

# مسیر داده‌کاوی در پایتون: یادگیری کتابخانه نامپای

تألیف

دکتر عبدالرسول مستاجران  
عضو هیئت علمی دانشگاه خوارزمی

دکتر محسن بهلولی زفره  
دکتر سپاهه محمدی



دانشگاه خوارزمی

تهران ۱۴۰۳

|                         |                                                                                                              |
|-------------------------|--------------------------------------------------------------------------------------------------------------|
| سرشناسه                 | : مستاجران، عبدالرسول، ۱۳۶۱-                                                                                 |
| عنوان و نام پدیدآور     | : مسير داده کاوی در پایتون: یادگیری کتابخانه نامپای/تالیف عبدالرسول مستاجران، محسن بهلولی زفره، سهامه محمدی. |
| مشخصات نشر              | : تهران: دانشگاه خوارزمی، انتشارات، ۱۴۰۳.                                                                    |
| مشخصات ظاهری            | : ۲۷۲ ص.                                                                                                     |
| شابک                    | : ۹۷۸-۶۲۲-۴۸۱۵-۲۷-۹ ریا ۲۸۰۰۰۰۰                                                                              |
| وضعیت فهرست نویسی       | : فیا                                                                                                        |
| موضوع                   | : پایتون (زبان برنامه نویسی کامپیوتر) Python (Computer program language)<br>داده کاوی Data mining            |
| شناسه افزوده            | : بهلولی زفره، محسن، ۱۳۶۳- . محمدی، سهامه، ۱۳۶۶- . دانشگاه خوارزمی. انتشارات.                                |
| رده بندی کنگره          | : ۷۳/QA۷۶                                                                                                    |
| رده بندی دیویی          | : ۱۳۳/۰۰۵                                                                                                    |
| شماره کتابشناسی ملی     | : ۱۹۲۴۸۱۵                                                                                                    |
| اطلاعات رکورد کتابشناسی | : فیا                                                                                                        |

www.ketab.ir



|                |                                                     |
|----------------|-----------------------------------------------------|
| عنوان کتاب     | : مسير داده کاوی در پایتون: یادگیری کتابخانه نامپای |
| تألیف          | : عبدالرسول مستاجران، محسن بهلولی زفره، سهامه محمدی |
| ناشر           | : دانشگاه خوارزمی                                   |
| چاپ و صحافی    | : دانشگاه خوارزمی                                   |
| طراح جلد       | : فاطمه منظور                                       |
| نوبت و سال چاپ | : اول، بهمن ۱۴۰۳                                    |
| شابک           | : ۹۷۸-۶۲۲-۴۸۱۵-۲۷-۹                                 |
| شمار           | : ۵۰۰ نسخه                                          |
| قیمت           | : ۲۸۰۰۰۰۰ ریا                                       |

کلیه حقوق مادی و معنوی این اثر متعلق به انتشارات دانشگاه خوارزمی است.  
آدرس: تهران، خ شهید مفتاح، شماره ۴۳، کدپستی ۱۴۹۱۱-۱۵۷۱۹ تلفن مرکز پخش: ۸۸۳۱۱۸۶۶  
www.khu.ac.ir pub@khu.ac.ir

# پیش‌گفتار

در دنیای امروز که داده‌ها به طور فزاینده‌ای اهمیت پیدا می‌کنند، علم داده و داده‌کاوی به زمینه‌هایی تبدیل شده‌اند که محققان و علاقمندان از رشته‌های مختلف به آن گرایش پیدا کرده‌اند. این زمینه‌ها به ما امکان می‌دهند تا بینش‌های ارزشمندی از مجموعه داده‌های بزرگ استخراج کنیم، که منجر به تصمیم‌گیری‌های آگاهانه‌تر، نوآوری و پیشرفت در رشته‌های مختلف می‌شود.

در تمامی نقشه‌راه‌های یادگیری داده‌کاوی و علم داده یکی از مهارت‌های لازم تسلط بر برنامه‌نویسی است که پر استفاده‌ترین زبان برنامه‌نویسی در این زمینه زبان پایتون است. این زبان امروزه به طور وسیعی در زمینه‌های مختلف از جمله داده‌کاوی، یادگیری عمیق، هوش مصنوعی، تحلیل داده‌های بورس و بازارهای مالی و ... مورد استفاده قرار می‌گیرد. علاوه بر زبان اصلی پایتون، کتابخانه‌های بنیادی در این زبان وجود دارد که تسلط بر آنها نیازمند هر فرد فعال در زمینه داده است. برخی از این کتابخانه‌های بنیادی شامل نامپای، پانداس، Matplotlib، scikit-learn و ... هستند. کتاب‌های فارسی قابل توجهی در آموزش زبان پایتون وجود دارند. اما در خصوص کتابخانه‌های پایتون منبع به زبان فارسی آنچنانی وجود ندارد. هدف نویسندگان فراهم کردن یک مسیر یادگیری از این کتابخانه‌هاست که با کتابخانه نامپای شروع می‌شود. این کتابخانه برای ایجاد و کار با آرایه‌ها استفاده می‌شود. ویژگی مهم آرایه‌های نامپای سرعت بسیار بالاتر عملیات بر روی آنها نسبت به دنباله‌های پایتون است. ضمن اینکه بسیاری دیگر از کتابخانه‌ها از نامپای استفاده می‌کنند.

