

Cezar Valentin Tomescu: Biodiversitatea florei și vegetației ecosistemelor natural din bazinul râului Suceava. Casa Cărții de știință, Cluj Napoca, 2016, 604 p.



Profesorul Victor Stănescu obișnuia să spună, la prima oră de curs de dendrologie, că la disciplina respectivă ne vom cunoaște circa o sută de noi colegi, care nu ne vor trăda niciodată. Totuși, de fiecare dată când vrem să vorbim la modul serios de biodiversitatea vegetală, ne lovim de cel puțin două praguri: 1) nu cunoaștem speciile și 2) limbajul foarte standardizat care ne este furnizată informația, accesibilă unui grup restrâns de specialiști. Faptul că o carte se adresează unui grup restrâns de oameni nu o face mai puțin valoroasă, comparativ cu lucrările ce se bucură de o largă audiență. Doar că promovarea acestui gen de literatură este mai dificilă, deoarece cartea în sine nu este totdeauna un mediator eficient între planta sau insecta în cauză și cel ce vrea să o cunoască: ne-am simți mult mai siguri dacă ne-ar lua

cineva de mână și ne-ar spune „poftim, asta este!”. Cartea colegului nostru Cezar Tomescu vine în întâmpinarea dorinței și nevoii de a carta din punct de vedere naturalistic bazinul Sucevei, temă căreia îi sunt dedicate ultimele două secțiuni, prima fiind caracterizarea fizico-geografică a zonei. Avem, în sfârșit, o sinteză a cercetărilor floristice din zonă, cercetări ce au început la finele secolului XVIII.

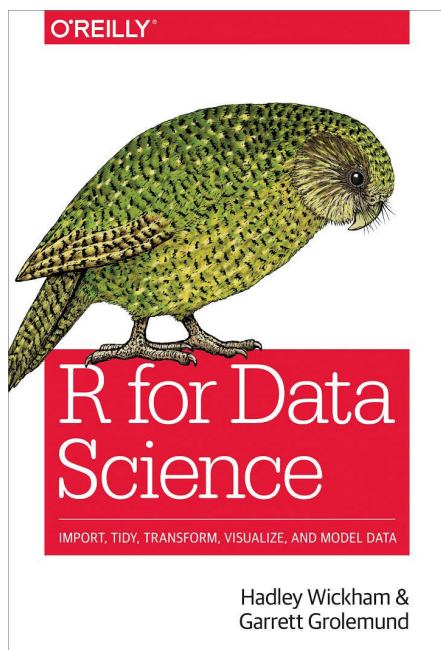
Conspectul floristic al fiecărei încrengături este detaliat la nivel de clasă, ordin, familie și specie; pentru fiecare specie autorul oferă referințele bibliografice de bază, precum și locațiile în care specia respectivă poate fi găsită. Istoricul cercetărilor fitocenologice este detaliat la începutul celei de-al treilea capitol, din care aflăm că primul studiu fitosociologic a fost realizat în 1902, odată cu prima hartă geobotanică a țării. În spiritul aceleiași rigori științifice, autorul prezintă pe rând fiecare asociație relevantă de specii, în raport cu răspândirea, condițiile staționale, compoziția floristică și fitocenotică, spectrul bioformelor și spectrul indicilor ecologici. Nu știu dacă și de unde s-a inspirat autorul, dar acest modul în care a organizat informația dovedește o deosebită capacitate de sinteză, trăsătură comună marilor profesori de botanică.

Marian Drăgoi

Wickham, H., Golemund, G., R for Data Science: Import, Tidy, Transform, Visualize, and Model data. O'Reilly Media, Inc., 2016, 522 p. ISBN 978-1491910399.

Cartea “R pentru știința datelor: import, ordonare, transformare, vizualizare și modelarea datelor” actualizează modalitățile de abordare a prelucrării statistice a datelor cu limbajul statistic R, utilizând în acest sens experiența celor doi autori, membri ai echipei de dezvoltare a RStudio (principala interfață). Cartea conține 25 de capitole, structurate în șase părți și ilus-

trate prin titluri bazate pe verbe (“A explora”, “A transforma”, “A programa”, “A modela”, “A comunica”) care, conform autorilor, doresc să sugereze o abordare bazată pe acțiune.



Prima parte debutează cu o prezentare a bibliotecii `ggplot2` pentru vizualizarea datelor (dezvoltată de primul autor), urmată de un cod de bune practici cu privire la notațiile și stilul de programare. Modul în care trebuie desfășurată o prelucrare adecvată a datelor este prezentat în capitolul dedicat bibliotecii `dplyr` (același autor de mai sus), dedicată unei prelucrări mai facile a datelor de tip “data frame” (tabele de variabile de diferite tipuri) și care aduce îmbunătățiri substanțiale modului de prelucrare a datelor. Aici este introdusă și utilizarea “tuburilor” (*pipe*), utile pentru fluidizarea și compactarea codului R. Analiza exploratorie a datelor propune diferite metode de vizualizare a variabilelor continue sau categoricale cu `ggplot2`. Ultimul capitol se concentrează pe descrierea fluxului de lucru și a caracteristicilor mediului de programare RStudio.

