

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
"Казанский (Приволжский) федеральный университет"  
Институт вычислительной математики и информационных технологий



УТВЕРЖДАЮ  
Проректор по образовательной деятельности КФУ  
Проф. Д. А. Таюрский



\_\_\_\_\_» \_\_\_\_\_ 20\_\_ г.

*подписано электронно-цифровой подписью*

## Программа дисциплины

Язык Python и анализ данных

Направление подготовки: 01.03.02 - Прикладная математика и информатика

Профиль подготовки: Системное программирование

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

## 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) Абдуллин А.И.

### 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  |
|------------------|--|
| ПК-1             | Способность собирать, обрабатывать и интерпретировать данные современных научных исследований, необходимые для формирования выводов по соответствующим научным исследованиям |
| ПК-2             | Способность понимать, совершенствовать и применять современный математический аппарат  |
| ПК-3             | Способность критически переосмысливать накопленный опыт, изменять при необходимости вид и характер своей профессиональной деятельности                                       |
| ПК-4             | Способность работать в составе научно-исследовательского и производственного коллектива и решать задачи профессиональной деятельности  |
| ПК-5             | Способность осуществлять целенаправленный поиск информации о новейших научных и технологических достижениях в сети Интернет и из других источников                           |

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

- синтаксис и управляющие конструкции языка Python;
- основные стандартные модули и библиотеки в Python.
- особенности разработки прикладных программ на языке Python;
- принципы разработки собственных модулей и библиотек.
- специализированные библиотеки для анализа данных

Должен уметь:

- разрабатывать прикладные программы на языке программирования Python;
- использовать как стандартные, так и дополнительные модули, расширения и пакеты;
- создавать собственные модули и импортировать их
- создавать собственные приложения и библиотеки.

Должен владеть:

- навыками разработки приложений с внешними источниками данных (текстовыми файлами, xml-файлами, csv-файлами);
- практическими навыками использования набора библиотек языка Python для прикладных задач в области анализа данных;
- навыками разработки приложений и собственных библиотек на языке Python.

Должен демонстрировать способность и готовность:

- применять полученные знания в своей дальнейшей профессиональной деятельности в области анализа данных с использованием программирования на Python и специализированных библиотек.

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

Данная дисциплина (модуль) включена в раздел "Б1.В.ДВ.13 Дисциплины (модули)" основной профессиональной образовательной программы 01.03.02 "Прикладная математика и информатика (Системное программирование)" и относится к дисциплинам по выбору.

Осваивается на 3 курсе в 5, 6 семестрах.

### 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 4 зачетных(ые) единиц(ы) на 144 часа(ов).

Контактная работа - 72 часа(ов), в том числе лекции - 0 часа(ов), практические занятия - 0 часа(ов), лабораторные работы - 72 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 72 часа(ов).

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

Форма промежуточного контроля дисциплины: отсутствует в 5 семестре; зачет в 6 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  | Семестр | Виды и часы контактной работы, их трудоемкость (в часах) |                      |                     | Самостоятельная работа |
|----|--|---------|--|----------------------|---------------------|------------------------|
|    |  |         | Лекции   | Практические занятия | Лабораторные работы |                        |
| 1. | Тема 1. Введение в программирование на языке Python.   | 5       | 0  | 0                    | 18                  | 9                      |
| 2. | Тема 2. Обработка данных. Массивы и векторные вычисления. Построение графиков и визуализация данных.                 | 5       | 0  | 0                    | 18                  | 9                      |
| 3. | Тема 3. Специализированные библиотеки Python для анализа данных.   | 6       | 0  | 0                    | 18                  | 36                     |
| 4. | Тема 4. Сбор данных Python. Извлечение данных с веб-страниц (web-scraping). Использование открытых данных через API. | 6       | 0  | 0                    | 18                  | 18                     |
|    | Итого  |         | 0  | 0                    | 72                  | 72                     |

#### 4.2 Содержание дисциплины (модуля)

##### Тема 1. Введение в программирование на языке Python.

Введение в Python. Синтаксис языка. Базовые типы данных: числа, строки, списки, кортежи, словари, множества. Функции. Итераторы и генераторы. Классы и объекты. Декораторы. Ввод-вывод. Обработка исключений. IPython, Jupyter Notebook. Подключение библиотек, создание собственных модулей. Элементы функционального программирования (lambda, map, zip, reduce, filter).

Чтение и запись данных в текстовом формате.

##### Тема 2. Обработка данных. Массивы и векторные вычисления. Построение графиков и визуализация данных.

Основы NumPy: многомерные массивы и векторные вычисления.

Индексирование и вырезание. Универсальные функции: быстрые поэлементные операции над массивами. Обработка данных с применением массивов.

Методы булевых массивов. Сортировка.

Устранение дубликатов и другие теоретико-множественные операции.

Файловый ввод-вывод массивов.

Линейная алгебра. Генерация случайных чисел

Визуализация данных в Python. Обзор библиотек: matplotlib, pandas, seaborn, plotly, bokeh. Базовые типы визуализаций: графики, столбчатые диаграммы, гистограммы, точечные диаграммы (scatter plots), ящики с усами. Комбинирование различных графических элементов. Построение интерактивных диаграмм с помощью plotly и bokeh.

