

Аналитический отчет по данным Netflix

Структура отчета:

1. Введение
2. Описание данных
3. Этапы анализа
4. Визуализации и результаты/рекомендации

1. Введение

Главная цель исследования – провести детальный разбор данных Netflix для определения основных трендов и ключевых особенностей, которые определяют стратегию платформы и поведение ее подписчиков.

2. Описание данных

- Файл: `netflix_titles.csv`
- Количество строк: 8807
- Количество колонок: 12
- Список колонок:
 - `Show_id` - номер в списке
 - `Type` - тип (Movie/TV Show)
 - `Title` - название
 - `Director` - продюссер
 - `Cast` - актерский состав
 - `Country` - страна выпуска
 - `Date_added` - дата добавления на платформу
 - `Release_year` - год выпуска
 - `Rating` - возрастной рейтинг
 - `Duration` - длительность (минуты/сезоны)
 - `Listed_in` - жанры
 - `Description` - описание

3. Этапы анализа

1. Предобработка данных

Цель: Подготовить данные для дальнейшего анализа и визуализации.

Выполненные действия:

- Преобразовано значение `date_added` в формат даты и времени для корректной работы с временными данными.
- Удалены пропуски в критически важных столбцах для повышения качества анализа.
- Добавлен столбец `release_decade`, содержащий десятилетие выпуска контента.
- Добавлен столбец месяца выпуска для анализа сезонности добавления контента.
- Разделена продолжительность (`duration`) на:
 - `duration_value` – числовое значение (кол-во минут или сезонов);
 - `duration_type` – тип продолжительности (минуты или сезоны).

2. Базовый анализ

Цель: Получить общее представление о структуре и распределении контента на платформе.

Выполненные действия:

- Определено общее количество фильмов и сериалов в датасете.
- Определён год с наибольшим количеством выпущенных фильмов.
- Найден год, в котором на платформу было добавлено больше всего контента.
- Составлен Топ-10 стран по объёму контента.
- Построено распределение контента по рейтингам (например, G, PG, R и т.д.).

3. Продвинутый анализ

Цель: Выявить ключевые закономерности и лидеров по различным параметрам контента.

Выполненные действия:

- Определен режиссер с наибольшим количеством фильмов и сериалов на платформе.
- Составлен ТОП-10 режиссеров с наибольшим количеством работ.
- Определен актер, который чаще всего встречается в актерском составе.
- Составлен ТОП-10 популярных актеров, наиболее часто появляющихся в контенте.
- Выявлены топ-5 самых популярных жанров по количеству фильмов и сериалов.
- Рассчитана средняя продолжительность фильмов (в минутах).
- Определено среднее количество сезонов для телесериалов.

4. Визуализация данных

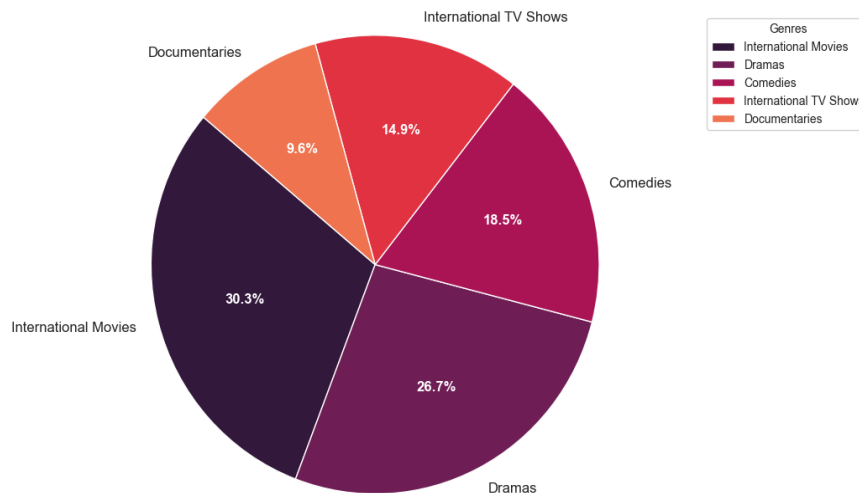
Цель: Наглядно представить ключевые показатели и распределения контента на платформе.

