

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
СТАВРОПОЛЬСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ**

УТВЕРЖДАЮ

Директор/Декан
института экономики, финансов и
управления в АПК
Гунько Юлия Александровна

«__» _____ 20__ г.

Рабочая программа дисциплины

Б1.О.17.04 Анализ данных

38.03.02 Менеджмент

Управление бизнесом

бакалавр

очная

1. Цель дисциплины

Задачами освоения дисциплины "Анализ данных" является освоение методологии исследования данных в разных сферах деятельности, включая их сбор и обработку с применением различных инструментов (Excel, Python (библиотеки Pandas, Matplotlib), SQL), а также визуализацию (Superset, Datalens) и интерпретацию с помощью статистических методов

2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Процесс изучения дисциплины направлен на формирование следующих компетенций ОП ВО и овладение следующими результатами обучения по дисциплине:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Перечень планируемых результатов обучения по дисциплине
ОПК-5 Способен использовать при решении профессиональных задач современные информационные технологии и программные средства, включая управление крупными массивами данных и их интеллектуальный анализ.	ОПК-5.3 Применяет технологии управления крупными массивами данных и их интеллектуальный анализ при решении профессиональных задач	знает умеет владеет навыками
ОПК-6 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности.	ОПК-6.3 Использует программно-технические средства обработки данных в профессиональной деятельности	знает умеет владеет навыками

3. Место дисциплины в структуре образовательной программы

Дисциплина «Анализ данных» является дисциплиной обязательной части программы. Изучение дисциплины осуществляется в 5 семестре(-ах).

Для освоения дисциплины «Анализ данных» студенты используют знания, умения и навыки, сформированные в процессе изучения дисциплин:

Технологическая (проектно-технологическая) практика

Базы данных и SQL

Управление данными

Интернет-технологии

Разработка программных приложений

Конфигурирование в информационных системах

Web-программирование

Технологии разработки web-ресурсов

Системы поддержки принятия решений

Экспертные системы Цифровые технологии в профессиональной деятельности

Технологическая (проектно-технологическая) практика
Базы данных и SQL
Управление данными
Интернет-технологии
Разработка программных приложений
Конфигурирование в информационных системах
Web-программирование
Технологии разработки web-ресурсов
Системы поддержки принятия решений
Экспертные системыВысшая математика
Технологическая (проектно-технологическая) практика
Базы данных и SQL
Управление данными
Интернет-технологии
Разработка программных приложений
Конфигурирование в информационных системах
Web-программирование
Технологии разработки web-ресурсов
Системы поддержки принятия решений
Экспертные системыМетоды оптимальных решений
Технологическая (проектно-технологическая) практика
Базы данных и SQL
Управление данными
Интернет-технологии
Разработка программных приложений
Конфигурирование в информационных системах
Web-программирование
Технологии разработки web-ресурсов
Системы поддержки принятия решений
Экспертные системыWeb программирование
Технологическая (проектно-технологическая) практика
Базы данных и SQL
Управление данными
Интернет-технологии
Разработка программных приложений
Конфигурирование в информационных системах
Web-программирование
Технологии разработки web-ресурсов
Системы поддержки принятия решений
Экспертные системыИнформационные технологии
Технологическая (проектно-технологическая) практика
Базы данных и SQL
Управление данными
Интернет-технологии
Разработка программных приложений
Конфигурирование в информационных системах
Web-программирование
Технологии разработки web-ресурсов
Системы поддержки принятия решений
Экспертные системыОзнакомительная практика

Технологическая (проектно-технологическая) практика
 Базы данных и SQL
 Управление данными
 Интернет-технологии
 Разработка программных приложений
 Конфигурирование в информационных системах
 Web-программирование
 Технологии разработки web-ресурсов
 Системы поддержки принятия решений
 Экспертные системы 1С: Кадровое дело/производство
 Технологическая (проектно-технологическая) практика
 Базы данных и SQL
 Управление данными
 Интернет-технологии
 Разработка программных приложений
 Конфигурирование в информационных системах
 Web-программирование
 Технологии разработки web-ресурсов
 Системы поддержки принятия решений
 Экспертные системы Цифровые финансы
 Технологическая (проектно-технологическая) практика
 Базы данных и SQL
 Управление данными
 Интернет-технологии
 Разработка программных приложений
 Конфигурирование в информационных системах
 Web-программирование
 Технологии разработки web-ресурсов
 Системы поддержки принятия решений
 Экспертные системы Статистика

Освоение дисциплины «Анализ данных» является необходимой основой для последующего изучения следующих дисциплин:

Выполнение и защита выпускной квалификационной работы
 Преддипломная практика

4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу с обучающимися с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины «Анализ данных» в соответствии с рабочим учебным планом и ее распределение по видам работ представлены ниже.

