COP 5725: Database Management Systems

Project Deliverable 3

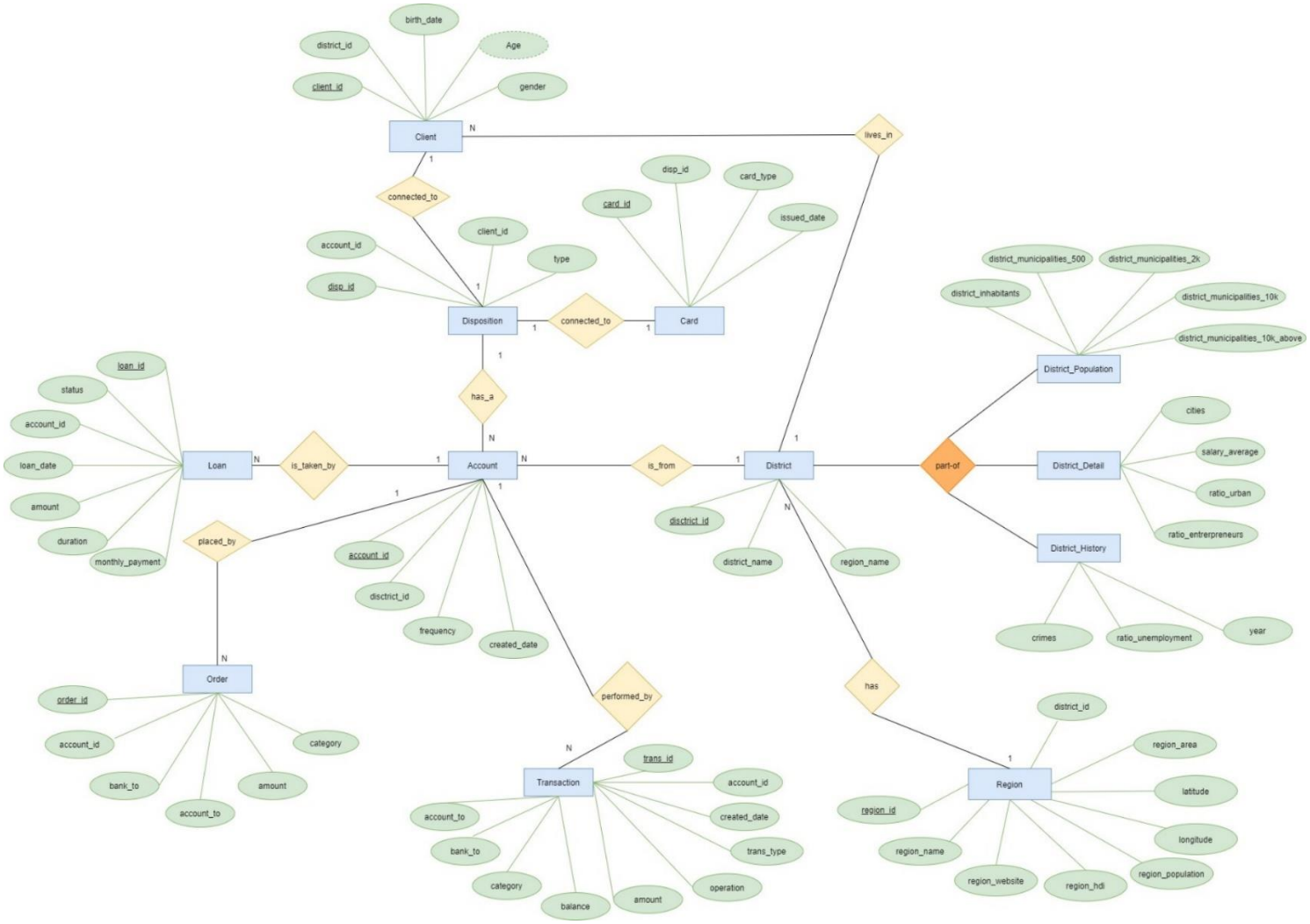
Group 25

Shangde Gao - gao.shangde@ufl.edu
Srija Gurijala - srijagurijala@ufl.edu
Dimitrios Melissourgos - dmelissourgos@ufl.edu
Andrei Sura - asura@ufl.edu
Mukul Yadav - mchand.yadav@ufl.edu

Czech Bank Financial Data Analysis and Demographics

Contents

ER Diagram



Note: There are no changes in the ER diagram except from the addition of the table “region”, which is connected to the district table by district_id with a 1-to-n relationship.

