

תרגול 12 – Analysis of Variance

תרגול זה עוסק בקובץ הנתונים "MultiWeightliftingScores.csv".

בקובץ זה מופיעים ביצועי נבחרות סין, ארה"ב ורוסיה (הדימוניים) בתחרות הרמת משקולות.

עליכם למצוא האם ממוצעי נבחרות מדינות אלו (היינו תוחלות ביצועי הנבחרות) שונים זה מזה או קרובים זה לזה.

לצורך כך:

1. העלו את הקובץ לטבלת נתונים ב-R (dataframe).
2. נסחו השערה סטטיסטית מתאימה לשאלה האמורה למעלה.
3. הציגו את אחוזוני התפלגות הנתונים בכל קבוצה בתצוגת boxplot. האם אתם עדים למופעים חריגים במי מהקבוצות?
4. חשבו את התפלגות הנתונים המשותפת על ידי החסרת הממוצע הקבוצתי התואם לכל אחד מהנתונים. בדקו את הנחות המודל לבדיקת ההשערה הסטטיסטית שניסחתם.
5. באם מתקיימות הנחות מודל, חשבו את השונות הבין מדגמית, את השונות התוך מדגמית, את האומדן הסטטיסטי (F) , ואת הסיכוי לדחיית השערת האפס בטעות (P_v) . לצורך חישובים אלו יש למצוא תחילה את גודל המדגם של קבוצה, את ממוצעי המדגמים, את השונות המדגמיות ואת הממוצע הכולל של כל הנתונים שנמדדו.
6. נסחו את מסקנתכם – האם ניתן לדחות או לקבל את השערת האפס? מה ניתן להסיק מכך?
7. ערכו את המבחן התואם לחישוב שביצעתם בשאלה 5 ב-R ובדקו את תוצאותיכם.