

# 234124 - מבוא לתכנות מערכות

## תרגיל בית 3

### סמסטר אביב 2021

תאריך פרסום:

תאריך הגשה:

משקל התרגיל: 5%

מתרגלים אחראים: נלסון גולדנשטיין, רוברט שהלה

מענה לשאלות בנוגע לתרגיל יינתן אך ורק בפורום התרגיל בפיאצה או בשעות הקבלה. לפני פרסום שאלה בפורום אנא בדקו אם כבר נענתה – מומלץ להיעזר בכלי החיפוש שהוצגו במצגת האדמיניסטרציה בתרגול הראשון.

## 1. הערות כלליות

- יש להגיש את תרגיל הבית בזוגות בלבד.
- שימו לב: לא תינתנה דחיות במועד הגשת התרגיל פרט למקרים חריגים. תכננו את הזמן בהתאם.
- כל חומר נלווה לתרגיל נמצא על השרת בתיקייה `~mtm/public/2021b/ex3`.
- קראי את התרגיל עד סופו לפני שאתם מתחילים לממש. חובה להתעדכן בעמוד ה-F.A.Q של התרגיל, הכתוב שם מחייב.
- העתקות קוד בין סטודנטים ובפרט גם העתקות מסמסטרים קודמים תטופלנה. עם זאת – מומלץ ומבורך להתייעץ עם חברים על ארכיטקטורת המימוש.
- מומלץ מאוד לכתוב את הקוד בחלקים קטנים ולבדוק שהוא עובד באמצעות שימוש בטסטים קטנים שתכתבו בעצמכן. לא נדרש מכם בתרגיל להגיש טסטים, אך כידוע, כולנו בני אדם – רצוי לבדוק את התרגיל שלכם היטב כולל מקרי קצה, כי אתו תידרשו לכך.

## 2. הקדמה

מטרת תרגיל זה היא היכרות עם תכנות ב-Python. התרגיל מורכב משני חלקים:

1. שליפת נתונים מקובץ טקסט באמצעות Python.
2. עבודה עם מחרוזות

## 3. חלק רטוב

### 3.1. טיפול בקובץ טקסט באמצעות Python

#### 3.1.1. שליפת נתונים מקובץ טקסט

בעקבות שדרוג מערכת המזגנים בפקולטה למדעי המחשב, הסמסטר אין תקציב לבדוק מבחנים. ועל כן, הפקולטה מבקשת את עזרתכם לפתח מערכת כדי להוזיל את בדיקת המבחנים ע"י חלוקת ציונים בצורה אוטומטית.

המערכת תקבל קובץ המכיל את מידע על כל הסטודנטים בקורס. הקובץ נתון בפורמט הבא. שימו לב כי עלולים להיות כמה רוחים בין כל מילה בשורה אבל פסיק אחד בין כל שדה:

`<id (int)>, <name (string)>, <semester (int)>, <homework avg (int)>`

מובטח שבשדה המסומן (int) יופיע רק ספרות.

לדוגמא, במקרה של 2 רשומים, הקובץ יראה כך:

`39401830, Zeev Jabotinsky, 2, 78`

`29441133, Joseph Trumpeldor, 1, 99`

אירעה תקלה במערכת, והשתבש חלק מהמידע בכמה שורות. שורה נחשבת כתקינה כאשר:

- תעודת זהות ייחודית בת 8 ספרות בלי אפסים בהתחלה. 01234567 הוא לא תקין כי יש 0 בהתחלה.
- שם חוקי הוא מחרוזת של אותיות [A-Z, a-z] ורווחים בלבד.
- סמסטר הוא מספר הגדול או שווה ל-1.
- ממוצע הציונים בתרגילי בית הוא גדול מ-50 וקטן שווה ל-100.

בנוסף, יש סטודנטים המופיעים מספר פעמים בקובץ, כלומר אותה תעודת זהות יכולה להופיע יותר מפעם אחת בקובץ. אנחנו מעוניינים להתייחס רק למופע האחרון התקין של כל ת.ז.

עליכם לייצר קובץ פלט מתוקן המכיל את הציון הסופי של כל סטודנט בפורמט הבא:

`<id (int)>, <homework avg (int)>, <final grade (int)>`

סטודנטים שהשורה שלהם בקובץ הקלט היא לא תקינה לא יקבלו ציון, כלומר לא יופיעו בקובץ הפלט. בנוסף, כמו שאמרנו רק המופע האחרון התקין של כל ת.ז יקבל ציון (כלומר אין כפילויות בקובץ הפלט).

קובץ הציונים הסופיים צריך להיות ממויין בסדר עולה, לפי מספר תעודת הזהות שלהם, עם רוח אחד ופסיק בין כל שדה.

אופן חישוב הציון הסופי (final grade) בקורס מת"מ:

הציון הסופי הוא הממוצע (ערך שלם תחתון) בין שתי הספרות האחרונות של מספר ת.ז וציון תרגילי הבית (homework avg).

לדוגמה: עבור ת.ז. 123456789 וציון ת.ב. 94:

$$final\ grade = \left\lfloor \frac{89 + 94}{2} \right\rfloor = 91$$

כדי לבצע זאת, עליכם לכלול בקובץ **gradesCalc.py** את הפונקציה בשם `final_grade` שמקבלת כקלט את ה-`path` לקובץ פרטי הסטודנטים המקורי (`input_file`) ואת ה-`path` לקובץ בו ייכתב הפלט (`output_file`) – קובץ הציונים הסופיים (אם הקובץ כבר קיים, יש לדרוס אותו). אין צורך להתייחס למקרים בהם קובץ הקלט אינו קיים או בפורמט שונה ממה שהוגדר. בנוסף, על הפונקציה להחזיר (`return`) את ממוצע הציון הסופי של הקורס כ- `int` (ערך שלם תחתון).

