



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное  
учреждение высшего образования  
**«Дальневосточный федеральный университет»**  
(ДФУ)

**ШКОЛА ЭКОНОМИКИ И МЕНЕДЖМЕНТА**

«СОГЛАСОВАНО»  
Руководитель ОП

(подпись)

Е.А. Тюрина  
(Ф.И.О. рук. ОП)

« 15 » января 20 21 г.

«УТВЕРЖДАЮ»

Заведующий департамента управления  
на основе данных (Data Driven Management Department)  
(название кафедры)

(подпись)

Ю. Д. Шмидт  
(Ф.И.О. зав. каф.)

« 15 » января 2021 г.

**РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ**

*анализ данных в R (продвинутый уровень)*

**Направление подготовки 38.04.01 Экономика**

*Магистерская программа «Экономические исследования»*

**Форма подготовки очная**

курс 2 семестр 3

лекции 8 час.

практические занятия 18 час.

лабораторные работы 0 час.

в том числе с использованием МАО лек. 0 /пр. 0 /лаб. 0 час.

всего часов аудиторной нагрузки 26 час.

в том числе с использованием МАО 0 час.

самостоятельная работа 82 час.

в том числе на подготовку к зачету час.

контрольные работы (количество)

курсовая работа / курсовой проект не предусмотрены

зачет с оценкой 3 семестр

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта по направлению подготовки 38.04.01 Экономика, утвержденного приказом Министерства науки и высшего образования российской Федерации от 11.08.2020 № 939.

Рабочая программа обсуждена на заседании департамента управления на основе данных (Data Driven Management Department) ШЭМ ДФУ № 4 от «15» января 2021 г.

Зав. департамента управления

на основе данных (Data Driven Management Department),

д-р экон. наук, проф.

Составитель

канд. экон. наук, доцент

Ю.Д. Шмидт

Е.В. Кочева

Владивосток  
2021

**I. Рабочая программа пересмотрена на заседании департамента управления на основе данных (Data Driven Management Department)**

Протокол от «\_\_\_\_\_» \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**II. Рабочая программа пересмотрена на заседании департамента управления на основе данных (Data Driven Management Department)**

Протокол от «\_\_\_\_\_» \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

## **Аннотация к рабочей программе дисциплины «Анализ данных в R (продвинутый уровень)»**

Учебный курс «Анализ данных в R (продвинутый уровень)» предназначен для студентов направления подготовки 38.04.01 Экономика профиля «Экономические исследования».

Дисциплина «Анализ данных в R (продвинутый уровень)» включена в состав вариативной части блока «Дисциплины (модули)», часть, формируемая участниками образовательных отношений.

Общая трудоемкость дисциплины составляет 3 зачетные единицы, 108 часов. Учебным планом предусмотрены лекционные занятия (8 часов), практические занятия (18 часов), самостоятельная работа студентов (82 часов). Дисциплина реализуется на 2 курсе в 3 семестре.

Дисциплина «Анализ данных в R (продвинутый уровень)» основывается на знаниях, умениях и навыках, полученных в результате изучения дисциплин «Skills for Study in Economics and Management (Академические навыки в области экономики и менеджмента)», «Статистика» и «Эконометрика», позволяет подготовить студентов к освоению ряда таких дисциплин, как «Аналитика и визуализация данных в Power BI», «Методы исследований», «Геоинформационные системы в экономике», а также подготовить к научно-исследовательской работе, к прохождению практики по получению профессиональных умений и опыта в аналитической деятельности и проектно-экономической деятельности, к преддипломной практике.

Содержание дисциплины состоит из четырех разделов и охватывает следующий круг вопросов:

1. Предобработка данных. Работа с data frame. Описательные статистики. Работа с пропущенными значениями. Циклы. Создание собственных функций. Функции семейства apply.

2. Работа с данными. Работа с данными в библиотеке dplyr. Преобразование данных в пакете Data.table.

3. Визуализация данных. Грамматика ggplot2. Facet как способ визуальной группировки данных. Динамическая визуализация в пакете plotly.

4. Многомерные методы анализа данных. Метод классификации данных – кластерный анализ. Логистическая регрессия. Метод главных компонент.

**Цель** – усвоение студентами теоретических знаний, формирование умений и практических навыков в области обработки и анализа данных, современных математических подходов к проблемам принятия решения в сложных ситуациях.

#### **Задачи:**

- формирование у студентов теоретических знаний статистического исследования различных социально-экономических процессов на макро- и микроуровне, на основе познания конкретных методов и методик анализа экономических показателей.

- выработка компетенций у будущих магистров, необходимых для успешного применения широкого инструментария статистического анализа данных при решении профессиональных задач, практической деятельности, а также при принятии управленческих решений.

- формирование навыков проведения статистических расчетов с применением компьютерной техники и современных статистических пакетов прикладных программ.

Для успешного изучения дисциплины «Анализ данных в R (продвинутый уровень)» у обучающихся должны быть сформированы следующие предварительные компетенции:

- способность выбрать инструментальные средства для обработки экономических данных в соответствии с поставленной задачей;

- способность эконометрические модели, интерпретировать полученные результаты;

- способность собирать необходимые данные, подготавливать информационный обзор.

- способность использовать для решения аналитических и исследовательских задач современные технические средства и информационные технологии.

В результате изучения данной дисциплины у обучающихся формируются следующие профессиональные (индикаторы компетенций).