Выполненные действия:

- Гистограмма распределения по рейтингам (например, G, PG, R) для понимания возрастной структуры контента.
- Диаграмма количества фильмов и сериалов для визуального сравнения категорий контента.
- Pie chart топ-5 жанров, позволяющий наглядно оценить популярные направления фильмов и сериалов.
- Barplot топ-10 актеров, которые чаще всего встречаются в актерских составах.
- Barplot топ-10 режиссеров, с наибольшим количеством работ на платформе.
- Barplot топ-10 стран по объему выпущенного контента, показывающий географическое распределение производства.
- Histogram распределения фильмов по длительности, для анализа продолжительности и выявления трендов.

4. Визуализации и результаты

1. Топ-5 самых популярных жанров на Netflix (по количеству контента)



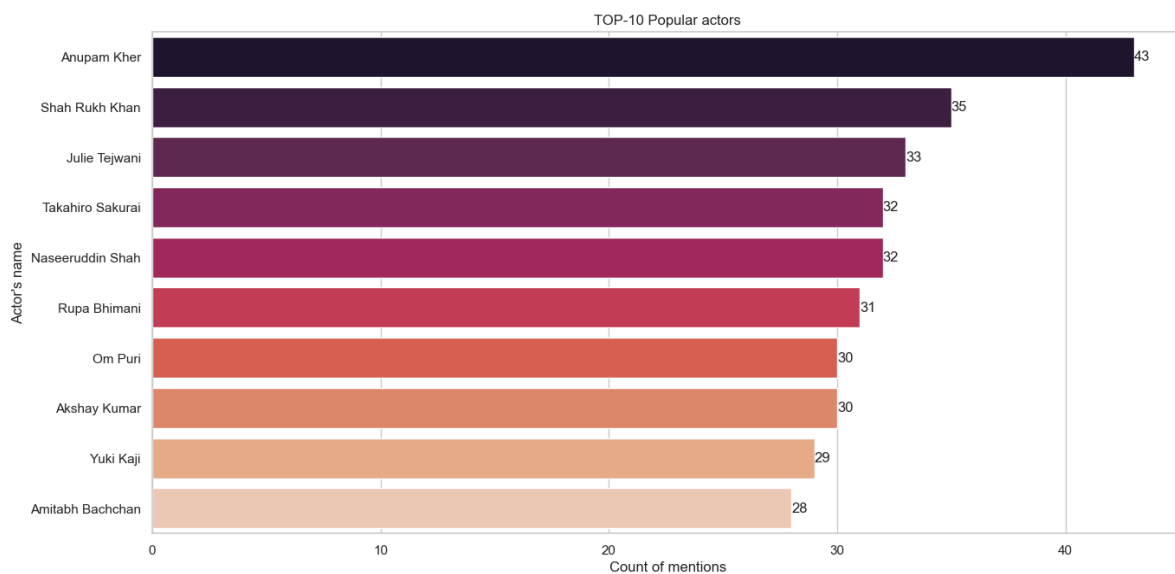
Основные наблюдения:

- Наибольшая доля приходится на International Movies (30,3%), что показывает сильное международное направление платформы.
- Dramas (26,7%) и Comedies (18,5%) также занимают значительную часть, что свидетельствует о высокой популярности этих жанров среди зрителей.
- International TV Shows (14,9%) и Documentaries (9,6%) занимают меньшие сегменты, что указывает на узкую, но стабильную аудиторию.

Рекомендации / инсайты для бизнеса:

- Основной фокус контента можно делать на международные фильмы и драмы, так как они формируют большую часть статистики.
- Комедии остаются значимой категорией для удержания широкой аудитории.
- Документальные фильмы и международные сериалы – перспективная ниша для продвижения, можно выделять их в отдельные подборки или акции.

