תרגיל בית 1

מיאס גנטוס -213461692

פאיסל עומרי –

שלב 1 :

1. קיבלתי את כל המסמכים הנמצאים באותה תיקיה דרך ואז עשיתי ניתוח על השם של הקובץ ושמרתי אותו ב list of dictionary , כך ש הוא המספר XX,

הוא אומר אם הוא מליאה או ועדה, הוא התוכן ו טקסט שעוזר בתכנות ולמצאו איפה השגיות.

פונקציה get\_all\_docx\_in\_current\_foleder .

1. אם המסמך ועדה:

ראשון אנחנו נמצא את מה הדבר הראשון שידברו עליו כי אנחנו נקלוט טקסט משמה עלה ,עושים את זה על ידי ההופעה הראשונה של 'סדר' ו'יום' באותו שורה ,כי אז זו הסתפרות גבוהה שזה אומר סדר הדיבור של המסמך.

אנחנו משנים את המדבר כאשר מופיע הסימן ":" (יש לנו בעיה בו כי זה אומר שאם אין ":" אז אנחנו לא נשנה את המדבר אפילו אם יש מדבר חדש ואנחנו לא יכולנו להתמודד עם זה אבל לרוב המקרים יש ":" שזה לא ישפיע כל כך על התוצר הסופי).

איך עשיתי ניקוי לשם?

1. אם יש לך סוגריים אז לא תקרא את הסוגריים ומה יש בניהם (בעיה אם השם האמיתי בין הסוגריים).
2. אם יש לנו מה שהוא שמפריד אבל עם רווחים (כמו " – ") אז תקרא רק את החלק הראשון (בעיה אם השם הוא בחלק השני, אם אין רווחים ואם השם עצמו יש בו סמל מפריד עם רווחים).
3. אם יש קיצורים (הקיצורים יהיו תמיד עם הסמל ) אז לא תקרא.
4. לא יכולתי לנקות את התפקיד אם הוא לא קצור והוא אינו בין סוגריים (כמו חבר הכנת במקום הח"כ)
6. ניקוי משפטים:

פשוט עשיתי אם האות לא בעברית או שהיא סמני פיסוק אז תמחק ואם יש משפט שהוא רק סמני פיסוק אז תמחק אותו, ואם יש שני סמני פיסוק צד לצד(אפילו אם יש רווח ביניהם) אז תמחק אותם.