Семестр	Трудоемкость час/з.е.	Контактная работа с преподавателем, час			Самостоятельная работа, час	Контроль, час	Форма промежуточной аттестации (форма контроля)
		лекции	практические занятия	лабораторные занятия			
5	108/3	18	36		54		За
в т.ч. часов: в интерактивной форме		4	8				

Семестр	Трудоемкость час/з.е.	Внеаудиторная контактная работа с преподавателем, час/чел					
		Курсовая работа	Курсовой проект	Зачет	Дифференцированный зачет	Консультации перед экзаменом	Экзамен
5	108/3			0.12			

5. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

№	Наименование раздела/темы	Семестр	Количество часов					Формы текущего контроля успеваемости и промежуточной аттестации	Оценочное средство проверки результатов достижения индикаторов компетенций	Код индикаторов достижения компетенций
			всего	Лекции	Семинарские занятия		Самостоятельная работа			
					Практические	Лабораторные				
1.	1 раздел. Раздел 1. Основы анализа данных									
1.1.	1. Введение в науку о данных	5	8	2	6		4			
1.2.	2. Инструменты обработки данных	5	8	2	6		4			
1.3.	3. Визуализация данных	5	8	2	6		4	КТ 1	Задачи	
2.	2 раздел. Раздел 2. Трансформация данных									
2.1.	4. Анализ и преобразование данных	5	8	2	6		4			
2.2.	5. Работа с временными рядами	5	8	2	6		2			
2.3.	6 Системы управления базами данных	5	8	2	6		8			
2.4.	7 Проектирование структурированных данных	5	2	2			8			
3.	3 раздел. Раздел 3. Сбор и обработка больших данных									
3.1.	8. Запросы на языке SQL	5	2	2			12			
3.2.	9. NoSQL хранилища	5	2	2			8			
	Промежуточная аттестация		За							
	Итого		108	18	36		54			
	Итого		108	18	36		54			

5.1. Лекционный курс с указанием видов интерактивной формы проведения занятий

Тема лекции (и/или наименование раздел) (вид интерактивной формы проведения занятий)/ (практическая подготовка)	Содержание темы (и/или раздела)	Всего, часов / часов интерактивных занятий/ практическая подготовка
1. Введение в науку о данных	1. Введение в науку о данных	2/-
2. Инструменты обработки данных	2. Инструменты обработки данных	2/-
3. Визуализация данных	3. Визуализация данных	2/-
4. Анализ и преобразование данных	4. Анализ и преобразование данных	2/-
5. Работа с временными рядами	5. Работа с временными рядами	2/-

рядами		
6 Системы управления базами данных	6 Системы управления базами данных	2/-
7 Проектирование структурированных данных	7 Проектирование структурированных данных	2/-
8. Запросы на языке SQL	8. Запросы на языке SQL	2/2
9. NoSQL хранилища	9. NoSQL хранилища	2/-
Итого		18

5.2.1. Семинарские (практические) занятия с указанием видов проведения занятий в интерактивной форме

Наименование раздела дисциплины	Формы проведения и темы занятий (вид интерактивной формы проведения занятий)/(практическая подготовка)	Всего, часов / часов интерактивных занятий/ практическая подготовка	
		вид	часы
1. Введение в науку о данных		Пр	6/2/-
2. Инструменты обработки данных	2. Инструменты обработки данных	Пр	6/2/-
3. Визуализация данных	3. Визуализация данных	Пр	6/-/-
4. Анализ и преобразование данных	4. Анализ и преобразование данных	Пр	6/2/-
5. Работа с временными рядами	5. Работа с временными рядами	Пр	6/2/-
6 Системы управления базами данных		Пр	6/2/-
Итого			

5.3. Курсовой проект (работа) учебным планом не предусмотрен

5.4. Самостоятельная работа обучающегося

Темы и/или виды самостоятельной работы	Часы
	4
	4

	4
	4
	2
	8
	8
	12
	8

6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Учебно-методическое обеспечение для самостоятельной работы обучающегося по дисциплине «Анализ данных» размещено в электронной информационно-образовательной среде Университета и доступно для обучающегося через его личный кабинет на сайте Университета. Учебно-методическое обеспечение включает:

1. Рабочую программу дисциплины «Анализ данных».
2. Методические рекомендации для организации самостоятельной работы обучающегося по дисциплине «Анализ данных».
3. Методические рекомендации по выполнению письменных работ (задачи) (при наличии).
4. Методические рекомендации по выполнению контрольной работы студентами заочной формы обучения (при наличии)
5. Методические указания по выполнению курсовой работы (проекта) (при наличии).

Для успешного освоения дисциплины, необходимо самостоятельно детально изучить представленные темы по рекомендуемым источникам информации:

№ п/п	Темы для самостоятельного изучения	Рекомендуемые источники информации (№ источника)		
		основная (из п.8 РПД)	дополнительная (из п.8 РПД)	метод. лит. (из п.8 РПД)
1	1. Введение в науку о данных.			
2	2. Инструменты обработки данных.			
3	3. Визуализация данных.			
4	4. Анализ и преобразование данных.			
5	5. Работа с временными рядами.			
6	6 Системы управления базами данных.			
7	7 Проектирование структурированных данных.			
8	8. Запросы на языке SQL.			
9	9. NoSQL хранилища.			

7. Фонд оценочных средств (оценочных материалов) для проведения промежуточной аттестации обучающихся по дисциплине «Анализ данных»

7.1. Перечень индикаторов компетенций с указанием этапов их формирования в процессе освоения образовательной программы

Индикатор компетенции (код и содержание)	Дисциплины/элементы программы (практики, ГИА), участвующие в формировании индикатора компетенции	1		2		3		4	
		1	2	3	4	5	6	7	8
ОПК-5.3:Применяет технологии управления крупными массивами данных и их интеллектуальный анализ при решении профессиональных задач	Преддипломная практика								x
ОПК-6.3:Использует программно-технические средства обработки	Высшая математика	x	x						
	Информационные технологии	x							

Индикатор компетенции (код и содержание)	Дисциплины/элементы программы (практики, ГИА), участвующие в формировании индикатора компетенции	1		2		3		4	
		1	2	3	4	5	6	7	8
данных в профессиональной деятельности	Методы оптимальных решений			x					
	Преддипломная практика								x
	Статистика			x					
	Цифровые технологии в профессиональной деятельности	x			x				

7.2. Критерии и шкалы оценивания уровня усвоения индикатора компетенций, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Оценка знаний, умений и навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций по дисциплине «Анализ данных» проводится в форме текущего контроля и промежуточной аттестации.

