

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ИНГУШСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»**

**ФИЗИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра «Информационные системы и технологии»**

СОГЛАСОВАНО

УТВЕРЖДАЮ

Руководитель образовательной программы

Декан физико-математического факультета

_____/М.Х. Мальсагов

_____/ М.Х. Мальсагов

от «27» апреля 2026г.

от «28» апреля 2026г

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.16 Технология анализа и визуализации данных

Направление подготовки

09.03.02 Информационные системы и технологии

Направленность (профиль подготовки)

Технологии искусственного интеллекта и анализа данных

Квалификация выпускника

Бакалавр

Форма обучения

Очная

Магас, 2026г.

1. Цели и задачи освоения дисциплины «Технология анализа и визуализации данных»

Цель дисциплины:

Изучение основ визуализации количественных и качественных данных, формирование умений применения основных средств визуализации данных различных типов.

Задачи дисциплины:

- формирование представлений о принципах и методах визуализации данных, а также структуре данных
- формирование умений визуализации данных при помощи основных компьютерных средств
- формирование умений подготовки презентации с использованием визуализации данных

Код и наименование профессионального стандарта	Обобщенные трудовые функции			Трудовые функции		
	Код	Наименование	Уровень квалификации	Наименование	Код	Уровень (подуровень) квалификации
06.015 Специалист по информационным системам.	С	Выполнение работ и управление работами по созданию (модификации) и сопровождению ИС, автоматизирующих задачи организационного управления и бизнес-процессы.	6	Определение первоначальных требований заказчика к ИС и возможности их реализации в ИС на этапе предконтрактных работ	С/01.6	6
				Документирование существующих бизнес-процессов организации заказчика (реверс-инжиниринг бизнес-процессов организации)	С/07.6	6
				Разработка модели бизнес-процессов заказчика	С/08.6	6
				Разработка архитектуры ИС	С/14.6	6
				Проектирование и дизайн ИС	С/16.6	6
				Разработка баз данных ИС	С/17.6	6

2. Место учебной дисциплины в структуре основной профессиональной образовательной программы

Дисциплина «Технология анализа и визуализации данных» относится к профессиональному циклу дисциплин. Для освоения данной дисциплины

необходимы знания, умения и компетенции, полученные обучающимися при изучении курса «Информатика», «Языки программирования». Дисциплина обеспечивает изучение дисциплин профессионального и специального циклов.

В результате освоения дисциплины обучающийся должен:

Знать:

- основные понятия, функции, состав и принципы работы систем анализа и визуализации данных;
- архитектуру и принципы работы инструментов и платформ для анализа и визуализации данных;
- особенности построения и функционирования различных систем анализа данных, включая базы данных и протоколы обмена данными;
- принципы обработки, анализа и визуализации данных.

Уметь:

- проектировать и реализовывать решения для анализа и визуализации данных;
- использовать языки программирования и специализированные инструменты для анализа данных (Python, R, SQL);
- производить настройку и конфигурирование платформ и инструментов для анализа и визуализации данных;
- интегрировать различные источники данных и обеспечивать их взаимодействие;
- использовать инструменты для моделирования и симуляции данных.

Владеть:

- основными методами и инструментами анализа и визуализации данных;
- навыками программирования и отладки программ для анализа данных;
- способами конфигурирования и настройки инструментов и платформ для анализа данных;

- методами диагностики и устранения ошибок в системах анализа и визуализации данных.

Процесс изучения дисциплины направлен на формирование следующих компетенций:

ОПК-2	ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности.	ОПК-2.1. Знать: современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности. ОПК-2.2. Уметь: выбирать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности. ОПК-2.3. Иметь навыки: применения современных информационных технологий и программных средств, в том числе отечественного производства, при решении задач профессиональной деятельности.
ОПК-4	ОПК-4. Способен участвовать в разработке технической документации, связанной с профессиональной деятельностью с использованием стандартов, норм и правил.	ОПК-4.1. Знать: основные стандарты оформления технической документации на различных стадиях жизненного цикла информационной системы. ОПК-4.2. Уметь: применять стандарты оформления технической документации на различных стадиях жизненного цикла информационной системы. ОПК-4.3. Иметь навыки: составления технической документации на различных этапах жизненного цикла информационной системы.
	ПК-6 Способен создавать и поддерживать системы искусственного интеллекта на основе нейросетевых моделей и методов.	ИД-1 ПК-6 Осуществляет оценку и выбор моделей искусственных нейронных сетей и инструментальных средств для решения поставленной задачи.

