

۱۴۹۴۵۴۷
۹۶۷۱

به نام خدا



آموزش زبان برنامه نویسی

www.ketab.ir

مؤلف

دکتر بهنام آبابایی

هرگونه چاپ و تکثیر از محتویات این کتاب بدون اجازه کتبی
ناشر ممنوع است. متخلفان به موجب قانون حمایت حقوق
مؤلفان، مصنفان و هنرمندان تحت پیگرد قانونی قرار می گیرند.

آموزش زبان برنامه نویسی R

مؤلف: دکتر بهنام آبابایی

ناشر: مؤسسه فرهنگی هنری دیباگران تهران

حروفچینی: صفحه آرای: شبنم هاشم زاده

طرح روی جلد: دانش فرسای

چاپ: دانشجو

نوبت چاپ: اول

تاریخ نشر: ۱۳۹۶

تیراژ: ۱۰۰ جلد

قیمت: ۱۸۰۰۰۰ ریال

شابک: ۹۷۸-۶۰۰-۱۲۴-۵۸۸-۶

نشانی واحد فروش: تهران، میدان انقلاب،

خ کارگر جنوبی، روبروی پاساژ مهستان،

پلاک ۱۲۵۱

سرشناسه: آبابایی، بهنام، ۱۳۶۲-

عنوان و نام پدید آور: آموزش زبان برنامه نویسی R / مؤلف: بهنام آبابایی

مشخصات نشر: تهران- دیباگران تهران- ۱۳۹۶

مشخصات ظاهری: ۱۵۴ ص. مصور.

شابک: ۹۷۸-۶۰۰-۱۲۴-۵۸۸-۶

وضعیت فهرست نویسی: فیفا

نوع: آر (زبان برنامه نویسی کامپیوتر)

موضوع: R (computer program language)

موضوع: آمار - داده پردازی

رشته: Statistics-Data processing

رده بندی کنگ: ۳۹۶ : ۴۲۴ / ۴۵ / ۲۷۶ QA

رده بندی دیویی: ۵۱۹ / ۵۲۸ / ۱۳

شماره کتابشناسی ملی: ۲۰۷۶۰۰۶

تلفن: ۰۰۴۶۱۰۶۶۴-۲۲۰۸۵۱۱۱ کد پستی: ۱۳۱۴۹۸۳۱۸۵

پست الکترونیکی: bookmarket@mftmail.com

فروش اینترنتی:

www.mftshop.com

www.mftbook.ir

نشانی اینستاگرام: Dibagaran_publishing

نشانی تلگرام: @mftbook

فهرست مطالب

فصل ۱: معرفی و مقدمات ۱۳

- ۱-۱ محیط R ۱۳
- ۲-۱ کاربردهای R در محاسبات آماری ۱۴
- ۳-۱ استفاده از R به صورت تعاملی ۱۴
- ۴-۱ یک جلسه آموزشی ۱۶
- ۵-۱ راهنمای R برای کار با توابع ۱۶
- ۶-۱ دستورهایی R و حساسیت به اندازۀ حروف ۱۷
- ۷-۱ فراخوانی و اصلاح دستورهایی قبلی ۱۷
- ۸-۱ اجرای دستورها از یا انتقال خروجی‌ها به یک فایل ۱۸
- ۹-۱ دوام داده‌ها و حذف اشیاء ۱۸

فصل ۲: تغییرات ساده؛ اعداد و بردارها ۲۱

- ۱-۲ بردارها و نسبت‌دهی ۲۱
- ۲-۲ محاسبات بر روی بردارها ۲۲
- ۳-۲ تولید رشته‌های عددی ۲۴
- ۴-۲ بردارهای منطقی ۲۵
- ۵-۲ مقادیر مفقود ۲۶

- ۲۶ ۶-۲ بردارهای کاراکتری
- ۲۷ ۷-۲ انتخاب و تغییر بخشی از یک مجموعه داده
- ۲۹ ۸-۲ انواع دیگر اشیاء

