

# Cypher Workshop: Analyze a Financial Network

This is a workshop on using Cypher in Kùzu to analyze a financial network. The aim of this exercise is to demonstrate that in certain scenarios, a graph is the best abstraction to answer questions about the data (as opposed to relational tables). We will show some Cypher queries (in Kùzu) and SQL queries (in DuckDB) that answer questions about the data to demonstrate this.

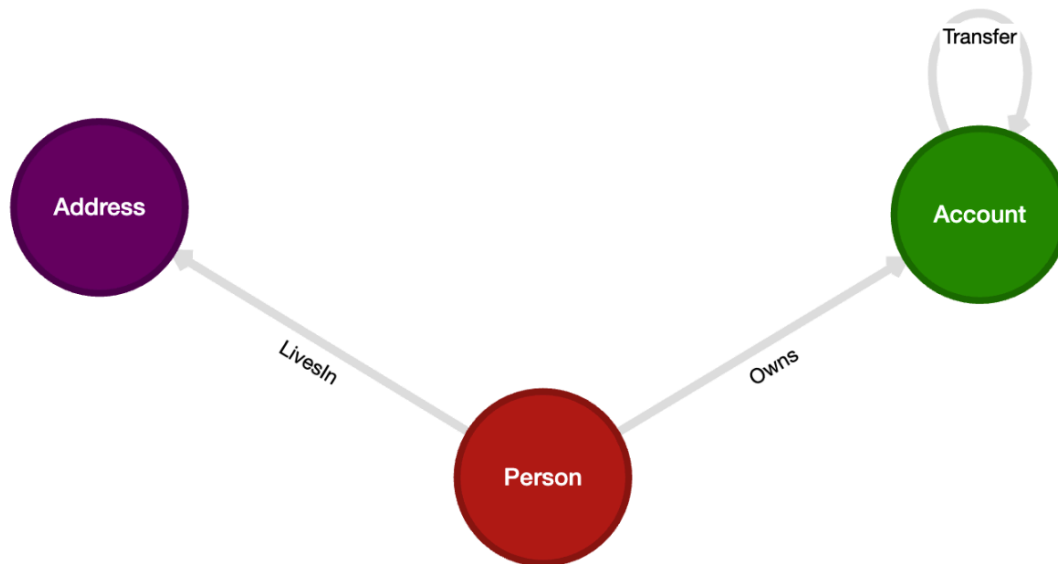
## Problem statement

Imagine you are an investigator at an organization that tracks financial crimes. Two email addresses have been flagged by law enforcement agencies, and both are now under suspicion for their potential involvement in fraudulent activities.

Your task is to analyze some data that consists of money transfers between individuals to assist in the ongoing investigation. The findings from this analysis will be crucial for an upcoming court case. To achieve this, you will delve into the dataset to uncover hidden patterns, connections, and transactions that could provide evidence of their illicit activities.

## Dataset

A dataset of persons, accounts, addresses and transfers between accounts is used in this demo. It can be represented pictorially as follows:

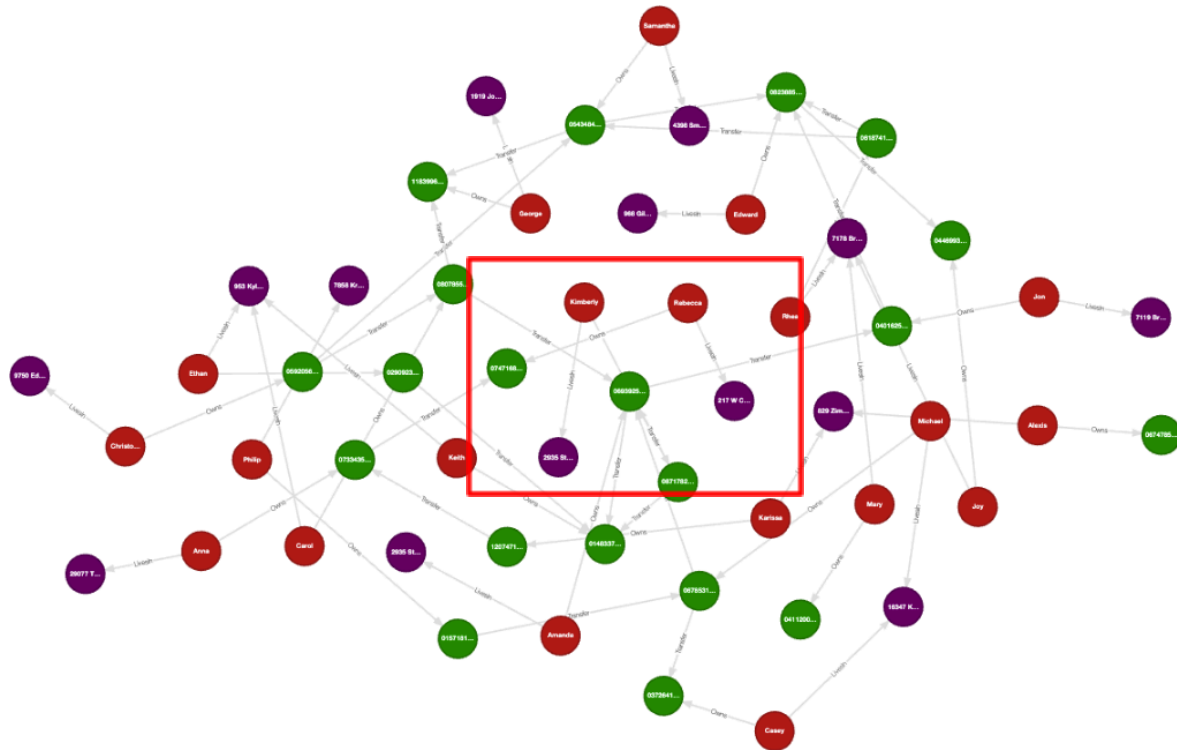


- 21 nodes of type **Person**
- 21 nodes of type **Account** (each person has exactly one account)
- 15 nodes of type **Address**
- 42 relationships of type **Transfer**, where the transfers are in one direction only (for reasons of simplicity)

# Cypher Workshop: Analyze a Financial Network

## Graph visualization

The entire graph as such has interesting structures and is small enough to visualize all at once in Kùzu explorer.



The portion of the graph in the middle shows that one account has more incoming transactions than the others. However, that knowledge alone is not sufficient to find evidence of suspicious activity.

## Data modelling

The graph schema shown above is based on the following input files.

- person.csv
- account.csv
- transfer.csv

From an RDBMS perspective, we just have three tables. The person table has an address column and the account table is connected to the person table via the person ID. The transfer table contains the source and destination account IDs, and the amount transferred.

From the three tables we begin with, we are able to separate out the required columns for our node and relationship tables in Kùzu, to give us the following six tables for our graph analysis:

- Node table: **Person**
- Node table: **Address**

# Cypher Workshop: Analyze a Financial Network

- Node table: **Account**
- Relationship table: **Owns** (between Person and Account)
- Relationship table: **LivesIn** (between Person and Address)
- Relationship table: **Transfer** (between Account and Account)

## DDL

The DDL commands for DuckDB and Kùzu are provided along with the data. You can run them via the CLI shell editor to populate the required tables in either database.

## Queries

Your goal is to write the appropriate Cypher and (where possible), SQL queries to answer the below questions.

#	Query
Q1	Find all possible <b>direct</b> transfers between two accounts owned by persons with emails <a href="mailto:georodaw366@hotmail.com">georodaw366@hotmail.com</a> and <a href="mailto:ezimmerman@yahoo.com">ezimmerman@yahoo.com</a>
Q2	Find <b>all possible</b> connections of type <b>Transfer</b> , including indirect ones up to length $k = 5$ , between the accounts owned by <a href="mailto:georodaw366@hotmail.com">georodaw366@hotmail.com</a> and <a href="mailto:ezimmerman@yahoo.com">ezimmerman@yahoo.com</a> . You can try $k > 5$ to also see how the number of paths increases rapidly.
Q3	Find the <b>shortest</b> connections of type <b>Transfer</b> between the accounts owned by <a href="mailto:georodaw366@hotmail.com">georodaw366@hotmail.com</a> and <a href="mailto:ezimmerman@yahoo.com">ezimmerman@yahoo.com</a>
Q4	Find all connections <b>of any type</b> between the persons <a href="mailto:georodaw366@hotmail.com">georodaw366@hotmail.com</a> and <a href="mailto:ezimmerman@yahoo.com">ezimmerman@yahoo.com</a> . We are searching for any possible shortest paths between these two persons, not just the transfers between accounts owned by them.
Q5	Find 3 persons who have all transferred money to each other (in at least one direction).
Q6	Find an important account that has the most incoming transactions