Текущий контроль проводится в течение семестра с целью определения уровня усвоения обучающимися знаний, формирования умений и навыков, своевременного выявления преподавателем недостатков в подготовке обучающихся и принятия необходимых мер по её корректировке, а также для совершенствования методики обучения, организации учебной работы и оказания индивидуальной помощи обучающемуся.

Промежуточная аттестация по дисциплине «Анализ данных» проводится в виде Зачет.

За знания, умения и навыки, приобретенные студентами в период их обучения, выставляются оценки «ЗАЧТЕНО», «НЕ ЗАЧТЕНО». (или «ОТЛИЧНО», «ХОРОШО», «УДОВЛЕТВОРИТЕЛЬНО», «НЕУДОВЛЕТВОРИТЕЛЬНО» для дифференцированного зачета/экзамена)

Для оценивания знаний, умений, навыков и (или) опыта деятельности в университете применяется балльно-рейтинговая система оценки качества освоения образовательной программы. Оценка проводится при проведении текущего контроля успеваемости и промежуточных аттестаций обучающихся. Рейтинговая оценка знаний является интегрированным показателем качества теоретических и практических знаний и навыков студентов по дисциплине.

Состав балльно-рейтинговой оценки студентов очной формы обучения

Для студентов очной формы обучения знания по осваиваемым компетенциям формируются на лекционных и практических занятиях, а также в процессе самостоятельной подготовки.

В соответствии с балльно-рейтинговой системой оценки, принятой в Университете студентам начисляются баллы по следующим видам работ:

№ контрольной точки	Оценочное средство результатов индикаторов достижения компетенций		Максимальное количество баллов
5 семестр			
КТ 1	Задачи		0
Сумма баллов по итогам текущего контроля			0
Посещение лекционных занятий			20
Посещение практических/лабораторных занятий			20
Результативность работы на практических/лабораторных занятиях			30
Итого			70
№ контрольной точки	Оценочное средство результатов индикаторов достижений компетенций	Максимальное количество баллов	Критерии оценки знаний студентов
5 семестр			
КТ 1	Задачи	0	

Критерии и шкалы оценивания результатов обучения на промежуточной аттестации

При проведении итоговой аттестации «зачет» («дифференцированный зачет», «экзамен») преподавателю с согласия студента разрешается выставлять оценки («отлично», «хорошо», «удовлетворительно», «зачет») по результатам набранных баллов в ходе текущего контроля успеваемости в семестре по выше приведенной шкале.

В случае отказа – студент сдает зачет (дифференцированный зачет, экзамен) по приведенным выше вопросам и заданиям. Итоговая успеваемость (зачет, дифференцированный зачет, экзамен) не может оцениваться ниже суммы баллов, которую студент набрал по итогам текущей и промежуточной успеваемости.

При сдаче (зачета, дифференцированного зачета, экзамена) к заработанным в течение семестра студентом баллам прибавляются баллы, полученные на (зачете, дифференцированном зачете, экзамене) и сумма баллов переводится в оценку.

Критерии и шкалы оценивания ответа на зачете

По дисциплине «Анализ данных» к зачету допускаются студенты, выполнившие и сдавшие практические работы по дисциплине, имеющие ежемесячную аттестацию и без привязке к набранным баллам. Студентам, набравшим более 65 баллов, зачет выставляется по результатам текущей успеваемости, студенты, не набравшие 65 баллов, сдают зачет по вопросам, предусмотренным РПД. Максимальная сумма баллов по промежуточной аттестации (зачету) устанавливается в 15 баллов

Вопрос билета	Количество баллов
Теоретический вопрос	до 5
Задания на проверку умений	до 5
Задания на проверку навыков	до 5

Теоретический вопрос

5 баллов выставляется студенту, полностью освоившему материал дисциплины или курса в соответствии с учебной программой, включая вопросы рассматриваемые в рекомендованной программой дополнительной справочно-нормативной и научно-технической литературы, свободно владеющему основными понятиями дисциплины. Требуется полное понимание и четкость изложения ответов по экзаменационному заданию (билету) и дополнительным вопросам, заданных экзаменатором. Дополнительные вопросы, как правило, должны относиться к материалу дисциплины или курса, не отраженному в основном экзаменационном задании (билете) и выявляют полноту знаний студента по дисциплине.

4 балла заслуживает студент, ответивший полностью и без ошибок на вопросы экзаменационного задания и показавший знания основных понятий дисциплины в соответствии с обязательной программой курса и рекомендованной основной литературой.

3 балла дан недостаточно полный и недостаточно развернутый ответ. Логика и последовательность изложения имеют нарушения. Допущены ошибки в раскрытии понятий, употреблении терминов. Студент не способен самостоятельно выделить существенные и несущественные признаки и причинно-следственные связи. Студент может конкретизировать обобщенные знания, доказав на примерах их основные положения только с помощью преподавателя. Речевое оформление требует поправок, коррекции.