۳۱ فصل ۳: اشیاء، مُدها و صفات کیفی

- ۳۱ ۱-۳ صفات کیفی ذاتی: length و mode
- ۳۲ ۲-۳ تغییر طول اشیاء
- ۳۳ ۳-۳ مشاهده و تطبیق صفات
- ۳۴ ۴-۳ کلاس یک شی

۳۵ فصل ۴: فاکتورهای مرتب‌شده و مرتب‌نشده

- ۳۵ ۱-۴ یک مثال خاص
- ۳۶ ۲-۴ تابع apply() و آرایه‌های ناصاف
- ۳۸ ۳-۴ فاکتورهای مرتب‌شده

۳۹ فصل ۵: آرایه‌ها و ماتریس‌ها

- ۳۹ ۱-۵ آرایه‌ها
- ۴۰ ۲-۵ فهرست‌سازی از آرایه‌ها: بخش‌هایی از آرایه
- ۴۰ ۳-۵ ماتریس اندیس‌ها
- ۴۲ ۴-۵ تابع array()
- ۴۳ ۱-۴-۵ بردارهای مخلوط و آرایه‌های ریاضی: قاعده بازچرخش
- ۴۴ ۵-۵ حاصل ضرب خارجی دو آرایه
- ۴۵ ۶-۵ محاسبه ترانزپوز یک آرایه

۴۵	۷-۵ کار با ماتریس‌ها
۴۵	۷-۵-۱ ضرب ماتریس‌ها
۴۶	۷-۵-۲ معادلات خطی و تعادل‌ها
۴۷	۷-۵-۳ مقادیر مشخصه و بردارهای مشخصه
۴۷	۷-۵-۴ تجزیه مقدار تکین و دترمینان‌ها
۴۸	۷-۵-۵ برازش منحنی‌ها به روش حداقل مربعات و تجزیه QR
۴۹	۸-۵ توابع cbind() و rbind()
۵۰	۹-۵ کاربرد تابع c() بر روی آرایه‌ها
۵۰	۱۰-۵ جداول فراوانی از فاکتورها

فصل ۶: لیست‌ها و دیتافریم‌ها ۵۳

۵۳	۶-۱ لیست‌ها
۵۴	۶-۲ ایجاد و تغییر لیست‌ها
۵۵	۶-۲-۱ الحاق لیست‌ها
۵۵	۶-۳ دیتافریم‌ها
۵۶	۶-۳-۱ ایجاد دیتافریم‌ها
۵۶	۶-۳-۲ توابع attach() و detach()
۵۷	۶-۳-۳ کار با دیتافریم‌ها
۵۸	۶-۳-۴ الصاق لیست‌ها به آدرس جستجو
۵۸	۶-۳-۵ مدیریت مسیر جستجو

فصل ۷: خواندن اطلاعات از فایل‌ها ۶۱

۶۱	۷-۱ تابع read.table()
۶۲	۷-۲ تابع scan()

- ۶۳ ۳-۷ دسترسی به مجموعه داده‌های داخلی R
- ۶۴ ۴-۷ بارگذاری داده‌ها از سایر بسته‌های R
- ۶۴ ۵-۷ ویرایش داده‌ها

۶۷ فصل ۸: توزیع‌های احتمالاتی

- ۶۷ ۱-۸ مجریه‌های از جداول آماری
- ۶۸ ۲-۸ بررسی توزیع احتمالاتی داده‌ها
- ۷۴ ۳-۸ آزمون‌های یک و دو نمونه‌ای

۷۹ فصل ۹: گروه‌بندی، دسته‌ها و دستورهای شرطی

- ۷۹ ۱-۹ عبارات گروه‌بندی شده
- ۷۹ ۲-۹ گزاره‌های کنترلی
- ۷۹ ۱-۲-۹ گزاره‌های شرطی: if
- ۸۰ ۲-۲-۹ ساختارهای تکراری: while و repeat

۸۳ فصل ۱۰: نوشتن توابع شخصی

- ۸۳ ۱-۱۰ مثال‌های ساده
- ۸۵ ۲-۱۰ ایجاد عملگرهای جفتی جدید
- ۸۵ ۳-۱۰ آرگومان‌های نامگذاری شده و مقادیر پیش‌فرض
- ۸۶ ۴-۱۰ کاربرد آرگومان "..."
- ۸۷ ۵-۱۰ نسبت‌دهی درون توابع
- ۸۷ ۶-۱۰ مثال‌های پیشرفته‌تر
- ۸۷ ۱-۶-۱۰ فاکتورهای کارایی در طرح‌های بلوکی

