START OF QUIZ
Student ID:
22039382,Amal,Cenith

Question 1

Topic: Lecture 2
Source: Lecture 2

As we increase the size of a corpus, the frequency of Hapax Legomena generally increases. Would the frequency of function words like ”the” or ”is” also increase? Why or why not? (1)

Question 2

Topic: Lecture 3
Source: Lecture 3

Why do we not care about the extra space required to create a reverse index? (2 reasons) (1)

Question 3

Topic: Lecture 2
Source: Lecture 2

If a language has a highly synthetic morphology (many affixes), would you expect it to have a higher or lower Type-Token Ratio (TTR) than a language with less rich morphological structure? Briefly explain why. (1)

Question 4

Topic: Lecture 1
Source: Lecture 1

Why is the .split() method useful when working with sentences or phrases? (1)

Question 5

Topic: Lecture 3
Source: Lecture 3

How does a defaultdict differ from a regular dictionary in Python? (2 differences) (1)

Question 6

Topic: Lecture 4
Source: Lecture 4

Attributive adverbs are a type of adverb that provides ”flavour” to speech verbs (example: ”she said quickly”; ”he spoke loudly”). They are often frowned upon in formal writing, because they can be replaced with other verbs: ”blurted” or ”shouted”, in the example. Write a quick function that finds them in the Brown corpus, and reports how many sentences in 1000 have them. (2)

Question 7

Topic: Lecture 4
Source: Lecture 4

In French, negation is often indicated by ”ne ... pas” (ie, ”je ne parle pas” - ”I am not speaking”; ”tu ne conduis pas” - ”You are not driving”, etc.). However, in speech, one of the two is often dropped: ”je ne parle.” or ”tu conduis pas.”. Using this information, how would you determine whether a corpus was composed of written or spoken French? You don't need to write the code, but explain the logic that you would use to come to this conclusion. (2)

Question 8

Topic: Lecture 1
Source: Lecture 1

Write a function that capitalizes the first letter of each word in a string, without using the .title() method or any external libraries. What are some assumptions that you are making? (2)

Question 9

Topic: Long
Source: Lecture 1

Write a function that validates if a string matches a phone number format, such as (123) 456-7890. What types of invalid inputs should the function check for? Are there edge cases we would be willing to accept? How would we handle those? Write 3 test cases - 2 that should pass, and one that should fail. (3)

END OF QUIZ