

به نام خدا



آموزش زبان برنامه نویسی R

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 و صفحه آرای: شبنم هاشم زاده

طرح روی جلد: داریوش فرسای

چاپ: درج عقیق

نوبت چاپ: سوم

تاریخ نشر: ۱۳۹۸

تیراژ: ۵۰ جلد

قیمت: ۵۲۰۰۰۰ ریال

شابک: ۹۷۸-۶۰۰-۱۲۴-۵۸۸-۶

نشانی واحد فروش: تهران، میدان انقلاب،

خ کارگر جنوبی، روبروی پاساژ مهستان،

پلاک ۱۲۵۱

تلفن: ۲۲۰۸۵۱۱۱-۶۶۴۱۰۰۴۶

کد پستی: ۱۳۱۴۹۸۳۱۸۵

فروش اینترنتی:

www.mftbook.ir

www.dibagarantehran.com

نشانی تلگرام: @mftbook

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: فیبا

موضوع: آر(زبان برنامه نویسی کامپیوتر)

موضوع: R(computer program language)

موضوع: آمار - داده پرداز

موضوع: Statistics-Data processing

رده بندی کنگره: ۱۳۹۶ آ۴۲/۴۵/۲۷۶ QA

رده بندی دیویی: ۵۱۹/۵۰۲۸۵۵۱۳۳

شماره کتابشناسی ملی: ۶۶۷۹۴۴۴

نشانی اینستاگرام: Dibagaran_publishing

فهرست مطالب

فصل 1: معرفی و مقدمات 13

- 13 1-1 محیط R
- 14 2-1 کاربردهای R در مطالعات آماری
- 14 3-1 استفاده از R به صورت تعاملی
- 16 4-1 یک جلسه آموزشی
- 16 5-1 راهنمای R برای کار با توابع
- 17 6-1 دستورهایی R و حساسیت به اندازه حروف
- 17 7-1 فراخوانی و اصلاح دستورهایی قبلی
- 18 8-1 اجرای دستورها از یا انتقال خروجی‌ها به یک فایل
- 18 9-1 دوام داده‌ها و حذف اشیاء

فصل 2: تغییرات ساده؛ اعداد و بردارها 21

- 21 1-2 بردارها و نسبت‌دهی
- 22 2-2 محاسبات بر روی بردارها
- 24 3-2 تولید رشته‌های عددی
- 25 4-2 بردارهای منطقی
- 26 5-2 مقادیر مفقود

- 26 6-2 بردارهای کاراکتری
- 27 7-2 انتخاب و تغییر بخشی از یک مجموعه داده
- 29 8-2 انواع دیگر اشیاء

31 فصل 3: اشیاء، مُدها و صفات کیفی

- 31 1-3 صفات کیفی ذاتی: mode و length
- 32 2-3 تغییر طول اشیاء
- 33 3-3 مشاهده و تنظیم کردن صفات
- 34 4-3 کلاس یک شیء

35 فصل 4: فاکتورهای مرتب‌شده و مرتب‌نشده

- 35 1-4 یک مثال خاص
- 36 2-4 تابع apply() و آرایه‌های ناصاف
- 38 3-4 فاکتورهای مرتب‌شده

39 فصل 5: آرایه‌ها و ماتریس‌ها

- 39 1-5 آرایه‌ها
- 40 2-5 فهرست‌سازی از آرایه‌ها: بخش‌هایی از آرایه
- 40 3-5 ماتریس اندیس‌ها
- 42 4-5 تابع array()
- 43 1-4-5 بردارهای مخلوط و آرایه‌های ریاضی: قاعدهٔ بازچرخش
- 44 5-5 حاصل ضرب خارجی دو آرایه
- 45 6-5 محاسبهٔ ترانهادهٔ یک آرایه

45	7-5 کار با ماتریس‌ها
45	1-7-5 ضرب ماتریس‌ها
46	2-7-5 معادلات خطی و تعادل‌ها
47	3-7-5 مقادیر مشخصه و بردارهای مشخصه
47	4-7-5 تجزیه مقدار تکین و دترمینان‌ها
48	5-7-5 برازش منحنی‌ها به روش حداقل مربعات و تجزیه QR
49	8-5 توابع <code>rbind()</code> و <code>cbind()</code>
50	9-5 کاربرد تابع <code>c()</code> بر روی آراییه‌ها
50	10-5 جداول فراوانی از فاکتورها

فصل 6: لیست‌ها و دیتافریم‌ها

53	1-6 لیست‌ها
54	2-6 ایجاد و تغییر لیست‌ها
55	1-2-6 الحاق لیست‌ها
55	3-6 دیتافریم‌ها
56	1-3-6 ایجاد دیتافریم‌ها
56	2-3-6 توابع <code>attach()</code> و <code>detach()</code>
57	3-3-6 کار با دیتافریم‌ها
58	4-3-6 الصاق لیست‌ها به آدرس جستجو
58	5-3-6 مدیریت مسیر جستجو

فصل 7: خواندن اطلاعات از فایل‌ها

61	1-7 تابع <code>read.table()</code>
62	2-7 تابع <code>scan()</code>

- 63 3-7 دسترسی به مجموعه داده‌های داخلی R
- 64 4-7 بارگذاری داده‌ها از سایر بسته‌های R
- 64 5-7 ویرایش داده‌ها

67 فصل 8: توزیع‌های احتمالاتی

- 67 1-8 مجموعه‌ای از جداول آماری
- 68 2-8 بررسی توزیع احتمالاتی داده‌ها
- 74 3-8 آزمون‌های یک و دو نمونه‌ای

