

## פרויקט מסכם - קורס תכנות מתקדם בפייתון

---

מגישים:

ארז כהן (ת.ז: 316098219)

תמיר מלינק (ת.ז: 315856542)

ליהיא דודסקו (ת.ז: 315560425)

---

נושא הפרויקט המסכם הינו: מערכת מידע המהווה כלי לניתוח נתונים אקדמיים של סטודנטים.  
המערכת מבוססת על DATA שאנחנו יצרנו ובה נתונים רנדומליים המציגה חיתוכים שבעזרתם ניתן ללמוד ולהפיק לקחים על מנת לשפר את מוסד הלימודים. המערכת כאמור מיועדת להנהלה, למרצים, ולעוד גורמים רלוונטיים אחרים מהמוסד האקדמי.

המערכת מוצגת ע"י ממשק ובו מוצגים הנתונים בצורה גרפית ע"פ בחירה בקבוצה המתאימה- החוג הרלוונטי (סינון לפי סוג הפקולטה והשנה). הנתונים המוצגים נשאלים כשאלות החוקרות קשר מסוים בין נושאים שבענינו ההצלחה ביניהם יכולה לספק מידע שימושי לשיפור המערכת הלימודית (לדוג': הבדלי ממוצע ציונים בין מסלול בוקר למסלול ערב. ישנו פירוט מלא על כל השאלות בהמשך המסמך) באמצעותם יהיה ניתן לקבל החלטות משמעותיות להמשך.

להלן רשימת הנושאים אשר נלמדו בקורס והשתמשנו בהם בעבודה באופן משמעותי:

1. TKINTER – מערכת המידע הינה מערכת מבוססת ע"ג ממשק. ישנה חשיבות רבה לאופן בו המערכת בנויה ומוצגת למשתמש. היה לנו חשוב ליצור ממשק אינטואיטיבי שיאפשר שימוש נוח למשתמש תוך הקפדה על רף גבוה של ויזואליות.
2. MATPLOTLIB - חלק בלתי נפרד ממערכת מידע הוא החלק הויזואלי של הצגת הנתונים. על מנת להקל על המשתמש בעיבוד המידע, עלינו להמחיש את הנתונים ולהציגם בצורה ברורה. לכן, חלק מרכזי מהפרויקט שלנו נשען על השימוש בנושא זה – הצגת התוצאות של כל שאלתה מבוסס על שימוש בנושא זה.
3. PANDAS - מטרת מערכת המידע שלנו היא להציג כל פעם (על פי בחירת המשתמש) מידע מתבקש בנושאים מסויימים. הקובץ מידע אותו יצרנו למטרה זו הינו קובץ עמוס במידע ולא כל המידע רלוונטי לכל שאלה ומחקר. על כן, היה עלינו לבחור בכלי אשר יאפשר לנו לנתח את הנתונים הרבים ביעילות ובקלות.

4. קבצים- על מנת להפעיל את מערכת המידע שלנו, יצרנו data המתורגמת לקובץ אקסל. בקובץ שיצרנו ישנו כל המידע הדרוש לקבלת תמונת מצב מלאה על נתוני הסטודנטים ובעזרת הנושאים האחרים שבחרנו לעבוד איתם אפשר לעבד אותו להפיק מהקובץ את המקסימום. בנוסף יש שימוש בקבצים בממשק עצמו- ישנה אפשרות לפתוח ולשמור קבצים מסוימים שרלוונטים למערכת המידע.

יש לציין שבמהלך העבודה ראינו לנכון לשפר את היכולות הגרפיות המוצגות בממשק וכן נעזרנו באתרים מסוימים כמו matplotlib על מנת לעשות זאת.

#### חלוקת משימות ועבודת צוות:

עבודתנו על הפרויקט דרשה מאיתנו לעבוד ביחד ולחוד. ראשית, חשבנו יחדיו על רעיון, פיתחנו אותו וכתבנו את המשמעויות התכנותיות של כל אלמנט בפרויקט. ניסחנו יחדיו 14 שאילתות המהוות בעינינו את הנושאים המרכזיים שצריכים להיות על הפרק עבור משתמשי המערכת. חילקנו בינינו את השאילתות וכל אחד מחברי הקבוצה בנה את השאילתות שלו על בסיס הDATA, כאמור כל שאילתה מהווה פונקציה. לאחר בניית כלל הפונקציות נפגשנו על מנת לאחד הכל ולפתח את הממש עליו יוצגו הנתונים. לבסוף הצלחנו לפתח ממשק בו בקלות רבה מוצגים נתונים באופן ברור וחד משמעי שניתן להסיק בעזרתם מסקנות ולקחים על מנת לשפר את המוסד הלימודי.