**במקרה שקובץ הקלט ריק, או מכיל רק שורות לא תקינות, קובץ הפלט שהפונקציה תייצר יהיה ריק ותחזיר 0 כממוצע הציון הסופי של הקורס.**

חתימת הפונקציה שעליכם לממש היא:

`final_grade(input_path: str, output_path: str) -> int`

(שימו לב שליד כל ארגומנט לפונקציה מופיע הטיפוס שלו. פייתון 3 נותנת לנו אופציה לציין ליד כל ארגומנט את הטיפוס שאנחנו מצפים לקבל למטרות נוחות ותיעוד. ה-`Interpreter` לא אוכף את קבלת הטיפוס המצויין, ולא יזרוק `TypeError` או כל שגיאה אחרת, במקרה שיועבר טיפוס אחר מזה שצויין.)

לדוגמה, בהינתן הקלט הבא:

```
32305629 , Michael Scott, 8 , 75
33401550, Jim Halpert, 19, 46
01401830, Kevin Malone , 5, 67
12345678, Toby Flenderson , 6, 99
27804448 , Cre100ed Bratton, 8, 82
```

נרצה שהפלט שנייצר ייראה כך:

```
12345678, 99, 88
32305629, 75, 52
```

**טיפ:** השתמשו במתודות `split` ו-`join` של מחרוזות, ניתן לראות את התיעוד של הפונקציות בכתובת הבאה:  
<https://docs.python.org/3/library/stdtypes.html#str.split>  
<https://docs.python.org/3/library/stdtypes.html#str.join>

בנוסף, מותר להשתמש בכל מתודה של טיפוסים נתונים המובנים בפייתון. אסור לעשות `imports` נוספים.

הערות:

- בכל הפונקציות בשאלה, אין לשנות את קובץ הקלט.
- ניתן להניח שכל קובץ הקלט קיים ובפורמט הרצוי.
- אפשר להניח שהארגומנטים שמועברים לפונקציות תקינים מבחינת טיפוס.
- מסופק לכם טסט בשם `test.py`. קבצי הפלט הרצויים נמצאים תחת התיקייה `tests`.

## 3.2. עבודה עם מחרוזות

תממשו פונקציה המקבלת שתי מחרוזות  $s_1$  ו-  $s_2$  ובודקת האם ניתן להרכיב את המחרוזת  $s_1$  מאותיות המחרוזת  $s_2$ . הפונקציה תחזיר True אם כן, אחרת False (bool).  
חתימת הפונקציה שעליכם לממש:

`check_strings(s1: str, s2: str) -> bool`

המחרוזות הינן מעל הא"ב האנגלי, כלומר כל האותיות a-zA-Z. אבל הפונקציה תתעלם מאותיות גדולות, כלומר, פעולת הפונקציה היא (case insensitive).

לדוגמה עבור המחרוזת הבאה  $s_2 = "abcabc"$ :

- (1)  $s_1 = "aabbcc"$ , הפונקציה מחזירה True.
- (2)  $s_1 = "caba"$ , הפונקציה מחזירה True.
- (3)  $s_1 = "aaa"$ , הפונקציה מחזירה False, מכיוון שב-  $s_2$  יש רק שני "a" ולא שלושה.

עוד דוגמה עבור המחרוזת הבאה  $s_2 = "baNaNaN"$ :

- (1)  $s_1 = "naanb"$ , הפונקציה מחזירה True.
- (2)  $s_1 = "ananas"$ , הפונקציה מחזירה False.
- (3)  $s_1 = "bannnn"$ , הפונקציה מחזירה False, מכיוון שב-  $s_2$  יש רק שני "n" ולא שלושה.

## 4.הגשה

הערות:

וודאו שאתם מריצים פייתון גרסה 3.6. שימו לב כי זאת לא גרסה ברירת מחדל על השרת. כדי להריץ פייתון 3.6 השתמשו בפקודה `python3.6`.

עם התרגיל אתם מקבלים את הקבצים: `test.py`, `gradesCalc.py`.

בנוסף לטסט `test.py` עם קבצי הקלט והפלט `input`, `out` שנמצאים תחת התיקייה `tests`.

אתם נדרשים להשלים את `gradesCalc.py`.

את ההגשה הרטובה יש לבצע דרך אתר הקורס, תחת **Electronic Submit <- HW3 <- Assignments**

**יש להגיש רק את הקובץ:** `gradesCalc.py` בתוך תיקיית `ZIP`, לא בתוך תיקיות פנימיות.

- מותר להגיש את התרגיל מספר פעמים, רק ההגשה האחרונה נחשבת.
- על מנת לבטח את עצמכם נגד תקלות בהגשה האוטומטית שימרו את קוד האישור עבור ההגשה. עדיף לשלוח גם לשותפה. כמו כן שימרו עותק של התרגיל על חשבון ה-csl3 שלכם לפני ההגשה האלקטרונית ואל תשנו אותו לאחריה (שינוי הקובץ יגרור שינוי חתימת העדכון האחרון).
- \*\*\* כל אמצעי אחר לא יחשב הוכחה לקיום הקוד לפני ההגשה.