

**САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ**

**Факультет Программной Инженерии и Компьютерной Техники**

**Лабораторная работа №3  
по дисциплине  
"Технологии программирования"**

**Выполнили:**

Студенты гр. Р3411

Ефремов Р.В.,

Леванович А.А.

**Преподаватели:**

Оголюк А.А.

# Код

```
#!/usr/bin/python2
# -*- coding: utf-8 -*-
import sys
from operator import itemgetter
from bs4 import BeautifulSoup
import re
# Вход: nameYYYY.html, Выход: список начинается с года, продолжается имя-ранг в алфавитном
# порядке.
# '2006', 'Aaliyah 91', Aaron 57', 'Abigail 895', ' и т.д.
def extr_name(filename):
    res = []
    year = filename[4:9]
    with open(filename, "r") as file:
        html_str = file.read()
        soup = BeautifulSoup(html_str, 'html.parser')
        soup = soup.find_all("table")[2]
        trs = soup.find_all("tr")
        for tr in trs[2:-1]:
            ths = tr.find_all("td")
            rank = ths[0].contents[0]
            male_name = ths[1].contents[0]
            female_name = ths[2].contents[0]
            res.append(str(male_name) + " " + str(rank))
            res.append(str(female_name) + " " + str(rank))
    res.sort()
    return [year] + res
def cut_num_at_end(name_with_num):
    m = re.search('(.+) \d+', name_with_num)
    return m.group(1)
def main():
    args = sys.argv[1:]
    if not args:
        print('use: [--file] file [file ...]')
        sys.exit(1)
    name_in_files_num = {}
    # для каждого переданного аргументом имени файла, вывести имена extr_name
    for arg in args:
        result = extr_name(arg)
        names_list = [cut_num_at_end(x) for x in result[1:]]
        for name in names_list:
            name_in_files_num[name] = name_in_files_num.get(name, 0) + 1
        print("for file " + arg + " names are: " + ", ".join(names_list))
    # напечатать ТОП-10 муж и жен имен из всех переданных файлов
    names_corteges_list = [(name, name_in_files_num[name]) for name in name_in_files_num]
    names_corteges_list.sort(key=itemgetter(0))
    names_corteges_list.sort(key=itemgetter(1), reverse=True)
    print("TOP-10 names (Num in files, Name):")
    print("\n".join([str(c[1]) + ", " + c[0] for c in names_corteges_list]))
if __name__ == '__main__':
    main()
```