

הגורם האנושי באיסוף נתונים – 096275

סמסטר חורף תשפ"ה 2024-2025

שיעורי בית #2

בחרו קובץ נתונים עליו תעבדו. על קובץ הנתונים לכלול נתונים שכוללים **התנהגות של בני אדם** הקשורה למזון. ניתן להשתמש בקובץ הנתונים שנאסף על ידכם בשיעורי הבית הקודמים באמצעות webscraping, להוריד באמצעות webscraping כל קובץ נתונים אחר, או להשתמש בקובץ נתונים קיים (ניתן למצוא קבצים כאלו באתרים כמו kaggle, מומלץ לפלטר על פי food).

במידה ובחרתם קובץ נתונים קיים, מלאו את פרטיו באקסל – וודאו שאין זוג נוסף שמבצע את המטלה על אותו קובץ נתונים. במסגרת המטלה תבצעו אקספלורציה ראשונית של נתונים.

1. בחרו 3 משתנים שמעניינים אתכם (מלפחות שתי סקאלות מדידה. שימו לב, אחד מהמשתנים לפחות חייב לתאר התנהגות של בני אדם), תארו במשפט כל אחד מהו כל משתנה ומה הוא מודד, ציינו את סקאלות המדידה ודווחו סטטיסטיקה תיאורית **רלוונטית** (ממוצע, חציון, סטיית תקן, שכיח וכו').
2. וויזואליזציה: הציגו כל אחד משלושת המשתנים באמצעות טבלאות ו/או גרפים רלוונטיים. תארו בקצרה את המידע המרכזי שמוצג.
3. זהו נתונים חסרים וערכים קיצוניים.
 1. האם ניתן לאפיין דפוסים של חוסרים (למשל, משתנים רגישים, אוכלוסיות מסוימות)? נמקו.
 2. כיצד נכון לדעתכם להתמודד עם החוסרים? הסבירו.
 3. הסבירו מדוע דרך ההתמודדות שהצעתם מתאימה.

הוראות הגשה

1. יש להגיש את העבודה מוקלדת בקובץ WORD (David, 12, רווח 1.5 בין שורות, שוליים סטנדרטיים)
2. היקף המטלה עד שני עמודים (כולל גרפים)
3. ההגשה בזוגות (רק אחד מבני הזוג מגיש) במודל בלבד עד השעה 23:59 ב-03.12.24
4. שם קובץ ההגשה יהיו HW2_ID1_ID2

בהצלחה!