- ۸۸ ۱۰-۶-۲ حذف تمامی نام‌ها از یک آرایه پرنیت‌شده
- ۸۹ ۱۰-۷ قلمرو
- ۹۲ ۱۰-۸ شخصی‌سازی یک محیط
- ۹۳ ۱۰-۹ کلاس‌ها، توابع عمومی و شی‌گرایی

فصل ۱۱: مدل‌های آماری

- ۹۷ ۱۱-۱ تعریف مدل‌های آماری: فرمول‌ها
- ۹۸ ۱۱-۱-۱ مثال‌ها
- ۱۰۰ ۱۱-۲ مدل خطی
- ۱۰۰ ۱۱-۳ توابع عمومی برای استخراج اطلاعات مدل‌های آماری
- ۱۰۲ ۱۱-۴ تحلیل واریانس و پایه مدل‌ها
- ۱۰۳ ۱۱-۴-۱ جدول‌های ANOVA
- ۱۰۳ ۱۱-۵ به‌روزرسانی مدل‌های برازش داده شده
- ۱۰۴ ۱۱-۶ مدل‌های خطی تعمیم‌یافته
- ۱۰۵ ۱۱-۶-۱ خانواده‌ها
- ۱۰۶ ۱۱-۶-۲ تابع $glm()$
- ۱۰۶ ۱۱-۶-۱-۲ خانواده مدل‌های گوسی
- ۱۰۷ ۱۱-۶-۲-۲ خانواده مدل‌های دوجمله‌ای
- ۱۰۹ ۱۱-۶-۲-۳ خانواده مدل‌های پواسون
- ۱۰۹ ۱۱-۷ مدل‌های غیرخطی حداقل مربعات و حداکثر درست‌نمایی
- ۱۰۹ ۱۱-۷-۱ حداقل مربعات
- ۱۱۲ ۱۱-۷-۲ حداکثر درست‌نمایی
- ۱۱۲ ۱۱-۸ برخی مدل‌های غیراستاندارد

- ۱۱۶ ۱-۱۲ دستورهای گرافیکی سطح بالا
- ۱۱۶ ۱-۱-۱۲ تابع `plot()`
- ۱۱۷ ۲-۱-۱۲ نمایش داده‌های چندمتغیره
- ۱۱۸ ۳-۱-۱۲ نمایش اشیاء گرافیکی
- ۱۱۹ ۴-۱-۱۲ آرگومان‌های توابع گرافیکی سطح بالا
- ۱۲۰ ۵-۱-۱۲ دستورهای گرافیکی سطح پایین
- ۱۲۱ ۱-۲-۱۲ نمایش دایره ریاضی
- ۱۲۲ ۲-۲-۱۲ فرمت‌های بصری هرشی
- ۱۲۲ ۳-۱۲ تعامل با گراف‌ها
- ۱۲۴ ۴-۱۲ کاربرد پارامترهای گرافیکی
- ۱۲۴ ۱-۴-۱۲ تغییرات دائمی: تابع `persist()`
- ۱۲۵ ۲-۴-۱۲ تغییرات موقتی: آرگومان‌های تابع `gcf()`
- ۱۲۶ ۵-۱۲ لیست پارامترهای گرافیکی
- ۱۲۶ ۱-۵-۱۲ اجزاء گرافیکی
- ۱۲۷ ۲-۵-۱۲ محورها و تیک‌ها
- ۱۲۸ ۳-۵-۱۲ حاشیه‌ها
- ۱۲۹ ۴-۵-۱۲ استفاده همزمان از چند تصویر
- ۱۳۱ ۶-۱۲ برنامه‌های راه‌اندازی ابزارهای گرافیکی
- ۱۳۲ ۱-۶-۱۲ استفاده از فایل‌های گرافیکی PostScript
- ۱۳۳ ۲-۶-۱۲ استفاده همزمان از چند ابزار گرافیکی
- ۱۳۴ ۷-۱۲ گراف‌های پویا

