценочные материалы для проведения промежуточной аттестации размещены на сайте ТГУ в разделе «Информация об образовательной программе» – <https://www.tsu.ru/sveden/education/eduop/>.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в LMS «Data-Diving».

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине (<https://www.tsu.ru/sveden/education/eduop/>).

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

- Бизли Д. Python. Исчерпывающее руководство / Д. Бизли. – Санкт-Петербург: Изд-во Питер, 2023. –368 с.

- Бизли Д. Python. Книга рецептов / Бизли Д., Джонс Б. К. - Москва : ДМК Пресс, 2019. - 646 с. URL: <https://e.lanbook.com/img/cover/book/131723.jpg> (дата обращения: 01.06.2023). - Режим доступа: для зарегистрированных пользователей.
- Свейгарт Э. Автоматизация рутинных задач с помощью Python: практическое руководство для начинающих: [пер. с англ.] / Э. Свейгарт. [и др.] - Москва: Диалектика, 2019. - 584 с.
- Мэтиз Э. Изучаем Python: программирование игр, визуализация данных, веб-приложения / Э. Мэтиз. – Санкт-Петербург: Питер, 2017. – 491 с.

б) дополнительная литература:

- Чакон С. Pro Git / С. Чакон, Б. Страуб. – Санкт-Петербург: Питер, 2016. – 496 с.
- МакКинни Уэс. Python и анализ данных: Первичная обработка данных с применением pandas, NumPy и Jupiter. – Москва: ДМК Пресс, 2015. – 484 с.
- Абдрахманов М.И. Python. Визуализация данных. Matplotlib. Seaborn. Mayavi. – Devpractice Team, 2020. – 412 с.
- Даббас, Элиас. Интерактивные дашборды и приложения с Plotly и Dash. Используем полноценный веб-фреймворк в Python на всю мощь – без JavaScript. – Москва: ДМК Пресс, 2022. – 306 с.
- Мартелли А. Python. Справочник : полное описание языка / Алекс Мартелли, Анна Рейвенскрофт, Стив Холден ; перевод с английского А. Г. Гузикевича. - 3-е изд.. - Санкт-Петербург [и др.] : Диалектика, 2019. - 892 с.
- Форта Б. Изучаем регулярные выражения : пер. с англ. / Бен Форта. - Москва : Диалектика, 2019. - 188 с.
- Browning J. V. Pro Python Second Edition // by J. Burton Browning, Marty Alchin. // Springer eBooks. URL: <http://dx.doi.org/10.1007/978-1-4842-0334-7>

в) ресурсы сети Интернет:

- Основы Python Академия Яндекса <https://academy.yandex.ru/handbook/python>
- Real Python Tutorials <https://realpython.com/>

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

– Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office OneNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);

– публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).

– **Облачная среда для работы с кодом:**

Google Colab ([ссылка на официальные инструкции](#))

Требования: наличие браузера, подключения к сети Интернет, учетной записи google

ИЛИ

– **Стационарная среда для работы с кодом:**

Jupyter Lab в составе Anaconda ([ссылка на скачивание](#) среды Anaconda, [инструкция по установке](#) на Windows с официального сайта (на англ.яз., там же: инструкция для macOS и Linux), [перевод инструкции](#) на рус.яз.)

Требования: наличие браузера, подключение к сети Интернет

ОС: **Windows** не ниже 8 64-bit x86; **macOS** 10.13+ 64-bit x86 & M1; **Linux**, включая Ubuntu, RedHat, CentOS 7+ 64-bit x86, 64-bit aarch64 (AWS Graviton2), 64-bit

Power8/Power9, s390x (Linux on IBM Z & LinuxONE). (Для более старых версий операционных систем можно найти соответствующие установщики в архиве).

Место на диске: не менее 5 Гб для скачивания и установки (+ объем для установки всех необходимых библиотек)

Занимаемая оперативная память: от 3 Гб

б) информационные справочные системы:

– Электронный каталог Научной библиотеки ТГУ –
<http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>

– Электронная библиотека (репозиторий) ТГУ –
<http://vital.lib.tsu.ru/vital/access/manager/Index>

– ЭБС Лань – <http://e.lanbook.com/>

– ЭБС Консультант студента – <http://www.studentlibrary.ru/>

– Образовательная платформа Юрайт – <https://urait.ru/>

– ЭБС ZNANIUM.com – <https://znanium.com/>

– ЭБС IPRbooks – <http://www.iprbookshop.ru/>

14. Материально-техническое обеспечение

Занятия по учебной дисциплине проводятся с использованием дистанционных образовательных технологий. Каждый обучающийся обеспечен доступом к образовательной платформе <https://edu.data-diving.ru/>.

15. Информация о разработчиках

Никитина Татьяна Владимировна, ООО "Газпромнефть - Региональные продажи", ведущий специалист

Литовченко Дмитрий Евгеньевич, Яндекс Практикум, продуктовый аналитик, менеджер образовательных программ "Аналитик данных" / "Аналитик данных Плюс" / "Аналитик данных Bootcamp"