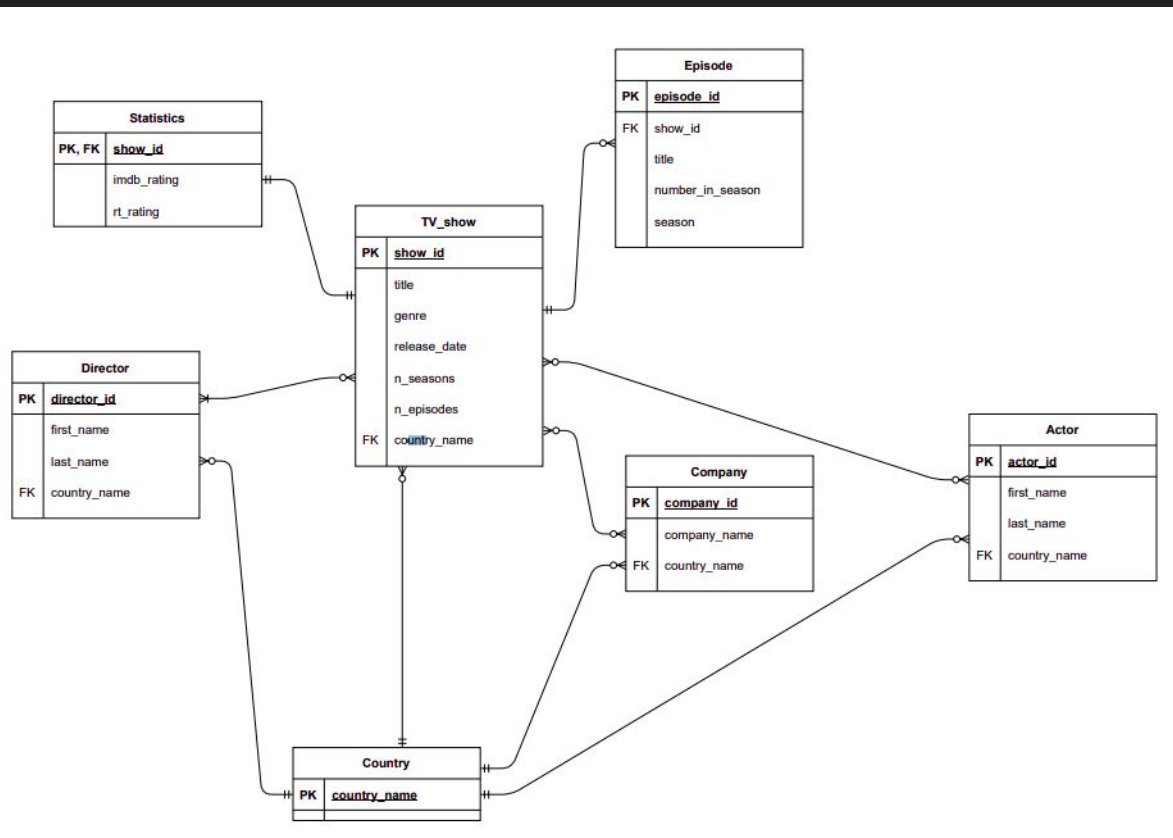


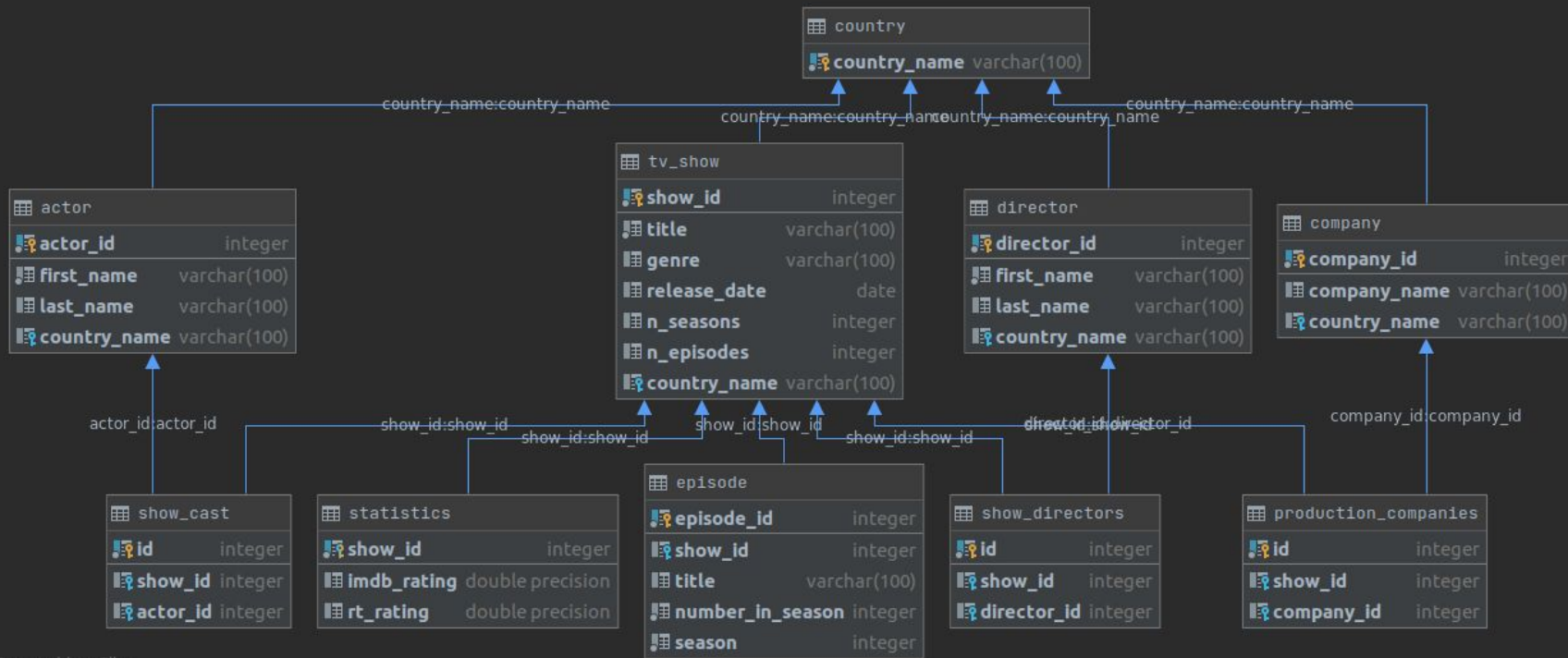
База данных “Сериалы Netflix”

Елизавета Демина

ER диаграмма



TR диаграмма



Создание и заполнение базы данных

	show_id	title	genres	release_year	n_seasons	n_episodes	country
1	1	Orange Is the New Black	{Comedy,Drama}	2013	7	91	USA
2	2	Dark	{Science fiction,Thriller,Myst...	2017	3	26	Germany
3	3	Cable Girls	{Period drama,Telenovela}	2017	5	42	Spain
4	4	The End of the F***ing World	{Dark comedy,Drama}	2017	2	16	UK
5	5	How to Sell Drugs Online (Fast)	{Comedy,Drama,Crime}	2019	2	12	Germany
6	6	Money Heist	{Crime,Drama,Thriller}	2017	2	31	Spain
7	7	The Queen's Gambit	{Period drama}	2020	1	7	USA
8	9	Fauda	{Action,Serial drama,Psycholog...	2015	3	36	Israel
9	10	Silver Spoon	{Crime drama,Psychological thr...	2015	3	40	Russia
10	12	Mindhunter	{Crime drama,Psychological thr...	2017	2	19	USA
11	17	Black Mirror	{Science fiction,Satire,Thrill...	2011	5	22	UK

	country_name
1	USA
2	Germany
3	Spain
4	UK
5	Israel
6	Russia

