Assignment 3

3.1

	5	4	3	4	3	2
E	5	4	3	4	3	2
S	4	3	2	3	2	3
U	3	2	1	2	3	4
O	2	1	0	1	2	3
H	1	0	1	2	3	4
#	0	1	2	3	4	5
	#	H	O	R	S	E

R	5	6	5	6	5	4	5
O	4	5	4	5	4	5	6
T	3	4	5	4	5	4	7
U	2	3	4	5	4	5	6
T	1	2	3	4	5	6	7
#	0	1	2	3	4	5	6
#		M	O	T	O	R	S

C	6	5	6	5	6	5	4	3
I	5	4	5	4	5	4	3	4
B	4	3	4	3	4	3	4	5
A	3	2	3	2	3	4	5	6
R	2	1	2	1	2	3	4	5
A	1	0	1	2	3	4	5	6
#	0	1	2	3	4	5	6	7
	#	A	E	R	O	B	I	C

3.2 Write a program to measure edit distance using the nltk library and compare the results.

Code

```python
str1 = input("str1 : ")
str2 = input("str2 : ")

e = edit_distance(str1,str2,2)
print("the edit distance of ",str1," and ",str2," is => ",e)
```

Result

```
D:\NLP>c:/python38/python.exe d:/NLP/test.py
str1 : baby
str2 : ruby
the edit distance of  baby  and  ruby  is =>  4
```

```
D:\NLP>c:/python38/python.exe d:/NLP/test.py
str1 : HOUSE
str2 : HORSE
the edit distance of  HOUSE  and  HORSE  is =>  2
```

```
D:\NLP>c:/python38/python.exe d:/NLP/test.py
str1 : TUTOR
str2 : MOTORS
the edit distance of  TUTOR  and  MOTORS  is =>  5
```

3.3 Write a program to apply edit distance to correct misspelled word on this sentence "I opan a bax ant reed a leter".

Code

```python
import nltk
from nltk import edit_distance
from nltk.tokenize import word_tokenize

f = open('1-1000.txt','r')
wordFromDict = [line.rstrip() for line in f.readlines()]
inputString = "I opan a bax ant reed a leter"

listString = word_tokenize(inputString)
newString = ""

for i in listString :
    minVal = 100
    word = ""
    for j in wordFromDict :
        val =  edit_distance(i,j)
        if val < minVal:
            minVal = val
            word = j
    newString += word + " "
    print(f"{i} can be replaced by ('{word}', {minVal})")

print("\nMispelled sentence is >> ",inputString)
print("New sentence is        >> ",newString)
```

Result

```
D:\NLP>c:/python38/python.exe d:/NLP/3-3.py
I can replaced by ('I', 0)
opan can replaced by ('open', 1)
a can replaced by ('a', 0)
bax can replaced by ('box', 1)
ant can replaced by ('and', 1)
reed can replaced by ('read', 1)
a can replaced by ('a', 0)
leter can replaced by ('letter', 1)

Mispelled sentence is >>  I opan a bax ant reed a leter
New sentence is        >>  I open a box and read a letter
```