

**Федеральное государственное образовательное бюджетное учреждение
высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ
РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)**

Тульский филиал Финуниверситета

Кафедра «Математика и информатика»

СОГЛАСОВАНО

ЗАО «ЛИМ»

Генеральный директор

В.В. Куликов

«19» декабря 2024 г.



УТВЕРЖДАЮ

Директор филиала

Г.В. Кузнецов

«24» декабря 2024 г.



Н.О. Козлова, Е.В. Манохин

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
ПО ДИСЦИПЛИНЕ**

ОСНОВЫ АНАЛИЗА И ВИЗУАЛИЗАЦИИ ДАННЫХ

для студентов, обучающихся по направлению подготовки
09.03.02 Информационные системы и технологии,
образовательная программа «Информационные технологии в экономике»,
профиль «Информационные технологии в управлении цифровой
экономикой»

Рекомендовано Ученым советом Тульского филиала
(протокол от 24 декабря 2024 г. № 22)

Одобрено кафедрой «Математика и информатика»
(протокол от 02 декабря 2024 г. № 5)

Тула 2024

«Основы анализа и визуализации данных»

КОД И НАИМЕНОВАНИЕ ОПОП:

09.03.02 Информационные системы и технологии

Образовательная программа - Информационные технологии в экономике

Профиль «Информационные технологии в управлении цифровой экономикой»

Общее количество заданий по дисциплине

Код компетенции	Наименование компетенции	Количество заданий
ПКП-3	Способен проводить исследования на всех этапах жизненного цикла программных систем	40
Всего:		40

Количество заданий по дисциплине в одном варианте

Код компетенции	Наименование компетенции	Количество заданий
ПКП-3	Способен проводить исследования на всех этапах жизненного цикла программных систем	20
Всего:		20

Типы заданий

Наименование типа заданий	Номер типа заданий
Задание комбинированного типа с выбором одного верного ответа из предложенных	1
Задание открытого типа	2
Задание открытого типа с развернутым ответом	3

Распределение заданий по компетенциям

Код компетенции	Наименование компетенции	Наименование индикаторов сформированности компетенции	Семестр	Номер задания
ПКП-3	Способен проводить исследования на всех этапах жизненного цикла программных систем	1. Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	6	1.1, 1.2, 2.1, 2.2, 3.1, 3.2, 4.1, 4.2, 5.1, 5.2, 6.1, 6.2, 7.1, 7.2, 8.1, 8.2, 9.1, 9.2, 10.1, 10.2
		2. Организует сбор и подготовку данных для систем машинного обучения.	6	11.1, 11.2, 12.1, 12.2, 13.1, 13.2, 14.1, 14.2, 15.1, 15.2, 16.1, 16.2, 17.1, 17.2, 18.1, 18.2, 19.1, 19.2, 20.1, 20.2

Типы, уровень сложности и время выполнения заданий

Код компетенции	Индикатор сформированности компетенции	Номер задания	Тип задания	Уровень сложности задания	Время выполнения задания (мин.)
ПКП-3	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	1.1, 1.2.	1	базовый	2
	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	2.1, 2.2	1	базовый	2
	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	3.1, 3.2	1	базовый	2
	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	4.1, 4.2	1	базовый	2
	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	5.1, 5.2	1	базовый	2
	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	6.1, 6.2.	2	повышенный	4
	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	7.1, 7.2	2	повышенный	4
	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	8.1, 8.2	2	повышенный	4
	Демонстрирует знания основных методов машинного обучения и интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.	9.1, 9.2	3	высокий	8
	Демонстрирует знания основных методов машинного обучения и	10.1, 10.2	3	высокий	8

	интеллектуального анализа данных, применяет готовые инструменты для создания интеллектуальных алгоритмов.				
	Организует сбор и подготовку данных для систем машинного обучения.	11.1, 11.2.	1	базовый	2
	Организует сбор и подготовку данных для систем машинного обучения.	12.1, 12.2	1	базовый	2
	Организует сбор и подготовку данных для систем машинного обучения.	13.1, 13.2	1	базовый	2
	Организует сбор и подготовку данных для систем машинного обучения.	14.1, 14.2	1	базовый	2
	Организует сбор и подготовку данных для систем машинного обучения.	15.1, 15.2	1	базовый	2
	Организует сбор и подготовку данных для систем машинного обучения.	16.1, 16.2.	2	повышенный	4
	Организует сбор и подготовку данных для систем машинного обучения.	17.1, 17.2	2	повышенный	4
	Организует сбор и подготовку данных для систем машинного обучения.	18.1, 18.2	3	высокий	8
	Организует сбор и подготовку данных для систем машинного обучения.	19.1, 19.2	3	высокий	8
	Организует сбор и подготовку данных для систем машинного обучения.	20.1, 20.2	3	высокий	8

Задания, позволяющие осуществлять оценку компетенций, установленных для дисциплины

Вариант 1

Задание 1.1.

