Netflix Database
Final Report

BADM 352 – Database Design and Management

Group 2 – SQL & Chill

Keemi Patel, Naiya Patel, Brian Deenihan, Eugene Hwang, Charlotte Shafer

Section 1: Context and Business Background Information

Netflix is a streaming service that started as an online DVD rental service in 1997. It entered into the streaming industry in 2007 and currently has over 260 million subscribers. Globally, Netflix has over 13,500 titles including a variety of movies, TV shows, and video games.

We chose Netflix for this project because we are all familiar with the customer-facing side of Netflix and were able to identify several interconnected entities. We also thought movies and TV shows would be an interesting focus. On top of that, this data can also be used to recommend new offerings on Netflix to customers given their watch history.

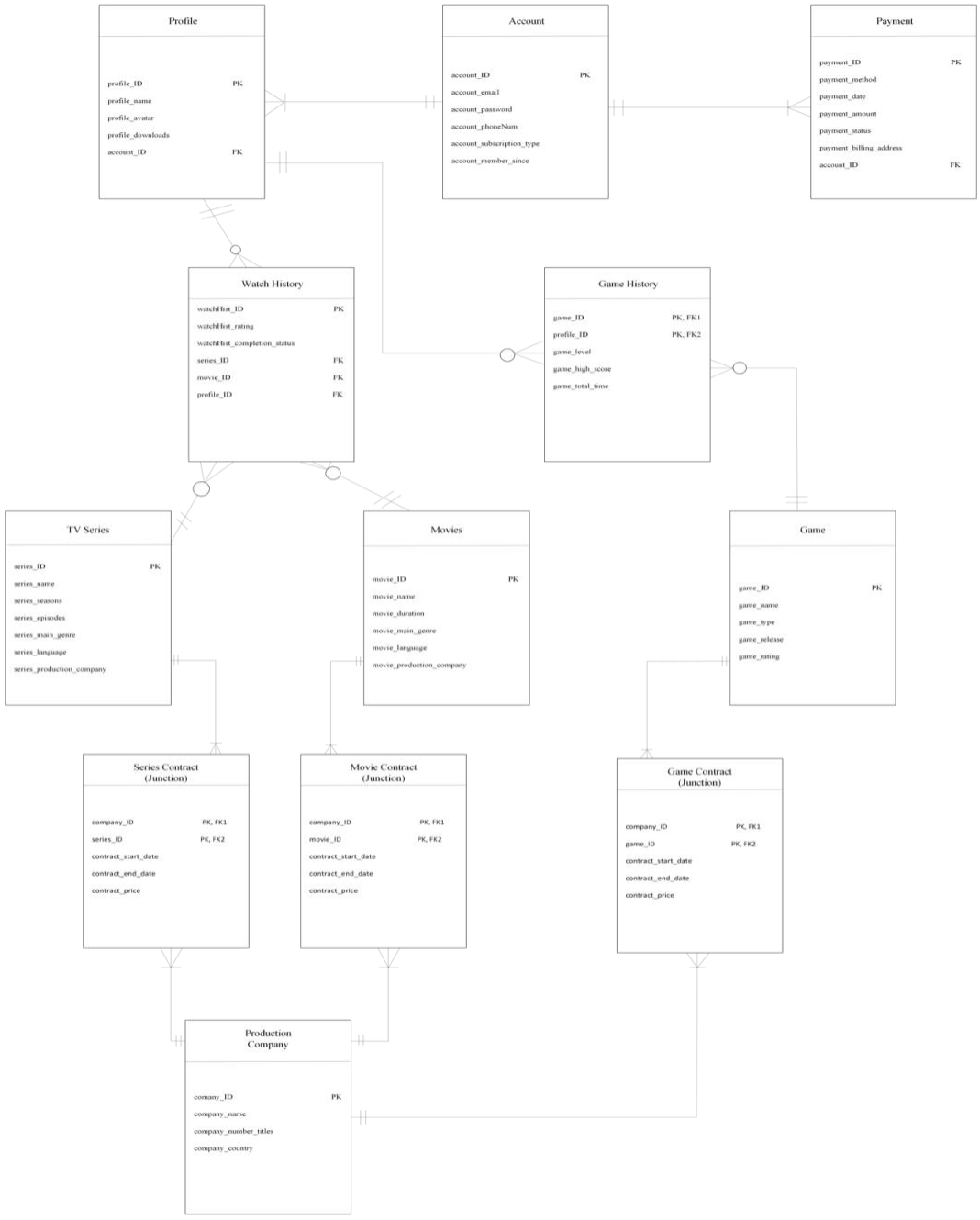
Section 2: Business Rules

Netflix (profile, account, TV shows, movies, production company, contracts, payment, games)

Each account can have multiple profiles and each profile is associated with one account. Each profile can play many games, each game can be played by multiple profiles. A production company can produce and license multiple shows, and a show can be produced by one or more production companies. A production company can also produce and license multiple movies, and a movie can be produced by one or more production companies. An account will have one subscription plan, either standard with ads, standard, or premium. Netflix also tracks and records user data, so each profile will have data such as the amount of time spent watching and whether the user marked the content as liked or disliked. An account can make multiple payments and each payment is associated with only one account.

Section 3: Relational Model and Underlying Assumptions

We had a total of 12 tables, which included any junction tables: Profile, Account, Payment, Watch History, Game History, TV Series, Movies, Game, Series Contract (junction), Movie Contract (junction), Game Contract (junction), and Production Company.



