

# Twitch Data Analysis using Neo4j and Python

## Neo4j Overview

Neo4j is a popular and powerful graph database management system designed for efficient storage, retrieval, and management of graph-structured data. Unlike traditional relational databases, which use tables to organize data, Neo4j utilizes a graph-based model to represent and handle data. In a Neo4j graph, data is organized as nodes, relationships, and properties. Neo4j employs Cypher queries, a specialized query language, to interact with and perform operations on the graph data.

## Dataset Description

The dataset used in this project revolves around Twitch, a popular live streaming platform. The data encompasses details about users, streams, games, languages, and teams associated with Twitch. The structure is graph-based, where nodes represent various entities, and edges depict the relationships or associations between these entities. Key relationships in the dataset include:

CHATTER: Represents chatter related to a stream.

HAS\_LANGUAGE: Denotes the language used.

HAS\_TEAM: Indicates a team affiliation.

MODERATOR: Represents moderators of a stream.

PLAYS: Denotes that a user plays a game.

VIP: Denotes VIP status of a user.

## Queries Explanation

Query 1: Retrieve All Users

- This query retrieves the names of all users from the Neo4j database.

#### Query 2: Retrieve Users with Highest Followers

- This query retrieves the names and follower counts of users with the highest number of followers.

#### Query 3: Retrieve All Languages

- This query retrieves the names of all languages used by streamers.

#### Query 4: Recommended Games by Streamer

- This query provides game recommendations for a specific streamer based on their game preferences.

#### Query 5: Popular Languages on Twitch

- This query identifies the popular languages used by streamers on Twitch.

#### Query 6: Get Moderators of a Streamer

- This query retrieves the moderators of a specific streamer.

#### Query 7: Retrieve Team Information

- This query retrieves team names, members, and games played for the teams with the highest total followers.

#### Query 8: Get Streams for a Game

- This query retrieves the names of streams associated with a specific game.