Код и формулировка компетенции	Индикаторы формирования компетенции
ПК-2. Способен готовить аналитические материалы в интересах органов государственной власти и организаций различных форм собственности	ПК-2.2. Прогнозирует динамику и объясняет взаимосвязи основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом

## I. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ:

**Цель** – усвоение студентами теоретических знаний, формирование умений и практических навыков в области обработки и анализа данных, современных математических подходов к проблемам принятия решения в сложных ситуациях.

### **Задачи:**

- формирование у студентов теоретических знаний статистического исследования различных социально-экономических процессов на макро- и микроуровне, на основе познания конкретных методов и методик анализа экономических показателей.

- выработка компетенций у будущих магистров, необходимых для успешного применения широкого инструментария статистического анализа данных при решении профессиональных задач, практической деятельности, а также при принятии управленческих решений.

- формирование навыков проведения статистических расчетов с применением компьютерной техники и современных статистических пакетов прикладных программ.

Код и формулировка компетенции	Индикаторы формирования компетенции	
ПК-2.2 - Прогнозирует динамику и объясняет взаимосвязи основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом	знает	основные понятия и теоретические положения статистического анализа данных; методы анализа, систематизации и обобщения информации, цели и пути их достижения, современные информационные ресурсы и технологии;
	умеет	применять научные термины, формулировать цели, анализировать и обобщать информацию, использовать современные информационные ресурсы и технологии при анализе, прогнозировании и оценке социально-экономических показателей;
	умеет	применять методы анализа, систематизации и обобщения данных, формулировать цели, задачи и пути их достижения, использовать информационные ресурсы и технологии при анализе и прогнозировании данных.

## II. ТРУДОЁМКОСТЬ ДИСЦИПЛИНЫ И ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ ПО ДИСЦИПЛИНЕ

Общая трудоемкость дисциплины составляет 3 зачётные единицы (108 академических часов).

### Структура дисциплины

№	Наименование раздела дисциплины	Семестр	Количество часов по видам учебных занятий и работы обучающегося					Формы промежуточной аттестации, текущего контроля успеваемости
			лекции	лабораторные работы	практические занятия	онлайн-курс	самостоятельная работа	
1	Тема 1. Предобработка данных	1	2	-	5	-	-	Контрольная работа
2	Тема 2. Работа с данными	1	2	-	5	-	-	Контрольная работа
3	Тема 3. Визуализация данных	1	2	-	5	-	-	Контрольная работа
4	Тема 4. Многомерные методы анализа данных	1	2	-	3	-	-	Контрольная работа
Итого:		-	8	-	18	-	-	-

### **III. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА**

#### **Тема 1. Предобработка данных (2 час.)**

Работа с data frame. Описательные статистики. Работа с пропущенными значениями. Циклы. Создание собственных функций. Функции семейства apply.

#### **Тема 2. Работа с данными (2 час.)**

5. Работа с данными в библиотеке dplyr. Преобразование данных в пакете Data.table.

#### **Тема 3. Визуализация данных (2 час.)**

Грамматика ggplot2. Facet как способ визуальной группировки данных. Динамическая визуализация в пакете plotly.

#### **Тема 4. Многомерные методы анализа данных (2 час.)**

Метод классификации данных – кластерный анализ. Логистическая регрессия. Метод главных компонент.

### **IV. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА И САМОСТОЯТЕЛЬНОЙ РАБОТЫ**

#### **Практические занятия (18 часов)**

##### **Занятие 1. Предобработка данных (5 час.)**

1. Работа с data frame.
2. Описательные статистики.
3. Работа с пропущенными значениями.
4. Циклы.
5. Создание собственных функций. Функции семейства apply.

##### **Занятие 2. Работа с данными (5 час.)**

1. Работа с данными в библиотеке dplyr.
2. Преобразование данных в пакете Data.table.

### **Занятие 3. Визуализация данных (5 час.)**

1. Грамматика ggplot2.
2. Facet как способ визуальной группировки данных.
3. Динамическая визуализация в пакете plotly.

### **Занятие 4. Многомерные методы анализа данных (3 час.)**

1. Метод классификации данных – кластерный анализ.
2. Логистическая регрессия.
3. Метод главных компонент.

## **V. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ**

Учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине «Анализ данных в R (продвинутый уровень)» включает в себя:

- план-график выполнения самостоятельной работы по дисциплине, в том числе примерные нормы времени на выполнение по каждому заданию;
- характеристика заданий для самостоятельной работы обучающихся и методические рекомендации по их выполнению;
- требования к представлению и оформлению результатов самостоятельной работы;
- критерии оценки выполнения самостоятельной работы.

<b>№ п/п</b>	<b>Дата / сроки выполнения</b>	<b>Вид самостоятельной работы</b>	<b>Примерные нормы времени на выполнение</b>	<b>Форма контроля</b>
1	Первая, вторая недели.	Подготовка к практическим занятиям. Изучение методических материалов, литературы.	7	Устный опрос (анализ и обсуждение актуальной терминологии).
2	Третья, четвертая недели.	Подготовка к практическим занятиям.	7	Устный опрос (анализ и обсуждение актуальной терминологии).



