

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

Б1.В.05

(индекс дисциплины)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Обработка и анализ данных

(наименование дисциплины)

по направлению подготовки (специальности)
38.04.01 Экономика

направленность (профиль)/специализация
Бизнес-аналитика

Форма обучения: заочная

Год набора: 2022

Общая трудоемкость: 5 ЗЕ

Распределение часов дисциплины по семестрам

Семестр	1	Итого
Форма контроля	Зачет с оценкой	
Вид занятий		
Лекции	4	4
Лабораторные		
Практические	4	4
Руководство: курсовые работы (проекты) / РГР		
Промежуточная аттестация	0,25	0,25
Контактная работа	8,25	8,25
Самостоятельная работа	168	168
Контроль	3,75	3,75
Итого	180	180

Рабочую программу составил(и):
профессор кафедры «Прикладная математика и информатика», доцент, д-р социол. наук,
Желнина Е. В.

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

Рецензирование рабочей программы дисциплины:



Отсутствует



Рецензент

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

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

38.04.01 Экономика

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

СОГЛАСОВАНО

Директор

Института финансов, экономики и управления

«__» _____ 20__ г.

(подпись)

М. О. Искосков

(И.О. Фамилия)

УТВЕРЖДЕНО

На заседании кафедры

«Прикладная математика и информатика»

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

1. Цель освоения дисциплины

Цель освоения дисциплины – формирование представления о современных методах и технологиях анализа данных, которые применяются на практике в различных областях науки и техники.

2. Место дисциплины в структуре ОПОП ВО

Данная дисциплина (учебный курс) относится к вариативным дисциплинам блока Б1 дисциплины (модули).

Дисциплины, учебные курсы, для которых необходимы знания, умения, навыки, приобретаемые в результате изучения данной дисциплины (учебного курса) – «Инвестиционный анализ и оценка проектов», «Компьютерное моделирование и планирование (продвинутый уровень)».

3. Планируемые результаты обучения

Формируемые и контролируемые компетенции (код и наименование)	Индикаторы достижения компетенций (код и наименование)	Планируемые результаты обучения
ПК-1 Способен собирать, анализировать и оценивать эффективность бизнес-анализа в организации	ПК-1.1. Проводит сбор, обработку, анализ и оценку данных	Знать: способы сбора и обработки данных, методы и технологии анализа и оценки данных
		Уметь: собирать и обрабатывать данные базовыми способами, анализировать и оценивать данные
		Владеть: навыками проведения разных видов анализа данных (корреляционный, кластерный, факторный анализ)
	ПК-1.2. Использует современные технические средства, информационно-коммуникационные технологии и программное обеспечение для сбора, обработки, анализа и оценки данных	Знать: современные технические средства, информационно-коммуникационные технологии и программное обеспечение, необходимое для осуществления процедур сбора, обработки, анализа и оценки данных
		Уметь: использовать современные технические средства, информационно-коммуникационные технологии и программное обеспечение при осуществлении процедур сбора, обработки, анализа и оценки данных
		Владеть: навыками реализации процедур сбора, обработки, анализа и оценки данных с использованием современных технических средств, информационно-коммуникационных технологий и программного обеспечения

4. Структура и содержание дисциплины

Модуль (раздел)	Вид учебной работы	Наименование тем занятий (учебной работы)	Семестр	Объем, ч.	Баллы	Интерак- тив, ч.	Формы текущего контроля (наименование оценочного средства)
Модуль 1. Основы анализа данных	Лек.	Введение в анализ данных. Современные технические средства, информационно-коммуникационные технологии и программное обеспечение для реализации процедур анализа данных. Подготовка и обработка данных для анализа	1	2	-		Собеседование Задания для оценки сформированности компетенций
	СР	Работа с лекционным материалом, подготовка задания по практической работе к защите	1	84	-		Собеседование Задания для оценки сформированности компетенций
	Пр	Практическая работа – Основы работы с Excel в процессах обработки и анализа данных – Подготовка данных для анализа – Обработка данных	1	2	-		Отчет по практической работе
Модуль 2. Методы, технологии и виды анализа данных	Лек.	Корреляционный и регрессионный анализ данных средствами Excel Кластерный и факторный анализ данных средствами Excel	1	2	-		Собеседование Задания для оценки сформированности компетенций
	СР	Работа с лекционным материалом, подготовка задания по практической работе к защите	1	84	-		Собеседование Задания для оценки сформированности компетенций
	Пр	Практическая работа – Визуализация данных средствами Excel – Исследовательский анализ данных средствами Excel	1	2	-		Отчет по практической работе
	Контроль	Зачет	1	4	-		Зачет с оценкой
Итого:				180	-		

5. Образовательные технологии

В рамках изучения дисциплины «Технологии прикладного анализа данных» предусмотрено использование следующих образовательных технологий:

- технология традиционного обучения: лекции и практические работы;
- интерактивные технологии: проблемная лекция.

6. Методические указания по освоению дисциплины

При освоении учебных тем необходимо:

- изучить учебный материал;
- акцентировать внимание на основных понятиях темы;
- понять рассмотренные в теоретическом материале основные способы работы с данными.

7. Оценочные средства

7.1. Паспорт оценочных средств

Семестр	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
1	ПК-1	Собеседование по Модулю 1, 2 Практические задания 1–5 Задания для оценки сформированности компетенций 1-50 Вопросы к зачету с оценкой (1-144)

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

7.2.1. _____ Собеседование _____ (наименование оценочного средства)

№ п/п	Вопросы для собеседования
Модуль 1	
1.	Какова роль анализа данных в современном обществе? Как связаны процесс принятия решения и анализ данных? Почему развитие анализа данных было неотъемлемой частью развития современной науки и техники?
2.	Какие виды анализа данных существуют? Охарактеризуйте каждый из них.
3.	Какие типы вероятностей были признаны с момента появления философии науки в 17 веке?
4.	Охарактеризуйте понятие «шкала измерения». Какие виды шкал измерения вы знаете. В каких случаях применяется тот или иной вид шкал измерения?
5.	Какова значимость анализа данных в различных областях науки и жизнедеятельности?
6.	Охарактеризуйте понятия выборки и генеральной совокупности. Как данные понятия взаимосвязаны?
7.	Что такое базовый период и отчетный период? Каковы отличия данных временных периодов? Приведите примеры