2 балла дан неполный ответ, представляющий собой разрозненные знания по теме вопроса с существенными ошибками в определениях. Присутствуют фрагментарность, нелогичность изложения. Студент не осознает связь данного понятия, теории, явления с другими объектами дисциплины. Отсутствуют выводы, конкретизация и доказательность изложения. Речь неграмотная. Дополнительные и уточняющие вопросы преподавателя не приводят к коррекции ответа студента не только на поставленный вопрос, но и на другие вопросы дисциплины.

1 балл дан неполный ответ, представляющий собой разрозненные знания по теме вопроса с существенными ошибками в определениях. Присутствуют фрагментарность, нелогичность изложения. Студент не осознает связь данного понятия, теории, явления с другими объектами дисциплины. Отсутствуют выводы, конкретизация и доказательность изложения. Речь неграмотная. Дополнительные и уточняющие вопросы преподавателя не приводят к коррекции ответа студента не только на поставленный вопрос, но и на другие вопросы дисциплины.

0 баллов - при полном отсутствии ответа, имеющего отношение к вопросу.

Задания на проверку умений и навыков

5 баллов Задания выполнены в обозначенный преподавателем срок, письменный отчет без замечаний. Работа выполнена в полном объеме с соблюдением необходимой последовательности.

4 балла Задания выполнены в обозначенный преподавателем срок, письменный отчет с небольшими недочетами.

2 баллов Задания выполнены с задержкой, письменный отчет с недочетами. Работа выполнена не полностью, но объем выполненной части таков, что позволяет получить правильные результаты и выводы.

1 баллов Задания выполнены частично, с большим количеством вычислительных ошибок, объем выполненной части работы не позволяет сделать правильных выводов.

0 баллов Задания выполнены, письменный отчет не представлен или работа выполнена не полностью, и объем выполненной части работы не позволяет сделать правильных выводов.

7.3. Примерные оценочные материалы для текущего контроля успеваемости и промежуточной аттестации по итогам освоения дисциплины «Анализ данных»

1. Что такое наука о данных?
 - a) Метод анализа текстов
 - b) Область знаний, основанная на обработке и анализе данных
 - c) Исключительно машинное обучение
 - d) Управление базами данных
2. Какой из этапов анализа данных включает очистку и подготовку данных?
 - a) Сбор данных
 - b) Моделирование
 - c) Обработка данных
 - d) Визуализация данных
3. Что из перечисленного не является задачей науки о данных?
 - a) Классификация
 - b) Регрессия
 - c) Агрегация данных
 - d) Линейная алгебра
4. Какой метод чаще всего используется для поиска скрытых закономерностей в данных?
 - a) Кластеризация
 - b) Ранжирование
 - c) Анализ выбросов
 - d) Анкеты и опросы
5. Что отличает анализ данных от традиционной статистики?
 - a) Только работа с большими данными
 - b) Широкий спектр методов и подходов
 - c) Использование исключительно числовых данных
 - d) Отсутствие необходимости в визуализации
6. Какая задача лучше всего подходит для классификации?
 - a) Определение будущих продаж
 - b) Определение пола по фотографии
 - c) Поиск ключевых слов в тексте
 - d) Вычисление средней заработной платы
7. Что является конечной целью анализа данных?
 - a) Построение графиков
 - b) Обнаружение инсайтов для принятия решений
 - c) Хранение данных
 - d) Создание таблиц
8. Какую задачу решает метод регрессии?
 - a) Определение категорий данных

- b) Прогнозирование количественных значений
 - c) Обнаружение аномалий
 - d) Сегментация данных
9. Что подразумевается под термином "большие данные"?
- a) Данные, которые невозможно хранить в реляционных базах данных
 - b) Большие файлы данных
 - c) Данные объемом свыше 1 ГБ
 - d) Данные, которые требуют специальных методов обработки
10. Какие основные этапы анализа данных?
- a) Визуализация и моделирование
 - b) Сбор, обработка, анализ, визуализация
 - c) Сбор и анализ
 - d) Моделирование и прогнозирование
11. Какой язык программирования наиболее популярен для анализа данных?
- a) Python
 - b) C++
 - c) Java
 - d) PHP
12. Какая библиотека Python предназначена для работы с массивами данных?
- a) Matplotlib
 - b) NumPy
 - c) Pandas
 - d) Seaborn
13. Что такое Jupyter Notebook?
- a) Язык программирования
 - b) Среда для написания и выполнения кода
 - c) Веб-сервер
 - d) База данных
14. Для чего используется библиотека Pandas?
- a) Для работы с многомерными массивами
 - b) Для работы с таблицами и временными рядами
 - c) Для работы с нейронными сетями
 - d) Для визуализации данных
15. Какую роль выполняет инструмент Apache Spark?
- a) Управление базами данных
 - b) Обработка больших объемов данных
 - c) Машинное обучение
 - d) Визуализация данных
16. Какой формат файлов чаще всего используется для хранения данных?
- a) .xls
 - b) .docx
 - c) .json
 - d) .png
17. Что позволяет делать язык R?
- a) Хранить данные
 - b) Строить интерактивные веб-сайты
 - c) Проводить статистический анализ
 - d) Создавать базы данных
18. Какая библиотека Python используется для визуализации данных?
- a) Pandas
 - b) Matplotlib
 - c) NumPy
 - d) TensorFlow
19. Что из перечисленного является инструментом для работы с большими данными?
- a) MongoDB
 - b) Apache Hadoop

c) SQLite

d) Redis

10. Какова цель визуализации данных?

A) Сжатие данных

B) Представление данных в наглядной форме

C) Хранение данных

D) Анализ алгоритмов

11. Какой тип диаграммы используется для отображения частей целого?