Relation Schemas

Region

region(region_id:number, district_id:number, region_name:varchar, region_area:float, region_population:number, region_hdi:float, region_website:varchar, latitude:number, longitude:number)

Account

account(account_id:number, district_id:number, frequency:char, created_date:date)

Client

client(client_id:number, district_id:number, birth_date:date, gender:char)

Disposition

disposition(disp_id:number, client_id:number, account_id:number, disp_type:char)

Card

card(card_id:number, disp_id:number, card_type:char, issued_date:date)

District

district(district_id:number, region_name:varchar, district_name:varchar)

District_Detail

district_detail(district_id:number, salary_average:float, ratio_urban:float, ratio_entrerpreneurs:float)

District_Population

district_population(district_id:number, inhabitants:number, municipalities_500:number, municipalities_2k:number, municipalities_10k:number, municipalities_10k_above:number)

District_History

district_history(district_id:number, crimes:number, ratio_unemployment:number, year:number)

Loan

loan(loan_id:number, account_id:number, loan_date:date, duration:number, payments:decimal, status:char)

Order

account_order(order_id:number, account_id:number, bank_to:char, account_to:number, amount:decimal, category:char)

Transaction

transaction(trans_id:number, account_id:number, created_date:date, trans_type:char, operation:char, amount:decimal, balance:decimal, category:char, bank_to:char, account_to:number)

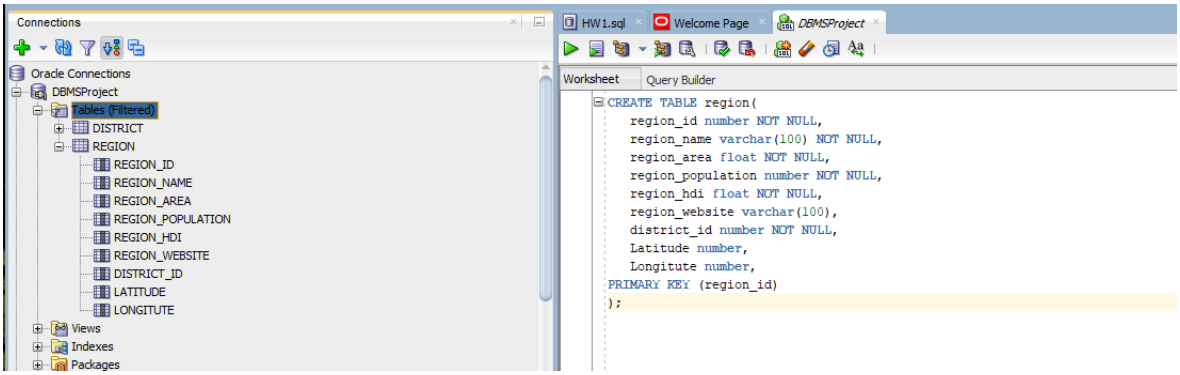
SQL Table Schemas

We have included a screenshot for every command and table created. The command can be seen in the main window, while the table created by the command, can be seen at the left panel.

Region

```
CREATE TABLE region(
  region_id number NOT NULL,
  region_name varchar(100) NOT NULL,
  region_area float NOT NULL,
  region_population number NOT NULL,
  region_hdi float NOT NULL,
  region_website varchar(100),
  district_id number NOT NULL,
  Latitude number,
  Longitute number,
PRIMARY KEY (region_id)
);
```

Note: We created this table by cross referencing data from other sources (Wikipedia).