№ п/п	Вопросы для собеседования
8.	Как взаимосвязаны понятия надежности и достоверности? Перечислите и охарактеризуйте основные меры обеспечения надежности статистического измерения.
9.	Какие виды представления количественных группировочных признаков существует в анализе данных? Определите и охарактеризуйте их.
10.	Охарактеризуйте формулу Стерджесса как одну из часто используемых статистических процедур для реализации статистической группировки.
11.	Почему применяется оформление данных табличным способом? Какие виды статистических таблиц применяются в статистике? Перечислите и охарактеризуйте каждый из видов.
12.	Перечислите и проанализируйте преимущества диаграммного представления данных. Каковы достоинства и недостатки графического представления данных? Какие существуют рекомендации по эффективному графическому представлению данных и информации?
13.	Почему графическое визуальное представление информации является важнейшим компонентом в понимании и выявлении закономерностей и тенденций?
14.	Что представляют собой абсолютные показатели вариации? Перечислите и охарактеризуйте наиболее часто использующиеся в статистике абсолютные показатели вариации.
15.	Что представляют собой относительные показатели вариации? Перечислите и охарактеризуйте наиболее часто использующиеся в анализе данных относительные показатели вариации.
16.	Что такое структурные средние вариации? Какое значение имеют структурные средние в анализе данных?
17.	Определите понятие среднего арифметического в анализе данных. Как оно рассчитывается? Перечислите его достоинства и недостатки. Приведите примеры использования среднего арифметического в анализе данных.
18.	Определите понятие моды в анализе данных. Как она рассчитывается? Перечислите ее достоинства и недостатки. Приведите примеры использования моды в анализе данных.
19.	Как представлены данные в нормальном распределении? Какие еще названия нормального распределения существуют?
20.	Опишите эмпирическое правило (правило 68-95-99,7), которое используется в анализе данных. Для чего оно применяется?
Модуль 2	
21.	Определите понятие корреляционного анализа как статистического метода. Что позволяет выяснить корреляционный анализ?
22.	Каковы границы коэффициента корреляции? Как трактуется значение коэффициента корреляции? Что отражает знак коэффициента корреляции? Как интерпретируется значение коэффициента корреляции? Чему равно абсолютное значение числа (модуль числа) коэффициента корреляции? Что говорит абсолютное значение коэффициента корреляции о величине корреляции?
23.	Что описывает коэффициент корреляции произведение-момент Пирсона, также известный как r Пирсона? Связь каких переменных может описать коэффициент корреляции Пирсона?
24.	Что измеряет коэффициент корреляции Спирмена? Какие характеристики выборки удовлетворяют применению коэффициента Спирмена?
25.	Что получится при возведении коэффициента корреляции в квадрат? Что измеряет данный расчетный показатель? В каких границах находится данный коэффициент?

№ п/п	Вопросы для собеседования
26.	Что получится при вычитании коэффициента детерминации из единицы (1)? Что измеряет данный расчетный показатель?
27.	Что представляет собой интерполяция? Что данная статистическая процедура позволяет получить исследователю?
28.	Что представляет собой экстраполяция? Что данная статистическая процедура позволяет получить исследователю?
29.	Чем отличаются друг от друга интерполяция и экстраполяция? Где применим тот или иной метод?
30.	На чем основано применение методов как интерполяции, так и экстраполяции? Какие показатели необходимо анализировать для реализации этих методов?
31.	Что понимают под основной тенденцией развития ряда динамики?
32.	Какие существуют статистические методы, осуществляющие прогнозирование временных рядов?
33.	Опишите метод простой скользящей средней. Как вычисляется простая скользящая средняя?
34.	Для каких целей могут быть использованы скользящие средние? Как ведут себя скользящие средние в отношении трендов?
35.	Как ведет себя экспоненциальное сглаживание в отношении экспоненциально уменьшающегося веса по мере старения наблюдения?
36.	Как называется подход машинного обучения, который моделирует человеческий мозг и состоит из ряда искусственных нейронов?
37.	Основной ценностью какого метода является не только оценка среднего объема спроса, но и оценка длины временного интервала между двумя ненулевыми требованиями, термин, называемый прерывистым спросом?
38.	Каковы этапы реализации метода Кростона? Охарактеризуйте каждый из них.
39.	Какой из специализированных методов прогнозирования обеспечивает ценность в определенных ограниченных обстоятельствах?
40.	Для чего применяются индексы в статистических исследованиях? В каком формате выражается значение статистического индекса? Перечислите и охарактеризуйте особенности и характеристики статистических индексов.

7.2.2 Практическая работа

(наименование оценочного средства)

Типовые примеры заданий

Практическое задание 1. «Генеральная и выборочная совокупности, их значение в анализе данных»

Задание:

1. Рассчитать объем и состав квотной пропорциональной выборки для генеральной совокупности по варианту (см табл. 1, 2).
2. Результат решения задачи представить в файле под именем «Фамилия_И.О._Задание_4.docx» или «Фамилия_И.О._Задание_4.xlsx».

Таблица 1 – Выбор варианта задания в зависимости от фамилии студента

№ варианта	1	2	3	4	5	6	7	8	9	10
Начальная буква фамилии студента	А	Б	В	Г	Д	Е	Ж	З	И	К
	Л	М	Н	О	П	Р	С	Т	У	Ф

	Ч	Ц	Ч	Ш	Щ	Э	Ю	Я		
--	---	---	---	---	---	---	---	---	--	--

Таблица 2 – Варианты заданий

№ варианта	Генеральная совокупность	Доверительный интервал	Ошибка выборки
1	Трудоспособное население г. Москва в 2021 году	95	4
2	Трудоспособное население Московской области в 2021 году	95	5
3	Трудоспособное население г. Самара в 2021 году	95	3
4	Трудоспособное население Самарской области в 2021 году	95	4
5	Трудоспособное население г. Москва в 2021 году	85	5
6	Трудоспособное население Московской области в 2021 году	85	3
7	Трудоспособное население г. Самара в 2021 году	85	4
8	Трудоспособное население Самарской области в 2021 году	85	5
9	Трудоспособное население г. Москва в 2021 году	97	3
10	Трудоспособное население Московской области в 2021 году	97	4

Рекомендации по выполнению задания

1) Предварительно следует ознакомиться с текстом учебника по данной теме, а также изучить рекомендуемую литературу.

2) Используйте статистические данные, имеющиеся на сайте Федеральной службы государственной статистики (<https://rosstat.gov.ru>).

3) Рассчитайте необходимый размер выборочной совокупности, исходя из размера генеральной совокупности и заявленных показателей доверительного интервала и ошибки выборки.

