CSC110 Fall 2022 Assignment 3: Loops, Mutation, and Applications

Jaeyong Lee

October 31, 2022

Part 1: Data Analysis with Toronto Health

Complete this part in the provided `a3_part1.py` file. Do **not** include your solutions in this file.

Part 2: Loops and Mutation Debugging Exercise

1. test_star_wars passed. test_legally_blonde and test_transformers failed.

2. test_legally_blonde raised an assertion error caused by an issue in the implementation of the function clean_text. While its job is to remove all punctuation and convert all letters to lower case, the function failed to convert all letters of the review to lowercase due to the way it called str.lower(text) in the middle of the function. Compared to how the function removed punctuation by performing variable reassignment on the variable text, calling str.lower(text) on the immutable string parameter 'text' does not change the value of 'text' at all. So when the function returns str.split(text), the text that it is splitting has not actually been converted to lowercase, although its punctuation has been properly removed thanks to the variable reassignment performed while removing punctuation. This is significant because the VADER word 'Amusing' is used in the review for Legally Blonde. Since it was not properly converted to lower case, the program missed it and it led to Legally Blonde's average intensity score being incorrect.

   test_transformers raised an assertion caused by an issue in the implementation of the function count_keywords. Within the for loop that iterates through the list of words in the review, the conditional statement 'if word not in occurrences_so_far' effectively causes the function to ignore repeated VADER words in the review. The value (AKA the 'count') of a VADER word in occurrences_so_far is only updated if the word is not already a key in occurrences_so_far. Since the review for Transformers repeats the word 'terrible' twice, which is a VADER word, the second use of the word 'terrible' was not counted for by the function and therefore, an incorrect average intensity was calculated as a result.

3. test_star_wars passed in spite of errors in the program because its review is coincidentally written in a way that makes it unharmed by the errors in the program. Firstly, none of its VADER words had a capital letter in it, so the error in the function clean_text, which didn't properly convert the review to all lower case letters, did not have an effect on the review for Star Wars. Secondly, the review for Star Wars includes no repeated VADER words, meaning it was unharmed by the error in the implementation of count_keywords that caused it to ignore repeated VADER words in a review. Hence, the average intensity score for Star Wars was calculated correctly and its test passed in spite of errors in the program.

Part 3: Chaos, Fractals, Point Sequences

Complete this part in the provided `a3_part3.py` starter file. Do **not** include your solutions in this file.