

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Ястребов Олег Александрович  
Должность: Ректор  
Дата подписания: 21.06.2022 16:57:12  
Уникальный программный ключ:  
ca953a0120d891083f939673078ef1a9896ae18a

Федеральное государственное автономное образовательное учреждение высшего образования

«Российский университет дружбы народов»

Факультет физико-математических и естественных наук

(наименование основного учебного подразделения (ОУП)-разработчика ОП ВО)

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Основы анализа больших данных

(наименование дисциплины/модуля)

**Рекомендована МССН для направления подготовки:**

02.03.01 Математика и компьютерные науки

(код и наименование направления подготовки/специальности)

**Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):**

Математика и компьютерные науки

(наименование (профиль/специализация) ОП ВО)

2022 г.

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Основы анализа больших данных» является приобретение знаний и навыков работы с большими данными в различных прикладных пакетах, в особенности в пакете R.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Основы анализа больших данных» направлено на формирование у обучающихся следующих компетенций (части компетенций): УК-1, УК-12, ОПК-1, ОПК-4, ПК-1.

Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1. Знает принципы сбора, отбора и обобщения информации, методика системного подхода для решения профессиональных задач
		УК-1.2. Умеет анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности
		УК-1.3. Владеет навыками научного поиска и практической работы с информационными источниками; методами принятия решений
УК-12	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	УК-12.1. Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных
ОПК-1	Способен консультировать и использовать фундаментальные	ОПК-1.1 Обладает базовыми знаниями, полученными в области математических и

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	знания в области математического анализа, комплексного и функционального анализа алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений, дискретной математики и математической логики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики в профессиональной деятельности	(или) естественных наук
		ОПК-1.2. Умеет использовать базовые знания, полученные в области математических и (или) естественных наук, в профессиональной деятельности
		ОПК-1.3. Имеет навыки выбора методов решения задач профессиональной деятельности на основе теоретических знаний
ОПК-4	Способен находить, анализировать, реализовывать программно и использовать на практике математические алгоритмы, в том числе с применением современных вычислительных систем	ОПК-4.1. Знает базовые основы современного математического аппарата, связанного с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности
		ОПК-4.2. Умеет использовать математический аппарат в профессиональной деятельности
		ОПК-4.3. Имеет практический опыт применения современного математического аппарата, связанного с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности
ПК-1	Способен разрабатывать и отлаживать программный код	ПК-1.1. Знает основы программирования; современные объектно-ориентированные языки программирования; современные структурные языки программирования; языки современных бизнес-приложений
		ПК-1.2. Умеет кодировать на языках программирования; тестировать результаты кодирования
		ПК-1.3. Владеет навыками разработки кода информационной системы; навыками верификации кода информационной системы

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Основы анализа больших данных» относится к обязательной части блока Б1 ОП ВО.

В рамках ОП ВО обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Основы анализа больших данных».

Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики	Последующие дисциплины/модули, практики <sup>1</sup>
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	-	Научно-исследовательская работа Преддипломная практика
УК-12	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	Обработка данных и визуализация	Компьютерный практикум по статистическому анализу данных Научно-исследовательская работа Преддипломная практика
ОПК-1	Способен консультировать и использовать фундаментальные знания в области математического анализа, комплексного и функционального анализа алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений, дискретной математики и математической	Алгебра Дискретная математика и математическая логика Математический анализ	Компьютерный практикум по статистическому анализу данных Научно-исследовательская работа Преддипломная практика

1 - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики	Последующие дисциплины/модули, практики <sup>1</sup>
	логики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики в профессиональной деятельности		
ОПК-4	Способен находить, анализировать, реализовывать программно и использовать на практике математические алгоритмы, в том числе с применением современных вычислительных систем	Основы программирования	Компьютерный практикум по статистическому анализу данных Научно-исследовательская работа Преддипломная практика
ПК-1	Способен разрабатывать и отлаживать программный код	Обработка данных и визуализация	Компьютерный практикум по статистическому анализу данных Научно-исследовательская работа Преддипломная практика

#### 4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Основы анализа больших данных» составляет 3 зачетных единицы.

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО

Вид учебной работы		ВСЕГО,	Семестр
		ак.ч.	3
<i>Контактная работа, ак.ч.</i>		54	54
в том числе:			
Лекции (ЛК)		18	18
Лабораторные работы (ЛР)		-	-
Практические/семинарские занятия (СЗ)		36	36
<i>Самостоятельная работа обучающихся, ак.ч.</i>		54	54
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>		-	-
<b>Общая трудоемкость дисциплины</b>	ак.ч.	<b>108</b>	<b>108</b>
	зач.ед.	<b>3</b>	<b>3</b>

