

DATABASES PROJECT

NETFLIX TV SHOWS AND MOVIES

SATTAROV MUKHAMED,
SHUMKARBEKOV NURSULTAN,
NURDINOV BAYAMAN,
RAIMZHANOV AKHMATAMIN

Databases



PROJECT NAME:

"Netflix Content Analytics Platform"

— an analytical system for researching the Netflix catalog.

BUSINESS/EDUCATIONAL TASKS :

- 1) Optimizing Netflix's Content Investment Strategy
- 2) Understanding Regional Content Demand
- 3) Enhancing Personalization and Recommendation Systems
- 4) Identifying Catalogue Gaps and Opportunities



DATABASES PROJECT



TWO DATASETS ARE USED:

- **titles.csv** — metadata about movies and TV series: **titles, years, descriptions, countries, genres, IMDb/TMDB ratings.**
- **credits.csv** — actors, directors, roles.



CHARACTERISTIC:

- **One title — many genres.**
- **One title — many countries.**
- **One title — many contributors (actors, directors, authors).**
- **One person contributes to many titles.**



Dataset Description (Kaggle — Netflix TV Shows and Movies)

DATABASES PROJECT

1. OPTIMIZING NETFLIX'S CONTENT INVESTMENT STRATEGY

PROBLEM: NETFLIX INVESTS HEAVILY IN PRODUCING MOVIES AND TV SHOWS, BUT NOT ALL GENRES OR FORMATS GENERATE EQUAL USER ENGAGEMENT.

- BUSINESS QUESTION: WHICH CONTENT TYPES (GENRES, COUNTRIES OF ORIGIN, AGE RATINGS) DRIVE THE HIGHEST ENGAGEMENT AND SUBSCRIBER GROWTH?

GOAL: IDENTIFY THE BEST-PERFORMING CONTENT CATEGORIES TO GUIDE PRODUCTION AND ACQUISITION STRATEGIES.

2. UNDERSTANDING REGIONAL CONTENT DEMAND

PROBLEM: NETFLIX SERVES A GLOBAL AUDIENCE WITH HIGHLY DIVERSE PREFERENCES.

BUSINESS QUESTION: WHICH TYPES OF CONTENT SHOULD BE PROMOTED, LOCALIZED, OR PRIORITIZED IN DIFFERENT REGIONS TO MAXIMIZE ENGAGEMENT?

GOAL: BUILD TAILORED MARKETING AND LOCALIZATION STRATEGIES FOR SPECIFIC COUNTRIES OR REGIONS.



DATABASES PROJECT

3. ENHANCING PERSONALIZATION AND RECOMMENDATION SYSTEMS

PROBLEM: EFFECTIVE PERSONALIZATION IS KEY TO KEEPING SUBSCRIBERS ACTIVE AND REDUCING CHURN.



BUSINESS QUESTION: HOW CAN METADATA FROM THE CATALOGUE (GENRES, KEYWORDS, CAST, COUNTRIES, MATURITY RATINGS) IMPROVE RECOMMENDATION ALGORITHMS?

GOAL: INCREASE AVERAGE WATCH TIME AND CUSTOMER SATISFACTION THROUGH BETTER RECOMMENDATION ACCURACY.

4. IDENTIFYING CATALOGUE GAPS AND OPPORTUNITIES

PROBLEM: NETFLIX MAY LACK CONTENT IN HIGH-DEMAND CATEGORIES COMPARED TO MARKET EXPECTATIONS.

BUSINESS QUESTION: WHICH GENRES, REGIONS, OR AGE GROUPS ARE UNDERREPRESENTED AND SHOULD BE EXPANDED TO ATTRACT NEW USERS?

GOAL: DISCOVER “CONTENT GAPS” AND GUIDE DECISIONS ON NEW CONTENT ACQUISITION OR PRODUCTION.



Column	Type	Description
id	INT (PK)	Internal ID
show_id	VARCHAR	ID from Netflix dataset
title	VARCHAR	Title name
type	VARCHAR	Movie / TV Show
description	TEXT	Content description
release_year	INT	Release year
date_added	DATE	Date added to Netflix
age_certification	VARCHAR	Age rating
runtime	INT	Duration (minutes)
seasons	INT	Number of seasons (for TV Shows)
imdb_id	VARCHAR	IMDb ID
imdb_score	DECIMAL	IMDb rating
imdb_votes	INT	Number of IMDb votes
tmdb_popularity	DECIMAL	TMDB popularity metric
tmdb_score	DECIMAL	TMDB rating

TABLE OF TITLES

Column	Type	Description
person_id	VARCHAR(50) PK	Person identifier
name	VARCHAR	Person's name

TABLE OF PERSONS

Column	Type	Description
credit_id	INT PK	Unique identifier of the credit record
title_id	INT FK	FK → titles.id
person_id	VARCHAR FK	FK → persons.person_id
role	VARCHAR	Role type: "ACTOR", "DIRECTOR", "WRITER",
character_name	VARCHAR	Character name (for actors)

Table of credits

Column	Type	Description
genre_id	INT PK	Unique genre identifier
name	VARCHAR	Genre name (UNIQUE)

TABLE OF GENRES

Column	Type	Description
id	INT PK	Unique ID of the title-genre relation
title_id	INT FK	FK → titles.id
genre_id	INT FK	FK → genres.genre_id

Table of title_genres

Column	Type	Description
id	INT PK	Unique ID of the title-country relation
title_id	INT FK	FK → titles.id
country_id	INT FK	FK → countries.country_id

TABLE OF COUNTRIES