4) Разделите генеральную совокупность на квоты в соответствии с полом (мужской, женский) и возрастом (18-25 лет, 26-35 лет, 36-45 лет, 46-55 лет, 56-79 лет). Вычислите доли (процент) каждой квоты.

5) В соответствии с определенными долями каждой половозрастной квоты вычислите конкретное количество людей по каждой квоте, которых необходимо опросить.

Практическое задание 2. «Понятие и процедура корреляционного анализа. Коэффициенты корреляции и результаты корреляционного анализа»

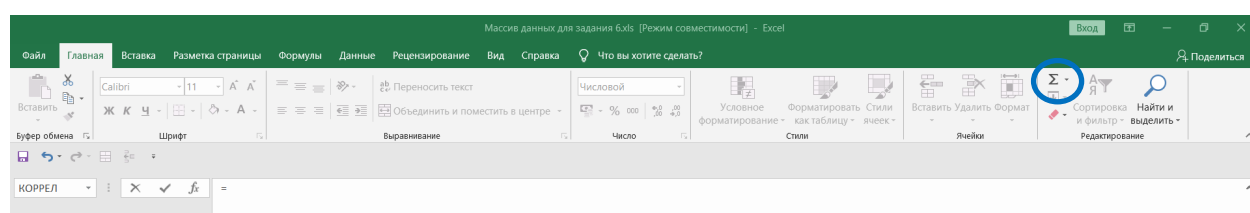
Задание:

1. В предложенном файле с массивом эмпирических данных рассчитать коэффициенты корреляции.
2. Интерпретировать получившиеся результаты
3. Результат решения задачи представить в файле под именем «Фамилия_И.О._Задание_4.docx» или «Фамилия_И.О._Задание_4.xlsx».

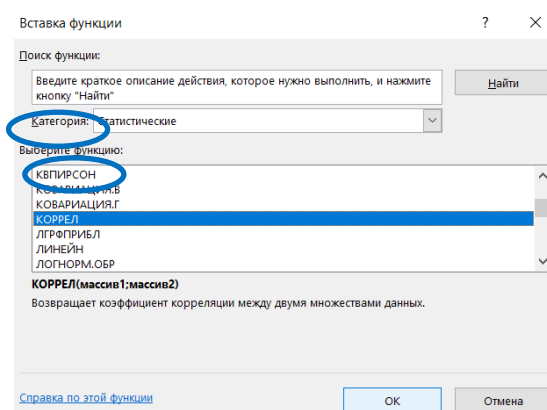
Рекомендации по выполнению задания

1) Предварительно следует ознакомиться с текстом учебника по данной теме, а также изучить рекомендуемую литературу.

2) В ячейке, в которой необходимо указать коэффициент корреляции, ставим знак «=». Затем нажимаем треугольник рядом с пунктом меню «Формулы»



3) В открывшемся окне выбираем категорию «Статистические» и функцию «КОРРЕЛ»



4) Выбираем массивы данных для расчета коэффициента корреляции.

Массив 1 – это данные исследуемой переменной (зависимой). В примере – вариант ответа «Онлайн оплата счетов».

Массив 2 – это независимая переменная. В примере – пол респондента.

Массив данных для задания 6.xls [Режим совместимости] - Excel

Файл Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Справка Что вы хотите сделать?

Вставить Вставить буфер обмена Шрифт Выравнивание Числовой Условное форматирование Форматировать как таблицу Стили Вставить Удалить Формат

Буфер обмена Шрифт Выравнивание Число Стили Ячейки Сортировка Найти и выделить Редактирование

MS = X ✓ ✗ =KORREL(C5:C346;N5:N346)

А В С D E F G H I J K L M N O P

2. Что нового появилось в Вашей жизни в период самоизоляции и?: онлайн оплата счетов

2. Что нового появилось в Вашей жизни в период самоизоляции и?: заказ продуктов на дом

2. Что нового появилось в Вашей жизни в период самоизоляции и?: доставка еды из ресторанов или кафе

2. Что нового появилось в Вашей жизни в период самоизоляции и?: покупка бытовых товаров, одежды, лекарств и т.д.

Аргументы функции

КОРРЕЛ

Массив1 C5:C346

Массив2 N5:N346

Возвращает коэффициент корреляции между двумя множествами данных.

Значение: -0,082

Справка по этой функции

OK Отмена

Worksheet

Укажите

5) Вычисляем коэффициент корреляции (кнопка OK или Enter на клавиатуре).

Массив данных для задания 6.xls [Режим совместимости] - Excel

Файл Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Справка Что вы хотите сделать?

Вставить Вставить буфер обмена Шрифт Выравнивание Числовой Условное форматирование Форматировать как таблицу Стили Вставить Удалить Формат

Буфер обмена Шрифт Выравнивание Число Стили Ячейки Сортировка Найти и выделить Редактирование

MS = X ✓ ✗ =KORREL(C5:C346;N5:N346)

А В С D E F G H I J K L M N O P

2. Что нового появилось в Вашей жизни в период самоизоляции и?: онлайн оплата счетов

2. Что нового появилось в Вашей жизни в период самоизоляции и?: заказ продуктов на дом

2. Что нового появилось в Вашей жизни в период самоизоляции и?: доставка еды из ресторанов или кафе

2. Что нового появилось в Вашей жизни в период самоизоляции и?: покупка бытовых товаров, одежды, лекарств и т.д.

2. Что нового появилось в Вашей жизни в период самоизоляции и?: опыт дистанционной работы/учебы

2. Что нового появилось в Вашей жизни в период самоизоляции и?: домашние тренировки

2. Что нового появилось в Вашей жизни в период самоизоляции и?: видеоконференции с родственниками, друзьями, коллегами

2. Что нового появилось в Вашей жизни в период самоизоляции и?: появилось новое хобби

2. Что нового появилось в Вашей жизни в период самоизоляции и?: Ничего

2. Что нового появилось в Вашей жизни в период самоизоляции и?: Психологические трудности

2. Что нового появилось в Вашей жизни в период самоизоляции и?: Сложности на работе

8. Укажите Ваш пол

9. Возраст по группам (1-до 25; 2-26-45; 3-46-59; 4-60 и старше)

12. Укажите Ваше материальное положение

1 Коэффициент корреляции по:

полу

возрасту

доходу

-0,082

Worksheet

Готово

6) Повторяем пункты 1–3 для всех зависимых переменных.

