

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Ястребов Олег Александрович  
Должность: Ректор  
Дата подписания: 07.06.2022 16:01:08  
Уникальный программный ключ:  
ca953a0120d891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение  
высшего образования «Российский университет дружбы народов»**

**Экономический факультет**

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

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Big Data Advanced**

(наименование дисциплины/модуля)

**Рекомендована МССН для направления подготовки/специальности:**

**38.00.00 Экономика и управление**

(код и наименование направления подготовки/специальности)

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

**Проектный анализ и моделирование в экономике**

(наименование (профиль/специализация) ОП ВО)

**2022 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Big Data Advanced» является обеспечение студентов необходимыми знаниями и навыками по работе с большими данными на основе нереляционных баз данных, в том числе обработке данных, их интерпретации, визуализации, построении моделей.

Основными задачами курса являются:

- изучение основных понятий, связанных с большими данными, их хранении и обработки;
- изучение принципов облачных и распределённых вычислений;
- овладение современными языками обработки больших данных;
- овладение основными библиотеками Python.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Big Data Advanced» направлено на формирование у обучающихся следующих компетенций (части компетенций):

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

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
ПК-5.1	Способен собирать, анализировать и использовать данные для решения аналитических и профессиональных задач, с применением современного программного обеспечения и информационных технологий	ПК-5.1.1. Анализирует решения с точки зрения достижения целевых показателей
		ПК-5.1.2. Применяет информационные технологии в объеме, необходимом для целей бизнес-анализа
		ПК-5.1.3. Анализирует принципиальные технические решения и технологии, предлагаемые для реализации инвестиционного проекта
ПК-5.2	Способен применять программные средства и статистические методы для анализа, обработки и представления данных в цифровой форме	ПК-5.2.1. Применяет информационные технологии в объеме, необходимом для профессиональных целей
		ПК-5.2.2. Формализует и проводит алгоритмизацию поставленных задач
		ПК-5.2.3. Представляет информацию с использованием программных средств и статистических методов в цифровой форме
ПК-5.6	Способен на основе описания экономических процессов и явлений строить экономико-математические, в том числе, эконометрические модели, анализировать и интерпретировать полученные результаты	ПК-5.6.1. Анализирует качество информации с точки зрения выбранных критериев
		ПК-5.6.2. Использует эконометрические методы прогнозирования на краткосрочную, среднесрочную и долгосрочную перспективу
		ПК-5.6.3. Строит стандартные логические и эконометрические модели, анализирует и интерпретирует полученные результаты

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Big Data Advanced» относится к части, формируемой участниками образовательных отношений блока Б1 ОП ВО.

В рамках ОП ВО обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Big Data Advanced».

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

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
ПК-5.1	Способен собирать, анализировать и использовать данные для решения аналитических и профессиональных задач, с применением современного программного обеспечения и информационных технологий	Введение в цифровую бизнес-аналитику, Бизнес-аналитика в SAP	Нет последующих дисциплин
ПК-5.2	Способен применять программные средства и статистические методы для анализа, обработки и представления данных в цифровой форме	Введение в цифровую бизнес-аналитику, Бизнес-аналитика в SAP	Нет последующих дисциплин
ПК-5.6	Способен на основе описания экономических процессов и явлений строить экономико-математические, в том числе, эконометрические модели, анализировать и интерпретировать полученные результаты	Сторителлинг	Нет последующих дисциплин

\* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

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

Общая трудоемкость дисциплины «Big Data Advanced» составляет 3 зачетные единицы.

