'''

The site SongLyrics contains the lyrics of songs.

In this exercise we scrape the words of a desired song.

with request package

Much of the main functionalities of the module are implemented through the class response, whose properties give

access to the HTTP response.

'''

import requests

artist = 'the-beatles'

song = 'two-of-us'

url = f'http://www.songlyrics.com/{artist}/{song}-lyrics/'

resp = requests.get(url)

text = resp.content.decode('utf-8').split('\n')

# After the text is extracted, it can be used regardless of its source. We note that the lyrics of the song always apear in a parragraph section with the id 'songLytricsDic',

# and they are ended by the term \_div\_.

# We can collect the lines that are part of the song.

song = []

song\_line = False

lyrics\_part\_identifier = '<p id="songLyricsDiv" class="songLyricsV14 iComment-text">'

for line in text:

if lyrics\_part\_identifier in line:

line = line.replace(lyrics\_part\_identifier, '')

song\_line = True

if song\_line:

if line.strip() == '</div>':

break

song.append(line.strip())

print(song)

# We remove the countless 'br' occurrences (and single < /p> occurrence) with a simple replace().

song = [line.replace('<br />','').replace('</p>','') for line in song]

# We conclude by filtering out lines with no text (originally contained a single br occurrence)

song = list(filter(len, song))

lyrics = '\n'.join(song)

print(lyrics)