Account

CREATE TABLE account (

account_id number NOT NULL,

district_id number NOT NULL,

frequency char(2) **NOT NULL**,

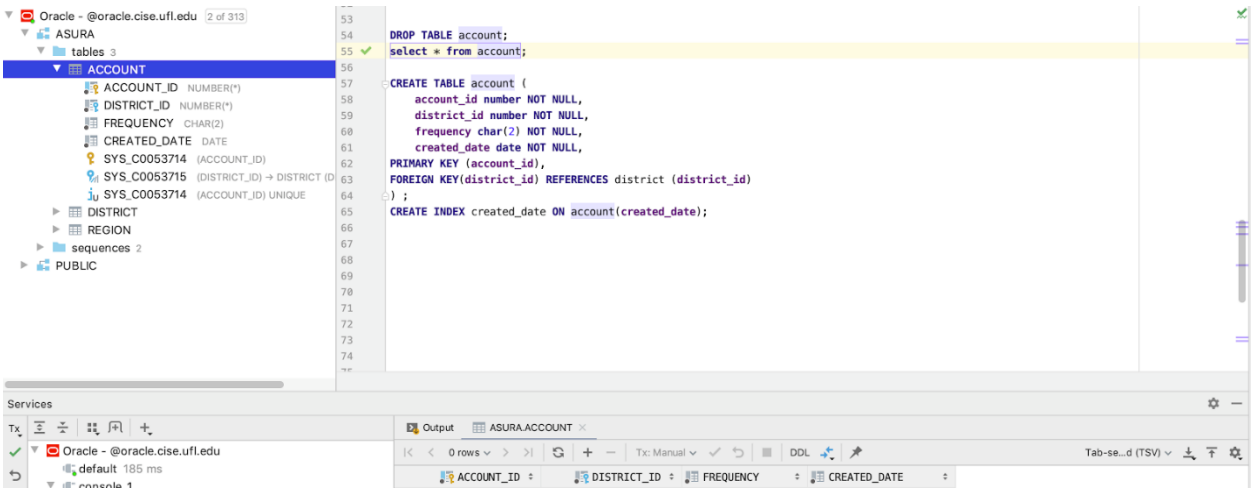
created_date date NOT NULL,

PRIMARY KEY (**account_id**),

FOREIGN KEY(**district_id**) **REFERENCES** district (**district_id**)

) ;

CREATE INDEX created_date **ON account**(**created_date**);



Client

CREATE TABLE client (

client_id number NOT NULL,

district_id number NOT NULL,

birth_date date NOT NULL,

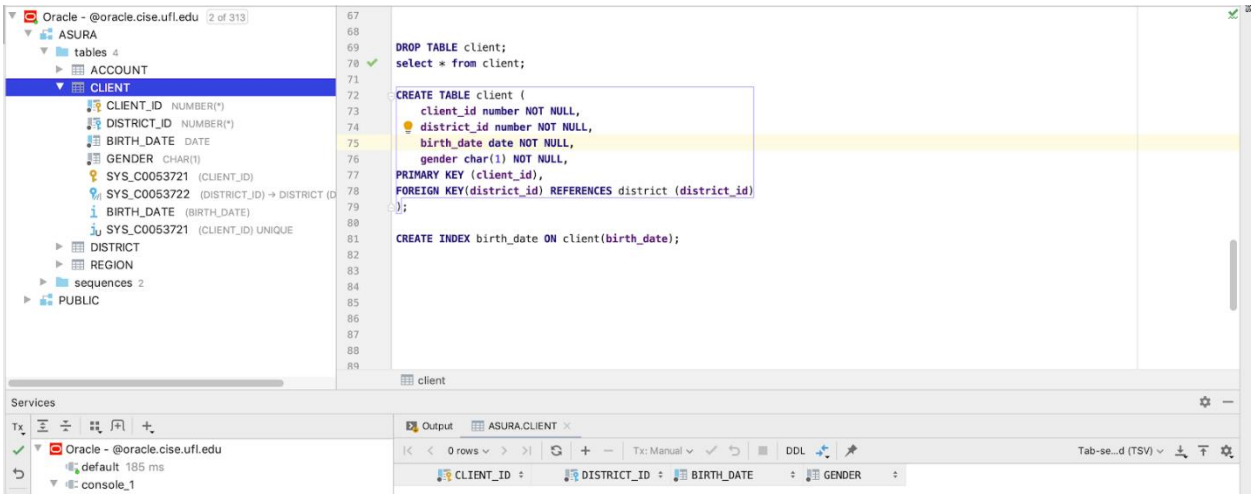
gender char(1) **NOT NULL**,

PRIMARY KEY (**client_id**),

FOREIGN KEY(**district_id**) **REFERENCES** district (**district_id**)

);

CREATE INDEX birth_date **ON client**(**birth_date**);



Disposition

CREATE TABLE disposition (

disp_id number NOT NULL,

client_id number NOT NULL,

account_id number NOT NULL,

disp_type char(1) **NOT NULL**,

PRIMARY KEY (**disp_id**),

FOREIGN KEY(**client_id**) **REFERENCES client** (**client_id**),

FOREIGN KEY(**account_id**) **REFERENCES account** (**account_id**)

);