Section 4: Database Tables

Here are the screenshots of each table in the database:

Account

account_ID	account_email	account_password	account_phoneNum	account_subscription_type	account_member_since
101	keemi2@gmail.com	passPat123	1234567890	Premium	2020-06-10
102	nayia3@gmail.com	passEL456	0987654321	Basic	2021-01-20
103	brian4@gmail.com	SHaPwd789	4567891230	Premium	2022-09-05
104	charlotte5@gmail.com	BaiPass321	7890123456	Basic	2019-11-15
105	eugene6@gmail.com	egene45Pass	3216549870	Premium	2023-03-01
NULL	NULL	NULL	NULL	NULL	NULL

Profile

profile_ID	profile_name	profile_avatar	profile_downloads	account_ID
1	Naiya	avatar1	5	101
2	Keemi	avatar2	6	102
3	Brian	avatar3	3	103
4	Charlotte	avatar4	2	104
5	Eugene	avatar5	6	105
NULL	NULL	NULL	NULL	NULL

Payment

payment_ID	payment_method	payment_date	Payment_amount	payment_status	payment_billing_address	account_ID
401	Credit Card	2024-01-01	14.99	Completed	123 Main St, New York, NY 10001	101
402	PayPal	2024-01-15	7.99	Completed	456 Oak Ave, Los Angeles, CA 90001	102
403	Credit Card	2024-02-01	14.99	Pending	789 Pine Rd, Chicago, IL 60601	103
404	Debit Card	2024-03-01	7.99	Completed	321 Maple Ln, Houston, TX 77001	104
405	Credit Card	2024-04-01	14.99	Completed	654 Elm St, Miami, FL 33101	105
NULL	NULL	NULL	NULL	NULL	NULL	NULL

Production Company

company_ID	company_name	company_number_titles	company_country
51	Warner Bros	120	USA
52	Paramount	80	USA
53	National Geo	60	USA
54	Telemundo	45	Mexico
55	CBS	75	USA
56	Pixar Animation Studios	50	USA
57	BBC Studios	65	UK
58	Studio Ghibli	40	Japan
59	Sony Pictures	85	USA
60	Bollywood Films	70	India
NULL	NULL	NULL	NULL

Game

game_ID	game_name	game_type	game_release	game_rating
601	Galaxy Fighters	Action	2022-06-15	4
602	Cooking Master	Simulation	2021-08-10	5
603	Detective Hunt	Mystery	2020-11-20	4
604	Space Racer	Racing	2019-05-25	3
605	Romance Quest	Adventure	2023-01-05	2
NULL	NULL	NULL	NULL	NULL