country

tv_show

- DDL скрипт: <https://github.com/de1iza/netflix-db/blob/main/ddl.sql>
- Заполнение данными: https://github.com/de1iza/netflix-db/blob/main/insert_data.sql

Заполнение базы данных

	company_id	company_name	country_name
1	1	Bambú Producciones	Spain
2	2	Lionsgate Television	USA
3	3	Wiedemann & Berg Television	Germany
4	4	Clerkenwell Films	UK
5	5	Dominic Buchanan	UK
6	8	bildundtonfabrik	Germany
7	9	Vancouver Media	Spain
8	10	Atresmedia	Spain
9	11	Flitcraft Ltd	USA
10	12	Wonderful Films	USA
11	13	Yes	Israel
12	14	Sreda	Russia
13	15	Denver and Delilah Productions	USA
14	16	Zeppotron	UK
15	17	House of Tomorrow	UK

company

	actor_id	first_name	last_name	country_name
1	1	Taylor	Schilling	USA
2	2	Laura	Prepon	USA
3	3	Louis	Hofmann	Germany
4	4	Maja	Schöne	Germany
5	5	Blanca	Suárez	Spain
6	6	Alex	Lawther	UK
7	7	Maximilian	Mundt	Germany
8	8	Úrsula	Corberó	Spain
9	9	Anya	Taylor-Joy	USA
10	10	Lior	Raz	Israel
11	11	Pavel	Priluchnyy	Russia
12	12	Jonathan	Groff	USA
13	13	Bryce Dallas	Howard	USA