CREATE INDEX disp_type **ON** disposition(**disp_type**);

Card

CREATE TABLE card (

card_id number NOT NULL,

disp_id number NOT NULL,

card_type char(2) **NOT NULL**,

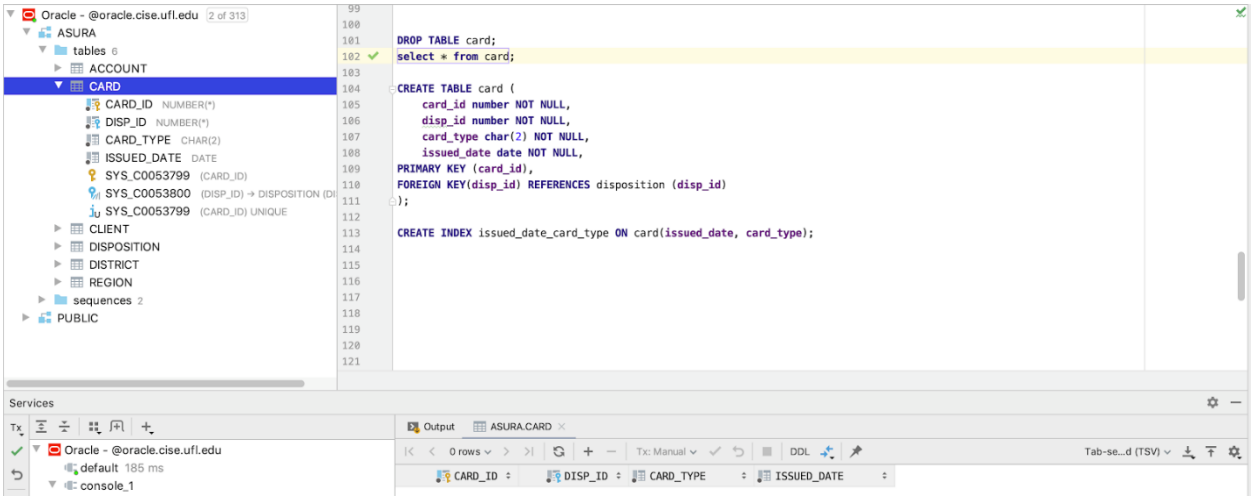
issued_date date NOT NULL,

PRIMARY KEY (**card_id**),

FOREIGN KEY(**disp_id**) **REFERENCES** disposition (**disp_id**)

);

CREATE INDEX issued_date_card_type **ON** card(**issued_date**, **card_type**);



District

CREATE TABLE district(

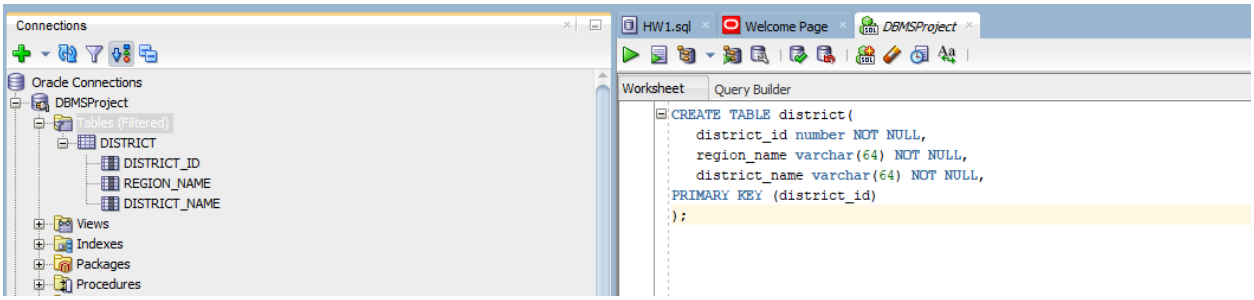
district_id number NOT NULL,

region_name varchar(64) **NOT NULL**,

district_name varchar(64) **NOT NULL**,

PRIMARY KEY (**district_id**)

);



