

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И. Т. ТРУБИЛИНА»**

ФАКУЛЬТЕТ ПРИКЛАДНОЙ ИНФОРМАТИКИ

УТВЕРЖДАЮ

Декан факультета
прикладной информатики
профессор



С.А. Курносов

27 марта 2020 г.

Рабочая программа дисциплины

Анализ данных
наименование дисциплины

Направление подготовки
38.03.05 Бизнес-информатика
шифр и наименование направления подготовки

Направленность
Архитектура предприятия
наименование направленности подготовки

Уровень высшего образования
Бакалавриат

Форма обучения
очная

Краснодар
2020

Рабочая программа дисциплины «Анализ данных» разработана на основе ФГОС ВО 38.03.05 Бизнес-информатика утвержденного приказом Министерства образования и науки РФ 11 августа 2016 г. № 1002.

Автор:

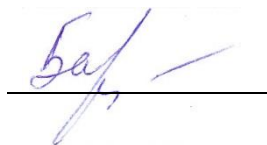
к.ф.-м.н., доцент



Д.А. Павлов

Рабочая программа обсуждена и рекомендована к утверждению решением кафедры системного анализа и обработки информации от 16.03.2020 г., протокол №8

Заведующий кафедрой
Д.э.н., профессор



Т.П. Барановская

Рабочая программа одобрена на заседании методической комиссии факультета прикладной информатики, протокол №7 от 27.03.2020.

Председатель
методической комиссии
к.п.н., доцент



Т.А. Крамаренко

Руководитель
основной профессиональной
образовательной программы
к.э.н., доцент



А.Е.Вострокнутов

1 Цель и задачи освоения дисциплины

Целью освоения дисциплины «Анализ данных» является изучение теоретических основ и методов анализа данных, применяемых при решении прикладных задач и использования их в современных информационных системах.

Задачи

- изучение методов статистического анализа данных;
- изучение методов интеллектуального анализа данных;
- изучение принципов организации и проведения аналитического исследования;
- решение практических задач, прогнозирование и выработка рекомендаций.

2 Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОПОП ВО

В результате освоения дисциплины формируются следующие компетенции:

ОПК-1 – способностью решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности

ПК-17 – способность использовать основные методы естественнонаучных дисциплин в профессиональной деятельности для теоретического и экспериментального исследования;

ПК-18 – способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования.

3 Место дисциплины в структуре ОПОП ВО

«Анализ данных» является дисциплиной базовой части ОПОП ВО подготовки обучающихся по направлению 38.03.05 «Бизнес-информатика», направленность «Архитектура предприятия».

4 Объем дисциплины (144 часа, 4 зачетные единицы)

Виды учебной работы	Объем, часов	
	Очная	Заочная
Контактная работа	65	-
в том числе:		
— аудиторная по видам учебных занятий	62	-
— лекции	32	-
— практические	-	-
— лабораторные	30	-
— внеаудиторная	3	-
— зачет		-
— экзамен	3	-
— защита курсовых работ		-
Самостоятельная работа	79	-
в том числе:		
— курсовая работа	-	-
— прочие виды самостоятельной работы	79	-
Итого по дисциплине	144	-

5 Содержание дисциплины

По итогам изучаемого курса студенты сдают экзамен.
Дисциплина изучается на 3 курсе, в 5 семестре.

Содержание и структура дисциплины по очной форме обучения

№ п/п	Наименование темы с указанием основных вопросов	Формируемые компетенции	Семестр	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)			
				Лекции	Практические занятия	Лабораторные занятия	Самостоятельная работа
1	Основы Python и анализа	ОПК-	5	4	-	6	12

№ п/п	Наименование темы с указанием основных вопросов	Формируемые компетенции	Семестр	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)			
				Лекции	Практические занятия	Лабораторные занятия	Самостоятельная работа
	данных Введение и основы синтаксиса. Списки и циклы. Операции с таблицами. Условия и функции. Библиотека Pandas для анализа данных. Этапы анализа данных. Предварительная обработка данных. Анализ данных и оформление результатов. Понятие и методы «Big Data»	1, ПК-17; ПК-18					
2	Предобработка данных Работа с пропусками. Изменение типов данных. Поиск дубликатов. Категоризация данных.	ОПК-1, ПК-17; ПК-18	5	2	-	4	15
3	Визуализация данных и исследовательский анализ данных Визуализация данных с помощью гистограмм и ящиков с усами. Изучение срезов данных. Нахождение взаимосвязей разных параметров в данных. Объединение таблиц. Получение выводов по сгруппированному данным.	ОПК-1, ПК-17; ПК-18	5	4	-	6	16
4	Статистический анализ данных Изучение объектов и их взаимосвязей методами статистики. Выборки и статистическая значимость. Выявление и обработка аномалий.	ОПК-1, ПК-17; ПК-18	5	6	-	8	16
5	Основы машинного обучения Модели машинного обучения. Оценки качества обучения.	ОПК-1, ПК-17; ПК-18	5	16	-	6	20
Итого				32	-	30	79

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

Методические указания (для самостоятельной работы)