Структура и содержание дисциплины

«Технология анализа и визуализации данных»

Структура дисциплины (модуля) Общая трудоемкость дисциплины составляет 4 зачетных единиц, 144 часов.

	Всего	Порядковый номер семестра		
		5		
Общая трудоемкость дисциплины, в том числе:	144			
Курсовой проект (работа)				
Аудиторные занятия всего В том числе:				
Лекции	18	+		
Практические занятия, семинары	16	+		
Лабораторные работы				
Самостоятельная работа	110	+		
Вид итоговой аттестации:				
Зачет/дифф.зачет		+		
К.С.Р.				
Экзамен				
Общая трудоемкость дисциплины	144			

Распределение учебных часов по темам и видам учебных занятий

Наименование разделов и тем	Объем часов		
	Лекции	Практические занятия	Самостоят.
Введение в анализ и визуализацию данных	4	2	15

Сбор и предобработка данных	4	4	15
Методы анализа данных	4	2	20
Инструменты и технологии визуализации данных	2	2	20
Проектирование и реализация систем анализа и визуализации данных	2	2	20
Практическое применение анализа и визуализации данных	2	2	20
Итого	18	16	110

3. Содержание учебной дисциплины

1. Введение в анализ и визуализацию данных

Содержание темы:

- Основные понятия и цели анализа и визуализации данных.
- Важность данных в современных информационных системах.
- Примеры успешного применения анализа и визуализации данных в различных отраслях.

Формы и методы проведения занятий:

- Лекции: Обзор понятий и целей анализа и визуализации данных.
- Дискуссии: Обсуждение примеров и важности анализа данных.

Практические занятия:

- Ознакомление с основными инструментами для анализа и визуализации данных (Excel, Google Sheets).
- Выполнение простых задач по созданию графиков и диаграмм.

2. Сбор и предобработка данных

Содержание темы:

- Методы сбора данных из различных источников.
- Очистка данных: устранение дубликатов, заполнение пропусков.
- Нормализация и стандартизация данных.

Формы и методы проведения занятий:

- Лекции: Теория сбора и предобработки данных.

- Примеры и кейс-стади: Рассмотрение реальных случаев сбора и очистки данных.

Практические занятия:

- Работа с инструментами для сбора данных (Python, библиотеки pandas).
- Выполнение задач по очистке и нормализации данных.

3. Методы анализа данных

Содержание темы:

- Основные методы и алгоритмы анализа данных: статистические методы, машинное обучение.
- Выявление тенденций и закономерностей.
- Прогнозирование и моделирование на основе данных.

Формы и методы проведения занятий:

- Лекции: Обзор методов анализа данных.
- Примеры и кейс-стади: Рассмотрение примеров применения методов анализа данных.

Практические занятия:

- Работа с библиотеками анализа данных (Scikit-learn, NumPy).
- Выполнение задач по анализу данных с использованием различных методов.

4. Инструменты и технологии визуализации данных

Содержание темы:

- Обзор инструментов визуализации данных (Tableau, Power BI, D3.js).
- Принципы создания эффективных визуализаций.
- Типы графиков и диаграмм и их применение.

Формы и методы проведения занятий:

- Лекции: Обзор инструментов и принципов визуализации данных.
- Примеры и кейс-стади: Рассмотрение успешных визуализаций данных.

Практические занятия:

- Работа с инструментами визуализации данных (Tableau, Power BI).
- Создание различных типов графиков и диаграмм на реальных данных.

5. Проектирование и реализация систем анализа и визуализации данных

Содержание темы:

- Принципы проектирования систем анализа и визуализации данных.
- Архитектурные особенности таких систем.
- Интеграция различных компонентов и источников данных.

