

נא להכיר: שפת התכנות R

יובל נוב, החוג לסטטיסטיקה, אוניברסיטת חיפה

האם גם לכם נמאס מהמגבלות ומהתקלות של אקסל? גם אתם לא רוצים לשלם סכומי עתק עבור תוכנות סטטיסטיות מסחריות כגון SAS? ובכן, יש אלטרנטיבה אשר צוברת תאוצה בעולם ובארץ. זוהי שפת התכנות R – תוכנה מקצועית, חינוכית ומודולרית, עם דגש מיוחד על סטטיסטיקה והסתברות.



הגירסה הראשונה של שפת R נוצרה בשנת 1993 בידי רוס איהקה (Ross Ihaka) ורוברט ג'נטלמן (Robert Gentleman), שני סטטיסטיקאים צעירים מאוניברסיטת אוקלנד בניו-זילנד. כבר מראשיתה, השפה אימצה את התחביר של שפת S, שהיא שפת תכנות סטטיסטית מסחרית שפותחה בשנות השבעים במעבדות חברת "בל" בארצות-הברית. השפה התפתחה במהירות, ותוך שנים ספורות זכתה להצלחה פנומנלית – בהתחלה בעיקר בעולם האקדמי, ולאחרונה יותר ויותר בתעשייה ובעולם העסקי.

איהקה וג'נטלמן, צמד המפתחים המקוריים של השפה, הם כיום רק שניים מבין עשרים חברי צוות הליבה של פיתוח R, המקדישים מזמנם ומניסיונם לשיפורה. עוזרים להם מאות מתנדבים מרחבי העולם אשר תורמים קוד, מדווחים על באגים (לעיתים אף מתקנים אותם), כותבים תיעוד, ומציעים הצעות לשיפור. סביב השפה צמחה קהילת משתמשים תוססת, שחבריה משתתפים בכנסים בין-לאומיים ייעודיים ל-R, מארגנים סדנאות הדרכה ללימוד השפה, ומתחזקים [אתרי אינטרנט ובלוגים](#) המוקדשים ל-R. הערכות שונות מדברות על כמה מיליוני משתמשים של R כיום, ומספרם גדל במהירות.

אחת הסיבות המרכזיות לפופולריות של R היא היותה שפה לא-מסחרית, המופצת בחינם לכל המעוניין. יותר ויותר חברות מבינות את פוטנציאל החיסכון הכספי האדיר הטמון ב-R, ומשלבות את השפה בתהליכי העבודה ובמוצרים שלהן. בין חברות אלה ניתן למנות חברות היי-טק כגון גוגל ואורקל, חברות תרופות כגון פייזר ומרק, ומוסדות פיננסיים כגון בנק אוף אמריקה. בסקר שנתי שעורך האתר kdnuggets.com (אחד מהאתרים העסקיים הגדולים ביותר של קהילת ניתוח הנתונים), שפת R תופסת באופן קבוע בשנים האחרונות את אחד משני המקומות הראשונים בקטגוריית התוכנה המועדפת לכריית נתונים.

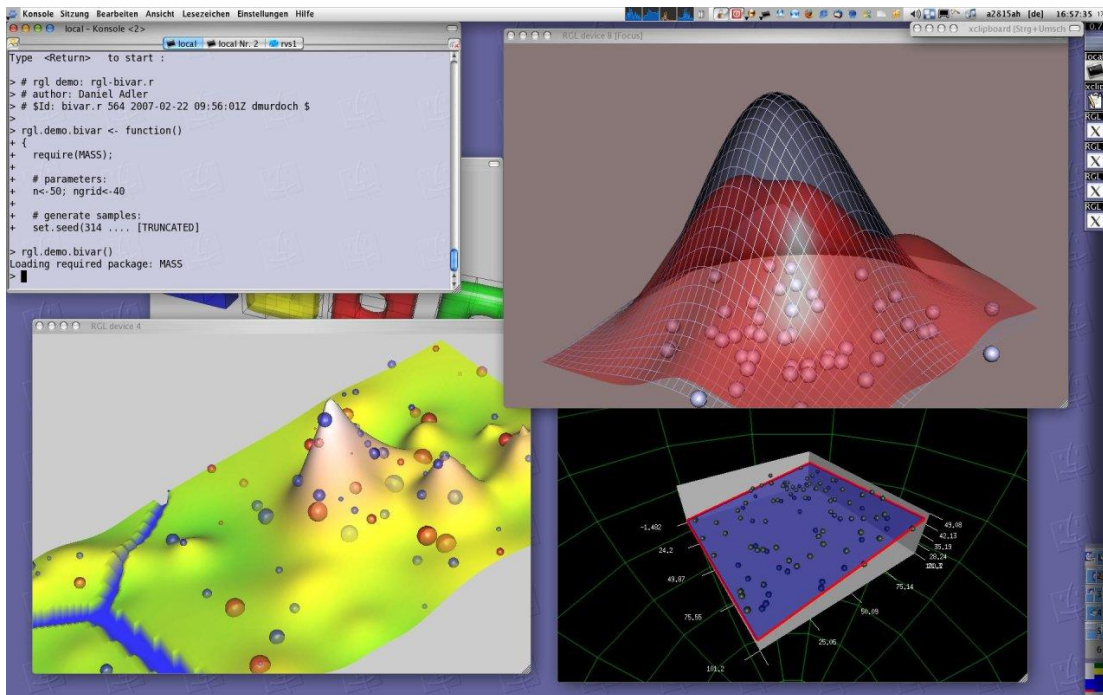
סיבה חשובה נוספת לפופולריות של השפה היא המודולריות והגמישות שלה. כל מי שפיתח שיטה חישובית מתמטית או סטטיסטית חדשה, יכול לתכנת אותה ב-R ואז להפיץ אותה ברבים בתור "חבילה" (הנקראת לעתים גם "ספריה"). משתמשים אחרים יכולים אז לטעון בקלות את החבילה הזו למחשבים שלהם, דרך מנגנון פשוט המובנה בתוכנה. [באתר המרכזי של R](#) ובאתרים אחרים נמצאות למעלה מ-5,000 חבילות שכאלה הניתנות להורדה בחינם, במגוון אדיר של תחומים: החל ממיקרו כלכלה, דרך חקר הקרקע, זיהוי דיבור, אופטימיזציה וגנטיקה, וכלה באסטרונומיה.