1. Грубич Т.Ю. Анализ данных : практикум / сост. Т. Ю. Грубич, Д. А. Павлов. - Краснодар: КубГАУ, 2015 – 193 с., 150 экз., http://edu.kubsau.ru/file.php/118/01_Metodichka_AD_2015_peredelannaja1.pdf

2. Павлов Д.А. Анализ данных : метод. рекомендации по организации контактной и самостоятельной работы / сост. Д. А. Павлов. – Краснодар : КубГАУ, 2020. – 31 с. – Режим доступа: http://edu.kubsau.ru/file.php/118/MRpoSR_AnalizDannyykh_novaja_.pdf

7 Фонд оценочных средств для проведения промежуточной аттестации

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

*Этап формирования компетенции соответствует номеру семестра

Номер семестра*	Этапы формирования и проверки уровня сформированности компетенций по дисциплинам, практикам в процессе освоения ОПОП ВО
<i>ОПК-1 способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности</i>	
1	Информатика
1, 2	Программирование
2	Электронная коммуникация
2	Программные и аппаратные средства информатики
2	Информационные технологии поддержки личной работы
2	Практика по получению первичных профессиональных умений и навыков, в том числе первичных умений и навыков научно-исследовательской деятельности
3	Объектно-ориентированное программирование
3	Информационные технологии
5	Базы данных
5	Анализ данных
6	Общая теория систем
6	Имитационное моделирование
7	Архитектура предприятия
7	Управление ИТ-сервисами и контентом
8	Электронный бизнес

Номер семестра*	Этапы формирования и проверки уровня сформированности компетенций по дисциплинам, практикам в процессе освоения ОПОП ВО
8	Информационная безопасность
8	Защита выпускной квалификационной работы, включая подготовку к защите и процедуру защиты
<i>ПК-17 - способность использовать основные методы естественнонаучных дисциплин в профессиональной деятельности для теоретического и экспериментального исследования</i>	
1	Дискретная математика
1	Линейная алгебра и аналитическая геометрия
12	Программирование
2	Математический анализ
2	Теория вероятностей и математическая статистика
3	Дифференциальные и разностные уравнения
3	Элементы теории нечетких множеств
3	Алгоритмы и структуры данных
4	Математическая экономика
4	Научно-исследовательская работа
4	Системы компьютерной математики
5	Анализ данных
5	Исследование операций
6	Общая теория систем
6	Имитационное моделирование
6	Системный анализ
6	Анализ сложных систем
6	Основы финансовых вычислений
6	Информационные системы в финансово-кредитной сфере
7	Инженерия знаний и интеллектуальные системы
8	Преддипломная практика
8	Защита выпускной квалификационной работы, включая подготовку к защите и процедуру защиты
<i>ПК-18 способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования</i>	
1	Дискретная математика
1	Линейная алгебра и аналитическая геометрия
1, 2	Программирование
2	Математический анализ
2	Теория вероятностей и математическая статистика
2	Практика по получению первичных профессиональных умений и навыков, в том числе первичных умений и навыков научно-исследовательской деятельности
3	Дифференциальные и разностные уравнения
3	Объектно-ориентированное программирование
3	Элементы теории нечетких множеств
3	Алгоритмы и структуры данных
4	Математическая экономика
4	Бухгалтерский и управленческий учет
4	Научно-исследовательская работа
4	Системы компьютерной математики

Номер семестра*	Этапы формирования и проверки уровня сформированности компетенций по дисциплинам, практикам в процессе освоения ОПОП ВО
5	Анализ данных
5	Исследование операций
5	Разработка приложений в среде Microsoft Office
5	Компьютерная графика
6	Общая теория систем
6	Имитационное моделирование
6	Разработка бизнес-приложений
6	WEB-программирование
6	Системный анализ
6	Анализ сложных систем
6	Основы финансовых вычислений
6	Информационные системы в финансово-кредитной сфере
7	Инженерия знаний и интеллектуальные системы
7	Современные методы и системы принятия решений
8	Информационная бизнес-аналитика
8	Разработка приложений для мобильных устройств
8	Разработка программ системного назначения
8	Преддипломная практика
8	Защита выпускной квалификационной работы, включая подготовку к защите и процедуру защиты

7.2 Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкалы оценивания

Планируемые результаты освоения компетенции (индикаторы достижения компетенции)	Уровень освоения				Оценочное средство
	неудовлетворительно (минимальный не достигнут)	удовлетворительно (минимальный, пороговый)	хорошо (средний)	отлично (высокий)	
<i>ОПК-1 способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности</i>					
Знать: - основные принципы применения ИКТ в профессиональной деятельности - основные требования по организации защиты информации	Уровень знаний ниже минимальных требований, имели место грубые ошибки	Минимально допустимый уровень знаний, допущено много грубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки, допущено несколько грубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок	Реферат, тест, лабораторная работа, кейс-задание, экзамен (вопросы и задания)
Уметь: - применять на практике основные	При решении стандартных задач не продемонстрированы	Продемонстрированы основные умения, ре-	Продемонстрированы все основные умения, ре-	Продемонстрированы все основные умения, ре-	