79 فصل 9: گروه‌بندی، حلقه‌ها و دستورهای شرطی

- 79 1-9 عبارات گروه‌بندی شده
- 79 2-9 گزاره‌های کنترلی
- 79 1-2-9 گزاره‌های شرطی: if
- 80 2-2-9 ساختارهای تکراری: for, repeat و while

83 فصل 10: نوشتن توابع شخصی

- 83 1-10 مثال‌های ساده
- 85 2-10 ایجاد عملگرهای جفتی جدید
- 85 3-10 آرگومان‌های نامگذاری شده و مقادیر پیش فرض
- 86 4-10 کاربرد آرگومان "..."
- 87 5-10 نسبت‌دهی درون توابع
- 87 6-10 مثال‌های پیشرفته‌تر
- 87 1-6-10 فاکتورهای کارایی در طرح‌های بلوکی

- 88 2-6-10 حذف تمامی نام‌ها از یک آرایه پرینت شده
- 89 7-10 قلمرو
- 92 8-10 شخصی‌سازی یک محیط
- 93 9-10 کلاس‌ها، توابع عمومی و شیء‌گرایی

97 فصل 11: مدل‌های آماری

- 97 1-11 تعریف مدل‌های آماری: فرمول‌ها
- 98 1-1-11 مثال‌ها
- 100 2-11 مدل‌های خطی
- 100 3-11 توابع عمومی برای استخراج اطلاعات مدل‌های آماری
- 102 4-11 تحلیل واریانس و مقایسه مدل‌ها
- 103 1-4-11 جدول‌های ANOVA
- 103 5-11 به‌روزرسانی مدل‌های برازش داده‌شده
- 104 6-11 مدل‌های خطی تعمیم‌یافته
- 105 1-6-11 خانواده‌ها
- 106 2-6-11 تابع $glm()$
- 106 1-2-6-11 خانواده مدل‌های گوسی
- 107 2-2-6-11 خانواده مدل‌های دو جمله‌ای
- 109 3-2-6-11 خانواده مدل‌های پواسون
- 109 7-11 مدل‌های غیرخطی حداقل مربعات و حداکثر درست‌نمایی
- 109 1-7-11 حداقل مربعات
- 112 2-7-11 حداکثر درست‌نمایی
- 112 8-11 برخی مدل‌های غیراستاندارد

- 116 1-12 دستورهای گرافیکی سطح بالا
- 116 1-1-12 تابع `plot()`
- 117 2-1-12 نمایش داده‌های چندمتغیره
- 118 3-1-12 نمایش اشیاء گرافیکی
- 119 4-1-12 آرگومان‌های توابع گرافیکی سطح بالا
- 120 2-12 دستورهای گرافیکی سطح پایین
- 121 1-2-12 نشانه‌گذاری ریاضی
- 122 2-2-12 فونت‌های برداری هِرشی
- 122 3-12 تعامل با گراف‌ها
- 124 4-12 کاربرد پارامترهای گرافیکی
- 124 1-4-12 تغییرات دائمی: تابع `par()`
- 125 2-4-12 تغییرات موقتی: آرگومان‌های توابع گرافیکی
- 126 5-12 لیست پارامترهای گرافیکی
- 126 1-5-12 اجزاء گرافیکی
- 127 2-5-12 محورها و تیک‌ها
- 128 3-5-12 حاشیه‌ها
- 129 4-5-12 استفاده همزمان از چند تصویر
- 131 6-12 برنامه‌های راه‌اندازی ابزارهای گرافیکی
- 132 1-6-12 استفاده از فایل‌های گرافیکی PostScript
- 133 2-6-12 استفاده همزمان از چند ابزار گرافیکی
- 134 7-12 گراف‌های پویا

135 **فصل 13: بسته‌ها**

136 1-13 بسته‌های استاندارد

136 2-13 بسته‌های توسعه‌دهندگان و منبع CRAN

136 3-13 فضاهای نام‌گذاری

139 **فصل 14: کار با سیستم‌های عامل**

139 1-14 فایل‌ها و پوشه‌ها

140 2-14 آدرس‌دهی فایل‌ها

141 3-14 دستورهای سیستم

142 4-14 فشردن‌سازی و آرشیوها

143 **فصل 15: یک جلسه آموزشی**

خط مشی کیفیت انتشارات مؤسسه فرهنگی هنری دیباگران تهران در عرصه کتاب‌هایی است که بتواند خواسته‌هایی 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 گام‌هایی هرچند کوچک برداشته و در انجام رسالتی که بر عهده داریم، مؤثر واقع شویم.

گسترده‌گی علوم و توسعه روزافزون 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 قدیمی‌ترین و راحت‌ترین راه دستیابی به اطلاعات و اطلاع‌رسانی، بیش از پیش روشن می‌نماید.

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 تلاش‌های مستمر خود برای رفع کمبودها و نیازهای موجود، منابعی پُر بار، معتبر و با کیفیت مناسب در اختیار علاقمندان قرار دهند.

کتابی که در دست دارید با همت "دکتر بهنام آبابایی" 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 دانش‌پژوه گرامی درخواست می‌نماید 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 می‌داند، 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 بوده، 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 قدرتمند 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 ترسیم 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 1990 توسط David M. Smith و Bill Venables نگاشته 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 بطور مفصل‌تر تشریح شده‌اند. اغلب کاربران جدید R این کتاب را از فصل پانزدهم («یک جلسه آموزشی») آغاز می‌کنند. این بخش، آشنایی لازم با فرم جلسات¹ R و از آن مهم‌تر برخی بازخوردهای سریع 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 12 مراجع کنند و نیازی به خواندن بقیه فصول ندارند.

¹ Sessions