actor

Заполнение базы данных

	episode_id	show_id	title	number_in_season	season
1	1	1	I Wasn't Ready	1	1
2	2	1	Tit Punch	2	1
3	3	2	Secrets	1	1
4	4	2	Lies	2	1
5	5	3	Chapter 1: Dreams	1	1
6	6	3	Chapter 2: Memories	2	1
7	7	4	<null>	1	1
8	8	4	<null>	2	1
9	9	5	Nerd Today, Boss Tomorrow		
10	10	5	Life's Not Fair, Get Used to It		
11	11	6	Do as Planned		
12	12	6	Lethal Negligence		
13	13	7	Openings		
14	14	7	Exchanges		
15	15	9	Episode 1.01		
16	16	9	Episode 1.02		
17	17	10	<null>		
18	18	10	<null>		
19	19	12	Episode 1		
20	20	12	Episode 2		
21	21	17	The National Anthem		
22	22	17	Fifteen Million Merits		

episode

	id	show_id	director_id
1	3	1	1
2	4	2	2
3	5	3	3
4	6	4	4
5	7	4	5
6	8	4	6
7	9	5	7
8	10	5	8
9	11	6	9
10	12	7	10
11	13	9	11
12	14	10	12
13	15	12	13
14	16	17	14

show_directors

	director_id	first_name	last_name	country_name
1	1	Jenji	Kohan	USA
2	2	Baran	bo Odar	Germany
3	3	Ramón	Campos	Spain
4	4	Jonathan	Entwistle	UK
5	5	Lucy	Tcherniak	UK
6	6	Lucy	Forbes	UK
7	7	Philipp	Käbbohrer	Germany
8	8	Matthias	Murmann	Germany
9	9	Álex	Pina	Spain
10	10	Scott	Frank	USA
11	11	Lior	Raz	Israel
12	12	Konstantin	Statskiy	Russia
13	13	Joe	Penhall	USA
14	14	Charlie	Brooker	UK

director

Заполнение базы данных

	show_id	imdb_rating	rt_rating
1	1	8	90
2	2	8.8	95
3	3	7.6	<null>
4	4	8.1	
5	5	7.9	
6	6	8.3	
7	7	8.6	
8	9	8.2	
9	10	7.7	
10	12	8.6	
11	17	8.8	

statistics

	id	show_id	actor_id
1	2	1	1
2	3	1	2
3	4	2	3
4	5	2	4
5	6	3	5
6	7	4	6
7	10	5	7
8	11	6	8
9	12	7	9
10	13	9	10
11	14	10	11
12	15	12	12
13	16	17	13

show_cast

	id	show_id	company_id
1	1	3	1
2	2	1	2
3	3	2	3
4	4	4	4
5	5	4	5
6	7	5	8
7	8	6	10
	10	7	11
	11	7	12
	12	9	13
	13	10	14
	14	12	15
	15	17	16
	16	17	17

production_companies

Запросы к базе данных

- Простые запросы: <https://github.com/de1iza/netflix-db/blob/main/selects.sql>
- Более сложные запросы: https://github.com/de1iza/netflix-db/blob/main/selects_advanced.sql

-- средний рейтинг по сериалам каждой страны

```
select country_name, avg(imdb_rating)
from shows.statistics join shows.tv_show
on statistics.show_id = tv_show.show_id
group by country_name;
```

	country_name	avg
1	Israel	8.2
2	Spain	7.95
3	UK	8.45
4	Germany	8.350000000000001
5	Russia	7.7
6	USA	8.4

-- режиссеры, создавшие сериалы с рейтингом >= 8.5

```
select first_name, last_name from shows.director join
(select director_id from shows.show_directors where show_id in
(select show_id from shows.statistics where imdb_rating >= 8.5)) as best_dirs
on best_dirs.director_id = director.director_id;
```

	first_name	last_name
1	Baran	bo Odar
2	Scott	Frank
3	Joe	Penhall
4	Charlie	Brooker

-- место сериала в рейтинге по оценке imdb среди сериалов этой же страны

```
select title, row_number() over(partition by country_name order by imdb_rating desc) as rank
from shows.tv_show join shows.statistics
on tv_show.show_id = statistics.show_id;
```

	title	rank
1	Dark	1
2	How to Sell Drugs Online (Fast)	2
3	Fauda	1
4	Silver Spoon	1
5	Money Heist	1
6	Cable Girls	2
7	Black Mirror	1
8	The End of the F***ing World	2
9	The Queen's Gambit	1
10	Mindhunter	2
11	Orange Is the New Black	3