در تالیف یک کتاب همواره دو نکته مورد توجه است. نکته اول اینکه با توجه به مهم بودن همه مطالب، مولف علاقمند است مطالب بیشتری را در کتاب قرار دهد و نکته دوم اینکه در تدوین یک کتاب خوب بایستی سطح کتاب و کاربران در نظر گرفته شود. بر این اساس سعی

شده است مطالب ضروری و مفید در کتاب گنجانده شود. بسیاری از منابع موجود در این زمینه، طیف وسیعی از مباحث کتابخانه نامپای را ارایه نکرده یا بطور ناقص و پراکنده بیان کرده‌اند، به همین دلیل در این کتاب مفاهیم اساسی نامپای با ارایه مثال‌های متنوع تشریح می‌شود. سعی شده است این کتاب به عنوان یک راهنمای جامع جهت یادگیری نامپای برای علم داده و داده‌کاوی طراحی شود. فصل‌های این کتاب شامل پیش‌نیازها، مفاهیم اولیه نامپای در ایجاد آرایه‌ها و ویژگی‌های آنها، اندیس‌گذاری و برش آرایه‌ها، عملیات‌های پردازش بر روی آرایه‌ها، توابع ریاضی و جبرخطی و آمار خواهد بود. اگرچه پیش‌نیاز مطالب ارائه شده در این کتاب آشنایی با مبانی برنامه نویسی، جبرخطی و نظریه ماتریس‌ها و مفاهیم آماری و ریاضی است، با این حال سعی شده است به طور مختصر مفاهیم مورد نیاز شرح داده شود. امید است خوانندگان گرامی، نظرها و پیشنهادها را از طریق ایمیل [datascience.farsibooks@gmail.com](mailto:datascience.farsibooks@gmail.com) با ما در میان گذاشته تا در چاپ‌های بعدی موجب غنی‌تر شدن این کتاب و کتاب‌های بعدی مسير داده‌کاوی گردد.

نویسندگان

تهران، زمستان ۱۴۰۳

# فهرست مطالب

|    |                                        |
|----|----------------------------------------|
| آ  | پیش گفتار                              |
| ۱  | ۱ پیش نیازها                           |
| ۱  | ۱.۱ نامپای چیست؟                       |
| ۴  | ۲.۱ نصب نامپای و محیط‌های توسعه پایتون |
| ۷  | ۳.۱ مروری بر دستورات پایتون            |
| ۷  | ۱.۳.۱ قواعد زبان پایتون                |
| ۱۱ | ۲.۳.۱ انواع داده و متغیرها در پایتون   |
| ۱۳ | ۳.۳.۱ عملگرهای پایتون                  |
| ۱۷ | ۴.۳.۱ گزاره‌های شرطی                   |
| ۲۰ | ۵.۳.۱ گزاره‌های تکرار                  |
| ۲۳ | ۶.۳.۱ تعریف تابع                       |
| ۲۴ | ۷.۳.۱ مباحث پیشرفته‌تر پایتون          |
| ۲۴ | تمرین‌ها                               |
| ۲۷ | ۲ مفاهیم اولیه نامپای                  |
| ۲۷ | ۱.۲ انواع داده در نامپای               |
| ۳۱ | ۲.۲ تبدیل نوع داده در نامپای           |
| ۳۴ | ۳.۲ ایجاد آرایه نامپای                 |
| ۳۴ | ۱.۳.۲ ایجاد آرایه دلخواه               |
| ۳۸ | ۲.۳.۲ ایجاد آرایه‌های خاص              |

|       |                                         |     |
|-------|-----------------------------------------|-----|
| ۳.۳.۶ | مثال در تحلیل داده‌های بازاریابی و فروش | ۲۵۰ |
| ۴.۳.۶ | مثال در تحلیل داده‌های حقوقی            | ۲۵۲ |
| ۵.۳.۶ | مثال در پردازش تصویر                    | ۲۵۳ |
|       | تمرین‌ها                                | ۲۵۶ |

منابع

۲۵۹

واژه‌نامه

۲۶۱