

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

Институт экономики и менеджмента

УТВЕРЖДАЮ:

Директор


Е. В. Нехода

« 23 » 05 20 24 г.

Рабочая программа дисциплины

Премодуль (Python и R для анализа данных)

по направлению подготовки

38.04.01 Экономика

Направленность (профиль) подготовки :

Анализ данных в экономике

Форма обучения

Очная

Квалификация

Магистр

Год приема

2024

СОГЛАСОВАНО:

Руководитель ОП

 Н.А. Скрыльникова

Председатель УМК

 М.В. Герман

Томск – 2024

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-1 Способен применять знания (на продвинутом уровне) фундаментальной экономической науки при решении практических и (или) исследовательских задач.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК 1.3 Обобщает и критически анализирует практические задачи.

2. Задачи освоения дисциплины

– Освоить основы языков программирования R и Python.

– Научиться выполнять операции по управлению данными с применением языков программирования R и Python.

– Научиться выполнять анализ данных в задачах поиска, привлечения, подбора и отбора персонала средствами языков R и Python.

3. Место дисциплины в структуре образовательной программы

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

Дисциплина относится к обязательной части образовательной программы.

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Первый семестр, зачет

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 3 з.е., 108 часов, из которых:

-лекции: 8 ч.

-практические занятия: 20 ч.

в том числе практическая подготовка: 20 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Основные понятия языков R и Python.

Синтаксис языков R и Python. Основные типы и структуры данных

Тема 2. Управление данными в R.

Операции импорта/экспорта, сортировки, фильтрации и агрегации данных; объединение данных.

Тема 3. Управление данными в Python.

Библиотеки numpy и pandas. Операции импорта/экспорта, сортировки, фильтрации и агрегации данных; объединение данных.

Тема 4. Графические возможности языков R и Python.

Графическая подсистема языка R. Команды построения графиков низкого и высокого уровня. Продвинутое графическое возможности языка R: пакеты ggplot и lattice. Команды для настройки и построения графиков в системе matplotlib.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости и выполнения домашних заданий и фиксируется в форме контрольной точки не менее одного раза в семестр.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет проводится в письменной форме по билетам, демонстрирующего освоение компетенции ОПК-11, индикатор достижения ИОПК-1.3 (Обобщает и критически анализирует практические задачи).

Билет содержит один теоретический вопрос. Продолжительность зачета 1,5 часа.

Примерный перечень теоретических вопросов:

1. Перечислите основные типы и структуры данных языка R.
2. Что такое атрибуты в языке R, каково их назначение? Перечислите основные атрибуты, их назначение и функции для работы с ними.
3. Перечислите основные отличия структур данных atomic vector, matrix, list и data frame языка R, назовите функции/методы, позволяющие преобразовать данные из одной структуры в другую.
4. Перечислите основные способы извлечения элементов из структур данных R, укажите их отличия и ограничения.
5. Перечислите основные инструменты импорта/экспорта данных в языке R.
6. Перечислите основные управляющие языковые конструкции языка R, приведите примеры их применения.
7. Перечислите основные команды пакета Graphics Base языка R, укажите их назначение и основные параметры.
8. Перечислите основные функции языка R для работы с данными (поиск и устранение пропущенных данных, поиск противоречий и дубликатов, извлечения подмножеств, фильтрации данных, соединения и агрегации данных).
9. Перечислите основные типы и структуры данных языка Python.
10. Перечислите способы определения функций на языке Python.
11. Перечислите основные свойства и методы класса ndarray библиотеки NumPy.
12. Перечислите методы библиотеки NumPy, реализующие файловый ввод/вывод, методы линейной алгебры, работы со случайными числами и вероятностными распределениями.
13. Перечислите основные структуры данных библиотеки pandas.
14. Перечислите основные возможности библиотеки pandas в решении задач, связанных с манипулированием данными (индексация, сортировка, фильтрация, редукция, обработка пропусков и противоречий).

Результат зачёта определяется оценками «зачтено» и «не зачтено».

Оценка «зачтено» выставляется при условии решения обучающимся задач для самостоятельной работы и развернутом ответе на теоретический вопрос.

Оценка «не зачтено» выставляется при условии отсутствия развернутого ответа на теоретический вопрос и/или отсутствии решений задач для самостоятельной работы.

11. Учебно-методическое обеспечение

- а) Электронный учебный курс по дисциплине в электронном университете «Moodle»
- б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

12. Перечень учебной литературы и ресурсов сети Интернет

- а) основная литература:

– 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.

– Бизли Д. Python. Книга рецептов / Бизли Д., Джонс Б. К.. - Москва : ДМК Пресс, 2019. - 646 с.

– Мاستицкий С. Э. Статистический анализ и визуализация данных с помощью R / Мастицкий С. Э., Шитиков В. К.. - Москва : ДМК Пресс, 2015. - 496 с.

б) дополнительная литература:

– Персиваль Г. Python. Разработка на основе тестирования. Повинуйся Билли-тестировщику, используя Django, Selenium и JavaScript / Персиваль Г.. - Москва : ДМК Пресс, 2018. - 622 с.

– Лучано Р. Python. К вершинам мастерства / Лучано Р.. - Москва : ДМК Пресс, 2016. - 768 с.

в) ресурсы сети Интернет:

– <http://www.machinelearning.ru/wiki/>

– <http://citforum.ru>

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

– Microsoft Office Standard 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.).

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

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения практических занятий, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

15. Информация о разработчиках

Богданов Александр Леонидович, к.т.н., доцент, ИЭМ ТГУ, доцент кафедры Информационных технологий и бизнес-аналитики.