

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
"Казанский (Приволжский) федеральный университет"  
Институт филологии и межкультурной коммуникации  
Высшая школа национальной культуры и образования им. Габдуллы Тукая



УТВЕРЖДАЮ  
Проректор по образовательной деятельности КФУ  
Проф. Д. А. Таюрский  
ДЕПАРТАМЕНТ  
ОБРАЗОВАНИЯ  
(ДО КФУ)



« 01 » ИЮНЯ 2021 г.

*подписано электронно-цифровой подписью*

## Программа дисциплины

Методы обработки Big Data

Направление подготовки: 44.03.05 - Педагогическое образование (с двумя профилями подготовки)

Профиль подготовки: Информатика и иностранный (английский) язык (в полилингвальной образовательной среде)

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2021

## Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО
2. Место дисциплины (модуля) в структуре ОПОП ВО
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся
4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий
  - 4.1. Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)
  - 4.2. Содержание дисциплины (модуля)
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
6. Фонд оценочных средств по дисциплине (модулю)
7. Перечень литературы, необходимой для освоения дисциплины (модуля)
8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)
12. Средства адаптации преподавания дисциплины (модуля) к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья
13. Приложение №1. Фонд оценочных средств
14. Приложение №2. Перечень литературы, необходимой для освоения дисциплины (модуля)
15. Приложение №3. Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Программу дисциплины разработал(а)(и): доцент, к.н. (доцент) Галимянов А.Ф. (кафедра билингвального и цифрового образования, Высшая школа национальной культуры и образования им. Габдуллы Тукая), Anis.Galimjanoff@kpfu.ru

**1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО**

Обучающийся, освоивший дисциплину (модуль), должен обладать следующими компетенциями:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОПК-9	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности
ПК-4	Способен использовать современные достижения отечественного и зарубежного методического наследия, современные методы и технологии обучения и диагностики в области информатики и английского языка

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

теоретические основы работы с большими данными

Должен уметь:

использовать инструментов и приложений данной дисциплины.

Должен владеть:

пониманием концепции теоретических основ дисциплины, основных инструментов и приложений данной дисциплины

**2. Место дисциплины (модуля) в структуре ОПОП ВО**

Данная дисциплина (модуль) включена в раздел "Б1.О.06.01.ДВ.18.02 Дисциплины (модули)" основной профессиональной образовательной программы 44.03.05 "Педагогическое образование (с двумя профилями подготовки) (Информатика и иностранный (английский) язык (в полилингвальной образовательной среде))" и относится к обязательной части ОПОП ВО.

Осваивается на 4 курсе в 8 семестре.

**3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся**

Общая трудоемкость дисциплины составляет 2 зачетных(ые) единиц(ы) на 72 часа(ов).

Контактная работа - 37 часа(ов), в том числе лекции - 18 часа(ов), практические занятия - 18 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 1 часа(ов).

Самостоятельная работа - 35 часа(ов).

Контроль (зачёт / экзамен) - 0 часа(ов).

Форма промежуточного контроля дисциплины: зачет в 8 семестре.

**4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий**

**4.1 Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)**

N	Разделы дисциплины / модуля	Се-местр	Виды и часы контактной работы, их трудоемкость (в часах)						Само-стоя-тельная ра-бота
			Лекции, всего	Лекции в эл. форме	Практи-ческие занятия, всего	Практи-ческие в эл. форме	Лабора-торные работы, всего	Лабора-торные в эл. форме	
1.	Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации	8	6	6	6	0	0	0	12

N	Разделы дисциплины / модуля	Се- местр	Виды и часы контактной работы, их трудоемкость (в часах)						Само- стоя- тель- ная ра- бота
			Лекции, всего	Лекции в эл. форме	Практи- ческие занятия, всего	Практи- ческие в эл. форме	Лабора- торные работы, всего	Лабора- торные в эл. форме	
2.	Тема 2. Технологии хранения и обработки Больших данных	8	6	6	6	0	0	0	12
3.	Тема 3. Аналитика в больших данных.	8	6	6	6	0	0	0	11
	Итого		18	18	18	0	0	0	35

#### 4.2 Содержание дисциплины (модуля)

##### Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации

Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации

Факторный анализ. Дискриминантный анализ. Кластерный анализ. Многомерное шкалирование. Методы контроля качества.