Game Contract

company_ID	game_ID	contract_start_date	contract_end_date	contract_price
52	603	2025-03-20	2025-09-20	105000
53	601	2025-01-05	2025-12-05	115001
57	604	2025-04-15	2025-10-15	112000
58	602	2025-02-10	2026-02-10	98000.2
60	605	2025-05-30	2026-05-30	120000
NULL	NULL	NULL	NULL	NULL

Game History

game_ID	profile_ID	game_level	game_high_score	game_total_time
601	1	10	1500	20
602	2	8	1200	15
603	3	5	800	10
604	4	12	1600	25
605	5	9	1300	18
NULL	NULL	NULL	NULL	NULL

Watch History

watchHist_ID	watchHist_rating	watchHist_completion_status	series_ID	movie_ID	profile_ID
701	5	98.2	201	NULL	1
702	4	28.4	NULL	502	2
703	3	50.6	202	NULL	3
704	5	68.9	NULL	505	4
705	2	22.2	204	NULL	5
NULL	NULL	NULL	NULL	NULL	NULL

Movies

movie_ID	movie_name	movie_duration	movie_main_genre	movie_language	movie_production_company
501	Eternal Horizons	120	Sci-Fi	English	Stellar Studios
502	Crimson Dusk	105	Drama	Spanish	CineLatino Productions
503	Neon Chase	110	Action	English	Velocity Films
504	Whispers of Time	95	Romance	French	Étoile Pictures
505	Shadow Protocol	130	Thriller	German	Nordlicht Entertainment
NULL	NULL	NULL	NULL	NULL	NULL

Movie Contract

company_ID	movie_ID	contract_start_date	contract_end_date	contract_price
56	501	2025-01-10	2025-12-10	100001
57	502	2025-02-20	2026-02-20	120001
58	503	2025-03-15	2025-09-15	95000.2
59	504	2025-04-05	2025-10-05	110000
60	505	2025-05-25	2026-05-25	130000
NULL	NULL	NULL	NULL	NULL

TV Series

series_ID	series_name	series_seasons	series_episodes	series_main_genre	series_language	series_production_company
201	The Chase	3	24	Thriller	English	Warner Bros
202	Galaxy Quest	5	60	Sci-Fi	English	Paramount
203	Foodie Life	2	16	Documentary	English	National Geo
204	Love Saga	4	40	Romance	Spanish	Telemundo
205	Detective Mind	6	72	Mystery	English	CBS
NULL	NULL	NULL	NULL	NULL	NULL	NULL

Series Contract

company_ID	series_ID	contract_start_date	contract_end_date	contract_price
51	201	2025-01-01	2025-12-31	50000.9
52	202	2025-02-15	2026-02-14	75000.2
53	203	2025-03-01	2025-09-30	62000.3
54	204	2025-04-01	2025-10-01	88000.2
55	205	2025-05-10	2026-05-09	93000.9
NULL	NULL	NULL	NULL	NULL

Section 5: Code for Creation of Tables and Insertion of Data

Creation:

Database

Create database naiyasp2_NETFLIX;

Account

```
create table Account(
account_id varchar(20) not null unique,
account_email varchar(100),
account_password varchar(20),
account_phoneNum int(10),
account_subscription_type char(30),
account_member_since date,
primary key (account_ID));
```

Payment

```
create table Payment(
payment_ID varchar(20),
payment_method char(50),
payment_date date,
Payment_amount float(4),
payment_status char(20),
payment_billing_address varchar(500),
account_ID varchar(20),
primary key (payment_ID));
```

Profile

```
create table Profile(
profile_ID varchar(20) not null unique,
profile_name char(50),
profile_avatar char(20),
```

profile_downloads int(10),
primary key (profile_ID),
foreign key (account_ID));

