

УЧЕБНАЯ ПРОГРАММА.

Основы программирования на языке R.

СОСТАВИТЕЛЬ.

Василий Викторович Гринев, доцент кафедры генетики Белорусского государственного университета, кандидат биологических наук.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Раздел 1. Введение в дисциплину

Тема 1.1. Язык и среда для программирования R. Проблема выбора языка программирования. История появления и развития R. Получение и установка R. Командная консоль R. Работа с R Commander, RStudio. Рабочее пространство R. Социальная инфраструктура R, сообщество R-разработчиков и R-пользователей. Сферы применения R в биологии и медицине.

Раздел 2. Базовые принципы организации языка R

Тема 2.1. Базовые принципы организации языка R. Лексика, синтаксис и семантика языка R. Встроенные операторы и функции. Константы и переменные.

Тема 2.2. Объекты R языка. Создание и работа с объектами. Базовые объекты: векторы, матрицы, таблицы, массивы, списки, факторы. Специальные классы объектов для работы с биологическими и медицинскими данными, их создание и использование. Атрибуты объектов.

Раздел 3. Работа с данными в R

Тема 3.1. Работа с данными в R. Типы данных. Импорт данных. Основы управления данными. Доступ к элементам векторов, матриц, таблиц, массивов и списков. Пропущенные значения. Аннотирование данных. Векторные и матричные операции. Представление даты и времени, временные ряды. Экспорт данных.

Тема 3.2. Организация вычислений в R. Оптимизация доступа к данным. Оптимизация структуры данных. Ветвления и циклы. R-функции. Аргументы функций. Ускорение вычислений. Векторизованные вычисления и векторизованные функции. Распараллеливание вычислений в R.

Тема 3.3. Графические возможности R. Базовые графические возможности. Расширение графических возможностей: подходы и их реализация. Специальные графические возможности для биологии и медицины.

Раздел 4. Работа с аппаратными средствами в R

Тема 4.1. Доступ к оперативной памяти компьютера и ее эффективный контроль. Работа с накопителями. Управление процессором. Удаленный доступ. Облачные вычисления. Специальные технические возможности R.

Раздел 5. Разработка приложений на R

Тема 5.1. Программный R-код. Правила написания программного кода. Отладка программного кода. Подключение внешних программных кодов. Взаимодействие с другими языками программирования.

Тема 5.2. Разработка пользовательских функций. Разработка пользовательских библиотек функций. Подключение внешних библиотек функций. Репозитории функций и библиотек функций.

Тема 5.3. Разработка готовых программных решений и приложений. Пайплайны. Работа с GitHub, CRAN и Bioconductor. Web-приложения на R.

ТЕМАТИКА ЛЕКЦИЙ И ЛАБОРАТОРНЫХ ЗАНЯТИЙ

Тематика лекций (20 часов)

- 1) Язык и среда для программирования R (2 часа).
- 2) Базовые принципы организации языка R (2 часа).
- 3) Объекты R языка (2 часа).
- 4) Работа с данными в R (2 часа).
- 5) Организация вычислений в R (2 часа).
- 6) Графические возможности R (2 часа).
- 7) Работа с аппаратными средствами в R (2 часа).
- 8) Разработка программных кодов на R языке (2 часа).
- 9) Разработка пользовательских R-функций и библиотек функций (2 часа).
- 10) Разработка приложений на R языке (2 часа).

Тематика лабораторных занятий (28 часов)

- 1) Установка и управление средой для программирования R (2 часа).
- 2) Работа с базовыми R-объектами (2 часа).
- 3) Работа со специальными R-объектами (2 часа).
- 4) Импорт/экспорт и управление данными в R (2 часа).
- 5) Векторные и матричные операции в R (2 часа).
- 6) Организация вычислений в R (2 часа).
- 7) Ускорение вычислений в R (2 часа).
- 8) Построение базовых типов графиков в R (2 часа).
- 9) Работа с аппаратными средствами в R (2 часа).
- 10) Разработка и отладка программных R-кодов (2 часа).
- 11) Разработка, оптимизация и тестирование пользовательских функций в R (2 часа).
- 12) Работа с библиотеками R-функций (2 часа).
- 13) Инфраструктура основных репозиториях R-функций (2 часа).
- 14) Инфраструктура репозитория Bioconductor и его использование в биологических и медицинских исследованиях (2 часа).

ИНФОРМАЦИОННАЯ ЧАСТЬ

Перечень основной литературы

- 1) *Кабаков Р. И.* R в действии. Анализ и визуализация данных в программе R / Р. И. Кабаков. – М.: ДМК Пресс, 2014. – 588 с.
- 2) *Маслицкий С. Э.* Статистический анализ и визуализация данных с помощью R / С. Э. Маслицкий, В. К. Шитиков. Хайдельберг – Лондон – Тольятти, 2014. – 401 с.
- 3) *Bell D.* R programming: A step-by-step guide for absolute beginners / D. Bell. – Guzzler Media LLC, 2019. – 163 p.
- 4) *Gillespie C.* Efficient R programming. First edition / C. Gillespie. – O'Reilly Media Inc., 2017. – 203 p.

Перечень дополнительной литературы

- 1) *Allesina S.* Computing skills for biologists: A toolbox / S. Allesina, M. Wilmes. – Princeton University Press, 2019. – 417 p.
- 2) *Beeley C.* Web application development with R using Shiny. Third edition / C. Beeley, S. Sukhdev. – Packt, 2018. – 145 p.
- 3) *Venables W. N.* An introduction to R / W. N. Venables, D. M. Smith, R Development Core Team. – CRAN, 2020. – 105 p.
- 4) *Wickham H.* R packages: Organize, test, document, and share your code. First edition / H. Wickham. – O'Reilly Media Inc., 2015. – 202 p.

Интернет-ресурсы

- 1) Гринев В. В. Авторские видеолекции на YouTube канале Grinev's Educational Channel (плейлист «R Programming Basics») (youtube.com/channel/UCYQ8QwQAX8ubVYYuegxNTYQ).
- 2) Маслицкий С. Э. R: Анализ и визуализация данных (r-analytics.blogspot.com).
- 3) Шитиков В. К., Маслицкий С. Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R (r-analytics.github.io/data-mining).
- 4) Экология с R (stok1946.blogspot.com).
- 5) Grolemond G. Hands-on programming with R (rstudio-education.github.io/hopr).
- 6) Introduction to R (by Petr Nazarov, Luxembourg Institute of Health; <http://edu.sablab.net/r2017/index.html>).
- 7) Peng R. D., Kross S., Anderson B. Mastering software development in R (bookdown.org/rdpeng/RProgDA).
- 8) Quick-R (www.statmethods.net).
- 9) R tutorial: An R introduction to statistics (www.r-tutor.com).
- 10) RDocumentation (www.rdocumentation.org).