

START OF QUIZ

Student ID:

010237352,Shan,Zhengyi

Question 1

Topic: Lecture 3

Source: Lecture 3

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

Question 2

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 3

Topic: Lecture 4

Source: Lecture 4

Why does the lexical diversity (type-to-token ratio) typically increase when analyzing smaller sub-corpora rather than larger ones? What does this suggest about the content of smaller texts? (1)

Question 4

Topic: Lecture 1

Source: Lecture 1

How would you convert a string into a list of characters? (1)

Question 5

Topic: Lecture 2

Source: Lecture 2

What role does linguistic annotation provide for corpora, specifically for computational linguistics? (1)

Question 6

Topic: Lecture 3

Source: Lecture 3

Lexicons are useful for initial text analysis but often lack the adaptability needed for advanced NLP tasks. Why is this the case? Provide at least 2 reasons with brief explanations.
(2)

Question 7

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 8

Topic: Lecture 1

Source: Lecture 1

You are given a sentence. Write a function to count how many words in the sentence start with a vowel, without using loops or list comprehensions. (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