

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Андрей Драгомирович Хлутков
Должность: директор
Дата подписания: 02.12.2024 23:48:09
Уникальный программный ключ:
880f7c07c583b07b775f6604a630281b13ca9fd2

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

Северо-Западный институт управления – филиал РАНХиГС

Кафедра экономики и финансов
(наименование кафедры)

УТВЕРЖДЕНО
Директор СЗИУ РАНХиГС
А.Д.Хлутков

**ПРОГРАММА БАКАЛАВРИАТА
«Бизнес-аналитика»**

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
реализуемой без применения электронного (онлайн) курса**

Б1.В.16 Системы обработки и анализа данных
(индекс, наименование дисциплины, в соответствии с учебным планом)

СОиАД

(краткое наименование дисциплины)

38.03.05 Бизнес-информатика
(код, наименование направления подготовки)

очная
(форма обучения)

Год набора – 2024

Санкт-Петербург, 2024г.

Автор–составитель:

Старший преподаватель кафедры Бизнес-информатики Гальченко Максим Иванович

Заведующий кафедрой бизнес-информатика

д.в.н., профессор

Наумов Владимир Николаевич

РПД по дисциплине Б1.В.16 «Системы обработки и анализа данных» одобрена на заседании кафедры бизнес-информатики. Протокол от 27.06.2024г. №10

СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы
2. Объем и место дисциплины в структуре образовательной программы
3. Содержание и структура дисциплины
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине
5. Оценочные материалы промежуточной аттестации по дисциплине
 6. Методические указания для обучающихся по освоению дисциплины
7. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине
 - 7.1. Основная литература
 - 7.2. Дополнительная литература
 - 7.3. Учебно-методическое обеспечение самостоятельной работы
 - 7.4. Нормативные правовые документы
 - 7.5. Интернет-ресурсы
 - 7.6. Иные источники
8. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

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

Дисциплина «Системы обработки и анализа данных» обеспечивает овладение следующими компетенциями:

Таблица 1.1

Код компетенции	Наименование компетенции	Код Компонента компетенции	Наименование компонента компетенции
ПКС-5	Способен решать задачи анализа больших данных с использованием существующих программных средств и технологий	ПКС-5.1	Применяет языки и системы обработки и анализа данных, статистические методы при решении задач предобработки и анализа данных
		ПКС-5.2	Применяет языки и системы анализа данных и машинного обучения, системы управления базами данных при работе с большими данными

В результате освоения дисциплины у студентов должны быть сформированы:

Таблица 1.2

ОТФ/ТФ (при наличии профстандарта)/ профессиональные действия	Код компонента компетенции	Результаты обучения
Анализ больших данных с использованием существующей организации методологической и технологической инфраструктуры/ Проведение аналитического исследования с применением технологий больших данных	ПКС-5.1; ПКС-5.2	на уровне знаний: <ul style="list-style-type: none"> - стандарты анализа данных, основные принципы хранения данных, хранения данных в СУБД, принципы работы с большими данными; - принципы разведочного анализа данных, ETL, основные виды диаграмм в анализе данных; - основные конструкции языков программирования R и Python, основные библиотеки для работы с данными, принципы работы в no-code/low-code приложениях при обработке данных.
		на уровне умений: <ul style="list-style-type: none"> - использовать знания о системах обработки данных и анализе данных в практической деятельности; на уровне навыков: <ul style="list-style-type: none"> - загружать данные, проводить их базовый анализ с использованием no-code/low-code приложений; - использовать R/Python для проведения базового анализа данных.

7. Объем и место дисциплины в структуре ОП ВО

Объем дисциплины

Общая трудоемкость дисциплины составляет 3 зачетных единицы 108 /81 акад/астр.часов.

Дисциплина реализуется частично с применением дистанционных образовательных технологий (далее – ДОТ).

Доступ к системе дистанционных образовательных технологий осуществляется каждым обучающимся самостоятельно с любого устройства на портале: <https://lms.ranepa.ru/>. Пароль и логин к личному кабинету / профилю предоставляется студенту в деканате.