		занятиям. Изучение методических материалов, литературы.		возможностей использования ППО). Контрольная работа
3	Пятая, шестая недели.	Подготовка к практическому занятию. Изучение методических материалов.	7	Демонстрация выполнения заданий, разбор ошибок.
4	Седьмая, восьмая недели.	Подготовка к практическому занятию. Подготовка к контрольной работе: изучение теоретических материалов	7	Работа в режиме дискуссии. Обсуждение типичных ошибок. Контрольная работа.
5	Девятая, десятая недели.	Подготовка к практическому занятию.	7	Устный опрос.
6	Одиннадцатая, двенадцатая недели.	Подготовка к практическому занятию. Подготовка к контрольной работе: изучение теоретических материалов.	7	Устный опрос (анализ и обсуждение). Контрольная работа.
7	Тринадцатая, четырнадцатая недели.	Подготовка к практическим занятиям.	7	Демонстрация выполнения заданий, разбор ошибок.
8	Пятнадцатая, шестнадцатая недели.	Подготовка к практическим занятиям. Изучение методических материалов. Подготовка к контрольной работе: изучение теоретических материалов	7	Устный опрос (анализ и обсуждение). Контрольная работа
9	Семнадцатая, восемнадцатая недели.	Доработка творческого задания и формирование пояснительной записки.	7	Проверка пояснительной записки, творческого задания
10	Восемнадцатая неделя	Подготовка к промежуточной аттестации.	9	Экзамен
ИТОГО			72	

## **Рекомендации по самостоятельной работе студентов**

В соответствии с учебным планом дисциплины предусмотрены следующие виды самостоятельной работы студентов:

- изучение рекомендованной литературы, подготовка к практическим занятиям, контрольным работам;
- выполнение самостоятельных практических заданий;
- подготовка отчетов по текущим работам.

## **Требования к оформлению и объему контрольной работы**

Решение практических задач должно быть представлено в виде аналитической записки в MS Word. Объем не более 5 страниц, шрифт Times New Roman, размер шрифта 14, межстрочный интервал – одинарный. Колонтитулы – 1,25 см; ориентация книжная, красная строка - 1,5 см.; автоперенос. Работа должна содержать иллюстративный материал. Формулы и уравнения следует выделять из текста в отдельную строку.

## **Шкала оценивания**

При проверке контрольного задания и подведении общих итогов предлагается следующая методика оценки: оценка ответов осуществляется по пятибалльной системе по следующей шкале.

Количество баллов	Оценка
100-86	«Отлично»
85-76	«Хорошо»
75-61	«Удовлетворительно»
менее 60	«Неудовлетворительно»

100-86 баллов выставляется, если студент/группа выразили своё мнение по сформулированной проблеме, аргументировали его, точно определив ее содержание и составляющие. Приведены данные отечественной и зарубежной литературы, статистические сведения, информация нормативно-правового характера. Продемонстрировано знание и владение навыком

самостоятельной исследовательской работы по теме исследования. Фактических ошибок, связанных с пониманием проблемы, нет.

85-76 баллов – работа студента/группы характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Для аргументации приводятся данные отечественных и зарубежных авторов. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет.

75-61 балл – проведен достаточно самостоятельный анализ основных этапов и смысловых составляющих проблемы; понимание базовых основ и теоретического обоснования выбранной темы. Привлечены основные источники по рассматриваемой теме. Допущено не более 2 ошибок в смысле или содержании проблемы

Менее 60 баллов – если работа представляет собой пересказанный текст без комментариев, анализа. Не раскрыта структура и теоретическая составляющая темы. Допущено три или более трех ошибок смыслового содержания раскрываемой проблемы.

## VI. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые разделы / темы дисциплины	Коды и этапы формирования компетенций		Оценочные средства	
				текущий контроль	промежуточная аттестация
1	Тема 1. Предобработка данных	ПК-2.2	знает	Собеседование (УО-1)	Собеседование (УО-1), вопросы к зачету: 1-4
			умеет	Контрольная работа (ПР-2)	
2	Тема 2. Работа с данными	ПК-2.2	знает	Собеседование (УО-1)	Собеседование (УО-1), вопросы к зачету: 5-8
			умеет	Контрольная работа (ПР-2)	
3	Тема 3. Визуализация данных	ПК-2.2	знает	Собеседование (УО-1)	Собеседование (УО-1), вопросы к зачету: 9-13
			умеет	Контрольная работа (ПР-2)	
4	Тема 4. Многомерные методы анализа данных	ПК-2.2	знает	Собеседование (УО-1)	Собеседование (УО-1), вопросы к зачету: 14-16
			умеет	Контрольная работа (ПР-2)	

Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков и (или) опыта деятельности, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в разделе ФОС.

## **VII. СПИСОК УЧЕБНОЙ ЛИТЕРАТУРЫ И ИНФОРМАЦИОННО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **Основная литература:**

1. Язык и среда программирования R : учеб. пособие / А.В. Золотарюк. – М.: ИНФРА-М, 2019. – 162 с. – Режим доступа: <http://znanium.com/catalog/product/978863>

2. Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 с. — ISBN 978-5-4497-0289-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/89404.html>

3. Дадян, Э. Г. Данные: хранение и обработка : учебник / Э.Г. Дадян. — Москва : ИНФРА-М, 2020. — 205 с. — (Среднее профессиональное образование). - ISBN 978-5-16-015663-7. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1045133>

4. Математическая статистика и анализ данных: учебное пособие / А. С. Мельниченко. М. : Издательский Дом МИСиС, 2018. — 45 с. – Режим доступа: <http://www.iprbookshop.ru/78563.html>

5. Плеханова, Т. И. Статистика : учебное пособие для СПО / Т. И. Плеханова, Т. В. Лебедева. — Саратов : Профобразование, 2020. — 418 с. — ISBN 978-5-4488-0660-5. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/92171.html>

6. Мельниченко А.С. Математическая статистика и анализ данных [Электронный ресурс]: учебное пособие/ Мельниченко А.С. – Электрон. текстовые данные.— М.: Издательский Дом МИСиС, 2018. – 45 с. – Режим доступа: <http://www.iprbookshop.ru/78563.html>.