Watch_History

create table Watch_History(
watchHist_ID varchar(20) not null unique,
watchHist_rating int,
watchHist_completion_status float(2),
series_ID varchar(20),
movie_ID varchar(20),
profile_ID varchar(20),
primary key (watchHist_ID),
foreign key (series_ID) references TV Series(series_ID),
foreign key (movie_ID) references Movies(movie_ID),
foreign key (profile_ID) references Profile (profile_ID));

Game_History

create table Game_History(
game_ID varchar(20) not null unique,
profile_ID varchar(20) not null unique,
game_level int,
game_high_score int,
game_total_time int,
primary key (game_ID, profile_ID),
foreign key (game_ID) references Game (game_ID),
foreign key (profile_ID) references Profile(profile_ID));

Game

create table Game(

game_ID varchar(20) not null unique,
game_name char(50),
game_type char(50),
game_release date,
game_rating int,
primary key (game_ID));

Movies

create table Movies(
movie_ID varchar(20) not null unique,
movie_name char(50),
movie_duration int,
movie_main_genre char(20),
movie_language char(20),
movie_production_company char(50),
primary key(movie_ID));

TV_Series

create table TV_Series(
series_ID varchar(20) not null unique,
series_name char(50),
series_seasons int,
series_episodes int,
series_main_genre char (20),
series_language char (20),
series_production_company char (50),
primary key(series_ID));

Production_Company

create table Production_Company(
company_ID varchar(20),

company_name char(100),
company_number_titles int (10),
company_country char (100),
primary key(company_ID));

Series_Contract

create table Series_Contract(
company_ID varchar(20),
series_ID varchar(20),
contract_start_date date,
contract_end_date date,
contract_price float(2),
primary key(company_ID, series_ID),
foreign key(company_ID) references Production_Company(company_ID),
foreign key(series_ID) references TV_Series(series_ID));

Movie_Contract

create table Movie_Contract(
company_ID varchar(20),
movie_ID varchar(20),
contract_start_date date,
contract_end_date date,
contract_price float(2),
primary key(company_ID, movie_ID),
foreign key(company_ID) references Production_Company(company_ID),
foreign key(movie_ID) references Movies(movie_ID));

Game Contract

create table Game_Contract(
company_ID varchar(20),
game_ID varchar(20),

contract_start_date date,
contract_end_date date,
contract_price float(2),
primary key(company_ID, game_ID),
foreign key(company_ID) references Production_Company(company_ID),
foreign key(game_ID) references Game(game_ID));

Insertion:

Account

INSERT INTO Account (account_ID, account_email, account_password, account_phoneNum, account_subscription_type, account_member_since) VALUES
('101', 'keemi2@gmail.com', 'passPat123', '1234567890', 'Premium', '2020-06-10'),
('102', 'nayia3@gmail.com', 'passEL456', '0987654321', 'Basic', '2021-01-20'),
('103', 'brian4@gmail.com', 'SHaPwd789', '4567891230', 'Premium', '2022-09-05'),
('104', 'charlotte5@gmail.com', 'BaiPass321', '7890123456', 'Basic', '2019-11-15'),
('105', 'eugene6@gmail.com', 'egene45Pass', '3216549870', 'Premium', '2023-03-01');

Payment

INSERT INTO Payment (
payment_ID, payment_method, payment_date, payment_amount, payment_status, payment_billing_address, account_ID) VALUES
('401', 'Credit Card', '2024-01-01', 14.99, 'Completed', '123 Main St, New York, NY 10001', 101),

('402', 'PayPal', '2024-01-15', 7.99, 'Completed', '456 Oak Ave, Los Angeles, CA 90001', 102),
('403', 'Credit Card', '2024-02-01', 14.99, 'Pending', '789 Pine Rd, Chicago, IL 60601', 103),
('404', 'Debit Card', '2024-03-01', 7.99, 'Completed', '321 Maple Ln, Houston, TX 77001', 104),
('405', 'Credit Card', '2024-04-01', 14.99, 'Completed', '654 Elm St, Miami, FL 33101', 105);

