

۲۰۸۸۲۹
۷۵۱۵۵



مؤسسه فرهنگی هنری
دیباگران تهران

به نام خدا



مؤسسه فرهنگی هنری
دیباگران تهران

راهنمای جامع برنامه نویسی

R

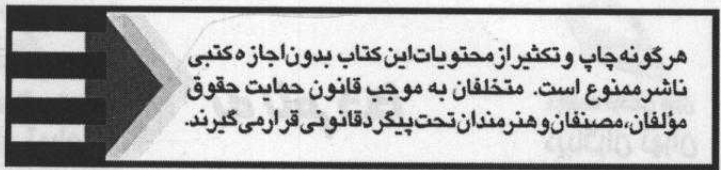
(مقدماتی و پیشرفته)

مؤلف:

دکتر سید مهدی صالحی

(عضو هیئت علمی دانشگاه نیشابور)

با همکاری: ریحانه عرفه



عنوان کتاب: راهنمای جامع برنامه نویسی R (مقدماتی و پیشرفته)

مؤلف: دکتر سید مهدی صالحی

ویراستار: ریحانه عرفه

ناشر: موسسه فرهنگی هنری دیباگران تهران

صفحه آرای: فروش عبداله

طراح جلد: داریوش فرسای

نوبت چاپ: اول

تاریخ نشر: ۱۳۹۸

چاپ و صحافی: صدف

تیراژ: ۱۰۰ جلد

قیمت: ۷۹۰۰۰ ریال

شابک: ۹۷۸-۶۲۲-۲۱۸-۲۵۳-۳

نشانی واحد فروش: تهران، میدان انقلاب،

خ کارگر جنوبی، روبروی پاساژ مهستان،

پلاک ۱۲۵۱

تلفن: ۲۲۰۸۵۱۱۱-۶۶۴۱۰۰۴۶

فروشگاههای اینترنتی دیباگران تهران :

WWW.MFTBOOK.IR

www.dibbook.ir

www.dibagaran-tehran.com

نشانی تلگرام: @mftbook

نشانی اینستاگرام دیبا dibagaran_publishing

هر کتاب دیباگران، یک فرصت جدید شغلی.

هرگوشی همراه، یک فروشگاه کتاب دیباگران تهران.

از طریق سایتها و اپ دیباگران، در هر جای ایران به کتابهای ما دسترسی دارید.

سرشناسه: صالحی، سید مهدی، ۱۳۶۵-
عنوان و نام پدیدآور: راهنمای جامع برنامه نویسی R
(مقدماتی و پیشرفته) / مؤلف: سید مهدی صالحی
ویراستار: ریحانه عرفه.
مشخصات نشر: تهران: دیباگران تهران: ۱۳۹۸
مشخصات ظاهری: ۳۰۲ص، مصور.
شابک: ۹۷۸-۶۲۲-۲۱۸-۲۵۳-۳
وضعیت فهرست نویسی: فیا
موضوع: آیین برنامه نویسی کامپیوتر
موضوع: R (computer programming language)
موضوع: آمار - داده داری
موضوع: Statistical data processing
شناسه افزوده: عرفه، ریحانه، ۱۳۷۶-
رده بندی کنگره: QA ۲۷۶/۴۵
رده بندی دیویی: ۵۱۹/۵۰۲۸۵۵۱۳۳
شماره کتابشناسی ملی: ۶۰۶۰۳۵۴