7. Дубина И.Н. Математико-статистические методы и инструменты в эмпирических социально-экономических исследованиях [Электронный ресурс]: учебное пособие/ Дубина И.Н.— Электрон. текстовые данные.— Саратов: Вузовское образование, 2018.— 415 с.— Режим доступа: <http://www.iprbookshop.ru/76234.html>.

8. Сальникова, К. В. Практические основы статистики и эконометрического моделирования : учебное пособие / К. В. Сальникова. — Саратов : Ай Пи Ар Медиа, 2020. — 385 с. — ISBN 978-5-4497-0427-6. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/91121.html>

9. Орлов, А. И. Эконометрика : учебное пособие / А. И. Орлов. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 676 с. — ISBN 978-5-4497-0362-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/89481.html>

10. Плеханова, Т. И. Статистика : учебное пособие для СПО / Т. И. Плеханова, Т. В. Лебедева. — Саратов : Профобразование, 2020. — 418 с. — ISBN 978-5-4488-0660-5. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/92171.html>

#### **Дополнительная литература**

1. Иванов, Ю. Н. Экономическая статистика : учебник / под ред. Ю.Н. Иванова. — 5-е изд., перераб. и доп. — Москва : ИНФРА-М, 2020. — 584 с. + Доп. материалы [Электронный ресурс]. — (высшее образование: Бакалавриат). — DOI 10.12737/7728. - ISBN 978-5-16-010399-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1063310>

2. Методы хранения и обработки данных: Учебник / Дадян Э.Г. - М.:НИЦ ИНФРА-М, 2018: 60x90 1/16 ISBN 978-5-16-107039-0 (online) - Режим доступа: <http://znanium.com/catalog/product/989190>

3. Бонцанини, М. Анализ социальных медиа на Python. Извлекайте и анализируйте данные из всех уголков социальной паутины на Python [Электронный ресурс] / М. Бонцанини ; пер. с англ. А. В. Логунова. — Электрон. дан. — Москва : ДМК Пресс, 2018. — 288 с. — Режим доступа: <https://e.lanbook.com/book/108129>.

4. Язык и среда программирования R: Учебное пособие / Золотарюк А.В. - М.:НИЦ ИНФРА-М, 2018. - 183 с.: 60x90 1/16. - (Высшее образование: Бакалавриат) ISBN 978-5-16-107182-3 (online) - Режим доступа: <http://znanium.com/catalog/product/997099>

5. Баженов Р.И. Интеллектуальные информационные технологии в управлении [Электронный ресурс]: учебное пособие/ Баженов Р.И.— Электрон. текстовые данные.— Саратов: Ай Пи Эр Медиа, 2018.— 117 с.— Режим доступа: <http://www.iprbookshop.ru/72801.html>.

6. Статистический анализ данных в MS Excel : учеб. пособие / А.Ю. Козлов, В.С. Мхитарян, В.Ф. Шишов. — М. : ИНФРА-М, 2019. — 320 с. — (Высшее образование: Бакалавриат). — [www.dx.doi.org/10.12737/2842](http://www.dx.doi.org/10.12737/2842). - Режим доступа: <http://znanium.com/catalog/product/987337>

7. Уэс Маккинли Python и анализ данных [Электронный ресурс]/ Уэс Маккинли— Электрон. текстовые данные.— Саратов: Профобразование, 2017.— 482 с.— Режим доступа: <http://www.iprbookshop.ru/64058.html>.

8. Маккинни, У. Python и анализ данных / У. Маккинни ; перевод с английского А. А. Слинкина. — 2-ое изд., испр. и доп. — Москва : ДМК Пресс, 2020. — 540 с. — ISBN 978-5-97060-590-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/131721>

9. Воскобойников, Ю. Е. Эконометрика в Excel. Модели временных рядов : учебное пособие / Ю. Е. Воскобойников. — 2-е изд., стер. — Санкт-Петербург : Лань, 2020. — 152 с. — ISBN 978-5-8114-4863-0. — Текст :

электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/126706>

10. Цыпин, А. П. Статистика в табличном редакторе Microsoft Excel : практикум для СПО / А. П. Цыпин, Л. Р. Фаизова. — Саратов : Профобразование, 2020. — 288 с. — ISBN 978-5-4488-0623-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/92170.html>

### **Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

- Федеральная служба государственной статистики [Электронный ресурс]. – Режим доступа: URL: <http://www.gks.ru>
- Центральная база статистических данных [Электронный ресурс]. – Режим доступа: URL: <http://www.gks.ru/dbscripts/Cbsd/DBInet.cgi>
- Единая межведомственная информационно-статистическая система [Электронный ресурс]. – Режим доступа: URL: <http://www.fedstat.ru/>

### **Программное обеспечение**

В процессе обучения магистрантов по дисциплине «Анализ данных в R (продвинутый уровень)» возможно использование следующих программных продуктов:

- Microsoft Office Excel;
- RStudio.

## **VIII. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

Реализация дисциплины « Анализ данных в R (продвинутый уровень)» предусматривает следующие виды учебной работы: лекции, лабораторные и практические занятия, самостоятельную работу студентов, текущий контроль и промежуточную аттестацию.