##### Тема 3. Специализированные библиотеки Python для анализа данных.

Введение в анализ табличных данных в Python. Пакет pandas. Объекты Series (последовательность) и DataFrame (таблица). Чтение-запись данных в различных форматах. Запросы к таблицам: выборка строк/столбцов по заданным критериям. Переформатирование данных: очистка,

преобразование, слияние, изменение формы.

Фильтрация отсутствующих данных. Агрегирование данных и групповые операции. Основы работы с временными рядами.

#### **Тема 4. Сбор данных Python. Извлечение данных с веб-страниц (web-scraping). Использование открытых данных через API.**

Сбор и подготовка данных. Извлечение данных с веб-страниц (web-scraping). Библиотека beautifulsoup. Работа с динамическими сайтами с помощью Selenium. Массовый скреппинг с помощью scrapy. Работа со структурированными данными: JSON и XML. Разбор XML с помощью lxml.objectify. Взаимодействие с HTML и Web API.

#### **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 образования и науки Российской Федерации от 5 апреля 2017 года №301)

Письмо Министерства образования Российской Федерации №14-55-996ин/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

библиотека pandas для анализа данных - [pandas.pydata.org](http://pandas.pydata.org)

библиотека для эффективной работы с многомерными массивами данных - [numpy.org](http://numpy.org)

дистрибутив Python вместе с основными библиотеками для анализа данных и пакетным менеджером conda - [anaconda.com](http://anaconda.com)

дистрибутивы и документация для языка программирования Python - [python.org](http://python.org)

#### **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)

библиотека pandas для анализа данных - <http://pandas.pydata.org/>

библиотека для эффективной работы с многомерными массивами данных - <http://www.numpy.org/>

дистрибутив Python вместе с основными библиотеками для анализа данных и пакетным менеджером conda - <http://www.anaconda.com/>

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

| Вид работ              | Методические рекомендации  |
|------------------------|--|
| лабораторные работы    | Чтение конспекта лекций и учебника должно сопровождаться практическим решением задач на основании теоретических положений дисциплины, для чего рекомендуется завести специальную тетрадь. Если студент видит несколько путей для решения задачи, то он должен сравнить их и выбрать из них самый оптимальный (минимальное количество строк кода). Полезно до начала программирования составить блок-схему. Решения задач при выполнении лабораторных работ следует излагать подробно, обосновывать каждый этап решения, исходя из теоретических положений курса. Программный код снабжать комментариями, придерживаться рекомендации PEP8. |
| самостоятельная работа | Самостоятельная работа над учебным материалом является составной частью обучения студента. По математическим курсам она складывается из чтения конспекта лекций и учебника, решения практических задач, самопроверки и выполнения контрольных заданий. Кроме этого, студент может обращаться с вопросами к преподавателю для получения устной или письменной консультации.   |
| зачет                  | На зачетах выясняется прежде всего отчётливое знание практических аспектов программы курса. При подготовке к зачету студент должен повторно изучить конспекты лекций и рекомендованную литературу, просмотреть решения основных задач, решенных самостоятельно и на лабораторных работах, семинарах, а также составить письменные ответы на все вопросы, вынесенные на зачет.  |

## 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.

Мультимедийная аудитория.

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 01.03.02 "Прикладная математика и информатика" и профилю подготовки "Системное программирование".

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

Направление подготовки: 01.03.02 - Прикладная математика и информатика

Профиль подготовки: Системное программирование

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

#### Основная литература:

1.Маккинли У., Python и анализ данных [Электронный ресурс] / Уэс Маккинли - М. : ДМК Пресс, 2015. - 482 с. - ISBN 978-5-97060-315-4 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785970603154.html>

2.Гуриков С. Р. Основы алгоритмизации и программирования на Python : учеб. пособие / С.Р. Гуриков. - М. : ФОРУМ : ИНФРА-М, 2018. - 343 с. - (Высшее образование: Бакалавриат). - Режим доступа: <http://znanium.com/catalog/product/924699>

3.Саммерфилд М., Python на практике [Электронный ресурс] / Марк Саммерфилд - М. : ДМК Пресс, 2014. - 338 с. - ISBN 978-5-97060-095-5 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785970600955.html>

#### Дополнительная литература:

1.Волкова В.М., Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python [Электронный ресурс]: учебное пособие / Волкова В.М. - Новосибирск : Изд-во НГТУ, 2017. - 74 с. - ISBN 978-5-7782-3183-2 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785778231832.html>

2.Прохоренок Н. А. и др. Python 3. Самое необходимое: Пособие / Прохоренок Н.А., Дронов В.А. - СПб:БХВ-Петербург, 2016. - 464 с. ISBN 978-5-9775-3631-8 - Режим доступа: <http://znanium.com/catalog/product/944129>

3.Сузи Р. А. Python: Пособие / Сузи Р.А. - СПб:БХВ-Петербург, 2015. - 759 с. ISBN 978-5-9775-1417-0 - Режим доступа: <http://znanium.com/catalog/product/939857>



Приложение 3  
к рабочей программе дисциплины (модуля)  
Б1.В.ДВ.13 Язык Python и анализ данных

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

Направление подготовки: 01.03.02 - Прикладная математика и информатика

Профиль подготовки: Системное программирование

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

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.