

# ניתוח נתונים עם R-52414

## תרגיל 1

### תכנות בסיסי, לולאות, גרפים

#### דגשים להגשת התרגיל

יש להגיש את קובץ הפלט מסוג pdf של Rmarkdown. וקובץ markdown (קוד) נוסף.

פלט ה-markdown צריך להיות עם הסברים של הפעולות שנעשו וכתורות לחלקים בקוד. יש להציג את פלט הקוד, עבור השאלות המודגשות-אין להציג את הקוד!

על הגרפים להיות אינפורמטיביים ככל האפשר, כלומר יש לשים כותרות לגרף ולצירים. להוסיף מקרא בעת הצורך וכדומה.

על הקוד להיות קריא ככל הניתן. כלומר שהחלוקה לצ'אנקים תהיה על סמך הפעולות שמתבצעות. יש להוסיף הערות והסברים גם בתוך הקוד. שמות המשתנים צריכים להיות מובנים ובהקשר לתוכן המשתנה.

אין הגבלה על כמות החבילות שניתן לטעון. יש לבצע את טעינת החבילות בצ'אנק הראשון של הקוד ולא להדפיס את פלט הקוד (בעזרת `echo=FALSE`). עם זאת יש לספק הסבר קטן לכל חיבלה שהשתמשתם ולאיזה צורך היא משמשת. לדוגמא: החבילה `readxl` משמשת לקריאת קובץ אקסל.

תחת כותרת ה-author יש לכתוב ת.ז בלבד! אין להגיש את התרגיל בצורה שמית.

#### רקע על הנתונים

במעבדה זו אנו באים לבדוק האם קיים קשר בין שכר המורים לציוני התלמידים במבחני pisa. לצורך לענות על השאלה מצורפים מספר קבצים.

#### קובץ pisa

מבחן הפיזה (<http://www.oecd.org/pisa/test/>) הינו מבחן סטנדרטי העובר במדינות המתקדמות. בקובץ נמצאים במספר גיליונות הציונים של מדינות שונות בשלושת הנושאים העיקריים של מבחן הפיזה: מדעים, שפה ומתמטיקה. בנוסף לכך לכל שורה מצורפת שנת הבחינה.

#### בקובץ sal

בקובץ מצורפים שני גיליונות נתוני השכר של מורים בעלי 15 שנות ותק במדינות שונות. בגיליון הראשון ניתן למצוא את השכר של כל מדינה בשנת בסיס מסוימת ואת התפתחותו לאורך השנים. בגיליון השני נמצא טבלה המכילה את המדינות, את אינדקס השכר בשנת 2017 (כפי שמופיע בטבלה הראשונה) ואת השכר הנומינלי בדולרים של כל מדינה.

ייתכן ובזמן העבודה על הנתונים תצטרכו לעשות מניפולציות כלשהן.

#### משימות

##### 1. נתוני פיזה (30)

1.1. יש לטעון את נתוני הפיזה.

1.1. יש להציג את שלושת המדינות המובילות בכל תחום בשנת 2015. איזו מדינה מובילה את הדירוגים?

1.3. יש לאחד את שלושת הטבלאות לטבלה אחת תוך שימוש בפקודה join/merge, לטבלה יש לקרוא בשם pisa. יש לשנות את שם העמודות של pisa כך שיתאמו לתוכן העמודות, כלומר לעמודת המדינה יש לקרוא בשם מדינה וכן הלאה.

1.4. יש ליצור משתנה חדש שיכלול את ממוצע המבחן של כל מדינה בכל שנה.

1.5. יש להציג את שלושת המדינות המובילות בממוצע הבחינה בשנת 2006 ובשנת הבחינה האחרונה.

## 2. נתוני שכר מורים(30)

2.1. יש לטעון את נתוני שכר המורים

2.2. יש להמיר את שכר-האינדקס של כל מדינה לשכר בדולרים. י

2.3. יש להוסיף שורה של שכר בינלאומי ממוצע בכל שנה, קראו לממוצע השכר בין המדינות כפי שקוראים למשתנה זה בטבלת הציונים ("International Average (OECD)").

2.4. בשנת 2010 לאיזו מדינה יש את השכר הכי קרוב לממוצע, לאיזו מדינה יש את הפער הכי גדול מלמעלה ומלמטה לממוצע?

2.5. נא להציג רשימה של 5 המדינות שבהן השכר בדולרים בשנת 2015 הוא הגבוהה ביותר. האם זה תואם לדירוגים המדינות המובילות במבחן פיזה?

2.6. יש להציג את חמשת המדינות ששכר המורים עלה בהן בשיעור הכי גבוהה מ-2005 ועד 2017. יש להציג את השכר בדולרים בשנת הבסיס וסוף התקופה. ואת שינויי השכר באחוזים. מי המדינה עם שינוי שכר המורים הגבוהה ביותר? האם התוצאות מפתיעות?

## 3. בחינת השפעת השכר על הציון במבחן(30)

3.1. יש לאחד את הנתונים לטבלה אחת.

3.2. יש לצייר ארבע גרפי נקודות של הציון בכל קטגוריה, וגרף אחד לציון הממוצע אל מול השכר. האם יש קשר?

3.3. יש לבחור גרף אחד ולהוסיף לו קו רגרסיה. סמנו את ישראל ואת ממוצע ה-OECD בצבע אחר.

3.4. כתבו סיכום של הממצאים. העלו את הנקודות החשובות לדעתכם העולות מהניתוח.

## 4. שאלה פתוחה(10)

4.1. חפשו בנתונים מידע נוסף שלא התייחסנו אליו עד עכשיו, משהו שיספר סיפור חדש ומעניין. ניתן לצייר בגרף או להציג בטבלה.