Выберите правильное определение анализа данных:

- А) Процесс сбора, очистки, преобразования, анализа и интерпретации данных для получения ценной информации.
- Б) Процесс сбора и обработки данных для создания баз данных.
- В) Процесс визуализации данных с помощью диаграмм и графиков.
- Г) Процесс создания прогнозных моделей для предсказания будущих событий.

Ответ:

Задание 2.1.

Что такое визуализация данных?

- А) Создание моделей данных для предсказания будущих событий.
- Б) Процесс преобразования данных в визуальные представления для облегчения понимания и анализа.
- С) Процесс сбора и очистки данных.
- Д) Процесс создания баз данных.

Ответ:

Задание 3.1.

Что такое корреляция между двумя переменными?

- А) Зависимость между двумя переменными, при которой изменение одной переменной приводит к изменению другой.
- Б) Случайное совпадение между двумя переменными.
- В) Отсутствие связи между двумя переменными.
- Г) Разница между двумя переменными.

Ответ:

Задание 4.1.

Что такое outlier (выброс) в данных?

- А) Значение данных, которое значительно отличается от других значений в наборе данных.
- Б) Среднее значение данных в наборе данных.
- В) Значение данных, которое наиболее часто встречается в наборе данных.
- Г) Максимальное значение данных в наборе данных.

Ответ:

Задание 5.1.

Какая из следующих диаграмм лучше всего подходит для отображения тренда временных данных:

- А) Гистограмма
- Б) Диаграмма рассеяния
- В) Круговая диаграмма
- Г) Линейный график

Ответ:

Задание 6.1.

Какой из следующих типов графиков лучше всего подходит для отображения распределения данных?

Ответ:

Задание 7.1.

С помощью какой диаграммы осуществляют визуализацию рассеивания данных через квартиль?

Ответ:

Задание 8.1.

Какая библиотека в Python наиболее часто используется для анализа данных?

Ответ:

Задание 9.1.

Три примера целей создания дашборда для линейных менеджеров.

Ответ:

Задание 10.1.

Охарактеризуйте Google Analytics как инструмент бизнес-аналитики.

Ответ:

Задание 11.1.

Какая из следующих задач относится к data mining:

- А) Создание веб-сайта для онлайн-магазина.
- Б) Разработка приложения для мобильного телефона.
- В) Исследование влияния маркетинговой кампании на продажи.
- Г) Проектирование архитектуры базы данных.

Ответ:

Задание 12.1.

Что такое big data?

- А) Маленький набор данных.
- Б) Набор данных, который слишком велик для обработки традиционными методами.
- В) Набор данных, который содержит информацию о пользователях.
- Г) Набор данных, который содержит только текстовую информацию.

Ответ:

Задание 13.1.

Что такое Machine Learning (машинное обучение)?

- А) Процесс обучения компьютеров на основе данных для автоматизации задач.
- Б) Процесс разработки программного обеспечения.
- В) Процесс сбора и очистки данных.
- Г) Процесс визуализации данных.

Ответ:

Задание 14.1.

Что такое глубокое обучение (Deep Learning)?

- А) Подвид машинного обучения, который использует многослойные нейронные сети.
- Б) Процесс обучения компьютеров на основе данных для автоматизации задач.
- В) Процесс сбора и очистки данных.
- Г) Процесс визуализации данных.

Ответ:

Задание 15.1.

Что такое искусственный интеллект (AI)?

- А) Процесс обучения компьютеров на основе данных для автоматизации задач.
- Б) Системы, которые имитируют когнитивные функции человека, такие как обучение, решение проблем и принятие решений.
- В) Процесс сбора и очистки данных.
- Г) Процесс визуализации данных.

Ответ:

Задание 16.1.

Какая диаграмма используется для выявления связи между двумя количественными переменными?

Ответ:

Задание 17.1.

Что такое "целевая переменная" в модели?

Ответ:

Задание 18.1.

Приведите не менее трех ключевых принципов работы с дашбордами.

Ответ:

Задание 19.1.

Охарактеризуйте аналитический дашборд.

Ответ:

Задание 20.1.

Раскройте структуру интерфейса дашборда.

Ответ:

Вариант 2**Задание 1.2.**

Какая из следующих задач относится к анализу данных:

- А) Создание веб-сайта для онлайн-магазина.
- Б) Разработка приложения для мобильного телефона.
- В) Исследование влияния маркетинговой кампании на продажи.
- Г) Проектирование архитектуры базы данных.

Ответ:

Задание 2.2.