2. Топ-10 самых популярных актеров (Barplot)



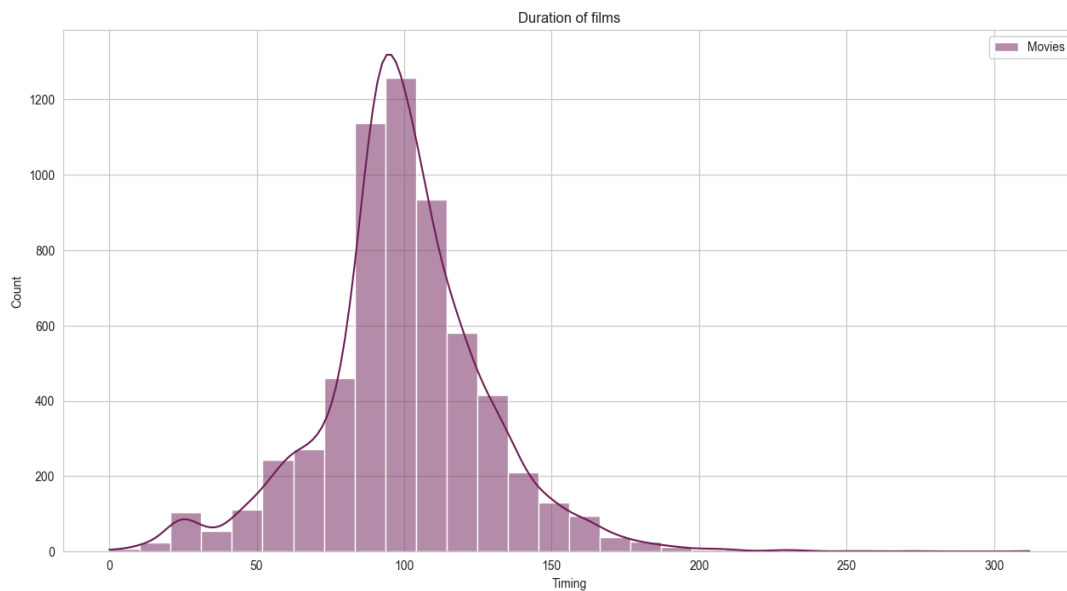
Основные наблюдения:

- В топ-10 актеров в основном индийские актеры – Netflix активно продвигает контент для азиатского рынка.
- Явный лидер: Аннупам Кхер (Anupam Kher) с значительным отрывом возглавляет рейтинг с 43 упоминаниями. Это более чем на 20% выше, чем у следующего за ним актера.
- Высокая концентрация узнаваемости: Первая тройка актеров (Аннупам Кхер, Шах Рукх Кхан, Джули Теджвани) формирует отдельную группу лидеров с результатами выше 35 упоминаний.
- Международное представительство: В топ-10 присутствуют не только индийские актеры (что ожидаемо, судя по именам), но и японские (Такахиро Сакурай, Юки Кадзи), что указывает на разнообразие аудитории или контента в исследуемой выборке.

Рекомендации / инсайты для бизнеса:

- Фокус на регионы: высокий процент индийских актеров = Netflix целится в Азию. Это значит, что для роста на других рынках (например, Европа, ЛатАм) стоит увеличить локальный контент.
- Фокус на лидерах: При планировании маркетинговых кампаний, продвижения контента или партнерств рекомендуется в первую очередь рассматривать сотрудничество с Аннупамом Кхером, Шах Рукх Кханом и Джули Теджвани, так как их имена гарантированно привлекут наибольшее внимание целевой аудитории.
- Учет международной аудитории: Наличие японских актеров озвучивания (Такахиро Сакурай, Юки Кадзи) в топ-10 – важный инсайт. Это прямое указание на наличие активной и заинтересованной аудитории, знакомой с аниме или японским медиа-контентом. Рекомендуется изучить эту целевую группу подробнее и учитывать ее предпочтения.

3. Распределение длительности фильмов



Основные наблюдения :

- Распределение длительности фильмов явно не нормальное, с выраженным правосторонним смещением.
- Подавляющее большинство фильмов (мода) длится около 100 минут.
- Существует "длинный хвост" — небольшое количество фильмов со длительностью более 150-200 минут

Рекомендации / инсайты для бизнеса:

- Формат контента: При создании или приобретении контента основной фокус должен быть на фильмах стандартной длины (~100 мин.), так как они являются самым популярным и ожидаемым форматом у аудитории.
- Платформенные стратегии: Длинные фильмы (>150 мин.) – это нишевый продукт. Их стоит продвигать для аудитории, целенаправленно ищущей "киноэпопеи" или авторское кино, а не для массового потребления.
- Пользовательский опыт: Для сервисов потокового вещания это прямое указание на необходимость функции паузы или "продолжить просмотр", особенно для фильмов из "длинного хвоста".