#### ספריות חיצוניות נדרשות:

```
import pandas as pd

from tkinter import *

from tkinter import ttk

import numpy as np

import matplotlib.pyplot as plt

import tkinter.filedialog as FD

from PIL import ImageTk,Image

import matplotlib.patches as mpatches

from mpl_toolkits.mplot3d import Axes3D
```

**הערה:** ע"מ להריץ את הקוד, יש לשמור את כלל קבצי הפרוייקט באותה תיקייה בנתיב יחסי.

אנחנו מרוצים מעבודתנו כצוות ומהרמה המקצועית שהגענו אליה במהלך הקורס ובפרט בעבודה.

ראינו לנכון לצרף פירוט קצר על כלל הפונקציות שבנינו לטובת בהירות בנושא. ראשית, מסך הבית של המערכת מאפשר למשתמש לבחור פקולטה מסוימת עליה הוא רוצה לקבל מידע. יש לציין שאנו בנינו את המערכת למוסד האקדמי שנקר, אך ורק לסטודנטים בשנה א' בפקולטה להנדסה, במחלקה להנדסת תעשייה וניהול אך כמובן שניתן להתאים את זה ע"י העתקת הקוד לכלל המגמות במוסד.

מיד לאחר הבחירה בפקולטה נפתח חלון חדש בו המשתמש יכול לשלוף מידע ולבצע פעולות שונות המחולקות ע"פ הקטגוריות הבאות:

File, Candidate students, Year A, Year B, Year C, Year D

#### פעולות:

- חיפוש מידע על סטודנט קיים: ישנה אפשרות למצוא מידע על סטודנט ספציפי על פי התז שלו - לאחר הקלדת תז בחלונית המתאימה למשתמש יוצגו שלושה גרפים אשר מייצגים את הישגיו ונתוניו בתחומים שונים כמו: בגרריות, קורסים אקדמיים כגון פייתון, חדו"א ועוד. במידה והסטודנט לא קיים תקפוץ הודעה באדום ותתריע על כך למשתמש.
- הרשמת סטודנט חדש: ישנה אפשרות לרשום סטודנטים חדשים דרך המערכת - יצרנו טופס הרשמה הכולל פרטים נדרשים רלוונטיים(תז, פקולטה רצויה, הישגים במקצועות הרלוונטיים לקבלה ועוד). המידע על הסטודנטים החדשים נשמר כקובץ טקסט וקל לשליפה במידת הצורך.
- חישוב ציון סכם (ע"ב ציון פסיכומטרי ובגרריות): משתמש אשר רוצה לבדוק את סיכויי קבלתו ללימודים יכול להשתמש במחשבון הקבלה המתכלל את ציוני הפסיכומטרי וסכם הבגרריות של המועמד. המחשבון מציג לסטודנט העתידי האם הוא עובר (כיתוב בירוק)/ לא עובר (כיתוב באדום) את רף הקבלה ע"פ ציון סכם משוקלל שהוגדר מראש על ידינו (כרגע ציון סכם עובר הינו ציון מעל 70, והוא נבחר רנדומלית, בפועל יהיה ניתן לשנות רף זה ולהתאימו לדרישות של כל מוסד אקדמי).
- פתיחת קבצים נוספים מהמחשב: לקחנו בחשבון שבעת השימוש במערכת שלנו המשתמש ירצה להיעזר במקביל במסמכים פרטיים הקיימים אך ורק אצלו במחשב. על כן, באמצעות המידע שלמדנו על שימוש בקבצים בקורס יצרנו פלטפורמה המאפשרת פתיחת ושמירת קבצים מהמחשב ואף יצרנו להם מקשי קיצור (שמירה - CTRL S , פתיחה - CTRL O) על מנת לייעל את השימוש.

לאחר המידע הבסיסי על הסטודנט, ישנן 14 שאילות המוצגות באמצעות גרפים:

1. מספר תלמידים מצטיינים בין השנים 2015-2020 : זיהוי מגמת מצטיינים באקדמיה הוא חלק בלתי נפרד מניתוח הנתונים שעל העומדים בראש המחלקות השונות לבצע. לטובת עיבוד והצגת נתונים אלו, בנינו פונקציה המשכללת את עמודות ציון סמסטר א' וציון סמסטר ב' לכל התלמידים בין השנים הדרושות. הגדרנו רף הצטיינות של ממוצע שנתי עם ציון של **85 ומעלה**. הפלט הגרפי מבוסס על שימוש ב-MATPLOTLIB והינו מערכת צירים שמציגה בגרף עמודות את הכמות המספרית של כל הסטודנטים המצטיינים בכל שנה.
2. ממוצע שנתי של סטודנטים בשנה א' בין השנים 2015-2020: כשרוצים לבצע מחקר מקיף על קבוצת אנשים בעלת מכנה משותף (למשל, סטודנטים שנה א' להנדסה תעשייה וניהול) יש לזהות ראשית האם ישנה יציבות בנתונים. על כן הרכבנו את השאילתה הזו ואת ההצגה הגרפית של תוצאותיה תוך שימוש בפונקציה היוצרת ממוצע, MATPLOTLIB.
3. הקשר בין ממוצע הישגים בפיזיקה 0, פיזיקה 1 ופיזיקה 2: רוב הסטודנטים בפקולטה להנדסה עוברים מכינה לקראת הלימודים ע"מ ליצור בסיס ידע לקורסי המשך. אחד מהמקצועות הנלמדים במכינה הוא חדו"א 0 - קורס יסוד בדרך לחדו"א 1 וחדו"א 2. ראינו לנכון לבחון את הקשר בין הממוצע של כל אחד מהקורסים בין השנים 2015-2020. יצרנו פונקציה המחשבת את הממוצע של קורס תוך הפרדת נתונים על פי השנים. הנתונים מוצגים בגרף עמודות.
4. הקשר בין ממוצע הישגים בחדו"א 0, חדו"א 1 וחדו"א 2: כפי שמפורט בסעיף הקודם, הסטודנטים עוברים קורס יסוד בפיזיקה 0 המכין אותם לקראת קורסי ההמשך בפיזיקה. את ניתוח הנתונים בשאילתה זו ביצענו בצורה דומה לניתוח הנתונים שנעשה בסעיף הקודם - חישוב ממוצע של הקורס ע"פ פונקציה שבנינו תוך הפרדת נתונים על בסיס שנת הלימוד. הנתונים מוצגים בגרף עמודות.
5. צפי נשירה של סטודנטים בשנה א' בין השנים 2015-2020: כמו שנכון לזהות מגמות הצטיינות(שאילתה 1), יש לזהות מגמות נושרים אפשריים. חקירת נתון זה לאורך טווח שנים יכול לעזור בזיהוי מגמות בקרב הסטודנטים וכמו כן לזהות נק' תורפה/לחיזוק בתחום ההוראה. בנינו פונקציה המזהה את הסטודנטים שממוצעם השנתי **נמוך מ-60** והגדרנו אותם כסטודנטים בפוטנציאל נשירה. נתוני שאילתה זו מופיעים בגרף עמודות אשר מציג נתון מספרי של סטודנטים אפשריים לנשירה בכל שנה בין השנים 2015-2020.

6. הקשר בין ממוצע ציונים לממוצע סכם קבלה בין השנים 2015-2020: כל סטודנט מתקבל ללימודים עם סכם קבלה מסוים ויוצא מהשנה הראשונה עם ממוצע ציונים משלו. רצינו לבחון האם יש קשר מגמתי בין הציונים. בנינו פונקציה אשר מחשבת את ממוצע הציונים השנתי בכל שנה ופונקציה אשר מחשבת את ממוצע סכם הקבלה של הסטודנטים בכל שנה. השונו ויזואלית בין הנתונים באמצעות גרף עמודות המציג את ההבדלים.
7. הבדלי הישגים בין מסלול בוקר למסלול ערב במקצועות ליבה (חדו"א, פיזיקה ופייתון): בחוג להנדסת תעשייה וניהול תחת הפקולטה להנדסה יש אפשרות ללמוד במסלול בוקר/ערב. אנו מאמינים כי ההשוואה בין מסלולי הלמידה והקפדה על צמצום פערי ציונים בין השניים היא חשובה ועל כן יש לחקור אותה. הגדרנו בקובץ DATA שלנו עמודה המשקפת את חלוקת הסטודנטים בין מסלול בוקר וערב (0- בוקר, 1- ערב). בהתבסס על עמודה זו יצרנו פונקציה המחשבת את ממוצע הציונים במקצועות הליבה בין המסלולים ובין השנים 2015-2020.
8. השפעת השתתפות בקורס אנגלית על ממוצע הציונים השנתי: על כל סטודנט להגיע לרמת פטור באנגלית עד שנה ג' באקדמיה. ישנם הרבה סטודנטים אשר מעדיפים לקחת על עצמם את הנושא כבר מהשנה הראשונה. מכיוון שמדובר בעוד קורס אשר התווסף למערכת עמוסה כשלעצמה, זיהינו נקודה חשובה להשפעה - האם הקורס באנגלית משפיע/לא משפיע על הישג הסטודנט. הגדרנו בקובץ DATA עמודה המשקפת את חלוקת הסטודנטים (0 - משתתף בקורס אנגלית, 1 - לא משתתף). בהתבסס על חלוקה זו חישובנו באמצעות פונקציה את הממוצעים של הסטודנטים המשתתפים ולא משתתפים בקורס אנגלית והשונו ביניהם באמצעות גרף מקווקו.
9. הקשר בין רמת בגרות במתמטיקה (יח"ל) להישגים בקורסי חדו"א (0,1,2): הסטודנטים מגיעים עם רקע תיכוני שונה אל האקדמיה. על אף שדרישות הסף זהות עבור כולם, ראינו לנכון לחקור האם יש הבדל בהישגי הסטודנטים באקדמיה לבין רמת היחידות שביצעו במתמטיקה בתיכון. בנינו עמודה בקובץ DATA המחלקת את הסטודנטים לפי היח"ל (3,4,5). בהתבסס על עמודה זו חילקנו את ממוצעי הסטודנטים בקורסי החדו"א. הנתונים מוצגים ויזואלית בגרף עמודות המציג את השוני בין השנים 2015-2020.
10. הקשר בין רמת הבגרות בפיזיקה (יח"ל) להישגים בקורסי פיזיקה (0,1,2): בהמשך לסעיף הקודם, השונו גם בין הישגי הסטודנטים בקורסי הפיזיקה על בסיס רמת הפיזיקה שלמדו בתיכון. בנינו עמודה בקובץ DATA אשר מחלקת את הסטודנטים על פי רמת יחידות הפיזיקה שביצעו בתיכון (3,4,5). לאחר מכן בנינו פונקציה אשר מחשבת את הממוצע בכל קורס פיזיקה של כל קבוצת סטודנטים (סטודנטים אשר ביצעו 3 יח"ל, 4 יח"ל וכו'..). הנתונים מוצגים באמצעות גרף עמודות המשקף את ההבדלים בין השנים 2015-2020.