Какая из следующих диаграмм лучше всего подходит для сравнения долей категорий в наборе данных:

- А) Гистограмма
- Б) Диаграмма рассеяния
- В) Круговая диаграмма
- Г) Линейный график

Ответ:

Задание 3.2.

Какой метод используется для выявления зависимости между двумя переменными:

- А) Регрессионный анализ
- Б) Кластерный анализ

- В) Анализ главных компонент
- Ш) Метод опорных векторов

Ответ:

Задание 4.2.

Какая из следующих задач относится к визуализации данных:

- А) Создание модели для предсказания цен на акции.
- Б) Разработка алгоритма для классификации изображений.
- В) Создание интерактивной карты с данными о населении городов.
- Г) Проектирование базы данных для хранения данных о клиентах.

Ответ:

Задание 5.2.

Что такое data mining?

- А) Процесс сбора и очистки данных.
- Б) Процесс визуализации данных с помощью диаграмм и графиков.
- В) Процесс поиска скрытых закономерностей и знаний в больших объемах данных.
- Г) Процесс создания прогнозных моделей для предсказания будущих событий.

Ответ:

Задание 6.2.

Какой метод визуализации данных можно использовать для сравнения нескольких объектов?

Ответ:

Задание 7.2.

Что означает "нормализация" данных?

Ответ:

Задание 8.2.

Что такое "тренд" в анализе временных рядов?

Ответ:

Задание 9.2.

Сформулируйте, в чем заключается правило простоты построения дашборда.

Ответ:

Задание 10.2.

Охарактеризуйте сервис Яндекс.Метрика как инструмент аналитики.

Ответ:

Задание 11.2.

Какая из следующих диаграмм лучше всего подходит для представления иерархических данных:

- А) Гистограмма
- Б) Диаграмма рассеяния
- В) Дендрограмма
- Г) Линейный график

Ответ:

Задание 12.2.

Какая из следующих технологий используется для обработки big data:

- А) Microsoft Word
- Б) Apache Hadoop
- В) Google Docs
- Г) Microsoft Excel

Ответ:

Задание 13.2.

Какая из следующих задач относится к машинному обучению:

- А) Разработка алгоритма для классификации изображений.
- Б) Разработка алгоритма для распознавания речи,
- В) Исследование влияния маркетинговой кампании на продажи.
- Г) Разработка алгоритма для игры в шахматы

Ответ:

Задание 14.2.

Какая из следующих задач относится к глубокому обучению:

- А) Разработка алгоритма для классификации изображений.
- Б) Разработка алгоритма для распознавания речи,
- В) Исследование влияния маркетинговой кампании на продажи.
- Г) Разработка алгоритма для игры в шахматы

Ответ:

Задание 15.2.

Какая из следующих задач относится к искусственному интеллекту:

- А) Разработка алгоритма для классификации изображений.
- Б) Разработка алгоритма для распознавания речи,
- В) Исследование влияния маркетинговой кампании на продажи.
- Г) Разработка алгоритма для игры в шахматы

Ответ:

Задание 16.2.

Очистка и нормализация данных относятся к методу ...

Ответ:

Задание 17.2.

Какая диаграмма больше всего подходит для отражения данных, имеющих три переменные характеристики?

Ответ:

Задание 18.2.

Сравните операционный и стратегический дашборды.

Ответ:

Задание 19.2.

Приведите не менее трех ключевых преимуществ дашбордов как инструментов визуализации данных.

Ответ:

Задание 20.2.

Раскройте структуру дашборда (из каких частей он состоит).

Ответ:

Ключи к оцениванию заданий

Вариант 1

№ задания	Верный ответ	Критерии оценивания
1.1	А	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
2.1	Б	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
3.1	А	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
4.1	А	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
5.1	Г	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
6.1	Гистограмма	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
7.1	«Ящик с усами»	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
8.1	Библиотека Pandas.	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
9.1	Возможные варианты ответов. 1. Отслеживать ежедневные, еженедельные или ежемесячные объемы продаж и сравнивать их с предыдущими периодами. 2. Контролировать уровни запасов. 3. Анализировать эффективность рекламных кампаний. 4. Следить за доходами и расходами. 5. Оценивать выполнение планов по отделам или отдельным сотрудникам.	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
10.1	Google Analytics — это бесплатный инструмент для анализа веб-трафика, предоставляющий подробную статистику о посетителях сайтов. Он активно используется для отслеживания таких ключевых метрик как Acquisition (привлечение трафика), Behavior (поведение пользователей) и Conversions (конверсии). Кроме того, сервис анализирует скорость загрузки страниц, эффективность рекламных кампаний, частоту и продолжительность посещений, а также распространение контента через репосты.	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
11.1	В	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
12.1	Б	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
13.1	А	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие

14.1	А	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
15.1	Б	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
16.1	scatter plot (точечная диаграмма)	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
17.1	"Целевая переменная" – это переменная, значение которой пытаются определить (предсказать).	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
18.1	1. Дашборд должен быть сфокусирован на основных показателях, важных для его функциональности. Избегайте перегрузки ненужной информацией. 2. Предоставлять пользователям возможность взаимодействовать с данными, например, изменять периоды времени или применять фильтры. 3. Обеспечить интуитивно понятный и удобный в использовании дизайн дашборда, чтобы пользователи могли без труда находить необходимые данные. Избегайте излишней визуальной перегрузки. 4. Использовать типы диаграмм и графиков, которые лучше всего передают нужные данные. 5. Данные на дашборде должны постоянно обновляться, чтобы всегда предоставлять актуальную информацию.	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
19.1	Аналитический дашборд помогает ознакомиться с тенденциями и сделать выводы. Как правило они создаются для конкретного бизнес-подразделения. Аналитики работают с ними, чтобы документировать отклонения в данных и отслеживать их причины. Например, это может быть дашборд, показывающий изменение количества пользователей определенного продукта в течение недели. Он может отображать средний биллинг, часть клиентов, показатели товарооборота у конкурентов.	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
20.1	Структура интерфейса дашборда 1. Заголовок (или хедер) - это название. Он может содержать наименование отчета или компании, логотип и т.д. 2. Модули - это элементы информационной панели отчета: ○ график - отображает вертикальное изменение показателей;	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие

	○ диаграмма - представлена в нескольких видах, отображает изменения различных категорий показателей. ○ индикатор - отслеживает численное увеличение или уменьшение основного показателя; ○ карта - показывает, где сосредоточены наиболее главные параметры, причем важность области зависит от размера или цвета точки; ○ таблица - двухмерный формат отображения данных. Названия и показатели представлены в строках и столбцах. 3. Футер (нижний колонтитул) - так называемый подвал, где расшифровываются выводы и ключевые значения.	
--	--	--

Вариант 2

№ задания	Верный ответ	Критерии оценивания
1.2	В	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
2.2	В	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
3.2	А	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
4.2	В	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
5.2	В	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
6.2	Линейчатый график	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
7.2	Приведение данных к общему масштабу	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
8.2	Долгосрочное направление или паттерн изменений	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
9.2	Дашборд должен быть простым и ясным: лучше меньше, да лучше. Основные показатели стоит разместить на главном экране, а более детальную информацию - в раскрывающихся элементах.	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
10.2	Яндекс.Метрика — это бесплатный инструмент аналитики, который позволяет оценивать посещаемость сайта и анализировать поведение пользователей. Сервис предоставляет отчеты по различным показателям, карты кликов и скроллинга, отслеживает звонки пользователей, а также обеспечивает функции, такие как тепловые карты, вебвизор, отслеживание рекламных кампаний и продвинутую отчетность с возможностями сегментации и сводной статистики.	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
11.2	В	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
12.2	Б	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
13.2	А	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
14.2	Б	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
15.2	Г	1 балл – полное совпадение с верным ответом 0 баллов – неверный ответ или его отсутствие
16.2	Очистка и нормализация данных относятся к методу предобработки данных	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
17.2	Пузырьковая диаграмма	2 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
18.2	Операционный дашборд отображает изменения в бизнес-данных. Примером корпоративного дашборда является график Яндекс-Метрики, с помощью которого	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие

	можно увидеть, как изменилась посещаемость сайта и что на нее повлияло. В конечном счете, пользователь сам выбирает, какой период времени ему смотреть на графике. Стратегический дашборд этакий взгляд с высоты птичьего полета на общую ситуацию и отдельные показатели, помогающий выявить проблемы и устранить их. Например, дашборд о лояльности сотрудников позволит увидеть, насколько лояльны работники компании, а также отследить изменения в различных группах.	
19.2	1. Все на одном экране. Важные данные отображаются в одном визуальном пространстве, не нужно листать презентации или переключать страницы в интерфейсе. Все находится прямо перед вами. 2. Динамика. Вся важная информация, например, текущие данные о доставке и сезонных продажах, содержатся на интерактивной доске. 3. Оперативность. Эксперты не упускают из виду, что в определенной категории продажи идут медленно или что пользователи часто сообщают о проблемах в работе системы.	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие
20.2	Дашборд состоит из трех частей: «граббер» осуществляет сбор данных из привязанных баз; «парсер» занимается обработкой и сортировкой данных; «пользовательский интерфейс» выводит на экран итоговые визуализации, отчеты и статистику.	3 балла – полное совпадение с верным ответом 1 балл – правильный ответ, без обоснования 0 баллов – неверный ответ или его отсутствие