A) Столбчатая диаграмма

B) Линейная диаграмма

C) Круговая диаграмма

D) Точечная диаграмма

12. Что из перечисленного НЕ является инструментом визуализации данных?

A) Tableau

B) Power BI

C) Jupyter Notebook

D) PostgreSQL

13. Какая библиотека Python чаще всего используется для построения графиков?

A) NumPy

B) pandas

C) Matplotlib

D) SQLAlchemy

14. Что из следующего лучше всего подходит для представления временных рядов?

A) Линейный график

B) Тепловая карта

C) Гистограмма

D) Круговая диаграмма

15. Какой тип диаграммы лучше всего использовать для отображения распределения данных?

A) Точечная диаграмма

B) Гистограмма

C) Линейный график

D) Столбчатая диаграмма

16. Как называется процесс уменьшения количества точек на диаграмме без потери сути?

A) Нормализация

B) Агрегация

C) Интеграция

D) Уменьшение размерности

17. Что такое интерактивная визуализация данных?

A) Визуализация с возможностью взаимодействия пользователя с данными

B) Статическая визуализация

C) Программирование визуальных эффектов

D) Использование 3D графики

18. Какая диаграмма лучше всего подходит для анализа зависимости между двумя переменными?

A) Точечная диаграмма

B) Круговая диаграмма

C) Линейный график

D) Гистограмма

19. Для чего используется тепловая карта?

A) Для отображения корреляций между данными

B) Для представления временных данных

C) Для анализа сезонности

D) Для отображения категорий

20. Что означает термин "очистка данных"?

- A) Удаление данных
- B) Устранение ошибок и пропусков в данных
- C) Создание новых переменных
- D) Преобразование форматов данных

21. Какой метод используется для обработки пропущенных данных?

- A) Удаление строк с пропусками
- B) Интерполяция данных
- C) Заполнение средними значениями
- D) Все перечисленные

22. • Что такое нормализация данных?

- A) Приведение данных к диапазону от 0 до 1
- B) Уменьшение объема данных
- C) Увеличение точности данных
- D) Удаление выбросов

23. • Какой метод применяется для уменьшения размерности данных?

- A) Метод главных компонент (PCA)
- B) Кластеризация
- C) Регрессия
- D) Нормализация

24. • Что такое категоризация данных?

- A) Преобразование числовых данных в категориальные
- B) Сортировка данных
- C) Упорядочивание данных
- D) Объединение данных

25. • Что такое выбросы в данных?

- A) Нормальные значения в данных
- B) Исключительные или аномальные значения
- C) Упущенные данные
- D) Минимальные значения

26. • Для чего используется стандартизация данных?

- A) Для приведения данных к одному масштабу
- B) Для удаления лишних данных
- C) Для объединения данных
- D) Для создания моделей

27. • Какой метод используется для выявления скрытых закономерностей в данных?

- A) Регрессия
- B) Кластеризация
- C) Тестирование гипотез
- D) Вычисление средних

28. • Какой формат данных чаще всего используется для работы с большими объемами данных?

- A) JSON
- B) CSV
- C) Excel
- D) TXT

29. • Что такое фича-инжиниринг?

- A) Процесс создания новых признаков на основе существующих данных
- B) Процесс удаления лишних данных
- C) Процесс агрегирования данных

30. Для чего используется индекс в СУБД?

- A) Для улучшения производительности запросов
- B) Для хранения данных
- C) Для сжатия данных
- D) Для объединения таблиц

"Разработка информационной системы для обработки, трансформации и визуализации данных (название набора данных)"

Наборы данных (варианты курсовой работы)

1. Малые и средние предприятия в сфере лёгкой промышленности за 2016–2023 гг.
 2. Загрязнение поверхностных вод в России: ежемесячные данные о высоком и экстремально высоком загрязнении водных объектов за 2008—2021 гг.
 3. Показатели результативности российских вузов: объединенные данные за 2013-2017 годы
 4. Загрязнение атмосферы в России: объем выбросов в воздух в разрезе муниципальных образований за 2019-2020 гг.
 5. Населенные пункты России: численность населения и географические координаты
 6. Федеральный реестр медицинских организаций
 7. Статистические данные о занятости и безработице среди населения по возрастным группам
 8. Демографическая ситуация по субъектам РФ
 9. История изменений муниципалитетов в России: наименования, коды ОКТМО и входящие в состав населенные пункты
 10. Пожары в России: данные о местах и типах природных пожаров за 2012–2021 гг.
 11. Показатели валового регионального продукта в России: объединенные данные за 1998—2019 гг.
 12. «Зеленые» инвестиции: затраты на охрану окружающей среды в регионах РФ по классификации СЕРА
 13. Площадки для выгула (дрессировки) собак
 14. Спортивные площадки
 15. Залы тренажерные
- др.)?
1. Что такое наука о данных и как она отличается от статистики?
 2. Какие ключевые этапы включает процесс анализа данных?
 3. Каковы основные задачи науки о данных (кластеризация, классификация, регрессия и др.)?
 4. В чем заключается разница между описательной и предсказательной аналитикой?
 5. Каковы преимущества и вызовы работы с большими данными?
 6. Что такое машинное обучение и как оно связано с анализом данных?
 7. Какие роли и навыки требуются специалисту по анализу данных?
 8. Какие категории инструментов используются для анализа данных (настольные программы, языки программирования и платформы)?
 9. Чем различаются инструменты для обработки данных, такие как Python и R?
 10. Какие библиотеки Python наиболее популярны для анализа данных и зачем они используются?
 11. Какова роль SQL в обработке данных?
 12. В чем преимущества облачных решений для анализа данных?
 13. Какие инструменты наиболее подходят для визуализации данных и почему?
 14. Как сравнить популярные платформы для работы с большими данными (Hadoop, Spark)?
 15. Какие типы визуализаций данных существуют и когда их следует использовать?
 16. Что такое диаграмма рассеяния и в каких ситуациях она полезна?
 17. В чем разница между столбчатой диаграммой и гистограммой?
 18. Как строить временные ряды и какие особенности их визуализации нужно учитывать?
 19. Что такое интерактивная визуализация данных и каковы ее преимущества?
 20. Какие инструменты используются для создания сводных таблиц и диаграмм?
 21. В чем ключевые отличия между Tableau, Power BI и Google Data Studio?
 22. Что такое очистка данных и почему это важный этап анализа?
 23. Как определить и устранить пропуски в данных?
 24. Какие методы нормализации данных существуют и когда их нужно применять?
 25. Что такое агрегация данных и как она используется в аналитике?
 26. Как категоризировать данные и какие методы преобразования категориальных данных существуют?
 27. Что такое масштабирование данных и зачем оно нужно в анализе данных?