А В С D E F G H I J K L M N O P

2. Что нового появилось в Вашей жизни в период самоизоляции и?: онлайн оплата счетов

2. Что нового появилось в Вашей жизни в период самоизоляции и?: заказ продуктов на дом

2. Что нового появилось в Вашей жизни в период самоизоляции и?: доставка еды из ресторанов или кафе

2. Что нового появилось в Вашей жизни в период самоизоляции и?: покупка бытовых товаров, одежды, лекарств и т.д.

2. Что нового появилось в Вашей жизни в период самоизоляции и?: опыт дистанционной работы/учебы

2. Что нового появилось в Вашей жизни в период самоизоляции и?: домашние тренировки

2. Что нового появилось в Вашей жизни в период самоизоляции и?: видеоконференции с родственниками, друзьями, коллегами

2. Что нового появилось в Вашей жизни в период самоизоляции и?: появилось новое хобби

2. Что нового появилось в Вашей жизни в период самоизоляции и?: Ничего

2. Что нового появилось в Вашей жизни в период самоизоляции и?: Психологические трудности

2. Что нового появилось в Вашей жизни в период самоизоляции и?: Сложности на работе

8. Укажите Ваш пол

9. Возраст по группам (1-до 25; 2-26-45; 3-46-59; 4-60 и старше)

12. Укажите Ваше материальное положение

1 Коэффициент корреляции по:

полу

возрасту

доходу

-0,082

7) Интерпретируем полученный результат.

Практическое задание 3. «Особенности и преимущества факторного анализа»

Задание:

1. Решить задачу. Провести факторный анализ прибыли предприятия от реализации одного вида продукции.

Исходные данные:

Показатели	План	Факт
Объем реализации продукции, шт.	20000	16000
Цена реализации, ден. ед.	170	200
Себестоимость продукции, ден. ед.	140	150

2. Результат решения задачи представить в файле под именем «Фамилия_И.О._Задание_4.docx» или «Фамилия_И.О._Задание_4.xlsx».

Рекомендации по выполнению задания

1) Предварительно следует ознакомиться с текстом учебника по данной теме, а также изучить рекомендуемую литературу.

2) Для расчета прибыли необходимо понимать, что это разница между выручкой и себестоимостью, что можно представить как произведение объема производства на разницу между ценой и себестоимостью одной единицы товара. Рассчитываем прибыль как по плану, так и по факту.

3) Вычислить разницу между фактической и плановой прибылью.

4) Провести анализ факторов, приведших к разнице между фактической и плановой прибылью. Обратите внимание, что в данном случае указано три фактора: объем реализации продукции, цена реализации и себестоимость продукции.

5) Оцените значимость каждого фактора.

Практическое задание 4. «Понятие, сфера применения и алгоритм проведения кластерного анализа»

Задание:

1. Постройте дендрограмму, иллюстрирующую результаты иерархического кластерного анализа. Тема кластерного анализа, а также данные для кластерного анализа подбираются студентом самостоятельно. Дендрограмма должна содержать не менее 5 уровней иерархии.

2. Результат решения задачи представить в файле под именем «Фамилия_И.О._Задание_4.docx» или «Фамилия_И.О._Задание_4.xlsx».

Рекомендации по выполнению задания

1) Предварительно следует ознакомиться с текстом учебника по данной теме, а также изучить рекомендуемую литературу.

2) Подберите тему для проведения иерархического кластерного анализа. Лучше всего, если данная тема будет коррелировать с Вашей профессиональной деятельностью или с направлением подготовки в вузе.

3) Последовательно выделяйте группы из кластеров верхнего уровня, формируя кластеры более низкого уровня.

4) Опишите и интерпретируйте полученный результат: сколько уровне в итоге получилось, какой основной кластер, сколько групп и т. д.

Практическое задание 5. «Методы проверки гипотез о взаимосвязи переменных» Задание:

1. Создать документ и сохранить его под именем Фамилия_И.О._Задание_6.xlsx.
2. Используя данные, в приведенных ниже таблицах, постройте различные виды диаграмм:
 - круговая
 - точечная
 - столбчатая
 - ленточная

Что нового появилось в Вашей жизни в период пандемии? / Пол

№ пп	Варианты ответов	8. Укажите Ваш пол		Всего
		муж	жен	
1	онлайн оплата счетов	28	32	60
2	заказ продуктов на дом	28	28	56
3	доставка еды из ресторанов или кафе	34	34	68
4	онлайн покупка бытовых товаров, одежды, лекарств и т.д.	16	58	74
5	опыт дистанционной работы/учебы	70	148	218
6	домашние тренировки	36	78	114
7	видеоконференции с родственниками, друзьями, коллегами	34	74	108
8	появилось новое хобби	22	44	66
9	Ничего	1	3	4
10	психологические трудности	0	2	2
11	сложности на работе	1	1	2
Всего		124	208	332

3. Убедитесь, что диаграммы содержат все необходимые элементы графического представления данных: заголовок, подписи данных, названия осей и т.д.
4. Используя данные, в приведенных ниже таблицах, постройте «ящик с усами» для указанной переменной.

Что нового появилось в Вашей жизни в период пандемии?

№ пп	Варианты ответов	Ответы		Процент наблюдений
		N	%	
1	онлайн оплата счетов	60	7,8%	18,1%

2	заказ продуктов на дом	56	7,3%	16,9%
3	доставка еды из ресторанов или кафе	68	8,8%	20,5%
4	онлайн покупка бытовых товаров, одежды, лекарств и т.д.	74	9,6%	22,3%
5	опыт дистанционной работы/учебы	218	28,2%	65,7%
6	домашние тренировки	114	14,8%	34,3%
7	видеоконференции с родственниками, друзьями, коллегами	108	14,0%	32,5%
8	появилось новое хобби	66	8,5%	19,9%
9	Ничего	4	0,5%	1,2%
10	психологические трудности	2	0,3%	0,6%
11	сложности на работе	2	0,3%	0,6%
Всего		772	100,0%	232,5%

Рекомендации по выполнению задания

- 1) Используйте только те статистические данные, которые предоставлены в задании.
- 2) Убедитесь, что в построенных диаграммах присутствует либо легенда, либо подписи данных.
- 3) При построении диаграмм учитывайте все требования к ним, указанные в тексте учебника и рекомендуемой литературе.
- 4) Правильно оформите подписи к каждой диаграмме.

