



работа на уроке (14.04)

1. 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`.

- а. Переведите TSV формат в CSV и сохраните как `princess_corpus.csv`.
 - б. Посчитайте, сколько реплик за фильм произносит каждая принцесса и сколько реплик в среднем произносят принцессы каждого периода.
 - в. Составьте для каждого мультфильма словарь, где будут записаны реплики каждого героя. Кто произносит больше всего реплик? А кто — больше всего слов?
2. У вас есть датасет супергероев (`SuperHeroes.csv`).

- а. Каких героев больше: злых или добрых? В этом вам поможет `Alignment`.
- б. Найдите среднее значение роста (`Height`) всех плохих людей.
NB! Обратите внимание, что у части персонажей рост может быть не указан.
- в. Пользователь вводит *имя супергероя* и интересующий *показатель* (ака название столбца, например `Height`), и печатать в ответ найденную информацию на

пересечении строки супергероя и столбца признака в случае, если поиск оказался успешен. Если искомый супергерой не нашёлся или значение показателя неизвестно нужно печатать в ответ строку "*Не найдено*".

NB! Имя стоит искать не по полному совпадению, а по подстроке.

- d. Напишите программу, которая принимает три параметра, задаваемых пользователем, и возвращает список супергероев, подходящих под заданные критерии.

NB! Обратите внимание, что числовые параметры могут задаваться не только точными (`150`), но и с плюсом (`150+` — 150 или больше) или минусом (`150-` — 150 или меньше).

Входные данные

```
Введите критерий 1: Height
Введите значение: 190+
Введите критерий 2: Alignment
Введите значение: good
Введите критерий 3: Publisher
Введите значение: Marvel Comics
```

Выходные данные

```
Agent Zero
Corsair
Cyclop
Hawkeye
Ronin
Shatterstar
Toxin
Valkyrie
Vision
Vision II
Ardina
Box III
Deathlok
...
```

(это не весь вывод по запросу из входных данных)