

# תרגיל רענון - ניתוח נתונים

## הסבר

לתכנת זה קצת כמו לרכוב על אופניים – אם לא מתאמנים יש תאונות וזה כואב...  
לכן נעשה תרגיל פשוט בניתוח נתונים.

## הבעיה:

כולם מדברים על איך עושים כסף מהשקעות בבורסה, אז היום ננסה לעזור למי שרוצה.

נחקור את הדרך הכי טובה להשקיע במדד (תיכף מסביר) שנקרא S&P 500 (תיכף מסביר), או בכינוי חיבה על ידי משקיעים ישראלים "סנופי".

נתחיל מרקע כללי, זו בעיקר העשרה, אין ממש צורך להבין לעומק מניות ומדדים: קודם כל כמה מושגים פשוטים:

1. חברה ציבורית – חברה ציבורית היא חברה שנסחרת בבורסה אחת או יותר וכל אחד כמעט (הציבור) יכול לקנות חלק ממנה. טסלה, מיקרוסופט ובנק הפועלים למשל, כולן חברות ציבוריות. חברת 'רפא"ל' למשל – איננה חברה ציבורית.
2. מניה – מניה היא נייר ערך המייצג בעלות בחלק מחברה. למשל אם יש לי מניות בחברת מיקרוסופט, אני למעשה הבעלים של חלק מהחברה, יחד עם כל שאר מחזיקי המניות. לצורך התרגיל – כל מה שאתם צריכים להבין שהערך של מניה מעיד בדרך כלל (אבל לא תמיד) על המצב הכלכלי של החברה, כשהחברה גדלה המנייה עולה ולהיפך.
3. אינדקס \ מדד – אינדקס או מדד הוא שקלול של מספר מניות יחד, למשל אינדקס תל אביב 100 מייצג ממוצע משוקלל של 125 מניות של החברות הגדולות ביותר בבורסת תל אביב. בדרך כלל מדדים מהסוג הזה נותנים לנו מושג על מצב הבורסה והכלכלה – כשמדד תל אביב 125 עולה משמעותית זה בדרך כלל סימן שהכלכלה הישראלית מתחזקת (ולהיפך).
4. מדד S&P 500 – מדד הכולל 500 חברות גדולות הנסחרות בבורסות בארה"ב, במדד זה נמצאות רוב חברות הטכנולוגיה הגדולות בעולם כמו אפל, טסלה, גוגל ומיקרוסופט, אבל לא רק חברות טכנולוגיה, זהו מדד משוקלל ולכן החברות הגדולות שבו משפיעות יותר על המדד מאשר החברות הקטנות.

השקעה במדד לעומד השקעה במניה בודדת נחשבת השקעה פחות מסוכנת, שכן גם אם חברה אחת או שתיים נקלעת למשבר פתאומי, השינוי במדד יהיה קטן יחסית. בתרגיל זה נחקור מהי אסטרטגיית ההשקעה הטובה ביותר במדד S&P 500.

## המשימה:

לרשותכם ההיסטוריה של מדד S&P 500 בעשר שנים האחרונות, עליכם לכתוב תכנה (קצרה) שתמצא בכל התקופה הזו מה התקופה (בהנחה שאני קונה) משקיע פעם אחת ומוכר את כל מה שקניתי פעם אחת בדיוק) שבא היה כדאי להשקיע כדי לקבל את הרווח הגבוה ביותר.

## היסטוריה של המדד:

ניתן למצוא את ההיסטוריה של מדד ה S&P 500 על ידי חיפוש פשוט בגוגל. לנוחותכם – הורדתי את כל ההיסטוריה לקובץ הנמצא כאן:

[https://github.com/weiss-gal/data\\_science\\_project/blob/main/2023\\_2024/11th\\_grade/lessons/01\\_data\\_analysis/historical\\_data\\_spx.csv](https://github.com/weiss-gal/data_science_project/blob/main/2023_2024/11th_grade/lessons/01_data_analysis/historical_data_spx.csv)

הקובץ מכיל שורה לכל יום, ובה ארבעה ערכים – פתיחה, סגירה, נמוך וגבוה.