Partea a doua este dedicată procesării (pregătirii) datelor în vederea prelucrării statistice. Este prezentată biblioteca `tibble` (totodată și nume al noii generații de “data frame”-uri), un important pas înainte în prelucrarea acestora. Următorul capitol este dedicat importului datelor în mediul R, prin intermediul bibliotecii `readr`, o alternativă de cca. 10 ori mai rapidă decât abordarea clasică. Transformarea unui set de date se poate realiza cu ușurință cu ajutorul bibliotecii `tidyr`, cea care livrează formate adecvate în vederea prelucrării ulterioare sau a vizualizării. Utilizarea informațiilor relaționale după principiul bazelor de date, cu un limbaj asemănător cu SQL, însă ceva mai simplu, este prezentat în capitolul următor, folosind `dplyr`; sunt prezentate în mod detaliat și însoțite de exemple clare diferitele tipuri de asociere (*join*) dintre date. O nouă interpretare a lucrului cu șiruri de caractere este prezentată prin intermediul bibliotecii `stringr`. Operațiile generale cu factori, de asemenea simplificate, sunt prezentate cu `forcats`, în timp ce datele temporale sunt abordate cu `lubridate`.

Partea a treia, dedicată programării, se întoarce la avantajele utilizării tuburilor (*pipe*) în programare cu biblioteca `magrittr`, argumentând simplificarea codului cu exemple concrete. O parte importantă este dedicată funcțiilor, elemente cheie în utilizarea R. Puțin surprinzător pentru cărțile ce tratează utilizarea R, următorul capitol este dedicat unei prezentări exhaustive a datelor de tip vector, urmate de o serie de exemple cu privire la funcții, inclusiv metode inovative, precum înlocuitori pentru unele situații ale funcției *for*.

Partea a patra este considerată de către autori ca o reîntoarcere la modelarea datelor, după partea de pregătire a datelor prezentată anterior. Primele două capitole ale acesteia sunt dedicate prezentării modelr și a modului de utilizare a unor modele liniare simple, în timp ce ultimele două abordează modele bazate pe seturi de date mai complexe, autorii insistând pe instrumentele de vizualizare pentru o diagnoză a rezultatelor acestora.

Ultima parte a cărții (“A comunica”) este de-

dicată modului de prezentare finală a rezultatelor cu ajutorul limbajului R Markdown. Principiile de bază, modul în care se poate formata textul sau tabelele, depanarea unui document R Markdown, sau citarea bibliografiilor sunt prezentate în capitolul 25 (primul al acestei părți), căruia îi urmează o prezentare detaliată a ggplot2. Formatele finale ale R Markdown (documente, cărți, prezentări, pagini Web interactive) sunt prezentate în capitolul următor, ultimul capitol dând indicații asupra fluxului de lucru.

Venind dinspre un specialist în prezentare grafică (H.W.), cartea este redactată într-un limbaj accesibil, însoțit de numeroase exemple de cod și incluzând frecvent cu exerciții care să fixeze cunoștințele expuse, fiind prezentată într-un mod atractiv. În ciuda numărului mare de pagini, parcurgerea sa este una ușoară.

Cartea aduce la zi diferite aspecte referitoare la analiza datelor în R (importul și pre-procesarea datelor, elemente de stilistică a codului, integrarea codului cu textul în R Markdown pentru compilarea unui produs final), în general mai bune decât versiunile de bază existente.

Discuția comparativă cu alte biblioteci rămâne deschisă în ceea ce privește viteza de procesare (se pare, compensată cu versiunea ce include calcul paralel), însă nu și la compilarea bibliotecilor cu cod ce include *pipes*.

La nivel general, este vorba o schimbare de paradigmă - începută de ceva vreme - pe care o prezintă unul din membrii marcanți ai comunității R, cu o cunoaștere foarte bună a limbajului și a aplicației și, în același timp, și implicat în dezvoltarea principalei interfețe. Pentru utilizatorii care nu încapsulează frecvent în biblioteci codul utilizat, un upgrade de limbaj și instrumente poate îmbunătăți semnificativ productivitatea (atât prin simplificarea codului, cât și printr-o lizibilitate mai bună și, implicit, posibilități mai bune de depanare). Per total, cartea merită parcursă pentru că furnizează, într-un singur loc, o multitudine de informații și abordări noi cu privire la programarea în R, rămânând la latitudinea cititorului să aleagă, dintre soluțiile propuse, pe cele mai convenabile.

Marius Teodosiu