Планируемые результаты освоения компетенции (индикаторы достижения компетенции)	Уровень освоения				Оценочное средство
	неудовлетворительно (минимальный не достигнут)	удовлетворительно (минимальный, пороговый)	хорошо (средний)	отлично (высокий)	
<p>принципы применения ИКТ в профессиональной деятельности</p> <ul style="list-style-type: none"> - перечислять и давать общую характеристику видов и источников угроз безопасности; - оценивать источники угроз информационной безопасности для различных профессиональных областей; - использовать современные средства защиты информации 	<p>ваны основные умения, имели место грубые ошибки</p>	<p>шены типовые задачи с негрубыми ошибками, выполнены все задания, но не в полном объеме</p>	<p>шены все основные задачи с негрубыми ошибками, выполнены все задания в полном объеме, но некоторые с недочетами</p>	<p>шены все основные задачи с отдельными незначительными недочетами, выполнены все задания в полном объеме</p>	
<p>Владеть:</p> <ul style="list-style-type: none"> - технологиями использования современных ИКТ в рамках профессиональной деятельности - современными технологиями и средствами защиты информации 	<p>При решении стандартных задач не продемонстрированы базовые навыки, имели место грубые ошибки</p>	<p>Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами</p>	<p>Продемонстрированы базовые навыки при решении стандартных задач с некоторыми недочетами</p>	<p>Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов</p>	
<p>ПК-17 – способность использовать основные методы естественнонаучных дисциплин в профессиональной деятельности для теоретического и экспериментального исследования</p>					
<p>Знать:</p> <ul style="list-style-type: none"> - основные методы естественнонаучных дисциплин с целью 	<p>Уровень знаний ниже минимальных требований, имели место грубые ошибки</p>	<p>Минимально допустимый уровень знаний, допущено много негрубых оши-</p>	<p>Уровень знаний в объеме, соответствующем программе подготовки, до-</p>	<p>Уровень знаний в объеме, соответствующем программе подготовки, без</p>	<p>Реферат, тест, лабораторная работа, кейс-задание, экзамен (вопросы и задания)</p>

Планируемые результаты освоения компетенции (индикаторы достижения компетенции)	Уровень освоения				Оценочное средство
	неудовлетворительно (минимальный не достигнут)	удовлетворительно (минимальный, пороговый)	хорошо (средний)	отлично (высокий)	
их использования в профессиональной деятельности для теоретического и экспериментального исследования	ки	бок	пущено несколько негрубых ошибок	ошибок	
Уметь: - использовать основные методы естественнонаучных дисциплин в профессиональной деятельности для теоретического и экспериментального исследования	При решении стандартных задач не продемонстрированы основные умения, имели место грубые ошибки	Продемонстрированы основные умения, решены типовые задачи с негрубыми ошибками, выполнены все задания, но не в полном объеме	Продемонстрированы все основные умения, решены все основные задачи с негрубыми ошибками, выполнены все задания в полном объеме, но некоторые с недочетами	Продемонстрированы все основные умения, решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме	
Владеть: - практическими навыками использования основных методов естественнонаучных дисциплин в профессиональной деятельности для теоретического и экспериментального исследования.	При решении стандартных задач не продемонстрированы базовые навыки, имели место грубые ошибки	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами	Продемонстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов	
ПК 18 – способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования					
Знать: - современный математический аппарат и инструментальные средства для обработ-	Уровень знаний ниже минимальных требований, имели место грубые ошибки	Минимально допустимый уровень знаний, допущено много негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки, допущено несколько не-	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок	Реферат, тест, лабораторная работа, кейс-задание, экзамен (вопросы и задания)

Планируемые результаты освоения компетенции (индикаторы достижения компетенции)	Уровень освоения				Оценочное средство
	неудовлетворительно (минимальный не достигнут)	удовлетворительно (минимальный, пороговый)	хорошо (средний)	отлично (высокий)	
ки, анализа и систематизации информации.			грубых ошибок		
Уметь: - использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования.	При решении стандартных задач не продемонстрированы основные умения, имели место грубые ошибки	Продемонстрированы основные умения, решены типовые задачи с негрубыми ошибками, выполнены все задания, но не в полном объеме	Продемонстрированы все основные умения, решены все основные задачи с негрубыми ошибками, выполнены все задания в полном объеме, но некоторые с недочетами	Продемонстрированы все основные умения, решены все основные задачи с отдельными незначительными недочетами, выполнены все задания в полном объеме	
Владеть: - практическими навыками использования математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования.	При решении стандартных задач не продемонстрированы базовые навыки, имели место грубые ошибки	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами	Продемонстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов	

7.3 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, характеризующих этапы формирования компетенций в процессе освоения ОПОП ВО

Лабораторные работы

1. Грубич Т.Ю. Анализ данных : практикум / сост. Т. Ю. Грубич, Д. А. Павлов. - Краснодар: КубГАУ, 2015 – 193 с., 150 экз., http://edu.kubsau.ru/file.php/118/01_Metodichka_AD_2015_peredelannaja1.pdf

Примеры тестов

№1 (Балл 1)

Совокупность сведений, подлежащих хранению, передаче обработке и использованию в человеческой деятельности – это...