Освоение курса дисциплины « Анализ данных в R (продвинутый

уровень)» предполагает рейтинговую систему оценки знаний студентов и предусматривает со стороны преподавателя текущий контроль за посещением студентами лекций, подготовкой и выполнением всех практических работ с обязательным предоставлением отчета о работе, выполнением всех видов самостоятельной работы.

Промежуточной аттестацией по дисциплине «Анализ данных в R (продвинутый уровень)» является экзамен, который проводится в виде тестирования.

В течение учебного семестра обучающимся нужно:

- освоить теоретический материал (20 баллов);
- успешно выполнить аудиторные и контрольные задания (50 баллов);
- своевременно и успешно выполнить все виды самостоятельной работы (30 баллов).

Студент считается аттестованным по дисциплине «Анализ данных в R (продвинутый уровень)» при условии выполнения всех видов текущего контроля и самостоятельной работы, предусмотренных учебной программой.

Критерии оценки по дисциплине «Анализ данных в R (продвинутый уровень)» для аттестации на экзамене следующие: 86-100 баллов – «отлично», 76-85 баллов – «хорошо», 61-75 баллов – «удовлетворительно», 60 и менее баллов – «неудовлетворительно».

Пересчет баллов по текущему контролю и самостоятельной работе производится по формуле:

$$P(n) = \sum_{i=1}^m \left[ \frac{O_i}{O_i^{max}} \times \frac{k_i}{W} \right],$$

где:  $W = \sum_{i=1}^n k_i^n$  для текущего рейтинга;

$W = \sum_{i=1}^m k_i^n$  для итогового рейтинга;

$P(n)$  – рейтинг студента;

$m$  – общее количество контрольных мероприятий;

$n$  – количество проведенных контрольных мероприятий;

$O_i$  – балл, полученный студентом на  $i$ -ом контрольном мероприятии;



$O_i^{max}$  – максимально возможный балл студента по  $i$ -му контрольному мероприятию;

$k_i$  – весовой коэффициент  $i$ -го контрольного мероприятия;

$k_i^n$  – весовой коэффициент  $i$ -го контрольного мероприятия, если оно является основным, или 0, если оно является дополнительным.

Приступая к изучению дисциплины, студенту необходимо в первую очередь ознакомиться с содержанием РПД. Лекции имеют целью дать систематизированные основы научных знаний о изучаемом предмете, основных методах исследования, применяющихся в экономической статистике.

При изучении и проработке теоретического материала студентам необходимо:

- повторить законспектированный на лекционном занятии материал и дополнить его с учетом рекомендованной по данной теме литературы;
- при самостоятельном изучении теоретической темы сделать конспект, используя рекомендованные в РПД литературные источники.
- при подготовке к текущему и промежуточному контролю использовать материалы РПД.

Практические занятия проводятся с целью углубления и закрепления знаний, полученных на лекциях и в процессе самостоятельной работы над учебной и научной литературой. При подготовке к практическому занятию для студентов очной формы обучения необходимо:

- изучить, повторить теоретический материал по заданной теме;
- изучить материалы практического решения задач по заданной теме, уделяя особое внимание расчетным формулам;
- при выполнении домашних расчетных заданий, изучить, повторить типовые задания, выполняемые в аудитории.

Работа с учебной и научной литературой является главной формой самостоятельной работы и необходима при подготовке к устному опросу на практических, к контрольным работам, тестированию, экзамену. Она включает проработку лекционного материала – изучение рекомендованных

источников и литературы по тематике лекций. Конспект лекции должен содержать реферативную запись основных вопросов лекции, предложенных преподавателем схем (при их демонстрации), основных источников и литературы по темам, выводы по каждому вопросу. Конспект должен быть выполнен в отдельной тетради по предмету. Он должен быть аккуратным, хорошо читаемым, не содержать не относящуюся к теме информацию или рисунки.

Конспекты научной литературы при самостоятельной подготовке к занятиям должны быть выполнены также аккуратно, содержать ответы на каждый поставленный в теме вопрос, иметь ссылку на источник информации с обязательным указанием автора, названия и года издания используемой научной литературы. Конспект может быть опорным (содержать лишь основные ключевые позиции), но при этом позволяющим дать полный ответ по вопросу, может быть подробным. Объем конспекта определяется самим студентом.

В процессе работы с учебной и научной литературой студент может:

- делать записи по ходу чтения в виде простого или развернутого плана (создавать перечень основных вопросов, рассмотренных в источнике);
- составлять тезисы (цитирование наиболее важных мест статьи или монографии, короткое изложение основных мыслей автора);
- готовить аннотации (краткое обобщение основных вопросов работы);
- создавать конспекты (развернутые тезисы).

Работу с литературой следует начинать с анализа соответствующего раздела РПД, в котором перечислены основная и дополнительная литература, учебно-методические издания необходимые для изучения дисциплины и работы на практических занятиях.

Выбрав нужный источник, следует найти интересующий раздел по оглавлению или алфавитному указателю, а также одноименный раздел конспекта лекций или учебного пособия. В случае возникших затруднений в понимании учебного материала следует обратиться к другим источникам, где изложение может оказаться более доступным. Необходимо отметить, что

работа с литературой не только полезна как средство более глубокого изучения любой дисциплины, но и является неотъемлемой частью профессиональной деятельности будущего выпускника.

Методические указания по организации внеаудиторной самостоятельной работы на занятии способствуют организации последовательного изучения материала, вынесенного на самостоятельное освоение в соответствии с учебным планом, и имеет такую структуру как:

- тема;
- вопросы и содержание материала для самостоятельного изучения;
- форма выполнения задания;
- алгоритм выполнения и оформления самостоятельной работы;
- критерии оценки самостоятельной работы.

