

Университет ИТМО

Практическая работа №1
по дисциплине «Визуализация и моделирование»

Автор: Ефимов Павел Леонидови

Поток: ВИМ 1.1

Группа: К3220

Факультет: ИКТ

Преподаватель: Чернышева А.В.

Санкт-Петербург, 2021 г.

Датасет: <https://www.kaggle.com/tmdb/tmdb-movie-metadata>

Датасет содержит 5 тысяч фильмов с TMDb. В датасете собраны такие данные как: бюджет, компания производитель, дата выхода, прибыль, средний балл.

Название столбца	Данные, хранящиеся в столбце	Тип данных
budget	Бюджет фильма	Целое число
genres	Жанры	Строка
homepage	Сайт фильма	Строка
id	Номер фильма	Целое число
keywords	Ключевые слова	Строки
original_language	Язык оригинала	Строка
original_title	Оригинальное название	Строка
overview	Описание	Строка
popularity	Популярность	Число с плавающей точкой
production_companies	Компания производитель	Строка
production_countries	Страна производитель	Строка
release_date	Дата выхода	Дата
revenue	Прибыль	Целое число
runtime	Время показа на экране	Целое число
spoken_languages	Язык озвучки фильма	Строка
status	Статус	Строка
tagline	Слоган	Строка
title	Название	Строка
vote_average	Средняя оценка	Число с плавающей точкой
vote_count	Количество голосов	Целое число

Задачи решаемые датасетом:

1. Нахождение самых популярных: серий фильмов, режиссеров, жанров
2. Определение лучшей даты выхода фильма
3. Вывод зависимости от средней оценки, жанра и режиссера фильма на прибыль
4. Нахождение самых популярных, прибыльных студий