- 1 данные
- 2 обработка информации
- 3 информация

№2 (1)

Средняя вероятность доступности данных в произвольный момент времени – это признак классификации:

- 1 по работе с конфиденциальной информацией
- 2 по уровню конфиденциальности
- 3 по требуемой степени безотказности

№3 (1)

Как переводится Knowledge Discovery in Databases (KDD)

- 1 «раскопка» данных
- 2 извлечение знаний из баз данных
- 3 тиражирование знаний

№4 (1)

Укажите основные виды моделей:

- 1 эмпирические
- 2 теоретические
- 3 смешанные
- 4 статистические
- 5 экспериментальные

№5 (1)

С какими формами данных, как правило, работают методы анализа данных:

- 1 структурированные и слабоструктурированные
- 2 неструктурированные
- 3 структурированные
- 4 слабоструктурированные
- 5 неструктурированные и слабоструктурированные

Примеры тем рефератов

1. ABC-анализ
2. Анализ многомерных данных
3. Подготовка данных к анализу
4. Проблемы измерения, возникающие при выборе способа анализа данных
5. Комплексный оперативный анализ информации хранилищ данных

Пример кейс-задания

Принять участие в соревновании Kaggle (<https://www.kaggle.com/c/house-prices-advanced-regression-techniques>) о предсказании цен на недвижимость.

Задание:

Скачать датасет;

Использовать любую модель (например, линейную регрессию) для построения предсказания.

Вопросы и задания для проведения промежуточного контроля (экзамена)

Компетенция: ОПК-1 – способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности»

Вопросы к экзамену

1. Понятие анализа данных. Методология анализа данных.
2. Подходы к моделированию.
3. Принципы анализа данных.
4. Типы данных. Методы анализа данных.
5. Этапы статистического исследования.
6. Задачи и методы статистического анализа данных. Постановка задачи классификации и представление результатов
7. Методы построения правил классификации
8. Алгоритм построения 1-правил для классификации
9. Метод Naïve Bayes
10. Методы построения деревьев решений
11. Методика "Разделяй и властвуй"
12. Алгоритм ID3
13. Алгоритм C4.5
14. Алгоритм покрытия
15. Методы построения математических функций
16. Корреляционный анализ Для двух переменных
17. Корреляционный анализ Для произвольного числа переменных
18. Регрессионный анализ
19. Метод наименьших квадратов
20. Нелинейные методы

Компетенция: «ПК-17 – способность использовать основные методы естественнонаучных дисциплин в профессиональной деятельности для теоретического и экспериментального исследования»

Вопросы к экзамену

1. Метод опорных векторов
2. Ядра
3. Поиск ассоциативных правил формальная постановка задачи

4. Представление результатов ассоциативных правил
5. Алгоритм Apriori
6. Свойство анти-монотонности
7. Описание алгоритма Apriori
8. Секвенциальный анализ Постановка задачи
9. Алгоритм AprioriALL
10. Алгоритм GSP Генерация кандидатов
11. Алгоритм GSP. Подсчёт поддержки кандидатов
12. Иерархия данных. Таксономия
13. Кластеризация. Типы алгоритмов
14. Классификация алгоритмов кластеризации
15. Иерархические алгоритмы
16. Представление результатов иерархического алгоритма
17. Возможные подходы к статистическому анализу данных
18. Аналитический подход к моделированию
19. Информационный подход к моделированию
20. Процесс анализа

Компетенция: «ПК-18 – способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования»

Вопросы к экзамену

1. Предобработка данных
2. Основные объекты Pandas
3. Группировка данных
4. Сводные таблицы
5. Извлечение данных по условиям
6. Агрегирующие функции
7. Объединение таблиц
8. Предварительная обработка данных
9. Визуализация данных
10. Библиотека Matplotlib для визуализации данных
11. Библиотека seaborn для визуализации данных
12. Библиотека Plotly для визуализации данных
13. Работа с библиотекой NumPy
14. Математические операции с матрицами NumPy
15. Линейная алгебра с numpy
16. Метрика качества оценки алгоритмов
17. Accuracy
18. Precision

- 19. Recall
- 20. ROC
- 21. Матрица исходов.

Практические задания для экзамена

В рамках практического задания для оценки освоения компетенций ОПК-1, ПК-17, ПК-18 обучающемуся предлагается выполнить следующее кейс-задание:

Ситуация:

В вашем распоряжении данные сервиса Яндекс.Недвижимость — архив объявлений о продаже квартир в Санкт-Петербурге и соседних населённых пунктов за несколько лет. Нужно научиться определять рыночную стоимость объектов недвижимости. Ваша задача — установить параметры. Это позволит построить автоматизированную систему: она отследит аномалии и мошенническую деятельность.

По каждой квартире на продажу доступны два вида данных. Первые вписаны пользователем, вторые — получены автоматически на основе картографических данных. Например, расстояние до центра, аэропорта, ближайшего парка и водоёма.