Самостоятельная работа как вид деятельности студента многогранна. В качестве форм самостоятельной работы при изучении дисциплины «Анализ данных в R (продвинутый уровень)» предлагаются:

- работа с научной и учебной литературой;
- подготовка к практическому занятию;
- более глубокое изучение с вопросами, изучаемыми на практических занятиях;

- подготовка индивидуального и групповых заданий
- подготовка к тестированию и экзамену;

Задачи самостоятельной работы:

- обретение навыков самостоятельной научно-исследовательской работы на основании анализа текстов литературных источников и применения различных методов исследования;

- выработка умения самостоятельно и критически подходить к изучаемому материалу.

Технология СР должна обеспечивать овладение знаниями, закрепление и систематизацию знаний, формирование умений и навыков. Апробированная

технология характеризуется алгоритмом, который включает следующие логически связанные действия студента:

- чтение текста (учебника, пособия, конспекта лекций);
- конспектирование текста;
- решение задач и упражнений;
- ответы на контрольные вопросы;
- составление планов и тезисов ответа.
- рекомендации по написанию группового домашнего задания

Тема задания выбирается в соответствии с интересами студента. По результатам работы группа студентов сдает преподавателю отчет о проделанной работе. Важно, чтобы в отчете, во-первых, были освещены как естественнонаучные, так и социальные стороны проблемы; а во-вторых, представлены теоретические положения и конкретные примеры. Особенно приветствуется использование собственных примеров из окружающей студентов жизни.

Отчет должен основываться на проработке нескольких дополнительных к основной литературе источников. Как правило, это научные монографии или статьи. План отчета должен быть авторским. Все приводимые в отчете факты и заимствованные соображения должны сопровождаться ссылками на источник информации. Недопустимо просто скомпоновать реферат из кусков заимствованного текста. Все цитаты должны быть представлены в кавычках с указанием в скобках источника и страницы.

Отчет оформляется в соответствии с требованиями ДВФУ к оформлению курсовых и выпускных квалификационных работ. Сам текст работы желательно подразделить на разделы. Завершают отчет разделы «Заключение» и «Список использованной литературы». В заключении должны быть представлены основные выводы, ясно сформулированные в тезисной форме и, обычно, пронумерованные. Список литературы должен быть составлен в полном соответствии с действующим стандартом (правилами), включая особую расстановку знаков препинания.

## IX. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Учебные занятия по дисциплине проводятся в помещениях, оснащенных соответствующим оборудованием и программным обеспечением. Перечень материально-технического и программного обеспечения дисциплины приведен в таблице.

Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
<p>690922, г. Владивосток, остров Русский, полуостров Саперный, поселок Аякс, 10, корп. G, ауд. G507, учебная аудитория для проведения занятий лекционного типа; учебная аудитория для проведения занятий семинарского типа (практических занятий); учебная аудитория для курсового проектирования (выполнения курсовых работ); учебная аудитория для текущего контроля и промежуточной аттестации</p>	<p>40 посадочных мест, автоматизированное рабочее место преподавателя, переносная магнитно-маркерная доска, Wi-Fi Ноутбук Acer ExtensaE2511-30VO Экран с электроприводом 236*147 см Trim Screen Line; Проектор DLP, 3000 ANSI Lm, WXGA 1280x800, 2000:1 EW330U Mitsubishi; Подсистема специализированных креплений оборудования CORSA-2007 Tuarex; Подсистема видеокоммутации; Подсистема аудиокоммутации и звукоусиления; акустическая система для потолочного монтажа SI 3CT LP Extron; цифровой аудиопроцессор DMP 44 LC Extron.</p>	

- Специализированные пакеты программ: *RStudio, MS Excel*.
- Материалы лекций в виде презентаций.
- Учебная аудитория с мультимедийным проектором и экраном.

В читальных залах Научной библиотеки ДВФУ предусмотрены рабочие места для людей с ограниченными возможностями здоровья, оснащены дисплеями и принтерами; оборудованные портативными устройствами для чтения плоскочечатных текстов, сканирующими и читающими машинами, видеоувеличителем с возможностью регуляции цветовых спектров; увеличивающими электронными лупами и ультразвуковыми маркировщиками.

В целях обеспечения специальных условий обучения инвалидов и лиц с ограниченными возможностями здоровья ДВФУ все здания оборудованы пандусами, лифтами, подъемниками, специализированными местами, оснащенными туалетными комнатами, табличками информационно-навигационной системы.

## **Х. ФОНДЫ ОЦЕНОЧНЫХ СРЕДСТВ**

Фонды оценочных средств включают в себя: перечень форм оценивания, применяемых на различных этапах формирования компетенций в ходе освоения дисциплины модуля, шкалу оценивания каждой формы, с описанием индикаторов достижения освоения дисциплины согласно заявленных компетенций, примеры заданий текущего и промежуточного контроля.

Заключение работодателя на ФОС отражено в экспертном заключении работодателя на основную профессиональную образовательную программу.

### **Паспорт фонда оценочных средств**

<b>Код и формулировка компетенции</b>	<b>Индикаторы формирования компетенции</b>	
ПК-2.2 - Прогнозирует динамику и объясняет взаимосвязи основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом	знает	основные понятия и теоретические положения статистического анализа данных; методы анализа, систематизации и обобщения информации, цели и пути их достижения, современные информационные ресурсы и технологии;
	умеет	применять научные термины, формулировать цели, анализировать и обобщать информацию, использовать современные информационные ресурсы и технологии при анализе, прогнозировании и оценке социально-

		экономических показателей;
	умеет	применять методы анализа, систематизации и обобщения данных, формулировать цели, задачи и пути их достижения, использовать информационные ресурсы и технологии при анализе и прогнозировании данных.

