Topic:	Python Warmup
Techniques:	Use of Python strings
Reading:	lecture notes
Collaboration Policy:	The lab should be completed **working in pairs**
Submission Deadline:	Wednesday, 30 January 2019

Students: _____

To receive full credit for this lab, you must answer at least 10 of the following questions, and explain briefly what python command(s) you used to determine the answer.

1 How long is this sequence?

2 What is the first basepair of the sequence?

3 What is the 2000th basepair of the sequence (that is, the one with index 1999)?

4 What is the last basepair of the sequence?

5 What are the first 10 characters of the sequence?

6 What are the last 10 characters of the sequence?

7   How many times does the character C appear in the sequence?

8   The GC-content of a sequence is the percentage of basepairs that are either G or C. What is the GC-content of this sequence?

9   The pattern CCAAT is a particular motif known as a "CAT box". How many times does this motif appear in the sequence?

10 What is the index at which the first occurrence of the pattern CCAAT begins?

11 What is the index at which the second occurrence of the pattern CCAAT begins?

12 What is the index at which the last occurrence of the pattern CCAAT begins?

13 Consider initial prefixes of the sequence, such as the first three characters GTT. That particular prefix occurs 164 times in the sequence. What is the shortest prefix that does not occur anywhere else? (Given the techniques we've learned so far, you will likely need to resort to some trial and error.)

14 What is the largest number of consecutive occurrences of A that can be found in the sequence? (Again, lacking more advanced programming techniques, some trial and error can be used.)