Задание:

Инструкция по выполнению проекта

Шаг 1. Откройте файл с данными и изучите общую информацию

Путь к файлу: /datasets/real_estate_data.csv

Шаг 2. Предобработка данных

- определите и изучите пропущенные значения:
 - о для некоторых пропущенных значений можно предположить логичную замену. Например, если человек не указал число балконов — скорее всего, их нет. Такие пропуски правильно заменить на 0. Для других типов данных нет подходящего значения на замену. В этом случае правильно оставить эти значения пустыми. Отсутствие значения — тоже важный сигнал, который не нужно прятать;
 - о заполните пропуски, где это уместно. Опишите, почему вы решили заполнить пропуски именно в этих столбцах и как выбрали значения;
 - о укажите причины, которые могли привести к пропускам в данных.
- приведите данные к нужным типам:
 - о поясните, в каких столбцах нужно изменить тип данных и почему.

Шаг 3. Посчитайте и добавьте в таблицу:

- цену квадратного метра;
- день недели, месяц и год публикации объявления;
- этаж квартиры; варианты — первый, последний, другой;

- соотношение жилой и общей площади, а также отношение площади кухни к общей.

Шаг 4. Проведите исследовательский анализ данных и выполните инструкции:

- Изучите следующие параметры: площадь, цена, число комнат, высота потолков. Постройте гистограммы для каждого параметра.

- Изучите время продажи квартиры. Постройте гистограмму. Посчитайте среднее и медиану. Опишите, сколько обычно занимает продажа. Когда можно считать, что продажи прошли очень быстро, а когда необычно долго?

- Уберите редкие и выбивающиеся значения. Опишите, какие особенности обнаружили.

- Какие факторы больше всего влияют на стоимость квартиры? Изучите, зависит ли цена от квадратного метра, числа комнат, удалённости от центра. Изучите зависимость цены от того, на каком этаже расположена квартира: первом, последнем или другом. Также изучите зависимость от даты размещения: дня недели, месяца и года.

- Выберите 10 населённых пунктов с наибольшим числом объявлений. Посчитайте среднюю цену квадратного метра в этих населённых пунктах. Выделите населённые пункты с самой высокой и низкой стоимостью жилья. Эти данные можно найти по имени в столбце 'locality_name'.

- Изучите предложения квартир: для каждой квартиры есть информация о расстоянии до центра. Выделите квартиры в Санкт-Петербурге ('locality_name'). Ваша задача — выяснить, какая область входит в центр. Создайте столбец с расстоянием до центра в километрах: округлите до целых значений. После этого посчитайте среднюю цену для каждого километра. Постройте график: он должен показывать, как цена зависит от удалённости от центра. Определите границу, где график сильно меняется — это и будет центральная зона.

- Выделите сегмент квартир в центре. Проанализируйте эту территорию и изучите следующие параметры: площадь, цена, число комнат, высота потолков. Также выделите факторы, которые влияют на стоимость квартиры (число комнат, этаж, удалённость от центра, дата размещения объявления). Сделайте выводы. Отличаются ли они от общих выводов по всему городу?

Шаг 5. Напишите общий вывод

Оформление: Выполните задание в Jupyter Notebook. Заполните программный код в ячейках типа code, текстовые пояснения — в ячейках типа markdown. Примените форматирование и заголовки.

Описание данных

- `airports_nearest` — расстояние до ближайшего аэропорта в метрах (м)
- `balcony` — число балконов
- `ceiling_height` — высота потолков (м)
- `cityCenters_nearest` — расстояние до центра города (м)

- days_exposition — сколько дней было размещено объявление (от публикации до снятия)
- first_day_exposition — дата публикации
- floor — этаж
- floors_total — всего этажей в доме
- is_apartment — апартаменты (булев тип)
- kitchen_area — площадь кухни в квадратных метрах (м²)
- last_price — цена на момент снятия с публикации
- living_area — жилая площадь в квадратных метрах(м²)
- locality_name — название населённого пункта
- open_plan — свободная планировка (булев тип)
- parks_around3000 — число парков в радиусе 3 км
- parks_nearest — расстояние до ближайшего парка (м)
- ponds_around3000 — число водоёмов в радиусе 3 км
- ponds_nearest — расстояние до ближайшего водоёма (м)
- rooms — число комнат
- studio — квартира-студия (булев тип)
- total_area — площадь квартиры в квадратных метрах (м²)
- total_images — число фотографий квартиры в объявлении

Пояснение: апартаменты — это нежилые помещения, не относящиеся к жилому фонду, но имеющие необходимые условия для проживания.

7.4 Методические материалы, определяющие процедуры оценивания знаний, умений и навыков и опыта деятельности, характеризующих этапы формирования компетенций

Контроль освоения дисциплины «Анализ данных» проводится в соответствии Положением системы менеджмента качества КубГАУ 2.5.1 «Текущий контроль успеваемости и промежуточная аттестация обучающихся».

Критерии оценки лабораторной работы

Оценка **«отлично»** выставляется в том случае, когда обучающийся правильно и полностью выполнил основное задание и, возможно при необходимости, дополнительное задание лабораторной работы, ответил правильно на теоретические вопросы, на дополнительные вопросы. Показал отличные знания и умения при выполнении лабораторной работы в рамках учебного материала.

Оценка **«хорошо»** выставляется в том случае, когда обучающийся правильно и полностью выполнил задание лабораторной работы, ответил на теоретические вопросы с небольшими неточностями, на большинство дополнительных вопросов также, возможно, допуская незначительные ошибки. Показал достаточно хорошие знания и умения при выполнении лабораторной работы в рамках учебного материала.

Оценка **«удовлетворительно»** выставляется в том случае, когда обучающийся правильно выполнил задание лабораторной работы, ответил на

теоретические вопросы с существенными неточностями. Показал минимальные удовлетворительные знания и умения при выполнении лабораторной работы в рамках учебного материала.

Оценка **«неудовлетворительно»** выставляется в том случае, когда обучающийся неправильно выполнил задание лабораторной работы, не ответил на теоретические вопросы. Показал недостаточный уровень знаний и умений при выполнении лабораторной работы в рамках учебного материала.

Критерии оценки при проведении тестирования

Оценка **«отлично»** выставляется при условии правильного ответа студента не менее чем 85 % тестовых заданий;

Оценка **«хорошо»** выставляется при условии правильного ответа студента не менее чем 70 % тестовых заданий;

Оценка **«удовлетворительно»** выставляется при условии правильного ответа студента не менее 51 %; .

Оценка **«неудовлетворительно»** выставляется при условии правильного ответа студента менее чем на 50 % тестовых заданий.

Критерии оценки реферата

Критериями оценки реферата являются: новизна текста, обоснованность выбора источников литературы, степень раскрытия сущности вопроса, соблюдения требований к оформлению.

Оценка **«отлично»** — выполнены все требования к написанию реферата: обозначена проблема и обоснована её актуальность; сделан анализ различных точек зрения на рассматриваемую проблему и логично изложена собственная позиция; сформулированы выводы, тема раскрыта полностью, выдержан объём; соблюдены требования к внешнему оформлению.

Оценка **«хорошо»** — основные требования к реферату выполнены, но при этом допущены недочёты. В частности, имеются неточности в изложении материала; отсутствует логическая последовательность в суждениях; не выдержан объём реферата; имеются упущения в оформлении.

Оценка **«удовлетворительно»** — имеются существенные отступления от требований к реферированию. В частности: тема освещена лишь частично; допущены фактические ошибки в содержании реферата; отсутствуют выводы.

Оценка **«неудовлетворительно»** — тема реферата не раскрыта, обнаруживается существенное непонимание проблемы или реферат не представлен вовсе.

Критерии оценки кейс-задания

Результат выполнения кейс-задания оценивается с учетом следующих критериев:

- полнота проработки ситуации;
- полнота выполнения задания;

- новизна и неординарность представленного материала и решений;
- перспективность и универсальность решений;
- умение аргументировано обосновать выбранный вариант решения.

Если результат выполнения кейс-задания соответствует обозначенному критерию, обучающемуся присваивается один балл (за каждый критерий по 1 баллу).

Оценка «отлично» – при наборе 5 баллов.

Оценка «хорошо» – при наборе 4 баллов.

Оценка «удовлетворительно» – при наборе 3 баллов.

Оценка «неудовлетворительно» – при наборе 2 баллов.

Критерии оценки знаний при проведении экзамена

Оценка «отлично» выставляется обучающемуся, который обладает всесторонними, систематизированными и глубокими знаниями материала учебной программы, умеет свободно выполнять задания, предусмотренные учебной программой, усвоил основную и ознакомился с дополнительной литературой, рекомендованной учебной программой. Как правило, оценка «отлично» выставляется обучающемуся усвоившему взаимосвязь основных положений и понятий дисциплины в их значении для приобретаемой специальности, проявившему творческие способности в понимании, изложении и использовании учебного материала, правильно обосновывающему принятые решения, владеющему разносторонними навыками и приемами выполнения практических работ.

Оценка «хорошо» выставляется обучающемуся, обнаружившему полное знание материала учебной программы, успешно выполняющему предусмотренные учебной программой задания, усвоившему материал основной литературы, рекомендованной учебной программой. Как правило, оценка «хорошо» выставляется обучающемуся, показавшему систематизированный характер знаний по дисциплине, способному к самостоятельному пополнению знаний в ходе дальнейшей учебной и профессиональной деятельности, правильно применяющему теоретические положения при решении практических вопросов и задач, владеющему необходимыми навыками и приемами выполнения практических работ.

Оценка «удовлетворительно» выставляется обучающемуся, который показал знание основного материала учебной программы в объеме, достаточном и необходимым для дальнейшей учебы и предстоящей работы по специальности, справился с выполнением заданий, предусмотренных учебной программой, знаком с основной литературой, рекомендованной учебной программой. Как правило, оценка «удовлетворительно» выставляется обучающемуся, допустившему погрешности в ответах на экзамене или выполнении экзаменационных заданий, но обладающему необходимыми знаниями под руководством преподавателя для устранения этих погрешностей, нарушающему последовательность в изложении учебного материала и испытывающему затруднения при выполнении практических работ.

