**Assignment 1**

Szabados Dávid

***Fitness of the input dataset to the purpose***

The selected dataset is an artificial data from a Czech bank. As the purpose is to have a dataset, in which simple or even complex work can be done, a bank dataset should fulfill this.

Link to the selected dataset:

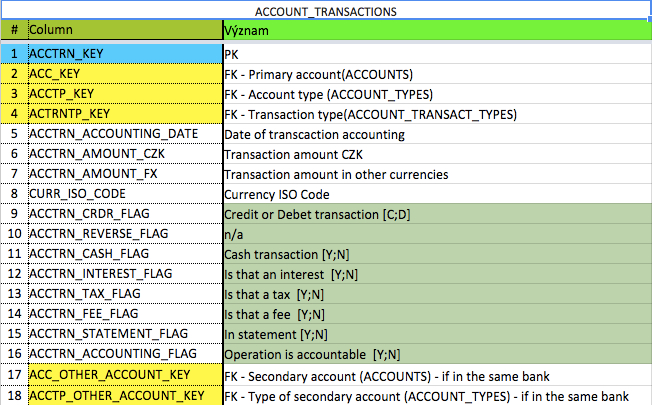
<https://relational.fit.cvut.cz/dataset/CS>

***Complexity of the input data set***

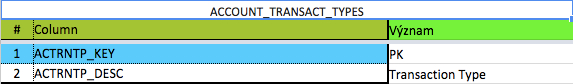
The data set uses 7 tables to represent the complexity of a bank’s information about accounts, transaction etc. and the relationships between them.

The following data tables are available for this assignment:

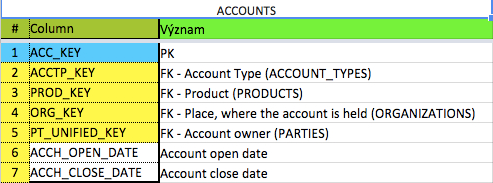
*Account Transactions:*



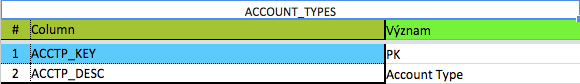
*Account Transaction Types:*



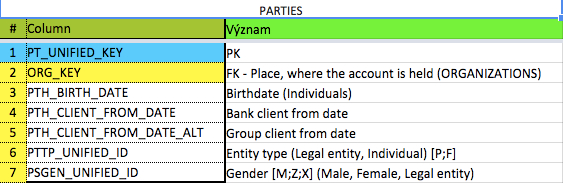
*Accounts:*



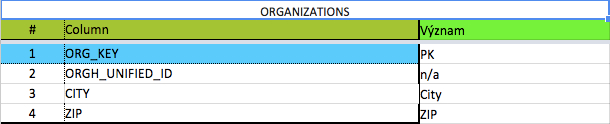
*Account Types:*



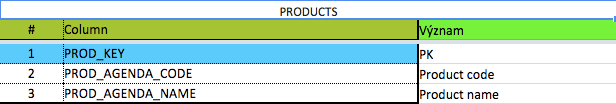
*Parties:*



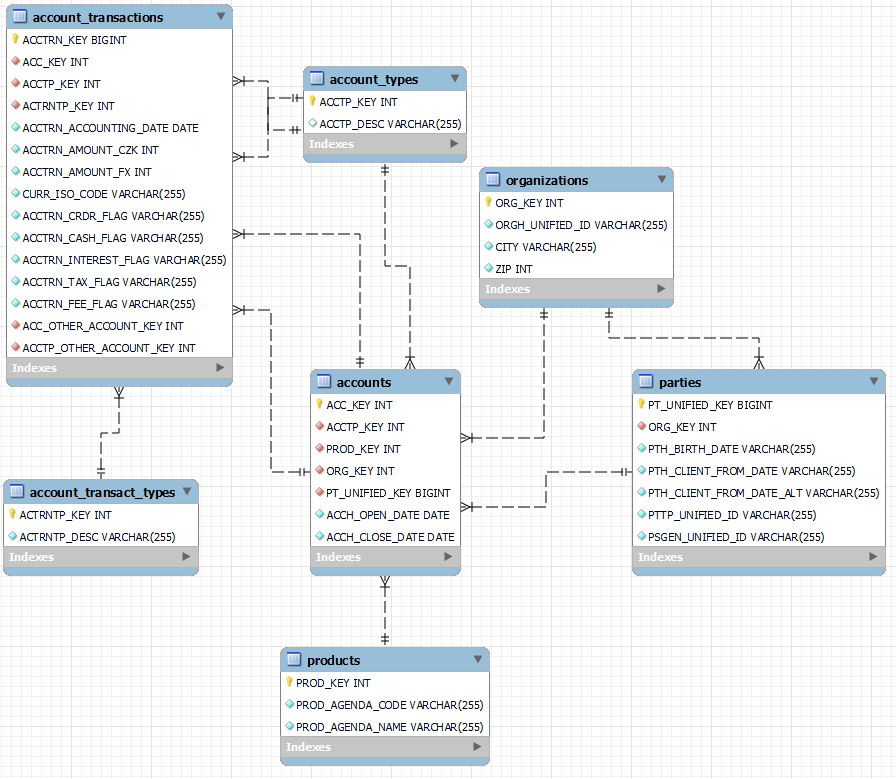
*Organizations:*



*Products:*



*Entity-Relationships Diagram:*



***Execution of the operational data layer***

The tables have been made with the columns exactly alike as it could be seen on the previous pictures.

The loading of the data has been made by the

LOAD DATA LOCAL INFILE “PATH/FILE\_NAME.csv”

command, after setting local\_infile to True, at the beginning of the script.

***Analytics plan***

During the assignment, I will use most of the knowledge we learnt during class. Mainly I want to focus on making programs and asking questions about the Accounts Transactions and the Accounts table, with the help of the remaining ones.

Some example questions:

In which city most of the accounts are being held?

Which accounts have been closed already?

***Execution of the analytical data layer***

Views:

Cash\_Transactions:

Shows every transaction, which is a cash transaction.

Closed\_Accounts:

Shows every account, that has been closed (aka close date is not 3000-01-01).

Legal\_Entity\_AccOP\_Date:

Shows the opening date of those accounts, where the owner is a legal entity.

Most\_Used\_City:

Shows that City, where most of the accounts are being held.

Procedures:

Transaction\_Amount:

Returns those transactions, where the amount is at least as much, as the given number. (Because these are transactions, the + and – signs are only the direction of the transaction, thus we speak in absolutes.)

Create\_Accounts\_INFO:

Creates a new table, which shows information about live the accounts, like the holder’s ID, Account type, location etc.

Event:

Create\_Live\_Accounts:

Calls the “Create\_Accounts\_INFO” Procedures every 1 minute. In other words, it updates the “ACCOUNTS\_INFO” Table.

***ETL***