

Документ подписан простой электронной подписью
 Информация о владельце:
 ФИО: Петренко Вячеслав Иванович
 Должность: и.о. директора Института цифрового развития
 Дата подписания: 01.06.2023 15:14:42
 Уникальный идентификатор:
 b47f96ee365a16e170f041bf7b0b10000e7ad4d

Аннотация дисциплины

Наименование дисциплины	Методы интеллектуального анализа данных
Содержание	Статистические основы анализа данных. Роль анализа данных в современном мире. Научные исследования. Построение системы анализа данных. Источники данных. Сбор данных. Подготовка данных. Генеральная совокупность и выборка. Jupyter Notebook: интерактивные вычисления и среда. Типы переменных. Количественные и номинативные переменные. Меры центральной тенденции. Меры изменчивости. Квартили распределения и график box-plot. Нормальное распределение. Центральная предельная теорема. Основы NumPy: массивы и векторные вычисления. Доверительные интервалы для среднего. Идея статистического вывода, р-уровень значимости. Работа в интерактивной оболочке IPython. HTML-блокнот в IPython (Jupyter Notebook). Пользовательский интерфейс нового поколения JupyterLab. Назначение пакета NumPy. Сравнение NumPy с MATLAB. Основы работы с Pandas. Массивы NumPy. Векторизация вычислений в NumPy. Введение в структуры данных Pandas. Объект Series. Объект DataFrame. Индексные объекты. Базовая функциональность Pandas. Переиндексация DataFrame. Удаление элементов из оси. Доступ по индексу, выборка и фильтрация. Переформатирование данных в Pandas. Арифметические операции и выравнивание данных. Применение функций и отображение. Сортировка и ранжирование. Индексы по осям с повторяющимися значениями. Обработка отсутствующих данных. Фильтрация отсутствующих данных. Замена отсутствующих данных. Иерархическое индексирование. Уровни переупорядочения и сортировки. Чтение и запись данных, форматы файлов, поддерживаемые Pandas. Сводная статистика по уровню. Работа со столбцами DataFrame. Комбинирование и слияние наборов данных. Слияние объектов DataFrame как в базах данных. Слияние по индексу. Конкатенация вдоль оси. Комбинирование перекрывающихся данных. Устранение дубликатов. Преобразование данных с помощью функции или отображения. Замена значений. Переименование индексов осей. Построение графиков и визуализация. Дискретизация и раскладывание. Обнаружение и фильтрация выбросов. Перестановки и случайная выборка. Вычисление индикаторных переменных. Чтение и запись данных в текстовом формате. Данные в формате JSON. Данные в формате XML. Двоичные форматы данных. Форматы данных BigData. Данные в формате Avro. Агрегирование данных и групповые операции Pandas. Данные в формате OCR. Данные в формате Parquet. Визуализация данных с помощью matplotlib. Встроенные средства визуализации pandas. Высокоуровневые средства визуализации seaborn. Средства интерактивной визуализации. Визуализация данных с помощью пакета Bokeh. Визуализация данных с помощью пакета Plotly. Визуализация данных с помощью пакета Altair Работа с временными рядами в Pandas. Механизм GroupBy. Выборка столбца или подмножества столбцов. Группировка с помощью словарей и объектов Series. Группировка с помощью функций. Группировка по уровням индекса. Агрегирование данных. Применение функций, зависящих от столбца, и нескольких функций. Возврат агрегированных данных в "неиндексированном" виде.

Результаты освоения дисциплины	Показывает способность применять информационно-коммуникационные технологии при поиске необходимой информации Способен применять современные принципы поиска, хранения, обработки, анализа и представления в требуемом формате информации Умеет решать конкретные задачи обработки данных с помощью современных средств автоматизации Обеспечивает информационную безопасность при использовании информационных технологий в профессиональной деятельности Целесообразно использует информационно-коммуникационные технологии при поиске необходимой информации Проектирует решение задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений Способен применять современные интерактивные программные комплексы для выполнения и редактирования текстов, изображений и чертежей Использует современные средства автоматизации разработки и выполнения конструкторской документации Способен решать конкретные задачи профессиональной деятельности с помощью современных программных средств подготовки конструкторско-технологической документации
Трудоемкость, з.е.	5 з.е.
Формы отчетности	Экзамен
Перечень основной и дополнительной литературы, необходимой для освоения дисциплины	
Основная литература	Горелов, В.И. Анализ статистических данных. Электронный ресурс : практикум / Т.Н. Ледашева / В.И. Горелов. - Анализ статистических данных, 2020-06-16. - Москва : Российская международная академия туризма, Университетская книга, 2015. - 120 с. - Книга находится в базовой версии ЭБС IPRbooks. - ISBN 978-5-98699-151-1 Крутиков, В. Н. Анализ данных : учебное пособие / В.Н. Крутиков, В.В. Мешечкин ; Министерство образования и науки Российской Федерации ; Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Кемеровский государственный университет». - Кемерово : Кемеровский государственный университет, 2014. - 138 с. : ил. - http://biblioclub.ru/. - Библиогр. в кн. - ISBN 978-5-8353-1770-7 Огерчук, А. А. Анализ и визуализация данных лабораторного эксперимента по физике. Электронный ресурс / Огерчук А. А. : учебное пособие для обучающихся по образовательным программам высшего образования по направлениям подготовки, входящих в состав укрупненных групп направлений подготовки 15.00.00 машиностроение, 24.00.00 авиационная и ракетно-космическая техника, 27.00.00 управление в технических системах. - Оренбург : ОГУ, 2018. - 119 с. - Рекомендовано ученым советом.
Дополнительная литература	Анализ данных качественных исследований : лабораторный практикум : направление подготовки 39.03.01 Социология; Профиль "Социальная структура, социальные институты и изменения". Бакалавриат / сост. А. П. Истомина ; ФГАОУ ВО Сев.-Кав. федер. ун-та. - Ставрополь : СКФУ, 2018. - 108 с. : ил. - Прил.: с. 100-107 Дятлов, А. В. Анализ данных в социологии. Учебник / А. В. Дятлов, Д. А. Гугуева. - Ростов-на-Дону, Таганрог : Издательство федерального университета, 2018. - 226 с. - Книга находится в премиум-версии ЭБС IPRbooks. - ISBN 978-5-9275-2690-1 Карпенко, О. М. Высшее образование в странах с развитой рыночной экономикой : анализ образовательной статистики и глобальных рейтингов в сфере образования / О. М. Карпенко, М. Д. Бершадская. - Высшее образование в странах мира.

[illegible]