Формы и методы проведения занятий:

- Лекции: Теория проектирования систем анализа и визуализации данных.
- Примеры и кейс-стади: Рассмотрение реальных систем и их архитектуры.

Практические занятия:

- Разработка прототипа системы анализа и визуализации данных.
- Интеграция различных источников данных и создание единой системы.

6. Практическое применение анализа и визуализации данных

Содержание темы:

- Примеры успешного применения анализа и визуализации данных в бизнесе, науке и других областях.
- Обсуждение реальных кейсов и получение инсайтов.

Формы и методы проведения занятий:

- Лекции: Обзор реальных кейсов применения анализа данных.
- Дискуссии: Обсуждение примеров и выводов из реальных кейсов.

Практические занятия:

- Разработка собственных проектов по анализу и визуализации данных на основе реальных данных.
- Презентация результатов и выводов.

Экзаменационные вопросы по дисциплине «Технология анализа и визуализации данных»

1. Какие основные этапы включает в себя процесс анализа данных?

2. Опишите основные методы сбора данных из различных источников.
3. Какие существуют методы предобработки данных и почему они важны?
4. Что такое очистка данных и какие основные техники используются для очистки данных?
5. Какова роль нормализации и стандартизации данных в анализе данных?
6. Какие статистические методы применяются в анализе данных?
7. Объясните основные концепции машинного обучения и его роль в анализе данных.
8. Каковы основные этапы создания модели машинного обучения?
9. В чем различие между обучением с учителем и без учителя?
10. Что такое гиперпараметры модели и как они влияют на результаты машинного обучения?
11. Какие существуют инструменты для визуализации данных?
Приведите примеры.
12. Каковы основные принципы создания эффективных визуализаций данных?
13. Какие типы графиков и диаграмм используются для визуализации данных и в каких случаях?
14. Опишите процесс создания интерактивной визуализации данных.
15. Каковы основные функции и возможности таких инструментов, как Tableau и Power BI?
16. В чем заключаются основные различия между различными типами графиков (столбчатые, линейные, круговые)?
17. Какие существуют методы оценки качества данных перед их анализом?
18. Как интегрировать различные источники данных для их анализа и визуализации?

19. Какие существуют методы и инструменты для обработки больших данных (Big Data)?
20. Опишите процесс проектирования системы анализа и визуализации данных.
21. Какие архитектурные особенности характерны для систем анализа данных?
22. Каковы основные шаги в разработке прототипа системы анализа данных?
23. Объясните роль и использование библиотек pandas и NumPy в анализе данных.
24. Каковы особенности работы с временными рядами в анализе данных?
25. В чем заключается процесс моделирования данных?
26. Какие существуют методы оценки результатов анализа данных?
27. Как визуализация данных помогает в принятии управленческих решений?
28. Приведите примеры успешного применения анализа и визуализации данных в бизнесе.
29. Какие ключевые навыки необходимы для эффективного анализа и визуализации данных?
30. Какие новые тенденции и технологии появляются в области анализа и визуализации данных?

7. Перечень основной и дополнительной учебной литературы, Необходимой для освоения дисциплины

Основная литература

1. Березовская, Е.А. Работа с сервисом бизнес-аналитики Yandex DataLens: учеб. пособие / С.В. Крюков; Южный федер. ун-т; Е.А. Березовская. — Ростов-на-Дону : Изд-во ЮФУ, 2022. — 94 с.: ил. — ISBN 978-5-9275-4119-5. — URL: <https://lib.rucont.ru/efd/812194> (дата обращения: 16.02.2024)
2. Колоколов, А. Заставьте данные говорить: Как сделать бизнес-дашборд в Excel. Руководство по визуализации данных: практическое руководство / А. Колоколов. - Москва: Альпина ПРО, 2023. - 244 с. - ISBN 978-5-206-00079-5. - Текст: электронный. - URL: <https://znanium.ru/catalog/product/2140915> (дата обращения: 11.04.2024).

Дополнительная литература

1. Полковникова, Н. А. Анализ и визуализация данных в Microsoft Excel в примерах и задачах: практическое пособие / Н. А. Полковникова. - Москва; Вологда: Инфра-Инженерия, 2023. - 172 с. - ISBN 978-5-9729-1485-2. - Текст: электронный. - URL: <https://znanium.com/catalog/product/2092453> (дата обращения: 11.04.2024).

