**Assignment 1.1**

1. Every word ends with ‘ess’

|  |
| --- |
| **Code** |
|  |
| **ผลการทำงาน** |
|  |

2. Every word contains substring ‘pen’ or ‘men’

|  |
| --- |
| **Code** |
|  |
| **ผลการทำงาน** |
|  |

3. Every word contains substrings started by ‘un’ or ‘pre’

|  |
| --- |
| **Code** |
|  |
| **ผลการทำงาน** |
|  |

4. Every word presents only 2-digit numbers

|  |
| --- |
| **Code** |
|  |
| **ผลการทำงาน** |
|  |

5. Every word with 4 length

|  |
| --- |
| **Code** |
|  |
| **ผลการทำงาน** |
|  |

**Assignment 1.2**

Students have to list the frequency of the tokens appearing in the data. Results should be sorted decently, and the length of the token should be over 2 alphabets. The number of sentences must present. Note that, stop words will not included. The lemmatization and stemming are not necessary.

|  |
| --- |
| **News.txt** |
|  |
| **Code** |
|  |
| **ผลการทำงาน** |
|  |