11. אחוז הסטודנטים בשנה א' הנכשלים במקצועות ליבה (חדו"א, פיזיקה ופייתון): בנינו פונקציה אשר מחשבת את ממוצע הסטודנטים בכל שנה במקצועות הליבה. כל סטודנט שממוצעו השנתי במקצוע מסוים **נמוך מ-60** מוגדר כנכשל. בעזרת פונקציה נוספת חישבנו מהו אחוז הנכשלים בכל מקצוע מסך הסטודנטים בכל שנה בין השנים 2015-2020. הנתונים מוצגים גרפית באמצעות מערכת צירים ייחודית - בציר ה-Y ישנם מקצועות ליבה, בגרף ה-X ישנם השנים. הנתונים הגרפים של ההצלבה בין מקצוע לשנה (לדוג': אחוז נכשלים בחדו"א 1 ב-2017) מופיעים במשבצת וכך נוצרת מעין "מפה" אשר משקפת בגרף אחד את מגמתיות הנכשלים בשלושת מקצועות הליבה בין השנים 2015-2020. כמו כן, ישנו מקרא לצד הגרף המראה טווח אחוז הנכשלים בטווח צבע של הצבע ירוק (אחוז נכשלים נמוך - ירוק מאוד בהיר, אחוז נכשלים גבוה - ירוק כהה).

12. מספר הסטודנטים בכל שנה בין השנים 2015-2020: כחלק מהנתונים על כל סטודנט בקובץ DATA אותו יצרנו, מצוינת השנה בה הסטודנט למד. כתבנו פונקציה אשר מחלקת בין הסטודנטים ובעזרתה אפשר לקבל נתון מספרי על הסטודנטים בכל שנה. נתונים אלו מוצגים באמצעות גרף עוגה (החלק הבולט בגרף העוגה - שנת 2019 - מספר הסטודנטים הגדול ביותר בין השנים 2015-2020).

13. הצגת התפלגות ממוצע ציונים שנתי באמצעות גרף היסטוגרמי: גרף היסטוגרמי מציג את הנתונים הנחקרים באמצעות מלבנים בתוך מערכת צירים. על בסיס ההגדרות של כל ציר, השטח של המלבנים מהווה את התפלגות ניתוח הנתונים. במערכת שלנו, ציר X מציג את הממוצע השנתי (0-100) וציר ה-Y מציג את כמות הסטודנטים. בצורה זו, נוח מאוד לראות את התפלגות ממוצעי הציונים בין הסטודנטים (השטח הכחול שיוצרים המלבנים סוכם 500 סטודנטים - מספר הסטודנטים שלמדו בין השנים הנחקרות).

14. הצגת תלת מימד של ממוצעי הציונים בסמסטר א', סמסטר ב' וממוצע שנתי: יצרנו מערכת תלת מימד (שלושה צירים) אשר מציגה בצורה שונה ואף נוחה יותר את הבדלי הממוצעים בין הסמסטרים והממוצע השנתי במהלך השנים 2015-2020. הנתונים מוצגים כנקודות אדומות במערכת הצירים והגרף ניתן להזזה ע"י המשתמש - בעזרת העכבר אפשר להציג את המערכת התלת מימדית מזוויות שונות בהתאם לרצון ונוחות המשתמש.