Перечень ресурсов информационно-телекоммуникационной сети

«Интернет»

1. Электронно-библиотечная система "ZNANIUM.COM"
2. Электронно-библиотечная система "РУКОНТ"
3. Open Academic Journals Index (OAJI). Профессиональная база данных - Режим доступа: <http://oaji.net/>
4. Президентская библиотека им. Б.Н.Ельцина (база данных различных профессиональных областей) - Режим доступа: <https://www.prilib.ru/>
5. Информационно-справочная система "Консультант Плюс" - Режим доступа: <http://www.consultant.ru/>

Перечень информационных технологий

Для проведения лекционных и лабораторных занятий рекомендуется использовать программное обеспечение: операционная система Windows 10, обслуживающие программы и среды разработки программ по выбору преподавателей.

Электронная поддержка дисциплины

При изучении дисциплины для проработки всех тем и выполнения заданий по всем темам студенты могут использовать различные учебно-методические материалы, размещаемые в электронном виде преподавателями на файловом ftp- сервере, в хранилище полнотекстовых материалов, а также в электронной образовательной среде, которая предполагает также возможность обмена информацией с преподавателем для подготовки заданий. Доступ студентов к студенческому файловому серверу, хранилищу полнотекстовых материалов, электронной образовательной среде осуществляется с использованием с использованием учетных записей студентов.

Сведения о переутверждении программы на очередной учебный год и регистрации изменений

Учебный год	Решение кафедры (№ протокола, дата)	Внесенные изменения	Подпись зав. кафедрой

Рабочая программа дисциплины Б1.В.16 «Технология анализа и визуализации данных» составлена в соответствии с требованиями ФГОСВО по направлению подготовки 09.03.02- «Информационные системы и технологии», утвержденного приказом Министерства образования и науки Российской Федерации от «19» сентября 2017 г. № 926(ред. от 8.02.2021).

Программу составили: ассистент кафедры «Информационные системы и технологии» Евлоев И.Т.

Программа одобрена на заседании кафедры «Информационные системы и технологии»

Протокол № 8 от «27» апреля 2026 года

Программа одобрена Учебно-методической комиссией физико-математического факультета

Протокол № 6 от «28» апреля 2026 года

**Сведения о переутверждении программы на очередной учебный год и
регистрации из- менений**

Учебный год	Решение кафед- ры (№ протокола, дата)	Внесенные изменения	Подпись зав. кафед- рой

Приложение

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.16 Технология анализа и визуализации данных

Направление подготовки

09.03.02 Информационные системы и технологии

Направленность (профиль подготовки)

Технологии искусственного интеллекта и анализа данных

Квалификация выпускника

Бакалавр

Форма обучения

Очная

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Перечень формируемых компетенций

Паспорт фонда оценочных средств

ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности.

ОПК-4. Способен участвовать в разработке технической документации, связанной с профессиональной деятельностью с использованием стандартов, норм и правил.

ПК-6. Способен создавать и поддерживать системы искусственного интеллекта на основе нейросетевых моделей и методов.

№ п/п	Контролируемые разделы (темы) дисциплины	Код Контролируемой компетенции (или ее части)	Наименование оценочного средства
1.	Тема 1: Введение в анализ и визуализацию данных Основные понятия и цели анализа и визуализации данных. • Важность данных в современных информационных системах. Примеры успешного применения анализа и визуализации данных в различных отраслях. .	ОПК-2, ОПК-4, ПК-6	Тест
2.	Тема 2: Сбор и предобработка данных Содержание темы: Методы сбора данных из различных источников. Очистка данных: устранение дубликатов, заполнение пропусков. Нормализация и стандартизация данных.	ОПК-2, ОПК-4, ПК-6	Тест
3.	Тема 3: 3. Методы анализа данных Содержание темы: Основные методы и алгоритмы анализа данных: статистические методы, машинное обучение. Выявление тенденций и закономерностей. Прогнозирование и моделирование на основе данных.	ОПК-2, ОПК-4, ПК-6	Коллоквиум