פתיחה וסגירה – הערך של המדד בתחילת ובסוף יום המסחר (לא רלוונטי אלינו בתרגיל הזה)

נמוך וגבוה – הערך הכי נמוך והכי גבוה שאליו הגיע המדד באותו יום

אפשר להוריד את הקובץ למחשב ולפתוח אותו באמצעות אקסל (אף על פי שזהו אינו בדיוק קובץ אקסל, כפי שאסביר מיד)

### איך קוראים את הקובץ בפייתון

אם תשימו לב הסימנים של שם הקובץ היא csv, זהו קיצור של Comma separated values או בשפת הקודש – "ערכים המופרדים בפסיקים". זהו פורמט מאוד נפוץ בתחום מדעי המידע – מעין "אקסל לעניים" ובוא כל שורה בנויה מעמודות המופרדות בפסיקים, כאשר השורה הראשונה היא שורת הכותרות של העמודות והשורה השנייה והלאה הם הערכים. תוכלו לראות איך באמת בנוי הקובץ אם תפתחו אותו באמצעות Notepad או אם תבחרו בלינק ששמתי לכם להציג אותו במצב "Code", זה נראה ככה:

data\_science\_project / 2023\_2024 / 11th\_grade / lessons / 01\_data\_analysis / historical\_data\_spx.csv

weiss-gal Add files via upload

Preview Code Blame 2537 lines (2537 loc) • 100 KB

1	Date,Close/Last,Open,High,Low
2	11/22/2023,4556.62,4553.04,4568.43,4545.05
3	11/21/2023,4538.19,4538.77,4542.14,4525.51
4	11/20/2023,4547.38,4511.70,4557.11,4510.36
5	11/17/2023,4514.02,4509.55,4520.12,4499.66
6	11/16/2023,4508.24,4497.08,4511.99,4487.83
7	11/15/2023,4502.88,4505.30,4521.17,4495.31
8	11/14/2023,4495.70,4458.97,4508.67,4458.97
9	11/13/2023,4411.55,4406.66,4421.76,4393.82

כדי לקרוא אותו נשתמש בספרייה מתאימה, יתכן וכדי להתקין אותו תצטרכו להפעיל את הפקודה הבאה בשורת הפקודה:

```
Command Prompt
Microsoft Windows [Version 10.0.19045.3693]
(c) Microsoft Corporation. All rights reserved.

C:\Users\weiss>pip install csv_
```

הספרייה מאפשרת לכם לקרוא את הקובץ שורה אחרי שורה כאשר בכל פעם תקבלו את השורה כרשימה של ערכים, ערך לכל עמודה בטבלה.

דוגמא לשימוש בספרייה ניתן למצוא כאן:

<https://docs.python.org/3/library/csv.html>

שימו לב – הפקודה `csv.reader()` מחזירה אובייקט שניתן לקרוא אותו באמצעות לולאת `for` (כמו בדוגמא) או לקרוא ממנו שורה שורה באמצעות הפקודה `next()`.

קריאה ל-`next()` על האובייקט "מזיזה" אתכם שורה קדימה בקובץ, ניתן להשתמש בה כדי לדלג על השורה הראשונה (שמכילה את הכותרות ולא כל כך רלוונטית אליכם)

#### חישוב הרווח:

בתרגיל הזה אנחנו מדברים על רווח יחסי, הנוסחא היא:

(מחיר המכירה – מחיר הקנייה) \ מחיר הקנייה

מחיר קנייה \ מחיר מכירה: בגלל שהמחיר משתנה בכל יום נלך לפי ההנחה שכאשר אנחנו קונים זה תמיד יהיה במחיר הגבוה של אותו יום וכאשר אנחנו מוכרים זה תמיד יהיה במחיר הנמוך של אותו יום.

#### הנחייה:

- א. כדאי מאוד לעבוד בשלבים – קודם כל לראות שאתם מצליחים לקרוא את הקובץ, אחר כך להוציא את הנתונים הרלוונטים ורק בסוף לפתור את הבעיה.
- ב. זה מאוד לא יעיל לקרוא את הקובץ הרבה פעמים – כדאי לעשות לולאה אחת שקוראת את הקובץ ושומרת את הנתונים שתצטרכו ואז עליהם לבצע את כל הפעולות שתרצו
- ג. אל תנסו למצוא פתרון אלגנטי או מתמטי מתקדם או יעיל – אתם בסך הכל צריכים למצוא מה השילוב של יום קנייה ויום מכירה שיתן את הרווח הכי גבוה. זה לגמרי לגיטימי לבדוק את כל האפשרויות
- ד. יש שורות שבהן שלושה ערכים מתוך ארבעה הם 0.00 – תדלגו עליהן

#### התוצאה המבוקשת:

אחרי שהתכנית מצאה את יום הקנייה ויום המכירה האידאליים, היא תדפיס תשובה בסגנון הבא:

The best timing is to buy at: 13-1-2014 and sell at 23-7-2020.

The profit would be 50%