

## מבוא לניתוח נתונים – תרגיל מספר 2

- תאריך הגשה: 20.4
- ההגשה לבד או בזוגות. במקרה של הגשה בזוגות -רק אחד מבני הזוג מגיש, אך יש להקפיד לרשום את השם ות.ז. של שני הסטודנטים שהגישו. מי שלא רשום על אף עבודה לא יקבל ציון, ולא יוכל להצטרף לעבודה בדיעבד.
- במידה וסעיף מסוים לא רץ – הציון על סעיף זה הינו אפס.
- יש להראות את הקוד דרכו הגעתם לפתרון. סעיף עם תשובה נכונה אך ללא קוד יקבל ציון אפס.
- התרגיל הינו אישי וניתן לבצעו בזוגות בלבד. אין להעביר פתרונות או חלקי פתרונות בין סטודנטים בכיתה. העברת התרגיל על כל המשתמע מכך לאנשים אחרים שאינם רשומים לקורס אסורה. כל הנ"ל הינם עבירות משמעת אשר יועברו לבחינת ועדת המשמעת של האוניברסיטה.
- הקפידו להגיש עבודה ברורה עם הסברים. בכל שלב, יש להסביר בבירור על איזה סעיף אתם עונים. על סעיף ללא הסברים יורדו לפחות חצי מהנקודות.
- יש להגיש קובץ אחד בלבד, מסוג ipynb.
- איחורים בהגשה מכל סיבה שאינה נמצאת בתקנון הרשמי של האוניברסיטה: על כל יום איחור בהגשה יורדות 5 נקודות בציון התרגיל באופן אוטומטי.

יש להמשיך עם אותו הקובץ שיצרתם בתרגיל מס' 1. כלומר – עם מדינה מקובץ החיסונים שמתחילה באותה האות באנגלית כמו השם של אחד מחברי הצוות. כל השאלות הבאות מתייחסות אך ורק למדינה שבחרתם (ולא לקובץ הכללי). **יש להגיש את התרגיל בתוך תבנית קובץ ה ipynb המצ"ב.**

1. (5%) כתבו איזו מדינה בחרתם והציגו את 3 השורות הראשונות בקובץ שלכם.
2. (10%) עבור איזו עמודה יש הכי הרבה נתונים חסרים? כמה נתונים חסרים יש עבור עמודה זו?
3. (10%) הציגו רק את השורות שעבורן יש ערכים חסרים לעמודה daily\_vaccinations
4. (10%) מלאו ערכים חסרים עבור העמודה daily\_vaccinations. הסבירו באיזו שיטה השתמשתם למלא את הערכים החסרים, ומדוע אתם חושבים שזו השיטה הנכונה לעמודה זו.
5. (15%) צרו הסטוגרמה משותפת ל: people\_vaccinated ול total\_vaccinations
6. (10%) צרו boxplot עבור העמודה people\_fully\_vaccinated\_per\_hundred. בידקו מהו הקו הירוק בתוך ה boxplot. האם הוא מתאים לחציון, ממוצע, שכיח או סטיית התקן של העמודה?
7. (20%) הוסיפו עמודה month המייצגת את החודש עבור כל רשומה. הציגו את people\_vaccinated\_per\_hundred ואת daily\_vaccinations המקסימליים החודשיים. כלומר, הציגו טבלה שנראית בסגנון הבא:

```
people_vaccinated_per_hundred  daily_vaccinations
month
```

8. (10%) מיינו את הנתונים שהצגתם בסעיף 7 לפי people\_vaccinated\_per\_hundred כך שיוצגו מהערך הקטן ביותר לגדול ביותר (במקרה והערכים שלכם כבר ממוינים כברירת מחדל, הציגו את הדרך בה יש למיין).
9. (10%) נקודות מתנה. נקודות אלו ירדו אוטומטית במידה ויוגש ערעור על הציון.

בהצלחה!!!

צוות הקורס