4.	Тема 4: . Инструменты и технологии визуализации данных Содержание темы: Обзор инструментов визуализации данных (Tableau, Power BI, D3.js). Принципы создания эффективных визуализаций. Типы графиков и диаграмм и их применение.	ОПК-2, ОПК-4, ПК-6	Коллоквиум
5.	Тема 5: Проектирование и реализация систем анализа и визуализации данных Содержание темы: Принципы проектирования систем анализа и визуализации данных. Архитектурные особенности таких систем. Интеграция различных компонентов и источников данных.	ОПК-2, ОПК-4, ПК-6	Тест, коллоквиум
6.	Тема 6 Практическое применение анализа и визуализации данных Содержание темы: Примеры успешного применения анализа и визуализации данных в бизнесе, науке и других областях. Обсуждение реальных кейсов и получение инсайтов.	ОПК-2, ОПК-4, ПК-6	Коллоквиум

Типовой тест промежуточной аттестации

- Какие этапы включает в себя процесс анализа данных?
 - а) Сбор данных
 - б) Предобработка данных
 - в) Анализ данных
 - г) Все вышеперечисленные +
- Какие методы предобработки данных существуют?
 - а) Очистка данных
 - б) Нормализация данных
 - в) Стандартизация данных
 - г) Все вышеперечисленные +
- Какие основные типы графиков используются для визуализации данных?
 - а) Столбчатые
 - б) Линейные
 - в) Круговые
 - г) Все вышеперечисленные +
- Какие основные методы анализа данных существуют?
 - а) Статистические методы
 - б) Машинное обучение

- в) Временные ряды
 - г) Все вышеперечисленные +
5. Что такое машинное обучение?
- а) Процесс обучения компьютерных систем на основе данных и алгоритмов, позволяющий им выполнять задачи без явного программирования
 - б) Процесс создания и обработки больших объемов данных
 - в) Процесс создания моделей для визуализации данных
 - г) Ответ "а" +
6. Какие инструменты используются для визуализации данных?
- а) Tableau
 - б) Power BI
 - в) D3.js
 - г) Все вышеперечисленные +
7. Какой инструмент широко используется для анализа данных на языке Python?
- а) NumPy
 - б) pandas
 - в) Matplotlib
 - г) Все вышеперечисленные +
8. Что такое предобработка данных?
- а) Процесс преобразования данных в удобный для анализа формат
 - б) Процесс создания данных
 - в) Процесс визуализации данных
 - г) Ответ "а" +
9. Какие методы используются для анализа временных рядов?
- а) ARIMA
 - б) LSTM
 - в) Prophet
 - г) Все вышеперечисленные +
10. Какие основные этапы проектирования системы анализа данных?
- а) Архитектурное проектирование
 - б) Интеграция данных
 - в) Разработка прототипа
 - г) Все вышеперечисленные +
11. Какие алгоритмы применяются в статистическом анализе данных?
- а) t-тест
 - б) ANOVA
 - в) Кластерный анализ
 - г) Все вышеперечисленные +
12. Какие методы оценки качества модели машинного обучения существуют?
- а) Точность
 - б) Полнота
 - в) F1-мера
 - г) Все вышеперечисленные +
13. Какие этапы включает в себя процесс визуализации данных?
- а) Подготовка данных

- б) Создание визуализации
- в) Интерпретация результатов
- г) Все вышеперечисленные +

14. Какова роль визуализации данных в анализе данных?

- а) Помогает обнаруживать тенденции и паттерны в данных
- б) Облегчает интерпретацию результатов
- в) Способствует принятию управленческих решений
- г) Все вышеперечисленные +

15. Какие основные типы данных используются в анализе данных?

- а) Количественные
- б) Категориальные
- в) Временные ряды
- г) Все вышеперечисленные +

16. Что такое графовые базы данных?

- а) Тип базы данных, использующий графовые структуры для хранения и обработки данных
- б) Инструмент для визуализации данных
- в) Технология машинного обучения
- г) Ответ "а" +

17. Какие основные этапы включает в себя процесс создания модели машинного обучения?

- а) Подготовка данных
- б) Выбор алгоритма
- в) Обучение модели
- г) Все вышеперечисленные +

18. Какие методы используются для интерпретации результатов анализа данных?

