



работа на уроке (10.04)

1. У вас есть датасет с языками и их кодами в формате CSV (`language_codes.csv`). Каждая строка состоит из значений, разделенных запятыми:

- столбец `alpha2` : двухбуквенный код языка;
- столбец `English` : английское название языка.

Например, `av,Avaric` .

Напишите программу, которая принимала бы на вход имя и фамилию человека на латинице, а возвращала язык, код которого совпадает с первыми буквами имени и фамилии.

Ввод	Вывод
<code>>>> Kate Moss</code>	<code><<< Central Khmer</code>
<code>>>> Federico Fellini</code>	<code><<< Fulah</code>

2. У вас есть корпус с репликами из диснеевских мультфильмов в формате TSV (`princess_corpus.tsv`). Каждая строка состоит из значений, разделенных знаком табуляции:

- столбец `Disney_Period` : один из трех периодов Диснея `EARLY / MID / LATE` ;
- столбец `Text` : реплика;
- столбец `Speaker_Status` : один из трех вариантов `PRINCE / PRINCESS / NON-P` (реплика принца, принцессы или другого персонажа);
- столбец `Movie` : английское название мультфильма;
- столбец `Speaker` : английское название героя;
- столбец `Year` : год выпуска мультфильма;
- столбец `UTTERANCE_NUMBER` : номер реплики для каждого мультфильма;

Например, `EARLY yes, your majesty. NON-P Snow White guard 1937 15` .

- a. Переведите TSV формат в CSV и сохраните как `princess_corpus.csv`.
- b. Посчитайте, сколько реплик за фильм произносит каждая принцесса и сколько реплик в среднем произносят принцессы каждого периода.