فهرست مطالب

فصل اول / معرفی نرم افزار و زبان برنامه نویسی R..... ۱۹

۱-۱- مقدمه..... ۲۰

۲-۱- چرا R؟..... ۲۰

۳-۱- چند نکته اساسی درباره‌ی زبان برنامه نویسی R..... ۲۲

۴-۱- شروع کار با R..... ۲۳

۴-۱- نحوه‌ی اجرای برنامه‌های R..... ۲۳

۱-۱-۱- روش تعاملی..... ۲۳

۱-۲-۲- روش BATCH..... ۲۵

۲-۴-۱- آشنایی اجزای اصلی با منوهای نرم افزار..... ۲۵

۳-۴-۱- روش ها- جستجو در نرم افزار..... ۲۶

۵-۱- بسته‌های نرم افزار..... ۲۶

۱-۵-۱- نصب بسته‌های نرم افزار..... ۲۸

۲-۵-۱- فراخوانی بسته‌های نرم افزاری..... ۲۹

۳-۵-۱- به روز رسانی بسته‌های نرم افزاری..... ۲۹

۶-۱- مفاهیم پایه..... ۳۰

۱-۶-۱- محاسبات ساده‌ی ریاضی..... ۳۰

۲-۶-۱- متغیرها..... ۳۲

۱-۲-۶-۱- ایجاد متغیرها..... ۳۲

۲-۲-۶-۱- حذف کردن متغیرها..... ۳۳

۷-۱- انواع داده‌ها..... ۳۳

۱-۷-۱- داده‌های عددی (double یا numeric)..... ۳۴

۲-۷-۱- داده‌های صحیح (integer)..... ۳۴

۳-۷-۱- داده‌های موهومی (complex)..... ۳۵

۴-۷-۱- داده‌های رشته‌ای (character)..... ۳۵

۵-۷-۱- داده‌های منطقی (logical)..... ۳۶

فصل دوم / بردارها ۳۸

۳۹	۱-۲- مقدمه.....
۳۹	۲-۲- ایجاد بردارها.....
۳۹	۱-۲-۲- تابع <code>c()</code>
۴۱	۲-۲-۲- تابع <code>scan()</code>
۴۲	۳-۲-۲- ایجاد دنباله‌های حسابی.....
۴۳	۴-۲-۲- دنباله‌ی اعداد تکراری.....
۴۳	۳-۲- شناسایی و ویرایش درایه‌ها.....
۴۵	۴-۲- استخراج عناصر مشروط.....
۴۶	۵-۲- اعمال جبر - درایه به درایه.....
۴۸	۶-۲- معرفی برخی توابع مهم برای بردارها.....
۴۹	۱-۶-۲- توابع آماری.....
۴۹	۱-۱-۶-۲- تابع <code>summary()</code>
۴۹	۲-۱-۶-۲- تابع <code>sample()</code>
۵۰	۳-۱-۶-۲- توابع <code>ftable()</code> , <code>table()</code>
۵۰	۲-۶-۲- توابع ریاضی.....
۵۰	۱-۲-۶-۲- خانواده توابع <code>cum*</code>
۵۱	۲-۲-۶-۲- تابع <code>order()</code>
۵۲	۳-۲-۶-۲- توابع <code>round()</code> , <code>floor()</code> , <code>ceiling()</code>
۵۲	۴-۲-۶-۲- خانواده توابع <code>gamma()</code>
۵۳	۳-۶-۲- چند تابع و عملگر متفرقه.....
۵۳	۱-۳-۶-۲- تابع <code>which()</code>
۵۵	۲-۳-۶-۲- توابع <code>any()</code> و <code>all()</code>
۵۷	۳-۳-۶-۲- عملگر <code>'%in%'</code>
۵۸	۴-۳-۶-۲- تابع <code>names()</code>
۵۹	۷-۲- توابع برداری شده.....
۶۰	۸-۲- داده‌های گمشده.....
۶۱	۱-۸-۲- کار با داده‌های گمشده.....
۶۲	۹-۲- معرفی بیش‌تر مقادیر <code>NaN</code> , <code>Inf</code> و <code>NULL</code>
۶۳	۱۰-۲- توابع مهم در یک نگاه.....

فصل سوم / ماتریس‌ها و آرایه‌ها ۶۶

۶۷	۱-۳- مقدمه.....
۶۷	۲-۳- ایجاد ماتریس‌ها.....