№ п/п	Контролируемые разделы / темы дисциплины	Коды и этапы формирования компетенций	Оценочные средства		
			текущий контроль	промежуточная аттестация	
1	Тема 1. Введение в данные	ПК-2.2	знает	Собеседование (УО-1)	Собеседование (УО-1), вопросы к зачету: 1-4
			умеет	Контрольная работа (ПР-2)	
2	Тема 2. Статистический анализ выборочных данных	ПК-2.2	знает	Собеседование (УО-1)	Собеседование (УО-1), вопросы к зачету: 5-9
			умеет	Контрольная работа (ПР-2)	
3	Тема 3. Исследование статистических взаимосвязей	ПК-2.2	знает	Собеседование (УО-1)	Собеседование (УО-1), вопросы к зачету: 10-13
			умеет	Контрольная работа (ПР-2)	
4	Тема 4. Сравнение и создание групп	ПК-2.2	знает	Собеседование (УО-1)	Собеседование (УО-1), вопросы к зачету: 14-16
			умеет	Контрольная работа (ПР-2)	

### Шкала оценивания уровня сформированности компетенций

Код и формулировка компетенции (индикатора)	Этапы формирования компетенции		Критерии	Показатели
ПК-2.2 - Прогнозирует динамику и объясняет взаимосвязи основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом	знает (пороговый уровень)	основные понятия и теоретические положения изучаемых дисциплин; методы анализа, систематизации и обобщения информации, цели и пути их достижения	социально-значимые проблемы и процессы в определенной области знаний и профессиональной деятельности и методы их анализа	способность дать определения основных понятий предметной области исследования; способность перечислить и раскрыть суть методов прогнозирования; способность обосновать актуальность выполняемого задания или исследования; способность перечислить источники информации по методам и подходам к проведению исследований
	умеет (продвинутый)	применять научные		Умение работать с электронными

		термины, формулировать цели, анализировать и обобщать информацию	базами данных и библиотечными каталогами, умение применять известные методы моделирования и прогнозирования бизнес-процессов, умение	для исследования; способность оперировать научными определениями относительно объекта и предмета исследования;
	владеет (высокий)	инструментами и методами проведения научных исследований, методами экономико-математического моделирования и прогнозирования, обоснования эффективности бизнес-проектов с использованием пакетов прикладных программ	владение терминологией предметной области знаний, владение способностью сформулировать задание по научному исследованию, чёткое понимание требований, предъявляемых к содержанию и последовательности исследования, владение инструментами визуализации результатов научных исследований	способность применять терминологию предметной области исследования в устных ответах на вопросы и в письменных работах, способность сформулировать задание по научному исследованию; способность проводить самостоятельные исследования и представлять их результаты на обсуждение на круглых столах, семинарах, научных конференциях

### **Методические рекомендации, определяющие процедуры оценивания результатов освоения дисциплины**

**Текущая аттестация студентов.** Текущая аттестация студентов по дисциплине «Анализ данных в R (продвинутый уровень)» проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

Текущая аттестация по дисциплине «Анализ данных в R (продвинутый уровень)» проводится в форме контрольных мероприятий по оцениванию фактических результатов обучения студентов и осуществляется ведущим преподавателем.

Объектами оценивания выступают:

- учебная дисциплина (активность на занятиях, своевременность выполнения различных видов заданий, посещаемость всех видов занятий по аттестуемой дисциплине);

- степень усвоения теоретических знаний (активность в ходе обсуждений теоретических материалов, активное участие в дискуссиях с аргументами из дополнительных источников, внимательность, способность



задавать встречные вопросы в рамках дискуссии или обсуждения, заинтересованность изучаемыми материалами);

– уровень овладения практическими умениями и навыками по всем видам учебной работы (определяется по результатам контрольных работ, практических занятий);

– результаты самостоятельной работы.

**Промежуточная аттестация студентов.** Промежуточная аттестация студентов по дисциплине «Анализ данных в R (продвинутый уровень)» проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

**Вид промежуточной аттестации – зачет** (3 семестр), который выставляется по результатам собеседования, а также результатам текущей аттестации.

## **Зачетно-экзаменационные материалы**

### **Оценочные средства для промежуточной аттестации**

#### **Вопросы к зачету**

1. Форматы данных
2. Числовые векторы и факторы
3. Пропущенные данные
4. Матрицы и списки
5. Таблицы данных
6. Векторизированные вычисления
7. Описательные статистики
8. Создание функций
9. Циклы и условные операторы
10. Графические устройства
11. Графические опции
12. Интерактивность
13. Выбросы
14. Классификация без обучения
15. Классификация с обучением
16. Вероятность успеха или логистическая регрессия

Реализация ситуационных задач в RStudio.

### Критерии выставления оценки студенту на зачете с оценкой

Набранные в течение семестра баллы (рейтинговой оценки)	Оценка зачета/ экзамена (стандартная)	Требования к сформированным компетенциям
86-100	«зачтено»/ «отлично»	Оценка «зачтено»/«отлично» выставляется студенту, 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, владеет разносторонними навыками и приемами выполнения практических задач.
76-85	«зачтено»/ «хорошо»	Оценка «зачтено»/«хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.
61-75	«зачтено»/ «удовлетворительно»	Оценка «зачтено»/«удовлетворительно» выставляется студенту, 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 ответах на дополнительные вопросы.
менее 61	«не зачтено»/ «неудовлетворительно»	Оценка «не зачтено»/«неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.

