## Week 1 - Auto-Correct and Minimum Edit Distance

1. The minimum edit distance between the words ‘deep’ and ‘creepy’ is:
   1. 4 (You need to replace *d* for *c*, which counts for 2, insert *r* and insert *y*.)
2. What is NOT a valid example of an edit string operation?
   1. SWITCH a letter ‘Lusca’ --> ‘Lucas’ (Switching a letter is a valid operation ONLY when switching adjacent letters. In this case, there were two switches: *s* and *c*, after *s* and *a*.)
3. Autocorrect is only applicable when dealing with misspelled words.
   1. False. (Autocorrect can be used for words that do not make any sense for a particular sentence. For example, ‘Happy birthday deer friends’ is a correctly-spelled sentence, but the word ‘deer’ makes no sense – it should be ‘dear’.)
4. Given the corpus: “I am happy because I am doing quizzes.” and the following sentence: “I sm very good at solving quizzes.”, what is true?
   1. There is a unique correction for the misspelled word “sm”. (The correction would be the word “am”.)
5. What is true about the probabilistic model:
   1. Replacing a character costs more than deleting a character. (Replacing a word costs 2 whereas deleting it costs 1.)
   2. If is the number of times a word appears in a corpus and is the corpus size, then the probability of the word in the corpus is .
   3. The sentence “Happy birthday deer friends” would not have any word corrected. (Since the probabilistic model just looks at misspelled words, the above sentence would not be corrected.)
6. If we build a distance matrix for the following case: Source: Pie --> Target: Bye, what is the value of ?
   1. 5
7. If is the distance matrix for two words of the same size and is the matrix size for the minimum edit distance algorithm, then:
   1. (The first line will always have increasing values as we move to the right, because it is the cost from editing the null string.)
   2. The algorithm avoids use of brute force by implementing a dynamic programming approach. (Using previously computed cells to compute the next one is a dynamic programming method.)
8. What is NOT true about minimum edit distance?
   1. It is used to check if a word is misspelled.
9. The minimum edit distance calculation is more computationally expensive for a large corpus.
   1. False. (The minimum edit distance depends only on the editing cost and the two words that are being considered, not on any corpus or vocabulary.)
10. For the corpus “Autocorrect is a powerful tool and it is used on our computers”, the value for is:
    1. .