۶۷.....	matrix() تابع ۱-۲-۳
۷۱.....	dim() تابع ۲-۲-۳
۷۱.....	cbind() و rbind() توابع ۳-۲-۳
۷۲.....	diag() تابع ۴-۲-۳
۷۳.....	۳-۳-۳ فراخوانی و ویرایش درایه‌ها
۷۷.....	۴-۳-۳ اعمال جبری درایه به درایه
۷۹.....	۵-۳-۳ توابع و عملگرهای ماتریسی
۷۹.....	۱-۵-۳ عملگرهای ضرب
۸۱.....	۲-۵-۳ solve() تابع
۸۵.....	۶-۲-۳ آرایه‌ها
۸۷.....	۱-۶-۳ فراخوانی درایه‌ها، بردارها و ماتریس‌ها
۸۷.....	۳-۶-۳ اعمال جبری

فصل ۴: نام / عمل‌ها، داده‌های چارچوب‌دار و فهرست‌ها ۹۱

۹۲.....	۱-۴-۳ داده‌ها
۹۲.....	۲-۴-۳ عامل‌ها (FACTOR)
۹۲.....	۱-۲-۴ ایجاد عامل‌ها
۹۲.....	۱-۱-۲-۴ factor() تابع
۹۵.....	۲-۱-۲-۴ gl(), تابع
۹۵.....	۳-۱-۲-۴ cut() تابع
۹۶.....	۳-۴-۳ داده‌های چارچوب‌دار (DATAFRAME)
۹۶.....	۱-۳-۴ ایجاد داده‌های چارچوب‌دار
۹۷.....	۲-۳-۴ فراخوانی و ویرایش ستون‌ها، ردیف‌ها و درایه‌ها
۹۷.....	۱-۲-۳-۴ '[' تابع
۹۸.....	۲-۲-۳-۴ attach() تابع
۹۹.....	۳-۲-۳-۴ '\$' تابع
۱۰۰.....	۴-۲-۳-۴ subset() تابع
۱۰۰.....	۳-۳-۴ اعمال جبری درایه به درایه
۱۰۱.....	۴-۳-۴ برخی توابع مهم برای داده‌های چارچوب‌دار
۱۰۱.....	۱-۴-۳-۴ str() تابع
۱۰۲.....	۲-۴-۳-۴ merge() تابع
۱۰۳.....	۴-۴-۳ فهرست‌ها (LIST)
۱۰۳.....	۱-۴-۴ ایجاد فهرست‌ها
۱۰۳.....	۱-۱-۴-۴ list() تابع

۱۰۴.....	vector() تابع ۲-۱-۴-۴
۱۰۵.....	split() تابع ۳-۱-۴-۴
۱۰۷.....	۲-۴-۴-فرخوانی و ویرایش آیتم‌ها
۱۰۸.....	۳-۴-۴-اعمال جبری
۱۰۹.....	۴-۴-۴-فهرست‌های بازگشتی
۱۱۰.....	۵-۴-۴-هموار کردن یک فهرست
۱۱۳.....	۵-۴-شناسایی اشیاء معرفی شده
۱۱۴.....	۶-۴-فرخوانی و ذخیره‌ی داده‌ها در خارج از محیط R
۱۱۵.....	۱-۶-۴-ذخیره‌ی موقت داده‌ها با فرمت RData
۱۱۶.....	۲-۶-۴-ذخیره‌ی داده‌ها با قالب‌های معروف
۱۱۶.....	۱-۲-۴-قالب متنی
۱۱۷.....	۲-۲-۶-۴-تبدیل کلاس
۱۲۱.....	۲-۶-۴-داده‌ی نرم‌افزارهای آماری SPSS, SAS, Minitab, Stata
۱۲۲.....	۳-۶-۴-ذخیره‌سازی و خوانش مرحله‌ای به مرحله‌ی داده‌ها
۱۲۲.....	۱-۳-۶-۴-تابع cat()
۱۲۳.....	۲-۳-۶-۴-تابع readLines()
۱۲۴.....	۳-۳-۶-۴-تابع writeLines()

فصل پنجم / توابع پیشرفته ۱۲۶.....

