Практическое задание 3

В файле filmdeathcounts.csv сохранены как типичные характеристики фильмов, такие как год выпуска (Year), рейтинг MPAA (MPAA_Rating), жанр (Genre), режиссёр (Director), продолжительность в минутах (Length_Minutes), так и довольно необычная — число смертей героев на экране (Body_Count). Подробности об этом наборе данных можно почитать на странице с источником данных.

Задача 1

Загрузите данные из файла filmdeathcounts.csv.

Задача 2

Запросите help по функции str_count(). Воспользуйтесь этой функцией, чтобы посчитать число символов | в каждой ячейке в столбце Genre. Посчитайте максимальное число символов | в этом столбце и определите, чему равно максимальное число указанных жанров. Эта информация поможет вам понять, на сколько столбцов нужно будет разбить столбец с жанрами, чтобы получить жанры по отдельности.

Примечание. Функция str_count() будет воспринимать символ | как специальный оператор «или» в регулярных выражениях. Чтобы такого не происходило, его нужно экранировать – поставить перед ним \\, чтобы R относился к этому символу как к обычному символу в строке. Такую же операцию нужно будет выполнять, например, при подсчёте точек или круглых и квадратных скобок.

Задача 3

Используя информацию, полученную в задаче 2, разбейте столбец **Genre** на несколько столбцов так, чтобы каждый жанр был в отдельном столбце. Назовите новые столбцы по своему усмотрению.