

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Тольяттинский государственный университет»

Б1.В.ДВ.01.01
(индекс дисциплины)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Анализ данных

(наименование дисциплины)

по направлению подготовки
01.04.02 Прикладная математика и информатика

направленность (профиль)
Математическое моделирование

Форма обучения: очная

Год набора: 2022

Общая трудоемкость: 4 ЗЕ

Распределение часов дисциплины по семестрам

Семестр	2	Итого
Форма контроля	Экзамен	
Вид занятий		
Лекции	8	8
Лабораторные		
Практические	16	16
Руководство: курсовые работы (проекты) / РГР		
Промежуточная аттестация	0,35	0,35
Контактная работа	24,35	24,35
Самостоятельная работа	84	84
Контроль	35,65	35,65
Итого	144	144

Рабочую программу составил: профессор, доцент, д.техн.наук, Мкртычев С.В.

(должность, ученое звание, степень, Фамилия И.О.)

Рецензирование рабочей программы дисциплины:



Отсутствует



Рецензент

(должность, ученое звание, степень, Фамилия И.О.)

Рабочая программа дисциплины составлена на основании ФГОС ВО и учебного плана
направления подготовки

01.04.02 Прикладная математика и информатика

Срок действия рабочей программы дисциплины до «31» августа 2024 г.

УТВЕРЖДЕНО

На заседании кафедры «Прикладная математика и информатика»

(протокол заседания № 2 от «15» сентября 2021 г.).

1. Цель освоения дисциплины

Цель освоения дисциплины – формирование у студентов знаний о современных методах и технологиях анализа данных, а также навыков их применения на практике.

2. Место дисциплины в структуре ОПОП ВО

Дисциплины и практики, на освоении которых базируется данная дисциплина: Современные проблемы прикладной математики и информатики.

Дисциплины и практики, для которых освоение данной дисциплины необходимо как предшествующее: Математические и эвристические методы распознавания образов для анализа данных.

3. Планируемые результаты обучения

Формируемые и контролируемые компетенции (код и наименование)	Индикаторы достижения компетенций (код и наименование)	Планируемые результаты обучения
ПК-5 Способен управлять проектами в области ИТ в условиях неопределенностей с применением формальных инструментов моделирования и управления проблемами проекта	ПК-5.1 Знает основные принципы управления проектами в области ИТ в условиях неопределенностей с применением формальных инструментов моделирования и управления проблемами проекта	Знать: методы управления проектами в области ИТ в условиях неопределенностей Уметь: применять формальные инструменты моделирования и управления проблемами проекта Владеть: способами управления проектами в области ИТ в условиях неопределенностей
	ПК-5.2 Умеет грамотно использовать принципы управления проектами в области ИТ в условиях неопределенностей с применением формальных инструментов моделирования и управления проблемами проекта	Знать: принципы управления проектами с применением формальных инструментов моделирования и управления проблемами проекта Уметь: управлять проектами в области ИТ в условиях неопределенностей Владеть: : навыками управления проектами с применением формальных инструментов моделирования

Формируемые и контролируемые компетенции (код и наименование)	Индикаторы достижения компетенций (код и наименование)	Планируемые результаты обучения
	ПК-5.3 Владеет навыками управления проектами в области ИТ в условиях неопределенностей с применением формальных инструментов моделирования и управления проблемами проекта	Знать: принципы управления проектами в области ИТ в условиях неопределенностей Уметь: управлять проектами с применением формальных инструментов моделирования Владеть: навыками управления проектами в области ИТ в условиях неопределенностей с применением формальных инструментов моделирования и управления проблемами проекта

4. Структура и содержание дисциплины

Модуль (раздел)	Вид учебной работы	Наименование тем занятий (учебной работы)	Семестр	Объем, ч.	Баллы	Интерактив, ч.	Формы текущего контроля (наименование оценочного средства)
Модуль 1. Технологии анализа данных	Лек	Тема 1. Технология OLAP	2	2		-	-
	Лек	Тема 2. Технология Data mining	2	2		-	-
	Пр	1. Консолидация данных в MS Excel. 2. Создание сводной таблицы в MS Excel 3. Создание OLAP-куба в MS SQL Server. 4. Использование MS Query для работы с OLAP-кубами. 5. Кластеризация данных в MS Excel	2	10		-	Отчеты по практическим работам №№1-5
	Ср	Подготовка к практическим занятиям по темам модуля 1	2	42			-
Модуль 2. Программные средства для анализа данных	Лек	Тема 3. Применение табличного процессора Excel для анализа данных	2	2		-	-
	Лек	Тема 4. Архитектура и функциональность BI-систем	2	2		-	-
	Пр	6. Анализ данных «что-если» в MS Excel. 7. Использование надстройки «Поиск решения» MS Excel. 8. Анализ данных с помощью MS Power BI	2	6		-	Отчеты по практическим работам №№6-8
	Ср	Подготовка к практическим занятиям по темам модуля 2	2	42			-
	ПА		2	0,35		-	-
	Контроль		2	35,65		-	
Итого:				0			

