

МИНОБРНАУКИ РОССИИ
Федеральное государственное автономное образовательное
учреждение высшего образования "Пермский
государственный национальный исследовательский
университет"

Авторы-составители: **Гарафутдинов Роберт Викторович**
Радионова Марина Владимировна
Васёва Галина Сергеевна
Ильин Иван Вадимович

Рабочая программа дисциплины
PYTHON ДЛЯ АНАЛИЗА ДАННЫХ
Код УМК 101464

Утверждено
Протокол №8
от «17» апреля 2024 г.

Пермь, 2024

1. Наименование дисциплины

Python для анализа данных

2. Место дисциплины в структуре образовательной программы

Дисциплина входит в обязательную часть Блока « Б.1 » образовательной программы по направлениям подготовки (специальностям):

Направление подготовки: **38.03.05** Бизнес-информатика
направленность Бизнес-аналитика

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины **Python для анализа данных** у обучающегося должны быть сформированы следующие компетенции:

38.03.05 Бизнес-информатика (направленность : Бизнес-аналитика)

ОПК.2 Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения

Индикаторы

ОПК.2.1 Использует знания основных положений и концепций в области программирования, архитектуру языков программирования, основную терминологию и базовые алгоритмы, основные требования информационной безопасности для практического применения

ОПК.2.2 Анализирует типовые языки программирования, составляет программы

4. Объем и содержание дисциплины

Направление подготовки	38.03.05 Бизнес-информатика (направленность: Бизнес-аналитика)
форма обучения	очная
№№ триместров, выделенных для изучения дисциплины	7
Объем дисциплины (з.е.)	4
Объем дисциплины (ак.час.)	144
Контактная работа с преподавателем (ак.час.), в том числе:	56
Проведение лекционных занятий	28
Проведение лабораторных работ, занятий по иностранному языку	28
Самостоятельная работа (ак.час.)	88
Формы текущего контроля	Защищаемое контрольное мероприятие (2) Итоговое контрольное мероприятие (1)
Формы промежуточной аттестации	Экзамен (7 триместр)

5. Аннотированное описание содержания разделов и тем дисциплины

Python для анализа данных

Цель – изучение структур данных языка программирования Python и его библиотек для работы с большими массивами данных.

Тема 1. Введение в программирование. Знакомство с языком Python и инструментарием

Устройство компьютера. Алгоритм и программа. Высокоуровневый язык Python: архитектурные особенности, преимущества и недостатки, дзен Python. Версии Python. Интерактивный и пакетный режимы работы интерпретатора. Среды разработки на Python: IDLE, PyCharm, Google Collaboratory. Установка библиотек. Переменные. Комментарии. Операторы. Типы данных. Преобразование типов. Функции. Ввод-вывод. Математические операции и операции сравнения. Обработка исключений.

Тема 2. Основные алгоритмические структуры. Пользовательские функции

Условный оператор if. Операции булевой алгебры. Оператор цикла while. Описание функции. Аргументы функций: позиционные, именованные. Аргументы со значениями по умолчанию. Области видимости переменных. Документирование функций (докстринги).

Тема 3. Структуры данных. Индексированные и неиндексированные коллекции

Классы и объекты. Методы. Изменяемые и неизменяемые типы данных. Простое и глубокое копирование. Итерируемые объекты (контейнеры). Строка. Диапазон. Кортеж. Список. Оператор цикла for. Реализация двумерных массивов. Множества и словари.

Тема 4. Обработка последовательностей. Подробнее о функциях. Файловый ввод-вывод

Функции для обработки последовательностей (max, sum и др.). Сортировка. Параметр key. Нефиксированное количество аргументов функции. Локальные и глобальные переменные. Передача в функцию изменяемых и неизменяемых объектов. Анонимные (лямбда) функции. Функции filter и map. Механизмы list, set и dict comprehensions. Чтение и запись неструктурированных и структурированных (CSV, JSON) текстовых файлов. Сериализация объектов: модули pickle и json.

Тема 5. Библиотеки Python для обработки рядов и табличных данных

Библиотека NumPy. Многомерные массивы. Векторизованные (поэлементные) операции с массивами. Генерация данных из различных распределений. Операции линейной алгебры. Библиотека pandas. Структуры данных Series и DataFrame. Основные операции с ними. Библиотека Matplotlib. Основы построения диаграмм.

Контрольная точка 1

Лабораторная работа 1. Знакомство со средой разработки. Ввод-вывод в Python. Вычисление математических выражений.

Лабораторная работа 2. Логический тип данных и условный оператор.

