

## Data Manipulation and Transformation

### ACCOUNT Table

1. Convert the Date attribute into a yyyy-mm-dd by adding **23** in year format in Excel or SQL
  - 1993 -> 2017
  - 1994 -> 2018
  - 1995 -> 2019
  - 1996 -> 2020
  - 1997 -> 2021
2. Replace in frequency attribute “POPLATEK MESICNE” AS **Monthly Issuance**, “POPLATEKTYDNE” AS **Weekly Issuance**, and “POPLATEK POBRATU” AS **Issuance After a Transaction** in Excel or create a case statement in SQL.
3. Create a Custom Column Card\_Assigned and assign below :
  - Silver -> Monthly issuance
  - Diamond - weekly issuance
  - Gold - Issuance after a transaction

### CARD Table

1. Replace type attribute value “junior” as Silver, “Classic” as Gold, And “Gold” as Diamond by using replace in Excel or by using update in SQL.
2. Convert issued attribute into yyyy-mm-dd adding 23 in year.

### CLIENT Table

1. Convert bith\_number attribute to yyyy-mm-dd format and also create another column named sex by applying in bith\_number 0 for females and 1 for males.  
(=if(mod(bith\_number,2)=0, “Female”, “Male”) in excel or using case statement in SQL.  
For **Male** its in **YYMMDD** format and for female it is **YYMM+50DD** for **Women**

## Data Manipulation and Transformation

### DISTRICT Table

Change all column names and delete the attributes **a12** and

#### district.csv

Demographic

COLUMN NAME	TYPE	DESCRIPTION
# a1 ⓘ	integer	district code
T a2 ⓘ	string	district name
T a3 ⓘ	string	region
# a4 ⓘ	integer	no. of inhabitants
# a5 ⓘ	integer	no. of municipalities with inhabitants < 499
# a6 ⓘ	integer	no. of municipalities with inhabitants 500-1999
# a7 ⓘ	integer	no. of municipalities with inhabitants 2000-9999
# a8 ⓘ	integer	no. of municipalities with inhabitants >10000
# a9 ⓘ	integer	no. of cities
# a10 ⓘ	decimal	ratio of urban inhabitants
# a11 ⓘ	integer	average salary
# a12 ⓘ	decimal	unemployment rate '95
# a13 ⓘ	decimal	unemployment rate '96
# a14 ⓘ	integer	no. of entrepreneurs per 1000 inhabitants
# a15 ⓘ	integer	no. of committed crimes '95
# a16 ⓘ	integer	no. of committed crimes '96

## **Data Manipulation and Transformation**

### **LOAN Table**

1. Convert the Date Attribute into yyyy-mm-dd format adding 23 in year.
2. Convert Status Attribute value “A” as Contract Finished, “B” as Loan Not Paid, “C” as Running Contract, and “D” Client in debt.

### **DATA TRANSFORMATION**

Treating the null values in the Transaction table by years:

1. For 2016 - Raiffeisen bank
2. For 2017 - UniCredit Bank
3. For 2018 - Moneta Money Bank
4. For 2020 - Trinity Bank
5. For 2021 - 'Kameron bank