Образец выполнения задания

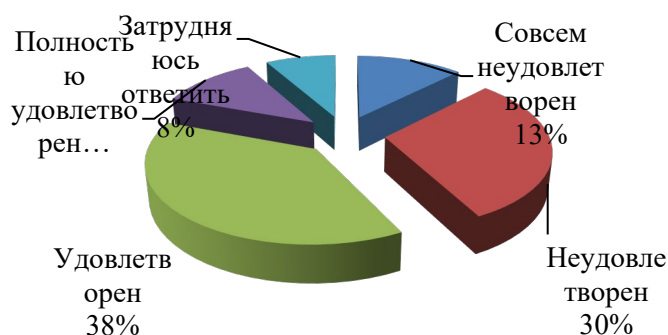


Рисунок 1 – Распределение ответов респондентов на вопрос об удовлетворенности благоустройством города (круговая диаграмма)

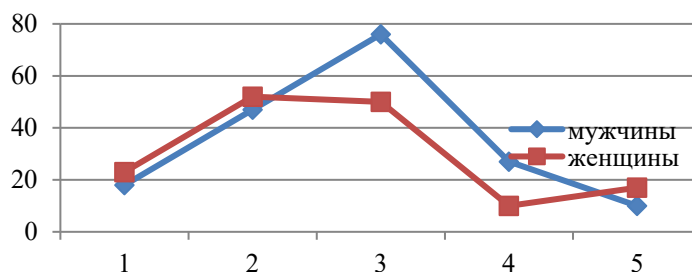


Рисунок 2 – Гендерное распределение ответов респондентов на вопрос об удовлетворенности благоустройством города (точечная диаграмма)

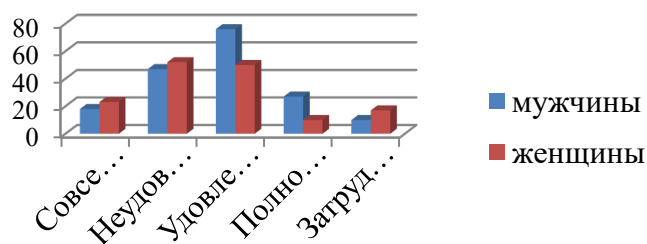


Рисунок 3 – Гендерное распределение ответов респондентов на вопрос об удовлетворенности благоустройством города (столбчатая диаграмма)

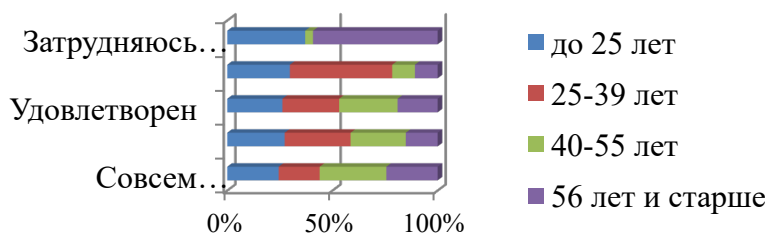


Рисунок 4 – Распределение по возрастным группам ответов респондентов на вопрос об удовлетворенности благоустройством города (ленточная диаграмма)

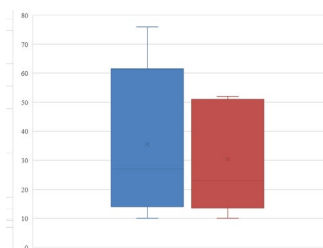


Рисунок 5 – Распределение ответов респондентов на вопрос об удовлетворенности благоустройством города (ящик с усами)

Критерии оценки:

- оценка «зачтено» выставляется студенту, если он ответил на три или два вопроса в полном объеме или частично (продемонстрировав знание материала на 50%);
- оценка «не зачтено» выставляется студенту, если он ответил только на один или не ответил ни на один вопрос.

Требования к оформлению

Отчет должен содержать подробное описание (включая иллюстративный материал) последовательности действий, проделанных студентом для выполнения заданий. Оформление отчета должно соответствовать методическому указанию рекомендациям, изложенным учебно-методическом пособии.

Критерии оценки:

Оценка за практические работы выставляется на основе письменного отчета студента. Преподаватель может потребовать студента исправить замечания по оформлению или содержанию отчета по практической работе.

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

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

7.2.3 Типовые задания для оценки сформированности компетенций

(наименование оценочного средства)

ПК-1 Способен собирать, анализировать и оценивать эффективность бизнес-анализа в организации

код и наименование компетенции

ОМ закрытого типа

Задание 1

Выберите один правильный вариант ответа.

Какой метод анализа данных работает путем моделирования отношений между зависимой переменной и одной или несколькими независимыми переменными?

- а) описательный анализ;
- б) анализ временных рядов;
- в) регрессионный анализ;
- г) дисперсионный анализ.

Правильный ответ: в.

Задание 2

Заполните пропуск: вставьте пропущенное слово.

Иерархическая кластеризация опирается на поиск иерархии кластеров, которая напоминает _____ структуру, называемую дендрограммой.

Правильный ответ: древовидную.

Задание 3

Выберите один правильный вариант ответа.

Какой вид анализа данных помогает определить, существует ли какая-либо связь между набором переменных и ведет к процедурам кластеризации и классификации?

- а) описательный анализ;
- б) факторный анализ;
- в) анализ временных рядов;
- г) дисперсионный анализ.

Правильный ответ: б.

Задание 4

Выберите один правильный вариант ответа.

Среди функций диагностической аналитики есть те, которые раскрывают скрытые связи путем рассмотрения событий, которые могли привести к выявленным аномалиям. Как называется эта функция?

- а) выявление аномалии;
- б) углубление в аналитику (обнаружение);
- в) определение причинно-следственных связей;
- г) сбор выборочной совокупности.

Правильный ответ: в.

ОМ открытого типа

Задание 5

Дайте развернутый ответ.

Что такое диагностический анализ?

Правильный ответ: диагностический анализ – это вид анализа данных (также называется анализом первопричин), который включает в себя такие процессы, как обнаружение данных, интеллектуальный анализ, а также детализация.

Задание 6

Дайте развернутый ответ.

Что такое генеральная совокупность?

Правильный ответ: генеральная совокупность – это вся группа, которую необходимо исследовать и о которой необходимо сделать статистические и аналитические выводы.

Задание 7

Дайте развернутый ответ.

Что такое иерархическая кластеризация?