Лабораторная работа 3. Оператор цикла while. Написание пользовательских функций.

Контрольная точка 2

Лабораторная работа 4. Итерируемые объекты (коллекции). Оператор цикла for. Сложные структуры данных, матрицы. Вложенные циклы.

Лабораторная работа 5. Обработка последовательностей. Продвинутое возможности пользовательских функций.

Лабораторная работа 6. Множества и словари. Файлы. Реализация сложного алгоритма.

Итоговое контрольное мероприятие

Лабораторная работа 7. Библиотека NumPy.

Лабораторная работа 8. Библиотека pandas. Тип Series.

Лабораторная работа 9. Библиотека pandas. Тип DataFrame.

Лабораторная работа 10. Визуализация данных с помощью библиотек Matplotlib и seaborn.

6. Методические указания для обучающихся по освоению дисциплины

Освоение дисциплины требует систематического изучения всех тем в той последовательности, в какой они указаны в рабочей программе.

Основными видами учебной работы являются аудиторные занятия. Их цель - расширить базовые знания обучающихся по осваиваемой дисциплине и систему теоретических ориентиров для последующего более глубокого освоения программного материала в ходе самостоятельной работы. Обучающемуся важно помнить, что контактная работа с преподавателем эффективно помогает ему овладеть программным материалом благодаря расстановке необходимых акцентов и удержанию внимания интонационными модуляциями голоса, а также подключением аудио-визуального механизма восприятия информации.

Самостоятельная работа преследует следующие цели:

- закрепление и совершенствование теоретических знаний, полученных на лекционных занятиях;
- формирование навыков подготовки текстовой составляющей информации учебного и научного назначения для размещения в различных информационных системах;
- совершенствование навыков поиска научных публикаций и образовательных ресурсов, размещенных в сети Интернет;
- самоконтроль освоения программного материала.

Обучающемуся необходимо помнить, что результаты самостоятельной работы контролируются преподавателем во время проведения мероприятий текущего контроля и учитываются при промежуточной аттестации.

Обучающимся с ОВЗ и инвалидов предоставляется возможность выбора форм проведения мероприятий текущего контроля, альтернативных формам, предусмотренным рабочей программой дисциплины. Предусматривается возможность увеличения в пределах 1 академического часа времени, отводимого на выполнение контрольных мероприятий.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации.

При проведении текущего контроля применяются оценочные средства, обеспечивающие передачу информации, от обучающегося к преподавателю, с учетом психофизиологических особенностей здоровья обучающихся.

7. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

При самостоятельной работе обучающимся следует использовать:

- конспекты лекций;
- литературу из перечня основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля);
- текст лекций на электронных носителях;
- ресурсы информационно-телекоммуникационной сети "Интернет", необходимые для освоения дисциплины;
- лицензионное и свободно распространяемое программное обеспечение из перечня информационных технологий, используемых при осуществлении образовательного процесса по дисциплине;
- методические указания для обучающихся по освоению дисциплины.

8. Перечень основной и дополнительной учебной литературы

Основная:

1. Гарафутдинов Р. В. Python для анализа данных: учебное пособие для студентов, обучающихся по направлениям подготовки бакалавров «Экономика», «Менеджмент», «Бизнес-информатика», «Торговое дело» / Р. В. Гарафутдинов. - Пермь: ПГНИУ, 2024, ISBN 978-5-7944-4096-6. - 276. - Библиогр.: с. 274-275
<https://elis.psu.ru/node/643488>

2. Маккинли, Уэс Python и анализ данных / Уэс Маккинли ; перевод А. Слинкина. — 2-е изд. — Саратов : Профобразование, 2019. — 482 с. — ISBN 978-5-4488-0046-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/88752.html>

Дополнительная:

1. Дональд, Р. Структуры данных в Python: начальный курс / Шихи Р. Дональд ; перевод А. В. Снастина. — Москва : ДМК Пресс, 2022. — 186 с. — ISBN 978-5-93700-110-8. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <https://www.iprbookshop.ru/125191.html>

2. Меликов, П. И. Изучаем основы Python. Практический курс для дата-аналитиков / П. И. Меликов. — Москва, Алматы : Ай Пи Ар Медиа, EDP Hub (Идипи Хаб), 2023. — 480 с. — ISBN 978-5-4497-2162-4, 978-601-81002-1-5. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <https://www.iprbookshop.ru/130920.html>