5. Образовательные технологии

В рамках учебного курса предусмотрены следующие образовательные технологии:

- технология традиционного обучения: лекции и практические работы, самостоятельная работа;
- технология проектного обучения: реализация и защита отчетов по практическим работам.

6. Методические указания по освоению дисциплины

6.1. Рекомендации по подготовке к лекционным занятиям

Изучение дисциплины требует систематического и последовательного накопления знаний, следовательно, пропуски отдельных тем не позволяют глубоко освоить предмет.

В ходе лекционных следует обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации. Задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

Студент может дополнить список использованной литературы современными источниками, не представленными в списке рекомендованной литературы, и в дальнейшем использовать собственные подготовленные учебные материалы при написании курсовых и выпускных квалификационных работ.

6.2. Рекомендации по подготовке к практическим занятиям

Студентам следует доводить каждую практическую работу до окончательного решения, демонстрировать понимание проведенных расчетов (анализов, ситуаций), в случае затруднений обращаться к преподавателю.

Для того чтобы практические занятия приносили максимальную пользу, необходимо помнить, что упражнение и решение задач проводятся по рассмотренному на лекциях материалу и связаны, как правило, с детальным разбором отдельных вопросов лекционного курса. Следует подчеркнуть, что только после усвоения лекционного материала с определенной точки зрения (а именно с той, с которой он излагается на лекциях) он будет закрепляться студентом на практических занятиях как в результате обсуждения и анализа лекционного материала, так и с помощью решения проблемных ситуаций, задач. При этих условиях студент не только хорошо усвоит материал, но и научится применять его на практике, а также получит дополнительный стимул (и это очень важно) для активной проработки лекции.

По результатам выполнения работы составляется отчет, который при необходимости нужно сопровождать комментариями, схемами, чертежами и рисунками.

Следует помнить, что выполнение каждой работы должно доводиться до окончательного логического ответа, которого требует условие, и по возможности с выводом. Полученный ответ следует проверить способами, вытекающими из существа данной задачи. Полезно также (если возможно) решать несколькими способами и сравнить полученные результаты. Решение задач данного типа нужно продолжать до приобретения твердых навыков в их решении.

6.3. Рекомендации по подготовке к экзамену

Подготовка к экзамену способствует закреплению, углублению и обобщению знаний, получаемых, в процессе обучения, а также применению их к решению практических задач. Готовясь к экзамену, студент ликвидирует имеющиеся пробелы в знаниях, углубляет, систематизирует и упорядочивает свои знания. На экзамене студент демонстрирует то, что он приобрел в процессе обучения по конкретной учебной дисциплине.

Необходимо ориентировать студентов на систематическую подготовку к занятиям в течение семестра, что позволит использовать время экзаменационной сессии для систематизации знаний.

7. Оценочные средства

7.1. Паспорт оценочных средств

Семестр	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
2	ПК-5	<i>Отчеты по практическим работам № № 1-8</i>

7.2. Типовые задания или иные материалы, необходимые для текущего контроля

7.2.1. Отчеты по практическим работам

(наименование оценочного средства)

Типовые примеры заданий

Практическая работа 2. Создание сводной таблицы в MS Excel.

Форма отчета по практической работе № 2. В отчет по практической работе должны быть включены следующие пункты:

- титульный лист;
- цель работы;
- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

Практическая работа 6. Анализ данных «что-если» в MS Excel.

Форма отчета по практической работе № 6. В отчет по практической работе должны быть включены следующие пункты:

- титульный лист;
- цель работы;
- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

Практическая работа 8. Анализ данных с помощью MS Power BI

Форма отчета по практической работе №8. В отчет по практической работе должны быть включены следующие пункты:

- титульный лист;
- цель работы;
- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

Оценка выполненной работы проводится по следующим критериям:

1. Наличие всей существенной информации по работе
2. Точность и полнота предоставляемых сведений
3. Непротиворечивость приводимой информации
4. Правильность интерпретаций и выводов, которые сделаны по результатам работы
5. Степень достижения студентом поставленной цели
6. Обоснованность применяемого решения
7. Грамотность (содержательная) используемых формулировок

Критерии оценки за отчеты по практическим работам:

– оценка «зачтено» ставится студенту, который продемонстрировал результаты выполнения практической работы, соответствующие поставленным задачам, и предоставил отчет, оформленный должным образом и содержащий краткое описание полученных результатов;

– оценка «не зачтено» ставится студенту, который не продемонстрировал результаты выполнения практической работы или не представил по ней отчет или представленный отчет не соответствует требованиям по оформлению.