- а) Построение выводов и рекомендаций на основе анализа данных
- б) Визуализация результатов
- в) Обратная связь от заказчика
- г) Все вышеперечисленные +

19. Какие ключевые навыки необходимы для успешного анализа данных?

- а) Знание языков программирования (Python, R)

б) Умение работать с данными и базами данных

- в) Навыки визуализации данных
- г) Все вышеперечисленные +

20. Что такое регрессионный анализ?

- а) Статистический метод анализа данных, используемый для изучения отношений между независимыми и зависимой переменными
- б) Метод предобработки данных
- в) Алгоритм машинного обучения
- г) Ответ "а" +

21. Какие методы используются для прогнозирования временных рядов?

- а) ARIMA
- б) LSTM
- в) Простые скользящие средние
- г) Все вышеперечисленные +

22. Какой инструмент широко используется для создания интерактивных визуализаций данных?

- а) D3.js
- б) Matplotlib
- в) Seaborn
- г) Ответ "а" +

23. Какие основные шаги включает в себя анализ данных?

- а) Загрузка данных
- б) Предобработка данных
- в) Моделирование данных
- г) Все вышеперечисленные +

24. Какие методы используются для кластерного анализа данных?

- а) K-means
- б) DBSCAN
- в) Hierarchical clustering
- г) Все вышеперечисленные +

25. Какие основные этапы включает в себя процесс подготовки данных?

- а) Очистка данных
- б) Нормализация данных
- в) Разделение данных на обучающую и тестовую выборки
- г) Все вышеперечисленные +

26. Какие методы используются для обработки пропущенных значений в данных?

- а) Удаление строк или столбцов с пропущенными значениями
- б) Заполнение пропущенных значений средним или медианным значением
- в) Использование алгоритмов машинного обучения для заполнения пропущенных значений
- г) Все вышеперечисленные +

27. Какой инструмент широко используется для анализа данных на языке R?

- а) ggplot2
- б) matplotlib
- в) seaborn
- г) Ответ "а" +

28. Какие алгоритмы машинного обучения используются для классификации данных?

- а) Логистическая регрессия
- б) Деревья решений
- в) Метод опорных векторов (SVM)
- г) Все вышеперечисленные +

29. Какой метод используется для оценки важности признаков в модели машинного обучения?

- а) Анализ компонент
- б) Перестановочная важность
- в) Отбор признаков с помощью регуляризации
- г) Ответ "б" +

30. Какие существуют методы оценки качества модели машинного обучения?

- а) Точность
- б) Полнота
- в) F1-мера
- г) Все вышеперечисленные +

Типовой вариант задания на контрольную работу

1. Объясните основные этапы анализа данных и опишите их последовательность.

Практические задания (70 баллов)

Ответьте на следующие вопросы и выполните задания:

1. Сбор и предобработка данных

- Загрузите набор данных "iris.csv" и проведите его предварительный анализ.
- Определите количество пропущенных значений в каждом столбце набора данных и заполните их медианным значением.

2. Анализ данных

- Постройте гистограмму распределения длины чашелистика для каждого вида ириса.
- Вычислите среднее, медиану и стандартное отклонение показателей "длина чашелистика" и "ширина чашелистика" для каждого вида ириса.

3. Визуализация данных

- а) Создайте scatter plot, отображающий зависимость длины чашелистика от ширины чашелистика для каждого вида ириса.
- б) Постройте box plot для каждого показателя ("длина чашелистика", "ширина чашелистика", "длина лепестка", "ширина лепестка") и сравните их распределения для каждого вида ириса.

