START OF QUIZ
Student ID:
71936512,Jimenez,Daniel

Question 1

Topic: Lecture 3
Source: Lecture 3

Describe the concept of the ”Minimum viable product”, and how it relates to using lexicons. (1)

Question 2

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 3

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 4

Topic: Lecture 2
Source: Lecture 2

Would a language with with lots of inflection have a higher or lower TTR than one with little inflection? Briefly explain. (1)

Question 5

Topic: Lecture 2
Source: Lecture 2

Why is it important to know when a corpus was constructed, and who constructed it? (1)

Question 6

Topic: Lecture 3
Source: Lecture 3

How does ”get” differ from a default dictionary (2 ways)? (1)

Question 7

Topic: Lecture 1
Source: Lecture 1

In class, we talked about how .isdigit() is insufficient for determining whether we can convert a string to a float. Write a short function "isfloat" that determines whether a provided string is a valid floating point number. (2)

Question 8

Topic: Lecture 1
Source: Lecture 1

Why is strip() such a useful function? (1)

Question 9

Topic: Coding
Source: Lecture 2

Imagine we have a large corpus in an unknown language. We don't have any ML tools to analyze the data. How might we determine the stopwords in our corpus? How might we test our theory of stopwords? (I'll make it easy on you - the tokens are space separated, and we have some way of separating sentences.) (3)

END OF QUIZ