Правильный ответ: иерархическая кластеризация – это тип кластеризации опирается на поиск иерархии кластеров, которая напоминает древовидную структуру, называемую дендрограммой; при иерархической кластеризации объекты, принадлежащие дочернему кластеру, также принадлежат родительскому кластеру.

Задание 8

Дайте развернутый ответ.

Что такое нечеткая кластеризация?

Правильный ответ: нечеткая кластеризация – это тип кластеризации предполагает, что каждый объект принадлежит каждому кластеру в определенной степени.

7.3. Оценочные средства для промежуточной аттестации по итогам освоения дисциплины

7.3.1. Вопросы к промежуточной аттестации

Семестр 1

№ п/п	Вопросы к зачету
1.	Какова роль анализа данных в современном обществе?
2.	Как связаны процесс принятия решения и анализ данных?
3.	Какие виды анализа данных существуют? Охарактеризуйте каждый из них.
4.	Почему развитие анализа данных было неотъемлемой частью развития современной науки и техники?
5.	Укажите, кто предвосхитил существование нормального распределения, которое играет центральную роль в анализе данных?
6.	Как бельгийский астроном Адольф Кетле назвал статистический подход, примененный им к социальным явлениям? Поясните, почему он выбрал именно это название.
7.	Кто из ученых ввел термин евгеника, чтобы выразить свою веру в то, что условия жизни человечества можно улучшить с помощью научно контролируемого

№ п/п	Вопросы к зачету
	разведения?
8.	Какие типы вероятностей были признаны с момента появления философии науки в 17 веке?
9.	Каким видится будущее анализа данных во время стремительного роста информационных технологий?
10.	Какие вненаучные сферы используют статистические инструменты?
11.	Какие существуют этические проблемы, связанные с неправильным или ненадлежащим использованием статистических инструментов?
12.	Какой смысл изначально вкладывался в понятия «статистика»?
13.	Приведите примеры наиболее раннего применения статистики.
14.	Какого ученого по праву называют отцом статистики? Почему?
15.	Перечислите и опишите основные характеристики анализа данных.
16.	Что такое статистические данные? Какова роль данных в статистической науке?
17.	Какое значение для анализа данных имеет изучение переменных, которые изменяются непредсказуемым образом?
18.	Охарактеризуйте понятие «шкала измерения». Какие виды шкал измерения вы знаете. В каких случаях применяется тот или иной вид шкал измерения?
19.	Какие способы анализа данных применяются? Укажите особенности каждого из способов анализа данных.
20.	Какова значимость анализа данных в различных областях науки и жизнедеятельности?
21.	Каковы преимущества статистического мышления и статистических операций при проведении исследований?
22.	Как анализ данных используется в психологии и образовании?
23.	Каких целей анализ данных должна достигать в исследованиях?
24.	Что является основным объектом статистического наблюдения?
25.	Что такое рабочая программа исследования? Каково ее предназначение?
26.	Охарактеризуйте понятия выборки и генеральной совокупности. Как данные понятия взаимосвязаны?
27.	Каким образом используется процентное измерение при изучении социально-экономических явлений и процессов?
28.	Что такое базовый период и отчетный период? Каковы отличия данных временных периодов? Приведите примеры.
29.	Что такое статистическое наблюдение? Приведите примеры выражения статистического наблюдения как числа и как качества.
30.	Укажите два наиболее важных аспекта точности статистического измерения.
31.	Как взаимосвязаны понятия надежности и достоверности?
32.	Перечислите и охарактеризуйте основные меры обеспечения надежности статистического измерения.
33.	Каковы три важные компонента надежности повторного тестирования?
34.	Охарактеризуйте Каппа-коэффициент, позволяющий определить надежность номинальных переменных.
35.	Проанализируйте надежность альфа как инструмент, позволяющий определить достоверность и точность измерения.
36.	Расскажите о Мета-анализе как об инструменте определения достоверности и точности измерения.
37.	Опишите алгоритм Байесовского анализа как метода обеспечения достоверности и точности статистического измерения.
38.	Проанализируйте значимость размера выборки, необходимого для оценки

№ п/п	Вопросы к зачету
	надежности статистического измерения.
39.	Определите и охарактеризуйте место статистической группировки в линейке методов анализа данных.
40.	В чем заключается суть статистической группировки?
41.	Что такое группировочный признак? Что определяет выбор основания группировки?
42.	В какой момент исследования необходимо идентифицировать типы статистических группировок?
43.	Какие виды представления количественных группировочных признаков существует в анализе данных? Определите и охарактеризуйте их.
44.	Какие виды статистических группировок выделяются по критерию сложности признака?
45.	Какой вид статистической группировки по критерию сложности признака применяется при изучении сложных социально-экономических феноменов, явлений и процессов?
46.	Опишите алгоритм проведения статистической группировки, подробнее охарактеризуйте его этапы.
47.	Охарактеризуйте формулу Стерджесса как одну из часто используемых статистических процедур для реализации статистической группировки.
48.	Назовите самое важное ограничение использования формулы Стерджесса.
49.	Какие виды статистических таблиц применяются в статистике? Перечислите и охарактеризуйте каждый из видов.
50.	Почему применяется оформление данных табличным способом?
51.	Какие ключевые элементы содержит статистическая таблица обобщенных данных?
52.	Какие формы может принимать статистическая таблица обобщенных данных в зависимости от характера используемого в ней подлежащего?
53.	Какие формы может принимать статистическая таблица обобщенных данных в зависимости от глубины разработки в ней сказуемого?
54.	Какие существуют правила оформления статистической таблицы обобщенных данных? Перечислите и охарактеризуйте каждое из них.
55.	Какие бывают итоговые в статистических таблицах обобщенных данных? Какие значения они могут принимать?
56.	Что такое единица измерения? Как она представлена в статистических таблицах обобщенных данных?
57.	В каких ситуациях в некоторых ячейках статистических таблиц обобщенных данных отсутствует информация? Как это оформляется непосредственно в ячейках?
58.	Перечислите и охарактеризуйте этапы работы со статистическими таблицами обобщенных данных.
59.	Какая информация о статистических таблицах обобщенных данных потребуется нам на этапе их формирования?
60.	Какая информация о статистических таблицах обобщенных данных потребуется нам на этапе их чтения?
61.	Какая информация о статистических таблицах обобщенных данных потребуется нам на этапе их анализа?
62.	Каковы преимущества графического представления статистических данных?
63.	Что такое статистический график?
64.	Перечислите и проанализируйте преимущества диаграммного представления данных.
65.	Каковы достоинства и недостатки графического представления данных?