۱۲۷.....	۱-۵-مقدمه
۱۲۷.....	۲-۵-تابع apply()
۱۳۳.....	۳-۵-تابع lapply()
۱۳۶.....	۴-۵-تابع sapply()
۱۳۸.....	۵-۵-تابع mapply()
۱۳۹.....	۶-۵-تابع tapply()
۱۴۳.....	۷-۵-تابع aggregate()
۱۴۵.....	۸-۵-تابع by()
۱۴۸.....	۹-۵-تابع get()
۱۵۰.....	۱۰-۵-تابع replicate()

فصل ششم / اصول برنامه‌نویسی R ۱۵۳.....

۱۵۴.....	۱-۶-مقدمه
۱۵۴.....	۲-۶-حلقه‌ها
۱۵۵.....	۱-۲-۶-حلقه‌ی for
۱۵۹.....	۲-۲-۶-حلقه‌ی while

۱۶۱.....	۳-۶- دستورات شرطی
۱۶۱.....	۱-۳-۶- دستور if
۱۶۲.....	۲-۳-۶- دستور if-else
۱۶۳.....	۳-۳-۶- تابع ifelse()
۱۶۴.....	۴-۳-۶- تابع switch()
۱۶۵.....	۴-۶- دستورات کنترلی
۱۶۸.....	۵-۶- تابع نویسی
۱۷۵.....	۱-۵-۶- توابع بازگشتی
۱۷۷.....	۲-۵-۶- تابع Vectorize()
۱۸۰.....	۳-۵-۶- مشخص کردن پیش فرض برای آرگومان‌ها
۱۸۲.....	۴-۵-۶- ساخت عملگرها و توابع جایگذاری
۱۸۵.....	۵-۶- متغیرهای محلی و سراسری
۱۹۸.....	۶-۶- گریزی به برنامه‌نویسی موازی
۱۹۸.....	۲-۶-۶- بسته snow
۲۰۲.....	۲-۶-۶- بسته parallel
۲۰۴.....	۳-۶-۶- بسته foreach
۲۰۵.....	۴-۶-۶- موازی یا سری: کدامیک؟!

فصل هفتم / کار با داده‌های سری

۲۰۷.....	۱-۷- مقدمه
۲۰۷.....	۲-۷- ساختار داده‌های متنی
۲۰۸.....	۳-۷- ایجاد کاراکترها و رشته‌های متنی
۲۰۹.....	۱-۳-۷- تابع character()
۲۱۰.....	۲-۳-۷- توابع paste() و paste0()
۲۱۱.....	۳-۳-۷- تابع toString()
۲۱۲.....	۴-۷- چاپ داده‌های متنی
۲۱۲.....	۱-۴-۷- تابع print()
۲۱۳.....	۲-۴-۷- تابع noquote()
۲۱۳.....	۳-۴-۷- تابع cat() برای کاراکترها
۲۱۵.....	۴-۴-۷- تابع sprintf()
۲۱۷.....	۵-۷- مدیریت داده‌های متنی
۲۱۸.....	۱-۵-۷- تابع nchar()
۲۱۹.....	۲-۵-۷- توابع tolower() و toupper()
۲۱۹.....	۳-۵-۷- تابع casefold()