Views

<https://github.com/de1iza/netflix-db/blob/main/views.sql>

```
-- функция, скрывающая все символы, начиная с позиции start в строке data
create or replace function shows.hide_data(data text, start integer)
returns text as $$
    declare
        res text := data;
    begin
        res := concat(left(data, start), repeat('*', length(data) - start));
        return res;
    end;
$$ language plpgsql;

-- представление содержит скрытые имена и фамилии актеров
create or replace view shows.actors_info as
select shows.hide_data( data: first_name, start: 2) as first_name,
       shows.hide_data( data: last_name, start: 2) as last_name,
       country_name
from shows.actor;
```

	first_name	last_name	country_name
1	Ta****	Sc*****	USA
2	La***	Pr****	USA
3	Lo***	Ho*****	Germany
4	Ma**	Sc****	Germany
5	Bl****	Su****	Spain
6	Al**	La*****	UK
7	Ma*****	Mu***	Germany
8	Ur****	Co*****	Spain
9	An**	Ta*****	USA
10	Li**	Ra*	Israel
11	Pa***	Pr*****	Russia
12	Jo*****	Gr***	USA
13	Br*****	Ho****	USA

Views

```
-- подсчитывает количество актеров/режиссеров из этой страны
-- type == 'actor' / 'director'
create or replace function shows.count_by_country(country_name text, type text)
returns integer as $$
declare
    res integer := 0;
begin
    if type = 'actor' then
        execute format('select count(*)
                        from shows.actor
                        where country_name = '%s'', country_name)

        into res;
    else
        execute format('select count(*)
                        from shows.director
                        where country_name = '%s'', country_name)

        into res;
    end if;
    return res;
end;
$$ language plpgsql;
```

```
-- количество снятых сериалов, известных актеров и режиссеров по странам
create or replace view shows.country_stats as
select
    country_name,
    count(tv_show.show_id) as n_shows,
    shows.count_by_country(country_name, type: 'actor') as n_actors,
    shows.count_by_country(country_name, type: 'director') as n_directors
from shows.statistics join shows.tv_show
on statistics.show_id = tv_show.show_id
group by country_name;
```

	country_name	n_shows	n_actors	n_directors
1	Israel	1	1	1
2	Spain	2	2	2
3	UK	2	1	4
4	Germany	2	3	3
5	Russia	1	1	1
6	USA	3	5	3

Функции и триггеры

https://github.com/de1iza/netflix-db/blob/main/functions_triggers.sql

```
-- функция для триггера, логирует изменения, записывая изменения в отдельную таблицу
create or replace function shows.log_change()
returns trigger as $logs$
begin
    create table if not exists shows.logs(
        operation char(1) not null,
        time timestamp not null,
        userid text not null,
        title text not null,
        genres text array,
        release_year integer,
        n_seasons integer,
        n_episodes integer
    );
    if (TG_OP = 'DELETE') then
        insert into shows.logs select 'D', now(), user, old.title, old.genres, old.release_year, old.n_seasons, old.n_episodes;
        return old;
    elseif (TG_OP = 'UPDATE') then
        insert into shows.logs select 'U', now(), user, new.title, new.genres, new.release_year, new.n_seasons, new.n_episodes;
        return new;
    elseif (TG_OP = 'INSERT') then
        insert into shows.logs select 'I', now(), user, new.title, new.genres, new.release_year, new.n_seasons, new.n_episodes;
        return new;
    end if;
    return null;
end;
@$logs$ language plpgsql;
```

-- триггер, вызывается при вставке / изменении / удалении, логирует изменения

```
create trigger log_update
after insert or update or delete on shows.tv_show
for each row
execute procedure shows.log_change();
```

Логирующий триггер

Функции и триггеры

```
-- функция для триггера, изменяет формат входных данных (названия стран)
create or replace function shows.process_country()
returns trigger as $$
begin
    if new.country_name = 'United Kingdom' then
        new.country_name := 'UK';
    elseif new.country_name = 'United States of America' then
        new.country_name := 'USA';
    end if;
    return new;
end;
$$ language plpgsql;

-- триггер, вызывается перед вставкой / изменением таблицы со странами, изменяет формат названия
create trigger process_input
before insert or update on shows.country
for each row
execute procedure shows.process_country();
```

Триггер обработки названий