№ п/п	Вопросы к зачету
66.	Какие существуют рекомендации по эффективному графическому представлению данных и информации?
67.	Каковы алгебраические принципы графического представления данных и информации?
68.	Для чего в анализе данных используется график, называемый «ящик с усами»?
69.	Что такое «полигон частот» в анализе данных?
70.	Каковы особенности линейного графика?
71.	Что представляет собой кумулята в анализе данных?
72.	Каковы особенности гистограммы в анализе данных?
73.	Что представляет собой круговая диаграмма?
74.	Каковы особенности пиктограмма в анализе данных?
75.	Чем интересно использование инфографики в анализе данных?
76.	Почему графическое визуальное представление информации является важнейшим компонентом в понимании и выявлении закономерностей и тенденций?
77.	Что представляет собой статистический показатель?
78.	Какое значение имеют статистические показатели в части проведения сравнений?
79.	В каком виде могут быть выражены статистические показатели?
80.	Обоснуйте значение статистических показателей в части связи статистических наблюдений с социальными или другими явлениями.
81.	Перечислите и охарактеризуйте основные категории частоты сбора данных.
82.	Как связаны цели и задачи статистического исследования и собираемые для этого типы данных и переменные?
83.	Что представляют собой абсолютные показатели вариации? Перечислите и охарактеризуйте наиболее часто использующиеся в статистике абсолютные показатели вариации.
84.	Что представляют собой относительные показатели вариации? Перечислите и охарактеризуйте наиболее часто использующиеся в анализе данных относительные показатели вариации.
85.	Что такое структурные средние вариации? Какое значение имеют структурные средние в анализе данных?
86.	Определите понятие среднего арифметического в анализе данных. Как оно рассчитывается? Перечислите его достоинства и недостатки. Приведите примеры использования среднего арифметического в анализе данных.
87.	Определите понятие моды в анализе данных. Как она рассчитывается? Перечислите ее достоинства и недостатки. Приведите примеры использования моды в анализе данных.
88.	Как представлены данные в нормальном распределении? Какие еще названия нормального распределения существуют?
89.	Почему нормально распределенные переменные очень широко распространены?
90.	Какими ключевыми характеристиками описываются нормальные распределения?
91.	Опишите эмпирическое правило (правило 68-95-99,7), которое используется в анализе данных. Для чего оно применяется?
92.	Какие значения необходимо знать для того, чтобы найти z-оценку значения?
93.	Что определяет среднее значение на кривой нормального распределения? Как влияет на кривую изменение среднего выборки?
94.	Что определяет стандартное отклонение на кривой нормального распределения? Как влияет на кривую изменение стандартного отклонения?
95.	Для чего производят преобразование нормального распределения в стандартное нормальное распределение?

№ п/п	Вопросы к зачету
96.	Для каких целей преобразовывают нормальное распределение в z-распределение?
97.	Что такое z-оценка? Что показывают различные виды z-оценки (положительная, отрицательная, нулевая)?
98.	Какие формы стандартного нормального распределения существуют?
99.	Какое влияние оказывает распределение данных выборки на значения центральной тенденции?
100.	Как группируются значения (данные) в нормальном распределении?
101.	Как группируются значения (данные) в асимметричном распределении?
102.	Что представляют собой данные временных рядов?
103.	Как строится график данных временных рядов?
104.	Каково назначение данных временных рядов? Приведите конкретные примеры.
105.	Чем является время в анализе временных рядов?
106.	Как данные временных рядов могут повлиять на прогнозирование?
107.	Как соотносится мониторинг и анализ временных рядов?
108.	Каким образом может происходить фиксация данных во временных рядах?
109.	Как называются измерения, собранные через регулярные промежутки времени, например, кластерный мониторинг и мониторинг работоспособности?
110.	Как называются измерения, собранные через нерегулярные временные интервалы, например, журналы и трассировки?
111.	Определите и охарактеризуйте линейный временной ряд.
112.	Определите и охарактеризуйте нелинейный временной ряд.
113.	Определите и охарактеризуйте неизменность как одну из характеристик временных рядов.
114.	Какое свойство отличает данные (показатели) временных рядов от реляционных данных?
115.	Для каких целей собираются, хранятся, визуализируются и анализируются данные временных рядов?
116.	Определите понятие корреляционного анализа как статистического метода. Что позволяет выяснить корреляционный анализ?
117.	Каковы границы коэффициента корреляции? Как трактуется значение коэффициента корреляции?
118.	Что отражает знак коэффициента корреляции?
119.	Как интерпретируется положительное значение коэффициента корреляции?
120.	Как интерпретируется отрицательное значение коэффициента корреляции?
121.	Чему равно абсолютное значение числа (модуль числа) коэффициента корреляции? Что говорит абсолютное значение коэффициента корреляции о величине корреляции?
122.	Что описывает коэффициент корреляции произведение-момент Пирсона, также известный как r Пирсона? Связь каких переменных может описать коэффициент корреляции Пирсона?
123.	Какие характеристики выборки удовлетворяют применению коэффициента Спирмена?
124.	Что измеряет коэффициент корреляции Спирмена?
125.	При каком виде зависимости каждая переменная изменяется в одном направлении с одинаковой скоростью во всем диапазоне данных.
126.	При каком виде отношений каждая переменная всегда изменяется только в одном направлении, но не обязательно с одинаковой скоростью?
127.	Что получится при возведении коэффициента корреляции в квадрат? Что измеряет данный расчетный показатель? В каких границах находится данный коэффициент?

