Introduction-to-Information-Retrieval-and-Text-Mining, HW 2

B07302230, 財金三 王博奕

1. 執行環境：Visual Studio Code
2. 程式語言：Python 3.8
3. 執行方式：
   1. Pip install nltk以便於接下來使用Porter’s algorithm。
   2. Pip install numpy 以計算consine similarity。
   3. Pip install math 以計算idf與tf-idf。
4. 作業處理邏輯說明：

本次作業我將接續上次作業所作的tokenize繼續往下做，但在本次作業中我多加了

我把所有.的都刪掉，所以像是一些專有名詞的縮寫會被我刪掉，例如U.S.