Таблица 2

Вид работы	Трудоемкость в акад. часах ауд./ЭО, ДОТ	Трудоемкость в астрон. часах ауд./ЭО, ДОТ
Общая трудоемкость	108	81
Контактная работа с преподавателем	38	28,5
Лекции		
Практические занятия	36	27
Лабораторные занятия		
Практическая подготовка		
Самостоятельная работа	70	52,5
Контроль		
Формы текущего контроля	Тест, устный опрос, практическое контрольное задание	
Форма промежуточной аттестации	Зачет с оценкой	

Место дисциплины в структуре ОП ВО

Дисциплина Б1.В.16 «Системы обработки и анализа данных» относится к части учебного плана, формируемой участниками образовательных отношений по направлению «Бизнес-аналитика» 38.03.05. Преподавание дисциплины «Системы обработки и анализа данных» основано на дисциплинах – Б1.О.10 «Основы информатики», Б1.О.11 «Программирование». В свою очередь она создаёт необходимые предпосылки для освоения программ таких дисциплин, как Б1.О.12 «Базы данных», Б1.О.14 «Объектно-ориентированный анализ и программирование», Б1.В.02 «Стандартизация, сертификация и управление качеством ИС и ИТ», Б1.В.07 «Проектирование информационных систем», Б2.В.28 «Архитектура предприятия» и ряда дисциплин по выбору студента.

Дисциплина изучается в 2-м семестре 1-го курса.

Формой промежуточной аттестации в соответствии с учебным планом является зачет с оценкой.

8. Содержание и структура дисциплины

Таблица 3

№ п/п	Наименование тем	Объем дисциплины, час.						Форма текущего контроля успеваемости и, промежуточ ной аттестации	
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР (ЭК)		
			Л	ЛР	ПЗ	КСР	СРО		СП
Тема 1.	Средства работы с большими данными. Системы управления базами данных	36			12		24		УО/ПКЗ/Т
Тема 2.	Инструментальные средства обработки данных. Методология low-code	34			12		22		УО/ПКЗ/Т
Тема 3.	Языки и пакеты анализа данных	36			12		24		УО/ПКЗ/Т
Контроль						2			
Промежуточная аттестация						2/1,5*			Зачет с оценкой
Всего (акад./астр. часы):		108/81			36/27	2/1,5	70/52,5		

Примечание:

- консультация перед Зачетом – 2 часа

Используемые сокращения:

Л – занятия лекционного типа (лекции и иные учебные занятия, предусматривающие

преимущественную передачу учебной информации педагогическими работниками организации и (или) лицами, привлекаемыми организацией к реализации образовательных программ на иных условиях, обучающимся);

ПЗ – практические занятия (виды занятия семинарского типа за исключением лабораторных работ);
КСР – индивидуальная работа обучающихся с педагогическими работниками организации и (или) лицами, привлекаемыми организацией к реализации образовательных программ на иных условиях (в том числе индивидуальные консультации);

СР – самостоятельная работа, осуществляемая без участия педагогических работников организации и (или) лиц, привлекаемых организацией к реализации образовательных программ на иных условиях;
СП – самопроверка;

СРО – самостоятельная работа обучающегося

контрольные работы (К), опрос (О), тестирование (Т). Выполнение задания (Зад)

Содержание дисциплины

Тема 1. Средства работы с большими данными. Системы управления базами данных

Данные и информация. Хранение данных, файлы JSON, XML и CSV. Базы данных, СУБД. Реляционные и нереляционные модели данных. Построение базы данных в реляционной СУБД. Типы данных и шкалы данных. Нереляционные (NoSQL) базы данных. Элементы технологии OLAP. Большие данные. хранилища данных, озёра данных. Медальная архитектура слоёв данных (medallion architecture). Схема «снежинка» и OBT. Стандарты моделирования (KDD, CRISP-DM, Data Science Life Cycle). CRISP-DM: этапы понимания бизнеса, понимания данных, подготовки данных. Разведочный анализ данных. Аномалии в данных.

Тема 2. Инструментальные средства обработки данных. Методология low-code

Общая характеристика методологии low-code. Классификация инструментальных средств low-code. Loginom community, KNIME Analytics Platform, Orange. Поток данных (workflow) KNIME. Установка компонент. Импорт данных. Преобразование данных. Визуализация данных. Метаузлы и компоненты. Работа с данными с помощью статистических пакетов JASP, jamovi.