District_Detail

CREATE TABLE district_detail(

district_id number NOT NULL,

salary_average float NOT NULL,

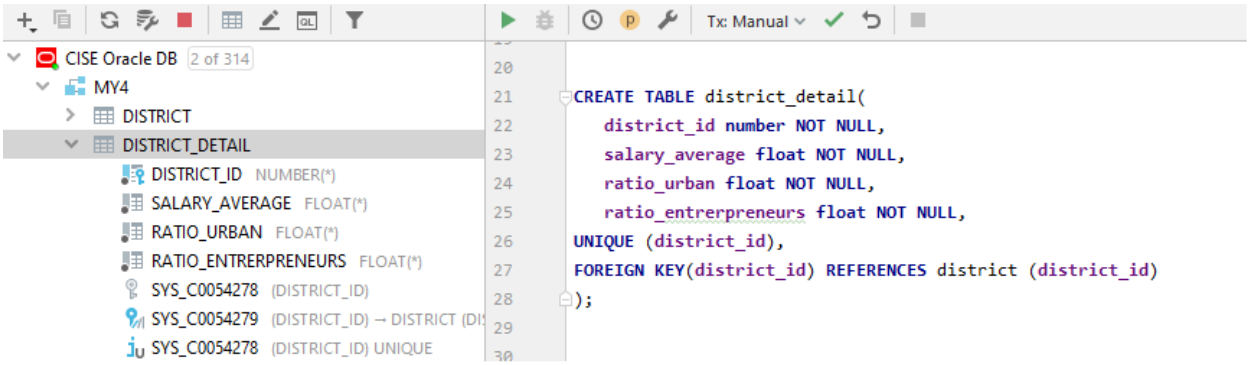
ratio_urban float NOT NULL,

ratio_entrerpreneurs float NOT NULL,

UNIQUE (**district_id**),

FOREIGN KEY(**district_id**) **REFERENCES** district (**district_id**)

);



District_Population

CREATE TABLE district_population(

district_id number NOT NULL,

inhabitants number NOT NULL,

municipalities_500 number NOT NULL,

municipalities_2k number NOT NULL,

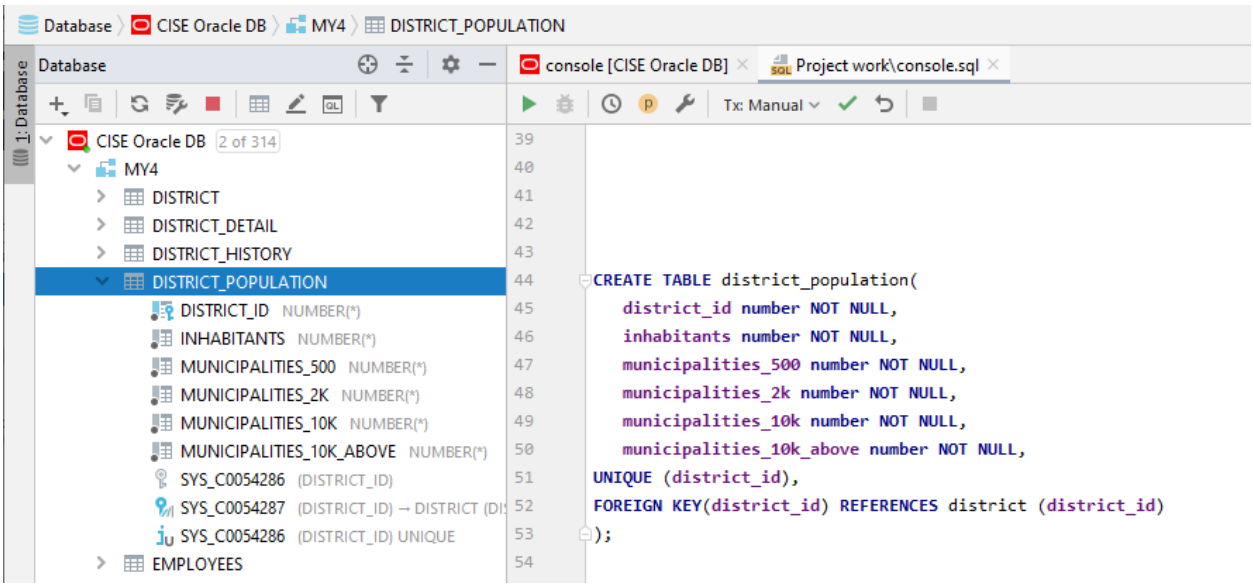
municipalities_10k number NOT NULL,

municipalities_10k_above number NOT NULL,

UNIQUE (**district_id**),

FOREIGN KEY(**district_id**) **REFERENCES** district (**district_id**)

);