Тема 2. Технологии хранения и обработки Больших данных

Основные направления развития методов обработки и хранения данных. Volume. Закон Мура. Velocity. Variety. Фреймворк Hadoop. Проблема хранения неструктуриро-ванных данных. Проблема преобразования данных. Семантические анализаторы. Само-обучающиеся автоматы.

##### Тема 2. Технологии хранения и обработки Больших данных

Тема 3. Программирование обработки и загрузки Больших данных

9 языков для Big Data (R, Python, Julia, Java, Scala, MATLAB, Go, Kafka, Hadoop). Фреймворки (Hadoop, Spark, Storm). Базы данных (Hive, Impala, Presto, Drill). Аналитиче-ские платформы (Rapid Miner, IBM SPSS Modeler, KNIME, Qlik Analytics Platform, STA-TISTICA Data Miner, Informatica Intelligent Data Platform, World Programming System, De-ductor, SAS Enterprise Miner). Прочие инструменты (Zookeeper, Flume, IBM Watson Analyt-ics, Dell EMC Analytic Insights Module, Windows Azure HDInsight, Microsoft Azure Machine Learning, Pentaho Data Integration, Teradata Aster Analytics, SAP BusinessObjects Predictive Analytics, Oracle Big Data Preparation).

##### Тема 3. Аналитика в больших данных.

Тема 4. Аналитика в больших данных

Аналитика Big Data - реалии и перспективы в России и мире. Технологии и ме-тоды анализа, которые используются для анализа Big Data( Data Mining; краудсорсинг; смешение и интеграция данных; машинное обучение; искусственные нейронные сети; распознавание образов; прогнозная аналитика; имитационное моделирование; простран-ственный анализ; статистический анализ; визуализация аналитических данных). Big data: применение и возможности. Решения на основе Big data. Рынок Big data в России. Big data в банках. Big data в бизнесе. Big data в маркетинге.

#### 5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Самостоятельная работа обучающихся выполняется по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Самостоятельная работа подразделяется на самостоятельную работу на аудиторных занятиях и на внеаудиторную самостоятельную работу. Самостоятельная работа обучающихся включает как полностью самостоятельное освоение отдельных тем (разделов) дисциплины, так и проработку тем (разделов), осваиваемых во время аудиторной работы. Во время самостоятельной работы обучающиеся читают и конспектируют учебную, научную и справочную литературу, выполняют задания, направленные на закрепление знаний и отработку умений и навыков, готовятся к текущему и промежуточному контролю по дисциплине.

Организация самостоятельной работы обучающихся регламентируется нормативными документами, учебно-методической литературой и электронными образовательными ресурсами, включая:

Порядок организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры (утвержден приказом Министерства науки и высшего образования Российской Федерации от 6 апреля 2021 года №245)

Письмо Министерства образования Российской Федерации №14-55-99бин/15 от 27 ноября 2002 г. "Об активизации самостоятельной работы студентов высших учебных заведений"

Устав федерального государственного автономного образовательного учреждения "Казанский (Приволжский) федеральный университет"

Правила внутреннего распорядка федерального государственного автономного образовательного учреждения высшего профессионального образования "Казанский (Приволжский) федеральный университет"

Локальные нормативные акты Казанского (Приволжского) федерального университета

#### 6. Фонд оценочных средств по дисциплине (модулю)

Фонд оценочных средств по дисциплине (модулю) включает оценочные материалы, направленные на проверку освоения компетенций, в том числе знаний, умений и навыков. Фонд оценочных средств включает оценочные средства текущего контроля и оценочные средства промежуточной аттестации.

