Unit 1 Assignment

**ANLY:520-51 (Fall 2016)**

Dean D’souza

H.U. ID: 168424

# Solutions:

1. **Try using the Python interpreter as a calculator, and typing expressions like 12 / (4 + 1).**
2. **Given an alphabet of 26 letters, there are 26 to the power 10, or 26 \*\* 10, ten-letter strings we can form. That works out to 141167095653376. How many hundred-letter strings are possible?**
3. **The Python multiplication operation can be applied to lists. What happens when you type ['Monty', 'Python'] \* 20, or 3 \* sent1?**
4. **Review 1 on computing with language. How many words are there in text2? How many distinct words are there?**
5. **Compare the lexical diversity scores for humor and romance fiction in 1.1. Which genre is more lexically diverse?**
6. **Produce a dispersion plot of the four main protagonists in Sense and Sensibility: Elinor, Marianne, Edward, and Willoughby. What can you observe about the different roles played by the males and females in this novel? Can you identify the couples?**
7. **Find the collocations in text5.**
8. **Consider the following Python expression: len(set(text4)). State the purpose of this expression. Describe the two steps involved in performing this computation.**
9. **Review 2 on lists and strings.**
10. **Define a string and assign it to a variable, e.g., my\_string = 'My String' (but put something more interesting in the string). Print the contents of this variable in two ways, first by simply typing the variable name and pressing enter, then by using the print statement.**
11. **Try adding the string to itself using my\_string + my\_string, or multiplying it by a number, e.g., my\_string \* 3. Notice that the strings are joined together without any spaces. How could you fix this?**
12. **Define a variable my\_sent to be a list of words, using the syntax my\_sent = ["My", "sent"] (but with your own words, or a favorite saying).**
13. **Use ' '.join(my\_sent) to convert this into a string.**
14. **Use split() to split the string back into the list form you had to start with.**
15. **Define several variables containing lists of words, e.g., phrase1, phrase2, and so on. Join them together in various combinations (using the plus operator) to form whole sentences. What is the relationship between len(phrase1 + phrase2) and len(phrase1) + len(phrase2)?**
16. **Consider the following two expressions, which have the same value. Which one will typically be more relevant in NLP? Why?**
17. **"Monty Python"[6:12]**
18. **["Monty", "Python"][1]**
19. **We have seen how to represent a sentence as a list of words, where each word is a sequence of characters. What does sent1[2][2] do? Why? Experiment with other index values.**
20. **(Q 29 in boo) We have been using sets to store vocabularies. Try the following Python expression: set(sent3) < set(text1). Experiment with this using different arguments to set(). What does it do? Can you think of a practical application for this?**