מעבר למנגנון החבילות, מתכנתים יכולים למזג את שפת R עם תוכנות אחרות בזכות העובדה ש-R היא שפת "קוד פתוח": קוד המקור שלה מפורסם בחינם ברבים, וכל מי שחפץ בכך יכול להעתיק אותו, לשנות אותו, לבנות עליו יישומים חדשים, ולהפיץ את התוצאה בכל דרך שירצה. כך למשל, חברות כגון Rstudio ו-Rcommander יצרו באופן זה "מעטפות" ידידותיות יותר ל-R, וחברת Revolution analytics פיתחה תוכנה מבוססת R המחקה את תוכנת SAS, אך עולה כמחצית המחיר.

בעולם המחקר האקדמי הסטטיסטי R הפכה מזמן לכלי העבודה החישובי הנפוץ והחשוב ביותר. טובי האוניברסיטאות בעולם (למשל הרווארד, סטנפורד וברקלי) מציעות קורסים בתכנות ב-R, לא רק במחלקות לסטטיסטיקה, אלא גם במחלקות למדעי המחשב ולהנדסה. גם באוניברסיטאות בארץ השפה צוברת תאוצה, וקורס תכנות ב-R הוא חובה בחלק מתוכניות הלימודים לתואר ראשון בסטטיסטיקה.

מבחינת סגנון התכנות, R דומה לשפות תכנות מתמטיות כגון Matlab, Maple ו-Mathematica. מצד אחד, זוהי שפת תכנות מלאה עם לולאות, תנאים, פונקציות, רקורסיה, וכדומה; אך מצד שני R היא פשוטה ונגישה יותר משפות כגון C, C++ או ג'אווה: לדוגמה, ניתן בקלות לעבוד עם וקטורים, מטריצות וטבלאות נתונים, ואין צורך להצהיר על משתנים או להקצות ולשחרר זיכרון.

ל-R יש יכולות גרפיות מתקדמות, העולות בהרבה על אלה של אקסל. ראו למשל את הדוגמה הבאה:



שפת R הולכת ותופסת מקום גם בעולם האקטואריה. הגירסה הבסיסית של R מצוידת כפי שהיא במגוון של פונקציות שימושיות לאקטוארים; בנוסף על כך, קבוצה של חוקרים קנדיים יצרו את [חבילת actuar](#), הכוללת כלים חישוביים וקבצי נתונים המיועדים במיוחד לעבודה באקטואריה. המגזין המקוון CompAct של ה-SOA (אגודת האקטוארים בארצות-הברית) מפרסם בקביעות טור בשם "[The R Corner](#)", המוקדש לשימושי R באקטואריה.

מי שמתעניין ב-R מוזמן להתקין את התוכנה על כל מחשב שרצה, לאחר שיריד אותה מ-[cran.r-project.org](#). האינטרנט מלא בחומר הדרכה (באנגלית) על R: [סרטונים ביוטיוב](#) (הנה למשל [סרטון המתמקד באקטואריה](#)), [ספרי הדרכה](#), [פורומים](#), ועוד. [האיגוד הישראלי לסטטיסטי](#) מארגן מפעם לפעם סדנאות להכרת R, וחומר לימודי בעברית על R ניתן למצוא [באתר הקורס](#) המתאים באוניברסיטת חיפה, מאת כותב שורות אלה.

בהצלחה!

נכתב במקור עבור "האקטואר" – עיתון אגודת האקטוארים בישראל, ספטמבר 2013