3. Протоद्याконов, А. В. Алгоритмы Data Science и их практическая реализация на Python : учебное пособие / А. В. Протоद्याконов, П. А. Пылов, В. Е. Садовников. — Москва, Вологда : Инфра-Инженерия, 2022. — 392 с. — ISBN 978-5-9729-1006-9. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/124000>

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

[//www.python.org/doc/](http://www.python.org/doc/) документация к языку Python

<https://numpy.org/doc/> документация к библиотеке NumPy

<https://pandas.pydata.org/docs/> документация к библиотеке pandas

<https://matplotlib.org/stable/index.html> документация к библиотеке Matplotlib

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

Образовательный процесс по дисциплине **Python для анализа данных** предполагает использование следующего программного обеспечения и информационных справочных систем:

- доступ в режиме on-line в Электронную библиотечную систему (ЭБС);
- доступ в электронную информационно-образовательную среду университета.
- IDE для Python версии 3 и выше.

При освоении материала и выполнения заданий по дисциплине рекомендуется использование материалов, размещенных в Личных кабинетах обучающихся ЕТИС ПГНИУ (student.psu.ru).

При организации дистанционной работы и проведении занятий в режиме онлайн могут использоваться:

система видеоконференцсвязи на основе платформы BigBlueButton (<https://bigbluebutton.org/>).

система LMS Moodle (<http://e-learn.psu.ru/>), которая поддерживает возможность использования текстовых материалов и презентаций, аудио- и видеоконтент, а так же тесты, проверяемые задания, задания для совместной работы.

система тестирования Indigo (<https://indigotech.ru/>).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Для лекционных занятий требуется аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением и маркерной доской.

Для проведения лабораторных занятий - компьютерный класс. Состав оборудования определен в Паспорте компьютерного класса (требуется доступ в Интернет и должна быть установлена IDE для Python).

Для групповых (индивидуальных) консультаций - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением и маркерной доской.

Для проведения текущего контроля - компьютерный класс. Состав оборудования определен в Паспорте компьютерного класса (требуется доступ в Интернет и должна быть установлена IDE для Python).

Самостоятельная работа студентов: аудитория, оснащенная компьютерной техникой с возможностью подключения к сети «Интернет», с обеспеченным доступом в электронную информационно-образовательную среду университета, помещения Научной библиотеки ПГНИУ.

Помещения научной библиотеки ПГНИУ для обеспечения самостоятельной работы обучающихся:

1. Научно-библиографический отдел, корп.1, ауд. 142. Оборудован 3 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

2. Читальный зал гуманитарной литературы, корп. 2, ауд. 418. Оборудован 7 персональными

компьютерами с доступом к локальной и глобальной компьютерным сетям.

3. Читальный зал естественной литературы, корп.6, ауд. 107а. Оборудован 5 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

4. Отдел иностранной литературы, корп.2 ауд. 207. Оборудован 1 персональным компьютером с доступом к локальной и глобальной компьютерным сетям.

5. Библиотека юридического факультета, корп.9, ауд. 4. Оборудована 11 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

6. Читальный зал географического факультета, корп.8, ауд. 419. Оборудован 6 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

Все компьютеры, установленные в помещениях научной библиотеки, оснащены следующим программным обеспечением:

Операционная система ALT Linux;

Офисный пакет Libreoffice.

Справочно-правовая система «КонсультантПлюс»

**Фонды оценочных средств для аттестации по дисциплине
Python для анализа данных**

**Планируемые результаты обучения по дисциплине для формирования компетенции.
Индикаторы и критерии их оценивания**

ОПК.2

Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
ОПК.2.2 Анализирует типовые языки программирования, составляет программы	Умеет создавать программы на Python и пользоваться библиотеками для анализа данных	Неудовлетворитель НЕУДОВЛЕТВОРИТЕЛЬНО выставляется студенту, который не знает Python и библиотеки для анализа данных. Заявленная часть компетенции в полной мере не сформирована. Имеющихся знаний, умений, навыков недостаточно для решения практических задач. Требуется повторное обучение. Удовлетворительн УДОВЛЕТВОРИТЕЛЬНО выставляется студенту, который не уверенно знает Python и библиотеки для анализа данных. Сформированы знания и умения в области ИИ, необходимые для дальнейшего обучения. Сформированность заявленной части компетенции соответствует минимальным требованиям. Хорошо ХОРОШО выставляется студенту, который в целом хорошо знает Python и библиотеки для анализа данных. Сформировано умение использовать соответствующие программные инструменты в стандартных ситуациях. Компетенция в целом соответствует требованиям. Отлично ОТЛИЧНО выставляется студенту, который отлично знает Python и библиотеки для анализа данных. Сформированность заявленной части компетенции превышает стандартные требования. Имеющихся знаний, умений, навыков и мотивации в полной мере достаточно для применения творческого подхода к решению сложных практических задач.