## 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы

Наименование раздела дисциплины	Содержание раздела (темы)	Вид учебной работы <sup>2</sup>
Раздел 1. Большие данные – основные понятия и задачи	Тема 1.1. Определение больших данных, области в которых используется анализ больших данных	ЛК
	Тема 1.2. Методы и инструменты анализа больших данных	ЛК
Раздел 2. Введение в пакет R	Тема 2.1. Начальные этапы работы в R.	ЛК, СЗ
	Тема 2.2. Операции над различными переменными, простейшая математика в R	ЛК, СЗ
	Тема 2.3. Операторы цикла и условия, создание функций в R	ЛК, СЗ
	Тема 2.4. Классы данных в R	ЛК, СЗ
	Тема 2.5. Ввод и вывод данных в R	ЛК, СЗ
	Тема 2.6. Графика в R	ЛК, СЗ
Раздел 3. Пакеты и функции в R, используемые в анализе больших данных.	Тема 3.1. Пакет bigmemory – создание, хранение, доступ и обработка сверхбольших матриц. Создание объекта класса big.matrix. Операции над объектом	ЛК, СЗ
	Тема 3.2. Пакет BGData – анализ данных, связанных с геномными исследованиями	ЛК, СЗ
	Тема 3.3. Пакеты bigSurvSGD и bigstatsr	ЛК, СЗ

## 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	Компьютер/ноутбук с доступом сети Интернет и электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams. Пакет R
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации,	Компьютер/ноутбук с доступом сети Интернет и электронно-образовательной среде

2 - заполняется только по ОЧНОЙ форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – семинарские занятия



Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
	оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.	Университета, браузер, ПО для просмотра PDF, MS Teams. Пакет R
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве численности обучающихся шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	ОС Windows / ОС Linux, пакет R
Для самостоятельной работы обучающихся	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	ОС Windows / ОС Linux, пакет R

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### *Основная литература:*

1. Роберт Кабаков. R в действии = R in Action. — ДМК-Пресс, 2014. — 588 с. — ISBN 978-5-947060-077-1
2. Хэдли Уикем, Гарретт Гроулмунд. Язык R в задачах науки о данных: импорт, подготовка, обработка, визуализация и моделирование данных = R for Data Science: Visualize, Model, Transform, Tidy, and Import Data. — Вильямс, 2017. — 592 с. — ISBN 978-5-9909446-8-8, 978-1-491-91039-9.
3. Норман Мэтлофф[en]. Искусство программирования на R. Погружение в большие данные. = The Art of R Programming: A Tour of Statistical Software Design.. — Питер, 2019. — 416 с. — ISBN 978-5-4461-1101-5.

### *Дополнительная литература:*

1. Шипунов А. Б. R — объектно-ориентированная статистическая среда. <http://ashipunov.info/shipunov/software/r/r-ru.htm>

### *Ресурсы информационно-телекоммуникационной сети «Интернет»:*

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров:

- Электронно-библиотечная система РУДН – ЭБС РУДН  
<http://lib.rudn.ru/MegaPro/Web>

- ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>
- ЭБС Юрайт <http://www.biblio-online.ru>
- ЭБС «Консультант студента» [www.studentlibrary.ru](http://www.studentlibrary.ru)
- ЭБС «Лань» <http://e.lanbook.com/>
- ЭБС «Троицкий мост»

## 2. Базы данных и поисковые системы:

- Comprehensive R Archive Network <https://cran.r-project.org/>
- поисковая система Яндекс <https://www.yandex.ru/>
- поисковая система Google <https://www.google.ru/>

*Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля<sup>3</sup>:*

1. Курс лекций по дисциплине «Основы анализа больших данных».
2. Задания для самостоятельной подготовки к проверочным работам.

## **8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ**

Оценочные материалы и балльно-рейтинговая система<sup>4</sup> оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Основы анализа больших данных» представлены в Приложении к настоящей Рабочей программе дисциплины.

### **РАЗРАБОТЧИКИ:**

Доцент кафедры прикладной информатики и теории вероятностей



И.С. Зарядов

Должность, БУП

Подпись

Фамилия И.О.

### **РУКОВОДИТЕЛЬ БУП:**

Зав. кафедрой прикладной информатики и теории вероятностей



К.Е. Самуйлов

Наименование БУП

Подпись

Фамилия И.О.

### **РУКОВОДИТЕЛЬ ОП ВО:**

Зав. кафедрой прикладной информатики и теории вероятностей



К.Е. Самуйлов

Наименование БУП

Подпись

Фамилия И.О.

3 - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины в ТУИС.

4 - ОМ и БРС формируются на основании требований соответствующего локального нормативного акта РУДН.