District_History

CREATE TABLE district_history(

district_id number NOT NULL,

crimes number NOT NULL,

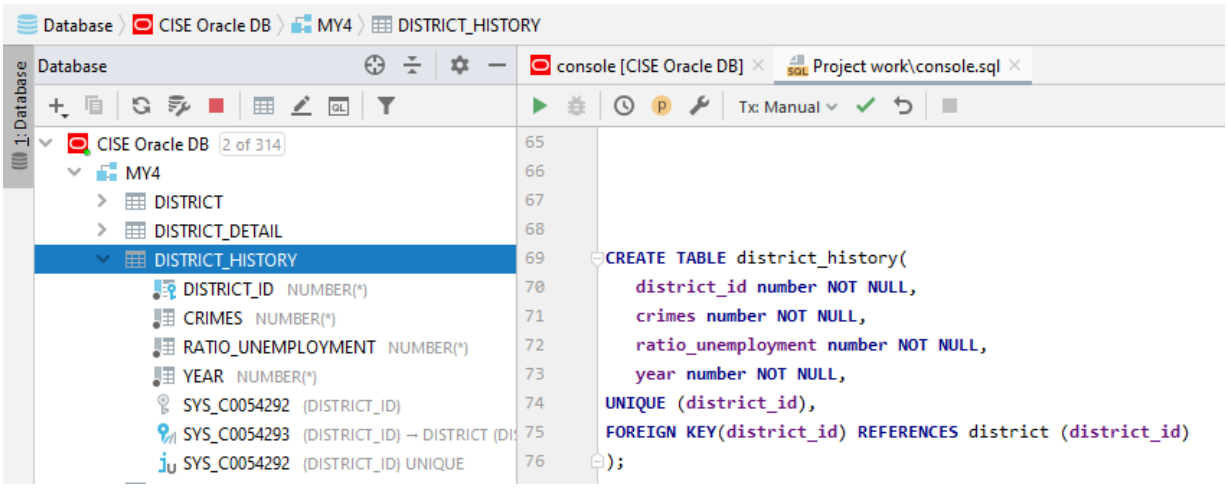
ratio_unemployment number NOT NULL,

year number NOT NULL,

UNIQUE (**district_id**),

FOREIGN KEY(**district_id**) **REFERENCES** district (**district_id**)

);



Loan

CREATE TABLE loan

(

loan_id number NOT NULL,

account_id number NOT NULL,

loan_date date NOT NULL,

duration number,

payments decimal(10, 2) **NOT NULL**,

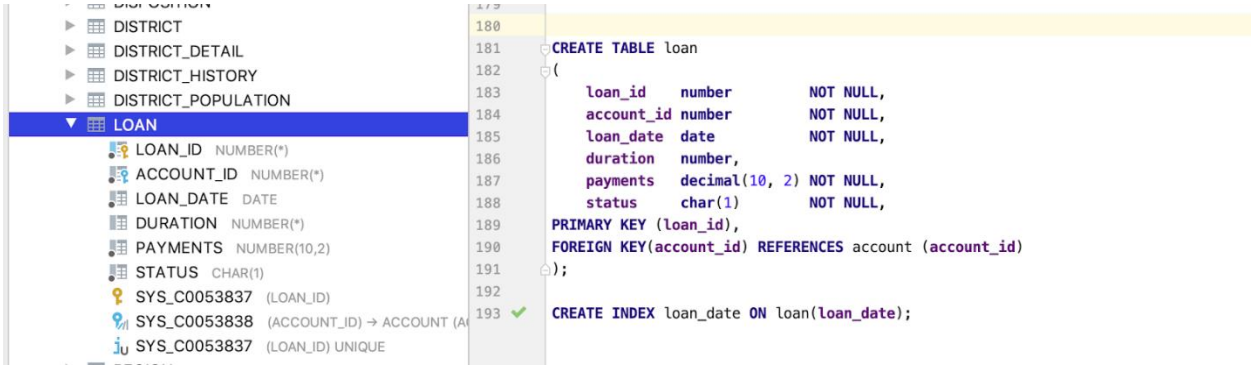
status char(1) **NOT NULL**,

PRIMARY KEY (**loan_id**),

FOREIGN KEY(**account_id**) **REFERENCES account** (**account_id**)

);

CREATE INDEX loan_date **ON** loan(**loan_date**);



