Activities: Dictionaries

1. Use the archive HIV.txt, which contains the genome of various HIV-viruses. It is organized as a text files with various lines. The odd lines are short and contain information about the virus. The even lines are very long because they contain the actual genome encoding. Print out on separate lines
   (a) the virus type
   (b) the number of characters in the virus description
   (c) the number of “A”, “C”, “G”, and “T” characters for each virus
   (d) the number of times two consecutive characters “AA”, “AC”, “AG”, “AT”, “CA”, … “TG” and “TT” appear in the genome.

2. “Insecure Securer Password Generation”: Write a function that takes a string that replaces each occurrence of a vowel with another one. You should use a dictionary to implement the substitution. `subdic = {"a": "o", "e": "i", …}`. Remember that you can use `if letter in subdic` in order to ascertain whether there is a substitution for a given letter.

3. Open the file “alice.txt” with encoding `"latin-1"`. Print out all words of length more than five characters that appear more than five times.

4. Open the file “alice.txt”. For all words in the file longer than five characters that appear more than five times, print the word and the list of line-numbers in which they appear.