

Name:sondos ali elsayed

Id: 20221442482

## Assignment 2:

### 1- Python code:

```
import nltk
from nltk.corpus import stopwords
from collections import Counter

# Download NLTK stopwords if not already downloaded
nltk.download('stopwords')

def process_text(file_path):
    # Read the contents of the file
    with open(file_path, 'r') as file:
        text = file.read()

    # Tokenize the text
    words = nltk.word_tokenize(text)

    # Remove stopwords
    stop_words = set(stopwords.words('english'))
    filtered_words = [word.lower() for word in words if word.lower()
not in stop_words]

    # Count the frequency of each word
    word_freq = Counter(filtered_words)

    return word_freq
```

```
def display_word_frequency(word_freq):  
    # Display word frequency count  
    for word, freq in word_freq.items():  
        print(f"{word}: {freq}")  
  
if __name__ == "__main__":  
    file_path = "random_paragraphs.txt"  
    word_freq = process_text(file_path)  
    display_word_frequency(word_freq)
```

## 2- create the Dockerfile to package this script and its dependencies into a Docker image:

```
# Use the official Python image as a base image  
FROM python:3.9  
  
# Set the working directory in the container  
WORKDIR /app  
  
# Copy the current directory contents into the container at /app  
COPY . /app  
  
# Install NLTK and download stopwords  
RUN pip install nltk && \  
    python -m nltk.downloader stopwords  
  
# Run the Python script when the container launches  
CMD ["python", "your_script_name.py"]
```

To build the Docker image, navigate to the directory containing your Dockerfile and run:

`docker build -t word-frequency .`

-For example, if the text in "random\_paragraphs.txt" contains:

```
This is a sample text with some random words. This text is used for demonstratio
```

-the output displayed in the console after running the Docker container would be something like:

```
sample: 1
text: 2
random: 1
words: 1
used: 1
demonstration: 1
purposes: 1
```

repository link:

<https://github.com/Sondosali/sondos>