В фонде оценочных средств содержится следующая информация:

- соответствие компетенций планируемым результатам обучения по дисциплине (модулю);
- критерии оценивания сформированности компетенций;
- механизм формирования оценки по дисциплине (модулю);
- описание порядка применения и процедуры оценивания для каждого оценочного средства;
- критерии оценивания для каждого оценочного средства;
- содержание оценочных средств, включая требования, предъявляемые к действиям обучающихся, демонстрируемым результатам, задания различных типов.

Фонд оценочных средств по дисциплине находится в Приложении 1 к программе дисциплины (модулю).

#### 7. Перечень литературы, необходимой для освоения дисциплины (модуля)

Освоение дисциплины (модуля) предполагает изучение основной и дополнительной учебной литературы. Литература может быть доступна обучающимся в одном из двух вариантов (либо в обоих из них):

- в электронном виде - через электронные библиотечные системы на основании заключенных КФУ договоров с правообладателями;

- в печатном виде - в Научной библиотеке им. Н.И. Лобачевского. Обучающиеся получают учебную литературу на абонементе по читательским билетам в соответствии с правилами пользования Научной библиотекой.

Электронные издания доступны дистанционно из любой точки при введении обучающимся своего логина и пароля от личного кабинета в системе "Электронный университет". При использовании печатных изданий библиотечный фонд должен быть укомплектован ими из расчета не менее 0,5 экземпляра (для обучающихся по ФГОС 3++ - не менее 0,25 экземпляра) каждого из изданий основной литературы и не менее 0,25 экземпляра дополнительной литературы на каждого обучающегося из числа лиц, одновременно осваивающих данную дисциплину.

Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля), находится в Приложении 2 к рабочей программе дисциплины. Он подлежит обновлению при изменении условий договоров КФУ с правообладателями электронных изданий и при изменении комплектования фондов Научной библиотеки КФУ.

#### 8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

BIG DATA AND ADVANCED ANALYTICS - <http://bigdataminsk.by>

Иллюстрированный самоучитель по SPSS - <http://www.spssbase.com/>

Официальный сайт российского офиса компании SPSS - <http://www.spss.ru>

#### 9. Методические указания для обучающихся по освоению дисциплины (модуля)

Вид работ	Методические рекомендации
лекции	Работа по лекциям включает в себя работу до лекции, работу во время лекции и работу после лекции. Студенты знакомы с учебным планом и преподаватель заранее сообщает тему следующей лекции. Студент должен ознакомиться с темой по материалам в сети Интернет, в виртуальной аудитории. Вопросы во время лекции поощряются по оценке преподавателя. После лекции материал прорабатывается и используется в лабораторных работах

Вид работ	Методические рекомендации
практические занятия	Практические работы выполняются по темам, определенным учебным планом. Легенды для конкретной работы предлагаются преподавателем. Каждая работа завершается отчетом. В отчете должны быть четко определены постановка задачи, используемый инструментарий, пути решения задачи, подробный ход решения задачи, выводы. Приветствуется обсуждение и возможные альтернативные варианты решения. Инструментарий зависит от имеющегося программного обеспечения.
самостоятельная работа	Самостоятельная работа включает в себя работу с лекционным материалом, подготовку к лабораторным работам и выполнение лабораторных работ вне аудитории, если это предлагается преподавателем, подготовку отчета, а также изучение нового материала по сети. Изучение нового материала по теме должно обязательно сопровождаться ознакомлением с новейшими достижениями, так как данная сфера относится к быстро развивающимся областям. Поэтому приветствуется включение в отчеты по лабораторным работам а также вопросы во время лекций по новейшим достижениям по изучаемой теме, это может поощряться преподавателем дополнительными баллами.
зачет	Зачет проводится в форме тестирования. Время тестирования варьируется так, чтобы на ответ на один вопрос отводился от одного до трех минут. Обычно тест открывается на сутки, количество попыток регламентируется преподавателем. Окончательная оценка ставится как арифметическое среднее оценки всех попыток, но может изменяться преподавателем.

### 10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем, представлен в Приложении 3 к рабочей программе дисциплины (модуля).

