

הגורם האנושי באיסוף נתונים – 096275

סמסטר חורף תשפ"ה 2024-2025

שיעורי בית #1

אספו נתונים באמצעות Web Scraping:

בחרו אתר המכיל ביקורות ודירוגים (גם לייק הוא דירוג) של מסעדות, חנויות מזון, מוצרי מזון, שפים וכו' שממנו תרצו להוריד נתונים.

1. תארו את קובץ הנתונים שעליו אתם עובדים: מאיפה אתר הקובץ נלקח (scraped), מי יצרו את קובץ הנתונים במקור וכיצד הנתונים נאספו על ידם. דווחו על כל מידע רלוונטי לאופן יצירת קובץ הנתונים.

2. כיצד בדקתם האם מותר לבצע Scraping מאתר זה?

3. הורידו את הנתונים ושמרו אותם כקובץ CSV. עליכם להגיע לפחות ל-4 משתנים ול-80 רשומות (אפשר בהחלט להוריד יותר). הגישו את הקוד בו השתמשתם בקובץ נפרד.

הוראות הגשה

1. יש להגיש את העבודה מוקלדת בקובץ WORD (David, 12, רווח 1.5 בין שורות, שוליים

סטנדרטיים) + קובץ הקוד בו השתמשתם + קובץ ה-CSV של הנתונים

2. יש לציין בראש העבודה את שמות המגישים ומספרי תעודת הזהות

3. היקף המטלה (WORD) עד חצי עמוד (ללא קובץ הקוד)

4. ההגשה בזוגות (רק אחד מבני הזוג מגיש)

5. ההגשה במודל בלבד עד השעה 23:59 ב-26.11.24

6. שם קובץ ההגשה הינו HW1_ID1_ID2

בהצלחה!