CIS 4930 NLP // HW #5 // Spring 2018

Date Assigned: February 11, 2018
Date Due: February 16, 2018

Submission Format

You will submit a soft copy of your solution using e-Learning (http://elearning.ufl.edu) by the end of the day (23:59 / 11:59 PM) on the assigned date (February 16). Submit one file, **hw5.py**.

Assignment

At the top of every solution file you submit this semester include: your name, section number, the assignment number, and the date due. Implement the Python to Complete these exercises.

Exercises

1. [3.25 (a) & (b)]: Pig Latin (http://en.wikipedia.org/wiki/Pig_Latin) is a simple translation of English text. Each word of the text is converted as follows: move any consonant (or consonant cluster) that appears at the start of the word to the end, then append ***ay***. For example: string → ingstray, idle → idleay, Voldemort → oldemortvay. Allow the user to enter the string of text to be converted into Pig Latin.

   (a) Implement a Python function to convert a word into Pig Latin.
   (b) Implement a Python function to convert a string of text (a sequence of words) into Pig Latin.

2. Create a function that receives a string, potentially a Gainesville address. Your code will query the user for a string and then call your function with the string entered. Your function will use regular expressions to detect whether the string is a valid Gainesville street address. The function will print to the screen whether or not the address given is valid. For the purposes of this problem, a valid Gainesville street address shall be only those addresses that conform to the following rules.

   a. The address starts with a *number* between **1** and **99999**.

   b. The initial *number* is followed by a single space and then a *directional indicator* which may be any of the following (potentially in lowercase): **N**, **NW**, **NE**, **SW**, **SE**, **S**, **W**, or **E**.

   c. The *directional indicator* is followed by a *numbered street* name, which will consist of a number between **1** and **999** and will be followed by (potentially in lowercase) **ST**, **ND**, **RD**, or **TH**.

d. The *numbered street* name will be followed by a single space and then the *road type* which may be any of the following (potentially in lowercase): **RD**, **AVE**, **DR**, **TERR**, **ST**, or **CIR**.

e. The *road type* is followed by "**, Gainesville, FL 326XX**" where XX can be any number between **00** and **19**.