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО для **ОЧНОЙ** формы обучения

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)	
		6	
Контактная работа, ак.ч.	108	108	
Лекции (ЛК)			
Лабораторные работы (ЛР)	34	34	
Практические/семинарские занятия (СЗ)			
Самостоятельная работа обучающихся, ак.ч.	56	56	
Контроль (экзамен/зачет с оценкой), ак.ч.	18	18	
Общая трудоемкость дисциплины	ак.ч.	108	108
	зач.ед.	3	3

#### 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

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

Наименование раздела дисциплины	Содержание раздела (темы)	Вид учебной работы*
Раздел 1. Инструменты больших данных.	Тема 1.1. Основы архитектуры хранения и обработки данных. Виды обработки данных. Инструменты работы с данными. Состояние данных. Облачные и распределенные вычисления. Культура сбора и источники данных.	ЛР
	Тема 1.2. Особенности Python. Арифметические операции. Переменные и именование переменных. Операторы сравнения, условия и условные конструкции. Логические операторы и логические значения.	ЛР
Раздел 2. Введение в типы данных в Python. Функции. Классы и их применение.	Тема 2.1. Простые типы данных. Списки. Кортежи. Множества. Циклы while и for. Понятие функций. Параметры функции. Области видимости.	ЛР
Раздел 3. Библиотеки numpy и scipy.	Тема 3.1. Принципы работы библиотек. Матрицы и вектора.	ЛР
Раздел 4. Основы pandas.	Тема 4.1. Основные операции в pandas. Функции и группировки. Сводные таблицы.	ЛР
Раздел 5. Визуализация данных.	Тема 5.1. Основные инструменты визуализации.	ЛР
	Тема 5.2. Визуализация большого массива данных. Принципы визуализации в Matplotlib и Seaborn.	ЛР

Наименование раздела дисциплины	Содержание раздела (темы)	Вид учебной работы*
Раздел 6. Основы описательной статистики при работе с большими данными.	Тема 6.1. Основные понятия в описательной статистике. Способы получения описательной статистики.	ЛР
	Тема 6.2. Обработка данных.	ЛР
Раздел 7. Центральная предельная теорема и статистический анализ в Python.	Тема 7.1. Центральная предельная теорема. Нормализация данных. Матрица корреляций. Распределение Стьюдента в больших данных.	ЛР
	Тема 7.2. Обработка данных.	ЛР
Раздел 8. Основные статистические тесты и проверка гипотез.	Тема 8.1. Проверка гипотез и построение моделей.	ЛР
Раздел 9. Основы ruspark.	Тема 9.1. Проверка гипотез в ruspark и визуализация данных.	ЛР

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

## 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.	
Лаборатория	Аудитория для проведения лабораторных работ, индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и оборудованием.	Перечень специализированного лабораторного оборудования: Windows, Microsoft Office, Anaconda Navigator, GoogleColab, доступ в Интернет
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.	

\* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

1. Виктор Майер-Шенбергер, Кеннет Кукьер Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / пер. с англ. И. Гайдюк. – Москва : Манн, Иванов и Фейбер, 2019, 231 с., ISBN 978-5-91657-936-9.

2. Shashank Tiwari Professional NoSQL. – Ind Wiley, 2021, 384 p., ISBN 978-0-470-94224-6.

3. Деви Силен, Арно Мейсман Основы Data Science и Big Data. Python и наука о данных. – СПб. : Питер, 2020. 336 с., ISBN 978-5-496-02517-1.

4. Фрэнкс, Билл. Укрощение больших данных: как извлекать знания из массивов информации с помощью глубокой аналитики / Билл Фрэнкс ; пер. с англ. Андрея Баранова. – Москва : Манн, Иванов и Фербер, 2020. – 349 с. : ил.; ISBN 978-5-00057-146-0.

5. Марц, Натан. Большие данные: принципы и практика построения масштабируемых систем обработки данных в реальном времени / Натан Марц, Джеймс Уоррен ; [пер. с англ. и ред. И. В. Берштейна]. – Москва : Вильямс, 2020. – 368 с. : ил., табл.; ISBN 978-5-8459-2075-1 : 300 экз.

6. Лесковец, Юре. Анализ больших наборов данных / Юре Лесковец, Ананд Раджараман, Джеффри Д. Ульман ; пер. с англ. Слинкин А. А. – Москва : ДМК Пресс, 2019. – 498 с. : ил.; ISBN 978-5-97060-190-7.