28. Как выполняется кодирование категориальных данных (например, one-hot encoding)?
29. Что такое временные ряды и где они применяются в анализе данных?
30. Каковы основные компоненты временных рядов (тренд, сезонность, шум)?
31. Что такое скользящее среднее и как его использовать для сглаживания временных рядов?
32. Какие методы прогнозирования временных рядов существуют?
33. Как проводится анализ сезонности во временных рядах?
34. В чем заключается разложение временного ряда и для чего оно нужно?
35. Каковы основные методы оценки модели временных рядов?
36. Что такое реляционная база данных и какие принципы лежат в ее основе?
37. Чем реляционные базы данных отличаются от нереляционных?
38. Каковы основные компоненты любой СУБД?
39. Какие типы данных поддерживаются в реляционных базах данных?
40. Что такое первичные и внешние ключи и зачем они нужны?
41. Какие популярные СУБД существуют и каковы их ключевые особенности (MySQL, PostgreSQL, Oracle)?
42. В чем разница между OLTP и OLAP системами?
43. Какие этапы включает проектирование базы данных?
44. Что такое нормализация данных и зачем она нужна?
45. Что такое первая, вторая и третья нормальные формы?
46. Как разрабатывать ER-диаграмму для базы данных?
47. Что такое денормализация данных и в каких случаях ее применяют?
48. Как спроектировать схему базы данных для интернет-магазина?
49. Какие принципы нужно учитывать при проектировании базы данных для больших объемов данных?
50. Как выполнить базовый SELECT-запрос для выборки данных?
51. Что такое оператор WHERE и как использовать его для фильтрации данных?
52. Как выполнять сортировку данных с помощью оператора ORDER BY?
53. Как применять оператор JOIN для объединения данных из нескольких таблиц?
54. Какие существуют агрегатные функции в SQL и как они работают?
55. В чем заключается использование оператора GROUP BY и HAVING?
56. Как создавать подзапросы в SQL и в чем их преимущества?
57. Что такое таблицы в базе данных и как они структурированы?
58. Какова роль индексов в базах данных и как они помогают ускорить запросы?
59. Что такое представления (VIEW) в базе данных и зачем их использовать?
60. Как создавать и использовать хранимые процедуры в базе данных?
61. Что такое триггеры в базе данных и как они работают?
62. Какие типы ограничений могут быть наложены на данные в таблицах?
63. В чем различие между постоянными и временными таблицами в базе данных?
64. Что такое NoSQL базы данных и чем они отличаются от реляционных баз данных?
65. Какие типы NoSQL баз данных существуют (документные, key-value, графовые, столбцовые)?
66. В чем преимущества и недостатки NoSQL баз данных по сравнению с реляционными?
67. Как работает база данных MongoDB и для каких задач она используется?
68. Что такое Redis и какие случаи его использования наиболее популярны?
69. В чем особенности работы с данными в базах данных типа Cassandra?
70. Какой подход к горизонтальному масштабированию используют NoSQL базы данных?

8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

- а) Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

№	Наименование ресурса сети «Интернет»	Электронный адрес ресурса
1		

10. Методические указания для обучающихся по освоению дисциплины

Методические рекомендации по освоению дисциплины "Анализ данных"
Методические рекомендации по выполнению курсовая работа (проекта).

Лекционные занятия

Основа освоения дисциплины – лекция, целью которой является целостное и логичное рассмотрение основного материала курса. Вместе с тем значимость лекции определяется тем, что она не только способствует выработке логического мышления, но и способствует развитию интереса к пониманию современной действительности.

Задача студентов в процессе умелой и целеустремленной работы на лекциях – внимательно слушать преподавателя, следить за его мыслью, предлагаемой системой логических посылок, доказательств и выводов, фиксировать (записывать) основные идеи, важнейшие характеристики понятий, теорий, наиболее существенные факты. Лекция задает направление, содержание и эффективность других форм учебного процесса, нацеливает студентов на самостоятельную работу и определяет основные ее направления (подготовку к практическим занятиям, выполнение творческих заданий, рефератов, решение контекстных задач).