Тема 3. Языки и пакеты анализа данных

Репозитории данных. Kaggle, GitHub. Организация работы над проектом с GitHub. Язык R, интегрированная среда разработки Rstudio(Positron). Репозитории языка, пакеты. Проекты, скрипты и блокноты, рабочее пространство. Структуры данных. Основы синтаксиса. Импорт данных с клавиатуры, из буфера данных, excel, текстовых файлов, из баз данных и файлов статистических систем. Основы работы с данными. Классические наборы данных. Пакеты работы с данными python.

9. Материалы текущего контроля успеваемости обучающихся

Формы и методы текущего контроля успеваемости.

В ходе реализации дисциплины «Системы обработки и анализа данных» используются следующие методы текущего контроля успеваемости обучающихся:

Таблица 4.1

Тема (раздел)	Формы (методы) текущего контроля успеваемости
Тема 1. Средства работы с большими данными. Системы управления базами данных	Защита задания, тестирование, опрос

Тема 2. Инструментальные средства обработки данных. Методология low-code	Защита задания, тестирование, опрос
Тема 3. Языки и пакеты анализа данных	Защита задания, тестирование, опрос

Зачет проводится с применением следующих методов (средств):

Зачет проводится в компьютерном классе. Во время Зачета проверяются этапы освоения компетенции ПКС-1.1.

Во время проверки сформированности этапов компетенции ПКС-1.1. оцениваются: правильность ответов на поставленные вопросы, степень их полноты и обоснованности.

Промежуточная аттестация может проводиться устно в ДОТ/письменно с прокторингом/ тестирование с прокторингом. Для успешного освоения курса учащемуся рекомендуется ознакомиться с литературой, размещенной в разделе 6, и материалами, выложенными в ДОТ.

Материалы текущего контроля успеваемости обучающихся.

Тест

1. Верно ли утверждение, что файл в файле CSV может быть использована только запятая в качестве разделителя полей?
 1. Верно
 - 2. Неверно**
 3. Ответ зависит от операционной системы
2. Верно ли утверждение, что любой файл CSV может быть открыт в электронных таблицах?
 1. Верно
 - 2. Неверно**
 3. Ответ зависит от используемой электронной таблицы
3. Отдельный объект JSON заключается в скобки вида:
 - 1. {}**
 2. ()
 3. []
 4. <
4. К какому типу СУБД относится MongoDB?
 1. Колоночная
 2. Графовая
 - 3. Документ-ориентированная**
 4. Ключ-значение
5. К какому типу СУБД относится Neo4j?
 1. Колоночная
 - 2. Графовая**
 3. Документ-ориентированная
 4. Ключ-значение
6. Какому уровню представления данных в озёрах данных соответствует описание «Структуры таблиц на этом уровне соответствуют системным таблицам-источникам "как есть", а также любым дополнительным столбцам метаданных, которые содержат дату/время загрузки, идентификатор процесса и т.д.»?
 1. Платиновый
 2. Стальной
 - 3. Бронзовый**
 4. Серебряный
 5. Золотой
7. Какому уровню представления данных в озёрах данных соответствует описание

«Данные на уровне Gold в the lake house, как правило, организованы в готовые к использованию базы данных "для проекта"»?

1. Платиновый
 2. Стальной
 3. Бронзовый
 4. Серебряный
 5. Золотой
8. Какими преимуществами обладает подход ОВТ относительно схемы «снежинка»?
1. Простота
 2. Лучшая согласованность и простота в управлении
 3. Быстродействие
 4. Безопасность
9. Какой оператор присвоения может быть использован в базовом R?
1. <-
 2. =
 3. ->
 4. :=
 5. =:

5. Оценочные материалы промежуточной аттестации по дисциплине

Таблица 4.3.1

Код компетенции	Наименование компетенции	Код Компонента компетенции	Наименование компонента компетенции
ПКС-5	Способен решать задачи анализа больших данных с использованием существующих программных средств и технологий	ПКС-5.1	Применяет языки и системы обработки и анализа данных, статистические методы при решении задач предобработки и анализа данных
		ПКС-5.2	Применяет языки и системы анализа данных и машинного обучения, системы управления базами данных при работе с большими данными