۲۱۹.....	chartr() تابع ۴-۵-۷
۲۲۰.....	grep() تابع ۵-۵-۷
۲۲۱.....	gsub() تابع ۶-۵-۷
۲۲۲.....	abbreviate() تابع ۷-۵-۷
۲۲۲.....	substring() تابع ۸-۵-۷
۲۲۳.....	strsplit() تابع ۹-۵-۷
۲۲۷.....	۶-۷- توابع مجموعه‌ای در بردارهای کاراکتری
۲۲۷.....	union() تابع ۱-۶-۷
۲۲۷.....	intersect() تابع ۲-۶-۷
۲۲۷.....	setdiff() تابع ۳-۶-۷
۲۲۸.....	setequal() تابع ۴-۶-۷
۲۲۸.....	identical() تابع ۵-۶-۷
۲۲۸.....	%in% is.element() تابع ۶-۶-۷
۲۲۹.....	sort() تابع ۷-۶-۷
۲۲۹.....	۷-۷- پردازش متن با استفاده از بسته‌ی نم‌نزاری tm
۲۳۰.....	۱-۷-۷- ورود اطلاعات با استفاده از تابع corpus()
۲۳۱.....	۲-۷-۷- پیش‌پردازش متن
۲۳۲.....	۱-۲-۷-۷- tfidf() تابع
۲۳۲.....	۲-۲-۷-۷- tm_map() تابع
۲۳۵.....	۳-۷-۷- تحلیل متن
۲۳۵.....	۱-۳-۷-۷- ماتریس کلمات (Term-Document Matrix)
۲۳۶.....	۲-۳-۷-۷- استخراج کلمات با فراوانی مورد نظر
۲۳۶.....	۳-۳-۷-۷- حذف کلمات نادر
۲۳۷.....	۴-۳-۷-۷- بررسی روابط میان کلمات
۲۳۷.....	۴-۷-۷- ابر کلمات (Word Cloud)

فصل هشتم / آشنایی با محیط گرافیکی R ۲۴۰

۲۴۱.....	۱-۸- مقدمه
۲۴۱.....	۲-۸- گرافیک پایه
۲۴۱.....	۱-۲-۸- رنگ‌ها در R
۲۴۲.....	۲-۲-۸- نمودارهای آماری (توصیفی)
۲۴۲.....	۱-۲-۲-۸- نمودار دایره‌ای
۲۴۴.....	۲-۲-۲-۸- نمودار میله‌ای
۲۴۶.....	۳-۲-۲-۸- نمودار جعبه‌ای

۲۴۷.....	۴-۲-۲-۸- نمودار مستطیلی (هیستوگرام)
۲۴۸.....	۵-۲-۲-۸- نمودار پراکنش
۲۵۱.....	۶-۲-۲-۸- نمودار چندک-چندک
۲۵۲.....	۷-۲-۲-۸- نمودار برآورد چگالی کرنل
۲۵۳.....	۳-۲-۸- نمودارهای ریاضی
۲۵۳.....	۱-۳-۲-۸- منحنی توابع
۲۵۴.....	۲-۳-۲-۸- رسم پاره خط
۲۵۵.....	۳-۳-۲-۸- چند ضلعی
۲۵۶.....	۴-۲-۸- نمودارهای سه بُعدی
۲۵۶.....	۱-۴-۲-۸- رویه (زرفنما)
۲۵۷.....	۲-۴-۲-۸- نمودار خط تراز (کانتور)
۲۵۸.....	۲-۸- ویرایش و سفارشی کردن نمودارها
۲۵۸.....	۱-۵-۸- افزودن نقاط به نمودار
۲۵۸.....	۲-۵-۸- افزودن خطوط به نمودار
۲۵۹.....	۵-۲-۸- ایجاد مینه شطرنجی
۲۵۹.....	۴-۵-۲-۸- مشخصات نقاط نمودار
۲۶۰.....	۵-۵-۲-۸- شش‌مضلعی جعبه‌ای
۲۶۳.....	۶-۵-۲-۸- افزودن جعبه را نما نمودار
۲۶۴.....	۷-۵-۲-۸- نوشتن متن در نمودار
۲۶۸.....	۸-۵-۲-۸- طراحی صفحه گرافیکی
۲۶۹.....	۶-۲-۸- ذخیره‌ی نمودارها
۲۶۹.....	۳-۸- بسته نرم‌افزاری ggplot2
۲۷۰.....	۱-۳-۸- آشنایی با بسته‌گرافیکی ggplot2
۲۷۰.....	۱-۱-۳-۸- لایه‌های ضروری
۲۷۰.....	۲-۱-۳-۸- لایه‌های اختیاری
۲۷۱.....	۲-۳-۸- تابع ggplot()
۲۷۱.....	۱-۲-۳-۸- نمودار میله‌ای
۲۷۳.....	۲-۲-۳-۸- هیستوگرام
۲۷۴.....	۳-۲-۳-۸- نمودار پراکنش
۲۷۵.....	۴-۲-۳-۸- نمودار جعبه‌ای
۲۷۷.....	۳-۳-۸- بهبود کارایی نمودارها
۲۷۹.....	۴-۳-۸- ذخیره‌ی نمودارها