Order

CREATE TABLE account_order (

order_id number NOT NULL,

account_id number NOT NULL,

bank_to char(2) **NOT NULL**,

account_to number NOT NULL,

amount decimal(10,2) **NOT NULL**,

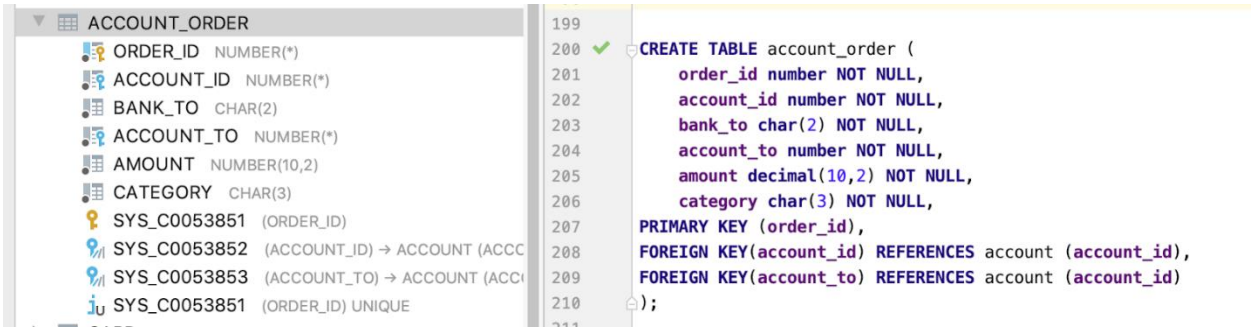
category char(3) **NOT NULL**,

PRIMARY KEY (**order_id**),

FOREIGN KEY(**account_id**) **REFERENCES account** (**account_id**),

FOREIGN KEY(**account_to**) **REFERENCES account** (**account_to**)

);



Transaction

CREATE TABLE transaction (

trans_id number NOT NULL,

account_id number NOT NULL,

created_date date NOT NULL,

trans_type char(2) **NOT NULL**,

operation char(3) **NOT NULL**,

amount decimal(10,2) **NOT NULL**,

balance decimal(10,2) **NOT NULL**,

category char(3) **NOT NULL**,

bank_to char(2) **NOT NULL**,

account_to number NOT NULL,

PRIMARY KEY (**trans_id**),

FOREIGN KEY(**account_id**) **REFERENCES account** (**account_id**)

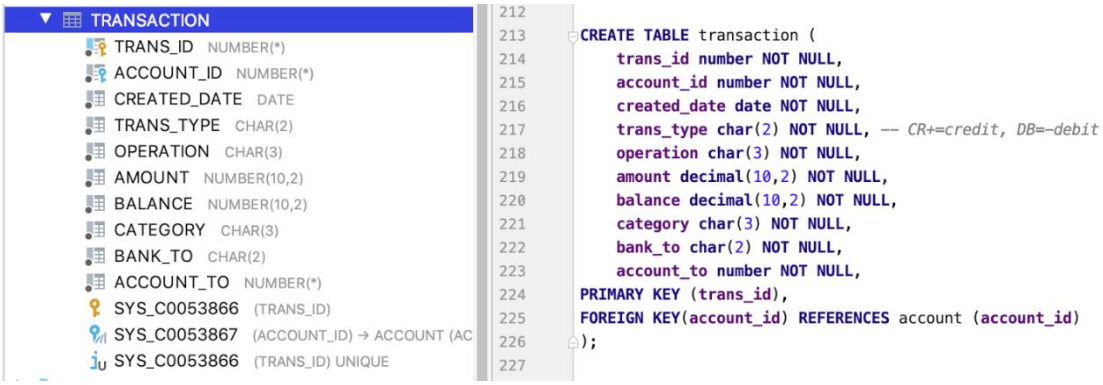
);

CREATE INDEX trans_created_date **ON transaction**(**created_date**, **trans_type**);

CREATE INDEX trans_amount **ON transaction**(**amount**);

CREATE INDEX trans_balance **ON transaction**(**balance**);

CREATE INDEX trans_category **ON transaction**(**category**);



Data Sources

The dataset to be used for drawing aforementioned inferences is hosted at data.world website. All of our data can be found in the links below.

- http://lisp.vse.cz/pkdd99/berka.htm

- https://data.world/lpetrocelli/czech-financial-dataset-real-anonymized-transactions/workspace/intro

- https://www.researchgate.net/post/Is_there_any_public_database_for_financial_transactions_or_at_least_a_synthetic_generated_data_set

- www.wikipedia.org