Показатели и критерии оценивания компетенций с учетом этапа их формирования

Таблица 4.3.2

Код компетенции	Показатель оценивания	Критерий оценивания
ПКС-5.1	1. Демонстрирует умения решать простейшие задачи анализа данных с использованием систем обработки и анализа данных 2. Демонстрирует умения использовать статистические пакеты, инструментальные средства low-code при решении задач анализа и обработки данных	1. Продемонстрированы знания основных типов данных, моделей данных, использования инструментальных средств их обработки 2. Правильно решены кейсы, направленные на обработку и анализ данных, использование репозитория.
ПКС-5.2	1. Демонстрирует умения	1. Продемонстрированы знания основных

Код компетенции	Показатель оценивания	Критерий оценивания
	работать с репозиториями, импортировать большие наборы данных 2. Демонстрирует умения решать простые задачи предобработки и визуализации больших данных, решать задачи их предобработки	типов данных, моделей данных, использования инструментальных средств их обработки 2. Правильно решены кейсы, направленные на обработку и анализ данных, использование репозитория при работе с большими данными.

Для оценки сформированности компетенции, знаний и умений, соответствующих данным компетенциям, используются контрольные вопросы и задачи.

Типовые вопросы, выносимые на зачет с оценкой:

Шкала оценивания.

Оценка результатов производится на основе балльно-рейтинговой системы (БРС). Использование БРС осуществляется в соответствии с приказом от 06 сентября 2019 г. №306 «О применении балльно-рейтинговой системы оценки знаний обучающихся».

Схема расчетов сформирована в соответствии с учебным планом направления, согласована с руководителем научно-образовательного направления, утверждена деканом факультета.

Схема расчетов доводится до сведения студентов на первом занятии по данной дисциплине, является составной частью рабочей программы дисциплины и содержит информацию по изучению дисциплины, указанную в Положении о балльно-рейтинговой системе оценки знаний обучающихся в РАНХиГС.

В случае если студент в течение семестра не набирает минимальное число баллов, необходимое для сдачи промежуточной аттестации, то он может заработать дополнительные баллы, отработав соответствующие разделы дисциплины, получив от преподавателя компенсирующие задания.

В случае получения на промежуточной аттестации неудовлетворительной оценки студенту предоставляется право повторной аттестации в срок, установленный для ликвидации академической задолженности по итогам соответствующей сессии.

Обучающийся, набравший в ходе текущего контроля в семестре от 51 до 70 баллов, по его желанию может быть освобожден от промежуточной аттестации.

Количество баллов	Оценка	
	прописью	буквой
96-100	отлично	А
86-95	отлично	В
71-85	хорошо	С
61-70	хорошо	Д
51-60	удовлетворительно	Е

Перевод балльных оценок в академические отметки «отлично», «хорошо», «удовлетворительно»

- «Отлично» (А) - от 96 по 100 баллов – теоретическое содержание курса освоено полностью, без пробелов необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные программой обучения учебные задания

выполнены, качество их выполнения оценено максимальным числом баллов.

- «Отлично» (В) - от 86 по 95 баллов – теоретическое содержание курса освоено полностью, без пробелов необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные программой обучения учебные задания выполнены, качество их выполнения оценено числом баллов, близким к максимальному.

- «Хорошо» (С) - от 71 по 85 баллов – теоретическое содержание курса освоено полностью, без пробелов, некоторые практические навыки работы с освоенным материалом сформированы недостаточно, все предусмотренные программой обучения учебные задания выполнены, качество выполнения ни одного из них не оценено минимальным числом баллов, некоторые виды заданий выполнены с ошибками.

- «Хорошо» (D) - от 61 по 70 баллов – теоретическое содержание курса освоено полностью, без пробелов, некоторые практические навыки работы с освоенным материалом сформированы недостаточно, большинство предусмотренных программой обучения учебных заданий выполнены, качество выполнения ни одного из них не оценено минимальным числом баллов, некоторые виды заданий выполнены с ошибками.

- «Удовлетворительно» (E) - от 51 по 60 баллов – теоретическое содержание курса освоено частично, но пробелы не носят существенного характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, большинство предусмотренных программой обучения учебных заданий выполнено, некоторые из выполненных заданий выполнены с ошибками.

10. Методические материалы для освоения дисциплины

Описание системы оценивания

Таблица 4.4

Оценочные средства (формы текущего и промежуточного контроля)	Показатели* оценки	Критерии** оценки
Опрос	Корректность и полнота ответов	Сложный вопрос: полный, развернутый, обоснованный ответ – 4 балла Правильный, но не аргументированный ответ – 2 балла Неверный ответ – 0 баллов Обычный вопрос: полный, развернутый, обоснованный ответ – 4 балла Правильный, но не аргументированный ответ – 2 балла Неверный ответ – 0 баллов. Простой вопрос: Правильный ответ – 2 балла; Неправильный ответ – 0 баллов
Тестирование	Процент правильных ответов на вопросы теста.	Максимальное количество баллов за один тест составляет 5.
Зачет с оценкой	В соответствии с балльно-рейтинговой системой на промежуточную аттестацию отводится 30 баллов. Зачет проводится по билетам. Билет содержит 2 вопроса и практическое задание. Максимально по каждому вопросу билета (заданию) начисляется 10 баллов	1-5 баллов за ответ, подтверждающий знания в рамках лекций и обязательной литературы, 6-10 баллов – в рамках лекций, обязательной и дополнительной литературы, 11-15 баллов – в рамках лекций, обязательной и дополнительной литературы, с элементами самостоятельного анализа.

Выполнение практической	<ul style="list-style-type: none"> • Выполнена обязательная часть; • Выполнена обязательная часть и задания для самостоятельной работы • Выполнена обязательная часть, задания для самостоятельной работы и дополнительные задания 	баллы начисляются от 1 до 5 в зависимости от полноты и правильности выполнения работы
Защита практической	<ul style="list-style-type: none"> • обоснованность решений • корректность выводов • аккуратность оформления • своевременность представления для защиты 	<p>Максимально 5 баллов</p> <p>1-2 баллов за ответ, подтверждающий знания в рамках лекций и обязательной части практической работы;</p> <p>3-4 балла за ответ, подтверждающий знания в рамках лекций, обязательной и самостоятельной части практической работы;</p> <p>5 баллов за ответ, подтверждающий знания в рамках лекций, обязательной, дополнительной литературы, с элементами самостоятельного анализа.</p>

Оценивание студентов на Зачете по дисциплине «Вычислительные системы, сети и телекоммуникации»

Баллы %	Критерии
100-85 «отлично»	Оценка «отлично» на Зачете выставляется обучающемуся, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, использует в ответе материал монографической литературы, правильно обосновывает принятое решение.
84-70 «хорошо»	– Оценка «хорошо» выставляется обучающемуся, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения, допускает неточности в увязывании теории с практикой.
69-51 «удовлетворительно»	– Оценка «удовлетворительно» выставляется обучающемуся, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала, испытывает затруднения при установлении связи теории и практики.
Менее 51 «неудовлетворительно»	– Оценка «неудовлетворительно» выставляется обучающемуся, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями устанавливает связь теории и практики.

11. Методические указания для обучающихся по освоению дисциплины

Рабочей программой дисциплины предусмотрены следующие виды аудиторных занятий: лекции, практические занятия, контрольные работы.

Описание учебной дисциплины и методика выполнения практических занятий имеются в ресурсах сети факультета. Подготовка к практической работе предусматривает

изучение теоретического материала. Перед выполнением практической работы необходимо внимательно ознакомиться с описанием практического задания, уяснить, в чем состоят её цель и заданные результаты. Выполнение каждой работы сопровождается оформлением. По результатам защиты работы выставляется оценка.

На лекциях рассматриваются наиболее сложный материал дисциплины. Лекция сопровождается презентациями, компьютерными текстами лекции, что позволяет студенту самостоятельно работать над повторением и закреплением лекционного материала. Для этого студенту должно быть предоставлено право самостоятельно работать в компьютерных классах в сети Интернет.

Тематика практических работ направлена на закрепление и углубление теоретических знаний, полученных студентами при самостоятельной работе, на экспериментальную проверку теоретических положений, выработку умений и практических навыков работы с компьютерной техникой. Студент должен научиться анализировать полученные результаты работы, сравнивать различные методы достижения поставленной цели и на их основе делать выводы.

С целью контроля сформированности компетенций разработан фонд контрольных заданий. Его использование позволяет реализовать балльно-рейтинговую оценку, определенную приказом от 28 августа 2014 г. №168 «О применении балльно-рейтинговой системы оценки знаний студентов».

Для активизации работы студентов во время контактной работы с преподавателем отдельные занятия проводятся в интерактивной форме. В основном, интерактивная форма занятий обеспечивается при проведении занятий в компьютерном классе.

Для работы с печатными и электронными ресурсами СЗИУ имеется возможность доступа к электронным ресурсам. Организация работы студентов с электронной библиотекой указана на сайте института (странице сайта – «Научная библиотека»).

Контрольные вопросы для подготовки к занятиям

Таблица 5

№ п/п	Наименование темы или раздела дисциплины	Контрольные вопросы для самопроверки
1	Тема 1. Средства работы с большими данными. Системы управления базами данных	<ol style="list-style-type: none"> 1. Дайте определение данных и информации 2. Опишите базовую структуру объекта JSON 3. Опишите базовую структуру XML файла 4. Опишите базовую структуру файла CSV 5. Приведите определения БД и СУБД 6. Опишите общую структуру РБД 7. Какие основные шкалы данных используются при анализе? 8. Опишите схему «снежинка», определите преимущества и недостатки подхода ОВТ 9. Опишите структуру, различия и области применения хранилищ данных и озёр данных 10. Приведите классификацию NoSQL СУБД 11. Опишите основные принципы технологии OLAP 12. Какие основные шаги предусмотрены стандартом KDD? 13. Какие основные шаги предусмотрены стандартом CRISP-DM? 14. Какие основные шаги предусмотрены стандартом Data Science Life Cycle? 15. Опишите действия выполняемые в CRISP-DM на этапах понимания бизнеса, понимания данных, подготовки данных 16. Приведите основные шаги разведочного анализа данных 17. Какие аномалии встречаются в данных?
2	Тема 2. Инструментальные средства обработки	<ol style="list-style-type: none"> 1. Приведите классификацию и парадигмы, применяемые в инструментах low-code 2. Какие возможности KNIME Analytics Platform для импорта и

	данных. Методология low-code	<p>экспорта данных вам известны?</p> <p>3. Какие возможности KNIME Analytics Platform для преобразования данных вам известны?</p> <p>4. Опишите основные типы данных и шкалы данных.</p> <p>5. Для каких целей служит и как строится гистограмма и столбчатая диаграмма в KNIME Analytics Platform?</p> <p>6. Для каких целей служит и как строится ящик с усами в KNIME Analytics Platform?</p> <p>7. Для каких целей служит и как строится диаграмма рассеяния в KNIME Analytics Platform?</p> <p>8. Для каких целей служит и как строится матрица диаграмм рассеяния в KNIME Analytics Platform?</p> <p>9. Как построить интерактивную и сводную таблицу в KNIME Analytics Platform?</p> <p>10. Чем различаются метаузел и компонент в KNIME Analytics Platform?</p> <p>11. Опишите принципы работы в JASP и jamovi</p>
3	Тема 3. Языки и пакеты анализа данных	<p>1. Как подключаются библиотеки R?</p> <p>2. Как происходит загрузка и установка дополнительных библиотек R?</p> <p>3. Опишите структуру и основные принципы работы со списками и векторами.</p> <p>4. Опишите структуру и основные принципы работы с фреймами данных.</p> <p>5. Покажите на примерах организацию ветвления в R</p> <p>6. Приведите синтаксис оператора цикла for в R</p> <p>7. Приведите синтаксис оператора цикла while в R</p> <p>8. Как импортируются данные из CSV файла в R?</p> <p>9. Как импортируются данные из Excel файла в R?</p> <p>10. Приведите синтаксис для построения гистограммы с помощью библиотеки ggplot2.</p> <p>11. Приведите синтаксис для построения ящика с усами и скрипичной диаграммы с помощью библиотеки ggplot2.</p> <p>12. Приведите синтаксис для построения диаграммы рассеяния с помощью библиотеки ggplot2.</p> <p>13. Приведите синтаксис для построения гистограммы с помощью библиотеки ggplot2.</p> <p>14. Как импортировать данные с помощью библиотеки pandas?</p> <p>15. Опишите основные принципы работы с объектами типа Dataframe.</p>

12. 7. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

7.1. Основная литература.