№ п/п	Вопросы к зачету
128.	Что получится при вычитании коэффициента детерминации из единицы (1)? Что измеряет данный расчетный показатель?
129.	Что представляет собой интерполяция? Что данная статистическая процедура позволяет получить исследователю?
130.	Что представляет собой экстраполяция? Что данная статистическая процедура позволяет получить исследователю?
131.	Чем отличаются друг от друга интерполяция и экстраполяция? Где применим тот или иной метод?
132.	На чем основано применение методов как интерполяции, так и экстраполяции? Какие показатели необходимо анализировать для реализации этих методов?
133.	Что понимают под основной тенденцией развития ряда динамики?
134.	Какие существуют статистические методы, осуществляющие прогнозирование временных рядов?
135.	Опишите метод простой скользящей средней. Как вычисляется простая скользящая средняя?
136.	Для каких целей могут быть использованы скользящие средние? Как ведут себя скользящие средние в отношении трендов?
137.	Как ведет себя экспоненциальное сглаживание в отношении экспоненциально уменьшающегося веса по мере старения наблюдения?
138.	Опишите авторегрессионную интегрированную скользящую среднюю (ARIMA) как статистический метод прогнозирования будущего
139.	Когда метод ARIMA работает лучше всего? При каком поведении данных? При каких характеристиках данных?
140.	Как называется подход машинного обучения, который моделирует человеческий мозг и состоит из ряда искусственных нейронов?
141.	Основной ценностью какого метода является не только оценка среднего объема спроса, но и оценка длины временного интервала между двумя ненулевыми требованиями, термин, называемый прерывистым спросом?
142.	Каковы этапы реализации метода Кростона? Охарактеризуйте каждый из них.
143.	Какой из специализированных методов прогнозирования обеспечивает ценность в определенных ограниченных обстоятельствах?
144.	Для чего применяются индексы в статистических исследованиях?
145.	Статистические индексы показывают общее относительное изменение или прямую измеримую цифру? Почему?
146.	В каком формате выражается значение статистического индекса?
147.	Перечислите и охарактеризуйте особенности и характеристики статистических индексов.
148.	Какой индекс формируется из отношения совокупной стоимости за определенный период к совокупной стоимости, найденной в базисном периоде?
149.	Какой индекс применяется для измерения изменений в объеме или количестве товаров, которые производятся, потребляются и продаются в течение оговоренного периода?
150.	Для чего в статистике используется индекс цен?
151.	Какие вы знаете области применения статистических индексов? Обоснуйте и приведите конкретные примеры.
152.	Каковы преимущества статистических индексов?
153.	Каковы ограничения статистических индексов?
154.	Что показывает индекс себестоимости продукции? Какова формула вычисления индекса себестоимости продукции?

№ п/п	Вопросы к зачету
155.	Что показывает индекс производительности труда? Какова формула вычисления индекса производительности труда?
156.	Что показывает индекс трудоемкости? Какова формула вычисления индекса трудоемкости? Чему равна величина индекса трудоемкости?

7.3.2. Критерии и нормы оценки

Семестр	Форма проведения промежуточной аттестации	Критерии и нормы оценки	
2	Зачет с оценкой	«отлично»	Студент сдал все практические работы и знает основные определения, методы анализа данных и владеет навыками их практического применения.
		«хорошо»	Студент сдал все практические работы и разбирается в основных терминах и понятиях.
		«удовлетворительно»	Студент сдал все практические работы и знает некоторые базовые принципы анализа данных.
		«неудовлетворительно»	Студент сдал не все практические работы или студент плохо владеет теоретическим материалом

8. Учебно-методическое и информационное обеспечение дисциплины

8.1. Обязательная литература

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
1.	Нестеров С.А.	Интеллектуальный анализ данных средствами MS SQL Server 2008	Учебное пособие	2016	ЭБС «IPRbooks» http://www.iprbookshop.ru/62813.html
2.	Пальмов С.В.	Интеллектуальный анализ данных	Учебное пособие	2017	ЭБС «IPRbooks» http://www.iprbookshop.ru/75376.html
3.	Воронова Л.И., Воронов В.И.	Machine Learning: регрессионные методы интеллектуального анализа данных	Учебное пособие	2018	ЭБС «IPRbooks» http://www.iprbookshop.ru/81325.html

8.2. Дополнительная литература

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
1.	Федин Ф.О., Федин Ф.Ф.	Анализ данных. Часть 1. Подготовка данных к анализу	Учебное пособие	2012	ЭБС «IPRbooks» http://www.iprbookshop.ru/26444.html
2.	Федин Ф.О., Федин Ф.Ф.	Анализ данных. Часть 2. Инструменты	Учебное пособие	2012	ЭБС «IPRbooks»

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно- методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
		Data Mining			http://www.iprbookshop.ru/26445.html
3.	Мельниченко А.С.	Математическая статистика и анализ данных	Учебное пособие	2018	ЭБС «IPRbooks» http://www.iprbookshop.ru/78563.html
4.	Истомина А.П.	Анализ данных качественных исследований	Практикум	2016	ЭБС «IPRbooks» http://www.iprbookshop.ru/66014.html
5.	Жуковский О.И.	Информационные технологии и анализ данных	Учебное пособие	2014	ЭБС «IPRbooks» http://www.iprbookshop.ru/72106.html

8.3. Перечень профессиональных баз данных и информационных справочных систем

Springer International Publishing, Part of Springer Science+Business Media [Электронный ресурс] – Springer International Publishing AG, 2020. Режим доступа к журн.: <http://link.springer.com> . – Загл. с экрана

8.4. Перечень программного обеспечения

№ п/п	Наименование ПО	Реквизиты договора (дата, номер, срок действия)
1.	Microsoft Access	До 01.07.2020
2.	Microsoft Visual Studio	До 01.07.2020
3.	MathCAD	MCD-7514-P/MCD-7503CP от 21.07.2009
4.	PyCharm	-

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

№ п/п	Наименование оборудованных учебных кабинетов, лабораторий, мастерских и др. объектов для проведения практических и лабораторных занятий, помещений для самостоятельной работы обучающихся (номер аудитории)	Перечень основного оборудования
1	Компьютерный класс УЛК-401	Компьютер (монитор 19", системный блок Pentium (R) Dual-Core E5500 2,8 GHz / 4 Gb / 500 Gb) - 12 шт, стол ученический - 7 шт., стол компьютерный -14шт., стол преподавательский -1 шт., стулья -35шт. Доска аудиторная(меловая) - 1 шт.
2	Компьютерный класс УЛК-407	Компьютер (монитор Samsung Sync Master 943n 19", системный блок Intel (R) Core 2 Quad 2,40 GHz 1 Gb) - 12 шт., стол лабораторный -10шт., стул -25 шт., доска 3-х секционная(меловая) -1 шт., стол преподавательский-1 шт.
3	Компьютерный класс УЛК-408	Компьютер (монитор 17", системный блок Intel (R) Celeron (R) 2,66 GHz / 1 Gb / 80 Gb) - 18шт., маршрутизатор 2801 Router-6 шт., коммутатор Catalyst-6 шт., экран/интерактивная доска Smart Board TB-1 шт., проектор Acer P1303W., стол преподавательский-1шт., стол ученический-13шт., стол компьютерный-18 шт., стул- 50 шт., доска аудиторная (маркерная)-1 шт.