Типовые задания для оценки сформированности компетенций

ПК-5 Способен управлять проектами в области ИТ в условиях неопределенностей с применением формальных инструментов моделирования и управления проблемами проекта
код и наименование компетенции

ОМ закрытого типа

Задание 1

Выберите один правильный вариант ответ.

Что такое аналитическая информация?

- а) совокупность знаний и сведений об организационной структуре предприятия, организации;
- б) совокупность знаний и сведений о менеджменте предприятия, организации;
- в) совокупность знаний и сведений, характеризующих деятельность предприятия, организации;
- г) все ответы правильные

Правильный ответ: в

Задание 2

Выберите один правильный вариант ответ.

В зависимости от среды аналитическая информация подразделяется на:

- а) простую и сложную;
- б) внутреннюю и внешнюю;
- в) отечественную и зарубежную;
- г) достоверную и недостоверную

Правильные ответы: б.

Задание 3

Выберите один правильный вариант ответ.

Количественная информация обычно представляется?

- а) в виде вербального описания;
- б) в графической форме;
- в) в табличной форме;
- г) в виде программного кода

Правильные ответы: в.

Задание 4

Выберите один правильный вариант ответ.

Данные в информационных системах выступают как

- а) способ интерпретации информации о внешнем мире в компьютерных системах;
- б) способ представления информации в определенной, фиксированной форме, пригодной для обработки, хранения и передачи;
- в) способ обработки информации о внешнем мире, пригодный для использования в компьютерных системах;
- г) все ответы правильные

Правильные ответы: б.

Задание 5

Выберите один правильный вариант ответ.

Что означает аббревиатура OLAP?

- а) Оперативная аналитическая обработка данных;
- б) Оперативная транзакционная обработка данных;
- в) Операционная система;
- г) Операционная оболочка

Правильные ответы: а.

Задание 6

Выберите один правильный вариант ответ.

Какой критерий определяет полноту, точность, своевременность и возможность интерпретации данных?

- а) количество данных;
- б) качество данных;
- в) достаточность данных;
- г) достоверность данных

Правильные ответы: б.

Задание 7

Выберите правильные ответы.

Как иначе называют данные низкого качества?

- а) грязные данные;
- б) сырые данные;
- в) примитивные данные;
- г) неструктурированные данные

Правильный ответ: а

Задание 8

Выберите один правильный вариант ответа.

Какая из представленных надстроек ТП Excel является средством анализа данных?

- а) Мастер суммирования;
- б) Поиск решения;
- в) Мастер подстановок;
- г) все ответы правильные

Правильный ответ: б.

Задание 9

Выберите правильные ответы.

К способам хранения многомерных данных относятся:

- а) MOLAP;
- б) ROLAP;
- в) HOLAP;
- г) OLTP

Правильные ответы: а,б,в.

Задание 10

Выберите один правильный вариант ответа.

Какая из представленных опций Excel используется для решения простых задач анализа "что-если"?

- а) Консолидация данных;
- б) Поиск решения;
- в) Подбор параметров;
- г) Фильтрация данных

Правильный ответ: в.

ОМ открытого типа

Задание 11

Дайте развернутый ответ.

Что такое консолидация данных?

Правильный ответ: это способ получения итоговой информации, при котором данные, расположенные в нескольких различных областях, объединяются в соответствии с выбранной функцией обработки.

Задание 12

Дайте развернутый ответ.

Что такое агрегация данных?

Правильный ответ: это тип процесса извлечения данных и информации, при котором данные ищутся, собираются и представляются в обобщенном формате на основе отчетов для достижения конкретных бизнес-целей или процессов и/или проведения анализа.

Задание 13

Дайте развернутый ответ.

Что представляет собой аналитический отчет.

Правильный ответ:

Аналитический отчет представляет собой комбинацию:

- метода получения информации;
- базы данных, из которой эта информация собирается (текущая или внешняя база данных);
- вида отчета (визуального представления полученной информации).

Задание 14

Дайте развернутый ответ.

Что такое сводная таблица?