### 11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Материально-техническое обеспечение образовательного процесса по дисциплине (модулю) включает в себя следующие компоненты:

Помещения для самостоятельной работы обучающихся, укомплектованные специализированной мебелью (столы и стулья) и оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду КФУ.

Учебные аудитории для контактной работы с преподавателем, укомплектованные специализированной мебелью (столы и стулья).

Компьютер и принтер для распечатки раздаточных материалов.

Компьютерный класс.

### 12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников - например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения - аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;

- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи;
- продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут;
- продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;
- продолжительности выступления обучающегося при защите курсовой работы - не более чем на 15 минут.

Программа составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению 44.03.05 "Педагогическое образование (с двумя профилями подготовки)" и профилю подготовки "Информатика и иностранный (английский) язык (в полилингвальной образовательной среде)".

*Приложение 2*  
*к рабочей программе дисциплины (модуля)*  
*Б1.О.06.01, ДВ.18.02 Методы обработки Big Data*

**Перечень литературы, необходимой для освоения дисциплины (модуля)**

Направление подготовки: 44.03.05 - Педагогическое образование (с двумя профилями подготовки)

Профиль подготовки: Информатика и иностранный (английский) язык (в полилингвальной образовательной среде)

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2021

**Основная литература:**

1. Много цифр. Анализ больших данных при помощи Excel / Форман Д.; Пер. с англ. Соколовой А. - М.: Альпина Паблиш, 2016. - 461 с.: 84x108 1/16 (Об-ложка) ISBN 978-5-9614-5032-3 <http://znanium.com/catalog/product/551044>
2. Анализ больших наборов данных / Юре Лесковец, Ананд Раджараман, Джеффри Д. Ульман; пер. с англ. А.А.Слинкина. - Москва : ДМК Пресс, 2016. - 498 с. - ISBN 978-5-97060-190-7 <http://znanium.com/catalog/product/1027845>
3. Данные: хранение и обработка : учебник / Э.Г. Дадян. - М. : ИНФРА-М, 2019. - 205 с. - (Высшее образование: Бакалавриат). - [www.dx.doi.org/10.12737/textbook\\_5cf8c7f2b8cdb8.06963680](http://www.dx.doi.org/10.12737/textbook_5cf8c7f2b8cdb8.06963680). <http://znanium.com/catalog/product/989190>

**Дополнительная литература:**

- 1 Базы данных: Учебник / Шустова Л.И., Тараканов О.В. - М.: НИЦ ИНФРА-М, 2016. - 304 с.: 60x90 1/16. - (Высшее образование: Бакалавриат) (Пере-плёт) ISBN 978-5-16-010485-0 <http://znanium.com/catalog/product/491069>
2. Социально-экономическое районирование в эпоху больших данных: монография-М.: НИЦ ИНФРА-М, 2018.-196 с. <http://znanium.com/catalog/product/923729>
3. Назарова, О.Б. Разработка реляционных баз данных с использованием CASE-средства ALL Fusion Data Modeler : учеб.- метод, пособие / О.Б. Назарова, О.Е. Масленникова. - 3-е изд., стер. - Москва : ФЛИНТА, 2019. - 73 с. - ISBN 978-5-9765-1601-4 <http://znanium.com/catalog/product/1034975>



*Приложение 3  
к рабочей программе дисциплины (модуля)  
Б1.О.06.01.ДВ.18.02 Методы обработки Big Data*

**Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем**

Направление подготовки: 44.03.05 - Педагогическое образование (с двумя профилями подготовки)

Профиль подготовки: Информатика и иностранный (английский) язык (в полилингвальной образовательной среде)

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2021

Освоение дисциплины (модуля) предполагает использование следующего программного обеспечения и информационно-справочных систем:

Операционная система Microsoft Windows 7 Профессиональная или Windows XP (Volume License)

Пакет офисного программного обеспечения Microsoft Office 365 или Microsoft Office Professional plus 2010

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI или Adobe Acrobat Reader DC

Kaspersky Endpoint Security для Windows

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, учебно-методические комплексы, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего образования (ФГОС ВО) нового поколения.