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
ОПК.2.1 Использует знания основных положений и концепций в области программирования, архитектуру языков программирования, основную терминологию и базовые алгоритмы, основные требования информационной безопасности для практического применения	Знает основную терминологию и базовые алгоритмы на Python	Неудовлетворител Выставляется студенту, который не знает базовые алгоритмические конструкции и структуры данных. Удовлетворительн Выставляется студенту, который не уверенно знает базовые алгоритмические конструкции и структуры данных. Хорошо Выставляется студенту, который в целом хорошо знает базовые алгоритмические конструкции и структуры данных. Сформировано умение использовать соответствующие программные инструменты в стандартных ситуациях. Компетенция в целом соответствует требованиям. Отлично Сформированы систематические знания и умения в области программирования. Имеющихся знаний, умений, навыков и мотивации в полной мере достаточно для применения творческого подхода к решению сложных практических задач.

Оценочные средства текущего контроля и промежуточной аттестации

Схема доставки : Базовая

Вид мероприятия промежуточной аттестации : Экзамен

Способ проведения мероприятия промежуточной аттестации : Оценка по дисциплине в рамках промежуточной аттестации определяется на основе баллов, набранных обучающимся на контрольных мероприятиях, проводимых в течение учебного периода.

Максимальное количество баллов : 100

Конвертация баллов в отметки

«отлично» - от 81 до 100

«хорошо» - от 61 до 80

«удовлетворительно» - от 42 до 60

«неудовлетворительно» / «незачтено» менее 42 балла

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ОПК.2.1 Использует знания основных положений и концепций в области программирования, архитектуру языков программирования, основную терминологию и базовые алгоритмы, основные требования информационной безопасности для практического применения	Контрольная точка 1 Защищаемое контрольное мероприятие	Умение установить и настроить среду выполнения программ на Python, запустить программу на выполнение, написать простейшую программу с вводом-выводом и выполнением математических операций. Умение реализовывать нелинейные алгоритмы на Python, в том числе с использованием пользовательских функций.
ОПК.2.2 Анализирует типовые языки программирования, составляет программы	Контрольная точка 2 Защищаемое контрольное мероприятие	Умение осмысленно применять различные структуры данных в программах на языке Python. Умение применять средства сложной обработки последовательностей в Python и работать с файлами.

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ОПК.2.2 Анализирует типовые языки программирования, составляет программы	Итоговое контрольное мероприятие Итоговое контрольное мероприятие	Знание специальные типов данных Python для хранения и обработки рядов и табличных данных. Умение использовать библиотеки NumPy, pandas, Matplotlib для загрузки, обработки и визуализации данных в рамках решения практических задач. Умение по содержательной постановке задачи выбрать необходимый набор инструментов, представляемых Python и его библиотеками, для решения поставленной задачи.

Спецификация мероприятий текущего контроля

Контрольная точка 1

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы самостоятельной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **30**

Проходной балл: **12.5**

Показатели оценивания	Баллы
Лабораторная работа 1. Знакомство со средой разработки. Ввод-вывод в Python. Вычисление математических выражений	10
Лабораторная работа 3. Оператор цикла while. Написание пользовательских функций	10
Лабораторная работа 2. Логический тип данных и условный оператор	10

Контрольная точка 2

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы самостоятельной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **30**

Проходной балл: **12.5**

Показатели оценивания	Баллы
Лабораторная работа 4. Итерируемые объекты (коллекции). Оператор цикла for. Сложные структуры данных, матрицы. Вложенные циклы	10
Лабораторная работа 6. Множества и словари. Файлы. Реализация сложного алгоритма	10
Лабораторная работа 5. Обработка последовательностей. Продвинутое возможности пользовательских функций	10

Итоговое контрольное мероприятие

Продолжительность проведения мероприятия промежуточной аттестации: **4 часа**

Условия проведения мероприятия: **в часы самостоятельной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **40**

Проходной балл: **17**

Показатели оценивания	Баллы
Лабораторная работа 7. Библиотека NumPy	10
Лабораторная работа 10. Визуализация данных с помощью библиотек Matplotlib и seaborn	10
Лабораторная работа 9. Библиотека pandas. Тип DataFrame	10
Лабораторная работа 8. Библиотека pandas. Тип Series	10