7. Андрианов И. А. Индексирование и поиск в последовательностях для больших баз данных: монография / И. А. Андрианов, А. Ф. Чернов ; М-во образования и науки Российской Федерации, Вологодский гос. ун-т. – Вологда : ВоГУ, 2021. – 167 с. : ил., табл.; 21 см.; ISBN 978-5-87851-503-0.

8. Big data and business analytics / ed. by Jay Liebowitz; forew. by Joe LaCugna. – Boca Raton [etc.] : CRC press, cop. 2020. – xx, 282 с. : ил.; 25 см.; ISBN 9781466565784.

9. Big data computing / ed. by Rajendra Akerkar. – Boca Raton [etc.]: CRC press, cop. 2014. – XXII, 542 с. : ил.; 24 см.; ISBN 9781466578371.

10. Data-driven innovation : big data for growth and well-being. – Paris: OECD, cop. 2020. - 452 с.: ил.; 28 см.; ISBN 978-92-64-22934-1

11. Доусон, Майкл. Програмируем на Python / Майкл Доусон; [пер. с англ.: В. Порицкий]. – Москва [и др.] : Питер, 2020. – 414 с. : ил., табл.; 24 см.; ISBN 978-5-496-01071-9

12. Маккинли, Уэс. Python и анализ данных / Уэс Маккинли; пер. с англ. Слинкин А. А. – Москва : ДМК Пресс, 2012. – 482 с., : ил.; ISBN 978-5-97060-315-4.

#### *Дополнительная литература:*

1. Swaroop С.Н. A byte of python. Учебное пособие по программирования на языке Python Открытый доступ по ссылке <https://wombat.org.ua/AByteOfPython/AByteofPythonRussian-2.01.pdf>

2. Python Data Analytics. Fabro Nelli, Rome Italy (2018) ISBN 978-1-4842-3912-4

3. PYTHON PROGRAMMING LANGUAGE FOR BEGINNERS. A Crash Course Guide with Tricks and Tools for Your First Approach to Learn and Programming with Python 2019 Clark Wes

#### *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:

- электронный фонд правовой и нормативно-технической документации  
<http://docs.cntd.ru/>

- поисковая система Яндекс <https://www.yandex.ru/>

- поисковая система Google <https://www.google.ru/>

- реферативная база данных SCOPUS  
<http://www.elsevierscience.ru/products/scopus/>

- <https://anaconda.org/anaconda/anaconda-navigator> - документация по работе со специализированным ПО

- <https://jupyter.org> – основная документация по среде для работы с большими данными

- <https://numpy.org> – основная документация по библиотеке Numpy

- <https://www.scipy.org> – основная документация по библиотеке Scipy

- <https://pandas.pydata.org> – основная документация по библиотеке Pandas

- <https://matplotlib.org> – основная документация по библиотеке Matplotlib

- <https://seaborn.pydata.org> – основная документация по библиотеке Seaborn

- <https://scikit-learn.org/stable/index.html> – основная документация по библиотеке scikit learn

- <https://spark.apache.org/docs/latest/> – основная документация по библиотеке pyspark

*Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля\*:*

1. Курс лекций по дисциплине «Big Data Advanced».

2. Лабораторный практикум по дисциплине «Big Data Advanced»

\* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**

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

Оценочные материалы и балльно-рейтинговая система\* оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Big Data Advanced» размещены в ТУИС  
<https://esystem.rudn.ru/course/index.php?categoryid=823>.

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

**РАЗРАБОТЧИКИ:**

**Доцент кафедры экономико-  
математического  
моделирования**



**Гомонов К.Г**

---

Должность, БУП

---

Подпись

---

Фамилия И.О.

**РУКОВОДИТЕЛЬ БУП:**

**Кафедра экономико-  
математического  
моделирования**



**Балашова С.А.**

---

Наименование БУП

---

Подпись

---

Фамилия И.О.

**РУКОВОДИТЕЛЬ ОП ВО:**

**Заведующая кафедрой  
экономико-математического  
моделирования**



**Балашова С.А.**

---

Должность, БУП

---

Подпись

---

Фамилия И.О.