Оценка «неудовлетворительно» выставляется обучающемуся, не знающему основной части материала учебной программы, допускающему принципиальные ошибки в выполнении предусмотренных учебной программой заданий, неуверенно с большими затруднениями выполняющему практические работы. Как правило, оценка «неудовлетворительно» выставляется обучающемуся, который не может продолжить обучение или приступить к деятельности по специальности по окончании университета без дополнительных занятий по соответствующей дисциплине.

8 Перечень основной и дополнительной литературы

Основная учебная литература:

1. Анализ данных : учеб. пособие / Т.А. Крамаренко, Т.Ю. Грубич, Д.А. Павлов, Т.В. Лукьяненко. – Краснодар : КубГАУ, 2018. – 126 с. Режим доступа:

https://edu.kubsau.ru/file.php/118/Uch_posobie_AnalizDannyykh_2017_na_pечат_368337_v1_.PDF

2. Федин Ф.О. Анализ данных. Часть 1. Подготовка данных к анализу [Электронный ресурс]: учебное пособие/ Федин Ф.О., Федин Ф.Ф.— Электрон. текстовые данные.— М.: Московский городской педагогический университет, 2012.— 204 с.— Режим доступа: <http://www.iprbookshop.ru/26444>.— ЭБС «IPRbooks», по паролю.

3. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining [Электронный ресурс]: учебное пособие/ Федин Ф.О., Федин Ф.Ф.— Электрон. текстовые данные.— М.: Московский городской педагогический университет, 2012.— 308 с.— Режим доступа: <http://www.iprbookshop.ru/26445>.— ЭБС «IPRbooks», по паролю.

Дополнительная учебная литература:

1. Жуковский О.И. Информационные технологии и анализ данных [Электронный ресурс] : учебное пособие / О.И. Жуковский. — Электрон. текстовые данные. — Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент, 2014. — 130 с. — 978-5-4332-0158-3. — Режим доступа: <http://www.iprbookshop.ru/72106>

2. Маккинли, Уэс Python и анализ данных / Уэс Маккинли ; перевод А. Слинкина. — 2-е изд. — Саратов : Профобразование, 2019. — 482 с. — ISBN 978-5-4488-0046-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/88752.html>

3. Нестеров, С. А. Интеллектуальный анализ данных средствами MS SQL Server 2008 / С. А. Нестеров. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 303 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/62813.html>

9 Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Перечень ЭБС

№	Наименование	Тематика	Ссылка
1.	IPRbook	Универсальная	http://www.iprbookshop.ru/
2.	Образовательный портал КубГАУ	Универсальная	https://edu.kubsau.ru/

Рекомендуемые интернет сайты:

– образовательный портал КубГАУ [Электронный ресурс]: Режим доступа: <http://edu.kubsau.local>

– материалы Национального Открытого Университета «Интуит» [Электронный ресурс]: Режим доступа: <http://www.intuit.ru>

10 Методические указания для обучающихся по освоению дисциплины

К нормативно-методическим документам и материалам, обеспечивающим качество подготовки обучающихся, относятся:

Пл КубГАУ 2.2.1 «Рабочая программа дисциплины, практики».

Пл КубГАУ 2.5.1 «Текущий контроль успеваемости и промежуточная аттестация обучающихся».

МИ КубГАУ 2.5.2 «Критерии оценки качества занятий».

Пл КубГАУ 2.5.4 «Контактная работа обучающихся с педагогическими работниками организации и (или) лицами, привлекаемыми организацией к реализации образовательных программ на иных условиях».

Перечень учебно-методической литературы по освоению дисциплины:

1. Грубич Т.Ю. Анализ данных : практикум / сост. Т. Ю. Грубич, Д. А. Павлов. - Краснодар: КубГАУ, 2015 – 193 с., 150 экз., http://edu.kubsau.ru/file.php/118/01_Metodichka_AD_2015_peredelannaja1.pdf

2. Павлов Д.А. Анализ данных : метод. рекомендации по организации контактной и самостоятельной работы / сост. Д. А. Павлов. – Краснодар : КубГАУ, 2020. – 31 с. – Режим доступа: https://edu.kubsau.ru/file.php/118/MRpoSR_AnalizDannyykh_novaja_.pdf

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем

Информационные технологии, используемые при осуществлении образовательного процесса по дисциплине позволяют: обеспечить взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети "Интернет"; фиксировать ход образовательного процесса, результатов промежуточной аттестации по дисциплине и результатов освоения образовательной программы; организовать процесс образования путем визуализации изучаемой информации посредством использования презентационных технологий; контролировать результаты обучения на основе компьютерного тестирования.

