

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Выборнова Любовь Алексеевна
Должность: Ректор
Дата подписания: 31.08.2026 09:30:57
Уникальный программный ключ:
c3b3b9c625f6c113afa2a2c42baff9e05a38b76e

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение высшего образования «Поволжский государственный университет сервиса»
(ФГБОУ ВО «ПВГУС»)

УТВЕРЖДЕНО

на заседании Высшей школы
интеллектуальных систем и
кибертехнологий

Протокол от 06.09.2025 № 2

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
(фонд оценочных средств)
для проведения промежуточной аттестации по дисциплине

«Анализ данных»

наименование дисциплины

по образовательной программе высшего образования –
программе

бакалавриат

«Информационные системы и технологии»

наименование образовательной программы

09.03.02 «Информационные системы и технологии»

шифр, наименование направления подготовки / специальности

Составит
или

Коновалова Анжелика Денисовна, старший
преподаватель, Высшая школа интеллектуальных
систем и кибертехнологий

ФИО, должность, структурное подразделение,
ученая степень, ученое звание

Тольятти
2025

1. Паспорт фонда оценочных средств (далее – ФОС)

1.1. Планируемые результаты обучения по дисциплине

Код компетенции	Наименование компетенции	Код и наименование индикаторов достижения компетенции	Основание (ПС) *для профессиональных компетенций
ОПК-6	Способен анализировать и разрабатывать организационно-технические и экономические процессы с применением методов системного анализа и математического моделирования	ИОПК-6.1. Применяет знания основ теории систем и системного анализа, дискретной математики, математической логики, теории алгоритмов, математического и имитационного моделирования для автоматизации задач принятия решений, анализа информационных потоков, расчета экономической эффективности и надежности информационных систем и технологий. ИОПК-6.2. Проводит инженерные расчеты основных показателей результативности создания и применения информационных систем и технологий.	–
ПК-5	Способен к выполнению работ по управлению, настройке, разработке баз данных ИС	ИПК-5.1. Осуществляет разработку структуры баз данных ИС в соответствии с архитектурной спецификацией	-

В результате изучения дисциплины обучающийся должен:

владеть навыками: разведочного и статистического анализа данных, построения и оценки аналитических моделей, расчета показателей, документирования и интерпретации результатов для поддержки принятия решений.

уметь: формализовать задачу анализа данных, исследовать и подготавливать данные, выбирать и применять статистические и математические методы, рассчитывать и интерпретировать показатели результативности, обосновывать выводы с учетом ограничений исходных данных и моделей.

знать: основы системного подхода к анализу данных; методы описательной статистики, статистического вывода, математического моделирования и оценки качества моделей; показатели результативности информационных систем и технологий, применяемые при аналитическом обосновании решений.

1.2. Содержание дисциплины

№	Тема (раздел дисциплины) (в соответствии с РПД)	Код компетенции
1	ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ. ЖИЗНЕННЫЙ ЦИКЛ АНАЛИТИЧЕСКОГО ИССЛЕДОВАНИЯ	ОПК-6 ПК-5
2	ИСТОЧНИКИ ДАННЫХ И ПОДГОТОВКА ДАННЫХ К АНАЛИЗУ	ОПК-6 ПК-5
3	ОПИСАТЕЛЬНАЯ СТАТИСТИКА И РАЗВЕДОЧНЫЙ АНАЛИЗ ДАННЫХ	ОПК-6 ПК-5

№	Тема (в соответствии с РПД)	(раздел дисциплины)	Код компетенции
4	МОДЕЛИРОВАНИЕ ДАННЫХ И СТРУКТУРА АНАЛИТИЧЕСКОЙ БАЗЫ ДАННЫХ		ОПК-6 ПК-5
5	SQL ДЛЯ АНАЛИЗА ДАННЫХ И АГРЕГИРОВАНИЕ ПОКАЗАТЕЛЕЙ		ОПК-6 ПК-5
6	СТАТИСТИЧЕСКИЙ ВЫВОД И ПРОВЕРКА ГИПОТЕЗ		ОПК-6 ПК-5
7	РЕГРЕССИОННЫЙ АНАЛИЗ И ПРОГНОЗИРОВАНИЕ		ОПК-6 ПК-5
8	КЛАССИФИКАЦИЯ, СЕГМЕНТАЦИЯ И ОЦЕНКА АНАЛИТИЧЕСКИХ МОДЕЛЕЙ		ОПК-6 ПК-5
9	КОМПЛЕКСНЫЙ АНАЛИТИЧЕСКИЙ ПРОЕКТ И ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ		ОПК-6 ПК-5

1.3. Система оценивания по дисциплине

Дисциплина изучается в 5 семестре.

Форма промежуточной аттестации по дисциплине – экзамен.

Форма промежуточной аттестации	Уровень сформированности	100-балльная шкала, %	Шкала освоения, %	5-балльная дифференцированная оценка/балл
экзамен	допороговый	ниже 61	ниже 61	«неудовлетворительно» / 2
	пороговый	61–85,9	61–69,9	«удовлетворительно» / 3
	пороговый	61–85,9	70–85,9	«хорошо» / 4
	повышенный	86–100	86–100	«отлично» / 5

2. Перечень оценочных материалов для проведения промежуточной аттестации

Контроль и оценка результатов освоения дисциплины осуществляется преподавателем в ходе текущего контроля успеваемости (в процессе проведения практических и лабораторных занятий, тестирования, защиты отчетов, выполнения курсовой работы).

В ходе проведения промежуточной аттестации осуществляется контроль и оценка результатов освоения компетенций.