Profile

INSERT INTO Profile (profile_ID, profile_name, profile_avatar, profile_downloads, account_ID) VALUES
(1, 'Naiya', 'avatar1', '5', 101),
(2, 'Keemi', 'avatar2', '6', 102),
(3, 'Brian', 'avatar3', '3', 103),
(4, 'Charlotte', 'avatar4', '2', 104),
(5, 'Eugene', 'avatar5', '6', 105);

Watch_History

INSERT INTO Watch_History (watchHist_ID, watchHist_rating, watchHist_completion_status, series_ID, movie_ID, profile_ID) VALUES
('701', 5, 98.2, '201', NULL, '1'),
('702', 4, 28.4, NULL, '302', '2'),
('703', 3, 50.6, '202', NULL, '3'),
('704', 5, 68.9, NULL, '305', '4'),
('705', 2, 22.2, '204', NULL, '5');

Game_History

INSERT INTO Game_History (game_ID, profile_ID, game_level, game_high_score, game_total_time) VALUES

('601', '1', 10, 1500, 20),
('602', '2', 8, 1200, 15),
('603', '3', 5, 800, 10),
('604', '4', 12, 1600, 25),
('605', '5', 9, 1300, 18);

Game

```
INSERT INTO Game (game_ID, game_name, game_type, game_release,
game_rating) VALUES
('601', 'Galaxy Fighters', 'Action', '2022-06-15', 4),
('602', 'Cooking Master', 'Simulation', '2021-08-10', 5),
('603', 'Detective Hunt', 'Mystery', '2020-11-20', 4),
('604', 'Space Racer', 'Racing', '2019-05-25', 3),
('605', 'Romance Quest', 'Adventure', '2023-01-05', 2);
```

Movies

```
INSERT INTO Movies (
   movie_ID,
   movie_name,
   movie_duration,
   movie_main_genre,
   movie_language,
   movie_production_company
)
VALUES
   ('501', 'Eternal Horizons', 120, 'Sci-Fi', 'English', 'Stellar Studios'),
   ('502', 'Crimson Dusk', 105, 'Drama', 'Spanish', 'CineLatino
Productions'),
   ('503', 'Neon Chase', 110, 'Action', 'English', 'Velocity Films'),
   ('504', 'Whispers of Time', 95, 'Romance', 'French', 'Étoile Pictures'),
   ('505', 'Shadow Protocol', 130, 'Thriller', 'German', 'Nordlicht
Entertainment');
```

TV_Series

INSERT INTO TV_Series (series_ID, series_name, series_seasons, series_episodes, series_main_genre, series_language, series_production_company) VALUES
('201', 'The Chase', 3, 24, 'Thriller', 'English', 'Warner Bros'),
('202', 'Galaxy Quest', 5, 60, 'Sci-Fi', 'English', 'Paramount'),
('203', 'Foodie Life', 2, 16, 'Documentary', 'English', 'National Geo'),
('204', 'Love Saga', 4, 40, 'Romance', 'Spanish', 'Telemundo'),
('205', 'Detective Mind', 6, 72, 'Mystery', 'English', 'CBS');

Production_Company

INSERT INTO Production_Company (company_ID, company_name, company_number_titles, company_country) VALUES
('51', 'Warner Bros', 120, 'USA'),
('52', 'Paramount', 80, 'USA'),
('53', 'National Geo', 60, 'USA'),
('54', 'Telemundo', 45, 'Mexico'),
('55', 'CBS', 75, 'USA'),
('56', 'Pixar Animation Studios', 50, 'USA'),
('57', 'BBC Studios', 65, 'UK'),
('58', 'Studio Ghibli', 40, 'Japan'),
('59', 'Sony Pictures', 85, 'USA'),
('60', 'Bollywood Films', 70, 'India');

Series_Contract

INSERT INTO series_contract (
company_ID,
series_ID,
contract_start_date,

```
    contract_end_date,
    contract_price
)
VALUES
    ('51', '201', '2025-01-01', '2025-12-31', 50000.88),
    ('52', '202', '2025-02-15', '2026-02-14', 75000.22),
    ('53', '203', '2025-03-01', '2025-09-30', 62000.34),
    ('54', '204', '2025-04-01', '2025-10-01', 88000.21),
    ('55', '205', '2025-05-10', '2026-05-09', 93000.87);
```

Movie_Contract

```
INSERT INTO movie_contract (
    company_ID,
    movie_ID,
    contract_start_date,
    contract_end_date,
    contract_price
)
VALUES
    ('56', '501', '2025-01-10', '2025-12-10', 100000.99),
    ('57', '502', '2025-02-20', '2026-02-20', 120000.88),
    ('58', '503', '2025-03-15', '2025-09-15', 95000.22),
    ('59', '504', '2025-04-05', '2025-10-05', 110000.11),
    ('60', '505', '2025-05-25', '2026-05-25', 130000.21);
```

Game_contract

```
INSERT INTO game_contract (
    company_ID,
    game_ID,
    contract_start_date,
    contract_end_date,
    contract_price
```

)
VALUES
('53', '601', '2025-01-05', '2025-12-05', 115000.99),
('58', '602', '2025-02-10', '2026-02-10', 98000.21),
('52', '603', '2025-03-20', '2025-09-20', 105000.43),
('57', '604', '2025-04-15', '2025-10-15', 112000.23),
('60', '605', '2025-05-30', '2026-05-30', 120000.22);

Section 6: Queries and Results

1. **Determine the most popular subscription plan**

```
9  SELECT account_subscription_type, COUNT(*) FROM Account
10 GROUP BY account_subscription_type
11 ORDER BY count(*) DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

account_subscription_type	COUNT(*)
Premium	3
Basic	2

2. **List the tv series that accounts with a premium subscription have watched**

```
14 SELECT DISTINCT TS.series_name
15 FROM TV_Series TS
16 INNER JOIN Watch_History WH ON TS.series_ID = WH.series_ID
17 INNER JOIN Profile P ON WH.profile_ID = P.profile_ID
18 INNER JOIN Account A ON P.account_ID = A.account_ID
19 WHERE A.account_subscription_type = 'Premium';
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

series_name
The Chase
Galaxy Quest
Love Saga

3. **List the company name(s) and movie name(s) of the movie(s) with the longest contract**

```
11 •  Select pc.company_name, m.movie_name
12    FROM Production_Company as pc INNER JOIN Movie_Contract as mc INNER JOIN Movies as m
13    ON pc.company_ID = mc.company_ID AND mc.movie_ID = m.movie_ID
14    WHERE DATEDIFF(mc.contract_end_date, mc.contract_start_date) =
15    (select max(datediff(contract_end_date, contract_start_date)) from Movie_Contract);
16
```

100% 62:13

Result Grid | Filter Rows: Search | Export:

company_name	movie_name
BBC Studios	Crimson Dusk
Bollywood Films	Shadow Protocol

4. **List the profile name, game name, and total game time of the entry in game history that has the longest total game time**

```
21 •  SELECT P.profile_name, G.game_name, GH.game_total_time
22    FROM Game_History GH
23    INNER JOIN Profile P ON GH.profile_ID = P.profile_ID
24    INNER JOIN Game G ON GH.game_ID = G.game_ID
25    WHERE GH.game_total_time = (
26        SELECT MAX(game_total_time) FROM Game_History
27    );
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

profile_name	game_name	game_total_time
Charlotte	Space Racer	25

5. **Fine the average game rating across all games**

```
185 •  select avg(game_rating) from Game;
186
```

100% 27:175

Result Grid | Filter Rows: Search | Expo

avg(game_ratin...
3.6000

6. **List the account information for accounts that include “pass” or “Pass” in their passwords**

