**Домашнее задание к лекции 5.«Regular expressions»**

Иногда при знакомстве мы записываем контакты в адресную книгу кое-как с мыслью, что "когда-нибудь потом все обязательно поправим". Копируем данные из интернета или из смски. Добавляем людей в разных мессенджерах. В результате получается адресная книга, в которой совершенно невозможно кого-то нормально найти: мешает множество дублей и разная запись одних и тех же имен.

Кейс основан на реальных данных из <https://www.nalog.ru/opendata/>, <https://www.minfin.ru/ru/opendata/>

Ваша задача: починить адресную книгу, используя регулярные выражения.  
Структура данных будет всегда:  
lastname,firstname,surname,organization,position,phone,email  
Предполагается, что телефон и e-mail у человека может быть только один.  
Необходимо:

1. поместить Фамилию, Имя и Отчество человека в поля lastname, firstname и surname соответственно. В записной книжке изначально может быть Ф + ИО, ФИО, а может быть сразу правильно: Ф+И+О;
2. привести все телефоны в формат +7(999)999-99-99. Если есть добавочный номер, формат будет такой: +7(999)999-99-99 доб.9999;
3. объединить все дублирующиеся записи о человеке в одну.

from pprint import pprint

# читаем адресную книгу в формате CSV в список contacts\_list

import csv

with open("phonebook\_raw.csv") as f:

rows = csv.reader(f, delimiter=",")

contacts\_list = list(rows)

pprint(contacts\_list)

# TODO 1: выполните пункты 1-3 ДЗ

# ваш код

# TODO 2: сохраните получившиеся данные в другой файл

# код для записи файла в формате CSV

with open("phonebook.csv", "w") as f:

datawriter = csv.writer(f, delimiter=',')

# Вместо contacts\_list подставьте свой список

datawriter.writerows(contacts\_list)

Домашнее задание сдается ссылкой на репозиторий [BitBucket](https://bitbucket.org/) или [GitHub](https://github.com/)

Не сможем проверить или помочь, если вы пришлете:

* архивы;
* скриншоты кода;
* теоретический рассказ о возникших проблемах.