Записывание лекции – творческий процесс. Запись лекции крайне важна. Это позволяет надолго сохранить основные положения лекции; способствует поддержанию внимания; способствует лучшему запоминанию материала. Важно уметь оформить конспект так, чтобы важные моменты были выделены графически, а главную информацию следует выделять в самостоятельные абзацы, фиксируя ее более крупными буквами или цветными маркерами. Конспект должен иметь поля для заметок. Это могут быть библиографические ссылки и, наконец, собственные комментарии. Для быстрой записи текста можно придумать условные знаки, при этом таких знаков не должно быть более 10–15. Условные обозначения придумывают для часто встречающихся слов (существует, который, каждый, точка зрения, на основании и т.п.).

Перед каждой лекцией необходимо внимательно прочитать материал предыдущей лекции. В рабочей тетради графически выделить: тему лекции, основные теоретические положения. Подготовленный студент легко следит за мыслью преподавателя, что позволяет быстрее запоминать новые понятия, сущность которых выявляется в контексте лекции. Повторение материала облегчает в дальнейшем подготовку к экзамену. Затем надо ознакомиться с материалом темы по учебнику, внести нужные уточнения и дополнения в лекционный материал. После усвоения каждой темы рекомендуется проверять свои знания, отвечая на контрольные вопросы по теме.

Лабораторные занятия

Целью лабораторных занятий является закрепление, расширение, углубление теоретических знаний, полученных на лекциях и в ходе самостоятельной работы, развитие познавательных способностей.

Являясь частью образовательного процесса, семинар преследует ряд основополагающих задач:

- работа с источниками, которая идет на уровнях индивидуальной самостоятельной работы и в ходе коллективного обсуждения;
- формирование умений и навыков индивидуальной и коллективной работы, позволяющих эффективно использовать основные методы исследования, грамотно выстраивать его основные технологические этапы (знакомство с темой и имеющейся по ней информацией, определение основной проблемы, первичный анализ, определение подходов и ключевых узлов механизма ее развития, публичное обсуждение, предварительные выводы);
- анализ поставленных проблем, умение обсуждать тему, высказывать свое мнение, отстаивать свою позицию, слушать и оценивать различные точки зрения, конструктивно

полемизировать, учиться думать, говорить, слушать, понимать, находить точки соприкосновения разных позиций, их разумного сочетания;

- формирование установок на творчество;
- диалог, внутренний и внешний; поиск и разрешение проблемы в рамках имеющейся о ней информации;

- поиск рационального зерна в самых противоречивых позициях и подходах к проблеме;
- открытость новому и принципиальную возможность изменить свою позицию и вытекающие из нее решения, в случае получения новой информации и связанных с ней обстоятельств сознательный отход от подготовленного к семинару текста во время своего, построенного на тезисном изложении фактов и мыслей, когда конспект привлекается лишь в том случае, когда надо привести какие-то факты.

Для эффективной работы на практическом занятии студенту необходимо учесть и выполнить следующие требования по подготовке к нему:

1. Внимательно прочитать, как сформулирована тема, определить ее место в учебном плане курса, установить взаимосвязи с другими разделами.

2. Познакомиться с целью и задачами работы на практическом занятии, обратив внимание на то, какие знания, умения и навыки студент должен приобрести в результате активной познавательной деятельности.

3. Проработать основные вопросы и проблемы (задания), которые будут рассматриваться и обсуждаться в ходе практического занятия.

4. Подобрать литературу по теме занятия; найти соответствующий раздел в лекциях и в рекомендуемых пособиях.

5. Добросовестно проработать имеющуюся научную литературу (просмотреть и подобрать информацию, сделать выписки (конспектирование узловых проблем), обработать их в соответствии с задачами практического занятия.

6. Обдумать и предложить свои выводы и мысли на основании полученной информации (предварительное осмысление).

7. Продумать развернутые законченные ответы на предложенные вопросы, предлагаемые творческие задания и контекстные задачи, опираясь на материал лекций, расширяя и дополняя его данными из учебника, дополнительной литературы, составить план ответа, выписать терминологию.

Видами заданий на практических занятиях:

- для овладения знаниями: чтение текста (учебника, первоисточника, дополнительной литературы), работа со словарями и справочниками, ознакомление с нормативными документами, учебно-исследовательская работа, использование аудио- и видеозаписей, компьютерной техники и Интернета и др.

- для закрепления и систематизации знаний: работа с конспектом лекции, обработка текста, повторная работа над учебным материалом (учебника, первоисточника, дополнительной литературы, аудио и видеозаписей, ответы на контрольные вопросы, аналитическая обработка текста, под-готовка мультимедиа сопровождения к защите рефератов, и др.

- для формирования умений: решение контекстных задач, подготовка к деловым играм, выполнение творческих заданий, анализ профессиональных умений с использованием аудио- и видеотехники и др.

Работа с научной и учебной литературой

Важнейшим средством информации, распространения знаний является книга. Работа с книгой состоит в том, чтобы облегчить специалистам возможность добывать из книги необходимые знания, отобрать нужную информацию наиболее эффективно и при возможно меньших затратах времени.

Приступая к изучению дисциплины необходимо внимательно просмотреть список основной и дополнительной литературы, определить круг поиска нужной информации. Если книг на одну тему несколько, то необходимо, прежде всего, просмотреть их, ознакомиться с оглавлением, содержанием предисловием, аннотацией или введением, характером и стилем изложения материала. Выбор необходимой литературы и периодики осуществляется самостоятельно, так как даже опытный библиограф не в состоянии учесть индивидуальные интересы.

Обучающийся должен внимательно изучить электронные каталоги и картотеки. Лаконичные

каталожные карточки несут богатую информацию: фамилия автора, название книги, его подзаголовок, научное учреждение, подготовившее издание, название издательства, год выхода книги, количество страниц. Обязательный справочный материал поможет вам в подборе необходимой литературы.

Изучение книги целесообразно начинать с предварительного знакомства с ней: просмотреть введение, оглавление, заключение, библиографию или список использованной литературы. Во введении или предисловии автор обычно формулирует задачи, которые ставятся в книге. Внимательно изучив оглавление, студент узнает общий план книги, содержание ее, а в научных трудах и основные мысли автора. К оглавлению полезно обращаться не только при предварительном знакомстве с книгой, но и в процессе повторного и выборочного чтения, завершения его.

После предварительного знакомства с книгой следует приступить к первому чтению, главная цель которого - понять содержание в целом. Это предварительное чтение - знакомство с книгой и выделение в ней всего того, что наиболее существенно и требует детальной проработки в другое время.

Следующим этапом является повторное чтение или чтение с проработкой материала - это критический разбор читаемого с целью глубокого проникновения в его сущность, конспектирования.

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства и информационных справочных систем (при необходимости).

11.1 Перечень лицензионного программного обеспечения

1. Kaspersky Total Security - Антивирус

11.3 Перечень программного обеспечения отечественного производства

1. Kaspersky Total Security - Антивирус

При осуществлении образовательного процесса студентами и преподавателем используются следующие информационно справочные системы: СПС «Консультант плюс», СПС «Гарант».

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

№ п/п	Наименование специальных помещений и помещений для самостоятельной работы	Номер аудитор ии	Оснащенность специальных помещений и помещений для самостоятельной работы
1	Учебная аудитория для проведения занятий всех типов (в т.ч. лекционного, семинарского, практической подготовки обучающихся), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации		
		106/ЭФ	Оснащение: специализированная мебель на 27 посадочных мест, рабочие станции 14 шт., проектор Epson EB-X18 – 1 шт., учебно-наглядные пособия в виде презентаций, подключение к сети «Интернет», доступ в электронную информационно-образовательную среду университета, выход в корпоративную сеть университета.

2	Помещение для самостоятельной работы обучающихся, подтверждающее наличие материально-технического обеспечения, с перечнем основного оборудования		
---	--	--	--

13. Особенности реализации дисциплины лиц с ограниченными возможностями здоровья

Обучающимся с ограниченными возможностями здоровья предоставляются специальные учебники и учебные пособия, иная учебная литература, специальные технические средства обучения коллективного и индивидуального пользования, предоставление услуг ассистента (помощника), оказывающего обучающимся необходимую техническую помощь, а также услуги сурдопереводчиков и тифлосурдопереводчиков.

а) для слабовидящих:

- на промежуточной аттестации присутствует ассистент, оказывающий студенту необходимую техническую помощь с учетом индивидуальных особенностей (он помогает занять рабочее место, передвигаться, прочитать и оформить задание, в том числе записывая под диктовку);

- задания для выполнения, а также инструкция о порядке проведения промежуточной аттестации оформляются увеличенным шрифтом;

- задания для выполнения на промежуточной аттестации зачитываются ассистентом;

- письменные задания выполняются на бумаге, надиктовываются ассистенту;

- обеспечивается индивидуальное равномерное освещение не менее 300 люкс;

- студенту для выполнения задания при необходимости предоставляется увеличивающее устройство;

в) для глухих и слабослышащих:

- на промежуточной аттестации присутствует ассистент, оказывающий студенту необходимую техническую помощь с учетом индивидуальных особенностей (он помогает занять рабочее место, передвигаться, прочитать и оформить задание, в том числе записывая под диктовку);

- промежуточная аттестация проводится в письменной форме;

- обеспечивается наличие звукоусиливающей аппаратуры коллективного пользования, при необходимости поступающим предоставляется звукоусиливающая аппаратура индивидуального пользования;

- по желанию студента промежуточная аттестация может проводиться в письменной форме;

д) для лиц с нарушениями опорно-двигательного аппарата (тяжелыми нарушениями двигательных функций верхних конечностей или отсутствием верхних конечностей):

- письменные задания выполняются на компьютере со специализированным программным обеспечением или надиктовываются ассистенту;

- по желанию студента промежуточная аттестация проводится в устной форме.

Рабочая программа дисциплины «Анализ данных» составлена на основе Федерального государственного образовательного стандарта высшего образования - бакалавриат по направлению подготовки 38.03.02 Менеджмент (приказ Минобрнауки России от 12.08.2020 г. № 970).

Автор (ы)

_____ ст. преп. КИИТ, Огур Максим Геннадьевич

Рецензенты

_____ доц. КИИТ, дэн Тамбиева Джаннет Алиевна

_____ доц. КИИТ, кэн Сорокин Анатолий Александрович

Рабочая программа дисциплины «Анализ данных» рассмотрена на заседании Кафедры инжиниринга IT-решений протокол № от г. и признана соответствующей требованиям ФГОС ВО и учебного плана по направлению подготовки 38.03.02 Менеджмент

Заведующий кафедрой _____ Шлаев Дмитрий Валерьевич

Рабочая программа дисциплины «Анализ данных» рассмотрена на заседании учебно-методической комиссии Институт экономики, финансов и управления в АПК протокол № от г. и признана соответствующей требованиям ФГОС ВО и учебного плана по направлению подготовки 38.03.02 Менеджмент

Руководитель ОП _____