### Оценочные средства для текущей аттестации ( типовые ОС по текущей аттестации и критерии оценки по каждому виду

## аттестации по дисциплине «Анализ данных в R (продвинутый уровень)»)

Текущая аттестация студентов заключается в выполнении практических и контрольных работ, а также устных бесед со студентами в течение семестра по пройденному материалу. Типовые задачи контрольных работ и практических занятий представлены ниже.

**Типовая задача 1.** Напишите функцию `NA.position`, которая выводит номера позиций пропущенных наблюдений в векторе. На вход функция получает числовой вектор с пропущенными значениями. Функция возвращает новый вектор с номерами позиций пропущенных значений.

Подсказка: `is.na(x)` отличный аргумент, позволяет найти пропущенные значения! Определите пропущенные значения через индексацию, как вариант!

**Типовая задача 2.** Напишите функцию `NA.counter` для подсчета пропущенных значений в векторе. На вход функция `NA.counter` должна принимать один аргумент - числовой вектор. Функция должна возвращать количество пропущенных значений.

**Типовая задача 3.** Напишите функцию `filtered.sum`, которая на вход получает вектор с пропущенными, положительными и отрицательными значениями и возвращает сумму положительных элементов вектора.

**Типовая задача 4.** Напишите функцию `outliers.rm`, которая находит и удаляет выбросы. Для обнаружения выбросов воспользуемся самым простым способом, используя график `Box plot`. Выбросами будем считать те наблюдения, которые отклоняются от первого или третьего квартиля больше чем на  $(1,5 * IQR)$ , где `IQR` - межквартильный размах. На вход функция получает числовой вектор `x`. Функция должна возвращать модифицированный вектор `x` с удаленными выбросами.

Подсказка: посмотрите справку по функции `quantile`. Не забудьте, что тип данных у вас числовой (`as.numeric`). На ретерне обратитесь к вектору и задайте условие по отбору нужных значений (без выбросов).

**Типовая задача 5.** Напишите функцию, которая подсчитывает и выводит в консоль те наблюдения, которые отклонились от среднего по вектору больше, чем на 2 стандартных отклонения. Если таких наблюдений нет, то функция вернет фразу “There are no outliers”. Таким образом, мы найдем выбросы!

**Типовая задача 6.** произвольным числом количественных переменных. Первая колонка в данных - это зависимая переменная, все остальные - предикторы. На первом этапе вы должны отобрать предикторы для модели. Функция возвращает в виде вектора коэффициенты линейной регрессии построенной только для отобранных предикторов (условие нормальности распределения). Если таких предикторов в данных не оказалось, то функция возвращает предупреждение “There are no normal variables in the data”.

**Типовая задача 7.** Для встроенных в R данных `AirPassengers` рассчитайте скользящее среднее с интервалом сглаживания равным 10. Напечатайте получившийся результат (первым значением в выводе должно быть среднее для элементов 1:10, во втором значении - среднее для элементов 2:11 и т.д., в последнем - среднее для элементов 135 :144)

**Типовая задача 8.** Напишите функцию `smart_hclust`, которая получает на вход `dataframe` с произвольным числом количественных переменных и число кластеров, которое необходимо выделить при помощи иерархической кластеризации. Функция должна в исходный набор данных добавлять новую переменную фактор - `cluster` – номер кластера, к которому отнесено каждое из наблюдений.

**Типовая задача 9.** В данных `swiss`, используя все переменные, выделите два кластера при помощи иерархической кластеризации и сохраните значение кластеров как фактор в переменную `cluster`. Затем визуализируйте взаимосвязь переменных `Education` и `Catholic` в двух выделенных кластерах.

**Типовая задача 10.** Напишите функцию `get_pca2`, которая принимает на вход `dataframe` с произвольным числом количественных переменных. Функция должна рассчитать, какое минимальное число главных компонент объясняет больше 90% изменчивости в исходных данных и добавлять значения этих компонент в исходный `dataframe` в виде новых переменных.

### Критерии оценки выполнения контрольных и практических работ

Результаты выполнения контрольных работ и практических занятий оцениваются по 100 бальной шкале, как процент правильно решенных задач.

#### Шкала оценивания практический заданий

Количество баллов	Оценка
100-86	«Отлично»
85-76	«Хорошо»
75-61	«Удовлетворительно»
менее 60	«Неудовлетворительно»

#### Критерии оценки устного опроса

Баллы (рейтинговой оценки)	Требования к сформированным компетенциям
86–100	Оценка «отлично» выставляется студенту, 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, владеет разносторонними навыками и приемами выполнения практических задач.
76–85	Оценка «хорошо» выставляется студенту, если он твердо знает материал,

	грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.
75–61	Оценка «удовлетворительно» выставляется студенту, 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 ответах на дополнительные вопросы.
менее 61	Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.

### **Оценочные средства**

**для проверки сформированности компетенций (индикаторов)**

**по дисциплине «Анализ данных в R (продвинутый уровень)»**

<b>Код и формулировка индикатора</b>	<b>Задание</b>
ПК-2.2 - Прогнозирует динамику и объясняет взаимосвязи основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом	Собрать данные и провести анализ (в соответствии с целями исследований). Результат представить в виде подробного отчета в R Markdown