

Лабораторная работа № 3

Перед началом выполнения каждой лабораторной работы, необходимо выполнить код в двух ячейках ниже. В случае если модуль `datetime` не установлен нужно установить. В последнем принте, вывести свою фамилию и инициалы.

In [1]:

```
!whoami
```

rogozhina\home

In [3]:

```
from datetime import datetime

current_time = datetime.now()
print(current_time)
print("Rogozhina N.A.") # написать здесь свою фамилию и инициалы
```

2022-12-13 20:10:19.071539

Rogozhina N.A.

Задания

Задание 1. Создайте файл с расширением `.txt`, содержащий несколько чисел, записанных через пробел. Напишите программу или функцию, которая подсчитывает и выводит на экран общую сумму чисел, хранящихся в этом файле.

In [2]:

```
with open('summa.txt', 'w', encoding = "utf-8") as f:
    f.write('35 45 1049')

with open('summa.txt', 'r') as f:
    a = f.read()
    a_split = a.split()
    A = 0
    print(a_split)
    for i in a_split:
        A += int(i)
print(A)
```

['35', '45', '1049']

1129

Задание 2. Напишите функцию `read_file`, которая принимает имя файла, и печатает его содержимое. Учитывайте, что содержимое файла может быть как на русском языке, так и на английском (т.е. не забываем про кодировку).

In [4]:

```
def read_file(file):
    with open(file, 'r', encoding = "utf-8") as f:
        for line in f.readlines():
            print(line)
read_file('test.txt')
```

test Hello, World!

Задание 3.

3.1. Напишите цикл, который выводит только запросы, содержащие набор букв «эпл». Внимательно просмотрите полученный список и найдите лишние слова.

```
queries = ["эпл айфоны", "купить эпл телефон", "лучшие смартфоны", "барон фон", "смартфон эпл айфон", "смартфоны 2019", "эплан", "фононая музыка", "эпл айфоны икс", "эпл айфон 64гб", "фон для фото", "купить эпл", "эпл айфон купить", "эплеренон купить", "смартфон где купить", "эплан показания", "смартфон huawei", "эпл"]
```

3.2. Проведите стемминг в списке `words`. Результат выведите на экран в следующем виде:

Исходное слово - ..., после стемминга - ...

3.3 Сформируйте список запросов с «эпл».

In [33]:

```
# 3.1
queries = ["эпл айфоны", "купить эпл телефон", "лучшие смартфоны", "барон фон", "смартфон эпл айфон", "смартфоны 2019", "эплан", "фононая"]
for query in queries:
    if 'эпл' in query:
        print(query)
```

эпл айфоны
купить эпл телефон
смартфон эпл айфон
эплан
эпл айфоны икс
эпл айфон 64гб
купить эпл
эпл айфон купить
эплеренон купить
эплан показания
эпл

In [3]:

```
# 3.2
from nltk.stem import SnowballStemmer
russian_stemmer = SnowballStemmer('russian')
words = ["эпл", "эплан", "эплеренон"]
for word in words:
    print(f"Исходное слово - {word}. После стемминга - {russian_stemmer.stem(word)}")
```

Исходное слово - эпл. После стемминга - эпл
Исходное слово - эплан. После стемминга - эпла
Исходное слово - эплеренон. После стемминга - эплеренон

In [42]:

```
# 3.3
queries = ["эпл айфоны", "купить эпл телефон", "лучшие смартфоны", "барон фон", "смартфон эпл айфон", "смартфоны 2019", "эплан", "фононая"]
words = ["эпл", "эплан", "эплеренон"]
apple = []
all_q = queries + words
for query in all_q:
    q = russian_stemmer.stem(query)
    for word in q.split(' '):
        if word == 'эпл':
            apple.append(query)
print(apple)
```

['эпл айфоны', 'купить эпл телефон', 'смартфон эпл айфон', 'эпл айфоны икс', 'эпл айфон 64гб', 'купить эпл', 'эпл айфон купить', 'эпл', 'эпл']

In []: