

Analyze Twitch Gaming Data

Twitch is the world's leading live streaming platform for gamers, with 15 million daily active users. Using data to understand its users and products is one of the main responsibilities of the Twitch Data Science Team.

In this project, you will be working with two tables that contain Twitch users' stream viewing data and chat room usage data.

Stream viewing data: stream table

Chat usage data: chat table

The Twitch Science Team provided this practice dataset.

Start by getting a feel for the `stream` table and the `chat` table.

Select the first 20 rows from each of the two tables.

What are the column names?

```
SELECT *  
FROM stream  
LIMIT 20;
```

```
SELECT *  
FROM chat  
LIMIT 20;
```

Database Schema	
chat	
name	type

time	DATETIME
device_id	TEXT
login	TEXT
channel	TEXT
country	TEXT
player	TEXT
game	TEXT

Rows: 3001
stream

name	type
time	DATETIME
device_id	TEXT
login	TEXT
channel	TEXT
country	TEXT
player	TEXT
game	TEXT
stream_format	TEXT
subscriber	TEXT

What are the unique games in the stream table?

```
SELECT DISTINCT game  
FROM stream;
```

What are the unique channels in the stream table?

```
SELECT DISTINCT channel  
FROM stream;
```

Aggregate Functions:

What are the most popular games in the `stream` table?

Create a list of games and their number of viewers using `GROUP BY`.

```
SELECT game, COUNT(*)  
FROM stream  
GROUP BY 1  
ORDER BY 2 DESC;
```

These are some big numbers from the game League of Legends (also known as LoL).

Where are these LoL stream viewers located?

Create a list of countries and their number of LoL viewers using `WHERE` and `GROUP BY`.

```
SELECT country, COUNT(*)  
FROM stream  
WHERE game = 'League of Legends'
```

GROUP BY country
ORDER BY COUNT(*) DESC;

The `player` column contains the source the user is using to view the stream (`site`, `iphone`, `android`, etc). Create a list of `players` and their number of streamers.

```
SELECT player, COUNT(*)  
FROM stream  
GROUP BY 1  
ORDER BY 2 DESC;
```

Create a new column named `genre` for each of the games.

Group the games into their genres: Multiplayer Online Battle Arena (MOBA), First Person Shooter (FPS), Survival, and Other.

Use `GROUP BY` and `ORDER BY` to showcase only the unique game titles.

```
SELECT game,  
CASE  
    WHEN game = 'League of Legends' THEN 'MOBA'  
    WHEN game = 'Dota 2' THEN 'MOBA'  
    WHEN game = 'Heroes of the Storm' THEN 'MOBA'  
    WHEN game = 'Counter-Strike: Global Offensive' THEN 'FPS'  
    WHEN game = 'DayZ' THEN 'Survival'  
    WHEN game = 'ARK: Survival Evolved' THEN 'Survival'  
    ELSE 'Other'  
END AS 'genre',
```

COUNT(*)

FROM stream

GROUP BY game

ORDER BY 3 DESC;

Let's write a query that returns two columns:

The hours of the time column

The view count for each hour

Lastly, filter the result with only the users in your country using a WHERE clause.

SELECT strftime('%H', time),

COUNT(*)

FROM stream

WHERE country = 'US'

GROUP BY 1;

The `stream` table and the `chat` table share a column: `device_id`.

Let's join the two tables on that column.

```
SELECT *  
FROM stream  
JOIN chat  
ON stream.device_id = chat.device_id;
```

CODE

```
SELECT *  
FROM stream  
LIMIT 20;
```

```
SELECT *  
FROM chat  
LIMIT 20;
```

```
--Unique games in stream table  
SELECT DISTINCT game  
FROM stream;
```

```
--Unique games in stream table  
SELECT DISTINCT channel  
FROM stream;
```

```
--the most popular games in the stream table  
SELECT game, COUNT(*)  
FROM stream  
GROUP BY 1  
ORDER BY 2 DESC;
```

```
--Where are these LoL stream viewers located?  
SELECT country, COUNT(*)
```

```
FROM stream
WHERE game = 'League of Legends'
GROUP BY country
ORDER BY COUNT(*) DESC;
```

```
--list of players and their number of streamers
SELECT player, COUNT(*)
FROM stream
GROUP BY 1
ORDER BY 2 DESC;
```

```
--Task #7 Group games by genre
```

```
SELECT game,
CASE
  WHEN game = 'League of Legends'
  THEN 'MOBA'
  WHEN game = 'Dota 2'
  THEN 'MOBA'
  WHEN game = 'Heroes of the Storm'
  THEN 'MOBA'
  WHEN game = 'Counter-Strike: Global Offensive'
  THEN 'FPS'
  WHEN game = 'DayZ'
  THEN 'Survival'
  WHEN game = 'ARK: Survival Evolved'
  THEN 'Survival'
  ELSE 'Other'
END AS 'genre',
COUNT(*)
FROM stream
GROUP BY game
ORDER BY 3 DESC;
```

```
--Return the secs of timestamp column
```

```
SELECT time,  
strftime('%S', time)  
FROM stream  
GROUP BY 1  
LIMIT 20;
```

--Task #10

```
SELECT strftime('%H', time),  
COUNT(*)  
FROM stream  
WHERE country = 'US'  
GROUP BY 1;
```

--Join stream table and chat table

```
SELECT *  
FROM stream  
JOIN chat  
ON stream.device_id = chat.device_id;
```