Перечень вопросов и заданий для подготовки к экзамену

ОПК-6. Способен анализировать и разрабатывать организационно-технические и экономические процессы с применением методов системного анализа и математического моделирования

1. Понятие анализа данных и его место в прикладной информатике, системной и бизнес-аналитике.
2. Описательный, диагностический, прогнозный и предписывающий анализ: назначение и различия.
3. Этапы жизненного цикла аналитического исследования. Требования к воспроизводимости результата.
4. Типы данных и шкалы измерения. Влияние типа данных на выбор аналитического метода.

5. Источники данных информационных систем. Понятие информационного потока в аналитической задаче.
6. Критерии качества данных: полнота, согласованность, точность, актуальность, уникальность.
7. Пропуски в данных: причины, способы выявления и стратегии обработки.
8. Дубликаты, противоречивые записи и нарушения форматов: выявление и устранение.
9. Аномальные значения и выбросы: способы обнаружения и риски автоматического удаления.
10. Преобразование признаков: кодирование, масштабирование, нормализация и агрегирование.
11. Меры центральной тенденции и положения: среднее, медиана, мода, квантили.
12. Меры вариации: размах, дисперсия, стандартное отклонение, межквартильный размах.
13. Распределения данных и их графическое исследование. Разведочный анализ данных.
14. Ковариация и корреляция. Корреляционная связь и ограничения причинной интерпретации.
15. Понятие показателя результативности информационной системы или процесса. Принципы расчета и интерпретации.
16. Генеральная совокупность, выборка и репрезентативность. Выборочные оценки.
17. Доверительный интервал и уровень доверия: назначение и интерпретация.
18. Нулевая и альтернативная статистические гипотезы. Уровень значимости и р-значение.
19. Ошибки первого и второго рода. Статистическая мощность.
20. Параметрические и непараметрические критерии: условия выбора.
21. Принципы статистического сравнения двух групп и интерпретация результата.
22. А/В-сравнение: постановка, метрика, гипотеза, риски множественных проверок и преждевременного вывода.
23. Задача регрессии. Простая и множественная линейная регрессия.
24. Интерпретация коэффициентов регрессионной модели и ограничения этой интерпретации.
25. Метрики качества регрессии: MAE, MSE, RMSE, R^2 .
26. Анализ остатков и диагностика регрессионной модели.
27. Обучающая и тестовая выборки. Кросс-валидация и контроль переобучения.
28. Задача классификации. Логистическая регрессия и дерево решений как базовые модели.
29. Матрица ошибок. Метрики accuracy, precision, recall и F1.
30. ROC-кривая и ROC-AUC: назначение и ограничения.
31. Задача кластеризации. Понятие меры расстояния и алгоритм k-средних.
32. Интерпретация сегментов и оценка практической полезности кластеризации.
33. Переобучение и недообучение аналитической модели. Способы обнаружения и снижения риска.
34. Интерпретируемость моделей и требования к обоснованию аналитического решения.
35. Документирование аналитического исследования: данные, преобразования, методы, метрики, ограничения и выводы.

ПК-5. Способен к выполнению работ по управлению, настройке, разработке баз данных ИС

36. Реляционная модель данных: сущности, атрибуты, домены и отношения.
37. Первичные и внешние ключи. Ссылочная целостность и ограничения данных.
38. Типы связей между сущностями и способы их реализации в структуре БД.
39. Функциональные зависимости и назначение нормализации данных.
40. Первая, вторая и третья нормальные формы. Практический смысл нормализации.
41. Архитектурная спецификация ИС как основание для разработки структуры базы данных.
42. Логическое и физическое проектирование базы данных: назначение и результаты этапов.
43. Операционная и аналитическая модели данных: различия и области применения.
44. Таблица фактов и таблицы измерений. Схема «звезда» для аналитической витрины.
45. Основные конструкции SQL для получения аналитической выборки: SELECT, WHERE, ORDER BY.
46. Соединения таблиц JOIN и их применение для объединения данных предметной области.
47. Агрегатные функции, GROUP BY и HAVING при расчете аналитических показателей.
48. Подзапросы и общие табличные выражения (CTE) в аналитических запросах.
49. Оконные функции SQL: ранжирование, накопительные показатели и сравнение периодов.
50. Требования к структуре аналитической БД, обеспечивающей воспроизводимый расчет показателей и подготовку данных для моделирования.

Типовые практические задания при промежуточной аттестации

1. Для заданного описания бизнес-процесса сформулировать аналитическую задачу, определить набор необходимых данных и не менее пяти измеримых показателей.
2. По фрагменту набора данных выявить проблемы качества, выбрать способы обработки и обосновать последовательность преобразований.
3. По представленным описательным статистикам и распределениям сформулировать выводы о состоянии показателя и возможных аномалиях.
4. Рассчитать и интерпретировать заданный показатель результативности информационной системы по исходной таблице наблюдений.
5. Для прикладной ситуации сформулировать статистические гипотезы, выбрать критерий и интерпретировать предоставленное p -значение.
6. По результатам двух регрессионных моделей выбрать предпочтительную с использованием MAE/RMSE/ R^2 и объяснить ограничения выбора.
7. По матрице ошибок рассчитать precision, recall и F1 и выбрать модель для заданной стоимости ошибок.
8. По архитектурной спецификации спроектировать структуру реляционной БД: сущности, атрибуты, ключи, связи и ограничения.
9. Нормализовать исходную таблицу до третьей нормальной формы и указать первичные и внешние ключи.
10. Написать SQL-запрос с JOIN и GROUP BY для расчета показателя по сегментам и периодам.
11. Написать SQL-запрос с оконной функцией для ранжирования или накопительного итога.
12. Спроектировать аналитическую витрину по схеме «звезда» для заданного набора показателей.
13. Разработать сквозную схему аналитического решения: источник → структура БД → подготовка → метод анализа → метрика → представление результата.