Правильный ответ: это интерактивная таблица, создаваемая пользователем на основе данных, содержащихся в списке, и отражающая обобщенные результаты обработки этих данных. Сводные таблицы используются для целенаправленного обобщения информации, представления взаимосвязей между категориями, изменения способа группировки данных, суммирования различных категорий данных и т.д.. Пример – сводная таблица Excel.

Задание 15

Дайте развернутый ответ.

Что такое OLAP?

Правильный ответ: это технология обработки данных, заключающаяся в подготовке суммарной (агрегированной) информации на основе больших массивов данных, структурированных по многомерному принципу. Реализации технологии OLAP являются компонентами программных решений класса BI.

Задание 16

Дайте развернутый ответ.

Что такое визуализация данных?

Правильный ответ: это процесс использования визуальных элементов, таких как диаграммы, графики или карты, для представления данных. Он переводит сложные, масштабные или числовые данные в визуальное представление, которое более удобно для поддержки принятия решений.

Задание 17

Дайте развернутый ответ.

Что такое хранилище данных?

Правильный ответ: это предметно-ориентированная, интегрированная, неизменяемая и поддерживающая хронологию электронная коллекция данных для обеспечения процесса принятия решений.

Задание 18

Дайте развернутый ответ.

Что такое витрина данных?

Правильный ответ: это массив узконаправленной информации, ориентированный на пользователей одной рабочей группы или департамента.

Задание 19

Дайте развернутый ответ.

Что такое ETL?

Правильный ответ: это общий термин для процессов «извлечение, преобразование, загрузка», которые происходят, когда данные переносят из нескольких систем в одно хранилище.

Задание 20

Дайте развернутый ответ.

Что представляет собой анализ «что-если» в Excel?

Правильный ответ: это процесс изменения значений в ячейках, который позволяет увидеть, как эти изменения влияют на результаты формул на листе. В Excel предлагаются средства анализа "что-если" трех типов: сценарии, таблицы данных и подбор параметров

7.3. Оценочные средства для промежуточной аттестации по итогам освоения дисциплины

7.3.1. Вопросы к промежуточной аттестации

Семестр 2

№ п/п	Вопросы к экзамену
1	Понятие аналитической информации
2	Современные технологии анализа данных
3	Концепция управления качеством информации
4	Уровни качества данных, оценка пригодности данных к анализу
5	Методы предобработки данных
6	Инструментарий предобработки данных в аналитическом приложении
7	Понятие ETL
8	Классификация программных продуктов для создания аналитических решений
9	Консолидация данных
10	Характеристики аналитических платформ
11	Обобщенная схема процесса консолидации данных
12	Понятие массива данных
13	Понятие хранилища данных
14	Детализированные и агрегированные данные, метаданные
15	Многомерное представление данных и многомерный куб
16	Понятие OLAP-системы
17	Требования, предъявляемые к OLAP-системам
18	Задачи и содержание оперативного анализа данных
19	Архитектура MOLAP
20	Архитектура ROLAP
21	Архитектура HOLAP
22	Концепция виртуальных хранилищ данных
23	Проблемы обработки больших объемов данных
24	Концепция управления качеством информации
25	Уровни качества данных, оценка пригодности данных к анализу
26	Предобработка данных и ее отличие от очистки
27	Типичный набор инструментов предобработки в аналитическом приложении
28	Фильтрация данных
29	Устранение дубликатов и противоречий данных
30	Трансформация данных. Цели и методы
31	Методы преобразования структур данных
32	Преобразование даты и времени
33	Группировка данных
34	Интеллектуальный анализ данных (Data mining)
35	Общие визуализаторы: графики, диаграммы, гистограммы, инфографика
36	Метод классификации
37	Метод кластеризации
38	Метод ассоциации
39	Использование MS Excel для обработки данных
40	Базовая терминология анализа данных, понятие модели и моделирования
41	Использование MS Excel для анализа данных
42	Методы сравнения данных в MS Excel

№ п/п	Вопросы к экзамену
43	Архитектура и функциональность современных аналитических ИС
44	Понятие BI-системы
45	Настройка «Поиск решения» MS Excel
46	Пакет анализа данных в MS Excel
47	Методы консолидации данных в MS Excel
48	Опция подбора параметра MS Excel
49	Анализ данных «что-если» в MS Excel
50	Создание сводной таблицы в MS Excel