- ۱۳۶ ۱-۱۳ بسته‌های استاندارد
- ۱۳۶ ۲-۱۳ بسته‌های توسعه‌دهندگان و منبع CRAN
- ۱۳۶ ۳-۱۳ فضاهای نام‌گذاری

فصل ۱۴: کار با سیستم‌های عامل

- ۱۳۹ ۱-۱۴ فایل‌ها، پوشه‌ها
- ۱۴۰ ۲-۱۴ آدرس‌دهی فایل‌ها
- ۱۴۱ ۳-۱۴ دستورهای سیستم
- ۱۴۲ ۴-۱۴ فشردن سازی و آشپزها

فصل ۱۵: یک جلسه آموزشی

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.

حمد و سپاس ایزد منان را که با الطاف بیکران خود این توفیق را به ما ارزانی داشت تا بتوانیم در راه ارتقای 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، مؤثر واقع شویم.

گسترده گی علوم و توسعه ررافزون 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 رایج 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 درصدد هستند تا با تلاش های مستمر خود برای رفع کمبودها و نیازهای موجود، منابعی پربار، معتبر و با کیفیت مناسب در اختیار علاقمندان قرار دهد.

کتابی که در دست دارید با همت "دکتر بهنام آبابا" و تلاش جمعی از همکاران انتشارات میسر گشته که شایسته است از یکایک این گرامیان تشکر و قدردانی کنیم.

کارشناسی و نظارت بر محتوا: زهره قزلباش

در خاتمه ضمن سپاسگزاری از شما دانش پژوه گرامی درخواست می نماید با 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

Publishing@mftmail.com

مقدمه مولف

گردآوری و ترجمه مطالب کتاب حاضر به نحوی صورت گرفته که درک مفاهیم را برای کاربرانی که شاید با مفاهیم بنیادی برنامه‌نویسی آشنایی نداشته باشند نیز ساده‌تر سازد. در عین حال، هرکجا که نیاز برده، مطالبی براساس تجارب شخصی مترجم به متون اصلی افزوده و یا کسر گردیده است. هدف اصلی این کتاب، ایجاد یک منبع فارسی مناسب برای آشنایی مقدماتی کاربران فارسی‌زبان با زبان برنامه‌نویسی C و متمند و منعطف R می‌باشد که در کشور ما کمتر مورد توجه قرار گرفته است. این زبان برنامه‌نویسی، به ویژه برای دانشجویان مقاطع تحصیلات تکمیلی و پژوهشگران حوزه‌های مختلف، ابزار قدرتمندی برای تحلیل داده‌ها و انجام تحلیل‌های متنوع آماری محسوب می‌گردد. مفاهیم پیچیده‌تر مثل ترجمه گراف‌های حرفه‌ای یا داده‌کاوی با استفاده از R در کتاب‌های بعدی مورد بحث قرار خواهند گرفت.

این کتاب، حاصل تجمیع و به روز رسانی تمامی مطالبی است که در سال ۱۹۹۰ توسط David M. Smith و Bill Venables گردآوری شدند و به تشریح زبان‌های برنامه‌نویسی S و S-Plus می‌پردازند. برخی تغییرات نسبت به نسخه‌های قدیمی‌تر کتاب اعمال شده تا تغییرات بین برنامه‌نویسی S و R مدنظر قرار گیرند و برخی مطالب نیز مورد غنای بیشتری شرح شده‌اند.

اغلب کاربران جدید R این کتاب را از فصل پانزدهم («یک جلسه آموشی») آغاز می‌کنند. این بخش، آشنایی لازم با فرم جلسات R¹ و از آن مهم‌تر برخی بازخوردها، سبب در مورد آنچه واقعاً در R اتفاق می‌افتد را فراهم می‌سازد. بسیاری از کاربران هم از R اساساً بعنوان یک ابزار گرافیکی استفاده می‌کنند. این کاربران می‌توانند به فصل ۱۲ مراجع کنند و نیازی به خواندن بقیه فصول ندارند.

¹ Sessions