```
187 •  select * from Account
188    where account_password like '%pass%' or account_password like '%Pass%';
189
```

100% 1:184

Result Grid | Filter Rows: Search | Edit: | Export/Import:

account_ID	account_email	account_password	account_phoneN...	account_subscription_t...	account_member_sin...
101	keemi2@gmail.com	passPat123	1234567890	Premium	2020-06-10
102	nayia3@gmail.com	passEL456	0987654321	Basic	2021-01-20
104	charlotte5@gmail.com	BaiPass321	7890123456	Basic	2019-11-15
105	eugene6@gmail.com	egene45Pass	3216549870	Premium	2023-03-01
NULL	NULL	NULL	NULL	NULL	NULL

7. **List the production company information from the production company with the least titles and the production company with the most titles**

```
44 •  SELECT * FROM Production_Company
45    WHERE company_number_titles in (SELECT MAX(company_number_titles) FROM Production_Company )
46    OR  company_number_titles IN (SELECT MIN(company_number_titles) FROM Production_Company );
47
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

company_ID	company_name	company_number_titles	company_country
51	Warner Bros	120	USA
58	Studio Ghibli	40	Japan
NULL	NULL	NULL	NULL

8. List the number of movies grouped by language

```
48 •  Select movie_language, count(*)
49    From Movies
50    Group by movie_language;
51
```

Result Grid | Filter Rows:

movie_language	count(*)
English	2
Spanish	1
French	1
German	1

9. List the account t ID for accounts with billing addresses in Illinois (written in billing address as 'IL')

```
10 •  select account_ID from Payment
11    where payment_billing_address like "% IL %";
12
```

100% | 45:11

Result Grid | Filter Rows: Search | Export:

account_ID
103

10. List the tv series information and tv series contract price for the series with the highest contract price

```
13 •  select S.*, C.contract_price from TV_Series as S inner join Series_Contract as C
14    on S.series_ID = C.series_ID
15    where C.contract_price in (select max(contract_price) from Series_Contract);
```

100% | 32:10

Result Grid | Filter Rows: Search | Export:

series_ID	series_name	series_seasons	series_episodes	series_main_genre	series_language	series_production_comp...	contract_pri...
205	Detective Mind	6	72	Mystery	English	CBS	93000.9

Section 7: Limitations and Scope of Extension

Given more time and resources, we would like to extend the scope of our database in three main ways: break the Watch History table into two tables — Movie Watch History and Series Watch History, add tables for internal aspects of Netflix, and add a table for projects that are in production. The Watch History table currently contains information of both TV series and movies that have been watched by profiles. This leads to many null values, since each row will contain a null value under the series_ID or movie_ID column, depending on which form of content the row relates to. Breaking up the Watch History table into Movie Watch History and Series Watch History would not only reduces the number of null values stored in the database, but also simplify queries related to attributes that may not be comparable between mediums. For example, the average completion status of the Watch History table would not carry much meaning given that movies and TV series typically have dramatically different lengths. Creating two tables would solve these problems. One limitation of our current model is that it only focuses on customer-facing aspects of Netflix. To expand the scope, we would want to add tables that represent internal divisions of Netflix, such as geographical divisions, departments, employees, and infrastructure providers. This would make our model more complex but would also provide additional data that could be useful for Netflix. One way to incorporate internal entities into the current model is adding relationships between geographical divisions and the TV Series, Movies, and Accounts tables. Depending on where an account is located, different movies and TV series will be available. Keeping track of the location of their users and in which region certain shows or movies are available would provide Netflix with more tailored data that could be used for recommendations or other analysis. Furthermore, we would like to add tables for projects that are currently in production, as this would help Netflix keep track of the status of their projects. Because finished and unfinished projects have varying attributes and relationships, creating entirely new tables to represent projects in production makes the most sense.

Section 8: AI Assistance

Generate sample data:

- Create user profiles such as names, avatars, and email addresses
- Generate movie titles, production company names, and games
- Provided us with varied durations and interactions for watch history and game history

Designing Scenarios:

- Created examples of information we can derive from the database such as finding “longest game time” or “top-watched movies”
- Helped us get started on creating our own scenarios for querying