7.3.2. Критерии и нормы оценки

Семестр	Форма проведения промежуточной аттестации	Критерии и нормы оценки	
2	Экзамен (устно)	«отлично»	Ставится студенту на экзамене, если он исчерпывающе и грамотно дал ответы на вопросы экзаменационного билета или при ответе допустил небольшую неточность на 1 вопрос, но при этом смог грамотно ответить на дополнительные вопросы ту, проявившему полные знания в рамках требований подготовки по дисциплине, усвоившему литературу, рекомендуемую программой и показавшему систематический характер знаний. В изложении материала и ответах на дополнительные вопросы допускаются небольшие неточности
		«хорошо»	Ставится студенту на экзамене, если он исчерпывающе и грамотно дал ответ на 1 вопрос экзаменационного билета, а на другой только тезисные высказывания или допустил небольшие неточности при ответе на вопросы экзаменационного билета и дал краткие ответы на дополнительные вопросы
		«удовлетворительно»	Ставится студенту на экзамене, если он не смог дать ответ на один из вопросов экзаменационного билета или

Семестр	Форма проведения промежуточной аттестации	Критерии и нормы оценки	
			ответил на все вопросы, но при этом ответы содержали только тезисные высказывания
		«неудовлетворительно»	Ставится студенту на экзамене, если он не дал ответ на вопросы экзаменационного билета или в ответе содержались фундаментальные ошибки

8. Учебно-методическое и информационное обеспечение дисциплины

8.1. Обязательная литература

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
1	Полубояров В. В.	Использование MS SQL Server Analysis Services 2008 для построения хранилищ данных	Курс лекций	2021	ЭБС "IPRbooks"
2	Чубукова И.А.	Data Mining	Учебное пособие	2020	ЭБС "IPRbooks"

8.2. Дополнительная литература

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
1	Башмакова Е. И.	Информатика и информационные технологии. Умный Excel 2016: библиотека функций	Учебное пособие	2020	ЭБС "IPRbooks"
2	Перфильев Д. А.	Информационно-аналитические технологии и системы	Учебное пособие	2013	ЭБС "IPRbooks"

8.3. Перечень профессиональных баз данных и информационных справочных систем

– Web of Science [Электронный ресурс] : мультидисциплинарная реферативная база данных. – Philadelphia: ClarivateAnalytics, 2016– . – Режим доступа : apps.webofknowledge.com. – Загл. с экрана. – Яз. рус., англ.

Scopus [Электронный ресурс] : реферативная база данных. – Netherlands: Elsevier, 2004– . – Режим доступа : scopus.com. – Загл. С экрана. – Яз. рус., англ.

Elibrary [Электронный ресурс] : научная электронная библиотека. – Москва : НЭБ, 2000– . – Режим доступа : elibrary.ru. – Загл. с экрана. – Яз. рус., англ.

8.4. Перечень программного обеспечения

№ п/п	Наименование ПО	Реквизиты договора (дата, номер, срок действия)
1	Windows	2013г., № 00179-40183-81808-ААОЕМ, бессрочный
2	Microsoft Office 13	№61935138 от 28.05.2012 (бессрочный)
3	СУБД MS SQL Server Express + среда MS SQL Studio (рус) 2008 R2 или выше	бесплатное ПО

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

№ п/ п	Наименование оборудованных учебных кабинетов, лабораторий, мастерских и др. объектов для проведения практических и лабораторных занятий, помещений для самостоятельной работы обучающихся (номер аудитории)	Перечень основного оборудования
	Компьютерный класс. Учебная аудитория для проведения занятий лекционного типа. Учебная аудитория для проведения занятий семинарского типа. Учебная аудитория для проведения лабораторных работ. Учебная аудитория для курсового проектирования (выполнения курсовых работ). Учебная аудитория для проведения групповых и индивидуальных консультаций Учебная аудитория для проведения занятий текущего контроля и промежуточной аттестации. (УЛК-408)	Компьютер (монитор 17", системный блок Intel (R) Celeron (R) 2,66 GHz / 1 Gb / 80 Gb), маршрутизатор 2801 Router, коммутатор Catalyst, экран/интерактивная доска Smart Board ТВ, проектор Acer P1303W., стол преподавательский, стол ученический, стол компьютерный, стул, доска аудиторная (маркерная).
	Компьютерный класс. Помещение для самостоятельной работы. Учебная аудитория для проведения занятий семинарского типа. Учебная аудитория для курсового проектирования (выполнения курсовых работ). Учебная аудитория для проведения групповых и индивидуальных консультаций. Учебная аудитория для проведения занятий текущего контроля и промежуточной аттестации (Г-401)	Стол� ученические, стулья ученические, ПК с выходом в сеть Интернет