Экзаменационные вопросы по дисциплине

1. Какие основные этапы включает в себя процесс анализа данных?
2. Опишите основные методы сбора данных из различных источников.
3. Какие существуют методы предобработки данных и почему они важны?
4. Что такое очистка данных и какие основные техники используются для очистки данных?
5. Какова роль нормализации и стандартизации данных в анализе данных?
6. Какие статистические методы применяются в анализе данных?
7. Объясните основные концепции машинного обучения и его роль в анализе данных.
8. Каковы основные этапы создания модели машинного обучения?
9. В чем различие между обучением с учителем и без учителя?
10. Что такое гиперпараметры модели и как они влияют на результаты машинного обучения?
11. Какие существуют инструменты для визуализации данных? Приведите примеры.
12. Каковы основные принципы создания эффективных визуализаций данных?
13. Какие типы графиков и диаграмм используются для визуализации данных и в каких случаях?
14. Опишите процесс создания интерактивной визуализации данных.
15. Каковы основные функции и возможности таких инструментов, как Tableau и Power BI?
16. В чем заключаются основные различия между различными типами графиков (столбчатые, линейные, круговые)?
17. Какие существуют методы оценки качества данных перед их анализом?
18. Как интегрировать различные источники данных для их анализа и визуализации?
19. Какие существуют методы и инструменты для обработки больших данных (Big Data)?
20. Опишите процесс проектирования системы анализа и визуализации данных.
21. Какие архитектурные особенности характерны для систем анализа данных?

22. Каковы основные шаги в разработке прототипа системы анализа данных?
23. Объясните роль и использование библиотек pandas и NumPy в анализе данных.
24. Каковы особенности работы с временными рядами в анализе данных?
25. В чем заключается процесс моделирования данных?
26. Какие существуют методы оценки результатов анализа данных?
27. Как визуализация данных помогает в принятии управленческих решений?
28. Приведите примеры успешного применения анализа и визуализации данных в бизнесе.
29. Какие ключевые навыки необходимы для эффективного анализа и визуализации данных?
30. Какие новые тенденции и технологии появляются в области анализа и визуализации данных?

Типовая практическая работа по дисциплине

Практическая работа: Исследование данных о потреблении электроэнергии

Цель работы: Проанализировать данные о потреблении электроэнергии и визуализировать результаты анализа.

Исходные данные: Набор данных "electricity_consumption.csv", содержащий информацию о потреблении электроэнергии в различных регионах за определенный период времени.

Задачи:

1. Загрузите набор данных и проведите его предварительный анализ.
2. Определите общий объем потребления электроэнергии за весь период.
3. Вычислите средний объем потребления электроэнергии в каждом регионе за весь период.
4. Постройте график изменения потребления электроэнергии по времени для каждого региона.
5. Определите месяцы с наибольшим и наименьшим объемом потребления электроэнергии.
6. Постройте круговую диаграмму, отражающую долю каждого региона в общем объеме потребления электроэнергии.

Ход работы:

1. **Предварительный анализ данных**
 - Загрузите набор данных "electricity_consumption.csv".
 - Изучите структуру данных: количество строк и столбцов, типы данных, наличие пропущенных значений.
 - Выведите первые несколько строк набора данных для ознакомления.
2. **Общий объем потребления электроэнергии**
 - Просуммируйте значения потребления электроэнергии по всем регионам за весь период.
 - Выведите полученный результат.

3. **Средний объем потребления электроэнергии**
 - Вычислите средний объем потребления электроэнергии для каждого региона.
 - Выведите полученные значения с указанием соответствующих регионов.
4. **График изменения потребления электроэнергии по времени**
 - Постройте график изменения потребления электроэнергии для каждого региона по времени.
 - Добавьте подписи к осям и легенду для каждого региона.
5. **Месяцы с наибольшим и наименьшим объемом потребления электроэнергии**
 - Определите месяцы с наибольшим и наименьшим объемом потребления электроэнергии.
 - Выведите полученные результаты.
6. **Круговая диаграмма доли каждого региона в общем объеме потребления электроэнергии**
 - Постройте круговую диаграмму, отражающую долю каждого региона в общем объеме потребления электроэнергии.
 - Добавьте подписи к секторам диаграммы и процентное значение для каждого региона.

Ожидаемые результаты:

- Загруженный и предобработанный набор данных.
- Общий объем потребления электроэнергии.
- Средний объем потребления электроэнергии для каждого региона.
- График изменения потребления электроэнергии по времени для каждого региона.
- Месяцы с наибольшим и наименьшим объемом потребления электроэнергии.
- Круговая диаграмма доли каждого региона в общем объеме потребления электроэнергии.