پیوست الف / تمرین‌ها ۲۸۱

تمرین‌های فصل اول ۲۸۲

تمرین‌های فصل دوم ۲۸۲

تمرین‌های فصل سوم ۲۸۴

تمرین‌های فصل چهارم ۲۸۶

تمرین‌های فصل پنجم ۲۸۷

تمرین‌های فصل ششم ۲۸۹

تمرین‌های فصل هفتم ۲۹۰

تمرین‌های فصل هشتم ۲۹۰

پیوست ب / آمار با توزیع‌های آماری ۲۹۲

نمایه ۲۹۶

مراجع ۳۰۱

خط مشی کیفیت انتشارات مؤسسه فرهنگی هنری دیباگران تهران در عرصه کتاب ملی است که بتواند خواسته های به روز جامعه فرهنگی و علمی کشور را تا حد امکان پوشش دهد. هر کتاب دیباگران تهران، یک فرصت جدید شغلی

حاجت و سیاست ایزد منان را که با الطاف بیکران خود این توفیق را به ما ارزانی داشت تا بتوانیم در راه ارتقاء دانش عمومی و فرهنگی این مرز و بوم در زمینه چاپ و نشر کتب علمی دانشگاهی، علوم پایه و به ویژه علوم کامپیوتر و انفورماتیک گام هایی هرچند کوچک برداشته و در انجام رسالتی که بر عهده داریم، مژده واقع داریم.

گسترده گی نرم افزار و سخت افزار روز افزون آن، شرایطی را به وجود آورده که هر روز شاهد تحولات اساسی چشمگیری در سطح جهان هستیم. این گسترش و توسعه نیاز به منابع مختلف از جمله کتاب را به عنوان قدیمی ترین و راحت ترین راه دستیابی به اطلاعات و اطلاع رسانی، بیش از پیش روشن می نماید.

در این راستا، واحد انتشارات مؤسسه فرهنگی هنری دیباگران تهران با همکاری جمعی از اساتید، مؤلفان، مترجمان، متخصصان، پژوهشگران، محققان و نیز پرسنل ورزیده و ماهر در زمینه امور نشر در صدد هستند تا با تلاش های مستمر خود برای رفع کمبودها و نیازهای موجود، منابعی پربار، معتبر و با کیفیت مناسب در اختیار علاقمندان قرار دهند.

کتابی که در دست دارید با همت "دکتر سیده زهرا قزلباش" و با همکاری سرکار خانم ریحانه عرفه " و تلاش جمعی از همکاران انتشارات میسر گشته که جایسته است از یکایک این گرامیان تشکر و قدردانی کنیم.

کارشناسی و نظارت بر محتوا: زهرا قزلباش

در خاتمه ضمن سپاسگزاری از شما دانش پژوه گرامی درخواست می نمایم تا مراجعه به آدرس dibagaran.mft.info (ارتباط با مشتری) فرم نظرسنجی را برای کتابی که در دست دارید تکمیل و ارسال نموده، انتشارات دیباگران تهران را که جلب رضایت و وفاداری مشتریان را هدف خود می داند، یاری فرمایید.

امیدواریم همواره بهتر از گذشته خدمات و محصولات خود را تقدیم حضورتان نماییم.

مدیر انتشارات

مؤسسه فرهنگی هنری دیباگران تهران

bookmarket@mft.info