1. Мастицкий С.Э., Шитиков В.К. (2014) Статистический анализ и визуализация данных с помощью R. – [Электронный ресурс]: Электронная книга, адрес доступа: <http://r-analytics.blogspot.com>
2. Сысоева М. В., Сысоев И. В. Программирование для «нормальных» с нуля на языке Python : Учебник. В двух частях. Часть 1 [Электронный ресурс]: / М.В. Сысоева, И.В. Сысоев ; отв. ред. В. Л. Черный. – 2-е изд., испр. и доп. – Москва : Базальт СПО; МАКС Пресс, 2023. – 184 с. [+ 4 с. вкл.] : ил. – (Библиотека ALT).

7.2. Дополнительная литература.

1. Zuur, Alain & Ieno, Elena & Elphick, Chris. (2010). Zuur AF, Ieno EN, Elphick CS.. A protocol for data exploration to avoid common statistical problems. *Methods Ecol Evol* 1: 3-14. *Methods in Ecology and Evolution*. 1. 3 - 14. 10.1111/j.2041-210X.2009.00001.x.
2. Барсегян, А. А. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.: ил. + CD-ROM — (Учебная литература для вузов)

7.3. Учебно-методическое обеспечение самостоятельной работы.

1. Положение об организации самостоятельной работы студентов федерального государственного бюджетного образовательного учреждения высшего образования «Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации» (в ред. приказа РАНХиГС от 11.05.2016 г. № 01-2211);
2. Положение о курсовой работе (проекте) выполняемой студентами федерального государственного бюджетного образовательного учреждения высшего образования «Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации» (в ред. приказа РАНХиГС от 11.05.2016 г. № 01-2211)

7.4. Нормативные правовые документы.

Не используются

7.5. Интернет-ресурсы.

СЗИУ располагает доступом через сайт научной библиотеки <http://nwapa.spb.ru/> к следующим подписным электронным ресурсам:

Русскоязычные ресурсы

Электронные учебники электронно - библиотечной системы (ЭБС) «Айбукс»

Электронные учебники электронно – библиотечной системы (ЭБС) «Лань»

Рекомендуется использовать следующий интернет-ресурсы

<https://textbook.rintro.ru/index.html>

,<https://stepik.org/course/4852/>

<https://stepik.org/course/129/>

<https://stepik.org/course/497/>

7.6. Иные источники.

<https://www.datascienceguide.org/index.html>

Navarro DJ and Foxcroft DR (2022). learning statistics with jamovi: a tutorial for psychology students and other beginners. (Version 0.75). DOI: 10.24384/hgc3-7p15 [Электронный ресурс]:

<https://davidfoxcroft.github.io/ljsj-book/learning-statistics-with-jamovi.pdf>

<https://arize.com/blog-course/>

13.8. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

Все практические занятия проводятся в компьютерном классе. Учебная дисциплина включает использование программного обеспечения Microsoft Office.

Методы обучения с использованием информационных технологий (компьютерное тестирование, демонстрация мультимедийных материалов).

Интернет-сервисы и электронные ресурсы (поисковые системы, электронная почта, профессиональные тематические чаты и форумы, системы аудио и видео конференций, онлайн энциклопедии, справочники, библиотеки, электронные учебные и учебно-методические материалы).

Система дистанционного обучения Moodle.

№ п/п	Наименование
1.	Компьютерные классы с персональными ЭВМ, объединенными в локальные сети с выходом в Интернет
2.	MSWord, MSExcel, jamovi, JASP, KNIME Analytics Platform, R language, Rstudio, Python, Anaconda
3.	Мультимедийные средства в каждом компьютерном классе и в лекционной аудитории
4.	Браузер, сетевые коммуникационные средства для выхода в Интернет