11.1 Перечень программного обеспечения

№	Наименование	Краткое описание
1	Windows	Операционная система
2	Office	Пакет офисных приложений
3	INDIGO	Система для проведения тестирования обучающихся
4	Anaconda 3	Система для работы с анализом данных (бесплатно для индивидуального использования)

11.2 Перечень современных профессиональных баз данных, информационных справочных и поисковых систем

№	Наименование	Тематика	Электронный адрес
1.	Научная электронная библиотека «eLIBRARY.RU»	Универсальная	https://elibrary.ru

11.3 Доступ к сети Интернет

Доступ к сети Интернет, доступ в электронную информационно-образовательную среду университета

12 Материально-техническое обеспечение для обучения по дисциплине

Планируемые помещения для проведения всех видов учебной деятельности

№ п/п	Наименование учебных предметов, курсов, дисциплин (модулей), практики, иных видов учебной деятельности, предусмотренных учебным планом образовательной программы	Наименование помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом, в том числе помещения для самостоятельной работы, с указанием перечня основного оборудования, учебно-наглядных пособий и используемого программного обеспечения	Адрес (местоположение) помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом (в случае реализации образовательной программы в сетевой форме дополнительно указывается наименование организации, с которой заключен договор)
1	2	3	4
1	Анализ данных	<p>Помещение №216 ЭК, посадочных мест — 30; площадь — 41,8кв.м; учебная аудитория для проведения занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.</p> <p>Технические средства обучения (компьютер персональный — 10 шт.); доступ к сети «Интернет»; доступ в электронную информационно-образовательную среду университета; специализированная мебель (учебная доска, учебная мебель) программное обеспечение: Windows, Office, INDIGO. Anaconda 3</p> <p>Помещение №215 ЭК, посадочных мест — 20; площадь — 44кв.м; учебная аудитория для проведения занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.</p> <p>Технические средства обучения (компьютер персональный — 10 шт.); доступ к сети «Интернет»; доступ в электронную информационно-образовательную среду университета; специализированная мебель (учебная доска, учебная мебель) программное обеспечение: Win-</p>	350044, Краснодарский край, г. Краснодар, ул. им. Калинина, 13

		<p>dows, Office, INDIGO. Anaconda 3</p> <p>Помещение №5 ЭК, посадочных мест — 20; площадь — 40,6кв.м; учебная аудитория для проведения занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.</p> <p>кондиционер — 1 шт.; технические средства обучения (сетевое оборудование — 1 шт.; компьютер персональный — 9 шт.); доступ к сети «Интернет»; доступ в электронную информационно-образовательную среду университета; специализированная мебель (учебная доска, учебная мебель) программное обеспечение: Windows, Office, INDIGO. Anaconda 3</p> <p>Помещение №310 ЭК, посадочных мест — 167; площадь — 157,1кв.м; учебная аудитория для проведения занятий лекционного типа.</p> <p>Сплит-система — 1 шт.; лабораторное оборудование (плеер — 1 шт.); специализированная мебель(учебная доска, учебная мебель); технические средства обучения, наборы демонстрационного оборудования и учебно-наглядных пособий (ноутбук, проектор, экран); программное обеспечение: Windows, Office.</p> <p>Помещение №310 ЭК, площадь — 3,6кв.м; помещение для хранения и профилактического обслуживания учебного оборудования.</p> <p>Лабораторное оборудование (плеер — 1 шт.); технические средства обучения (сетевое оборудование — 1 шт.; акустическая система — 1 шт.; микрофон — 2 шт.).</p> <p>Помещение №112 ЗР, посадочных мест — 96; площадь — 49,7кв.м; учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего</p>	
--	--	--	--

		<p>контроля и промежуточной аттестации.</p> <p>Специализированная мебель (учебная доска, учебная мебель); технические средства обучения, наборы демонстрационного оборудования и учебно-наглядных пособий (ноутбук, проектор, экран); программное обеспечение: Windows, Office.</p> <p>Помещение №4 ЭК, площадь — 31,1 кв.м; помещение для хранения и профилактического обслуживания учебного оборудования. кондиционер — 2 шт.; лабораторное оборудование (шкаф лабораторный — 1 шт.; набор лабораторный — 1 шт.); технические средства обучения (принтер — 1 шт.; проектор — 1 шт.; микрофон — 1 шт.; ибп — 4 шт.; сервер — 1 шт.; носитель информации — 1 шт.; компьютер персональный — 15 шт.).</p> <p>Помещение №4 ЭК, площадь — 9,1 кв.м; помещение для хранения и профилактического обслуживания учебного оборудования. сплит-система — 2 шт.; штатив — 1 шт.; лабораторное оборудование (шкаф лабораторный — 2 шт.; стенд лабораторный — 4 шт.); технические средства обучения (экран — 1 шт.; сетевое оборудование — 5 шт.; сервер — 6 шт.; компьютер персональный — 2 шт.).</p>	
2	Анализ данных	<p>Помещение №206 ЭК, посадочных мест — 20; площадь — 41 кв.м; помещение для самостоятельной работы.</p> <p>Технические средства обучения (компьютер персональный — 9 шт.); доступ к сети «Интернет»; доступ в электронную информационно-образовательную среду уни-</p>	350044, Краснодарский край, г. Краснодар, ул. им. Калинина, 13

		<p>верситета;</p> <p>специализированная мебель (учебная мебель).</p> <p>Программное обеспечение: Windows, Office, специали- зированное лицензионное и свободно распространяемое программное обеспечение, предусмотренное в рабочей программе</p>	
--	--	--	--